Biosynthesis and transport of terpenes

Hieng-Ming Ting

Thesis committee

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

Co-promotor
Dr A.R. van der Krol
Associate professor, Laboratory of Plant Physiology
Wageningen University

Other members
Prof. Dr A.H.J. Bisseling, Wageningen University
Prof. Dr M. Boutry, University of Louvain, Belgium
Dr M.A. Jongsma, Wageningen University
Dr M.H.A.J. Joosten, Wageningen University

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

Biosynthesis and transport of terpenes

Hieng-Ming Ting

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 24 March 2014
at 11 a.m. in the Aula.

Hieng-Ming Ting
Biosynthesis and transport of terpenes
184 pages.

PhD thesis, Wageningen University, Wageningen, NL (2014)
With references, with summaries in English and Dutch

ISBN 978-90-6173-892-9

To my beloved wife **Ya-Fen**
my lovely son **Teck-Yew**
my family in Malaysia

Contents

Chapter 1	General introduction	9
Chapter 2	The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of *CYP71AV1* type and relative gene dosage	23
Chapter 3	Experiments to address the potential role of LTPs in sesquiterpene emission	65
Chapter 4	Characterisation of aberrant pollen and ovule phenotypes associated with chromosomal translocations in two T-DNA insertion mutants of Arabidopsis	91
Chapter 5	Inhibition of vesicle transport during terpene biosynthesis causes proteasome malfunction	105
Chapter 6	General discussion	145
References		161
Summary		173
Samenvatting		175
Acknowledgements		177
Curriculum Vitae		180
Publications		181
Education statement		182

Chapter **1**

General introduction

Biological function of terpenes

Terpenes or terpenoids are a large and structurally diverse family of primary and secondary metabolites in plants (Trapp&Croteau, 2001). Some terpenes are only produced in minute amounts to function as plant phytohormones, such as gibberellic acid, abscisic acid, brassinosteroid, cytokinins and the strigolactones. Terpenes can also be produced in bulk amounts in plastids where they form (part of) the pigments functioning in photosynthesis, such as chlorophylls, plastoquinones and carotenoids. In addition, plants contain the structurally diverse monoterpenes, sesquiterpenes, diterpenes and triterpenes that function as secondary metabolites with important ecological functions in the interaction of plants with other organisms (Aharoni *et al.*, 2005, Pichersky&Gershenzon, 2002, Staniek *et al.*, 2013, Trapp&Croteau, 2001). Many of the terpenoids have biological activity in humans - as medicine or flavour and fragrance compounds - and therefore there are many efforts to elucidate and engineer the different steps in the biosynthesis of these compounds in plants (Staniek *et al.*, 2013).

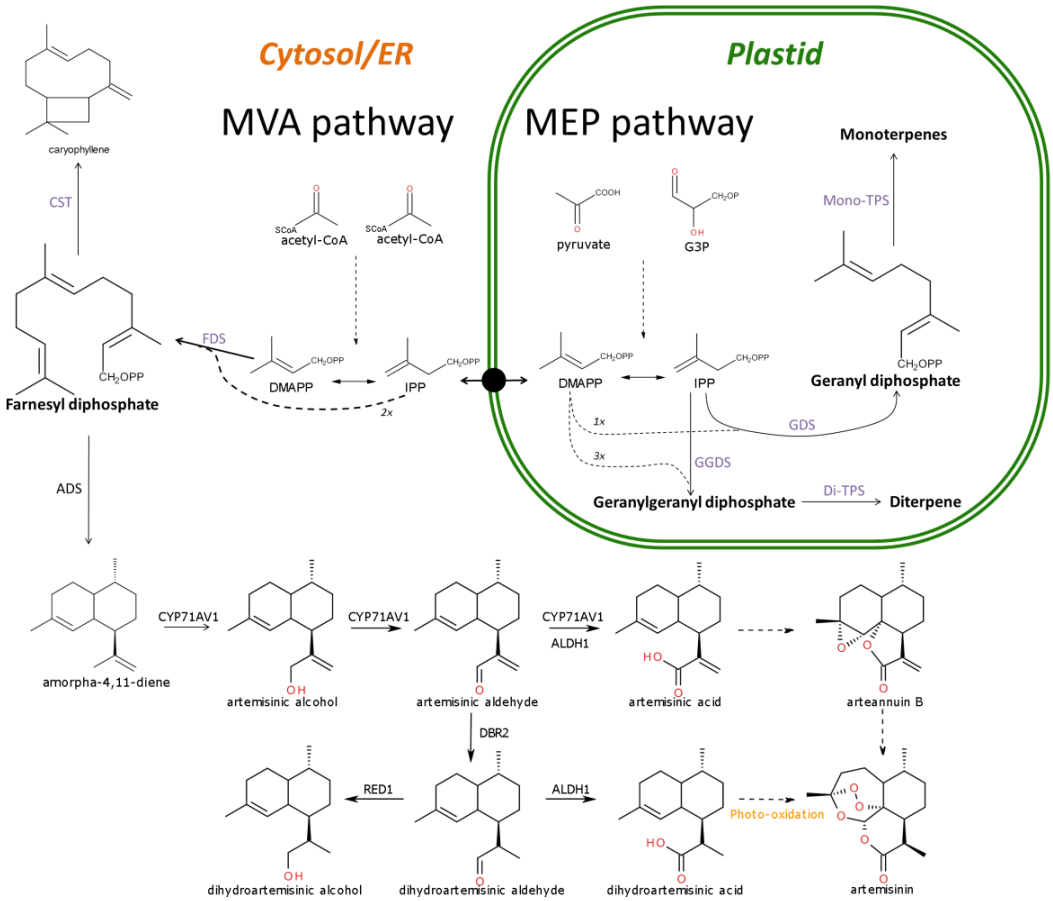

Figure 1. Compartmentation of terpene biosynthesis in the plant cell.

Terpene biosynthesis

Terpenes can be divided into hemiterpenes (C5), monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), sesterterpenes (C25) and triterpenes (C30), and all are synthesized from the condensation of the five-carbon isoprenoid precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP and DMAPP are synthesized by two independent pathways in different subcellular compartment, the cytosolic mevalonic acid (MVA) pathway and the plastidial methylerythritol phosphate (MEP) pathway (Fig. 1). The isoprenoid precursors are derived from acetyl-CoA in the MVA pathway, and from pyruvate and glyceraldehyde-3-phosphate in the MEP pathway. IPP and DMAPP are condensed by geranyl diphosphate synthase to form geranyl diphosphate (GPP, C10) in the plastids, to form the direct precursor for monoterpene biosynthesis. Geranylgeranyl diphosphate synthase also catalyses a condensation reaction with IPP and DMAPP in the plastid to form geranylgeranly diphosphate (GGPP, C20), which is the precursor for diterpenes and carotenoids. In the cytosol, IPP and DMAPP are condensed by farnesyl diphosphate synthase to form farnesyl diphosphate (FPP, C15) in the cytosol, the direct precursor for sesquiterpenes, triterpenes and sterol biosynthesis (Aharoni *et al.*, 2005, Bohlmann&Keeling, 2008, Nagegowda, 2010, Staniek *et al.*, 2013). In this thesis, I have mainly worked with two sesquiterpene biosynthesis pathways: the multi-enzyme pathway for artemisinin biosynthesis in *Artemisia annua* and the single enzymatic pathway of caryophyllene biosynthesis of Arabidopsis by caryophyllene synthase (CST) (Fig. 1).

Artemisinin is the sesquiterpene lactone endoperoxide produced by the plant *A. annua* L (Fig. 1). The biosynthesis of artemisinin starts with the cyclization of FPP to amorpha-4,11-diene by amorphadiene synthase (ADS) (Bouwmeester *et al.*, 1999, Mercke *et al.*, 2000). Amorpha-4,11-diene is subsequently oxidized by amorphadiene oxidase (AMO/CYP71AV1), a P450 enzyme, to artemisinic alcohol, artemisinic aldehyde and artemisinic acid (Ro *et al.*, 2006, Teoh *et al.*, 2006). In the low artermisinin production chemotype, artemisinic acid is likely spontaneously converted to arteanniun B. However, in the high artermisinin production chemotype, artemisinic aldehyde is reduced by artemisinic aldehyde reductase (DBR2), to the intermediate dihydroartemisinic aldehyde which is converted by aldehyde dehydrogenase 1 (ALDH1) to the final intermediate before artemisinin, dihydroartemisinic acid (Bertea *et al.*, 2005, Zhang *et al.*, 2008). The last step, conversion of dihydroartemisinic acid to artemisinin, is likely non-enzymatic and occurs spontaneously through photo-oxidation (Sy&Brown, 2002, Wallaart *et al.*, 1999). Recently, it was shown that dihydroartemisinic aldehyde can also be

converted to dihydroartemisinic alcohol by artemisinic aldehyde reductase (RED1) (Rydén *et al.*, 2010) (Fig. 1).

A. annua has two types of trichomes: non-glandular and glandular, secretory, trichomes. It has been shown that the artemisinin biosynthesis genes (*ADS*, *CYP71AV1*, *DBR2* and *ALDH1*) are expressed in both apical and sub-apical cells of the glandular secretory trichomes (Olofsson *et al.*, 2012). Hence, dihydroartemisinic acid is produced in the apical and sub-apical cells, while artemisinin is excreted into and stored in the subcuticular space of the glandular secretory trichomes (Duke *et al.*, 1994).

Caryophyllene is a common sesquiterpene floral volatile, which is a constituent of more than 50% of the angiosperm families' floral odor (Knudsen *et al.*, 2006). In Arabidopsis, for example, more than 40% of the total floral volatiles emitted by the flowers is represented by caryophyllene (Chen *et al.*, 2003). The biosynthesis of caryophyllene consists of a single step, catalysed by caryophyllene synthase (CST), which cyclizes FPP to caryophyllene (Fig. 1). CST-promoter:GUS analysis showed that Arabidopsis CST is mainly expressed in the stigma of the flower suggesting that caryophyllene is exclusively released from the stigmas (Chen *et al.*, 2003, Tholl *et al.*, 2005).

Metabolic engineering of terpene biosynthesis

Because of their commercial and ecological importance, there is strong interest in the possibilities to enhance or chance the production of terpenoids in plants. The assumption was that with some understanding of the biosynthesis pathways terpene biosynthesis can be engineered in the host plant itself or in heterologous expression systems.

Engineering of plants

Because many of the medicinal plants that contain commercially interesting terpenes often have low yields and are difficult to cultivate it may be of benefit to engineer the pathway of an interesting terpene in an alternative plant host, which is easy to grow and propagate. Tobacco is a fast-growing and high-biomass producing crop species that seems to be a suitable heterologous host for terpene engineering. For instance, monoterpene synthases from lemon (*Citrus limon* L. Burm. f.) (Lücker *et al.*, 2004); patchoulol synthases from *Pogostemon cablin* , ADS from *A. annua* and limonene synthase from lemon (Wu *et al.*, 2006); geraniol synthases from *Valeriana officinalis* and *Lippia dulcis* (Dong *et al.*, 2013) have been successfully transformed into tobacco. Furthermore, multiple attempts are under

way to engineer the artemisinin biosynthesis pathway in tobacco (Farhi *et al.*, 2011, Zhang *et al.*, 2011). Artemisinin was indeed detected in *N. tabacum* after transformation with multiple genes (five genes involved in the mevalonate and artemisinin pathways: *HMGR*, *ADS*, *DBR2*, *CYP71AV1* and *CPR*) combined into a single transformation vector. All of these multiple genes were separated and driven by different promoters and terminators. To increase the availability of precursor for artemisinin production, these multiple genes were targeted to different subcellular compartments, such as the mitochondria for amorphadiene synthase (Farhi *et al.*, 2011). This was the first heterologous host plant overexpressing the artemisinin biosynthetic genes, which can produce artemisinin. However, the amount of the artemisinin produced, 7 μg/g dry weight, is still much lower than *A. annua* (0.01~1g / g dry weight) and the results have not been confirmed by a peer reviewed publication. Hence, further studies are needed to improve and optimize artemisinin production by stable transformation of heterologous plant hosts.

Previously, artemisinin biosynthetic genes have been transiently expressed in *Nicotiana benthamiana* (van Herpen *et al.*, 2010) and stable in *Nicotiana tabacum* (Wu *et al.*, 2006, Zhang *et al.*, 2011). However, no artemisinin was detected in these *Nicotiana* species. Zhang et al did find artemisinin precursors amorphadiene, artemisinic alcohol and dihydroartemisinic alcohol but no artemisinic acid or dihydroartemisinic acid (Zhang *et al.*, 2011). The likely explanation for this is that the Wageningen group showed that these artemisinin precursors are strongly modified (oxidation, glycosylation, etc.) by endogenous *N. benthamiana* enzymes (van Herpen *et al.*, 2010). This shows that conversion to glycosides may be a problem connected with pathway expression in heterologous plant hosts. Also in other engineering studies this glycosylation occurred, not only of the end product as found by Van Herpen (van Herpen *et al.*, 2010) but also with pathway intermediates (Miettinen *et al.*, 2013). In the latter paper it was shown that a range of oxidation and glycosylation reactions drain intermediates from the strictosidine pathway that was transiently expressed in *N. benthamiana.* This resulted in only very low levels of free compounds being produced. A likely explanation for this is that these products in a heterologous host are not transported out of the cell for sequestration, as usually happens in the homologous host, where it cannot be further converted by detoxification enzyme activities.

Therefore, some attempts have been made to engineering terpene biosynthesis in the homologous plant host. One approach has been to boost the pathway providing the general precursor for terpene biosynthesis. In peppermint, the constitutive overexpression of the enzyme that catalyses the first committed step in the MEP pathway, 1-deoxy-D-xylulose-5-

phosphate reductoisomerase (DXR), in combination with an antisense construct for menthofuran synthase resulted in 44% increased essential oil yield compared with wild-type (Mahmoud&Croteau, 2001). Similarly, transgenic spike lavender plants overexpressing 1-deoxy-D-xylulose-5-phosphate synthase (DXS) from Arabidopsis, had up to 74% higher essential oil yield in flowers and up to 359% in leaves when compared with wild-type (Muñoz-Bertomeu *et al.*, 2006).

Also for artemisinin production in *A. annua* this has been attempted, by overexpression of key enzymes of the precursor pathway such as HMGR and FPS, resulting in a 1.2-fold and 3-fold increase in production levels, respectively (Aquil *et al.*, 2009, Chen *et al.*, 2000). Alternatively, or in addition, the expression of the artemisinin biosynthesis pathway genes may be boosted by overexpression. Overexpression of *ADS*, *CYP71AV1* and *CPR* in *A. annua* resulted in a 2~2.4-fold increase in artemisinin production (Jing *et al.*, 2008, Ma *et al.*, 2009). In another study, overexpression of *FPS*, *CYP71AV1* and *CPR* in *A. annua* resulted in a 3.6-fold increase in artemisinin production (Chen *et al.*, 2013). Recently, two AP2/ERF transcription factors from *A. annua*, AaERF1 and AaERF2, were identified which are involved in the coordinated expression of the different genes of the artemisinin biosynthesis pathways and it was demonstrated that these transcription factors can bind to the *ADS* and *CYP71AV1* promoters (Yu *et al.*, 2012). The content of artemisinin significantly increased in *A. annua* transgenic lines (1.6-fold increased artemisinin yield) which over-expressed AaERF1 and AaERF2 when compared to wild-type (Yu *et al.*, 2012).

Engineering of micro-organisms

Also *E. coli* and yeast have the basic metabolic pathways supplying the precursors for terpenoids. Upon engineering with plant genes they are able to synthesise terpenoids, such as amorphadiene, artemisinic acid, dihydroartemisic acid and a taxol precursor (Chandran *et al.*, 2011, Covello, 2008, Paddon *et al.*, 2013).

Introduction of the MVA pathway from yeast into *Escherichia coli* and co-expressing it with the *ADS* from *A. annua* resulted in the production of >25g/L amorphadiene in *E. coli* (Tsuruta *et al.*, 2009). The artemisinin biosynthetic genes *ADS* and *CYP71AV1* were also expressed in yeast, resulting in the production of artemisinic acid and dihydroartemisinic acid (Lenihan *et al.*, 2008, Ro *et al.*, 2006, Zhang *et al.*, 2008). Westfall and co-worker reported that production of >40g/L amorphadiene and some artemisinic acid was achieved in yeast over-expressing genes from the MVA pathway plus *ADS*, *CYP71AV1* and *CPR* (Westfall *et*

al., 2012). Artemisinic acid and dihydroartemisinic acid can subsequently be converted to artemisinin through a chemical process (Kopetzki *et al.*, 2013, Paddon *et al.*, 2013).

Metabolic engineering of *E. coli* by combined optimization of the MEP pathway and taxadiene synthase can produced taxadiene up to >1g/L in fed-batch fermentations (Ajikumar *et al.*, 2010). Conversion of taxadiene to taxadien-5α-ol is the next step in taxol formation, and >58 mg/L of taxadien-5α-ol was achieved when a CYP450 and CPR from *Taxus cuspidate* were co-expressed in the taxadiene producing *Saccharomyces cerevisiae* strain (Ajikumar *et al.*, 2010).

The role of terpene transport

As discussed above, the early metabolic engineering work in plants has shown that it is possible to make plants produce novel terpenoids. However, production levels are still far from economically relevant. This is due to a number of limitations, one of these is that understanding the production of terpenes not only involves the identification, characterization and engineering of the biosynthetic genes (Bohlmann&Keeling, 2008, Bouvier *et al.*, 2005, Tholl, 2006), but also how these (often very lipophilic) molecules are transported within the cell from one enzyme to the next and how the products are sequestered for example by transport from within the cell to the apoplast. On these latter aspects hardly anything is known so far and we therefore postulated that gaining knowledge about terpene transport would potentially allow us to improve the success of terpenoid metabolic engineering.

Transport between enzymes

Terpenes biosynthesis pathways are often distributed over several different subcellular compartments, such as plastids, mitochondria and ER. The precursors or intermediates therefore require transport possibly mediated by transporters to facilitate exchange between the different organelles. In the multi-enzymatic pathway of artemisinin biosynthesis, for example, the first committed step in the pathway is by ADS, which is localized in the cytosol (Kim *et al.*, 2008). In the next step, amorphadiene is converted to artemisinic alcohol, artemisinic aldehyde and artemisinic acid by CYP71AV1 which is supposed to be anchored to the cytoplasmic surface of the endoplasmic reticulum (ER) like all cytochrome P450s. At this CYP71AV1 step the pathway can also branch. The final product of efficient CYP71AV1 catalysis is artemisinic acid, which is the precursor for arteannuin B. Alternatively, the product artemisinic aldehyde may serve as substrate for an enzymatic reduction to

dihydroartemisinic aldehyde by DBR2, which is again localized in the cytosol (Zhang *et al.*, 2008). Dihydroartemisinic aldehyde is subsequently the substrate for an aldyhde dehydrogenase, ALDH1 which is supposedly localised in the cytosol as no signal peptide or organelle-targeting signal was found in the protein sequence of AaALDH1 (Teoh *et al.*, 2009). The product of AaALDH1 is dihydroartemisinic acid which is the precursor of artemisinin. In *A. annua* there are varieties with relatively high levels of artemisinin and low levels of arteannuin B (HAP chemotypes) and varieties with high arteannuin B and low levels of artemisinin (LAP chemotypes). The choice of transfer of the intermediate artemisinic aldehyde to either CYP71AV1 – for further conversion to artemsinic acid - or to DBR2 – for conversion to dihydroartemisic aldehyde – seems therefore an important determinant for having a HAP or LAP chemotype. As all enzymes of the pathway are present in both HAP and LAP chemotypes it needs further investigation to determine what causes the difference in flux through the two different branches of the pathway (Maes *et al.*, 2011).

Transport to the plasma membrane

Several studies have observed phytotoxicity of terpenes on plant cells (Chowhan *et al.*, 2013, Dayan *et al.*, 1999, Graña *et al.*, 2013) raising the question how plants protect themselves against these potential toxic compounds. This can be achieved by two alternative mechanisms: detoxification by conjugation to sugar groups or glutathion (GSH) and sequestration in the vacuole, or by sequestration in the apoplast by secretion across the plasma membrane. Indeed, many studies provide evidence for the latter, showing that terpenes are secreted and accumulated in the extracellular, sub-cuticular space of glandular trichomes (Dai *et al.*, 2010, Markus Lange&Turner, 2013, Tissier, 2012). Because biosynthesis of sesquiterpenes occurs in the cytosol and at the ER membrane where further modifications are catalysed by P450 enzymes, transport of the terpene molecules to the plasma membrane is required, before transport to the apoplast can occur. This may either be mediated by carrier proteins (e.g. like the LTP proteins for lipids) or by sequestration in vesicles or lipid droplets and fusion of these vesicles with the plasma membrane (vesicle pathway).

Transport over the plasma membrane to the apoplast

At the plasma membrane, the lipophilic terpenes may either diffuse passively through the membrane to the apoplastic space (passive pathway) (Caissard et al., 2004, Effmert et al., 2005, Niinemets et al., 2004) or specific plasma membrane transporters may be involved in the transport over the plasma membrane (transporter pathway).

ABC transporters are encoded by a large gene family in eukaryotes (e.g. 48 genes in human) (Dean, 2009) but especially in plants (e.g. 130 genes in Arabidopsis) (Kang *et al.*, 2011). ABC transporters function as exporter or importer of different type of substrates, including xenobiotics, peptides, lipids, steroids, inorganic acids, carboxylates, heavy metal chelates, chlorophyll catabolites and phytohormones across different biological membranes (e.g. plasma membrane, ER membrane vacuolar membrane; see review (Kang *et al.*, 2011)). ABC transporters have one or multiple so-called ATP binding cassettes and they function as ATP-driven pumps (Martinoia *et al.*, 2002). Some transporters do not use ATP but other nucleotides and the structural domains of the ABC transporters can therefore be divided by trans-membrane domains (TMD) and nucleotide binding domains (NBD). ABC transporters are divided into several subfamilies based on the number and the spatial distribution of these structural domains. All the pleiotropic drug resistance (PDR) transporters, belonging to the full size ABCG subfamily and consisting of two copies of the two basic elements, are found only in plants and fungi (Verrier *et al.*, 2008).

Several lines of evidence suggest that these ABCG/PDR transporters play a central role in plants, such as heavy metal resistance (Kim *et al.*, 2007), abiotic stresses tolerance (Kim *et al.*, 2010), pathogen defence (Stein *et al.*, 2006), auxin precursor transport (Růžička *et al.*, 2010), ABA transport (Kang *et al.*, 2010) and strigolactone transport (Kretzschmar *et al.*, 2012). Moreover, several ABCG transporters have been proven to be involved in the transport of specific terpenes. NpPDR1 from *Nicotiana plumbaginifolia* is localized in plasma membrane and is the first ABCG transporter reported to secrete an antifungal diterpene (sclareol) into the apoplast (Jasiński *et al.*, 2001). The homologs of NpPDR1, AtPDR12 from Arabidopsis and SpTUR2 from *Spirodela polyrrhiza*, have also been suggested to be involved in secretion of sclareol (Campbell *et al.*, 2003, Van Den Brûle *et al.*, 2002). Recently, another plasma membrane ABCG transporter, NtPDR1, was also shown to be involved in diterpene (sclareol, manool and the macrocyclic cembrene) transport to the apoplast, but not monoterpene (eucalyptol) transport in *N. tabacum* BY2 cells (Crouzet *et al.*, 2013).

In *A. annua* the artemisinin biosynthesis gene are expressed in both apical and sub-apical cells of the glandular secretory trichome (Olofsson *et al.*, 2012), and the end product artemisinin is stored in the subcuticular space of the glandular secretory trichomes (Duke *et al.*, 1994, Duke&Paul, 1993, Tellez *et al.*, 1999). Thus, it is hypothesised that an ABCG transporter must be involved in transport of artemisinin precursors and is localized in the plasma membrane of the apical cells of the glandular secretory trichome. During the course

of my PhD, several ABCG type transporters from *A. annua* were identified that supposedly play a role in the transport of DHAA from the apical glandular trichome cell to the subcuticular space of the trichome where DHAA is likely (photochemically) converted to artemisinin (Zhang *et al.*, 2012).

Caryophyllene is also a potential toxic compound as it can be converted to caryophyllene oxide which can react with proteins (Park *et al.*, 2011). The plant cell prevents high build-up of endogenous caryophyllene by emission of the volatile compound into the headspace of the plant. At the start of my PhD, it was not known how volatile (sesqui)terpenes are emitted by plants, e.g. whether carrier proteins are involved in bringing the molecule to the plasma membrane and/or whether transport proteins are involved in transport over the plasma membrane.

Lipid transfer proteins (LTPs)

Plant lipid transfer proteins (LTPs) are small, abundant, basic proteins having eight cysteine residues conserved in similar positions in the primary structure to form four disulphide bridges forming a hydrophobic cavity (Kader, 1996). LTPs have been found in monocots, dicots and gymnosperms. It has been shown that LTPs play a role in phospholipids and fatty acids transfer between membranes *in vitro* (Kader, 1996). LTPs can be categorised into two main families according to their molecular weight: Type-I LTPs have molecular masses of around 9 kDa and Type-II LTPs have molecular masses of 7 kDa (Douliez *et al.*, 2000). Most of the LTPs were proposed to be localised outside the cell as they contain an N-terminal secretion signal peptide (Kader, 1996, Wang *et al.*, 2012). However, recently HaAP10, an LTP from wheat, was demonstrated to be localized intracellularly in imbibing seeds (Pagnussat *et al.*, 2012). Several studies have shown that LTPs are involved in different biological roles such as pathogen defence and signalling (Jung *et al.*, 2003, Maldonado *et al.*, 2002), wax assembly and cutin deposition (Cameron *et al.*, 2006, Hollenbach *et al.*, 1997), cell wall formation (Nieuwland *et al.*, 2005) and pollen tube adhesion (Chae *et al.*, 2009). Additionally, LTPs are also involved in resistance against abiotic stresses, such as freezing and drought stress (Guo *et al.*, 2013). However, at the onset of my PhD, it was not clear whether the coincidence of LTP and terpene biosynthesis gene expression also means that LTPs are involved in terpene transport. Near the end of my thesis research a new publication provided more direct evidence that LTPs are involved in terpene transport. NtLTP1, which is expressed in glandular trichomes of tobacco, was shown to have lipid binding activity (Choi *et al.*, 2012). Moreover, NtLTP1 was shown to be limiting for the secretion of several

compounds, including a diterpene [cembratrienol, (1R,3S)-cembra-4,7,11,15-tetraen-3-ol, cembratriendiol and labda-8(17),13E-dien-15-al], alkane and aromatic dicarboxylic acid from the glandular trichomes, and overexpression of *NtLTP1* resulted in a 1.6~1.9-fold increase in diterpene secretion (Choi *et al.*, 2012). As discussed above, in *A. annua* artemisinin is located in the extracellular space of the glandular trichomes (Brown, 2010, Duke *et al.*, 1994, Duke&Paul, 1993, Tellez *et al.*, 1999). *LTP* cDNA sequences are highly represented in cDNA libraries of these glandular secretory trichomes (Bertea *et al.*, 2006). Moreover, also other studies have indicated a correlation between the site of terpene production and elevated expression of *LTPs* (Harada *et al.*, 2010, Lange *et al.*, 2000, Schilmiller *et al.*, 2010). The close association of *LTP* gene expression and terpene production and the fact that LTPs contain a hydrophobic pocket which may also accommodate lipophilic terpenes, and that there are examples of LTPs limiting diterpene secretion (Choi *et al.*, 2012) suggest it is likely that LTPs are somehow involved in terpene transport.

Vesicle transport

Within the cell, the intracellular trafficking of vesicles budding of the ER is important for the secretion of proteins to the plasma membrane and apoplast, as well as for transport from the ER to other cell organelles like peroxisomes and the vacuole (Jürgens, 2004). In the vesicular trafficking pathway, the cargo of vesicles is delivered to the target compartment by fusion of the cargo vesicle membrane with the membrane of the target compartment. This vesicle fusion is mediated by the SNARE (soluble *N*-ethylmaleimide sensitive factor attachment protein receptors) superfamily proteins. One v-SNARE (SNARE protein on the transport vesicle) pairs with three t-SNARE proteins (SNARE protein on the target membrane, including syntaxins) to form a stable heteromeric core complex which drives the membrane fusion (Lang&Jahn, 2008). v-SNAREs consist of long vesicle-associated membrane proteins (VAMPs) or ‘longins’ (containing an N-terminal longin domain) and short VAMPs or ‘brevins’ (Filippini et al., 2001). Plants are unique for having a widely expanded set of longin-type v-SNARE genes, which can be further categorized into three major groups: VAMP7-like, Ykt6-like and Sec22-like. Additionally, based on the phylogenetic analysis of protein sequence, plants have two major groups of VAMP7-like proteins: VAMP71 and VAMP72 (Fujimoto&Ueda, 2012). The VAMP71 family seems to play a role in vacuolar/late endosomal trafficking, while the VAMP72 family of v-SNAREs is localized to the trans Golgi network and is involved in exocytotic processes in plants (Kwon et al., 2008, Sanderfoot, 2007).

Proteins, lipids, membrane components and soluble cargo can be transported by vesicles between intracellular compartments and, when fusion to the plasma membrane occurs, to the apoplast (see review (Bassham&Blatt, 2008, Gerst, 1999)). Some histological studies indicate that vesicle transport may be involved in the secretion of terpenes. In grapevine flowers, the sesquiterpene synthase valencene synthase (VvValCS) is responsible for the major sesquiterpenoid volatiles in grapevine's flowers and immunolabeling analysis showed that VvValCS is localized in the outer edges of lipid vesicles in pollen grains (Martin *et al.*, 2009). This suggests that valencene is stored in and transported by these lipid vesicles. Moreover, emission of sesquiterpenes copaene and caryophyllene can be induced by heat treatment in *Sauromatum guttatum* flowers. This occurs in parallel with the release of vesicles from the ER and then fusion to the plasma membrane (Skubatz *et al.*, 1995). However, evidence for a direct causal relationship between the induced vesicle transport activity and terpene emission is lacking. Terpenes may be sequestered into the membrane of vesicles budding off from the ER and thus could be transported along with the protein secretion pathway to their destination. Alternatively, there could be dedicated vesicle trafficking for terpene transport within the cell. Furthermore, electron microscopy studies of the secretory cells of the glandular trichomes of *Prostanthera ovalifolia* have shown that the plastids in these cells are surrounded by vesicles and these vesicles do fuse with the plasma membrane. This may be of relevance to the transport of monoterpenes as the plastids are the subcellular compartment where monoterpene biosynthesis takes place (Gersbach, 2002).

Scope of the thesis

Many studies have been done to identify and characterise terpene biosynthetic genes and to use these genes for metabolic engineering (see reviews: (Covello, 2008, Farhi *et al.*, 2013, Staniek *et al.*, 2013)). Surprisingly, all the metabolic engineering is only done with enzymes, without any clue of how terpene transport and storage in or outside the cell occur. Hence, in this thesis I address different aspects of transport of terpenes in plants. Firstly, I study the issue of substrate transport between enzymes, including the regulation of intermediate transport between two different biosynthesis enzymes (CYP71AV1 and DBR2) that determines the resulting *A. annua* chemotype (LAP and HAP) (**Chapter 2**). To achieve this, I established the expression of the entire artemisinin biosynthetic pathway in *Nicotiana benthamiana*. Although I managed to get the pathway working in *N. benthamiana* and to

elucidate the mechanism behind the two chemotypes of *A. annua*, artemisinin was not produced. We hypothesize that this might be caused by the absence of (the right) transporters in *N. benthamiana*.

Therefore, I looked for candidate transporter genes in an *A. annua* trichome cDNA library. I found that *LTPs* are highly expressed in *A. annua* trichomes. Three *AaLTPs* were cloned and their protein sequences compared with Arabidopsis LTPs. In **Chapter 3**, three different functional assays were carried out to study the role of AaLTPs in relation to the production of the sesquiterpene artemisinin and its precursors. Moreover, I also studied the involvement of two Arabidopsis LTPs (LTP1 and LTP3) in sesquiterpene emission. Unfortunately, functional analysis of the interaction between LTPs and terpene synthases (TPS) in Arabidopsis was impossible, likely by chromosomal translocation in the *ltp1* and *tps21* T-DNA insertion lines. In **Chapter 4** I describe and discuss the aberrant pollen and ovule phenotype that is associated with the apparent chromosomal translocation in the *ltp1* and *tps21* lines.

Previous studies suggested that vesicle transport may also be involved in terpene secretion. The VAMP72 family of the v-SNAREs are involved in vesicle transport and MtVAMP721e from *Medicago truncatula* has been shown to be involved in the exocytotic pathway of vesicle transport. In **Chapter 5** I used an RNAi construct of *MtVAMP721e* to investigate the role of VAMP72 in terpene transport. Hereto I transiently expressed caryophyllene synthase (*CST*) together with *MtVAMP721e-RNAi* in *N. benthamiana*. To gain knowledge about the transcriptional regulation of the genes during expression of *CST* and inhibition of VAMP72, I did RNAseq analysis.

Finally, in **Chapter 6** I discuss the most important findings from this thesis and consider future perspectives of the study of terpene transport.

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.

New Phytologist (2013) 199: 352-366, doi: 10.1111/nph.12274

Abstract

The plant *Artemisia annua* that produces the anti-malaria compound artemisinin, 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 heterologous production platforms. We present the first systematic comparison of artemisinin biosynthesis genes to determine factors that contribute to artemisinic acid (AA) or dihydroartemisinic acid (DHAA) chemotype of agro-infiltrated leaves with ADS, *CYP71AV1*/*AMO*, *DBR2* and *ALDH1*. Results show that the enzyme activity of DBR2 and ALDH1 from the two chemotypes does not differ. The Amorphadiene Oxidase from HAP (AMOHAP) showed reduced activity compared to that from LAP chemotype (AMOLAP), which relates to a seven amino acid N-terminal extension in AMOLAP 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 transient expression in *Nicotiana benthamiana* show that AMOLAP not only displayed 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 saturated (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*.

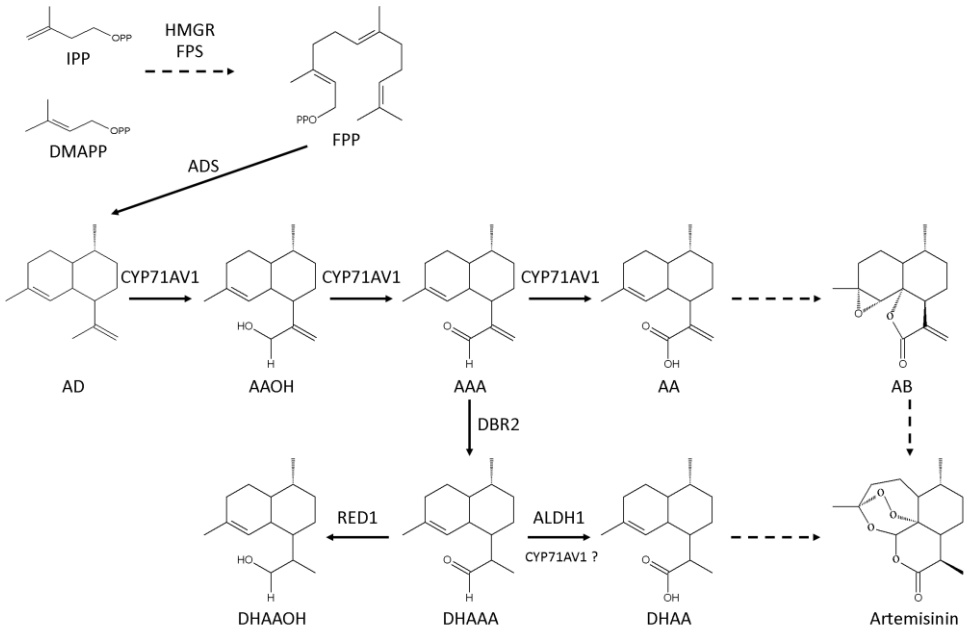

Figure 1. Artemisinin biosynthetic pathway in *Artemisia annua*.
IPP: isopentenyl diphosphate; DMAPP: dimethylallyl diphosphate; FPP: farnesyl diphosphate; AD: amorpha-4,11-diene; AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; AB: arteannuin B; DHAAOH: dihydroartemisinic 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 GGCCGCTAATTGGTACCGAAATG) 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 baseline 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 using 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_2PO_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), 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).

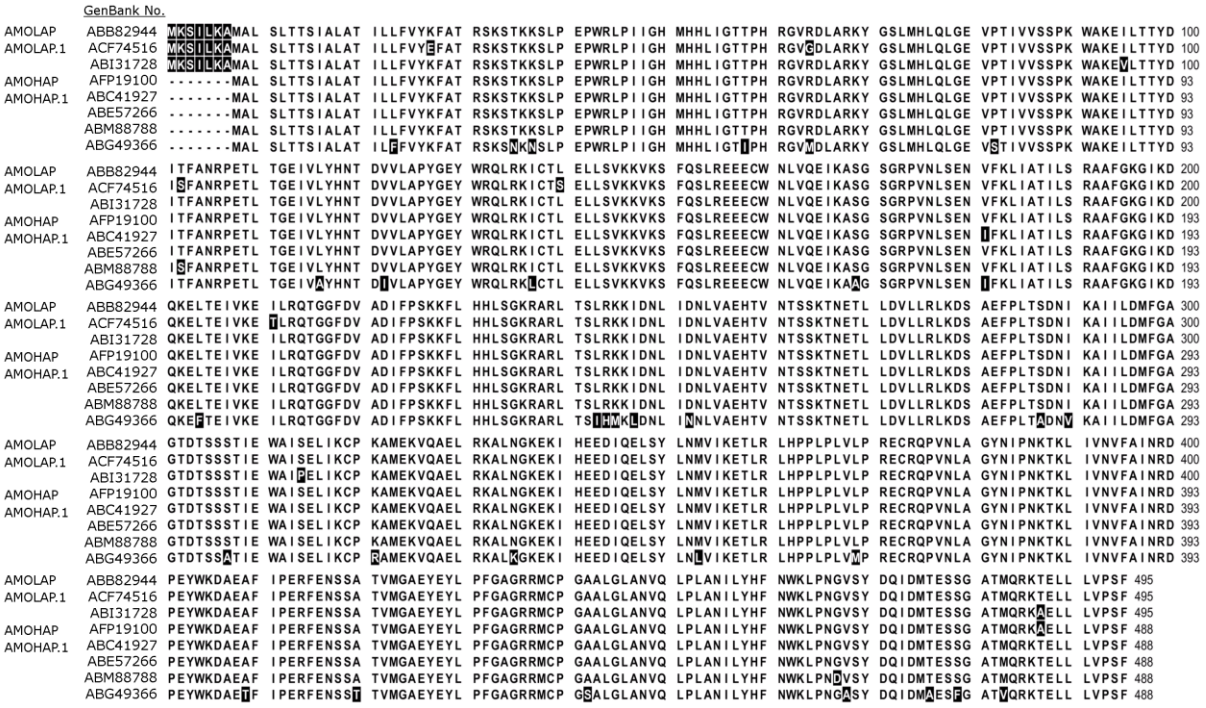

Figure 2. Alignment of the AMOLAP and AMOHAP deduced amino acid sequences.
All amino acid sequences used in the alignment were retrieved from GenBank (GenBank number given in front of each sequence). Two of the sequences were from confirmed LAP chemotypes: ABB82944, referred to as AMOLAP (Ro *et al.*, 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.

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).

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 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:YFP) (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.

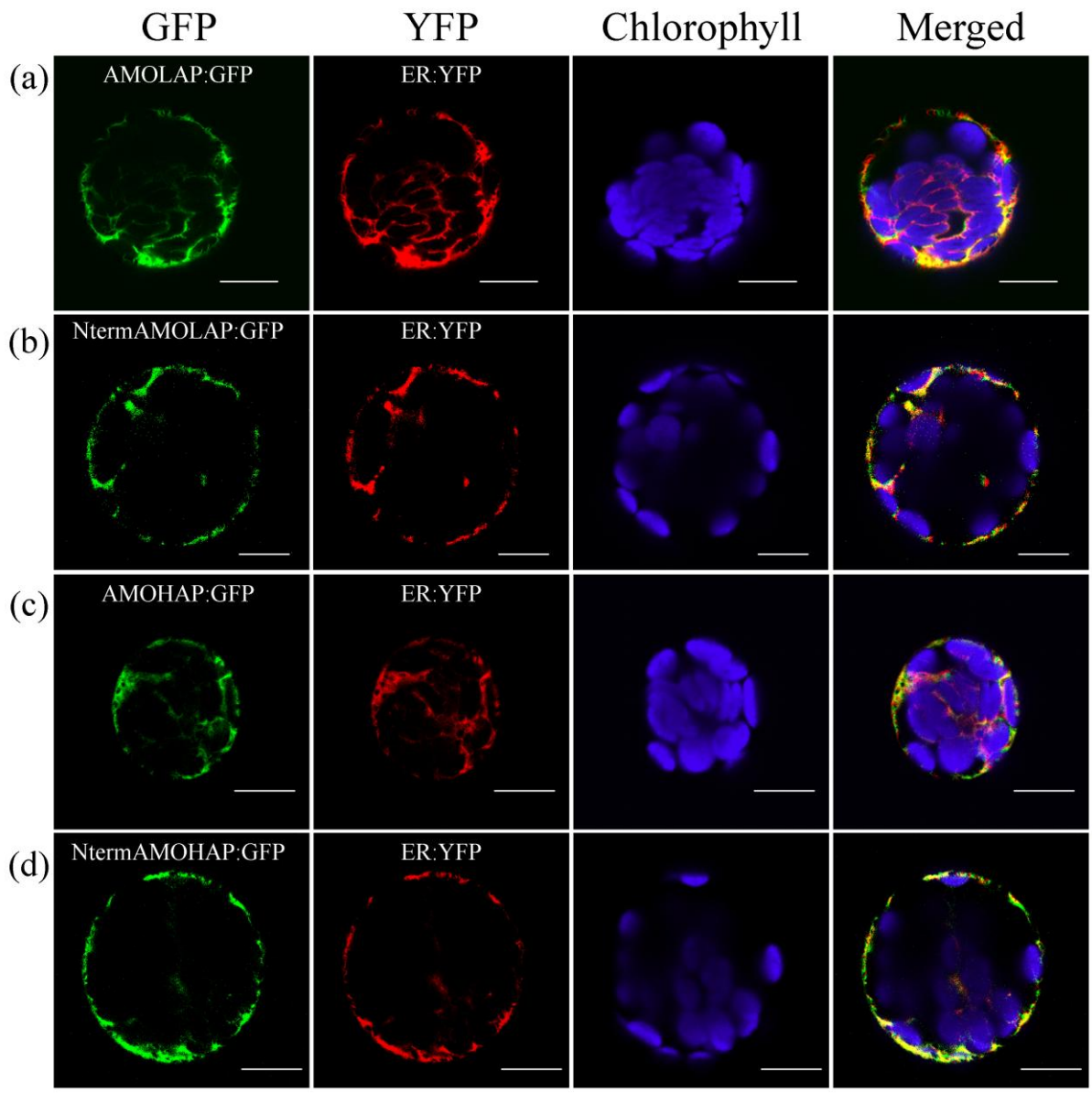

Figure 4. Subcellular localization of AMOLAP and AMOHAP.
Confocal microscopy analysis of *Arabidopsis* protoplasts co-transfected with (a) AMOLAP:GFP + ER:YFP; (b) NtermAMOLAP:GFP + ER:YFP; (c) AMOHAP:GFP + ER:YFP; (d) NtermAMOHAP:GFP + ER:YFP. Artificial colors were given to GFP fluorescence (green), 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 mitochondrial targeting signal, and overexpression was combined with a mitochondrial 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.

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-diglucoside, 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). 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.

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 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.

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

($ng\ g^{-1}$ FW)	AmFH+AMOLAP	AmFH+AMOHAP	AmFH+AMOLAP+DBR2	AmFH+AMOHAP+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.

Table 2. Conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaf extracts.

Ret (min)	Detected Mass(D)[a]	MS-MS fragments	Mol form	ΔMass (ppm)	Putative ID	Intensity[b]			
						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-CO_2-H]$^-$,543[M-Mal-H]$^-$, 381[M-Mal-Hex-H]$^-$	$C_{30}H_{46}O_{14}$	0	AAOH-Hex2-Mal (I)[c]	1096±288	2239±124	861±43	1378±246
29.35	629.2810	585[M-CO_2-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-CO_2-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[d]	$C_{25}H_{41}N_3O_8S$	1.0	AAA-GSH-H_2O	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	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)[e]	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-CO_2-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-CO_2-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-CO_2-H]$^-$,545[M-Mal-H]$^-$, 383[M-Hex-H]$^-$,221[M-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-CO_2-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-CO_2-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-CO_2-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-CO_2-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

Non-volatile metabolites with mass intensity higher than 500 in LC-QTOF-MS, which were significantly increased in leaves agro-infiltrated with *AmFH+AMOLAP*, *AmFH+AMOHAP*, *AmFH+AMOLAP+DBR2*, or *AmFH+AMOHAP+DBR2* were targeted for analysis by LC-QTOF-MS/MS fragmentation.

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; GSH: glutathione; (FA): formic acid adduct; Ret (min): retention time, in minutes; Mol form: molecular formula of the metabolite; ΔMass (ppm): deviation between the detected mass and real accurate mass, in ppm; Putative ID: putative identification of metabolite; ND: not detectable.

[a]Detected mass (D): The mass was detected in negative mode of LC-QTOF-MS.

[b]Peak intensities are the mean ± S.D. of three agro-infiltrated leaves.

[c]Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times).

[d]The ions of a number of representative GSH adducts in the negative ion mode (Dieckhaus *et al.*, 2005).

[e]AA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen *et al.*, 2010).

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

(mg kg^{-1} FW)	AmFH+AMOLAP	AmFH+AMOHAP	AmFH+AMOLAP+DBR2	AmFH+AMOHAP+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.

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.

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

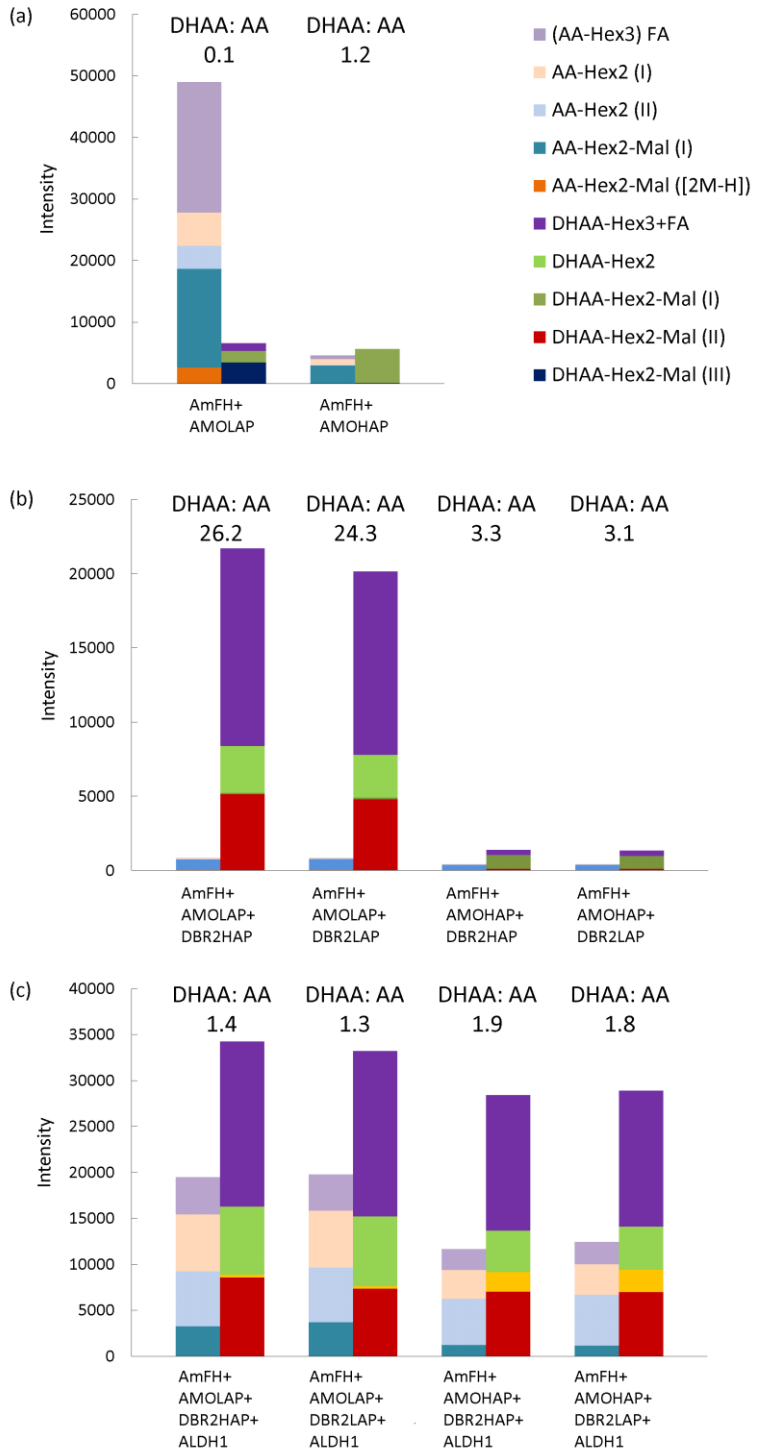

Figure 5. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) related compounds in leaves of *N. benthamiana* agro-infiltrated with artemisinin biosynthesis genes as identified by LC-QTOF-MS.
(a) Agro-infiltrated leaves with *AmFH+AMOLAP* and *AmFH+AMOHAP*. For all products identified by LC-QTOF-MS 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.

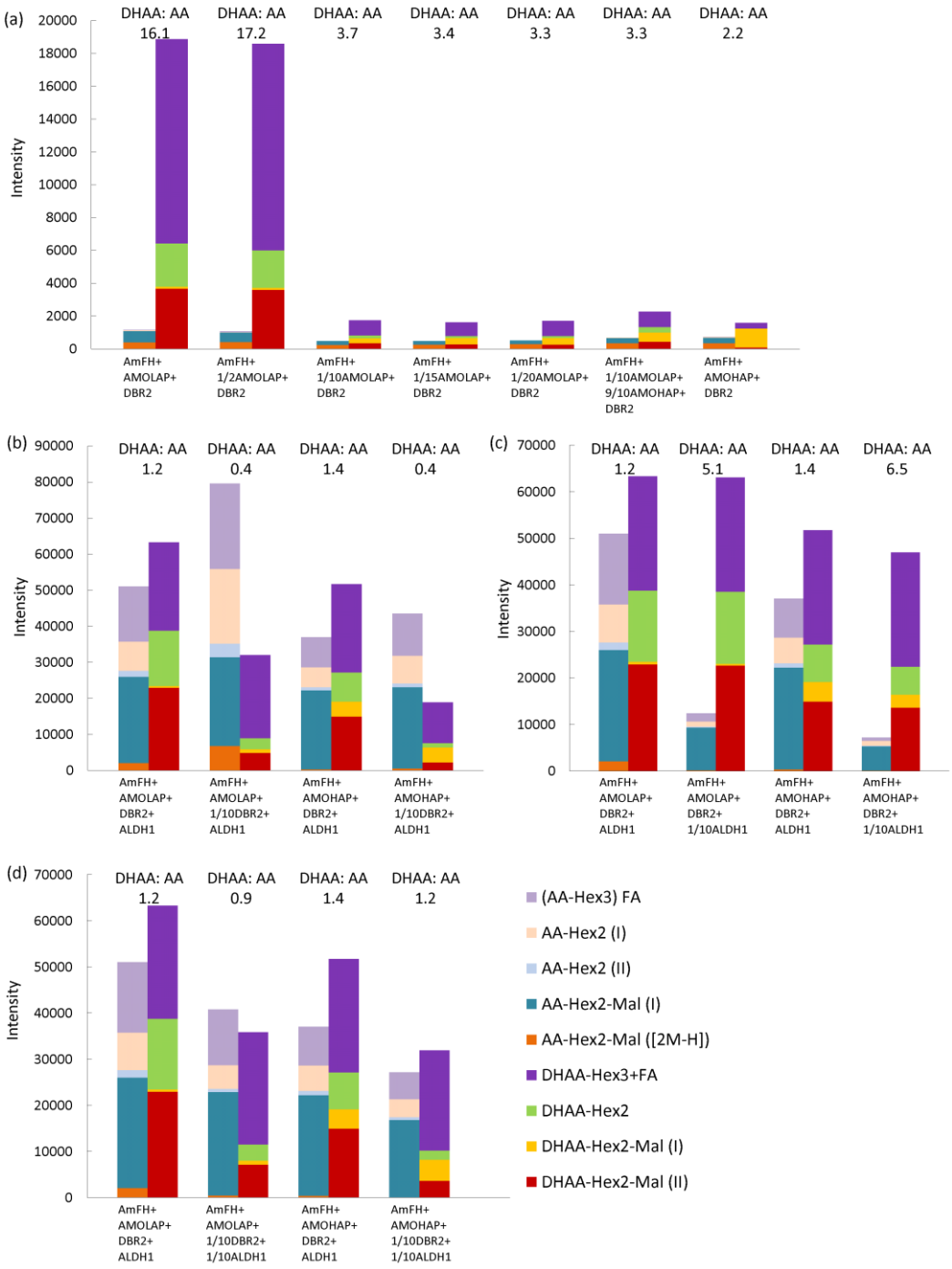

Figure 6. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) conjugated compounds in leaves of *N. benthamiana* agro-infiltrated with AMOLAP/HAP and different dosage of DBR2 and ALDH1.
(a) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of AMOLAP/HAP. For all 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.

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 expressing 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+AMOHAP* 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).

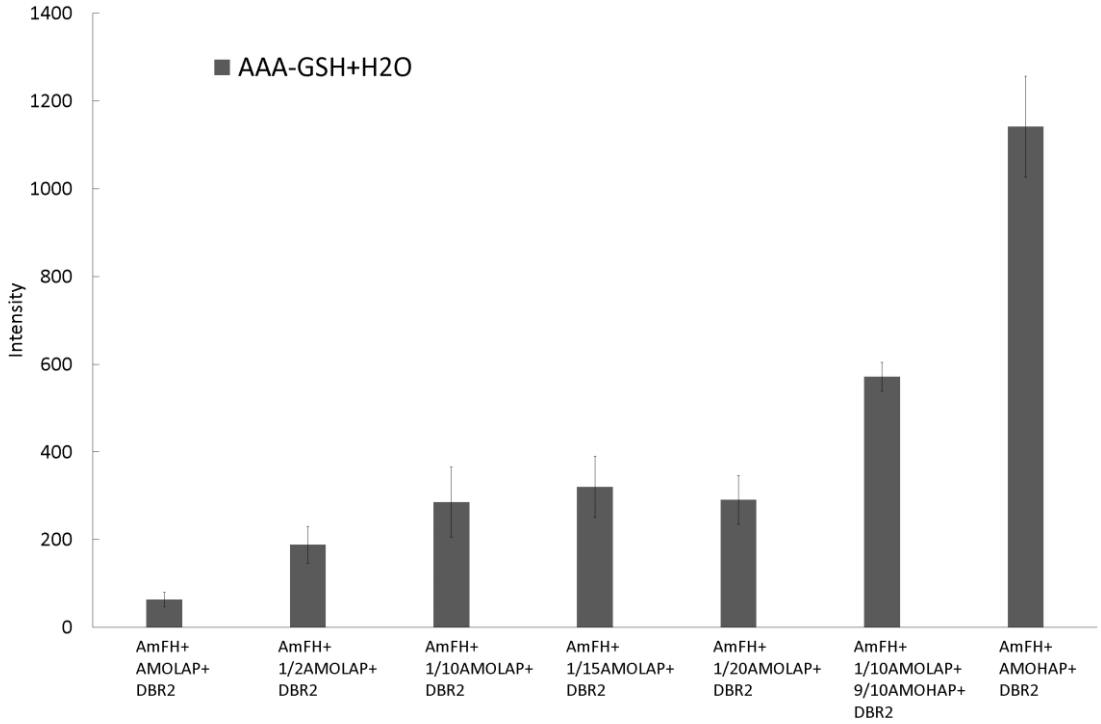

Figure 7. Artemisinic aldehyde **(**AAA) conjugated compounds in leaves of *N. benthamiana* agro-infiltrated with dilution of AMOLAP/HAP.
Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis. Peak intensities are the mean ± S.D. of three agro-infiltrated leaves.

Effects on product profile by DBR2 or ALDH1 dosage

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 AMOHAP 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 combination 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 glutathione 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.

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).

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.

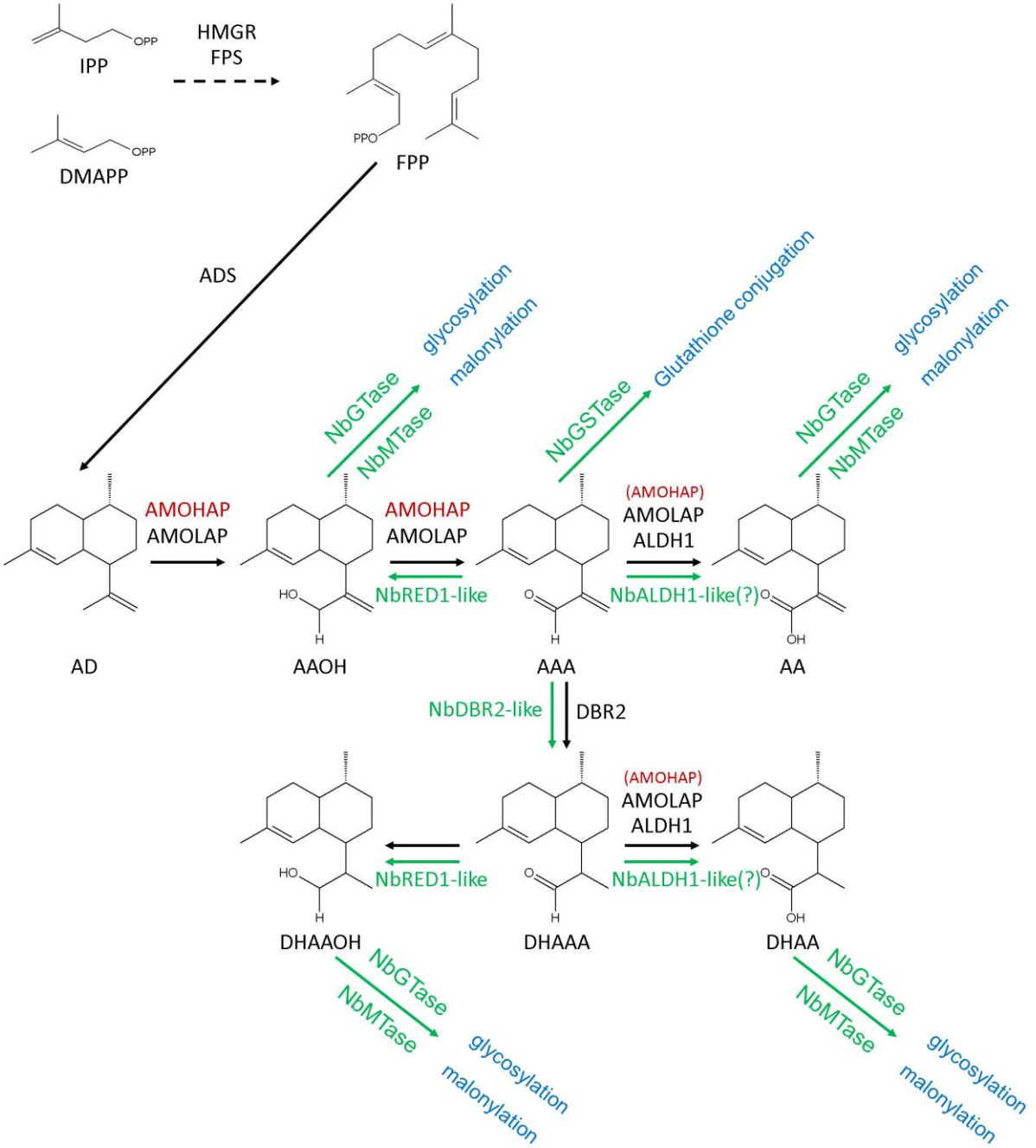

Figure 8. Schematic representation of the artemisinin precursor biosynthetic pathway in *Nicotiana benthamiana*.
NbDBR2-like: *Nicotiana benthamiana* artemisinic aldehyde double-bond reductase-like enzyme; NbRED1-like: *Nicotiana benthamiana* dihydroartemisinic aldehyde reductase 1-like enzyme; NbALDH1-like: *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.

Acknowledgements

H-M.T. was funded by the graduate school of Experimental Plant Sciences (EPS). T.v.H. and J.B. were supported through the NWO-CW/ACTS IBOS programme (053.63.305) which was co-sponsored by Dafra Pharma, Turnhout, Belgium. We thank Ric de Vos for helpful discussions on LC-QTOF-MS/MS data, Bert Schipper for assistance in LC-QTOF-MS analysis, Desalegn Woldes Etalo, Ting Yang and Lemeng Dong for help in MS data analysis and Qing Liu for support in the glutathione conjugation assay. We also would like to acknowledge Sajad Rashidi Manfard and Peter E. Brodelius (Linneaeus University, Sweden) for providing seeds of an *A. annua* LAP chemotype.

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'CGGC*GGATCC*ATGGCACTCTCACTGACCACTTCCA 3'	Bam HI
2	5'CGGC*GGTACC*CTAGAAACTTGGAACGAGTAACAACTCAGCCTTTC 3'	Kpn I
3	5'CGGC*GGATCC*ATGAAGAGTATACTAAAAGCAATGGCACTCTCACTG ACCAC 3'	Bam HI
4	5'CGGC*GGTACC*CTAGAAACTTGGAACGAGTAACAACTCAGTCTTTC 3'	Kpn I
5	5' AT*GGATCC*GTCTGAAAAACCAACCTTG3'	Bam HI
6	5' TATA*GCGGCCGC*TAGAGGAGTGACCCTT3'	Not I
7	5'GACAAA*TCTAGA*AAGATGAGCTCAGGAGCTAATGGAAG 3'	XbaI
8	5'CACAAA*GCGGCCGC*TTAAAGCCACGGGGAATCATA 3'	Not I
9	5'AT*GGATCC*ATGAAGAGTATACTAAAAGCAATGGCACTCTCACTGA3'	Bam HI
10	5'TAT*GGTACC*GTGACCAA TAATGGGAAGCCGCCAT 3'	Kpn I
11	5'AT*GGATCC*ATGGCACTCTCACTGACCACTTCCATTGCTCT3'	Bam HI
12	5'TAT*GGTACC*GTGACCAATAATGGGAAGTCGCCATGGCTC3'	Kpn I
13	5'GCC*GGTACC*ATGAGCAAGGGCGAGGAGCTGTTCA3'	Kpn I
14	5'TTAATT*GCGGCCGC*TTATTCCACATCTTTATACAGCTCGTCCATG3'	Not I
15	5'GACAGTGCT*GAATTC*CCATTAACATCTGATAACATTAAAG3'	EcoR I
16	5'CTCGCCCTTGCTCAT*GGTACC*CAAGAAACTTGGAACGAGT3'	Kpn I
17	5'ACAGTGCT*GAATTC*CCATTAACATCTGATAACATTAAAGC3'	EcoR I
18	5'CCTCGCCCTTGCTCAT*GGTACC*CAAGAAACTTGGAACGAG3'	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.

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[a]							
	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)[b]	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	68	67	1168	968	98	98	1055	1135
(AA-Hex3) FA	10	10	ND	ND	4923	5101	3318	3231
AA-Hex2 (I)[c]	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

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; ND: not detectable; GSH: glutathione; (FA): formic acid adduct.

[a]Peak intensities are the mean of three agro-infiltrated leaves.

[b]Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times).

[c]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 *AmFH+AMOLAP/HAP+DBR2*compared with different dilution of *AMOLAP/HAP* were analysed by LC-QTOF-MS.

Putative ID	Intensity[a]						
	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)[b]	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_2O	64	189	286	321	291	572	1142
(AA-Hex3) FA	39	80	ND	ND	ND	ND	3
AA-Hex2 (I)[c]	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

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; ND: not detectable; GSH: glutathione; (FA): formic acid adduct.

[a]Peak intensities are the mean of three agro-infiltrated leaves.

[b]Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times).

[c]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[a]							
	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)[b]	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	180	842	194	570	2018	4006	2127	4995
(AA-Hex3) FA	15298	23704	1764	12096	8416	11777	733	5854
AA-Hex2 (I)[c]	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

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; ND: not detectable; GSH: glutathione; (FA): formic acid adduct.

[a]Peak intensities are the mean of three agro-infiltrated leaves.

[b]Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times).

[c]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.

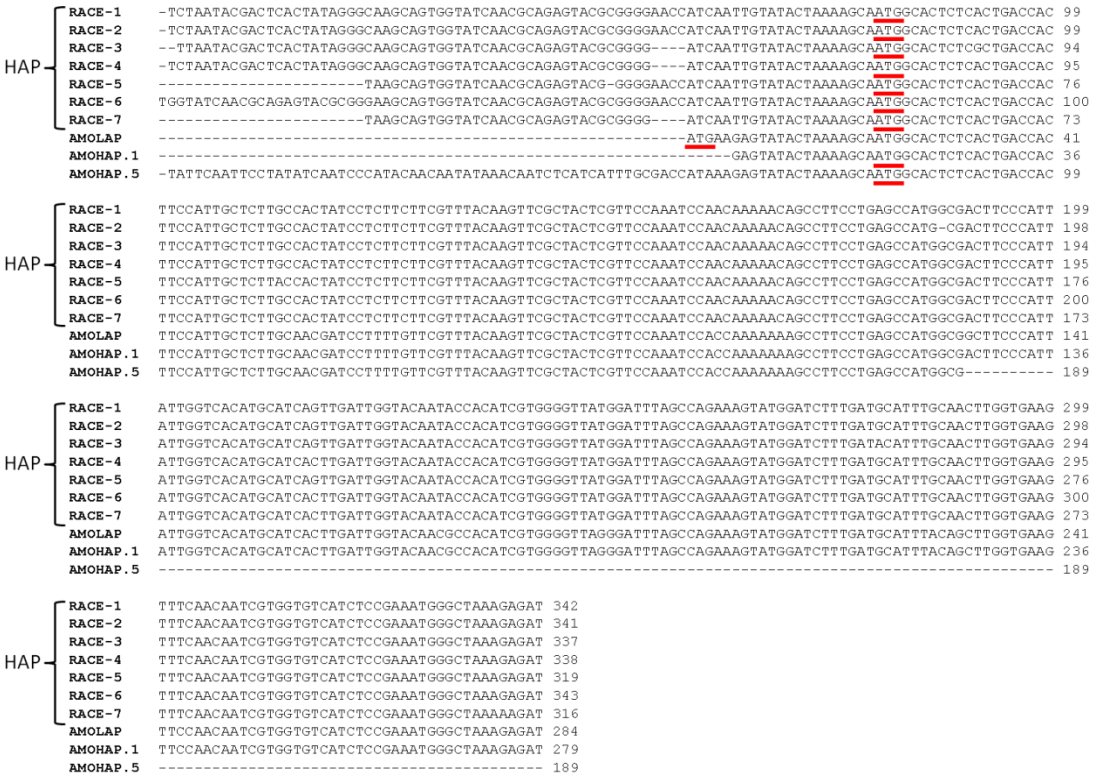

Figure S1. Alignment of the 5'RACE sequences of *AMOHAP* in Clustal W.
Seven 5'RACE clones from a HAP chemotype are aligned with *AMOLAP* (DQ268763), *AMOHAP.1* (DQ315671) and the recently published *CYP71AV1* promoter isolated from an *Artemisia annua* HAP chemotype, here referred to as *AMOHAP.5* (Wang *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.

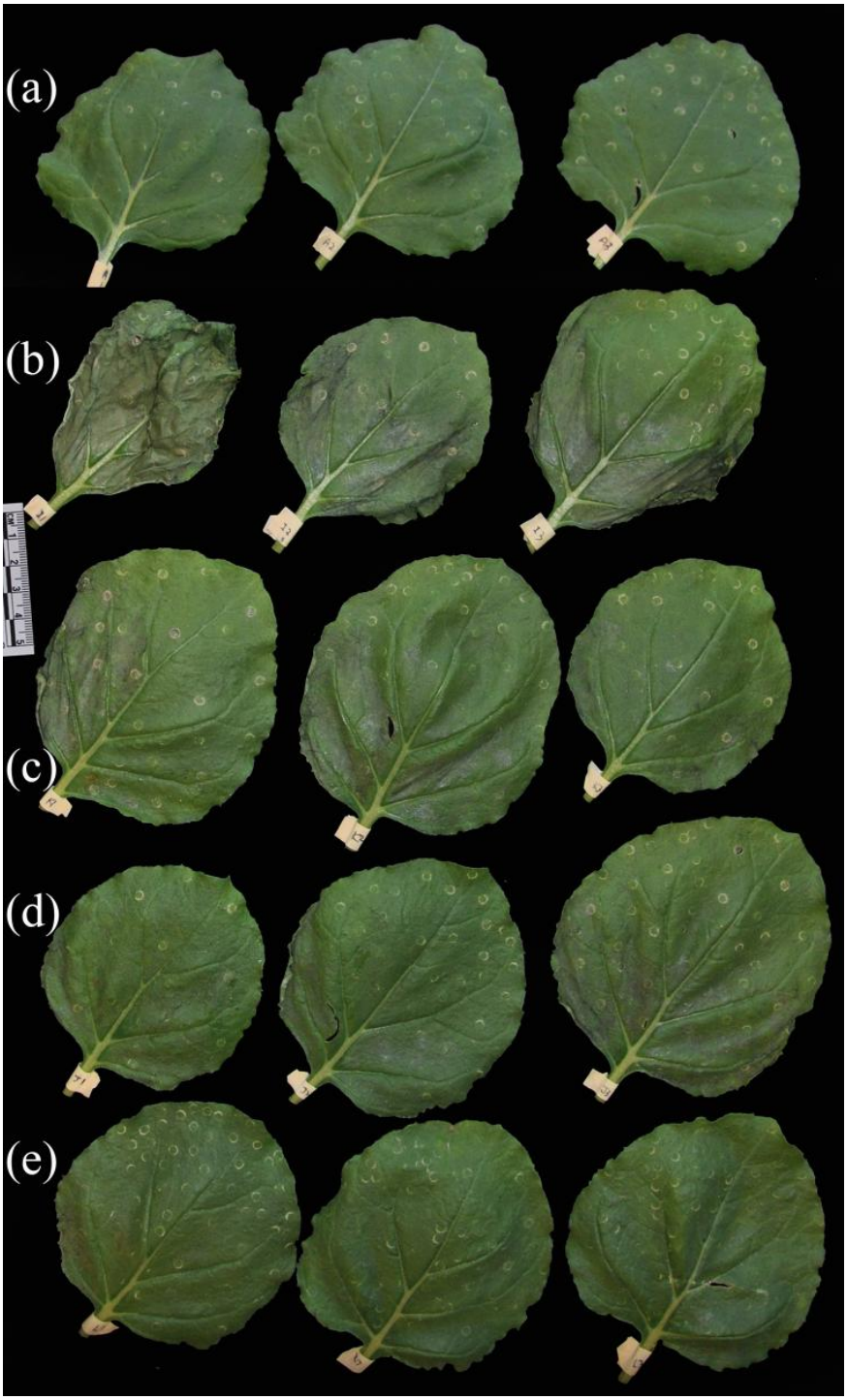

Figure S2. Necrotic symptoms in *Nicotiana benthamiana*.
Leaves agro-infiltrated with (a) pBin empty vector, (b) *AmFH+AMOLAP*, (c) *AmFH+AMOLAP+DBR2*, (d) *AmFH+AMOHAP*, and (e) *AmFH+AMOHAP+DBR2* seven days after infiltration.

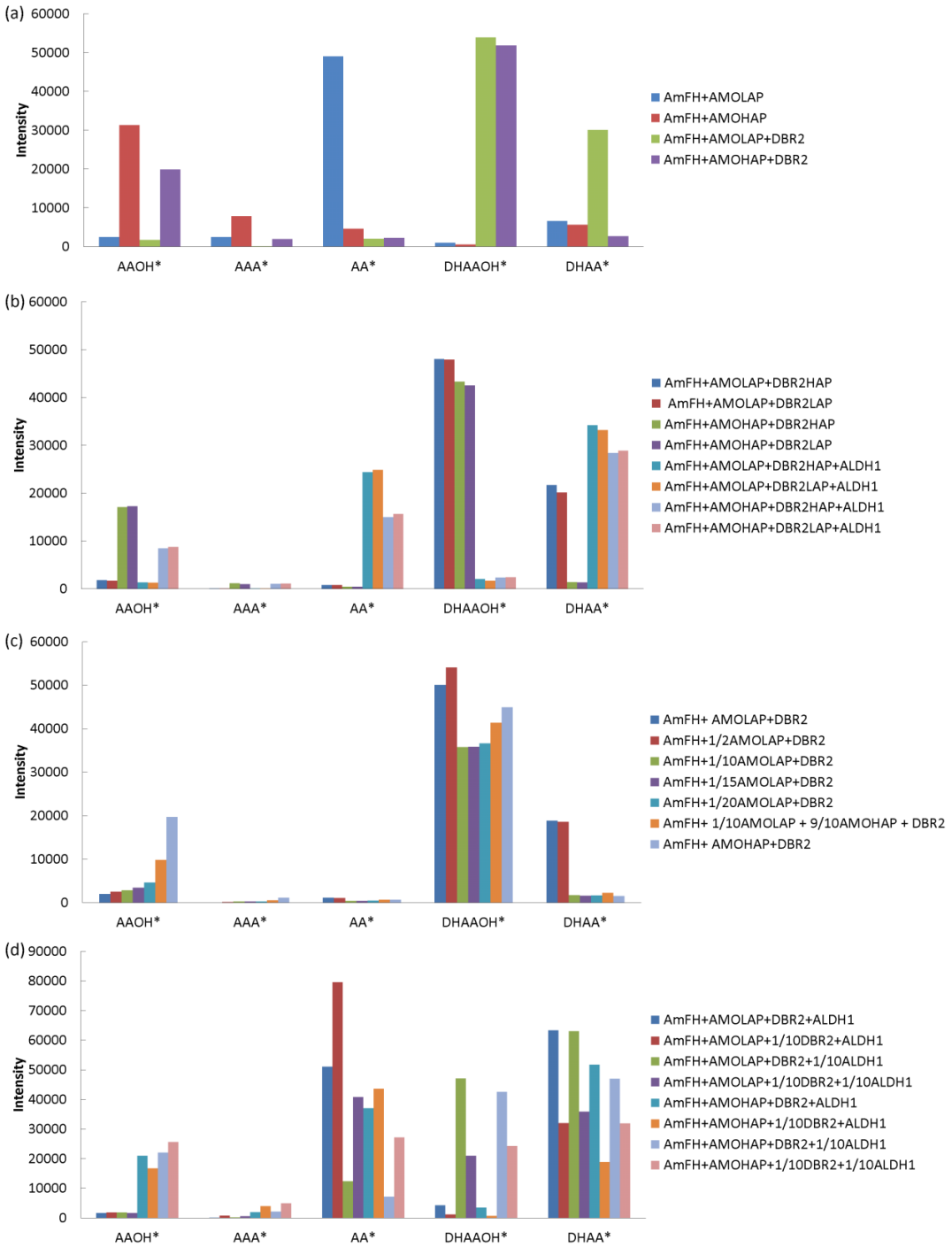

Figure S3. All conjugated compounds grouped according to free precursor detected in agro-infiltrated *Nicotiana benthamiana*.
(a) Agro-infiltrated leaves with *AmFH+AMOLAP/HAP* and *AmFH+AMOLAP/HAP+DBR2*. For all products identified by LC-QTOF-MS see Table 2.
(b) Agro-infiltrated leaves with *AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1*. 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)

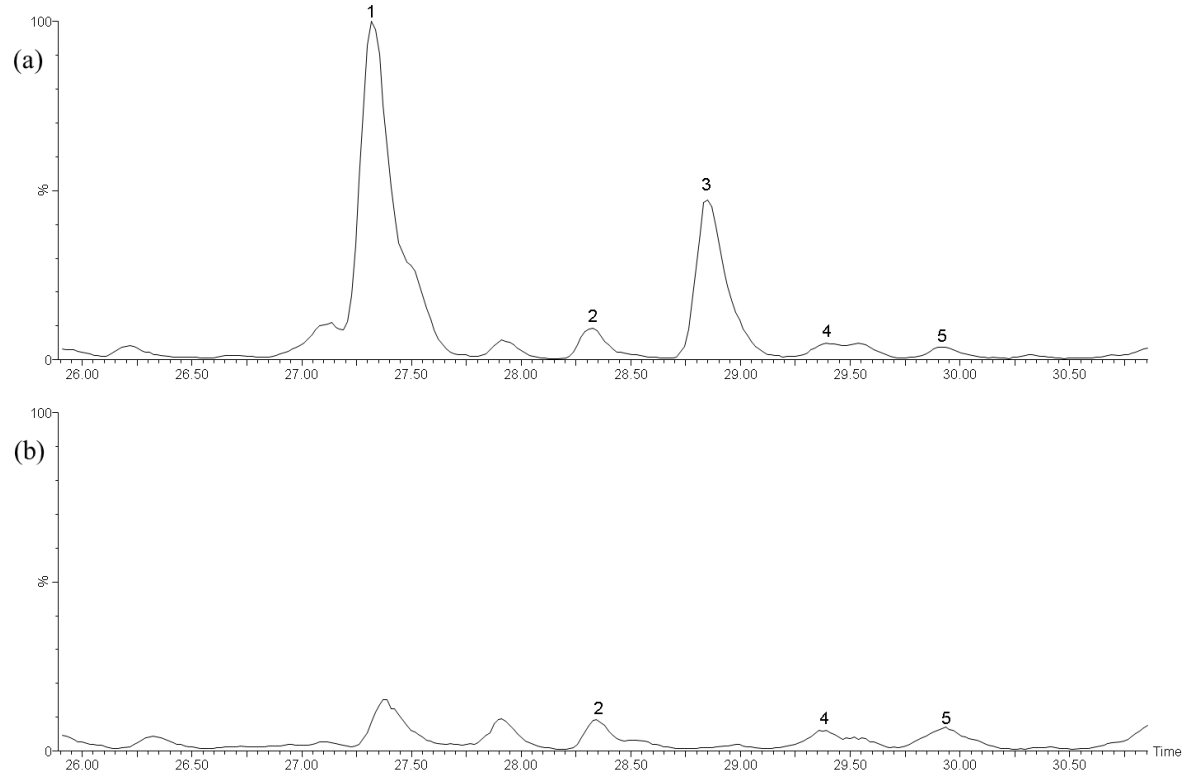

Figure S4. Malonylated glycosylated compounds are not cleaved by Viscozyme L.
(a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+DBR2* without Viscozyme-treatment.
(b) LC-QTOF-MS chromatogram (total ion count) of methanol 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)

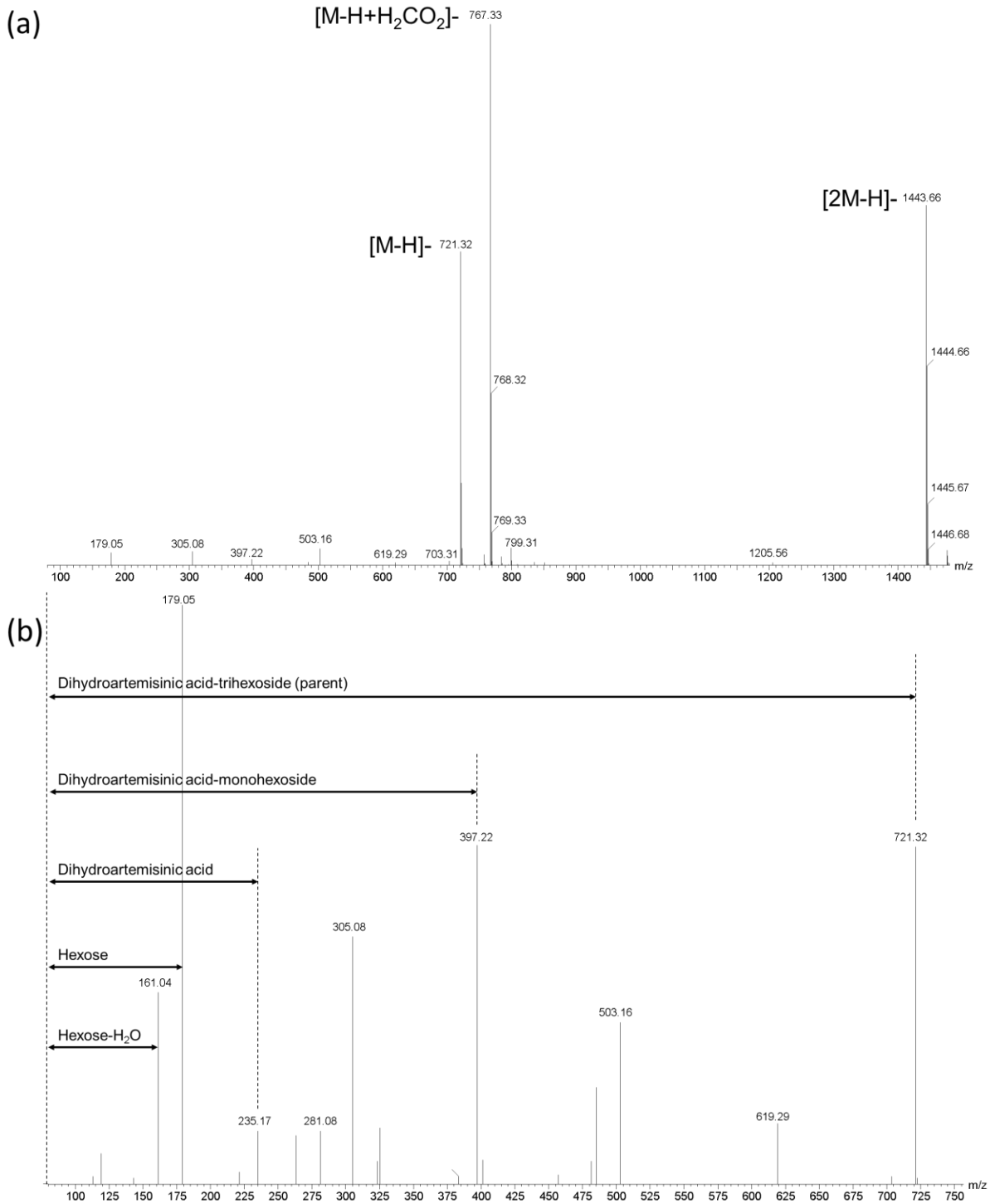

Figure S5. MS/MS spectrum and MS/MS fragmentation of DHAA-Hex3.
(a) MS/MS spectrum of the compound eluting at 23.87 min, DHAA-Hex3 (m/z 721.32;$[M-H]^-$).
(b) MS/MS fragmentation of mass 721.32 eluting at 23.87 min.
(DHAA: dihydroartemisinic acid, Hex: compound conjugated with hexose)

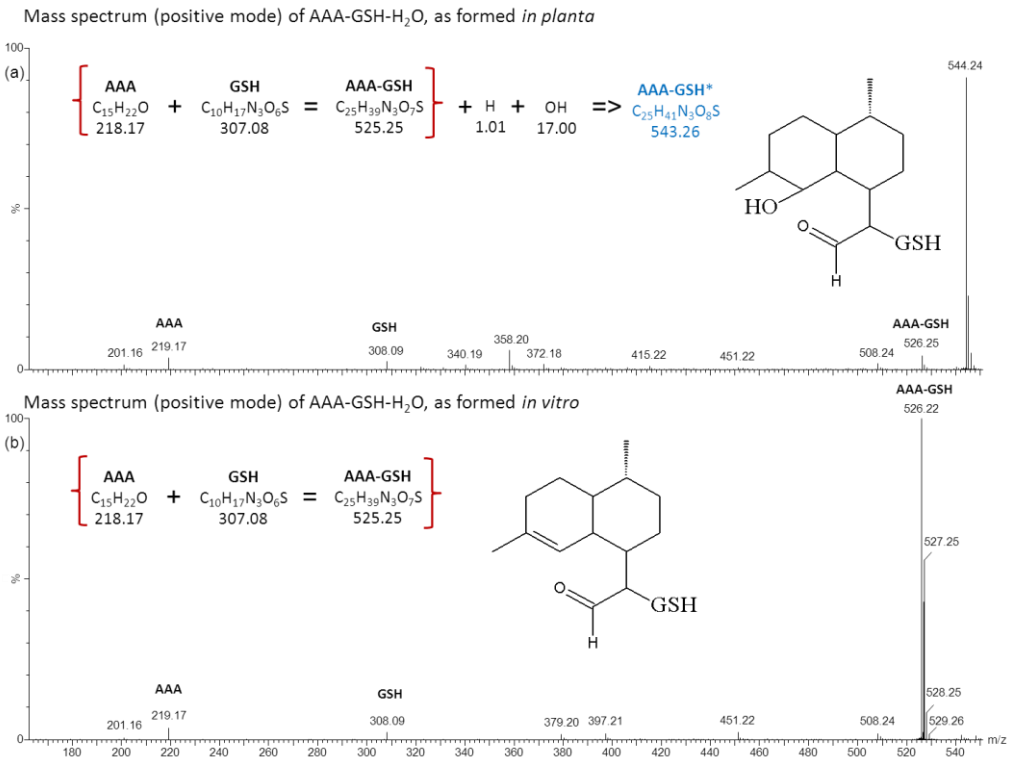

Figure S6. Comparison of *in vitro* and *in planta* formed AAA-GSH conjugates.
(a) Mass spectrum (positive mode) of AAA-GSH*, as formed *in planta* (Table 2). The mass suggests a AAA-GSH derivative with 18 D mass increase, suggesting an additional OH and H. This is most easily explained by a reduction 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)

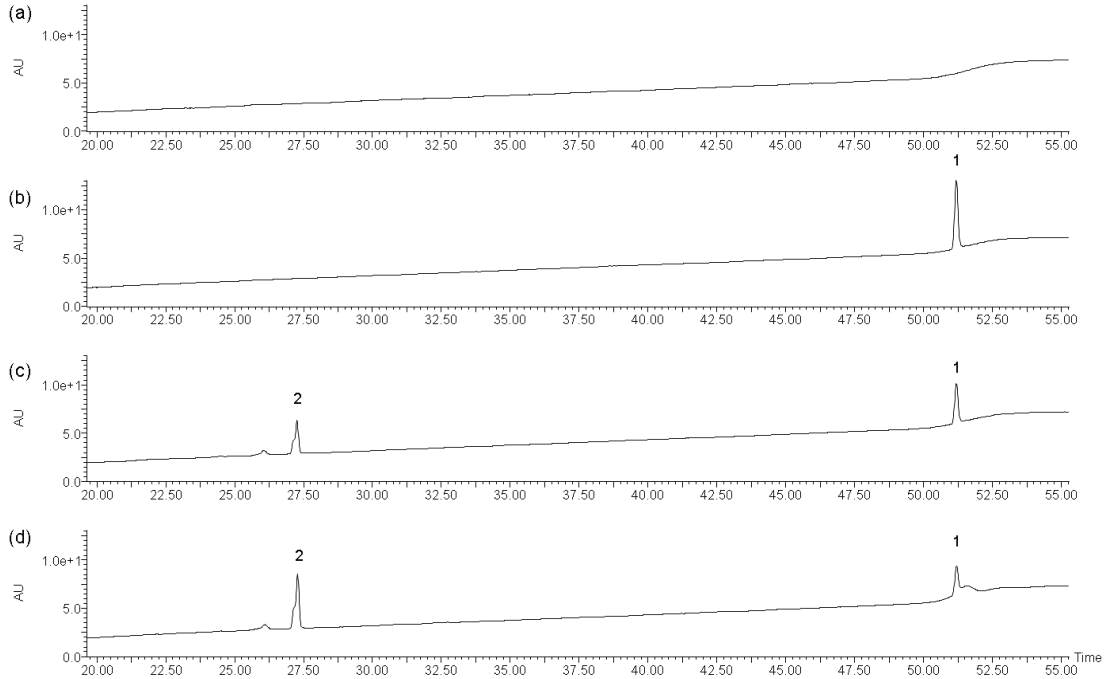

Figure S7. *In vitro* conjugation of artemisinic aldehyde (AAA) with glutathione.
LC-QTOF-MS analysis of (a) Glutathione (GSH) incubated with Glutathione S transferase (GST), (b) AAA+GST, (c) AAA+GSH and (d)AAA+GSH+GST.

Peak 1: AAA, Peak 2: AAA-GSH conjugate.

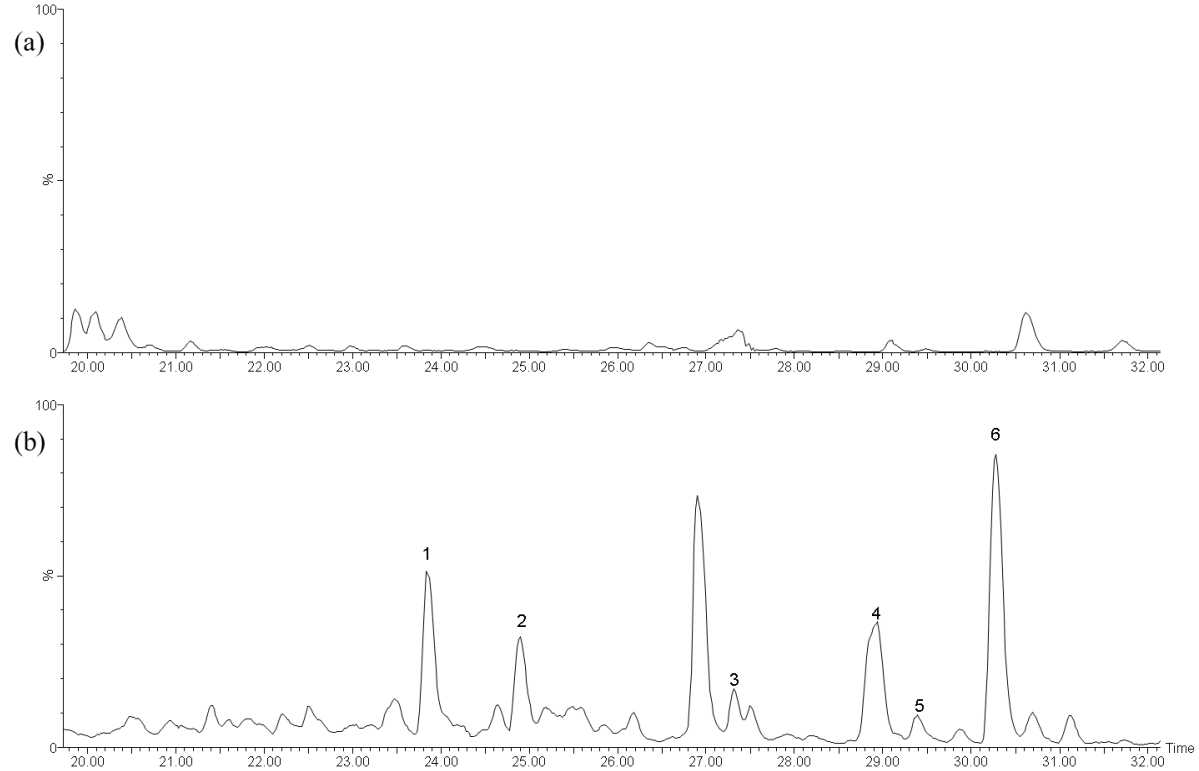

Figure S8. Absence of artemisinin precursor conjugates in *Artemisia annua* HAP chemotype.
(a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of flowers of *Artemisia annua* HAP chemotype.
(b) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+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)

Chapter **3**

Experiments to address the potential role of LTPs in sesquiterpene emission

Hieng-Ming Ting[1], Kim Boutilier[2], Francel W.A. Verstappen[1], Mariëlle Schreuder[1], Harro J. Bouwmeester[1] and Alexander van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, P.O. Box 658, 6700 AR Wageningen, The Netherlands.

[2]Plant Research International, P.O. Box 619, 6700 AP Wageningen, The Netherlands.

Abstract

In plants, there often is a strong correlation between high terpenoid biosynthesis activity and high expression of *lipid transfer protein* (*LTP*) genes. Here, we investigated the role of *LTP* genes in sesquiterpene emission in *Artemisia annua* and Arabidopsis. Three *LTP* genes were isolated from an *A. annua* trichome library. Subcellular localisation studies of the LTP proteins fused to GFP showed that these proteins are secreted and accumulate in the apoplast. The *A. annua LTP* genes were transiently co-expressed in *Nicotiana benthamiana* with the genes of the artemisinin biosynthesis pathway but failed to change the product profile. Subsequently, *AaLTP-RNAi* constructs were used to transform *A. annua* plants but transgenic lines were not obtained. Finally, the *AaLTPs* were cloned into expression vectors and used to transform Arabidiopsis. Arabidopsis does not make the sesquiterpene artemisinin, but flowers of Arabidopsis do emit the sesquiterpene caryophyllene. The AaLTPs did not increase caryophyllene emission in Arabidopsis while in one line emission was even reduced. Publicly available Arabidopsis expression data were screened for *LTPs* with floral expression and knockout mutants of these *LTPs* were screened for changes in floral volatile emission, but without success. Homology with the *A. annua LTPs* was used to select two Arabidopsis homologs (*LTP1* and *LTP3*) for a more detailed investigation. When knockout mutants of these *LTPs* were analysed for floral volatile emission. Arabidopsis *ltp1* and *ltp3* mutants both showed a 20% reduction in caryophyllene in the flower headspace, while emission of other floral volatiles was not affected. This suggests that these LTPs play a role in sesquiterpene emission. However, detailed expression profiling of *LTP1*, *LTP3*, and the sesquiTerPene Synthases, *TPS21* and *TPS11* using GUS reporter lines, showed only partial overlap in spatial expression pattern. We tried, but failed, to obtain the *ltp1 ltp3* homozygous double mutant, most likely due to chromosomal translocations in the *ltp1* insertion mutant. While the results suggest that LTPs are involved in sesquiterpene emission in Arabidopsis, more research is needed to obtain definitive proof.

Keywords: caryophyllene, lipid transfer protein (LTP), terpene synthase (TPS).

Introduction

Terpenoids are the most structurally varied class of secondary metabolites in plants. They are important in signalling between plants and in the interaction of plants with and defence against other organisms. Several studies have demonstrated an association between terpene biosynthesis and lipid transfer proteins (LTPs). For instance, LTP sequences are highly represented in cDNA libraries of cells which are active in terpenoid secretion (Bertea *et al.*, 2006, Harada *et al.*, 2010, Lange *et al.*, 2000, Schilmiller *et al.*, 2010). A high level of expression of LTP genes has also been observed in glandular trichomes of the medicinal plant *Artemisisa annua*, which is known for its production of the anti-malaria sesquiterpene artemisinin (Bertea *et al.*, 2006). Also, in plant-insect and plant-pathogen interactions the induction and emission of defensive terpenes seems to coincide with the expression of *LTPs*: insect feeding in Sitka spruce (*Picea sitchensis*) leads to induced terpene production as well as *LTP* expression (Friedmann *et al.*, 2007), and infection of rice by the rice blast fungus (*Magnaporthe grisea*) induces diterpene phytoalexin biosynthesis, which also coincides with induced *LTP* expression (Xiong *et al.*, 2001). Recently, *NtLTP1*, which is expressed in the glandular trichomes of tobacco, was shown to have lipid-binding activity. Changes in its expression through silencing or over-expression correlated with differences in diterpene secretion (Choi *et al.*, 2012).

Lipid transfer proteins (LTPs) are small ubiquitous and highly abundant plant proteins (~9 kDa) that have a high pI (~9) and contain eight cysteine residues at conserved positions. Three-dimensional structure analysis of several LTPs showed that LTPs have a hydrophobic cavity that is enclosed by four α-helices held together by four disulfide bonds between the eight cysteine residues (Charvolin *et al.*, 1999, Heinemann *et al.*, 1996, Shin *et al.*, 1995). LTPs can bind fatty acids and are characterized by their ability to transfer phospholipids between membranes in *in vitro* assays (Kader, 1996). Besides the few characteristic conserved features, LTPs are highly divergent within and between plant species. For example, 71 putative LTPs with highly divergent sequences and expression profiles have been identified in *Arabidopsis thaliana* (Arabidopsis) (Beisson *et al.*, 2003). The diversity in protein sequence and expression profiles of the individual members of the LTP family is reflected in the wide range of roles attributed to these proteins, but for many LTPs it is not known how they fulfil these functions (Arondel *et al.*, 2000, Kader, 1996, Ng *et al.*, 2012, Nieuwland *et al.*, 2005, Wang *et al.*, 2012). The coincidence of high terpene production and high LTP activity may be explained in two ways: LTPs may be involved in transport of lipid

molecules that provide a sink for the hydrophobic terpenes, or LTPs may act directly as carriers for terpenes through binding of terpenes to the hydrophobic cavity of LTP proteins.
In Arabidopsis, terpenes are mainly emitted from flowers, and recently it was shown that emission of the sesquiterpene (*E*)-β-caryophyllene is involved in protection against bacterial infection in flowers (Chen *et al.*, 2003, Huang *et al.*, 2012). In Arabidopsis flowers, terpene synthase 21, TPS21, is responsible for the biosynthesis of (*E*)-β-caryophyllene, α-humulene and α-copaene, while terpene synthase 11, TPS11, is responsible for biosynthesis of the remaining floral sesquiterpenes (α-barbatene, thujopsene, iso-bazzanene, β-barbatene, *E*-β-farnesene, β-acoradiene, β-chamigrene, α-zingiberene, α-cuprenene, α-chamigrene, cuparene, β-bisabolene, β-sesquiphellandrene, δ-cuprenene) (Tholl *et al.*, 2005). In addition, Arabidopsis flowers emit the monoterpenes β-myrcene, limonene and linalool, of which the formation is catalysed by terpene synthase 3, TPS03; terpene synthase 10, TPS10; terpene synthase 14, TPS14 and terpene synthase 24, TPS24 (Chen *et al.*, 2011, Chen *et al.*, 2003).

Here we used *Artemisia annua* and Arabidopsis to investigate the relation between terpenes and LTPs and to determine whether LTPs play a role in the production and/or emission of terpenes in plants. Three LTP genes were isolated from an *A. annua* trichome cDNA library and the subcellular targeting of these proteins was determined in a transient expression assay. The *A. annua LTP* genes were coexpressed with the genes of the artemisinin biosynthesis pathway but failed to change the product profile. Because we failed to obtain *A. annua* plants transformed with an *AaLTP-RNAi* construct we continued the investigation of the role of LTPs in terpene transport with the model plant Arabidopsis. A preliminary screen of twelve Arabidopsis *ltp* knockout mutants did not show an effect on headspace emission by pooled mutant inflorescences. Because LTP1 and LTP3 from Arabidopsis showed the highest homology to the LTPs from *A. annua*, we investigated the *ltp1* and *ltp3* T-DNA insertion mutants in more detail. Both mutants displayed about 20% reduction in caryophyllene emission in open flowers. Moreover, *ltp1* also showed a 40% reduction in caryophyllene accumulation in flower extracts. Although the reduction in emission is small, the results suggested a role for LTP proteins in terpene production and/or emission in flowers, particularly considering the possibility that LTPs may be redundant. A detailed expression analysis of *LTP1*, *LTP3*, and the two *TPS* genes, *TPS11* and *TPS21* that are mainly expressed in flowers showed that these *LTP* and *TPS* genes show partially overlapping and partially unique expression patterns, suggesting that if there is a role for LTPs in terpene emissions, that this is most likely not the only function of LTPs.

Materials and Methods

Isolation of LTP genes from Artemisia annua

It had been shown that 10% of *LTPs* were expressed in the glandular trichome cDNA library of *A. annua* (Bertea *et al.*, 2006). Three partial *LTP* candidate genes were expressed in that glandular trichome library. 5'-RACE and 3'-RACE were carried out using the SMART-RACE cDNA Amplification Kit (Clontech, USA) according to the manufacturer's protocol. SeqMan software was used to assemble the sequences of 5'-RACE and 3'-RACE products resulting in the full-length cDNA sequences of *AaLTP1*, *AaLTP2* and *AaLTP3* (Table S1).

Vector construction

35S:AaLTP1, 35S:*AaLTP2* and 35S:*AaLTP3*. The three *AaLTPs* were amplified from cDNA of *A. annua* flower using the primer pair AaLTP1-F/AaLTP1-R, AaLTP2-F/AaLTP2-R and AaLTP3-F/AaLTP3-R, respectively (Table S2). The primers introduced NcoI and NotI restriction sites which were used for cloning into ImpactVectorpIV1A_2.1, which contains the CaMV35S promoter and Rbcs1 terminator (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). The resulting pIV1A_2.1/AaLTP1, pIV1A_2.1/AaLTP2 and pIV1A_2.1/AaLTP3 were cloned into the pBinPlus binary vector using LR recombination (Invitrogen) (van Engelen et al., 1995). The pBinPlus construct containing the 35S:AaLTP1, 35S:AaLTP2 and 35S:AaLTP3 were transferred to *Agrobacterium tumefaciens* AGL-0 using electroporation.

35S:*AaLTP1-RNAi* and 35S:*AaLTP2-RNAi*. To construct an *AaLTP1-RNAi* and an *AaLTP2-RNAi* construct, DNA fragments were amplified using the primer pair attB1-AaLTP1-F/attB2-AaLTP1-R and attB1-AaLTP2-F/attB2-AaLTP2-R (Table S2) by using the pIV1A_2.1/AaLTP1 and pIV1A_2.1/AaLTP2 vector as template. The amplicons were subcloned into the pDONR221 vector (Invitrogen) using the BP clonase II enzyme mix (Invitrogen) to generate entry vectors pDONR221-*AaLTP1* and pDONR221-*AaLTP2*. In a subsequent step, the RNAi fragments were cloned into the pHELLSGATE8 vector through LR recombination (Invitrogen). The vector was transferred to *Agrobacterium tumefaciens* AGL-0 by electroporation.

35S:AaLTP1-GFP, 35S:*AaLTP2-GFP* and 35S:*AaLTP3-GFP*. The GFP coding sequence was amplified by PCR using primers EGFP_c-term-F/EGFP_c-term-R (Table S2) by using pBin-Egfp 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 cloning into ImpactVectorpIV1A_2.1 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 pair AaLTP1_gfp-F/AaLTP1_gfp-R, AaLTP2_gfp-F/AaLTP2_gfp-R and AaLTP3_gfp-F/AaLTP3_gfp-R (Table S2) 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 by 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 were cloned into the pBinPlus binary vector using LR recombination (Invitrogen) (van Engelen et al., 1995). The pBinPlus constructs containing the 35S:AaLTP1-GFP, 35S:AaLTP2-GFP and 35S:AaLTP3-GFP were transferred to *Agrobacterium tumefaciens* AGL-0 using electroporation.

Subcellular localisation studies of AaLTP-GFP proteins in transient expression assays

Individual *A. tumefaciens* strains carrying 35S:AaLTP1-GFP, 35S:AaLTP2-GFP and 35S:AaLTP3-GFP constructs were infiltrated into *Nicotiana benthamiana* leaves, which were harvested after seven days for confocal laser scanning as described (Dong *et al.*, 2013).

Quantification of artemisinin precursors by UPLC-MRM-MS

Targeted analysis of artemisinin precursors in agro-infiltrated *N. 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 previously described (Ting *et al.*, 2013).

Transformation of Artemisia annua

A. annua transformation with *AaLTP1-RNAi* and *AaLTP3-RNAi* was carried out in the Laboratory of Prof. Peter Brodelius (Linnaeus University, Sweden) as previously described (Wang *et al.*, 2011).

Transformation of Arabidopsis

35S:AaLTP1, 35S:AaLTP2 and 35S:AaLTP3 were each transformed into Arabidopsis (Col-0) using the *A. tumefaciens*-mediated floral dip method as described (Zhang *et al.*, 2006).

Transformed seeds were selected on Murashige and Skoog (MS) medium for kanamycin resistance and confirmed by PCR. Three independent homozygous lines of each constructs of the T3 generation were randomly chosen for further analysis.

Plant materials and growing conditions

All Arabidopsis lines are in the Col-0 and Ws background. Seeds of Arabidopsis T-DNA insertion mutant lines (At2g38540/LTP1 [SALK_134262], At5g59320/LTP3 [SALK_095248], At2g33470/GLTP [SALK_014537], At1g27950/LTP-like [SALK_072495], At2g27130/LTP-like [SALK_119487], At2g10940/LTP-like [SALK_083118], At1g70250/LTP-like [SALK_142707], At1g48750/LTP-like [SALK_144344], At5g44630/TPS11 [SALK_126868] and At5g23960/TPS21 [SALK_138212]) were obtained from NASC (Nottingham Arabidopsis Stock Centre). At3g08770/LTP6 (FLAG_137H03), At4g15160/LTP-like (FLAG_103F11) and At1g05450/LTP-like (FLAG_456G06) in the Ws background were obtained from PublicLines at Institut National de la Recherche Agronomique (http://dbsgap.versailles.inra.fr/publiclines/). At1g62790/LTP-like (GABI_173C09, Col-0 ecotype) was obtained from the GABI-Kat FST population (http://www.gabi-kat.de). Lines carrying the *LTP1::GUS* and *LTP3::GUS* transcriptional fusions were kindly provided by Dr. Keun Chae (Chae *et al.*, 2010). *TPS11::GUS* and *TPS21::GUS* lines were kindly provided by Dr. Dorothea Tholl (Tholl *et al.*, 2005). Plants were grown in soil in a greenhouse with a 16h light/8h dark photoperiod at 22°C.

Genotyping of T-DNA insertion lines

Genomic DNA was extracted from leaves of T-DNA insertion lines as in (Edwards *et al.*, 1991). The presence of the T-DNA insertion was verified by PCR on genomic DNA using the corresponding specific forward primer (LP), reverse primer (RP), and T-DNA specific primer (LB) (Supplementary Table S3).

Headspace trapping and analysis of volatiles by GC-MS

The flower volatile headspace was analysed by placing inflorescences in a vial with water to prevent dehydration, and placing the vial in a container from which volatiles were collected for 4 hours (between 11:00 am and 15:00 pm) using dynamic headspace trapping as described (Houshyani *et al.*, 2013). The headspace of isolated flowers was obtained from ten freshly opened flowers [floral stage 15; (Smyth *et al.*, 1990)] placed in a 20 ml glass vial from which volatiles were collected for 4 hours (between 11:00 am and 15:00 pm). The headspace

samples were analysed by GC-MS as described (Houshyani *et al.*, 2013). Terpenoids were identified by comparison of mass spectra and retention time with authentic standards: β-myrcene (Aldrich), limonene (Janssen Chimica) and caryophyllene (Sigma-Aldrich). Quantification of terpene levels was accomplished by determining the peak area of the characteristic *m/z* (69 and 93) for β-myrcene and limonene and the characteristic *m/z* (133) for caryophyllene and comparison with calibration curves.

GC-MS analysis of floral extracts

The terpene content of Arabidopsis flowers was measured by GC-MS analysis of extracts obtained by grinding 300 mg of flowers (stage 15) in liquid nitrogen followed by extraction with 2 ml dichloromethane. Water was removed from the extract using anhydrous Na_2SO_4 and the samples were analysed by gas chromatography (7890A, Agilent) on a 30-m x 0.25-mm i.d., 0.25-mm film thickness column with 5-m guard column (Zebron ZB5-MS, Phenomenex) coupled to a mass selective detector (model 5965c, Agilent). The gas chromatograph was programmed at an initial temperature of 45°C for 1 min, with a ramp of 10°C min^{-1} to 300°C. The injector was used in splitless mode and injection port temperature was 250°C. The He inlet pressure was controlled with electronic pressure control to achieve a constant column flow of 1.0 mL min^{-1}. The levels of caryophyllene were determined by comparison of peak area, mass spectra and retention time with authentic caryophyllene (Sigma-Aldrich). Variations in processing and injection volume were corrected by using an internal standard, *p*-cymene (Acros).

GUS staining assay

GUS activity staining was performed in 20 mM sodium phosphate buffer (pH 7.5) containing 0.1% v/v Triton X-100, 10 mM EDTA, 1 mM X-Gluc, and either 0.5 (*TPS21*) or 2.5 mM (*TPS11*, *LTP1* and *LTP3*) of each ferri- and ferrocyanide. A high concentration of iron was used to limit diffusion of indoyl intermediates of the GUS reaction in case of high levels of GUS expression (Block&Debrouwer, 1992). Samples were stained overnight and either examined directly after clearing in 70% ethanol or after fixation in a solution of 3:1 ethanol: acetic acid.

Results

Cloning and characterisation of three Artemisia annua LTP genes

Three *A. annua LTP* sequences were identified in an *A. annua* trichome cDNA library and the three cDNAs were isolated using gene specific primer pairs. The sequences of the *A. annua LTPs* (*AaLTPs*) were compared with those of Arabidopsis *LTPs* (*AtLTP*s). Phylogenetic analysis showed that *AaLTP1* is distinct from any homolog in Arabidopsis, while *AaLTP2* and *AaLTP3* are in the same clade, but are still distinct from, *AtLTP1* and *AtLTP3* (Fig. 1). LTPs have been shown to be located inside the cell as well as in the extracellular space (Carvalho *et al.*, 2004, Pagnussat *et al.*, 2012, Pagnussat *et al.*, 2009). All three *AaLTP* genes have a putative N-terminal secretion signal, suggesting that they are secreted to the apoplast in plants. To determine the subcellular localisation of the AaLTPs, GFP-fusion expression constructs were made and the subcellular localisation studied by transient expression of the AaLTP-GFP protein fusions in *N. benthamiana* leaves. The AaLTP proteins accumulated in the ER and the apoplast, often in a polar fashion and sometimes coordinated between individual cells, resulting in an expression pattern around a single cavity (Fig. 2).

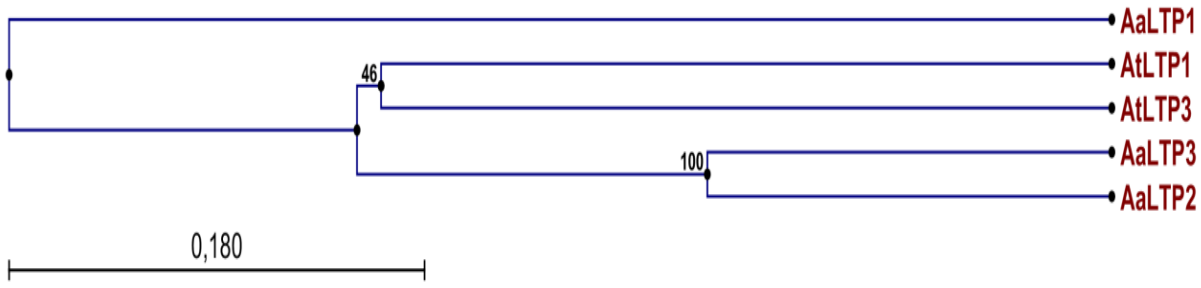

Figure 1. Phylogenetic analysis of AaLTPs and AtLTPs.
Phylogenetic tree was generated using the alignment and the UPGMA method. One hundred bootstrap replicates were used to assess the significance of the tree topology. The analysis involved 5 amino acid sequences (3 from *A. annua* and 2 from Arabidopsis). The scale bar indicates amino acid substitutions per site, and numbers indicate the branch support values in percentage.

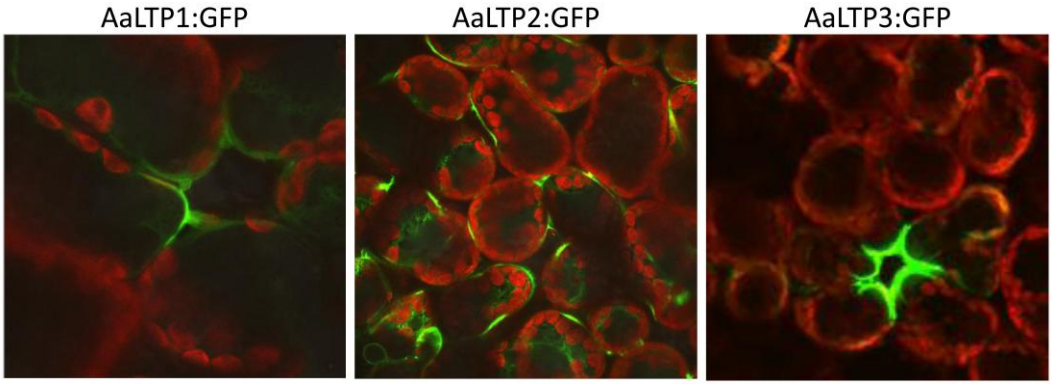

Figure 2. Confocal microscopy analysis of *Nicotiana benthamiana* leaves infiltrated with *AaLTP1:GFP, AaLTP2:GFP* or *AaLTP3:GFP*.

Testing the function of the AaLTPs in N. benthamiana, A. annua and Arabidopsis

To study the function of AaLTPs in artemisinin production, expression constructs were made of the *AaLTP* genes under control of the *CaMV 35S* promoter. The artemisinin biosynthesis pathway genes (*HMGR*, *FPS*, *ADS*, *AMOLAP* and *DBR2*; see **Chapter 2**) were expressed with and without the *AaLTP* expression constructs by agroinfiltration of *N. benthamiana* leaves. Comparison of the product profile at seven days post agro-infiltration showed that co-expression of the *AaLTPs* had no significant effect on the production of artemisinic acid (AA) and dihydroartemisinic acid (DHAA) (Table 1). To assess the role of the *AaLTPs* directly in *A. annua* itself, *AaLTP1-RNAi* and *AaLTP2-RNAi* constructs were made for transformation of *A. annua* (cloning of the *AaLTP3-RNAi* construct failed). Transformation of *A. annua* was done by the group of Prof. Peter E. Brodelius (Linnaeus University, Kalmar, Sweden). Unfortunately, these transformation efforts did not result in any transgenic lines due to problems with regeneration.

Table 1**.** 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	AmFH+AMOLAP+DBR2+AaLTPs
AA	51 ± 12	45 ± 1
DHAA	517 ± 141	525 ± 89

Results are means ± S.D. of three co-infiltrated leaves.
AmFH: mitochondrial targeted amorphadiene synthase (ADS), mitochondrial targeted farnesyl diphosphate synthase (FPS) and a cytosolic (truncated) 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) under control of the CaMV35S promoter (van Herpen *et al.*, 2010).
AMOLAP: amorphadiene oxidase (AMO) from *A. annua* LAP (low artemisinin production) chemotype under control of the CaMV35S promoter (**Chapter 2**).
DBR2: artemisinic aldehyde reductase (DBR2) from *A. annua* under control of the CaMV35S promoter (**Chapter 2**).
AaLTPs: AaLTP1, AaLTP2 and AaLTP3 from *A. annua* under control of the CaMV35S promoter.

Because of the problems with the characterisation of the AaLTPs in *A. annua* and the absence of a significant effect on the artemisinin pathway products upon transient expression in *N. benthamiana*, we switched to the model plant Arabidopsis. Arabidopsis was transformed with the three *AaLTP* expression constructs. Three independent transformed Arabidopsis lines with either the *AaLTP1*, *AaLTP2* or *AaLTP3* under the control of the constitutive *CaMV 35S* promoter were isolated and made homozygous. The effect of overexpression of the individual *AaLTPs* on sesquiterpene (caryophyllene) emission from isolated flowers [floral stage 15; (Smyth *et al.*, 1990)] was investigated. In none of these lines caryophyllene emission was significantly increased, but in one transgenic line (AaLTP2#15) caryophyllene emission was significantly reduced compared with wild-type (Fig. 3).

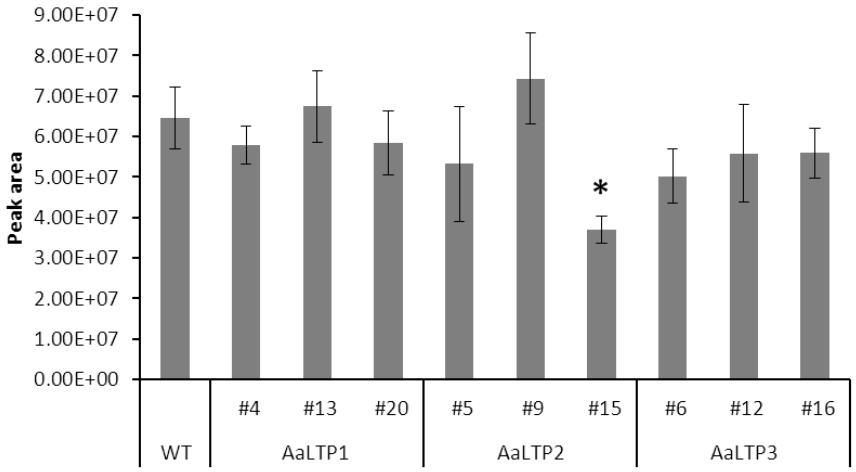

Figure 3. Caryophyllene emission from flowers of *AaLTP*-overexpressing Arabidopsis.
Peak area of caryophyllene produced by Arabidopsis flowers, as detected by headspace GC-MS analysis. Each bar represents the mean of three biological replicates +/- SE. (*) indicates significant difference (Student's t-test, P <0.05) relative to the wild-type (WT).
.

Floral volatile emission from Arabidopsis LTP mutants

Because of the limited success with the *AaLTPs* we decided to make better use of the Arabidopsis resources and study Arabidopsis *LTPs*. Arabidopsis terpene emission occurs mainly from the flowers so we decided to test if we could find a link between one of the Arabidopsis *LTPs* and terpene emission from the Arabidopsis flower. The main floral sesquiterpene, caryophyllene, is produced by flower-specific expression of *TPS21*, while the other, less abundant, floral sesquiterpenes are produced by the flower-specific expression of *TPS11* (Tholl *et al.*, 2005). The floral monoterpenes are produced by the four monoterpene synthases, AtTPS03, AtTPS10, AtTPS14 and AtTPS24 (Chen *et al.*, 2011, Chen *et al.*, 2003).

Because the terpene emission is largely limited to Arabidopsis flowers, we searched for Arabidopsis *LTPs* with expression in the flower using the Genevestigator online search tool (https://www.genevestigator.ethz.ch/). Twelve *LTP* genes showed an overlap with *TPS* gene expression in floral tissues (Chen *et al.*, 2003, Dudareva *et al.*, 1996, Tholl *et al.*, 2005). It is important to note that all these *LTPs* are also expressed in other tissues, where *TPS* genes are not expressed, indicating that terpene transport (if at all) is not their only function (Arondel *et al.*, 2000, Kader, 1996, Ng *et al.*, 2012, Nieuwland *et al.*, 2005, Wang *et al.*, 2012). Results of the *in silico LTP* and *TPS* expression studies are summarized in Figure 4. To investigate whether the *AtLTPs* with floral expression play a role in terpene emission, headspace analysis of the inflorescence volatiles was performed on twelve LTP/ LTP-like T-DNA lines: *gltp*, *ltp1/3/6/* and eight *ltp-like* mutants (see Materials and Methods) and the results compared with the volatiles of WT inflorescences. Analysis of the headspace produced in four hours by ten isolated inflorescences per genotype showed that the most abundant volatile is (*E*)-β-caryophyllene (produced by TPS21). The signal of the other sesquiterpenes (produced by TPS11) was too low to detect. Quantification of the volatiles by GC-MS showed large variation between replicates, most likely caused by variation in the number and developmental stage of the flowers in the pooled inflorescences (data not shown).

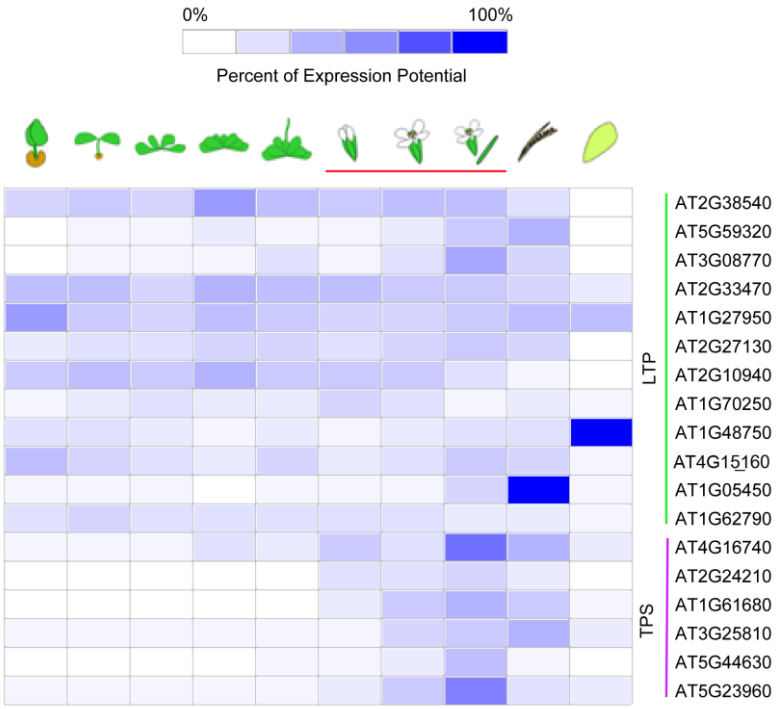

Figure 4. Genevestigator heat map of Arabidopsis *LTP* and *TPS* expression (transcript abundance) in different development stages.

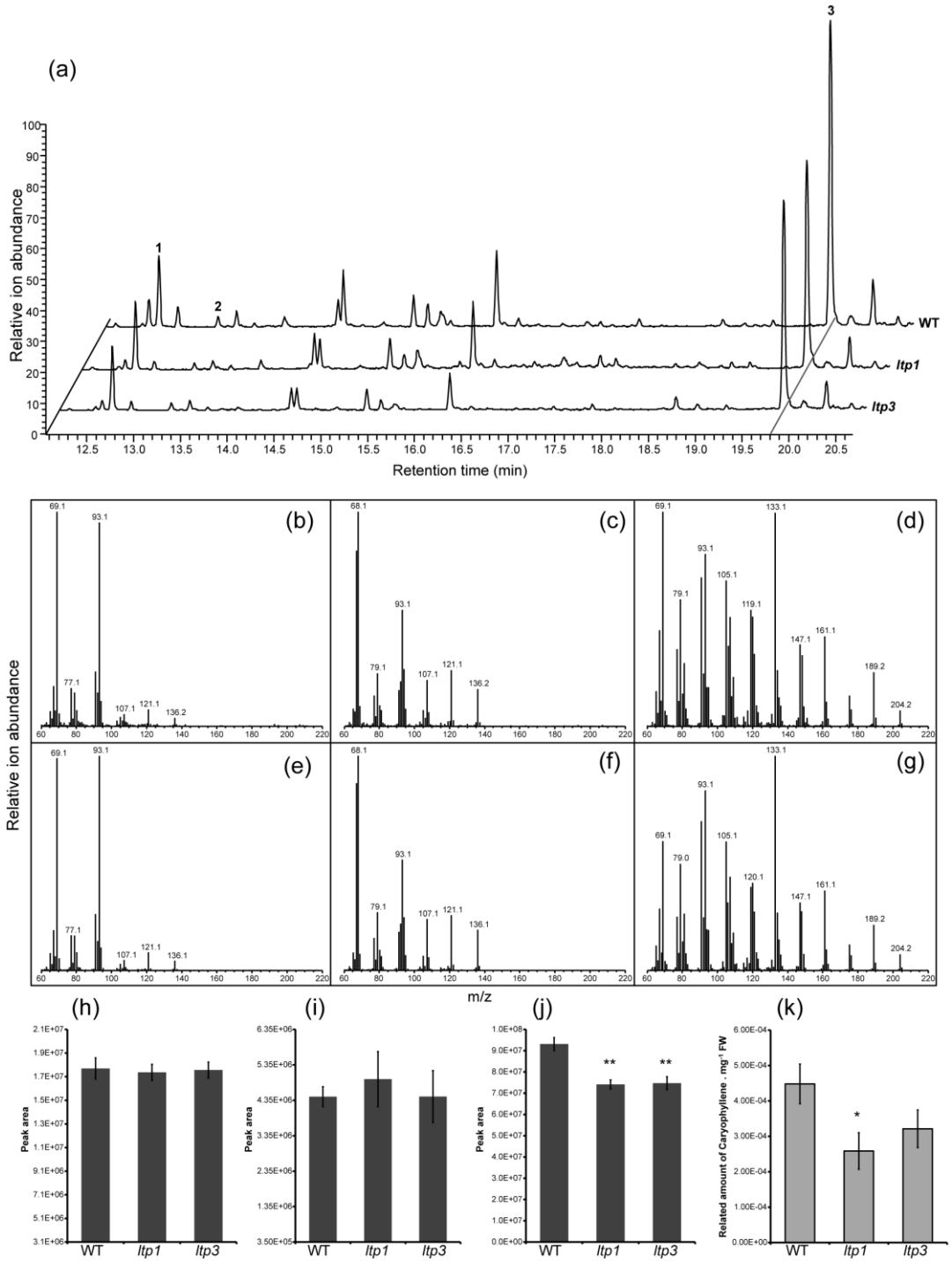

Figure 5. Sesquiterpene emission and content is reduced in *ltp1* and *ltp3* mutant flowers.
(a) GC-MS chromatograms showing the volatiles emitted from Arabidopsis WT, *ltp1* and *ltp3* flowers. 1: β-myrcene, 2: limonene, 3: caryophyllene.
(b-d) Mass spectra for terpenes emitted by Arabidopsis flowers: β-myrcene (b), limonene (c) and caryophyllene (d).
(e-g) Mass spectra for β-myrcene (e), limonene (f) and caryophyllene (g) standards.
(h-j) Peak area of β-myrcene (h), limonene (i) and caryophyllene (j) produced by Arabidopsis flowers, as detected by headspace GC-MS analysis. Each bar represents the mean of five biological replicates +/- SE. (**) indicates significant differences (Student's t-test, P <0.001) relative to the wild-type.
(k) Relative amount of caryophyllene in Arabidopsis flowers, as detected by solvent extract GC-MS analysis. Each bar represents the mean of three biological replicates +/- the standard error. (*) indicates significant differences (Student's t-test, P <0.05) relative to the wild-type (WT).

Reduced sesquiterpene emission and accumulation in isolated flowers of LTP mutants

Because *AtLTP1* and *AtLTP3* showed the closest homology to the AaLTPs we did a more detailed analysis of the floral emission of the Arabidopsis *ltp1* and *ltp3* mutants using flowers in one developmental stage, at the time of bud opening [floral stage 15; (Smyth *et al.*, 1990)] instead of pooled inflorescences. Using this approach, emission was less variable between replicates. Analysis of the emission from *ltp1* and *ltp3* mutant flowers now showed that caryophyllene emission from both mutants was significantly reduced by about 20% compared with that from WT flowers (20.3% reduction for *ltp1*, $P < 0.001$; 19.7% reduction for *ltp3*, $P < 0.001$; Figs. 5a, j). This experiment was repeated with new plants and similar results were obtained: a significant reduction in emission of caryophyllene (28.2% for *ltp1*, $P < 0.005$; 20.7% for *ltp3*, $P < 0.05$). In contrast, the emission of the floral monoterpenes limonene and β-myrcene was not significantly altered(Figs. 5a,h,i). These results suggest that LTP1 and LTP3 affect sesquiterpene, but not monoterpene emission from flowers.

To determine if caryophyllene was also reduced inside the *ltp1* and *ltp3* mutant flowers, isolated flowers were extracted for GC-MS analysis. Caryophyllene accumulation was significantly decreased by about 40% in *ltp1* flowers compared with WT flowers ($P < 0.05$; Fig. 5k), while in *ltp3* flowers there was a trend towards a reduction of 20% in caryophyllene accumulation, but this was not significant.

The relatively small effect of the *ltp1* and *ltp3* mutations on caryophyllene emission may be the result of functional redundancy in LTP function. Therefore, we tried to obtain *ltp1 ltp3* double mutant lines to test whether caryophyllene emission would be further reduced. However, no double mutants could be obtained because of chromosomal translocation in the *ltp1* mutant (see **Chapter 4**).

Partial overlap in floral expression patterns of LTP and TPS genes

The two Arabidopsis sesquiterpene synthases, *TPS11* (multi-sesquiterpene synthase) and *TPS21* [(*E*)-β-caryophyllene synthase], are highly expressed in flower organs (Chen *et al.*, 2011, Tholl *et al.*, 2005). Similarly, *LTP1* and *LTP3* (which affect caryophyllene emission from flowers; see above) are also highly expressed in flower organs (Chae *et al.*, 2010) but the details of the floral expression of these genes in flowers are missing. Therefore we determined the spatial expression pattern of *TPS21*, *TPS11*, *LTP1* and *LTP3* in flowers using GUS reporter lines (Chae *et al.*, 2010, Tholl *et al.*, 2005).

Both *TPS21::GUS* and *LTP1::GUS* showed strong expression in the stigma of the flowers (Figs. 6a, j, k). Only *LTP3* expression could be easily detected in pollen grains (Fig.

6e), although *TPS11* expression in pollen was observed after wounding of the anthers (Fig. 6h). *LTP1* was expressed in the sporophytic tissue of the anther (Fig. 6b). All four genes were expressed in the developing ovules, but only *TPS11* expression could be stained *in situ* in the carpel tissue. For the remaining three lines, GUS expression could only be observed when the ovules were excised from the carpels. *LTP1* and *LTP3* were expressed in the inner integuments of the ovule, particularly in the middle and micropylar regions (Figs. 6c, f). We routinely observed GUS expression in the embryo sac of *LTP1::GUS* and *LTP3::GUS* plants (Fig. 6f). The two *TPS* genes showed similar expression patterns, with both genes being expressed in the funuculus and in the micropylar region of the ovule integuments. Thus, the *LTP1* and *TPS21* expression patterns partially overlap in flowers and the micropylar region of the ovule integuments, and all the other sites where the *TPS* genes are expressed.

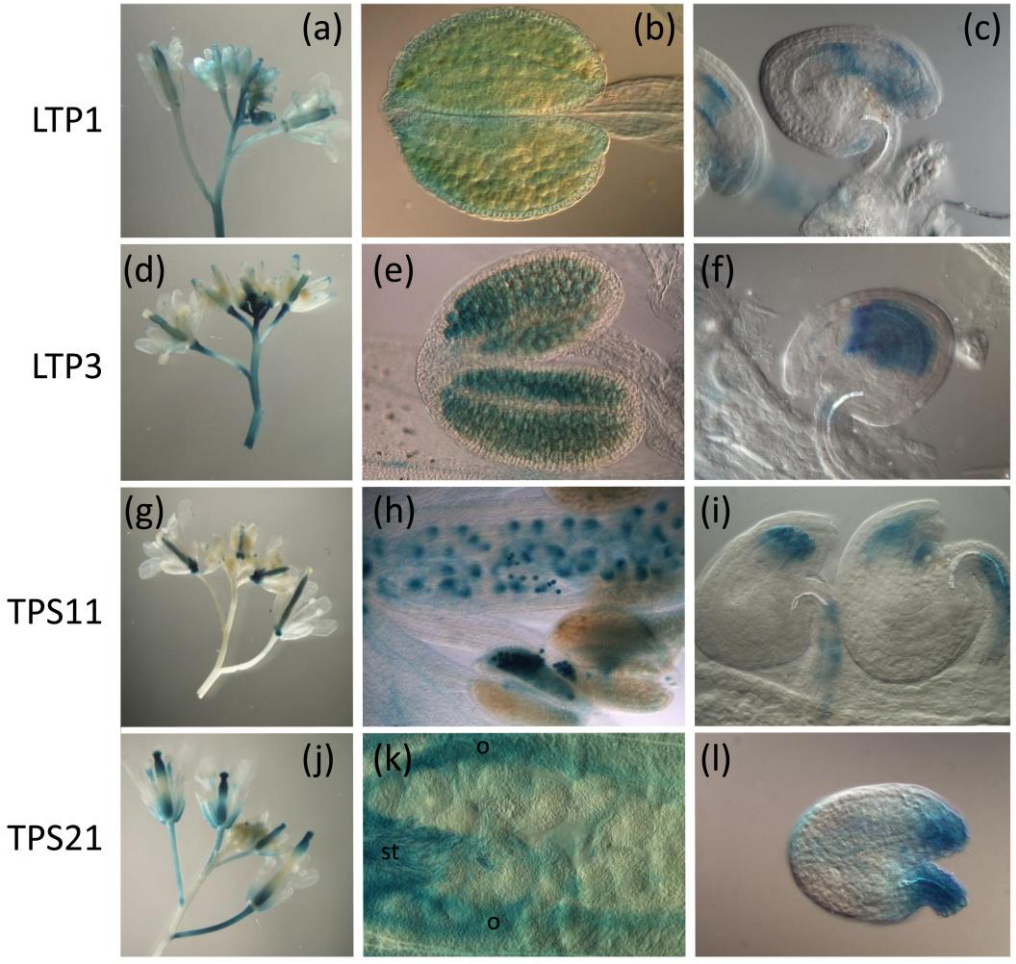

Figure 6. The lipid transfer protein and terpene synthase genes are expressed in developing flowers
(a), (d), (g), (j), inflorescences;
(b), (e), anthers;
(h) *TPS11::GUS* expression in the ovules and in the pollen of damaged anthers;
(k) *TPS21::GUS* expression in the style (st) and ovary (o) of the carpel,
(c), (f), (i) ,(l) ovules.

Discussion

In *Artemisia annua,* the artemisinin precursor, dihydroartemisinic acid (DHAA), is produced in the apical and sub-apical cells of the glandular trichomes (Olofsson *et al.*, 2012), while the end product, artemisinin, accumulates in the apoplastic space under the cuticle. Therefore, specific membrane transporters are likely involved in the sequestration of DHAA in the subcuticular space. Several studies have shown that *LTP* genes are highly represented in cDNA libraries of glandular trichomes, which is the site of high terpenoid biosynthesis activity, including in *A. annua* (Bertea *et al.*, 2006, Harada *et al.*, 2010, Lange *et al.*, 2000, Schilmiller *et al.*, 2010). In the present study, three *LTP* genes were isolated from *A. annua* (Fig.1) and shown to be secreted to the apoplast upon transient expression in plant cells (Fig. 2). The location of the LTP proteins is not evenly distributed in the apoplast but often seemed to be coordinated between cells, as if they were forming a cavity in the spongy parenchyma cells (Fig. 2), suggestive of secretory duct formation. However, we failed to prove a function of these AaLTPs in relation to the production of the sesquiterpene artemisinin in three different functional assays:

(1) We failed to obtain transformants of *A. annua* with an *LTP-RNAi* constructs due to regeneration problems, indicative of a vital function of LTPs in the process of regeneration. The closest homologs of the *A. annua LTP* genes in Arabidopsis are *LTP1* and *LTP3* (Fig.1). *AtLTP1* was shown to be expressed in the shoot apical meristem of Arabidopsis (Thoma *et al.*, 1994) and its expression is induced during somatic embryogenesis (Potocka *et al.*, 2012). Thus, knocking down *LTPs* in *A. annua* may negatively affect plant regeneration and hence result in lethality. In the future, this problem may be circumvented by making use of inducible *LTP-RNAi* constructs or using trichome-specific promoters which will avoid suppression of gene expression during the tissue culture regeneration phase.

(2) We also attempted to characterise the three *AaLTPs* by overexpression in Arabidopsis, but no significant change was detected in the Arabidopsis flower sesquiterpene emission when the *AaLTPs* were overexpressed (Fig. 3). There could be several explanations for this: (i) AaLTPs are not involved in sesquiterpene production/emissions, (ii) the AaLTPs are specific for artemisinin related products and do not recognize caryophyllene and other Arabidopsis (floral) terpenes or (iii) endogenous LTP expression in Arabidopsis flowers is already saturating. Remarkably, one of the *AaLTP* overexpressing lines showed significant lower caryophyllene emission. This could be related to disruption of a relevant gene by the T-

DNA insertion in this line or could be caused by silencing of the transgene and the corresponding endogenous Arabidopsis *LTP* gene(s).

(3) The *AaLTP* genes were also characterised in transient co-expression with the artemisinin biosynthesis pathway genes in *N. benthamiana*, but no effect on the product profile was observed (Table 1). Again this could mean that AaLTPs have no role in the secretion of artemisinin biosynthetic pathway products, but it could also be that endogenous *N. benthamiana LTP* gene expression is already saturating or that other factors are missing in *N. benthamiana*, to allow the LTPs to be functional. A good candidate for such a factor is an ABC transporter, which is likely involved in secretion of artemisinin biosynthetic pathway products in *A. annua*, but could be missing in *N. benthamiana*.

LTP1 and LTP3 seem to affect sesquiterpene emission from flowers

Because Arabidopsis offers more options to manipulate gene expression we next screened for LTP function in relation to sesquiterpene production in Arabidopsis. The inventory of *LTP* genes with floral expression (Fig. 4) resulted in a list of candidate LTP gene for which T-DNA insertion mutants were obtained. After selecting the homozygous lines, these different LTP mutants were screened for floral emission from pooled inflorescences. However, in none of these lines emission was significantly reduced. The headspace measurement of pooled inflorescences suffered from large variation between replicates, most likely due to variable number and stages of flowers on the pooled inflorescences. Because of the close homology to the *AaLTPs*, *ltp1* and *ltp3* mutants of Arabidopsis were analysed in more detail using headspace measurement from isolated freshly opened flowers. This approach resulted in less variation between replicates and in two independent experiments both the *ltp1* and *ltp3* mutant now displayed a significant 20-28% reduction in caryophyllene emission into the headspace (Fig. 5). The limited effect of the individual LTPs could be due to the redundancy in LTP function in the flower. To assess this, we tried to obtain an *ltp1 ltp3* double mutant but we failed to obtain such a double mutant due to chromosomal translocation problems with the *ltp1* mutant line as described in **Chapter 4**. In the future, such a double (or multiple) *ltp* mutant could be obtained using RNAi targeting multiple Arabidopsis *LTPs*. It is not impossible that the Arabidopsis line #15 with *AaLTP2* overexpression is actually already silenced for multiple endogenous Arabidopsis *LTPs* (Fig. 3), which could be further assessed in the future. Double knockdown lines of *LTP1* and *LTP3* may also be obtained using inducible artificial microRNAs (Schwab *et al.*, 2006) to investigate whether their effect on sesquiterpene emissions is additive. The relatively small effect on caryophyllene emission in

the *ltp1* and *ltp3* mutant could also be explained by different emission pathways (facilitated and non-facilitated). If there is a role of LTPs in caryophyllene emission this could be only for the facilitated pathway with the remaining emission occurring through a non-facilitated pathway. Another explanation for the reduced caryophyllene emission in the two mutants could be a pleiotropic effect on caryophyllene synthase expression. Analysis of the transcript level of caryophyllene synthase in *ltp1* and *ltp3* mutants could verify this. During this research a clear example was published where LTPs were shown to affect terpene accumulation. In tobacco, NtLTP1 was shown to control accumulation of the diterpene cembratrienol, (1*R*,3*S*)-cembra-4,7,11,15-tetraen-3-ol, cembratriendiol and labda-8(17),13*E*-dien-15-al in glandular trichomes (Choi *et al.*, 2012). *Ntltp1* mutant plants displayed a 50-85% reduction in the diterpene levels in trichomes, while overexpression of *NtLTP1* raised diterpene levels by 160-190% (Choi *et al.*, 2012).

Do LTP1 and LTP3 provide a lipid sink in flowers?

The effect of the *ltp1* and *ltp3* mutations on the sesquiterpene concentration in Arabidopsis flower extracts could possibly be explained by reduced lipid transfer function in the flower, resulting in a reduced lipid pool and thus reduced storage capacity for the hydrophobic terpenes. Indications of a close association between terpene biosynthesis activity and lipid storage comes from grape. In grape the sesquiterpene synthase, valencene synthase (VvValCS), catalyses the formation of the major sesquiterpenoid volatiles in the flowers (Martin *et al.*, 2009) and localizes to lipid vesicles in pollen grains. Although the lipid storage theory is attractive, LTP1 has been shown to be located in the apoplast (Federico *et al.*, 2005, Potocka *et al.*, 2012), while LTP3 and other LTP proteins are also predicted to be localized extra-cellularly (Yeats&Rose, 2008), as was demonstrated for the barley LTP6 protein (Federico *et al.*, 2005) and here for the AaLTPs (Fig. 2). An extracellular location of LTPs is not compatible with a role in intra-cellular lipid storage for sequestering of terpenes. A reduced apoplastic lipid pool could be an alternative explanation for the effect of the LTP mutations on caryophyllene levels in the flowers. However, a reduced apoplastic sink capacity would likely also affect monoterpene levels in floral extracts and moreover would be expected to enhance rather than reduce headspace emission. The headspace analysis of the *ltp1* and *ltp3* flowers shows that the emission of two other hydrophobic terpenes (the monoterpenes β-myrcene and limonene) was not affected, which is not consistent with a putative role of LTP1 and LTP3 in a general hydrophobic sink capacity in the flower. Alternatively, LTPs could play a role in transport of the terpenes across the membrane, such

as in the unloading of transporters. Although in the tobacco *Ntltp1* mutant both diterpenes and aromatic dicarboxylic acid and alkene levels were affected (Choi *et al.*, 2012) it is possible that the Arabidopsis LTP1 and LTP3 are substrate specific which could explain the reduced emission of sesquiterpenes but not monoterpenes in the *ltp1* and *ltp3* mutants.

LTPs are also expressed in tissues where TPS11 and TPS 21 are not present

The limited effect of *ltp1* and *ltp3* mutations on sesquiterpene accumulation and emission is consistent with the only partial spatial overlap in expression of these *LTPs* and *TPS21* and *TPS11* in flowers. Assuming that LTP1 and LTP3 are required for caryophyllene transport at the sites of *TPS21* expression, the 20% reduction in caryophyllene emission in the *ltp1* and *ltp3* mutant relates to a reduced emission from the tissues where both the *LTP* and *TPS* are expressed: the anthers and stigma in *ltp1* and the anthers and ovule in *ltp3* (Fig. 6). However, *TPS21* is also expressed at the base of the petals, where *LTP1* and *LTP3* do not show obvious expression. Other *LTP* genes with a role in floral caryophyllene production might be expressed in this region of the flower. TPS11 is responsible for emissions of a subset of sesquiterpenes, but emission of these compounds is so low that an effect of the LTP mutations could not be evaluated by our flower headspace GC-MS analysis. However, there is some overlap in expression of *TPS11* with *LTP1* and *LTP3*, with all of them being expressed in the ovule and anthers. Moreover, both *LTP1* and *LTP3* are also expressed in non-floral tissues. While the expression inventory of the *TPS11* and *TPS21* (Figs. 6, S1~S4) indicates that these genes are mainly expressed in the flower, we recently did find a vegetative phenotype for the *tps21* mutant, indicating that at least TPS21 also plays a role in vegetative tissue (van der Krol/Delatte unpublished results).

In summary, there are indications for a role for LTPs in terpene emission based on our results with the Arabidopsis *ltp1* and *ltp3* mutant, and maybe the Arabidopsis AaLTP2OE #15 line but also from recent literature (Choi *et al.*, 2012). The effects of LTPs on terpenes could be explained through lipid sink function but this explanation would not be compatible with the effect on caryophyllene emission butnot on other apolar monoterpenes. Results of the other experiments described in this chapter do not support a role for LTPs in artemisinin pathway product transport, but also do not exclude this role because of alternative explanations (e.g. saturating endogenous LTP levels, limiting additional factors) and clearly more research is needed to determine if and how LTPs contribute to terpene accumulation and emission.

Acknowledgements

H-M.T. was funded by the graduate school of Experimental Plant Sciences (EPS). We thank Desalegn Woldes Etalo for help in MS data analysis. We thank Merche Soriano Castan and Hui Li for support in GUS staining assay.

Supporting Information

Table S1. AaLTPs cDNA and deduced protein sequence.

		Sequence
AaLTP1	cDNA	ATGGTTGGAAAGGTTGTGTTGGTCGTAGCCATTTACTTC CTTGTGGTGGCTGGGCTACATGCAGTAGAAGGCGAGGTG ACATGCGATCAGGTTGTGAGCAACATGACGCCGTGTGTG ACCTACCTAACCAGTAGTGGGGATTCCGTACCCTCAGAT TGTTGTAGCGGTGTTAACTCACTAAACAATGCCGCTACA ACTACTGCTGACAAACAAGCTGCTTGCAAGTGCCTTGAA CAAAGTGCCTCTCAGTTATCGGATATCGACCTTGAGAAA GCTAGAAGCCTTCCGGGGAAATGCGGAGTCAACTTGCCT TATGAGATTAGCCCCACAACTGATTGCTCAACGATACAA TGA
	Protein	MVGKVVLVVAIYFLVVAGLHAVEGEVTCDQVVSNMTPCV TYLTSSGDSVPSDCCSGVNSLNNAATTTADKQAACKCLEQ SASQLSDIDLEKARSLPGKCGVNLPYEISPTTDCSTIQ
AaLTP2	cDNA	ATGGCAAGTATGACAATGAGGGTTTTATGTGTTATTGCG GCTTGCATGGTGGTGGTAGCACCATATGCCGAGGCTCTC TCATGTAGTGAAGTAACGAGCAAGTTGGCGCCATGCTTT AACTACCTAAAGTCTGGTGGTAAGGTGCCACCAGCATGT TGCGACGGAGTCAAGGGACTAAACTCCGCTGCTAAAAC GACCCCTGATAGAAAGACAGCATGCACTTGCATGAAGA GTGCTTATAAATCATACAATGGCATCAACGCTGATAATG CTGCTGGCCTTCCTGGCAAGTGTGGTGTTAATATTCCCTA CAAGATCAGCCTTAGCACCGACTGCAACAAGGTCAAGT GA
	Protein	MASMTMRVLCVIAACMVVVAPYAEALSCSEVTSKLAPCF NYLKSGGKVPPACCDGVKGLNSAAKTTPDRKTACTCMKS AYKSYNGINADNAAGLPGKCGVNIPYKISLSTDCNKVK
AaLTP3	cDNA	ATGGCAAGGATGGCAATGATTGTTTCATGTGTAATCGTG GCTTGTATGTTGGTAGCAGCACCCTATGCTGAGGCTATT AGCTGTGGTCAGGTGGCTAGTAGCTTGGCACCATGCCTT GGCTACCTACAAAAAGGTGGTGATGTGCCACCAGCATGT TGCAGTGGTGTAAAAGGACTCAATGACGCAGCTAAAAC AACCCCTGATCGTCAAACTGCCTGCACCTGCTTGAAGAA CGCTTATTCCGCCAACTCGGGCATTAGTTCCAGCAATGC CGCCGGCCTCCCTGGCAAGTGTGGTGTTAGCATCCCTTA CAAGATTAGCCCCGACACTGACTGCACCAAGGTGCAGTG A
	Protein	MARMAMIVSCVIVACMLVAAPYAEAISCGQVASSLAPCLG YLQKGGDVPPACCSGVKGLNDAAKTTPDRQTACTCLKNA YSANSGISSSNAAGLPGKCGVSIPYKISPDTDCTKVQ

Table S2. Primers used for vector constructs.

Primer	**Sequence (5' to 3') ***
AaLTP1-F	TT*CCATGG*TTGGAAAGGTTGTGTTGG
AaLTP1-R	*GCGGCCGC*TCATTGTATCGTTGAGCAAT
AaLTP2-F	TT*CCATGG*CAAGTATGACAATGAG
AaLTP2-R	*GCGGCCGC*TCACTTGACCTTGTTGCAGT
AaLTP3-F	TT*CCATGG*CAAGGATGGCAATGATTGT
AaLTP3-R	*GCGGCCGC*TCACTGCACCTTGGTGCAGT
attB1-AaLTP1-F	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**ATGGTTGGAAAGGTTGTG
attB2-AaLTP1-R	**GGGGACCACTTTGTACAAGAAAGCTGGGT**TCATTGTATCGTTGAGCAA
attB1-AaLTP2-F	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**ATGGCAAGTATGACAATGA
attB2-AaLTP2-R	**GGGGACCACTTTGTACAAGAAAGCTGGGT**TCACTTGACCTTGTTGCA
EGFP_c-term-F	GAAGGA*GCGGCCGC*GAGCAAGGGCGAGGAGCTG
EGFP_c-term-R	CACAAA*GAGCTC*TTTATACAGCTCGTCCATGC
AaLTP1_gfp-F	AACACC*GGATCC*AATGGTTGGAAAGGTTGTGTTGGTC
AaLTP1_gfp -R	GACAAA*GCGGCCGC*TTGTATCGTTGAGCAATCAGTTGTGG
AaLTP2_gfp-F	AACACC*GGATCC*AATGGCAAGTATGACAATGAGGGTTTTATG
AaLTP2_gfp -R	GACAAA*GCGGCCGC*CTTGACCTTGTTGCAGTCGGTG
AaLTP3_gfp-F	AACACC*GGATCC*AATGGCAAGGATGGCAATGATTGTTTC
AaLTP3_gfp -R	GACAAA*GCGGCCGC*CTGCACCTTGGTGCAGT

***** Restriction sites are shown in italics and *att* sites in bold.

Table S3. Primers used for T-DNA lines genotyping.

Primer	Sequence (5' to 3')
SALK_134262-LP	ATGCCAACTTCATCACTCCAG
SALK_134262-RP	CGAAAACATGACTTTCCAAATG
SALK_095248-LP	TCGATGCATAATCAAATCGTG
SALK_095248-RP	GTTCAAACACAATGGCTTTCG
SALK_126868-LP	GGTCACATGAAGACTGCCTTC
SALK_126868-RP	TGTTTGATTAATTAAATAGCCTTCTCC
SALK_138212-LP	CAAGAGGGATGCGATTTGTAC
SALK_138212-RP	TTCAACCATTTTGCTTCTTGC
SALK_119487-LP	TATCCGATCCTTCAGAACCTG
SALK_119487-RP	TGGTGTTTATTCATCGGGAAG
SALK_014537-LP	AGACACGACCACACCAATTTC
SALK_014537-RP	TCTGATCCCGACAAGTTCAAG
SALK_144344-LP	GTTTTTGGAAGGGTTTGATCC
SALK_144344-RP	GATAGGTACGAGGGAAATCCG
SALK_083118-LP	CAGTGGATTTGATATGGTTTTTG
SALK_083118-RP	CTGAAGTTAGGTGCTTGCGTC
SALK_072495-LP	ACTTTGTGTTTGGACTTTGCG
SALK_072495-RP	AGGGACAAAAACAAAAGCACC
SALK_142707-LP	CACGTTTTGTCCGTCTAGCTC
SALK_142707-RP	CCTTGCTCCAATCATCTCAAG
LBb1.3	ATTTTGCCGATTTCGGAAC
FLAG_103F11-LP	ACCCTCCTAAACCATCACCAC
FLAG_103F11-RP	TACACGTTGAGCAATGACCAC
FLAG_137H03-LP	TGGTTAAGGATCGGCTAAACC
FLAG_137H03-RP	TGTAAGGGAGATCGACACCAC
FLAG_456G06-LP	CCAAAAGTCAAAGTCAAAGCC
FLAG_456G06-RP	TGCGAGTTATGTGACGTCAAC
FLAG_LB	CGTGTGCCAGGTGCCCACGGAATAGT
GABI_173C09-LP	AGTTGGCAAGAAAACAAAACG
GABI_173C09-RP	AATCGGAGATCGGAGAAAATG
GABI_LB	ATATTGACCATCATACTCATTGC

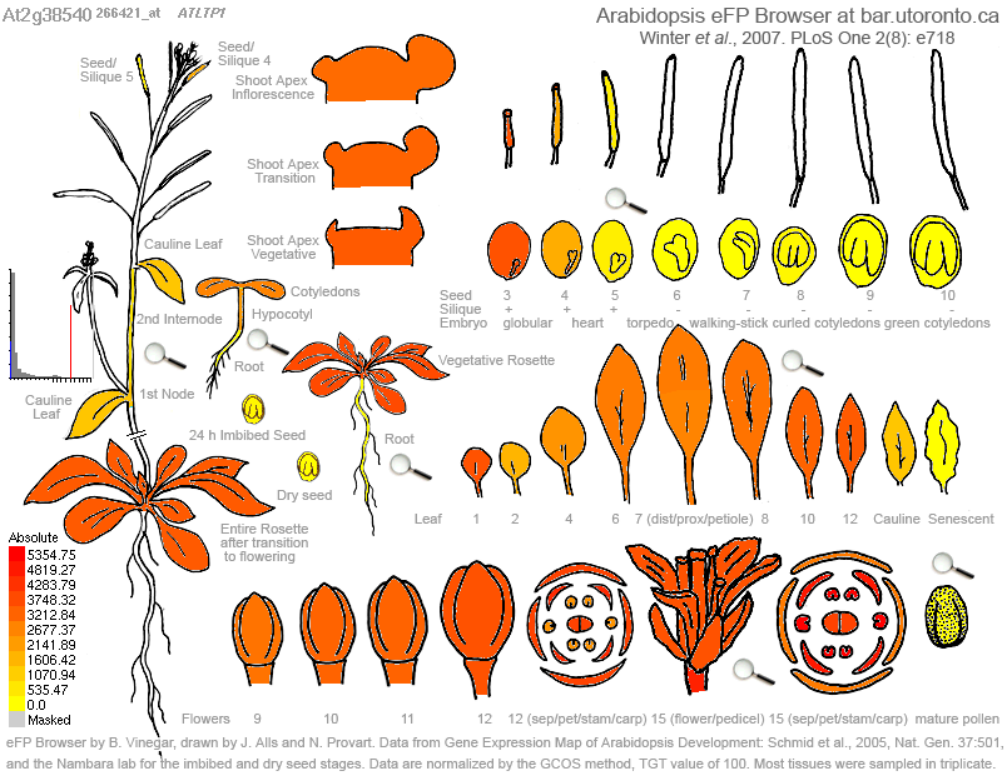

Figure S1. An eFP overview of expression pattern of *LTP1* (At2g38540) in various Arabidopsis tissues based on microarray data. Data were obtained from the Arabidopsis eFP browser (Winter *et al.*, 2007).

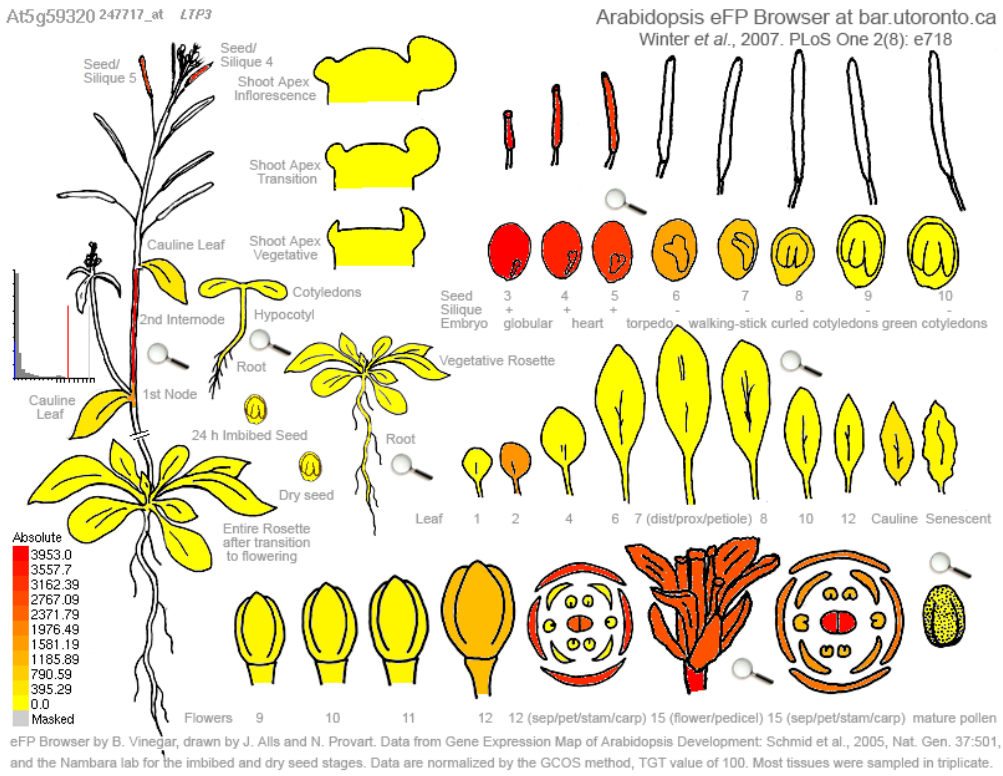

Figure S2. An eFP overview of the expression pattern of *LTP3* (At5g59320) in various Arabidopsis tissues based on microarray data. Data were obtained from the Arabidopsis eFP browser (Winter *et al.*, 2007).

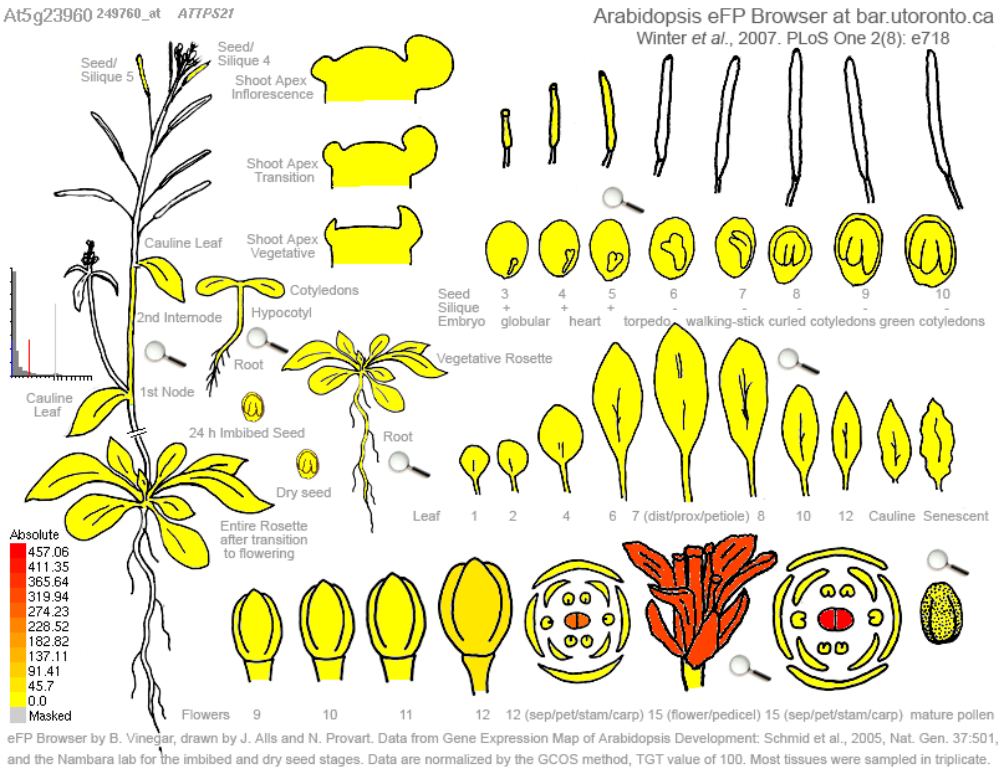

Figure S3. An eFP overview of the expression pattern of *TPS21* (At5g23960) in various Arabidopsis tissues based on microarray data. Data were obtained from the Arabidopsis eFP browser (Winter *et al.*, 2007).

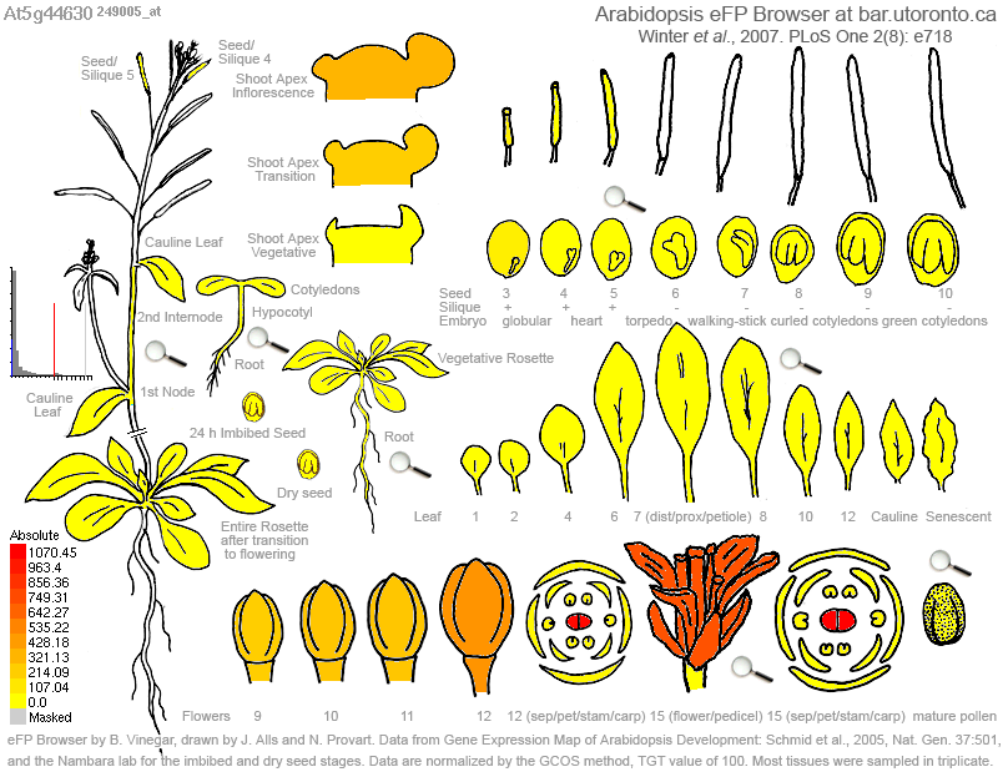

Figure S4. An eFP overview of the expression pattern of *TPS11* (At5g44630) in various Arabidopsis plant tissues based on microarray data. Data were obtained from the Arabidopsis eFP browser (Winter *et al.*, 2007).

Chapter 4

Characterisation of aberrant pollen and ovule phenotypes associated with chromosomal translocations in two T-DNA insertion mutants of Arabidopsis

Hieng-Ming Ting[1], Kim Boutilier[2], Mariëlle Schreuder[1] and Alexander van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, P.O. Box 658, 6700 AR Wageningen, The Netherlands.

[2]Plant Research International, P.O. Box 619, 6700 AP Wageningen, The Netherlands.

Abstract

It is known that T-DNA insertion mutant lines sometimes contain chromosomal translocations, which, upon crossing with a plant with a normal arranged genome, results in a seed development problem. In our research on the role of lipid transfer proteins (LTPs) in sesquiterpene transport, we used two LTP (*ltp1* and *ltp3*) and two sesquiterpene synthase (*tps21* and *tps11*) T-DNA insertion mutant lines. Two out of these four T-DNA insertion mutants displayed the characteristics of lines with chromosomal rearrangements (*ltp1* and *tps21*) as indicated by the problems in seed-set in the F1 progeny after crossing with different genotypes. Previously, only an aberrant pollen phenotype was described as a consequence of chromosomal translocation in a T-DNA line. Here we give a detailed description of both the aberrant pollen and an aberrant ovule phenotype associated with the apparent chromosomal translocation in the *ltp1* and *tps21* lines. The results show that pollen and ovule phenotypes are similar but not identical between different T-DNA insertion lines, indicating that different chromosomal rearrangements may result in subtle but distinct differences in the associated aberrant pollen and ovule phenotype.

Keywords: chromosomal translocations, gametophyte, lipid transfer protein (LTP), terpene synthase (TPS), T-DNA.

Introduction

In our research on a putative role of LTPs in terpene transport, we identified *ltp1* and *ltp3* T-DNA insertion mutants which both display a 20% reduction in caryophyllene emission in open flowers (**Chapter 3**). The small reduction in caryophyllene emission in the single mutants could be the result of functional redundancy between LTPs and therefore we tried to obtain *ltp1 ltp3* double mutants. Sesquiterpenes in Arabidopsis are mainly emitted by the flowers and are the result of the activity of terpene synthase 21 (producing (*E*)-β-caryophyllene, α-humulene and α-copaene) and terpene synthase 11 (producing α-barbatene, thujopsene, iso-bazzanene, β-barbatene, E-β-farnesene, β-acoradiene, β-chamigrene, α-zingiberene, α-cuprenene, α-chamigrene, cuparene, β-bisabolene, β-sesquiphellandrene, δ-cuprenene) (Tholl *et al.*, 2005). LTP and TPS genes have partially overlapping expression profiles and this suggests that there is a role for LTPs in terpene emissions (**Chapter 3**). In order to eliminate redundancy in sesquiterpene production in Arabidopsis flowers we first identified the *tps21* and *tps11* homozygous single mutants and subsequently tried to obtain the *tps21 tps11* double mutant, but failed. In the F1 progeny from *tps21*x*tps11* about 58% of the seeds failed to develop. However, also in the F1 progeny from *ltp1*x*ltp3* about 44% of seeds failed to develop. Backcrosses of the individual mutant lines with wild-type plants resulted in a similar phenotype as for the *ltp1*x*ltp3* and *tps21*x*tps11* crosses, with about 57% seed set for the backcross of *ltp1*xWT and about 43% seed set for the backcross of *tps21*xWT, while the F1 plants from the backcross of *ltp3*xWT and *tps11*xWT did not display a seed-set phenotype. Combined these results are consistent with the interpretation that chromosomal rearrangements have occurred in the *ltp1* and *tps21* T-DNA insertion lines. Analysis of the genetic interaction between the *LTP* and *TPS* genes is hence hindered by gametophytic defects in the *ltp1* and *tps21* T-DNA insertion lines upon crossing with other mutants or wild-type plants. The occurrence of chromosomal translocations, associated with the insertion of a T-DNA has been described before ((Clark&Krysan, 2010) and references therein). Moreover, previous analysis of Arabidopsis T-DNA insertion mutant populations resulted in an estimated frequency of occurrence varying from 12.5% (Castle *et al.*, 1993), 19% (Clark&Krysan, 2010) to 32% (Budziszewski *et al.*, 2001). Unfortunately in the present study, up to 50% of T-DNA (2 out of 4) of the insertion mutants assessed turned out to have a chromosomal translocation event. Plants that are heterozygous for a reciprocal chromosomal translocation are predicted to create ca. 50% non-viable gametes. By contrast, plants that are homozygous for a reciprocal translocation are genetically balanced and therefore are

predicted to produce gametes that are all viable. Because the *ltp1* and *tps21* mutants do not display a gametophytic defect when homozygous but only when they are crossed, we presume that these two T-DNA insertion mutants are homozygous for a reciprocal chromosomal translocation (Curtis *et al.*, 2009). In previous work the pollen phenotype associated with chromosomal translocation in T-DNA lines was described in detail (Clark&Krysan, 2010). Here we extend the pollen phenotype to two other T-DNA insertion lines and describe an aberrant ovule development phenotype associated with the putative chromosomal translocation in *ltp1* and *tps21*.

Materials and Methods

Plant materials and growth conditions

Seeds of Arabidopsis T-DNA insertion mutant lines (At2g38540/LTP1 [SALK_134262], At5g59320/LTP3 [SALK_095248], At5g44630/TPS11 [SALK_126868] and At5g23960/TPS21 [SALK_138212]) were obtained from NASC (Nottingham Arabidopsis Stock Centre). Plants were grown in soil in a greenhouse with a 16h light/8h dark photoperiod at 22°C.

Genotyping of T-DNA insertion lines

Genomic DNA was extracted from leaves of T-DNA insertion lines as described (Edwards *et al.*, 1991). The presence of the T-DNA insertion was verified by PCR on genomic DNA using the corresponding specific forward primer (LP), reverse primer (RP), and T-DNA specific primer (LB) (Supplementary Table S1).

Microscopy

Anthers and ovules were cleared in HCG (30 ml water, 80 g chloral hydrate, 10 ml glycerol) and studied by DIC microscopy using a Nikon OPTIPHOT microscope. Isolated pollen was stained with 1.25 μg/ml 4',6-diamidino-2-phenylindole (DAPI) (Custers, 2003) using a Zeiss Axioskop epifluorescence microscope (excitation wavelength, 400 nm; emission wavelength, 420 nm).

Results

Chromosomal translocations in ltp1 and tps21 T-DNA insertion lines

To study whether functional redundancy in LTPs is the cause of the relatively small effect of the *ltp1* and *ltp3* mutations on caryophyllene emission (**Chapter 3**), we tried to obtain *ltp1 ltp3* double mutant lines to see whether caryophyllene emission would then be further reduced. Reciprocal crosses between homozygous *ltp1* and *ltp3* mutant plants were made, but approximately 44% of the seeds of the double heterozygous plants (*LTP1/ltp1* and *LTP3/ltp3*) did not develop. The non-developed F2 seeds derived from reciprocal crosses were white, small and shrivelled, suggesting a failure in fertilisation rather than seed abortion. The genotype of the F2 plants derived from these crosses was determined by analysis of genomic DNA with primers specific for the T-DNA insertions into *LTP1* and *LTP3*. An anomalous segregation ratio for the F2 progeny of reciprocal crosses between *ltp1* and *ltp3* mutants was observed in which the expected double homozygous *ltp1/ltp1 ltp3/ltp3* genotypes were missing and all the heterozygous combinations were also reduced (Table 1).

Table1. Genotypes in the F2 progeny of *ltp1 LTP1*/*ltp3 LTP3* plants.

Genotype	No.	Expected (%)	Observed (%)
LTP1/LTP1 LTP3/LTP3	1	6.25	0.39
LTP1/ltp1 LTP3/LTP3	0	12.5	0
LTP1/LTP1 LTP3/ltp3	0	12.5	0
ltp1/ltp1 LTP3/LTP3	63	6.25	24.5
LTP1/LTP1 ltp3/ltp3	69	6.25	26.9
LTP1/ltp1 LTP3/ltp3	105	25.0	40.9
ltp1/ltp1 LTP3/ltp3	1	12.5	0.39
LTP1/ltp1 ltp3/ltp3	18	12.5	7.0
ltp1/ltp1 ltp3/ltp3	0	6.25	0
Total	257	100	100

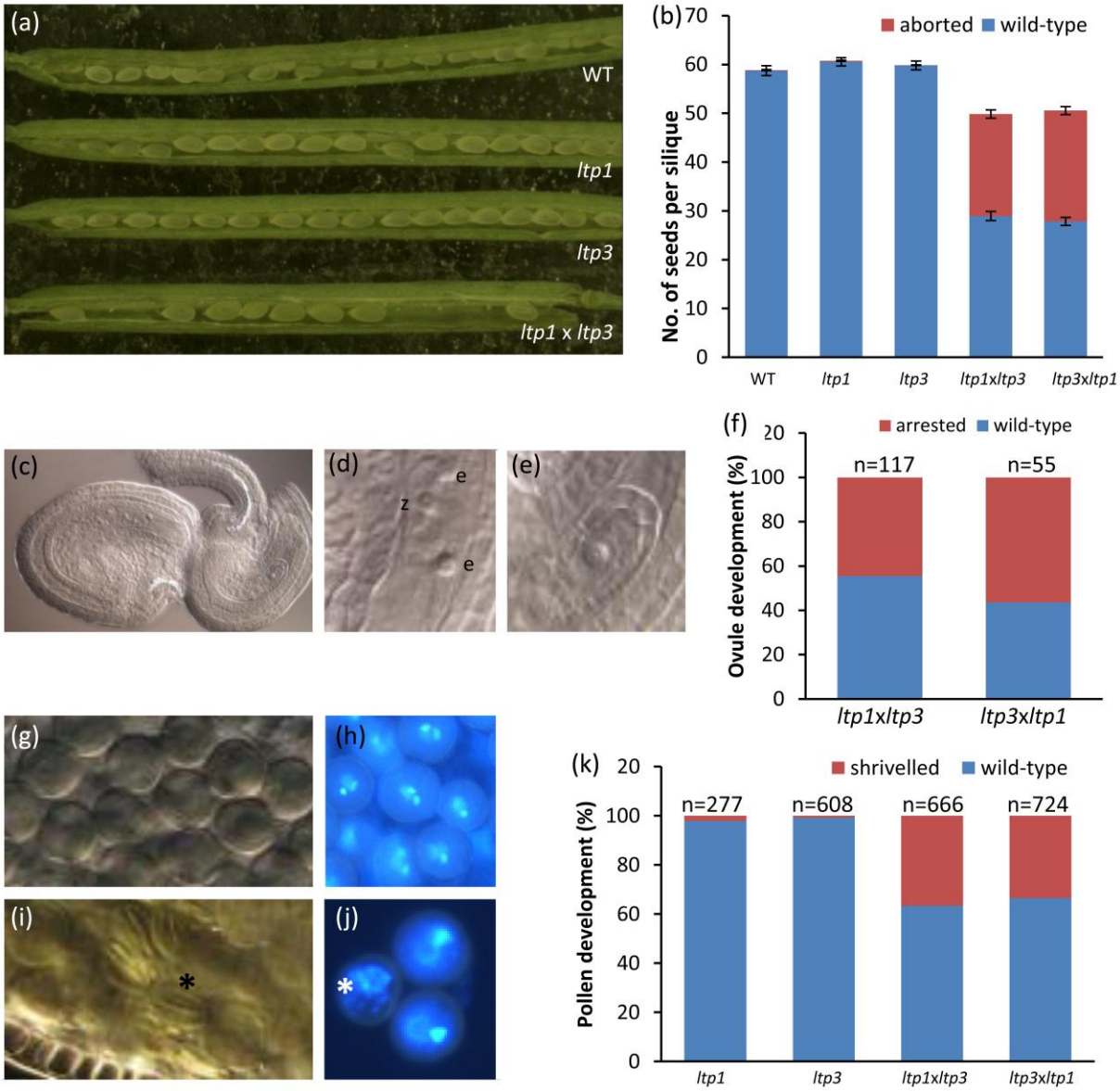

Figure 1. Development of *ltp1* and *ltp3* mutants

(a-b) Seed development in wild-type (WT) and *ltp* mutant plants. Siliques derived from reciprocal crosses between *ltp1* and *ltp3* show reduced seed set compared to the wild-type (WT) and single mutants. A total of 15 siliques were measured for wild-type and mutant plants.

(c-f) Ovule development in the progeny of crosses between *ltp1* and *ltp3*.

(c) A fertilised seed (left) and an under developed ovule (right) in the same silique.

(d) A close-up of the seed on the left in panel C showing a zygote (z) and endosperm nuclei (e).

(e) A close-up of the ovule on the right in panel C showing an arrested one-cell embryo sac.

(f) The proportion of arrested and wild-type ovules present at anthesis for plants derived from the cross *ltp1*x*ltp3* (n= 117) and *ltp3*x*ltp1* (n=55).

(g-k) Pollen development in the progeny of crosses between *ltp1* and *ltp3*.

(g) Wild-type anthers are filled with round pollen grains at the trinucleate stage (H).

(i) Plants derived from crosses between ltp1 and ltp3 contain many shrivelled pollen grains (*).

(j) Degenerating microspore nucleus (*) in the same anther as binucleate pollen grains.

(k) Proportion of well-developed and degenerated pollen at anthesis in WT and *ltp* plants.

Aberrant ovule and pollen development in the ltp1 ltp3 double heterozygous mutant

The ovule and pollen development were examined in double heterozygous *LTP/1ltp1 LTP3/ltp3* plants compared with WT and the homozygous *ltp1* and *ltp3* single mutants using tissue clearing and DIC microscopy. Ovule and seed development were phenotypically normal in WT and *ltp1* and *ltp3* lines. In contrast, about 50% of the ovules derived from the *LTP1/ltp1 LTP3/ltp3* and *LTP3/ltp3 LTP1/ltp1* heterozygous double mutants did not develop past the single cell gametophyte (FG1) stage of development. This defect could clearly be seen at anthesis, where WT ovules with embryo sacs containing an egg cell, two synergid cells and a central cell (FG6 stage) were observed side by side with ovules that contained a tear-shaped, single-celled embryo sac (Fig. 1). Subsequent seed and embryo development in the circa 50% viable progeny appeared phenotypically normal (data not shown).

Abnormal pollen development was also observed in anthers of *LTP1/ltp1 LTP3ltp3* and *LTP3/ltp3 LTP1/ltp1* plants. In both plants the anthers contained a large proportion of shrivelled pollen, a phenotype that was not observed in the WT, and *ltp1* and *ltp3* lines. The abnormal pollen grains were arrested at the uninucleate stage of development (Fig. 1), and unlike the ovules, were degenerated at anthesis. At anthesis, almost 40% collapsed and DAPI-negative pollen grains were observed in the heterozygous *ltp1/LTP1 ltp3/LTP3* double mutants, while only up to 2% aborted pollen grains were found in the single *ltp1* and *ltp3* mutants (Fig. 1). These results suggest that gametophyte development in *ltp1/LTP1 ltp3/LTP3* double mutants is perturbed shortly after meiosis. However, the genetic basis for the phenotype is unclear as only 25% of the haploid gametophytes should carry the double mutant allele combination (*ltp1 ltp3*) after independent segregation of the two loci.

Aberrant ovule and pollen development in the tps11 tps21 double heterozygous mutant

Since our results suggested a role for LTP1 and LTP3 in reproductive development, and they seem to be involved in sesquiterpene emission, we analysed whether TPS11 and TPS21 also affect reproduction. No obvious effect on pollen and ovule development was detected in the *tps21* and *tps11* single T-DNA insertion mutant lines (Fig. 2a). However, the double heterozygous *tps11/TPS11 tps21*/*TPS21* F1 progeny displayed a similar ovule and pollen phenotype as the *ltp1 ltp3* double heterozygous mutant, although the proportion of abnormal pollen and ovules was higher than in the crosses between the *ltp* mutants, and the ovules from *tps11/TPS11 tps21*/*TPS21* plants often proceeded to the two-celled embryo sac stage (Fig. 2).

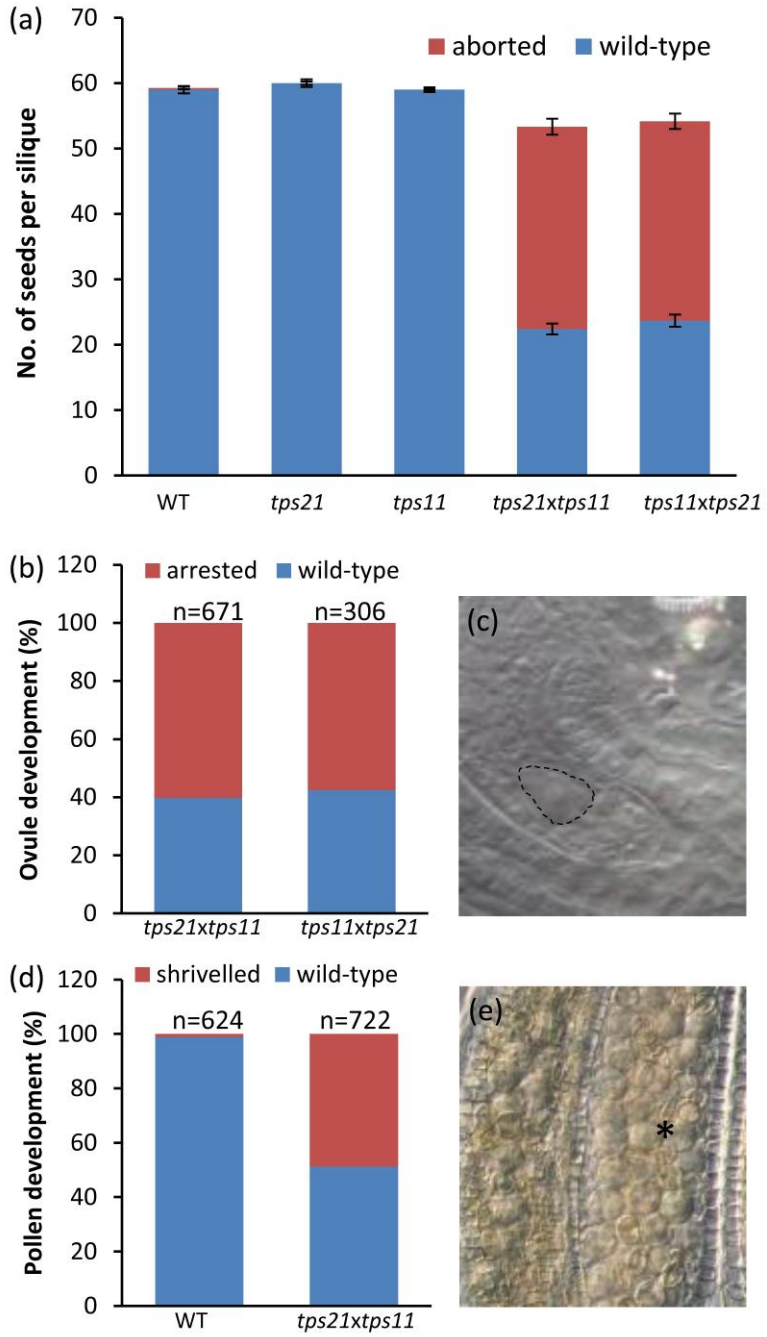

Figure 2. Development of *tps11* and *tps21* mutants
(a) Seed development in wild-type (WT) and *tps* mutant plants. Siliques derived from reciprocal crosses between *tps11* and *tps21* show reduced seed set compared to the wild-type (WT) and single mutants. A total of 15 siliques were measured for wild-type and mutant plants.
(b-c) Ovule development in the progeny of crosses between *tps11* and *tps21* plants.
(b) A high proportion of ovules show arrested development at anthesis.
(c) An ovule arrested at the one nucleate stage. The single cell is circled.
(d-e) Pollen development in the progeny of crosses between *tps11* and *tps21* plants.
(d) Proportion of well-developed and degenerated pollen at anthesis in WT and the progeny of a cross between *tps11* and *tps21* plants.
(e) Plants derived from crosses between *tps11* and *tps21* plants contain many shrivelled pollen grains (*).

Chromosomal rearrangements responsible for the gametophyte phenotype of LTP and TPS double mutants

Initially, these results suggested a link between LTP gene function, sesquiterpene levels and ovule/pollen development. However, the abnormal segregation of the gametophyte phenotypes prompted us to search for other explanations. Recently it was shown that a relatively large percentage of T-DNA insertion mutants display chromosomal translocations, which only become apparent upon crossing with a genotype that lacks the translocation (Clark&Krysan, 2010). The phenotype of both the *tps21*x*tps11* and the *ltp1*x*ltp3* F1 progeny plant showed an aberrant seed set (Figs 1, 2), which could be an indication of a chromosomal translocation event in some of the homozygous mutant lines. A chromosomal translocation in the mutant lines should also result in an aberrant seed set phenotype after a backcross to WT plants. Therefore, to test for chromosomal rearrangements in our *LTP* and *TPS* T-DNA insertion lines, the individual *ltp1, ltp3, tps11* and *tps21* homozygous mutant lines were backcrossed to WT and the number of F2 seeds in siliques of F1 plants assessed. We observed 100% seed set in the *ltp3/*LTP3 and *tps11*/TPS11 plants, but only about 60% in the *ltp1/*LTP1 and *tps21*/TPS21 siliques (Fig. 3). Hence, both *ltp1* and *tps21* display a phenotype in the cross with WT plants that is indicative of chromosomal translocations.

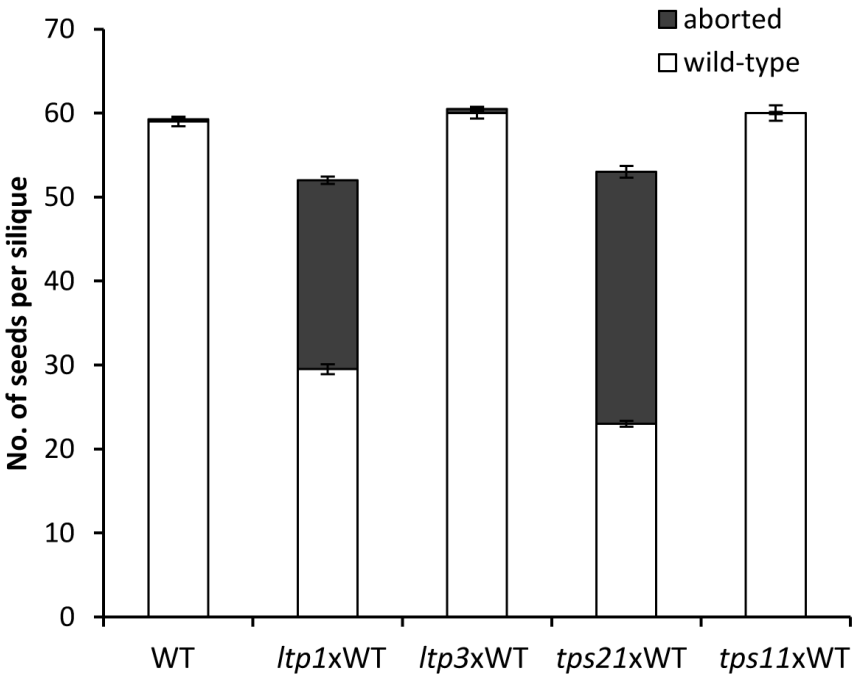

Figure 3. Abnormal seed set in backcrossed *LTP1* and *TPS21* T-DNA insertion lines. Reduced seed set was observed in *ltp1/LTP1* siliques and the *tps211/TPS21* siliques.

Discussion

Chromosomal translocation in T-DNA insertion mutants hindered LTP TPS interaction studies

Because of the significant, but limited effect on caryophyllene emission in the *LTP* single mutants, we tried to combine the *ltp1* and *ltp3* mutations to test whether the reductions in caryophyllene emission are additive. However, we were unable to generate double mutants due to aberrant ovule and pollen development in the double heterozygous F1 progeny. Unfortunately, initially this was not recognized as being related to chromosomal translocation events, also because the possibility that *LTP* genes have crucial functions in reproductive development is not without precedent. In rice (*Oryza sativa*), the LTP *OsC6*, has been proposed to transport the lipid precursor of sporopollenin, a component of the pollen coat, making *OsC6* essential for anther and pollen development (Zhang *et al.*, 2010). Similarly, an anther-specific LTP, CaMF2, plays a role in pollen viability in chili pepper (*Capsicum annuum* L.) (Chen *et al.*, 2011). In Arabidopsis, a stigma/style cysteine-rich adhesin (SCA)-like LTP, LTP5, is involved in pollen tube growth and may also play a role in adhesion-mediated guidance in the pistil (Chae *et al.*, 2009).

The cross of the *tps11* and *tps21* mutant resulted in a similar seed set phenotype, which initially strengthened the interpretation of an interaction between LTPs and TPSs, both affecting gametogenesis. However, the backcross of *ltp1* and *tps21* with WT plants also resulted in the pollen and ovule phenotype in the F1 progeny, suggesting (but not proving) that the observed phenotypes are not caused by loss-of-function of the LTP or TPS functionality, but rather by a chromosomal translocation in the T-DNA insertion mutant *ltp1* and *tps21*.

Differences in pollen and ovule phenotypes associated with chromosomal translocations

Detailed analysis of the aberrant ovule and pollen phenotype in the *ltp1* and *tps21* lines showed they are not exactly identical, which may be due to the difference in site or severity of the chromosomal translocation (Figs. 1, 2). Analysis of the SALK T-DNA insertion lines revealed up to 19% of chromosomal translocations (Clark&Krysan, 2010). While in the publication of Clark and Krysan, only the aberrant pollen phenotype was related to chromosomal translocations, here we also describe an aberrant ovule development phenotype that is associated with this T-DNA population. In addition, the pollen phenotype associated with the *ltp1* mutant is slightly different from that described by (Clark&Krysan, 2010). This

could be due to differences in the chromosomal translocation event in the T-DNA insertion lines. Alternatively, it is possible that there is an additional effect on pollen and ovule development related to the loss of LTP function.

Acknowledgements

H-M.T. was funded by the graduate school of Experimental Plant Sciences (EPS).

Supporting Information

Table S1. Primers used in this study.

Primer	Sequence (5’ to 3’)
SALK_134262-LP	ATGCCAACTTCATCACTCCAG
SALK_134262-RP	CGAAAACATGACTTTCCAAATG
SALK_095248-LP	TCGATGCATAATCAAATCGTG
SALK_095248-RP	GTTCAAACACAATGGCTTTCG
SALK_126868-LP	GGTCACATGAAGACTGCCTTC
SALK_126868-RP	TGTTTGATTAATTAAATAGCCTTCTCC
SALK_138212-LP	CAAGAGGGATGCGATTTGTAC
SALK_138212-RP	TTCAACCATTTTGCTTCTTGC
LBb1.3	ATTTTGCCGATTTCGGAAC

Chapter **5**

Inhibition of vesicle transport during terpene biosynthesis causes proteasome malfunction

Hieng-Ming Ting[1], Thierry L. Delatte[1], Pim Kolkman[1], Harro J. Bouwmeester[1] and Alexander van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, P.O. Box 658, 6700 AR Wageningen, The Netherlands.

Abstract

Plants produce numerous volatile terpenes which may serve different functions in biotic and abiotic stress responses. Much effort has been dedicated to the identification and characterization of the terpene biosynthetic genes, but little is known about how these lipophilic molecules are transported within the cell and from the cell into the apoplast. Here, we investigated whether vesicles play a role in terpene transport in the cell by inhibiting vesicle fusion with target membranes (mediated by VAMP72 proteins) in combination with ectopic production of the sesquiterpene caryophyllene and the monoterpenoid linalool. Vesicle fusion was inhibited by *Agrobacterium tumefaciens*-mediated transient expression of a *MtVAMP721e-RNAi* construct in *Nicotiana benthamiana*. Caryophyllene and linalool production were introduced by transient co-expression of a caryophyllene synthases and linalool synthase, respectively. Headspace analysis of the leaves showed that caryophyllene and linalool emission increased about 5-fold when *N. benthamiana* VAMP72 function was blocked. Intriguingly, a co-expressed DsRed protein accumulated to higher levels in these treatments. RNA sequencing analysis of the agro-infiltrated *N. benthamiana* leaves showed that differentially expressed genes belonged to two categories: (up-regulated) proteasome-related genes and (down-regulated) photosynthesis genes. We discuss this unexpected finding, the role of vesicle fusion in terpene transport and the implications for the improvement of terpene production in metabolic engineering.

Keywords: terpene transport, vesicle-associated membrane proteins (VAMP72), caryophyllene synthase, linalool synthase, proteasome, photosynthesis, *Nicotiana benthamiana*.

Introduction

Plants produce numerous terpenoids, which have many different functions, varying from plant hormones to abiotic stress protectants (e.g. heat protection) or in biotic stress responses (e.g. attraction of predators of plant herbivores) (Holopainen&Gershenzon, 2010). Bulk volatile isoprene and monoterpenoids are emitted by some tree species and seem to function in heat protection of the photosynthesis system (Vickers *et al.*, 2009). These small hydrophobic molecules are produced in the plastids and possibly sequester into the membranes. They are also emitted, however at present it is unknown how these molecules reach the plasma membrane that they need to pass to be emitted into the atmosphere. Plants also produce many sesquiterpenes and for these the first committed biosynthetic step occurs in the cytosol. For instance, the first step in the production of the anti-malaria agent artemisinin by amorphadiene synthase (ADS) produces the volatile amorphadiene and the first step in the production of the anti-cancer agent costunolide by germacrene A synthase (GAS) produces the volatile germacrene A. When the genes for *ADS* or *GAS* are transiently expressed in *Nicotiana benthamiana* leaves, amorphadiene and germacrene A are detected in the headspace of the leaves (Liu *et al.*, 2011, van Herpen *et al.*, 2010). If additional genes from these pathways, such as the ER-localized cytochrome P450s, are co-expressed, the volatile amorphadiene and germacrene A disappear from the headspace as they are converted to less volatile, oxidized intermediates. In addition, some of these products are also converted by endogenous *N. benthamiana* enzymes, such as cytochrome P450s and glycosyltransferases (Liu *et al.*, 2011, van Herpen *et al.*, 2010). Very little is known about the subcellular transport of the volatile products from the cytosol to the plasma membrane –for emission - or ER – for further conversion by for example cytochrome P450s - and the transport across the plasma membrane to the apoplast. As lipophilic compounds would automatically sequester into the membrane enclosing the different sub-cellular compartments, one option for transport of terpenes within the cell could be through vesicle transport between subcellular compartments and delivery of the terpenoids to target membranes by vesicle fusion. Hints that terpenoids may be transported by vesicles come from studies on *Plasmodium falciparum* for which it was shown that the sesquiterpene lactone artemisinin is transported from the red blood cell into the parasite (*Plasmodium falciparum*) via parasite-derived membrane vesicles (Eckstein-Ludwig *et al.*, 2003, Gershenzon&Dudareva, 2007).

The fusion of vesicles with target membranes is mediated by a group of proteins called SNAREs (Soluble *N*-ethylmaleimide sensitive factor Attachment protein REceptor)

(Lang&Jahn, 2008). One v-SNARE (SNARE protein on the transport vesicle) pairs with three t-SNARE proteins (SNARE proteins on the target membrane, including syntaxins) (Lang&Jahn, 2008), leading to membrane fusion between the two compartments. v-SNAREs consist of long vesicle-associated membrane proteins (VAMPs) or 'longins' (containing an N-terminal longin domain) and short VAMPs or 'brevins' (Filippini *et al.*, 2001). However, in plants only the longin-type v-SNAREs occur, and they can be further classified into three major groups: VAMP7-like, Ykt6-like and Sec22-like (Fujimoto&Ueda, 2012). A sub-group of the VAMP7-likes, the VAMP72 family, has been shown to be involved in exocytosis in plants (Kwon *et al.*, 2008, Sanderfoot, 2007).

Here we addressed the question whether vesicle transport is involved in transport of terpenoids within the cell and whether vesicle fusion plays a role in the emission of the terpenoids into the apoplast. As a model we used the volatile sesquiterpene caryophyllene, which is not easily converted to oxidized and/or glycosylated, non-volatile, compounds upon heterologous production in plants, and the monoterpenoid linalool. To evaluate the role of vesicle transport/fusion we combined transient expression of caryophyllene synthase (*CST*) and linalool synthase (*FaNES*) with a *SNARE-RNAi* construct to target *N. benthamiana VAMP72* genes.

To assess whether and how ectopic caryophyllene production and silencing of *SNARE* genes affect the transcriptional response in *N. benthamiana* RNA sequencing was used. This showed that the transient expression of *MtVAMP721e-RNAi* results quite specifically in down regulation of multiple *N. benthamiana VAMP72* genes. The response to *CST* expression, however, was more pleiotropic and for example also affected cell wall related processes. The combination of caryophyllene production and inhibition of VAMP72 function resulted in a much stronger transcriptional response, such as an increase in the expression of components of the 26S proteasome and reduction in expression of some photosynthetic genes. Results are discussed in the context of the different pathways that may be involved in terpene emission.

Materials and Methods

Vector construction

35S:*DsRed-RNAi*. The Gateway pK7GWIWG2(II)-*Q10*::*DsRED* binary vector which contain the red fluorescent marker DsRED1 under the constitutive Arabidopsis Ubiquitin10 promoter was obtained from Dr. Sergey Ivanov (Ivanov *et al.*, 2012). To construct a *DsRed-*

RNAi construct, a 223 bp fragment was amplified using the primer pair attB1-DsRed-F/attB2-DsRed-R (Table S1) using the pK7GWIWG2(II)-*Q10*::*DsRED* vector as template. The amplicon was subcloned into the pDONR221 vector (Invitrogen) using the BP clonase II enzyme mix (Invitrogen) to generate an entry vector pDONR221-DsRed. In a subsequent step, the RNAi fragments were cloned into the pK7GWIWG2(II)-*Q10*::*DsRED* vector through LR recombination (Invitrogen). The vector was transferred to *Agrobacterium tumefaciens* AGL-0 by electroporation.

35S:*MtVAMP721e-RNAi* and 35S:*EV-RNAi*. *MtVAMP721e-RNAi* and *EV-RNAi* which both contain the *DsRed* reporter gene were kindly provided by Dr. Sergey Ivanov and Prof. Ton Bisseling (Ivanov *et al.*, 2012).

35S:*CST*. Genomic DNA was extracted from *Arabidopsis thaliana* (L.) Heynh (Col.0) leaves by the cetyltrimethylammonium bromide (CTAB) method (Chen *et al.*, 2008). A 2160-bp genomic DNA fragment containing 6 introns from start codon to stop codon of caryophyllene synthase (At5g23960) was amplified by PCR using the primer pair CST-F/CST-R (Table S1). The primers used introduce NcoI and NotI restriction sites which were used for cloning into ImpactVectorpIV1A_2.1, which contains the CaMV35S promoter and Rbcs1 terminator (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). The resulting pIV1A_2.1/CST was cloned into the pBinPlus binary vector using LR recombination (Invitrogen) (van Engelen *et al.*, 1995). The pBinPlus construct containing the 35S:CST was transferred to *Agrobacterium tumefaciens* AGL-0 using electroporation.

35S:*FaNES*. The strawberry linalool/nerolidol synthase, *FaNES*, with a synthetic intron (Yang *et al.*, 2008) was cloned into ImpactVectorpIV1A_2.4, which contains a CaMV35S promoter, plastid targeting signal and Rbcs1 terminator (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm) (obtained from Jan G. Schaart, Plant Breeding, Wageningen University). pIV1A_2.4/FaNES was cloned into the pBinPlus binary vector using LR recombination (Invitrogen) (van Engelen *et al.*, 1995).

Identification of NbVAMP72 genes and similarity analysis

To identify the *N. benthamiana VAMP72* genes, we used Arabidopsis *VAMP72* genes to BLAST against the *N. benthamiana* genome sequence which resulted in 14 *NbVAMP72* genes. To investigate the similarity between Arabidopsis, *Medicago truncatula* Gaertn. and *N. benthamiana* VAMP72s, protein sequences were used to construct a phylogenetic tree, which was generated using the alignment and the UPGMA method. The percentage of sequence similarity between *MtVAMP721e* and *NbVAMP72s* was calculated using BLAST (Altschul *et al.*, 1997) at the NCBI website (http://blast.ncbi.nlm.nih.gov/).

Transient expression in N. benthamiana

Transient expression in leaves of *N. benthamiana* was done as previously reported (Ting *et al.*, 2013). The total dosage of *Agrobacterium tumefaciens* within each experiment was kept constant for the different gene combinations. *N. benthamiana* leaves were infiltrated with *Agrobacterium* containing the following expression constructs: (1) *pBin+DsRed-RNAi*, (2) *MtVAMP721e-RNAi+DsRed-RNAi*, (3) *CST+DsRed-RNAi*, (4) *CST+MtVAMP721e-RNAi*, (5) *CST+EV-RNAi*, (6) *FaNES+DsRed-RNAi*, (7) *FaNES+MtVAMP721e-RNAi*, (8) *FaNES+EV-RNAi* (six plants per construct, one leave per plant). Infiltrated leaves were harvested seven days post agroinfiltration. Per sample two leaves were pooled, resulting in three biological replicates for each treatment.

Headspace trapping and analysis of volatiles by GC-MS

The volatile headspace of the agro-infiltrated *N. benthamiana* leaves was analysed by placing leaves in a vial with water to prevent dehydration, and placing the vial in a 1 L glass jar from which volatiles were collected for 30 minutes (between 13:45 pm and 14:15 pm) using dynamic headspace trapping as described (Houshyani *et al.*, 2013). The headspace samples were analysed by GC-MS as described (Houshyani *et al.*, 2013). Terpenoids were identified by comparison of mass spectra and retention time with authentic standards: caryophyllene (Sigma-Aldrich) and linalool (Fluka). Quantification of terpene levels was accomplished by determining the peak area of the characteristic *m/z* (133 for caryophyllene and 71 + 93 for linalool).

RNA isolation

After harvest, leaves were immediately frozen in liquid nitrogen and ground to a fine powder for RNA isolation. RNA was extracted using the RNeasy® Plant Mini kit (Qiagen, Hilden, Germany) with on-column RNase-Free DNase digestion (Qiagen, Hilden, Germany). The

quantity and quality of RNA for qRT-PCR was determined using a NanoDrop (NanoDrop Technologies, USA) and agarose gel electrophoresis. For RNA-seq analysis the RNA quality was checked using an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA).

Quantitative reverse transcription PCR (qRT-PCR)

cDNA was transcribed from 1 µg of RNA using the iScript cDNA Synthesis kit (BioRad, USA) and diluted 20-fold before use as template for qRT-PCR. A gene specific primer pair NbVAMP72c-F/NbVAMP72c-R was used for PCR. GADPH was used for normalization (Torres-Barceló *et al.*, 2008) (Table S1). Each PCR reaction mixture contained 1 µL cDNA template, 2 µL of each primer (3 µM), 10 µL of iQ™ SYBR® Green Supermix master mix (Biorad) and 5 µL of deionized water. The PCR was performed in the iCycler iQ5 system (BioRad) using a three-step temperature program: (1) 95°C for 3 min, (2) 40 cycles of 95°C for 10 s, 55°C for 30 s and (3) 95°C for 1 min, followed by melting curve analysis from 65 to 95°C. Relative expression values were calculated using the efficiency δCt (cycle threshold) method (Livak&Schmittgen, 2001).

RNA-sequencing and data analysis

Construction of the cDNA libraries and subsequent Illumina paired-end sequencing (Illumina Hiseq^{TM} 2000) was performed by BGI, China. The reads were assembled by BGI, using the available *N. benthamiana* genome sequence as template (Bombarely *et al.*, 2012), which was carried out using the SOAPaligner/SOAP2 program with the default settings (Li *et al.*, 2009). The gene expression level was calculated using the RPKM method (Reads Per kb per Million reads) (Mortazavi *et al.*, 2008). We used a false discovery rate (FDR) ≤ 0.05 and the absolute value of log2Ratio ≥ 1 as the threshold to select significance of gene expression differences. DEGs (differentially expressed genes) were then used for gene ontology (GO) functional analysis, using Blast2GO (Conesa *et al.*, 2005). For pathway analysis, we mapped all DEGs to terms in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Ogata *et al.*, 1999). Mapman was used to visualize gene expression changes, using the tomato pathways (http://mapman.gabipd.org/web/guest/mapmanstore) (Thimm *et al.*, 2004).

Results

Expression of VAMP72-RNAi enhances transient caryophyllene production

To investigate whether vesicle transport is involved in the transport of the volatile caryophyllene out of the cell we used transient expression in *N. benthamiana*. For the production of caryophyllene we used *TPS21* (*CST*) from Arabidopsis (encoding a caryophyllene synthase) under control of the constitutive CaMV35S promoter. For inhibition of *VAMP72* gene expression in *N. benthamiana* we used a *MtVAMP721e-RNAi* expression construct which previously was used to inhibit vesicle transport in *Medicago truncatula* (Ivanov *et al.*, 2012). We identified fourteen *VAMP72* genes in *N. benthamiana* (Fig. S1) and *MtVAMP721e* shows 67-83% sequence identity to these *N. benthamiana VAMP72* genes (Fig. S2, Table S2). Control experiments showed that *MtVAMP721e-RNAi* is at a least able to reduce *VAMP72c* mRNA levels in *N. benthamiana* when transiently expressed by agroinfiltration (Fig. S2). Because in transient expression the high levels of RNAi produced may generate pleiotropic effects (e.g. by titration of Argonaute complexes) we also tested the effect of a *DsRed-RNAi* construct on transient caryophyllene production. All vectors containing an RNAi construct also contained *DsRed* (see Fig. 1) and in all infiltration experiments the relative dosage of *CST* was kept constant. All transient expression treatments were evaluated at seven days post agroinfiltration by harvesting the infiltrated leaves and measuring the amount of caryophyllene that is released into the headspace in 30 minutes in the light.

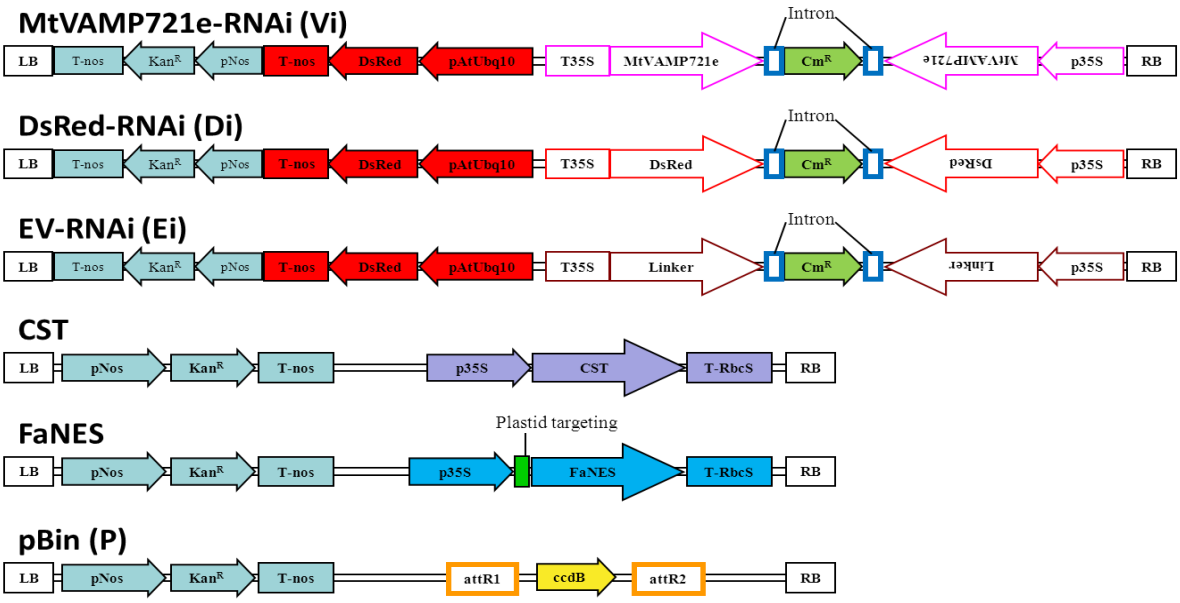

Figure 1. Maps of constructs used in this study.

Leaves transiently expressing *pBin+DsRed-RNAi* (*P+Di*) or *MtVAMP721e-RNAi+DsRed-RNAi* (*Vi+Di*) produced very low levels of caryophyllene, which must be attributed to endogenous *N. benthamiana* CST activity. However, leaves with *CST+DsRed-RNAi* (*C+Di*) produced substantial amounts of caryophyllene in the 30 minutes of headspace trapping (Fig. 2). We expected that, if vesicle transport is involved in caryophyllene transport inside the cell, inhibition of vesicle fusion by down regulation of *VAMP72* expression would decrease the emission of caryophyllene. However, when a similar dosage of *CST* was combined with *MtVAMP721e-RNAi* (*C+Vi*), caryophyllene emission was approximately 5-fold higher (Fig. 2). This indicates that inhibition of *N. benthamiana VAMP72* gene expression does affect the flux through the caryophyllene biosynthesis and/or emission pathway, but in an unexpected way. The experiment was repeated multiple times (n>5) yielding similar results.

Analysis of RNA-seq data

To better understand how the inhibition of vesicle transport, when combined with CST leads to enhanced caryophyllene emission we performed RNA-seq on the agroinfiltrated leaves. For this, the experiment was repeated (with similar results on caryophyllene emission) and RNA was extracted from the leaves seven days post-agroinfiltration of the combinations *P+Di*, *Vi+Di*, *C+Di* and *C+Vi*. Subsequently RNA was used for semi-quantitative analysis of the transcriptome by RNA-seq (see Materials and methods). Differentially expressed genes (DEGs) between the treatments (*Vi+Di*, *C+Di* or *C+Vi*) and the control *P+Di* were selected using a threshold of the absolute value of $\log_2$FoldChange≥1 and a false discovery rate (FDR)≤0.05. Inhibition of vesicle fusion (treatment *Vi+Di*) resulted in the up-regulation of only 144 and down-regulation of 214 genes. The transcriptional response to caryophyllene overproduction (treatment *C+Di*) was more pronounced, with 619 up-regulated and 769 down-regulated genes compared to the control (*P+Di*). In contrast, the specific transcriptional response to the combination of caryophyllene overproduction with inhibition of VAMP72 activity by *MtVAMP721e-RNAi* (treatment *C+Vi*) resulted in the up-regulation of 1353 genes and down regulation of 1822 genes compared with the control (*P+Di*), which means that the co-expression of *CST* and *MtVAMP721e-RNAi* causes a much larger transcriptional response than either treatment alone. Figure 3 shows the Venn diagrams of the overlap in transcriptional response in the different treatments (Fig. 3).

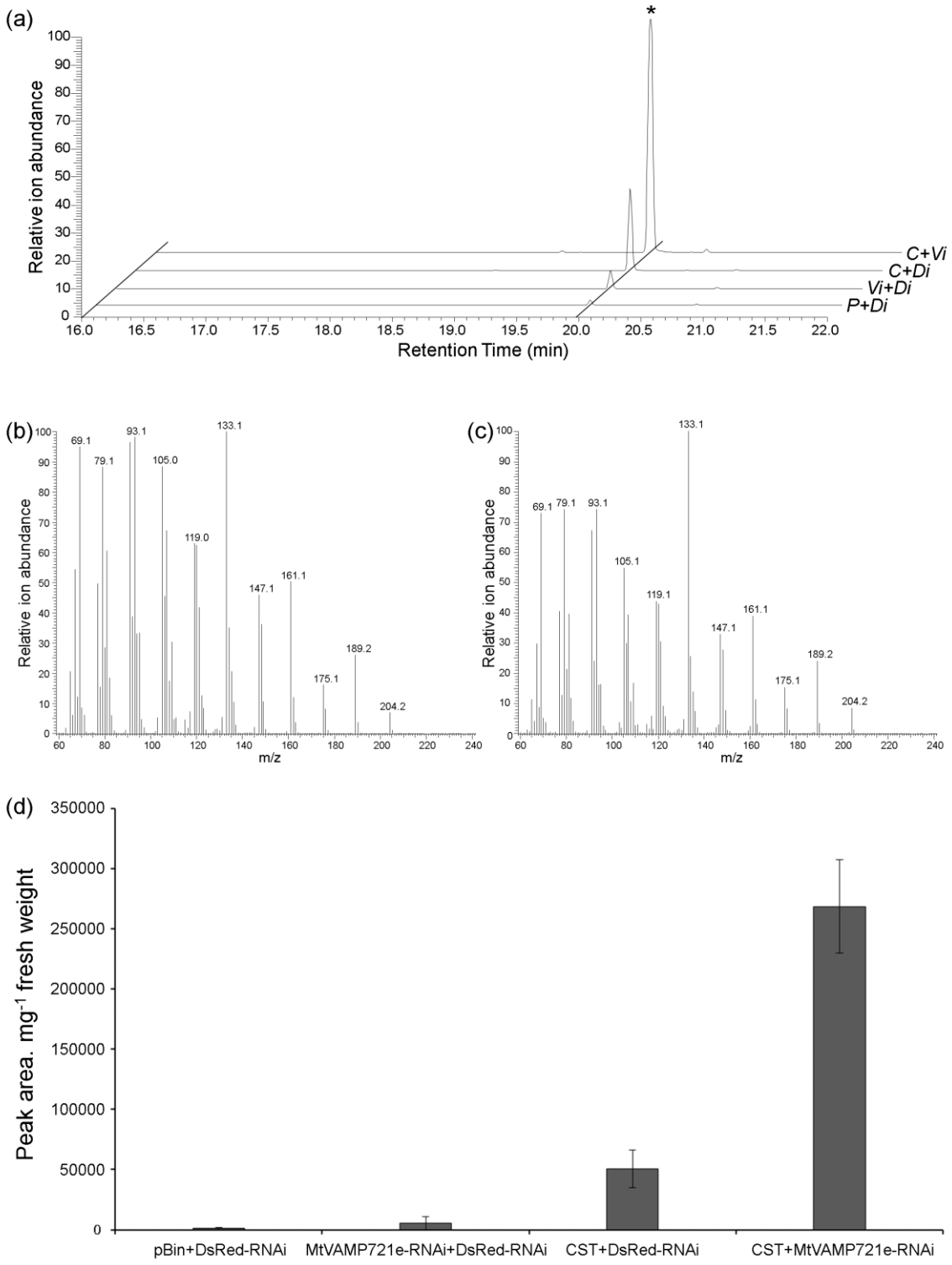

Figure 2. Caryophyllene emission from *N. benthamiana* leaves transiently expressing *pBin+DsRed-RNAi* (*P+Di*), *MtVAMP721e-RNAi+DsRed-RNAi* (*Vi+Di*), *CST+DsRed-RNAi* (*C+Di*) and *CST+MtVAMP721e-RNAi* (*C+Vi*).

(a) GC-MS chromatograms showing the caryophyllene emitted from *N. benthamiana* transiently expressing *P+Di*, *Vi+Di*, *C+Di* and *C+Vi*. *: caryophyllene.

(b) Mass spectrum of caryophyllene emitted by *N. benthamiana*.

(c) Mass spectrum of caryophyllene standard.

(d) Peak area of caryophyllene produced by leaves of *N. benthamiana*, as detected by headspace GC-MS analysis. Each bar represents the mean of three biological replicates +/- SE.

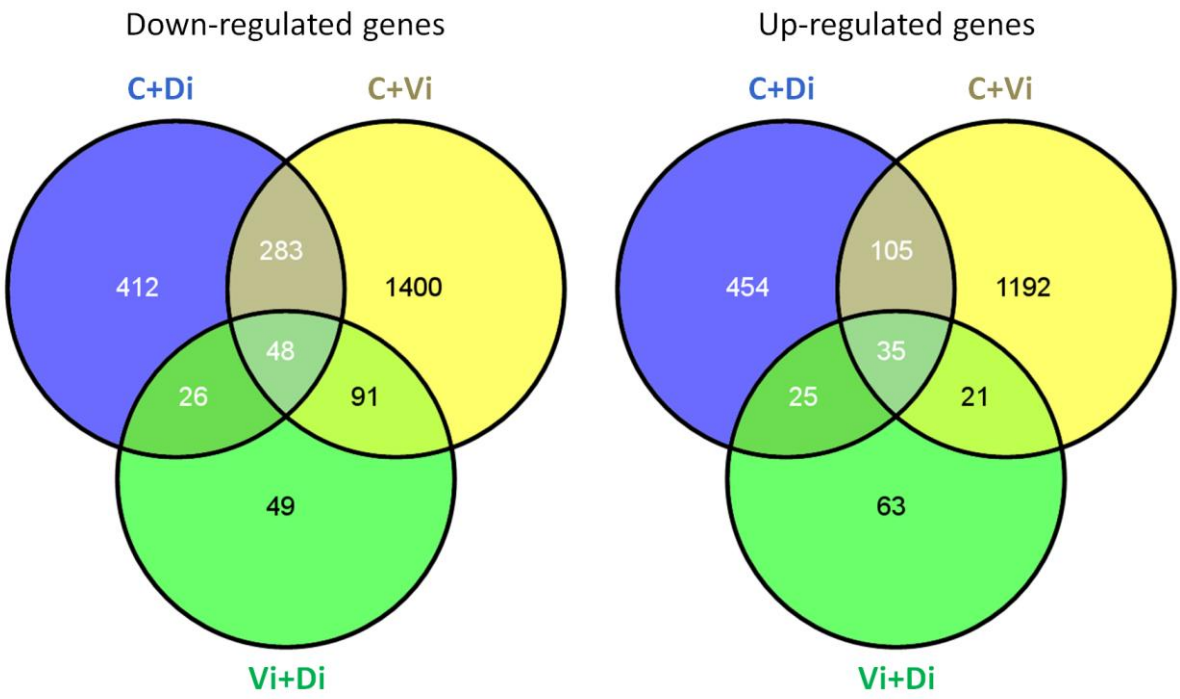

Figure 3. Venn diagrams showing common genes significantly up- and down-regulated in *N. benthamiana* leaves transiently expressing *CST* and *MtVAMP721e-RNAi*. Differentially expressed genes (DEGs) between the treatments (*Vi+Di*, *C+Di* or *C+Vi*) and the control *P+Di* (empty vector+ *DsRed-RNA*i) were filtered with a cut-off $\log_2$FoldChange≥1 and false discovery rate (FDR)≤0.05. *C+Di*, *CST+DsRed-RNAi*; *C+Vi*, *CST+MtVAMP721e-RNAi*; *V+Di*, *MtVAMP721e-RNAi+DsRed-RNA*i.

We also compared the expression of *DsRed* and *CST* in the different treatments. *DsRed* was a component of all treatments while the treatments *P+Di*, *Vi+Di* and *C+Di* also contained a *DsRed-RNAi* construct. Indeed, the expression level of *DsRed* was low in these treatments, and high in the *C+Vi* treatment (Fig. 4). *CST* was a component of the *C+Di* and *C+Vi* treatments, both having the same dosage of the *CST* expression construct in the agroinfiltration. Using the cut-off value $\log_2$FoldChange≥1, the *CST* transgene was not part of the DEGs set. The read counts of *CST* in the *C+Di* and *C+Vi* treatments indicate a slightly higher count in the *C+Vi* (1.6-fold) (Fig. 4). This is not sufficient to account for the measured increased emission of caryophyllene in this treatment (Fig. 2).

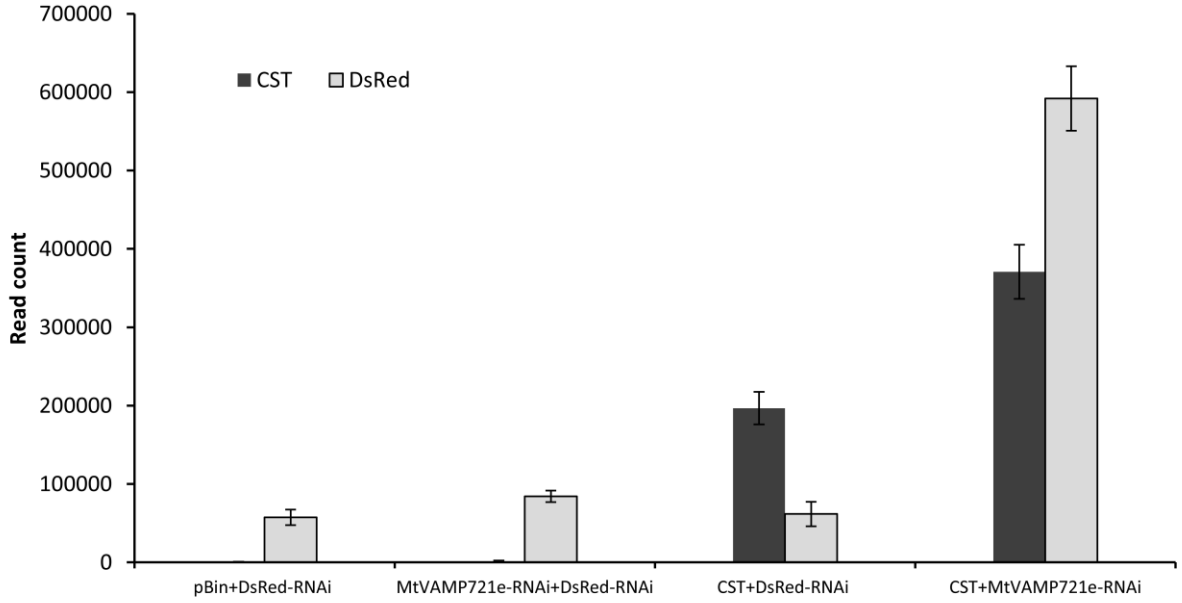

Figure 4. Total number of *CST* and *DsRed* reads in leaves upon agroinfiltration with *pBin+DsRed-RNAi*, *MtVAMP721e-RNAi+DsRed-RNAi*, *CST+DsRed-RNAi* and *CST+MtVAMP721e-RNAi*.

GO-term enrichment analysis

To get insight in the biological processes affected by the different treatments, a GO-term enrichment analysis was performed on the DEG sets, using P≤0.05 (Tables S3-S5). In addition, the DEGs of the different treatments were mapped to the reference canonical plant pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Tables S6-S8). For each of the three treatments, the top ten of pathways that were most affected are shown in Table 1. The pathways that were most strongly affected by the different treatments were also visualized via the MapMan tool (Thimm *et al.*, 2004). An overview of all DEGs between *Vi+Di*, *C+Di* and *C+Vi* versus control *P+Di* for cellular metabolism is shown in Figure 5. A list of all up- or down-regulated DEGs corresponding to MapMan functional categories is provided in Table S9.

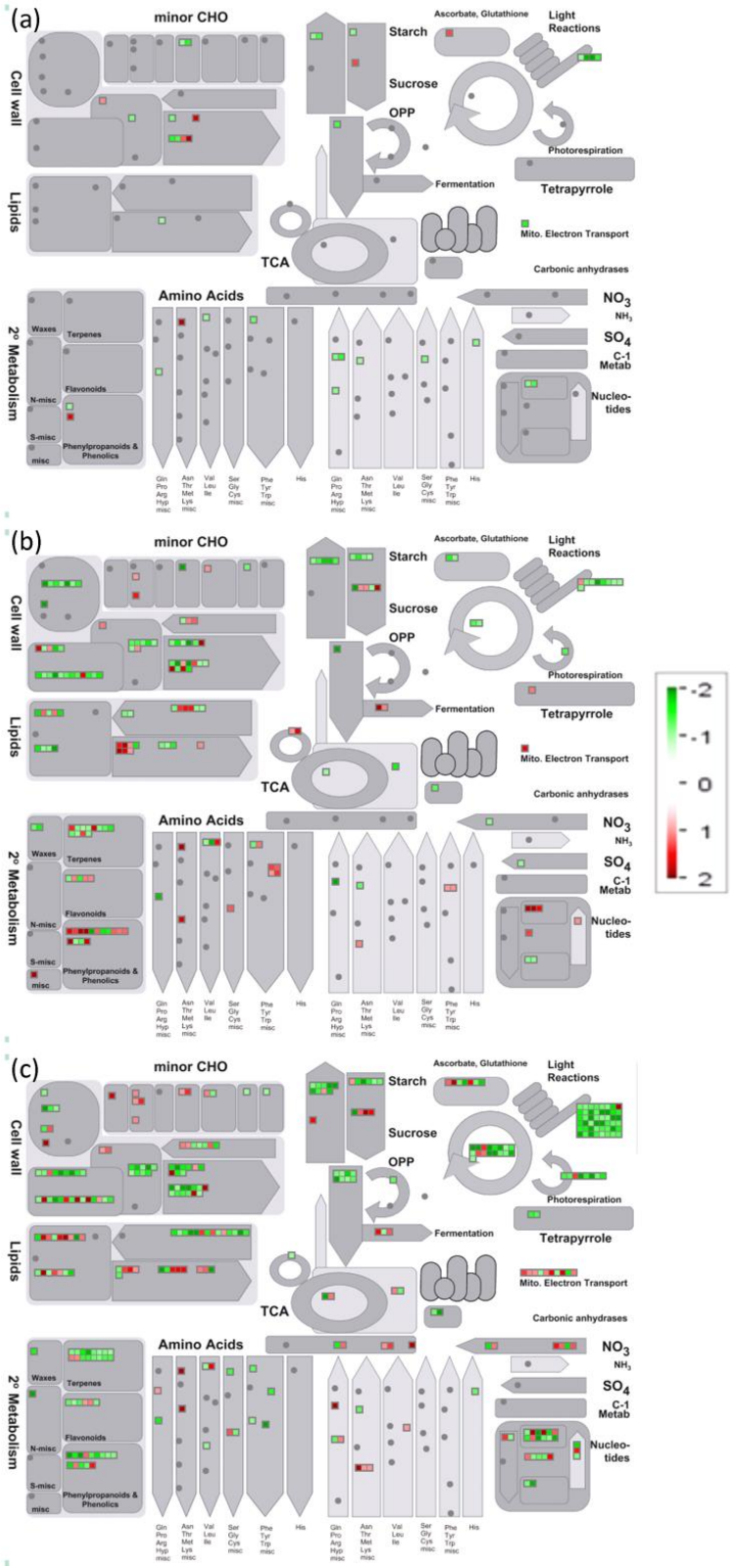

Figure 5. MapMan overview of cellular metabolism showing all differentially expressed genes (DEGs) between treatments: (a) *MtVAMP721e-RNAi+DsRed-RNAi*, (b) *CST+DsRed-RNAi*, (c) *CST+MtVAMP721e-RNAi* and control (*pBin+DsRed-RNAi*). Individual genes are represented by squares. Genes significantly up- or down-regulated (log_2FoldChange≥1 and FDR≤0.05) relative to the control (*pBin+DsRed-RNAi*) are indicated in red and green, respectively. Scale bars display log_2FoldChange. CHO, carbohydrates; OPP, oxidative pentose phosphate pathway; TCA, tricarboxylic acid cycle.

Transcriptional response to Vi+Di treatment

Analysis of the DEGs that are specific for *MtVAMP721e-RNAi* shows that the *VAMP72* genes of *N. benthamiana* are most strongly affected (as expected). As a result, the SNARE interaction pathway of vesicular transport shows up for this treatment as the pathway that is strongly affected (Table 1). Indeed, the expression of eleven out of the fourteen different *N. benthamiana VAMP72* genes was down regulated (Table S10), while other type VAMP genes did not show a significant change in expression. This indicates that the *M. truncatula MtVAMP721e-RNAi* targets multiple *VAMP72* genes in *N. benthamiana*, but no other related *VAMP* genes. Using the threshold of $Q \leq 0.05$, no other pathways were significantly affected by the inhibition of VAMP72 (Tables 1, S6).

Table 1. The top 10 considerably changed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of differentially expressed genes (DEGs) between treatments and control with significant statistical support ($Q \leq 0.05$)

Top #	MtVAMP721e-RNAi +DsRed-RNAi	CST+DsRed-RNAi	CST+MtVAMP721e-RNAi
1	SNARE interactions in vesicular transport (9)*	Glycosaminoglycan degradation (14)	Proteasome (75)
2		Glycosphingolipid biosynthesis-ganglio series (11)	Photosynthesis (49)
3		Biosynthesis of secondary metabolites (144)	Carbon fixation in photosynthetic organisms (39)
4		Other glycan degradation (20)	Carotenoid biosynthesis (39)
5		Starch and sucrose metabolism (41)	Other glycan degradation (27)
6		Metabolic pathways (230)	Glyoxylate and dicarboxylate metabolism (24)
7		Stilbenoid, diarylheptanoid and gingerol biosynthesis (22)	Metabolic pathways (489)
8		Carotenoid biosynthesis (19)	Biosynthesis of secondary metabolites (267)
9		Ascorbate and aldarate metabolism (15)	alpha-Linolenic acid metabolism (22)
10		Phenylpropanoid biosynthesis (33)	Ascorbate and aldarate metabolism (25)

* Number of DEGs with pathway annotation.

Transcriptional response to C+Di treatment

Analysis of the transcriptional response to *CST* indicates that ectopic production of caryophyllene is creating a substantial pleiotropic effect in the leaves. GO term enrichment analysis of the specific DEGs shows that glucan metabolic processes in cell periphery and cell wall are affected by the ectopic caryophyllene production (Table S5). In addition, ectopic production of caryophyllene had a strong negative effect on the expression of genes involved in glycosaminoglycan degradation, suggesting a reduced breakdown of glycosaminoglycans (Table S7). Projection of the data in Mapman shows that *CST* expression also resulted in down regulation of genes involved in cell wall biosynthesis (Fig. 5b), while biotic response, heat stress response, calcium signaling and development genes were significantly induced (Table S9).

Transcriptonal response to the C+Vi treatment

When caryophyllene production was combined with inhibition of VAMP72 function, some of the specific transcriptional responses to *CST* alone were abolished. For instance, compared with the *C+Di* treatment the *C+Vi* treatment does no longer show a significant increase in DEGs involved in biotic stress, heat stress and calcium signaling (Table S9). Most interestingly, the combination of *CST+MtVAMP721e-RNAi* (*C+Vi*) seems specifically to upregulate processes related to protein turnover by the 26S proteasome complex and downregulate processes related to photosynthesis (Table S3). This is also clear in the output of MapMan for the ubiquitin dependent degradation pathway, which shows up-regulation of almost every gene from the 26S proteasome only for the *C+Vi* treatment (Fig. 6), while for photosynthesis almost all genes were down-regulated (Fig. 5c). Also genes involved in cell wall-related processes were down-regulated in *C+Vi* (Fig. 5c). For instance, some genes of the phenylpropanoid pathway (mainly lignin biosynthesis) showed reduced expression (Table S9, Bin 16.2). Besides the many genes involved in cellular processes for which expression was down-regulated by the *C+Vi* treatment, a number of transcription factors were up-regulated by this treatment (e.g. AP2/EREBP, C2C2(Zn) CO-like, MADS box transcription factor family) (Table S9, Bins 27.3.3, 27.3.7 and 27.3.24). This may explain why so much more genes were regulated in *C+Vi* compared with the other treatments (*Vi+Di*, *C+Di*) in which just few transcription factors were induced (Fig. 3, Table S9).

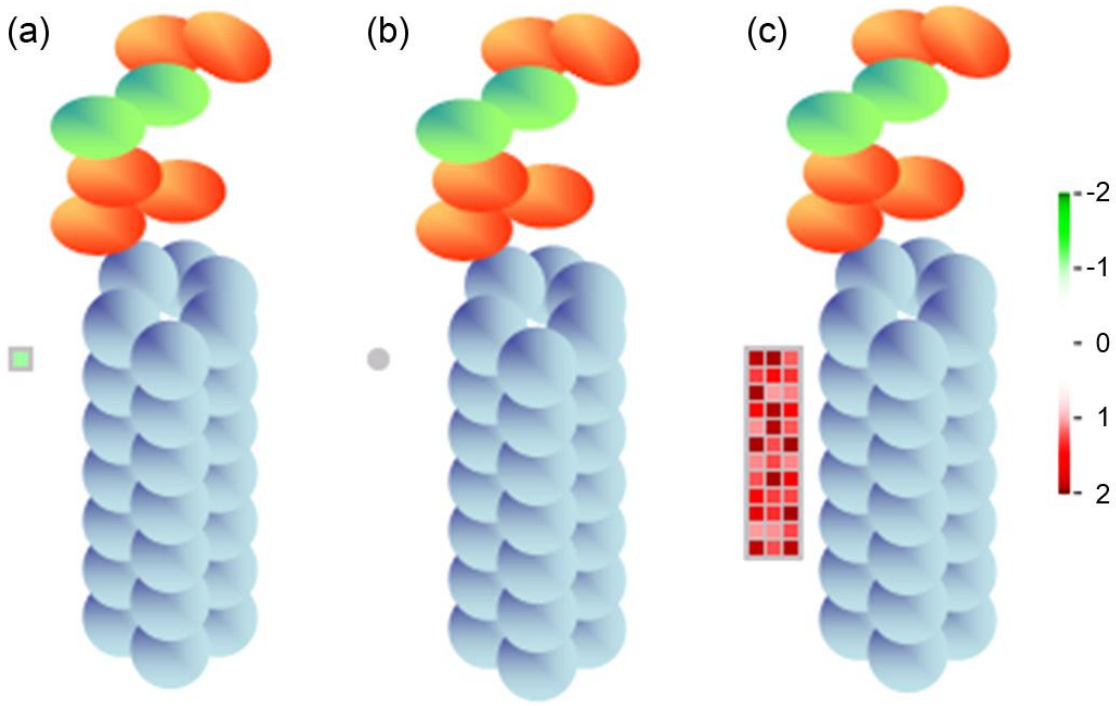

Figure 6. MapMan overview of the ubiquitin dependent degradation pathway showing all differentially expressed genes (DEGs) between treatments: (a) *MtVAMP721e-RNAi+DsRed-RNAi*, (b) *CST+DsRed-RNAi* and (c) *CST+ MtVAMP721e-RNAi* and control (*pBin+DsRed-RNAi*). Individual genes are represented by *small squares*. Genes significantly up- or down-regulated (log_2FoldChange≥1 and FDR≤0.05) relative to the control (*pBin+DsRed-RNAi*) are indicated in red and green, respectively. Scale bars display log_2FoldChange.

Indications of reduced proteasome activity in the C+Vi treatment

Results from the transcriptional response to (*C+Vi*) indicate a coordinated up regulation of genes of the 26S proteasome complex. From yeast and mammalian studies it is known that transcriptional control of proteasome genes is under feedback regulation by a transcription factor (NF-E2-related factor 2, Nrf2) which itself is targeted for degradation by the 26S proteasome (Kobayashi&Yamamoto, 2005). When proteasome function is impaired, the activity of this transcription factor increases and consequently so does the expression of proteasome genes (Meiners *et al.*, 2003). Although such feedback regulation has not been described for plants yet, the up regulation of proteasome gene expression in response to the *CST+MtVAMP721e-RNAi* treatment in *N. benthamiana* (Fig. 6c) could be indicative of reduced proteasome activity.

There are multiple observations that may confirm impaired proteasome function as a consequence of the *C+Vi* treatment. The first is the increased caryophyllene production in response to the *C+Vi* treatment. Although the transcriptional activity of *CST* was slightly higher in this treatment than in the *C+Di* treatment (Fig. 4), this cannot account for the 5-fold

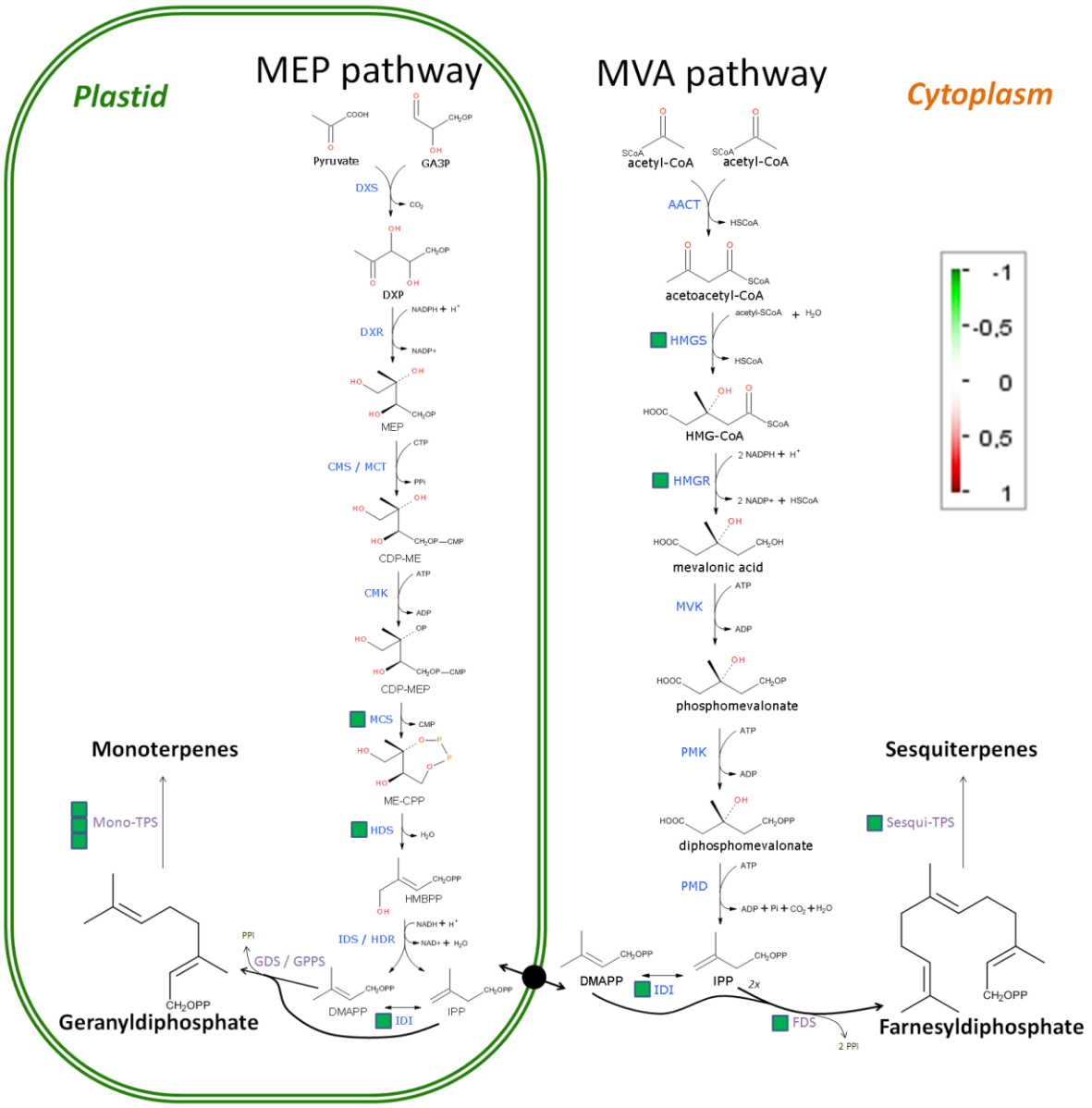

Figure 7. Expression pattern of terpene biosynthesis genes in *CST+MtVAMP721e-RNAi*. Individual genes are represented by squares. Genes significantly down-regulated (log_2FoldChange≥1 and FDR≤0.05) relative to the control (*pBin+DsRed-RNAi*) are indicated in green. Scale bars display log_2FoldChange.

increase in caryophyllene emission by leaves agro-infiltrated with *C+Vi*. Moreover, the *C+Vi* treatment resulted in down-regulation of the expression of multiple genes of the MVA pathway (*HMGS*, *HMGR*, *IDI*) and the MEP pathway (*MCS*, *HDS*, *IDI*), as well as the first step towards sesquiterpene biosynthesis (*FDS*) (Fig. 7). This effect was specific for the *CST+MtVAMP721e-RNAi* combination, and no down-regulation of these genes was observed for the expression of *CST* or *MtVAMP721e-RNAi* alone. The increased caryophyllene production in the *CST+MtVAMP721e-RNAi* treatment may therefore be most easily

explained by a decreased CST enzyme turnover, possibly as a consequence of reduced proteasome activity.

A second indication of impaired proteasome function in the *C+Vi* treatment may be the coordinated down-regulation of photosynthesis genes. Recently it was reported that the expression of photosynthesis genes is mainly controlled by the transcriptional repressor PIF3, for which protein stability is regulated by ubiquitination and proteasome-mediated degradation (Liu *et al.*, 2013). Impaired proteasome activity could result in stabilization of the *N. benthamiana* PIF3 homolog, resulting in enhanced suppression of photosynthesis genes.

Only the combination of C+Vi results in stabilization of DsRed protein

To get further information on protein stability in the different treatments, we analysed DsRed protein accumulation. First, we tested whether other RNAi constructs have an effect on transient CST activity, by comparing the effect of an *EV-RNAi*, (*Ei*, which only contains a small linker) or *DsRed-RNAi* with the effect of the *MtVAMP721e-RNAi* (*Vi*) on CST activity. Leaves transiently expressing *CST+DsRed-RNAi* or *CST+EV-RNAi* emitted similar levels of caryophyllene while the inhibition of the expression of *VAMP72* again stimulated caryophyllene emission (Fig. 8) just as before (Fig. 2). The effect of *Vi* on transient caryophyllene production is therefore specific and not due to some general effect of RNAi constructs.

Because both the *Ei* and *Vi* vector contain a *35S:DsRed* gene, we could also compare the accumulation of DsRed protein in the combinations *C+Ei* and *C+Vi*. Leaves infiltrated with *C+Vi* visually had a higher DsRed level than the empty vector control (Fig. 9a). Indeed, semi-quantification of the signal in the red and green components of the images - setting the average signal for the red (DsRED) and green (chlorophyll) channel of the *C+Ei* image at 100% - showed an increase of about 25% for the red signal in the *C+Vi* treatment (Fig. 9c), indicating reduced turnover of the DsRed protein in this treatment. In the same leaf the green signal was reduced by 10% (Figs 9a,c) which could be in accordance with the reduction in photosynthesis related gene expression (Fig. 5c). DsRed protein turnover is supposed to be mediated by the proteasome (Verkhusha *et al.*, 2003). These results thus support the assumption that the *C+Vi* treatment results in an inhibition of 26S proteasome activity (and as a consequence in an increase in proteasome-related gene expression, Fig. 6c).

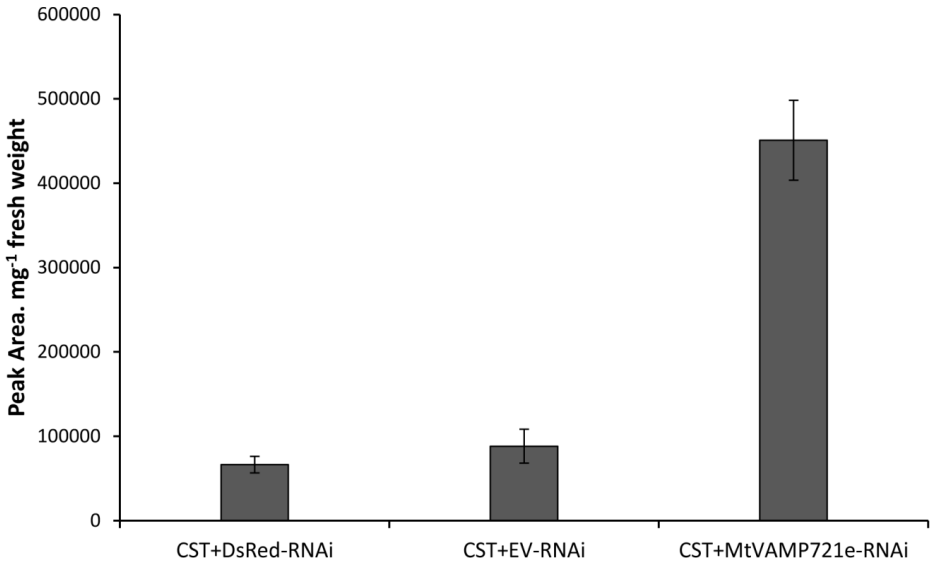

Figure 8. Caryophyllene emission in leaves upon transient expression in *N. benthamiana* of *CST+DsRed-RNAi*, *CST+EV-RNAi* and *CST+MtVAMP721e-RNAi*.

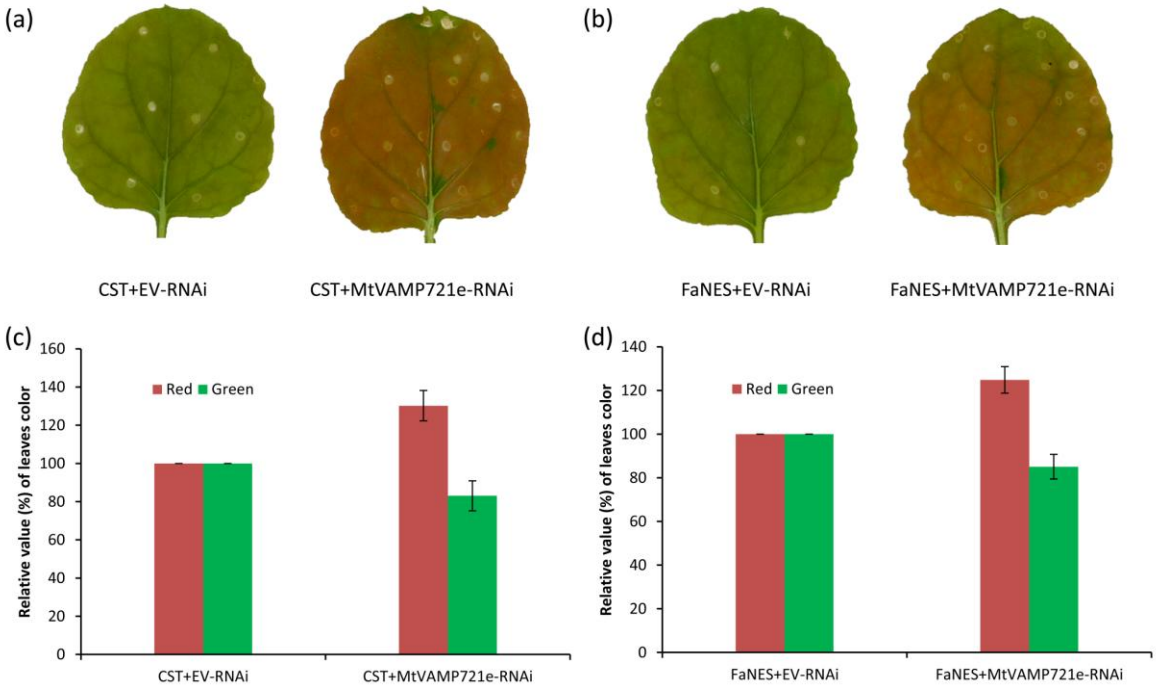

Figure 9. DsRed and chlorophyll abundance in leaves of *N. benthamiana* upon transient expression of caryophyllene synthase (*CST*) or linalool synthase (*FaNES*) with or without *MtVAMP721e-RNAi*.
(a) Image of leaves transiently expressing *CST+EV-RNAi* and *CST+MtVAMP721e-RNAi*, (b) Image of leaves transiently expressing *FaNES+EV-RNAi* and *FaNES+MtVAMP721e-RNAi*, (c) Red and green color in leaves transiently expressing *CST+EV-RNAi* and *CST+MtVAMP721e-RNAi*, (d) Red and green color in leaves transiently expressing *FaNES+EV-RNAi* and *FaNES+MtVAMP721e-RNAi*.

Also other terpene synthases increase protein stability

To assess whether the effect of *CST+MtVAMP721e-RNAi* is specific for caryophyllene, or whether it can also be elicited by the expression of other terpene synthases in combination with *VAMP72-RNAi*, we repeated the experiment with a linalool synthase (plastid targeted FaNES, cloned from strawberry). Transient expression of *FaNES* in *N. benthamiana* leaves resulted in the emission of the monoterpenoid linalool into the headspace (Fig. 10). Also in the case of linalool, the combination *FaNES+Vi* strongly enhanced linalool emission (Fig. 10). The transient expression of linalool synthase in *N. benthamiana* probably also results in the production of glycosylated linalool products. Preliminary GC-MS analysis of leaf extracts after treatment with a mix glycosidases (Viscozyme L) suggested that also the accumulation of glycosylated linalool products was boosted by the *FaNES+Vi* treatment compared with the *FaNES+Ei* treatment. Analysis of the leaves showed that, just as for caryophyllene (Figs. 9a,c), the combination of linalool production with inhibition of VAMP72 resulted in stabilization of DsRED, while chlorophyll levels were reduced (Figs. 9b,d). Combined, the results show that both sesquiterpenes and monoterpenes, in combination with inhibition of *VAMP72* genes cause an impairment in proteasome function. Apparently plants have a similar positive feedback regulation of proteasome gene expression as described for yeast and mammals, and this results in enhanced proteasome gene expression (Fig. 5).

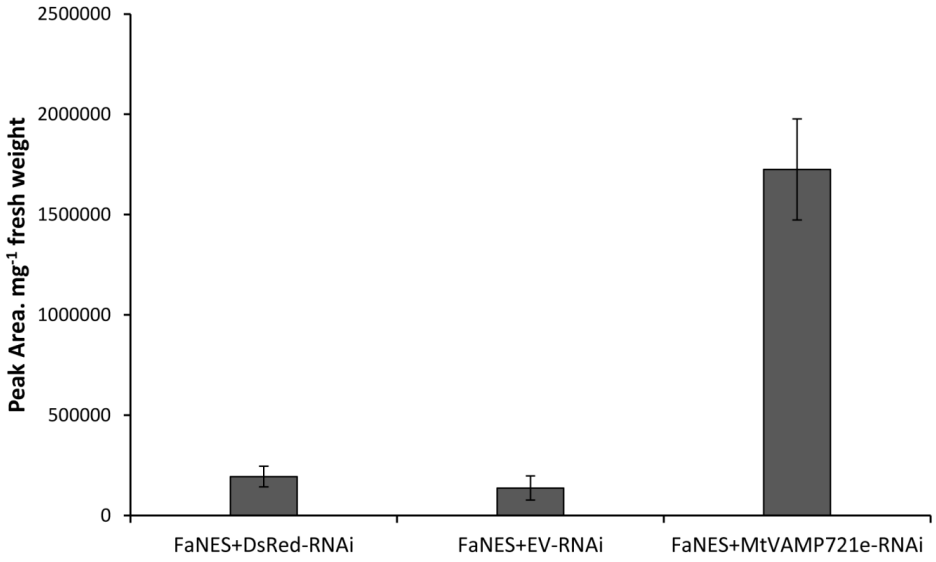

Figure 10. Linalool emission in leaves transiently expressing *FaNES+DsRed-RNAi*, *FaNES+EV-RNAi* and *FaNES+MtVAMP721e-RNAi*.

Discussion

Inhibition of vesicle fusion affects terpene emission capacity?

In this study, we investigated whether vesicles are involved in transport of volatile terpenes within the cell and whether vesicle fusion plays a role in their emission to the apoplast. A study on *Sauromatum guttatum* flowers showed fusion of vesicles from the ER with the plasma membrane and this correlated with the heat-induced release of sesquiterpenes (α-copaene and caryophyllene) (Skubatz *et al.*, 1995), hinting at the involvement of vesicle transport and fusion in terpene emission. These data suggest that terpenes may be stored in vesicles that are transported to the plasma membrane. Our results show that a heterologous *MtVAMP721e-RNAi* construct (*Vi*) is able to suppress multiple *VAMP72* genes of *N. benthamina* (Fig. S2, Table S10), but when *Vi* is combined with caryophyllene production this results in enhanced emission levels rather than a reduction as anticipated (Figs 2, 8). Similar results were obtained for the overexpression of the monoterpenoid linalool in combination with *Vi*. Also emission of linalool was not blocked but actually enhanced, while also accumulation of linalool glycosides was enhanced (Fig. 10). Combined, these results suggest that vesicle fusion, as controlled by the VAMP72 proteins, is not required for either the sesquiterpene or monoterpene emission pathway. However, there is a strong interaction between the ectopic production of the two terpenoids with the inhibition of VAMP72 function. There is a strong increase in terpene production, an increase in DsRed protein stability and an overall strong transcriptional response to the *C+Vi* treatment. All this makes it difficult to rule out the possibility that there is involvement of vesicle fusion in terpene emission.

Impaired proteasome activity in leaves agro-infiltrated with TPS+Vi

Although the RNA-seq data show a small increase in *CST* transgene expression in *C+Vi* compared to *C+Di* (Fig. 4), the expression of several genes of the MVA precursor pathway was down-regulated in *C+Vi*. Thus the transcription results do not explain the 5-fold increase in caryophyllene emission by transient expression of *C+Vi* in *N. benthamiana* (Figs. 2, 8). However, increased volatile emission could be explained if proteins in the biosynthesis pathway are more stable. An indirect indication that proteins may indeed be more stable in the *C+Vi* treatment is the increase in DsRed protein stability (Fig. 9). This increased protein stability of DsRed, and possibly of other proteins involved in the terpenoid pathway, including the proteins introduced through transient expression, may be the result of decreased

proteasome activity. From studies in mammals and yeast, it is known that proteasome gene expression is self-regulated by transcription factor(s) that themselves are targeted for proteasome mediated destruction (Kobayashi&Yamamoto, 2005, Villeneuve *et al.*, 2010). As a result, impairment in proteasome function causes a stabilization of these transcription factors, resulting in an up-regulation of the expression of proteasome-associated genes (Dreger *et al.*, 2010, Meiners *et al.*, 2003). Presumably, proteasome homeostasis in plants is under a similar feedback control and an impairment of proteasome function by the *C+Vi* treatment could therefore explain the up-regulation of the expression of proteasome-associated genes in our study (Fig. 6c, Table S9), possibly mediated by the stabilization of a *N. benthamiana* transcription factor. In general, the specific down- or up-regulation of gene expression in response to the *C+Vi* or *FaNES+Vi* treatments could thus be ascribed to stabilization of transcriptional repressors or activators, respectively. Indications that transcriptional repressors are stabilized in the *C+Vi* treatment comes from the coordinated suppression of photosynthesis genes (Fig. 5c, Table S9). In Arabidopsis it has been shown that the transcriptional repressor PIF3 is a key controller of the expression of a subset of photosynthesis-related genes (Liu *et al.*, 2013). If the presumed proteasome malfunction caused by *C+Vi* results in suppression of the *N. benthamiana* homolog of PIF3 we could expect a similar set of genes being reduced in expression. Indeed, the homologs of a substantial part of the photosynthesis genes that are targeted by PIF3 in Arabidopsis, also show reduced expression in the *C+Vi* treatment in *N. benthamiana* (Fig. 5c, Table S9).

Multiple pathways for terpene emission?

Hydrophobic compounds such as the sesquiterpene caryophyllene and the monoterpene linalool are likely to sequester in membranes and could therefore be passive passengers of vesicle trafficking. Terpene emission from plant cells may involve multiple pathways, e.g. (1) vesicle transport and fusion, (2) specific carrier proteins that transport to specific membrane transporters and (3) direct diffusion between the ER or plastidial (stromule) membrane and the plasma membrane, possibly through contact sites (Fig. 11a). An inhibition of VAMP72 activity possibly affects both the vesicle fusion pathway and the membrane transporter pathway, as the membrane transporter first needs to be inserted into the plasma membrane through vesicle fusion before it can transport anything. Inhibition of VAMP72 would then only leave the direct diffusion pathway as a means to emit volatile terpenes. Inhibition of VAMP72 could thus initially result in a higher accumulation of terpenes within the cell. Presumably this high endogenous terpene concentration causes inhibition of proteasome

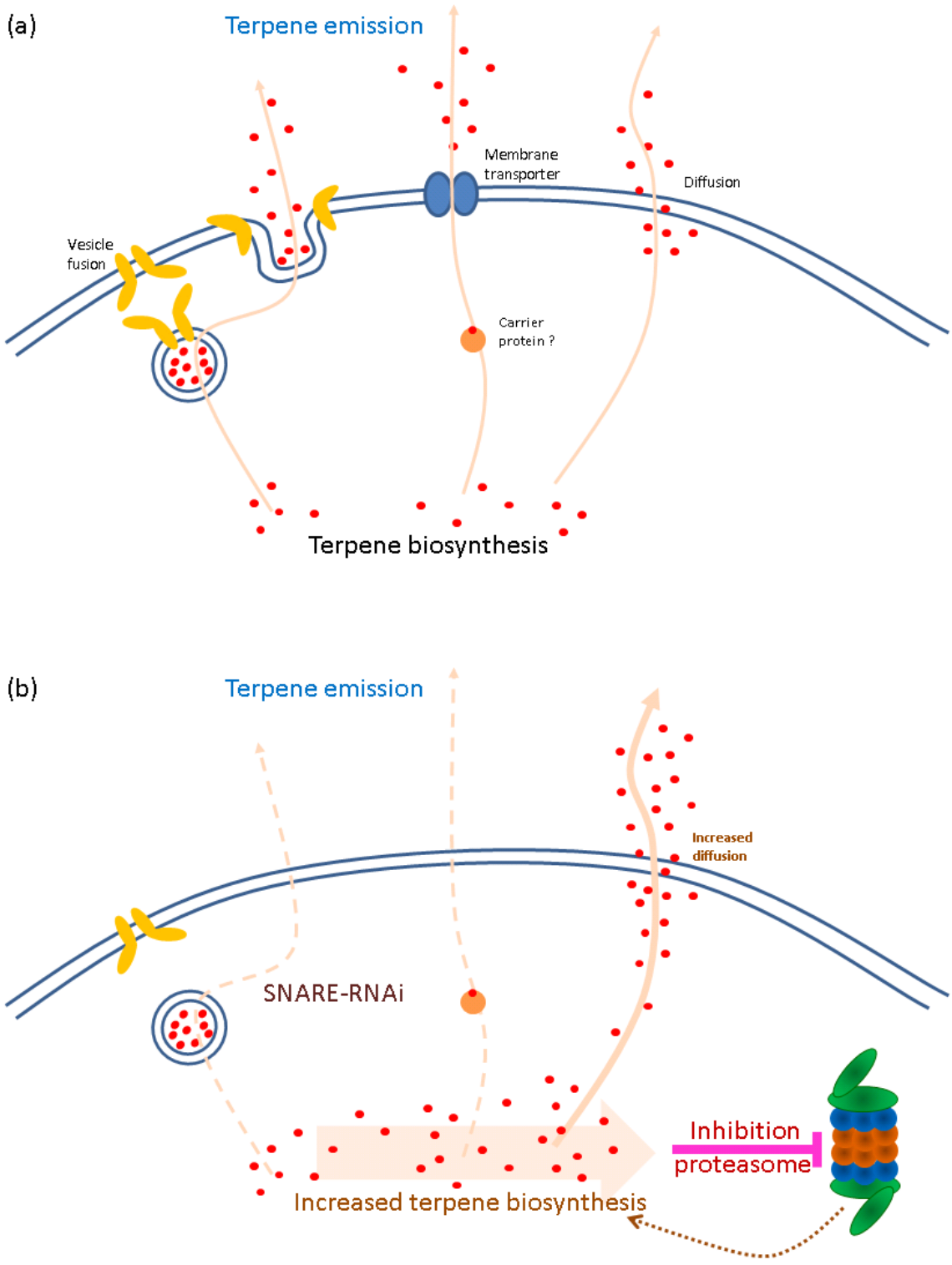

Figure 11. Model for terpene emission with and without inhibition of SNARE.
(a) Three possible pathways for terpene emission to the headspace: (1) vesicle transport and vesicle fusion, (2) specific carrier proteins for transport of terpenes to membrane transporters and (3) diffusion between ER or plastidal stromule membrane and plasma membrane contact sites.
(b) An inhibition of SNARE (VAMP72) activity could affect both the vesicle fusion pathway and the membrane transporter pathway. The higher endogenous terpene concentration causes inhibition of proteasome function, leading to a positive feedback on endogenous terpene accumulation. Thus, this increases the direct diffusion pathway.

function, through an as yet unknown mechanism. Due to the resulting increase in stability of proteins in the MVA pathway and the terpene synthases themselves the terpene production is increased, leading to even higher terpene accumulation and stronger inhibitory effect on the proteasome. However, at the same time the higher endogenous terpene content of the cell may result in a higher emission rate through the direct diffusion pathway which is independent of vesicle fusion (Fig. 11b).

Key to this hypothesis is the causal relationship between higher than normal endogenous terpene concentrations and the inhibition of the proteasome (Fig. 11b). In mammalian systems it has been shown that certain terpenes can indeed affect proteasome activity. For instance, Yang and coworkers showed that the triterpene, celastrol, can activate apoptosis triggered by inhibition of proteasome activity (Yang *et al.*, 2006). Schumacher and coworkers showed that the sesterterpene, heteronemin, also inhibits proteasome activity and induces apoptotic cell death (Schumacher *et al.*, 2010). Recently, it has been indicated that proteasome inhibitors can act as potential drug for cancer therapy (Orlowski&Kuhn, 2008). Several recent studies have suggested that these inhibitors - through a decrease in proteasome activity – induce apoptosis in the tumor cell (Adams, 2004) (Almond&Cohen, 2002, Kupperman et al., 2010). In future research we will investigate whether high levels of caryophyllene or linalool cause protein damage hence inhibiting proteasome functioning.

Concluding remarks

Our study shows for the first time that vesicle transport is involved in terpene (caryophyllene and linalool) emission. Inhibition of VAMP72, and thereby of vesicle fusion with the plasma membrane, resulted in the inhibition of normal terpene transport. The resulting accumulation of terpenes somehow decreases proteasome activity resulting in increased protein (transient expressed terpene synthase and DsRed) stability in *N. benthamiana*. Moreover, our finding about terpenes and proteasome activity can provide new insights for medical research, such as an investigation of novel leads for anticancer therapies.

Acknowledgments

H-M.T. was funded by the graduate school of Experimental Plant Sciences (EPS). We thank Dr. Sergey Ivanov and Prof. Ton Bisseling for providing 35S:*MtVAMP721e-RNAi* and 35S:*EV-RNAi* constructs. We thank Jan G. Schaart for providing the pIV1A_2.4/FaNES construct. We would like to thank Francel W.A. Verstappen, Iris Kappers, Geert Stoopen, Roland Mumm, Berhane Weldegergis and Peter Tóth for their assistance with headspace trapping and GC-MS analysis. We thank Desalegn Woldes Etalo for help with MS data analysis, Esmer Jongedijk for help with the terpene biosynthesis pathway figure and Ralph Bours for help with Image J software.

Supporting Information

Table S1. List of primers used for the vector construction and gene expression analysis by qRT-PCR.

Primer name	Primer sequence (5' to 3') *
CST-F	TTTT*CCATGG*GGAGTGAAGTCAACCGTCCA
CST-R	TTGGTT*GCGGCCGC*TCAAATGGGTATAGTTTCAA
*att*B1-DsRed-F	**GGGGACAAGTTTGTACAAAAAAGCAGGCT**TCGGGTGCTTCACGTACACCTT
*att*B2-DsRed-R	**GGGGACCACTTTGTACAAGAAAGCTGGGT**CGCTCCTCCAAGAACGTCATC
NbVAMP72c-F	ATGGAAAACATTGAGAAGGTTCTTGACC
NbVAMP72c-R	CAGTATACCTGAAGTCTTGTGCCTGG
GADPH-F	GGTGTCAAGCAAGCCTCTCAC
GADPH-R	GATGCCAAGGGTGGAGTCAT

***** Restriction sites are shown in italics and *att* sites in bold.

Table S2. Percentage of sequence identity (%) between *MtVAMP721e* and *NbVAMP72s*.

Gene	MtVAMP721e	NbVAMP72													
		a	b	c	d	e	f	g	h	i	j	k	l	m	n
MtVAMP721e	100	78	81	83	78	77	67	67	78	74	75	76	80	71	73

Table S3. Gene Ontology (GO) Cellular Component classification with *P*-value (if $P \leq 0.05$) (after Bonferroni Correction) of differentially expressed genes (DEGs) between treatments (*Vi+Di*, *C+Di* and *C+Vi*) with the control (*P+Di*).

Gene Ontology term (Cellular Component)	***Vi+Di***	***C+Di***	***C+Vi***
proteasome complex			1.10E-40
proteasome core complex			9.66E-17
proteasome regulatory particle			6.55E-07
proteasome accessory complex			6.55E-07
protein complex			6.42E-06
chloroplast thylakoid			2.20E-17
chloroplast part			7.68E-17
plastid thylakoid			1.51E-16
organelle subcompartment			1.82E-16
chloroplast			2.52E-16
thylakoid			7.82E-16
plastid part			1.64E-10
thylakoid part			1.02E-08
photosystem			2.12E-08
photosynthetic membrane			4.35E-07
plastid			3.60E-06
photosystem I			3.74E-06
chloroplast stroma			0.00027
plastid envelope			0.00042
photosystem II	0.04798		0.00092
plastid stroma			0.00284
extracellular region		0.00346	1.10E-06

external encapsulating structure		5.33E-11	0.00457
cell periphery		7.78E-11	0.00787
cell wall		7.63E-12	0.04836
intrinsic to membrane		0.00033	
anchored to membrane		0.01091	

*V+Di, MtVAMP721e-RNAi+DsRed-RNA*i; *C+Di, CST+DsRed-RNAi*; *C+Vi, CST+MtVAMP721e-RNAi; P+Di, pBin+DsRed-RNAi*.

Table S4. Gene Ontology (GO) Molecular Function classification with *P*-value (if $P \leq 0.05$) (after Bonferroni Correction) of differentially expressed genes (DEGs) between treatments (*Vi+Di*, *C+Di* and *C+Vi*) with the control (*P+Di*).

Gene Ontology term (Molecular Function)	***Vi+Di***	***C+Di***	***C+Vi***
pattern binding			0.02002
carbohydrate binding		0.00061	
iron ion binding		0.00815	
pattern binding		0.01052	
catalytic activity		0.00477	0.04256
endopeptidase activity			1.26E-10
peptidase activity			4.39E-08
peptidase activity, acting on L-amino acid peptides			6.12E-05
peptidase inhibitor activity	0.00356		
peptidase regulator activity	0.00402		
carbon-oxygen lyase activity, acting on polysaccharides			0.00402
transferase activity, transferring hexosyl groups		0.00021	
transferase activity, transferring glycosyl groups		0.00059	
hydrolase activity, acting on glycosyl bonds		0.0038	
galactosidase activity		0.00261	
hydrolase activity, hydrolyzing O-glycosyl compounds		0.01106	
carbohydrate transmembrane transporter activity	0.02814		
hexose phosphate transmembrane transporter activity	3.42E-06		
hexose transmembrane transporter activity	0.00017		
monosaccharide transmembrane transporter activity	0.00044		
organic cation transmembrane transporter activity	0.00766		
sugar transmembrane transporter activity	0.02697		

*V+Di, MtVAMP721e-RNAi+DsRed-RNA*i; *C+Di, CST+DsRed-RNAi*; *C+Vi, CST+MtVAMP721e-RNAi; P+Di, pBin+DsRed-RNAi*.

Table S5. Gene Ontology (GO) Biological Process classification with *P*-value (if $P \leq 0.05$) (after Bonferroni Correction) of differentially expressed genes (DEGs) between treatments (*Vi+Di*, *C+Di* and *C+Vi*) with the control (*P+Di*).

Gene Ontology term (Biological Process)	***Vi+Di***	***C+Di***	***C+Vi***
proteolysis			8.28E-12
proteolysis involved in cellular protein catabolic process			4.81E-12
modification-dependent protein catabolic process			4.82E-11
cellular protein catabolic process			1.45E-11

protein catabolic process			1.63E-11
modification-dependent macromolecule catabolic process			4.82E-11
cellular macromolecule catabolic process			1.32E-10
macromolecule catabolic process			1.98E-10
cellular catabolic process			0.00298
glucan metabolic process		0.00228	
polysaccharide metabolic process		0.00737	
carbon fixation			6.06E-10
photosynthesis			4.94E-05
photosynthesis, light reaction			0.00032
regulation of photosynthesis	0.01514		
regulation of photosynthesis, light reaction	0.00444		
regulation of metabolic process	0.03083		
regulation of generation of precursor metabolites and energy	0.00444		
regulation of molecular function	0.01061		
regulation of catalytic activity	3.94E-05		
regulation of hydrolase activity	0.00082		
regulation of phosphatase activity	0.01739		
regulation of dephosphorylation	0.01739		
negative regulation of catalytic activity	7.07E-06		
negative regulation of hydrolase activity	7.07E-06		
negative regulation of phosphatase activity	0.03655		
negative regulation of molecular function	0.04954		
triose phosphate transport	0.00017		
hexose phosphate transport	1.91E-05		
hexose transport	0.00094		
ammonium transport	0.00094		
monosaccharide transport	0.00157		
organic anion transport	0.04622		
response to stimulus			0.00075
response to carbohydrate stimulus	0.00208		
response to disaccharide stimulus	9.84E-06		
response to hexose stimulus	0.00817		
response to monosaccharide stimulus	0.00817		
recognition of pollen	0.01223		
cell killing	0.00052		

*V+Di, MtVAMP721e-RNAi+DsRed-RNA*i; *C+Di, CST+DsRed-RNAi*; *C+Vi, CST+MtVAMP721e-RNAi; P+Di, pBin+DsRed-RNAi*.

Table S6. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of differentially expressed genes (DEGs) between treatment (*Vi+Di*) with the control (*P+Di*).

Pathway ID	**Pathway enrich in DEGs of *Vi+Di***	**Q-value***
ko04130	SNARE interactions in vesicular transport	0.000284
ko00195	Photosynthesis	0.215656

ko00250	Alanine, aspartate and glutamate metabolism	0.717925
ko00966	Glucosinolate biosynthesis	0.717925
ko00903	Limonene and pinene degradation	0.717925
ko00945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	0.717925
ko00785	Lipoic acid metabolism	0.76418
ko00944	Flavone and flavonol biosynthesis	0.76418
ko00905	Brassinosteroid biosynthesis	0.76418
ko00330	Arginine and proline metabolism	0.76418
ko01100	Metabolic pathways	0.76418
ko00053	Ascorbate and aldarate metabolism	0.76418
ko00941	Flavonoid biosynthesis	0.76418
ko00290	Valine, leucine and isoleucine biosynthesis	0.76418
ko00260	Glycine, serine and threonine metabolism	0.76418
ko01110	Biosynthesis of secondary metabolites	0.76418
ko00750	Vitamin B6 metabolism	0.76418
ko00940	Phenylpropanoid biosynthesis	0.76418
ko03008	Ribosome biogenesis in eukaryotes	0.76418
ko00901	Indole alkaloid biosynthesis	0.76418
ko04712	Circadian rhythm - plant	0.76418
ko00910	Nitrogen metabolism	0.76418
ko00760	Nicotinate and nicotinamide metabolism	0.76418
ko00604	Glycosphingolipid biosynthesis - ganglio series	0.76418
ko00051	Fructose and mannose metabolism	0.76418
ko00460	Cyanoamino acid metabolism	0.76418
ko00340	Histidine metabolism	0.76418
ko00073	Cutin, suberine and wax biosynthesis	0.76418
ko00500	Starch and sucrose metabolism	0.76418
ko00740	Riboflavin metabolism	0.76418
ko03040	Spliceosome	0.76418
ko00960	Tropane, piperidine and pyridine alkaloid biosynthesis	0.76418
ko00562	Inositol phosphate metabolism	0.76418
ko00950	Isoquinoline alkaloid biosynthesis	0.76418
ko04075	Plant hormone signal transduction	0.771995
ko00770	Pantothenate and CoA biosynthesis	0.771995
ko00520	Amino sugar and nucleotide sugar metabolism	0.771995
ko00531	Glycosaminoglycan degradation	0.771995
ko00908	Zeatin biosynthesis	0.771995
ko00943	Isoflavonoid biosynthesis	0.771995
ko04146	Peroxisome	0.771995
ko00196	Photosynthesis - antenna proteins	0.782612
ko00410	beta-Alanine metabolism	0.782612
ko00600	Sphingolipid metabolism	0.823352
ko00240	Pyrimidine metabolism	0.823352
ko03022	Basal transcription factors	0.823352
ko00380	Tryptophan metabolism	0.823352

ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis	0.823352
ko00280	Valine, leucine and isoleucine degradation	0.823352
ko00350	Tyrosine metabolism	0.823352
ko02010	ABC transporters	0.834519
ko03020	RNA polymerase	0.834519
ko00860	Porphyrin and chlorophyll metabolism	0.834519
ko03060	Protein export	0.848752
ko00360	Phenylalanine metabolism	0.86703
ko00592	alpha-Linolenic acid metabolism	0.86703
ko03050	Proteasome	0.86703
ko04145	Phagosome	0.867805
ko00970	Aminoacyl-tRNA biosynthesis	0.867805
ko00480	Glutathione metabolism	0.867805
ko00052	Galactose metabolism	0.867805
ko00511	Other glycan degradation	0.867805
ko00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	0.867805
ko00904	Diterpenoid biosynthesis	0.867805
ko04070	Phosphatidylinositol signaling system	0.867805
ko00230	Purine metabolism	0.867805
ko03420	Nucleotide excision repair	0.867805
ko03015	mRNA surveillance pathway	0.878087
ko03013	RNA transport	0.878087
ko00906	Carotenoid biosynthesis	0.891253
ko04144	Endocytosis	0.944754
ko00040	Pentose and glucuronate interconversions	0.952611
ko04626	Plant-pathogen interaction	0.973664
ko04141	Protein processing in endoplasmic reticulum	0.973664
ko04120	Ubiquitin mediated proteolysis	0.973664
ko03010	Ribosome	0.996692

V+Di, *MtVAMP721e-RNAi+DsRed-RNA*i; *P+Di, pBin+DsRed-RNAi*.
***** pathways with Q-value ≤ 0.05 are significantly enriched in DEGs.

Table S7. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of differentially expressed genes (DEGs) between treatment (*C+Di*) with the control (*P+Di*).

Pathway ID	**Pathway enrich in DEGs of *C+Di***	**Q-value***
ko00531	Glycosaminoglycan degradation	3.73E-06
ko00604	Glycosphingolipid biosynthesis - ganglio series	6.69E-06
ko01110	Biosynthesis of secondary metabolites	6.73E-06
ko00511	Other glycan degradation	2.54E-05
ko00500	Starch and sucrose metabolism	2.73E-04
ko01100	Metabolic pathways	1.04E-03
ko00945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	2.07E-03
ko00906	Carotenoid biosynthesis	3.01E-03
ko00053	Ascorbate and aldarate metabolism	4.38E-03

ko00940	Phenylpropanoid biosynthesis	4.38E-03
ko00903	Limonene and pinene degradation	5.22E-03
ko00600	Sphingolipid metabolism	7.30E-03
ko00052	Galactose metabolism	7.62E-03
ko00520	Amino sugar and nucleotide sugar metabolism	1.77E-02
ko04626	Plant-pathogen interaction	3.16E-02
ko00130	Ubiquinone and other terpenoid-quinone biosynthesis	3.16E-02
ko00040	Pentose and glucuronate interconversions	4.25E-02
ko00944	Flavone and flavonol biosynthesis	4.25E-02
ko04075	Plant hormone signal transduction	4.25E-02
ko00360	Phenylalanine metabolism	4.25E-02
ko00943	Isoflavonoid biosynthesis	4.75E-02
ko00909	Sesquiterpenoid and triterpenoid biosynthesis	4.85E-02
ko00905	Brassinosteroid biosynthesis	4.85E-02
ko00564	Glycerophospholipid metabolism	5.06E-02
ko00480	Glutathione metabolism	7.03E-02
ko00904	Diterpenoid biosynthesis	8.11E-02
ko00908	Zeatin biosynthesis	8.11E-02
ko00941	Flavonoid biosynthesis	8.80E-02
ko04712	Circadian rhythm - plant	1.09E-01
ko00950	Isoquinoline alkaloid biosynthesis	1.36E-01
ko01040	Biosynthesis of unsaturated fatty acids	1.51E-01
ko00592	alpha-Linolenic acid metabolism	2.78E-01
ko00430	Taurine and hypotaurine metabolism	2.78E-01
ko04070	Phosphatidylinositol signaling system	2.86E-01
ko00603	Glycosphingolipid biosynthesis - globo series	2.87E-01
ko00350	Tyrosine metabolism	3.88E-01
ko00561	Glycerolipid metabolism	3.88E-01
ko00966	Glucosinolate biosynthesis	4.55E-01
ko00565	Ether lipid metabolism	4.55E-01
ko00562	Inositol phosphate metabolism	4.55E-01
ko00250	Alanine, aspartate and glutamate metabolism	4.96E-01
ko00591	Linoleic acid metabolism	4.96E-01
ko00460	Cyanoamino acid metabolism	4.97E-01
ko00073	Cutin, suberine and wax biosynthesis	4.97E-01
ko00310	Lysine degradation	5.15E-01
ko04140	Regulation of autophagy	5.18E-01
ko02010	ABC transporters	5.18E-01
ko00780	Biotin metabolism	5.18E-01
ko03030	DNA replication	5.18E-01
ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis	5.18E-01
ko00730	Thiamine metabolism	5.18E-01
ko00960	Tropane, piperidine and pyridine alkaloid biosynthesis	5.18E-01
ko00196	Photosynthesis - antenna proteins	5.18E-01
ko00410	beta-Alanine metabolism	5.18E-01

ko00330	Arginine and proline metabolism	5.18E-01
ko00590	Arachidonic acid metabolism	5.18E-01
ko00062	Fatty acid elongation	6.59E-01
ko00910	Nitrogen metabolism	6.89E-01
ko00770	Pantothenate and CoA biosynthesis	6.89E-01
ko00290	Valine, leucine and isoleucine biosynthesis	7.01E-01
ko00380	Tryptophan metabolism	7.44E-01
ko00510	N-Glycan biosynthesis	7.44E-01
ko00280	Valine, leucine and isoleucine degradation	7.44E-01
ko00710	Carbon fixation in photosynthetic organisms	7.44E-01
ko00071	Fatty acid metabolism	7.44E-01
ko04650	Natural killer cell mediated cytotoxicity	7.44E-01
ko00620	Pyruvate metabolism	7.44E-01
ko00340	Histidine metabolism	7.44E-01
ko00740	Riboflavin metabolism	7.91E-01
ko03430	Mismatch repair	7.91E-01
ko00660	C5-Branched dibasic acid metabolism	7.95E-01
ko00630	Glyoxylate and dicarboxylate metabolism	8.24E-01
ko04145	Phagosome	8.39E-01
ko00061	Fatty acid biosynthesis	8.39E-01
ko00195	Photosynthesis	8.44E-01
ko04144	Endocytosis	8.77E-01
ko00750	Vitamin B6 metabolism	9.14E-01
ko00640	Propanoate metabolism	9.18E-01
ko00100	Steroid biosynthesis	9.29E-01
ko00902	Monoterpenoid biosynthesis	9.33E-01
ko00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	9.35E-01
ko00051	Fructose and mannose metabolism	9.45E-01
ko00760	Nicotinate and nicotinamide metabolism	9.45E-01
ko00402	Benzoxazinoid biosynthesis	9.80E-01
ko00860	Porphyrin and chlorophyll metabolism	9.80E-01
ko00240	Pyrimidine metabolism	9.80E-01
ko00010	Glycolysis / Gluconeogenesis	1.00E+00
ko04710	Circadian rhythm - mammal	1.00E+00
ko00900	Terpenoid backbone biosynthesis	1.00E+00
ko04120	Ubiquitin mediated proteolysis	1.00E+00
ko03020	RNA polymerase	1.00E+00
ko00020	Citrate cycle (TCA cycle)	1.00E+00
ko03420	Nucleotide excision repair	1.00E+00
ko03410	Base excision repair	1.00E+00
ko00920	Sulfur metabolism	1.00E+00
ko00450	Selenocompound metabolism	1.00E+00
ko00650	Butanoate metabolism	1.00E+00
ko00030	Pentose phosphate pathway	1.00E+00
ko04141	Protein processing in endoplasmic reticulum	1.00E+00

ko00270	Cysteine and methionine metabolism	1.00E+00
ko00260	Glycine, serine and threonine metabolism	1.00E+00
ko04146	Peroxisome	1.00E+00
ko03022	Basal transcription factors	1.00E+00
ko00230	Purine metabolism	1.00E+00
ko03440	Homologous recombination	1.00E+00
ko04130	SNARE interactions in vesicular transport	1.00E+00
ko03060	Protein export	1.00E+00
ko03018	RNA degradation	1.00E+00
ko03013	RNA transport	1.00E+00
ko03008	Ribosome biogenesis in eukaryotes	1.00E+00
ko03015	mRNA surveillance pathway	1.00E+00
ko03040	Spliceosome	1.00E+00
ko03010	Ribosome	1.00E+00
ko00190	Oxidative phosphorylation	1.00E+00

C+Di, *CST+DsRed-RNAi*; *P+Di*, *pBin+DsRed-RNAi*.
***** pathways with Q-value ≤ 0.05 are significantly enriched in DEGs.

Table S8. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of differentially expressed genes (DEGs) between treatment (*C+Vi*) with the control (*P+Di*).

Pathway ID	**Pathway enrich in DEGs of *C+Vi***	**Q-value***
ko03050	Proteasome	7.60E-42
ko00195	Photosynthesis	3.82E-10
ko00710	Carbon fixation in photosynthetic organisms	7.91E-07
ko00906	Carotenoid biosynthesis	7.40E-05
ko00511	Other glycan degradation	6.34E-03
ko00630	Glyoxylate and dicarboxylate metabolism	6.98E-03
ko01100	Metabolic pathways	7.61E-03
ko01110	Biosynthesis of secondary metabolites	1.00E-02
ko00592	alpha-Linolenic acid metabolism	1.00E-02
ko00053	Ascorbate and aldarate metabolism	1.02E-02
ko00480	Glutathione metabolism	1.21E-02
ko00460	Cyanoamino acid metabolism	1.45E-02
ko00904	Diterpenoid biosynthesis	1.85E-02
ko03030	DNA replication	2.55E-02
ko00908	Zeatin biosynthesis	3.43E-02
ko04130	SNARE interactions in vesicular transport	3.93E-02
ko00196	Photosynthesis - antenna proteins	4.62E-02
ko00073	Cutin, suberine and wax biosynthesis	4.82E-02
ko00051	Fructose and mannose metabolism	7.46E-02
ko00562	Inositol phosphate metabolism	7.46E-02
ko00500	Starch and sucrose metabolism	7.67E-02
ko04626	Plant-pathogen interaction	9.40E-02
ko00950	Isoquinoline alkaloid biosynthesis	1.37E-01

ko00901	Indole alkaloid biosynthesis	2.06E-01
ko00940	Phenylpropanoid biosynthesis	2.06E-01
ko00360	Phenylalanine metabolism	2.69E-01
ko00944	Flavone and flavonol biosynthesis	3.18E-01
ko00040	Pentose and glucuronate interconversions	3.18E-01
ko00910	Nitrogen metabolism	3.18E-01
ko00520	Amino sugar and nucleotide sugar metabolism	3.18E-01
ko00250	Alanine, aspartate and glutamate metabolism	3.42E-01
ko00943	Isoflavonoid biosynthesis	3.42E-01
ko04146	Peroxisome	3.42E-01
ko00902	Monoterpenoid biosynthesis	3.42E-01
ko00905	Brassinosteroid biosynthesis	3.50E-01
ko00604	Glycosphingolipid biosynthesis - ganglio series	3.50E-01
ko00960	Tropane, piperidine and pyridine alkaloid biosynthesis	3.50E-01
ko04712	Circadian rhythm - plant	3.54E-01
ko00945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	3.54E-01
ko00591	Linoleic acid metabolism	3.80E-01
ko02010	ABC transporters	3.80E-01
ko00903	Limonene and pinene degradation	3.80E-01
ko03430	Mismatch repair	3.85E-01
ko00600	Sphingolipid metabolism	4.18E-01
ko00900	Terpenoid backbone biosynthesis	4.41E-01
ko00750	Vitamin B6 metabolism	4.71E-01
ko00590	Arachidonic acid metabolism	4.91E-01
ko00071	Fatty acid metabolism	4.98E-01
ko00130	Ubiquinone and other terpenoid-quinone biosynthesis	4.98E-01
ko03420	Nucleotide excision repair	5.37E-01
ko00941	Flavonoid biosynthesis	5.44E-01
ko00531	Glycosaminoglycan degradation	5.70E-01
ko00240	Pyrimidine metabolism	5.83E-01
ko00350	Tyrosine metabolism	6.06E-01
ko04075	Plant hormone signal transduction	6.06E-01
ko00430	Taurine and hypotaurine metabolism	6.57E-01
ko00030	Pentose phosphate pathway	7.05E-01
ko00052	Galactose metabolism	7.10E-01
ko00402	Benzoxazinoid biosynthesis	7.10E-01
ko04140	Regulation of autophagy	7.19E-01
ko00730	Thiamine metabolism	7.19E-01
ko00072	Synthesis and degradation of ketone bodies	7.19E-01
ko00330	Arginine and proline metabolism	8.07E-01
ko04070	Phosphatidylinositol signaling system	8.50E-01
ko00310	Lysine degradation	8.56E-01
ko04145	Phagosome	8.63E-01
ko03410	Base excision repair	8.84E-01
ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis	8.86E-01

ko00760	Nicotinate and nicotinamide metabolism	8.86E-01
ko00510	N-Glycan biosynthesis	8.91E-01
ko00740	Riboflavin metabolism	9.00E-01
ko00565	Ether lipid metabolism	9.16E-01
ko03440	Homologous recombination	9.35E-01
ko00785	Lipoic acid metabolism	9.65E-01
ko00670	One carbon pool by folate	9.70E-01
ko00966	Glucosinolate biosynthesis	9.72E-01
ko00640	Propanoate metabolism	9.72E-01
ko03008	Ribosome biogenesis in eukaryotes	9.72E-01
ko00380	Tryptophan metabolism	9.72E-01
ko00561	Glycerolipid metabolism	9.72E-01
ko00061	Fatty acid biosynthesis	9.82E-01
ko00260	Glycine, serine and threonine metabolism	9.89E-01
ko00340	Histidine metabolism	9.97E-01
ko00410	beta-Alanine metabolism	1.00E+00
ko00564	Glycerophospholipid metabolism	1.00E+00
ko00010	Glycolysis / Gluconeogenesis	1.00E+00
ko00450	Selenocompound metabolism	1.00E+00
ko00100	Steroid biosynthesis	1.00E+00
ko00280	Valine, leucine and isoleucine degradation	1.00E+00
ko00660	C5-Branched dibasic acid metabolism	1.00E+00
ko04141	Protein processing in endoplasmic reticulum	1.00E+00
ko00909	Sesquiterpenoid and triterpenoid biosynthesis	1.00E+00
ko00650	Butanoate metabolism	1.00E+00
ko00270	Cysteine and methionine metabolism	1.00E+00
ko03022	Basal transcription factors	1.00E+00
ko00860	Porphyrin and chlorophyll metabolism	1.00E+00
ko03450	Non-homologous end-joining	1.00E+00
ko00514	Other types of O-glycan biosynthesis	1.00E+00
ko04120	Ubiquitin mediated proteolysis	1.00E+00
ko00290	Valine, leucine and isoleucine biosynthesis	1.00E+00
ko00062	Fatty acid elongation	1.00E+00
ko00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1.00E+00
ko00620	Pyruvate metabolism	1.00E+00
ko04650	Natural killer cell mediated cytotoxicity	1.00E+00
ko00770	Pantothenate and CoA biosynthesis	1.00E+00
ko04710	Circadian rhythm - mammal	1.00E+00
ko00230	Purine metabolism	1.00E+00
ko01040	Biosynthesis of unsaturated fatty acids	1.00E+00
ko03060	Protein export	1.00E+00
ko03013	RNA transport	1.00E+00
ko03020	RNA polymerase	1.00E+00
ko00920	Sulfur metabolism	1.00E+00
ko03015	mRNA surveillance pathway	1.00E+00

ko00970	Aminoacyl-tRNA biosynthesis	1.00E+00
ko04144	Endocytosis	1.00E+00
ko00020	Citrate cycle (TCA cycle)	1.00E+00
ko03040	Spliceosome	1.00E+00
ko03018	RNA degradation	1.00E+00
ko00190	Oxidative phosphorylation	1.00E+00
ko03010	Ribosome	1.00E+00

C+Vi, CST+MtVAMP721e-RNAi; P+Di, pBin+DsRed-RNAi.
***** pathways with Q-value ≤ 0.05 are significantly enriched in DEGs.

Table S9. Wilcoxon rank sum test of differentially expressed genes (DEGs) between treatments (*Vi+Di*, *C+Di* and *C+Vi*) with the control (*P+Di*) using MapMan tool (*P*≤0.05).

Bin	**Name (description)**	**V+Di**		**C+Di**		**C+Vi**	
		Up	**Down**	**Up**	**Down**	**Up**	**Down**
1	PS (Photosynthesis)					5	68
1.1	PS.lightreaction					1	47
1.1.1	PS.lightreaction.photosystem II					1	23
1.1.1.2	PS.lightreaction.photosystem II.PSII polypeptide subunits	0	3	0	2	0	12
1.1.2	PS.lightreaction.photosystem I					0	9
1.1.2.2	PS.lightreaction.photosystem I.PSI polypeptide subunits					0	9
1.1.5	PS.lightreaction.other electron carrier (ox/red)					0	4
1.2	PS.photorespiration					1	7
1.3	PS.calvin cycle					3	14
1.3.2	PS.calvin cycle.rubisco small subunit					0	2
1.3.4	PS.calvin cycle.GAP					0	3
2.1	major CHO metabolism.synthesis			0	5	2	8
2.1.2	major CHO metabolism.synthesis.starch			0	5	1	8
10	cell wall			12	41	18	54
10.2	cell wall.cellulose synthesis					0	9
10.5	cell wall.cell wall proteins			0	8		
10.5.1	cell wall.cell wall proteins.AGPs			0	7		
10.5.1.1	cell wall.cell wall proteins.AGPs.AGP			0	7		
10.6	cell wall.degradation					4	18
10.7	cell wall.modification			1	11		
10.8	cell wall.pectin*esterases					1	8
10.8.1	cell wall.pectin*esterases.PME					1	7
11.1	lipid metabolism.FA synthesis and FA elongation					3	11
11.1.8	lipid metabolism.FA synthesis and FA elongation.acyl coa ligase			3	0		
11.9.2	lipid metabolism.lipid degradation.lipases			6	1		
11.9.2.1	lipid metabolism.lipid degradation.lipases.triacylglycerol lipase			5	1		
13.1.3	amino acid metabolism.synthesis.aspartate family			2	0	2	0
16	secondary metabolism					7	31
16.2	secondary metabolism.phenylpropanoids					1	8
17.2	hormone metabolism.auxin	3	0	9	3		
17.2.3	hormone metabolism.auxin.induced-regulated-responsive-activated	3	0	9	3		
17.5.1	hormone metabolism.ethylene.synthesis-degradation					13	4

17.6	hormone metabolism.gibberelin			0	4		
17.6.3	hormone metabolism.gibberelin.induced-regulated-responsive-activated			0	4		
20	stress	9	4	29	14		
20.1	stress.biotic			18	7		
20.2.1	stress.abiotic.heat	4	0	8	0		
23	nucleotide metabolism			7	2		
23.3	nucleotide metabolism.salvage			3	0		
23.3.3	nucleotide metabolism.salvage.NUDIX hydrolases			3	0		
26.21	misc.protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	3	0			0	6
26.28	misc.GDSL-motif lipase					3	13
26.9	misc.glutathione S transferases			6	0	9	3
27.3.3	RNA.regulation of transcription.AP2/EREBP, APETALA2/Ethylene-responsive element binding protein family					16	6
27.3.7	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family			3	0	6	1
27.3.24	RNA.regulation of transcription.MADS box transcription factor family					3	1
27.3.32	RNA.regulation of transcription.WRKY domain transcription factor family			7	0		
27.3.37	RNA.regulation of transcription.AS2,Lateral Organ Boundaries Gene Family			3	0		
28	DNA			4	17		
29	protein					165	158
29.4	protein.postranslational modification					40	50
29.5	protein.degradation					99	72
29.5.1	protein.degradation.subtilases					1	10
29.5.4	protein.degradation.aspartate protease			3	0	6	2
29.5.5	protein.degradation.serine protease					1	8
29.5.7	protein.degradation.metalloprotease					2	0
29.5.11	protein.degradation.ubiquitin					74	40
29.5.11.20	protein.degradation.ubiquitin.proteasome					36	0
30	signalling					47	101
30.2	signalling.receptor kinases			8	19	6	63
30.2.11	signalling.receptor kinases.leucine rich repeat XI			2	9	2	35
30.2.17	signalling.receptor kinases.DUF 26					1	19
30.3	signalling.calcium			19	5		
30.4	signalling.phosphinositides					5	0
30.6	signalling.MAP kinases					0	2
31	cell	2	6				
31.4	cell.vesicle transport	1	4				
33	development			20	12		
34	transport	6	18	22	60		
34.10	transport.nucleotides					1	4
34.19.2	transport.Major Intrinsic Proteins.TIP					0	3
34.13	transport.peptides and oligopeptides			1	7		
34.98	transport.membrane system unknown			2	0		
34.99	transport.misc			1	13		

C+Di, *CST+DsRed-RNAi*; *C+Vi*, *CST+MtVAMP721e-RNAi*; *V+Di*, *MtVAMP721e-RNAi+DsRed-RNA*i.

Table S10. Differential gene analysis of *NbVAMP72s* [*MtVAMP721e-RNAi+DsRed-RNAi (Vi+Di)* vs. *pBin+DsRed-RNAi (P+Di)*].

Gene ID	Gene name	Log_2FoldChange (Vi+Di / P+Di)
NbS00001259g0006.1	NbVAMP72a	-2.6
NbS00022342g0004.1	NbVAMP72b	-2.9
NbS00031386g0004.1	NbVAMP72c	-2.6
NbS00006341g0014.1	NbVAMP72d	-2.8
NbS00039776g0004.1	NbVAMP72e	-2.0
NbS00060438g0004.1	NbVAMP72f	-0.1
NbS00033654g0004.1	NbVAMP72g	-0.7
NbS00012174g0011.1	NbVAMP72h	-3.0
NbS00004540g0001.1	NbVAMP72i	*
NbS00053389g0007.1	NbVAMP72j	*
NbS00024620g0006.1	NbVAMP72k	-3.1
NbS00036546g0008.1	NbVAMP72l	*
NbS00004447g0302.1	NbVAMP72m	-2.5
NbS00060093g0001.1	NbVAMP72n	-2.4

***** No identify in the differential gene analysis.

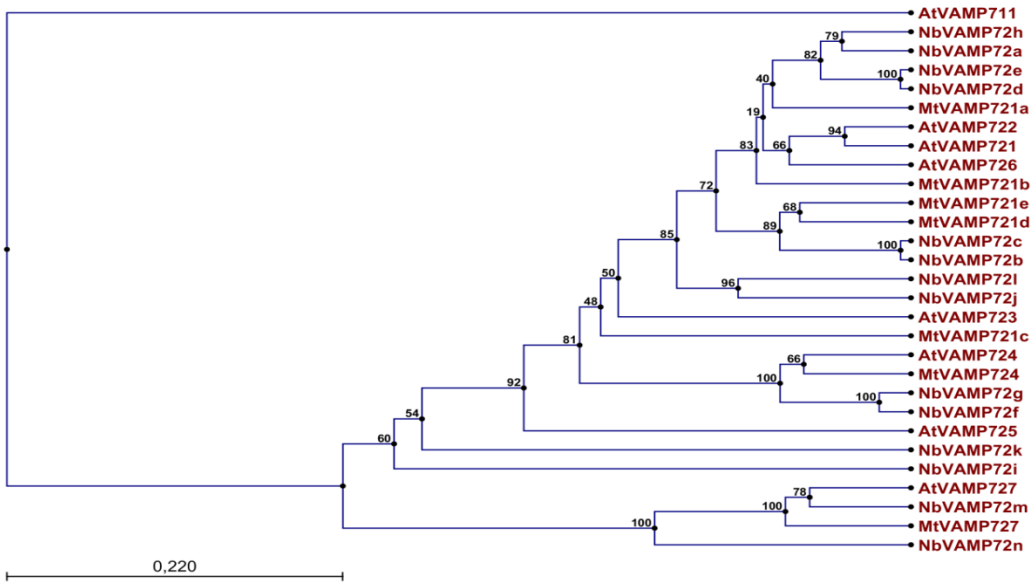

Figure S1. Phylogenetic analysis of MedVAMP72, AtVAMP72 and NbVAMP72.
Phylogenetic tree was generated using the alignment and the UPGMA method. One hundred bootstrap replicates were used to assess the significance of the tree topology. The analysis involved 29 amino acid sequences (7 from Medicago, 8 from Arabidopsis and 14 from *N. benthamiana*). The *scale bar* indicates amino acid substitutions per site, and *numbers* indicate the branch support values in percentage.

(a) > **NbS00031386g0004.1** Niben044Scf00031386:7464..11878 Vesicle associated membrane protein 7B [Solanum lycopersicum]

Length = 1,364
Score = 400 bits (202), Expect = 1e-110
Identities = 484/578 (83%), Frame = +1 / +1

```
MtVAMP721e: 62   ctgaatacactgaattcagcggaaacttcaactccatcgcttttcaatgccttcagaagc 121
                 ||||||||||  ||  |||||  ||  ||||||||||||||  ||||  |||  ||||||||||||||
NbVAMP72c : 243  ctgaatacacagatttcagtggtaacttcaactccatagcttatcagtgccttcagaagc 302

MtVAMP721e: 122  ttccagcttctaacaacaaattcacttacaattgtgatggccacaccttcaactacctcg 181
                 | ||  |||||  ||  |||||  ||||||||||  ||||||||  ||||||||||  |||||  |
NbVAMP72c : 303  tccctgcttccaataacaagttcacttacaactgtgatggtcacaccttcaattaccttg 362

MtVAMP721e: 182  tcgataatggctacacgtattgtgttgttgcggatgaaactgttggaagacaagtgccgg 241
                 ||||||||||||  |||  ||  |||||  |||||  ||  ||   |  ||||||||||||||  ||
NbVAMP72c : 363  tcgataatggcttcacatactgtgtggttgctgaggagtccgttggaagacaagttccaa 422

MtVAMP721e: 242  tggcttttctggagcgtgtgaaggatgattttgttgcgaaatatggtggtgggaaggcgt 301
                 | ||  |||  ||||||||||  ||||||||||||||   ||||||||||  ||||||||||
NbVAMP72c : 423  tagcatttttggagcgtgttaaggatgattttgtgtcgaaatatgggggtgggaaggctg 482

MtVAMP721e: 302  ctacagctgcacctaacagcttgaataaggaatttgggccaaaattgaaggaacatatgc 361
                 | ||||||  | ||  ||||||||||  ||||||||||  ||  ||  ||||||||  |||||||
NbVAMP72c : 483  ccacagctcctccaaacagcttgaacaaggaatttggacctaagttgaaggagcatatgc 542

MtVAMP721e: 362  agtattgtgttgatcacccggaggaggtgagtaagcttgctaaggtgaaagctcaggtgt 421
                 ||||||||  |||  ||  ||  ||  ||  |  ||  ||  ||||  ||||||||||  |||||  |
NbVAMP72c : 543  agtattgtgctgaccatcctgatgaaattagcaaaattgcgaaggtgaaagcccaggttt 602

MtVAMP721e: 422  ctgaagtgaaaggtgtcatgatggagaacattgaaaaggtcctggacaggggggaaaaga 481
                 | |||||  ||||||||  ||||||||  ||||||||  |||||  ||  |||  |  ||  |||||||
NbVAMP72c : 603  cggaagttaaaggtgttatgatggaaaacattgagaaggttcttgaccgtggagaaaaga 662
                                   F ---->
MtVAMP721e: 482  tagagcttttggttgacaagactgagaaccttcatcaccaggcacaagacttcaggaact 541
                 ||||||||  |  ||  ||  |||||  |||||||||||||||||||||||||||||||||  |
NbVAMP72c : 663  tagagcttcttgtggataagacagagaaccttcatcaccaggcacaagacttcaggtata 722

MtVAMP721e: 542  cagggaccaaaatccgcagaaaaatgtggctgcaaaatatgaagataaagctgattgtat 601
                 | ||  |||  |||||||  ||  |  ||||||||||||||  ||||||||  ||||||||  ||  |
NbVAMP72c : 723  ctggaacccaaatccggaggagaatgtggctgcaaaacatgaagatcaagctgatagtgt 782
                 <---- R
MtVAMP721e: 602  tggcaattttgatagctctgatcctcataattgtcctc 639
                 |||  ||  ||||  |||  ||||||||||||||||||||
NbVAMP72c : 783  tgggtatcctgattgctttgatcctcataattgtcctc 820
```

(b)

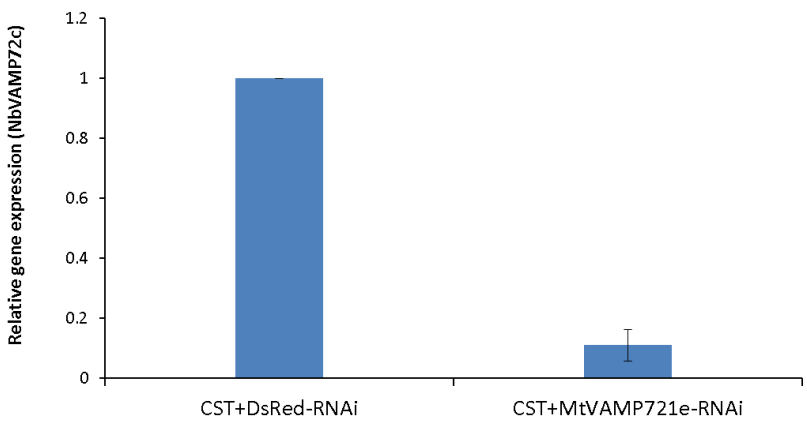

Figure S2. *MtVAMP721e-RNAi* can target *VAMP72* in *N. benthamiana*.
(a) Sequence alignment of *MtVAMP721e* and *NbS00031386g0004.1* (*NbVAMP72c*). Sequence of *MtVAMP721e* was blast in *N. benthamiana* data set (http://solgenomics.net/). Red characters represent the primers used for qRT-PCR.
(b) Relative gene expression of *NbVAMP72c* in leaves agroinfiltration with *CST+DsRed-RNAi* and *CST+MtVAMP721e-RNAi*. Data are reported as mean ± standard error (three different biological experiments).

Chapter **6**

General discussion

Terpenoids are the largest class of natural product that are produced by plants, with functions that range from a role in plant development to direct defence against pathogens and indirect defence against insects through the attraction of natural enemies (Aharoni *et al.*, 2005). The biosynthetic complexity of terpenoids varies from simple to very elaborate, and often involves enzymatic steps occurring in different subcellular locations and sequestration outside of the cell. Moreover, the production of many potentially toxic terpenoids takes place in special cells (e.g. glandular trichomes, laticifers) (Heinig *et al.*, 2013; Kutchan, 2005). Although it has been known for a long time that these complex terpenoid biosynthetic pathways must require several transport steps, the role of transport is almost completely unstudied. In this thesis, I addressed several aspects of transport in the biosynthesis of a number of terpenoids: what determines the choice between two branches of the artemisinin pathway (**Chapter 2**)? Is there a role for LTPs in terpene transport (**Chapter 3**)? Is vesicle transport involved in terpene production and emission (**Chapter 5**)? Because we encountered a problem in some of the crosses with our T-DNA insertion mutants and this was initially seen as a phenotype related to a function of LTPs and terpenes in gametogenesis, these phenotypes were intensively characterized. After that work was finished it turned out that the pollen and ovule phenotype was related to chromosomal translocations in some of the T-DNA lines and the phenotypes had nothing to do with the LTPs or terpenes. However, since there is still limited information on the phenotypes of different chromosomal translocation events, the results of the aberrant pollen and ovule study are reported in **Chapter 4**.

Subcellular compartmentalization in terpene biosynthesis

Sesquiterpene biosynthesis starts in the cytosol and primary sesquiterpene products may be further modified by cytochrome P450s, reductases and dehydrogenases. For the artemisinin biosynthesis pathway in *Artemisia annua* the first dedicated enzyme of the pathway, amorphadiene synthase (ADS), is located in the cytosol, while amorphadiene oxidase (AMO/CYP71AV1, a cytochrome P450 enzyme) is located on the cytoplasmic face of the ER membrane and the artemisinic aldehyde reductase (DBR2) and aldehyde dehydrogenase 1 (ALDH1) are cytosolic enzymes again (Kim *et al.*, 2008, Teoh *et al.*, 2009, Ting *et al.*, 2013, Zhang *et al.*, 2008). In the research in **Chapter 2**, I showed that there are slightly different forms of AMO/CYP71AV1 (AMOHAP and AMOLAP) that differ in their affinity for AAA. Because the AMOHAP has reduced affinity for AAA, this intermediate is more available for

DBR2, which catalyses a branch-point in the biosynthesis pathway. Therefore, in *A. annua* the AMO type (HAP or LAP) most likely contributes to the difference between high artemisinin producing (HAP) and low artemisinin producing (LAP) chemotypes by determining the flux through the different branches of the pathway. It would be interesting to see whether the flux towards the DBR2 branch of the pathway can be further enhanced, for instance by bringing DBR2 closer to the ER bound AMO through engineering of an ER-transmembrane domain on DBR2.

Altering the location of the enzyme can have big effects on the activity as the subcellular localisation of substrates may greatly differ. For instance, it has been shown that overexpression of the cytosolic patchoulol synthase (PTS, sesquiterpene synthase) and farnesyl diphospate synthase (FPS) when targeted to the plastids results in 100-fold higher patchoulol production compared with cytosolic located PTS and FPS (Wu *et al.*, 2006). In the context of the artemisinin pathway it is still of interest to investigate how the product of the ADS (amorphadiene) reaches the AMO P450 enzyme located on the ER and how the product of the AMO on the cytosolic face of the ER reaches the soluble DBR2. Also, these steps may be suitable for further engineering, for instance by attaching an ER transmembrane domain to ADS (and DBR2 as suggested above) to anchor them to the ER membrane. Future research will have to show whether a combined localisation of ADS and DBR2 to the ER more efficiently directs the flux towards the dihydro-branch of the pathway, that is towards dihydroartemisinic aldehyde (DHAAA) and dihydroartemisinic acid (DHAA).

Putative terpene transport pathways

Terpene intermediates and end products are often apolar and therefore may preferentially sequester in membranes or lipid vesicles. Volatile terpenes are emitted by plants, both from root, leaf and floral tissues (reviewed in (Dudareva *et al.*, 2004)) indicating that there must be one or more pathways by which these molecules reach the apoplast in plants. Different transport pathways have been described for (plant) cells for the transport of proteins to their different destinations, involving vesicles derived from the ER, Golgi and trans-Golgi network (TGN) and sorting of these vesicles based on cargo and SNAREs which are involved in directing fusion to target membranes. It could be that terpenes simply hitchhike along with this vesicle transport. At the plasma membrane (PM) the lipophilic terpenes may then either just diffuse over the membrane (passive, none-facilitated pathway) or specific transporters

may be involved in their transport over the membrane (facilitated transport pathway). It has been shown that a small, hydrophobic or nonpolar molecule can diffuse across the bilayer of the plasma membrane (Stein, 1986). As an alternative to transport by vesicles, terpenes may be transported through the cytosol by carrier proteins, e.g. soluble proteins with a hydrophobic pocket that can accommodate the apolar terpenes. LTPs are good candidate for such function and have been shown to be very pleiotropic in the type of molecule that they can bind (reviewed in (Lev, 2010)). Indeed some LTP proteins are not secreted to the extracellular space and operate inside the cell, while the function of LTPs may even be dual (both outside and inside the cell). Recently it was shown that HaAP10, an LTP from sunflower (*Helianthus annuus*) is localized in the apoplast in dry seeds, but relocalises to the cytosol upon germination (Pagnussat *et al.*, 2012). An overview of all the putative pathways by which hydrophobic terpenes may reach the extracellular space is shown in Figure 1. Within the putative trafficking routes, we can distinguish between transport through the cytosol and transport over the plasma membrane.

Pathway 1: Passive diffusion pathway: small, hydrophobic or nonpolar molecule can diffuse across the bilayer of the plasma membrane (Stein, 1986), while several recent studies also postulated that lipophilic terpenes may diffuse passively over the membrane to the apoplastic space (Caissard *et al.*, 2004, Effmert *et al.*, 2005, Niinemets *et al.*, 2004). Diffusion of lipophilic terpenes through the cytosol is not very efficient, but this part of the terpene transport may either be facilitated by endogenous carrier proteins (Pathway 6), by vesicle transport (Pathways 3, 4) or by ER-PM contact sites (Pathway 5).

Pathway 2: Besides passive diffusion over the plasma membrane, transport of terpenes over the plasma membrane may also be facilitated by membrane transport proteins. Indeed, for diterpenes, transporter proteins have been described and in different plants species these are ABC type membrane transporters (Campbell *et al.*, 2003, Crouzet *et al.*, 2013, Jasiński *et al.*, 2001, Van Den Brûle *et al.*, 2002). Moreover, in the follow up of this thesis work, transporters for artemisinin related compounds were identified (Boutry, Wang, vdKrol unpublished data). It is noteworthy that in order to reach the plasma membrane, the transporter proteins follow pathway 3, while also the LTPs in the apoplast reach this position through pathway 3 (see below).

The classical protein secretory pathway in the plant cell starts in the ER, where secreted proteins with N-terminal signal peptide are exported to the Golgi and subsequently, by sorting and trafficking of vesicles through the TGN reach different subcellular destinations, which can be the plasma membrane (Pathway 3) or multivesicular bodies (Pathway 4).

Pathway 3: Secreted proteins in vesicles that go to the plasma membrane are inserted into the membrane or are secreted into the apoplast after SNARE-mediated fusion of the vesicle membrane with the plasma membrane (reviewed in (Jürgens, 2004)). Terpenes are often modified by ER-localized cytochrome P450s, which significantly contributes to the structural diversity of terpenes (Weitzel&Simonsen, 2013). Terpenes may be transported from the ER by sequestering into vesicles budding from the ER, followed by transport through the Golgi and trans-Golgi network to the plasma membrane. Thus, terpenes could hitchhike along this vesicle secretion pathway and reach the plasma membrane, where they then still may be transported over the membrane by ABC transporters for which the entry site is located in the membrane.

Pathway 4: There are two identified and characterized endosomal compartments: the first one is the TGN or TGN-derived compartments that act as early endosomes. The TGN acts as a sorting station that classifies cargo from Golgi to either PM, cell wall, cell plate or vacuole (reviewed in (Reyes *et al.*, 2011)). The other endosomal compartments are the late endosomes or multivesicular bodies (MVBs) and/or paramural bodies (PMB). Multivesicular bodies (MVBs) are spherical endosomal organelles containing small vesicles formed by inward budding into the endosomal lumen, while PMBs are membranous structure located between the plasma membrane and the plant cell wall (Marchant&Robards, 1968). Morphological studies of the glandular trichomes of *Calceolaria adscendens* Lidl. (Scrophulariaceae) and *Pogostemon cablin* (Lamiaceace) by electron microscopy (SEM and TEM) identified abundant vesicles between the cell wall and plasma membrane (Guo *et al.*, 2013, Sacchetti *et al.*, 1999). Also in mammalians, MVBs have been shown to fuse with the plasma membrane in an exocytic manner, leading to a release of their contents including internal vesicles into the extracellular space. When terpenes would be included into MVBs and/or PMBs, the pathway involving these MVBs/PMBs could potentially deliver terpenes to the extracellular space.

Pathway 5: Several studies haves shown that there may be specific ER-PM contact sites which could be important for the transfer of terpenes. For yeast recently proteins were characterized that facilitate this ER-PM contact (Manford *et al.*, 2012). It has been shown that plastids in *Pogostemon cablin* (Lamiaceace) are very closely associated with the smooth-ER and plasma membrane (Guo *et al.*, 2013) which may be of importance for the emission of monoterpenes and diterpenes, which are produced in the plastids.

Pathway 6: Carrier mediated transport could be important for terpene transport in the cytosol. LTP proteins have also been located in the cytosol (Pagnussat *et al.*, 2012) and

terpene transport through the cytosol could be facilitated by LTPs or LTP-like proteins. LTP expression has been shown to be limiting for secretion of the diterpene sclareol (Choi *et al.*, 2012), indicating that they play a role in terpene accumulation. Moreover, the LTPs of *A. annua* are localised to the apoplast (**Chapter 3**), which could signify that there is also a carrier function for LTPs in the apoplast. Note that the presence of LTPs in the apoplast depends on the protein secretion pathway (Pathway 3), which is in turn is dependent on the function of SNARE proteins that mediate the fusion of vesicles of the protein secretory pathway with the plasma membrane.

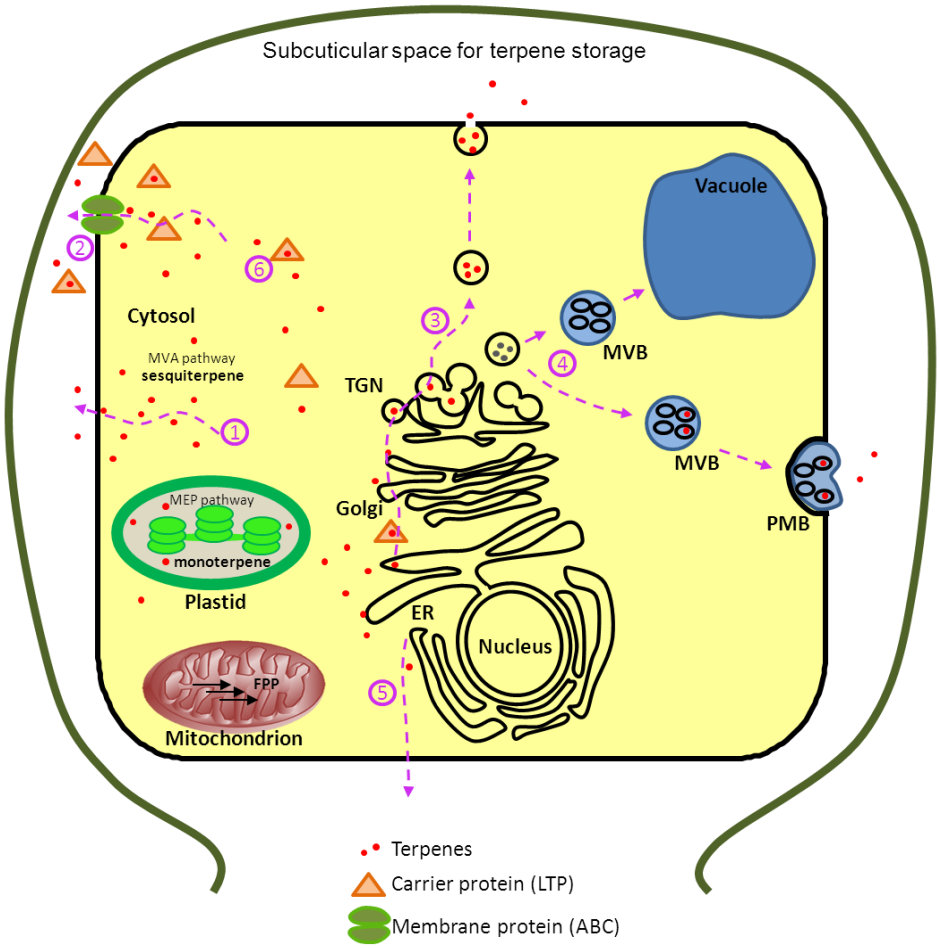

Figure 1. Different putative subcellular terpene transport pathways.
The scheme shows different putative pathways of terpenes transport to the extracellular space, based on literature of transport pathways in plants. (1) non-facilitated diffusion from cytosol to apoplast, possibly by vesicles (2) transfer through cytosol to membrane transporter (ABC proteins?) facilitating transport to apoplast; (3) transport along the vesicle sorting pathway from ER, Golgi and trans-Golgi network (TGN) to plasma membrane; (4) transport along the vesicle sorting pathway from ER, Golgi and TGN to multivesicular body (MVB) and paramural body (PMB). (5) transport from ER membrane to plasma membrane followed by non-facilitated or facilitated transport over the membrane. (6) transfer by the carrier protein (LTP) in cytosol or in apoplast.

Are LTPs involved in terpene production?

In contrast to terpene biosynthesis, very little is known about terpene transport within the cell and between cells. I investigated a putative role of LTPs in terpene transport based on the fact that *LTP* genes are often highly expressed in cells where terpene production is high. The functional assay for LTPs consists of a lipid transport/binding assay and these have shown that the hydrophobic pocket of the LTP proteins can harbour different types of hydrophilic lipids, but very little research has been done on binding studies of LTPs with terpenes. Because the crystal structure of some LTP proteins has been determined, *in silico* docking experiments with different terpene structures could possibly also show whether the LTPs isolated from *A. annua* can accommodate artemisinin intermediates and this would be an interesting future approach to take. When present inside the cell, LTPs may act as carrier protein for the translocation of terpenes to a membrane transporter. When present in the apoplast, they may also function as carrier protein. For instance, in Arabidopsis, LTPs have been characterized involved in the exine secretion in the tapetum of Arabidopsis flowers during pollen formation. It was shown that LTPs (either bound or not bound to exine precursors) are transported within the secretory vesicles derived from ER-TGN to the plasma membrane where they are released into the extracellular space (Huang *et al.*, 2013). A similar transport route could be envisioned for terpenes, being transported bound to LTPs through the cytosol and subsequently secreted into the apoplast.

LTPs belong to a multigene family. For instance, a total of 71 putative *LTPs* have been identified in the Arabidopsis genome (Beisson *et al.*, 2003). The high number of *LTP* genes suggests functional redundancy. This, combined with their promiscuous binding properties, may result in the absence of a clear phenotype in single LTP mutants. To avoid the functional redundancy and to prevent problems associated with T-DNA insertion mutants (**Chapter 4**), artificial microRNA (amiRNA) or RNA interference (RNAi) approaches could be used (Chuang&Meyerowitz, 2000, Miki *et al.*, 2005, Ossowski *et al.*, 2008). It would likely be best to drive such constructs by an inducible promoter to avoid problems associated with the regeneration of transformants with silenced *LTP* gene expression, which for example failed when we transformed *A. annua* with *LTP* RNAi constructs (**Chapter 3**). In this way, the expression of multiple *LTP* genes could be targeted for suppression at the same time and the effect on terpene emission (e.g. caryophyllene emission in Arabidopsis flowers) could be measured.

Recently, NtLTP1 which is expressed in glandular trichomes of tobacco, was shown to enhance diterpene secretion upon overexpression (Choi *et al.*, 2012), indicating that in tobacco *LTP* expression can be limiting for secretion capacity. However, in our tests with overexpression of the *AaLTPs* with the artemisinin pathway genes in *N. benthamiana* no enhanced transport to the apoplast of pathway products was detected (**Chapter 3**). In the end, this leaves us with very limited evidence that LTPs play a role in terpene transport, although both the results from the reduction of caryophyllene in Arabidopsis *ltp1* and *ltp3* mutants and the effect of overexpression of tobacco LTP on sclareol secretion do show significant effects on terpenes. The question remains how this effect relates to how LTPs actually function in the apoplast. Only recently, in the research following this thesis work this may have become more clear (see under Future Perspective below). However, there are still many types of experiments that can be done to further investigate the role of LTPs in terpene transport/accumulation (see Table 1).

Table 1. Suggestions for new experiments.

Complementation of Arabidopsis *ltp1* and *ltp3* mutants with *AaLTPs*
LTP gene silencing in *A. annua* with inducible-promoter
LTP gene silencing in *A. annua* with trichome-specific promoter
Test combinations of LTPs + PDRs (ABC transporters)
Use Virus-induced gene silencing in *A. annua* to transiently knock down *LTPs*
Test the overexpression of LTPs in combination with PDRs in *A. annua*

Is vesicle transport involved in terpene emission?

Next to the research on a role for LTPs in terpene transport, I also investigated the role for vesicle transport in terpene emission. The fusion of vesicle membranes to the plasma membrane is mediated by SNARE complexes (Table 2), which consist of proteins present in the vesicle membrane (v-SNAREs) and proteins present in the target membrane (t-SNAREs) (Lang&Jahn, 2008) and vesicle fusion may thus be manipulated by modifying these SNAREs. In this thesis, I used the transient expression of an Arabidopsis caryophyllene synthase gene (*CST*, *AtTPS21*/At5g23960) with or without inhibition of vesicle fusion by silencing the expression of the plant specific SNARE, *VAMP72* and found that instead of a decrease in

caryophyllene emission, this inhibition results in higher caryophyllene emission, most likely due to increased protein stability. These effects on protein stability will be discussed below. First I will discuss how the results may relate to vesicle fusion being involved in terpene emission. In *Medicago truncatula*, it has been shown that two highly homologous SNARE proteins, vesicle-associated membrane proteins of type VAMP72, are associated with exocytotic vesicles (Ivanov *et al.*, 2012). These Medicago VAMP72s were shown to be involved in the formation of the membrane interface with symbiotic *Rhizobium* bacteria. While silencing of these *MtVAMP72* genes with a *MtVAMP72-RNAi* construct had a minor effect on non-symbiotic plant development and nodule formation, it did block rhizobium induced symbiosis and AMF induced arbuscule formation, without affecting the root colonization by these micro-organisms (Ivanov *et al.*, 2012). I used the same *MtVAMP72-RNAi* construct to inhibit *VAMP* mediated vesicle trafficking in *N. benthamiana* in combination with overexpression of a caryophyllene biosynthesis gene (*CST*, *AtTPS21*/At5g23960). Transcriptional analysis confirmed that the effect of the *MtVAMP72-RNAi* construct was specific for *VAMP72* type genes in *N. benthamiana*. Because other type *VAMP* genes were not suppressed in expression by the *MtVAMP72-RNAi*, most likely only selected vesicles from Pathway 3 (Fig.1) were inhibited for membrane fusion. I also tested the effect of an *MtVAMP711-RNAi* construct on caryophyllene emission. The *VAMP71-RNAi* did affect caryophyllene emission, but the effect was smaller than with the *VAMP72-RNAi*: co-expression of the *MtVAMP711-RNAi* construct with *CST* resulted in about 3-fold increase in caryophyllene emission (unpublished data), while *MtVAMP72-RNAi* resulted in about 5-fold increase in caryophyllene emission compared with CST overexpression alone (**Chapter 5**). In Arabidopsis the different VAMP proteins locate to different subcellular locations (summarized in Table 2). For instance, the Arabidopsis VAMP72 proteins are located in PM, TGN, and early endosomes while the VAMP71 class proteins are located in late endosomes. Assuming similar location and function of the VAMP72 and VAMP71 class proteins in *N. benthemiana*, the *VAMP72-RNAi* is expected to target trafficking to the plasma membrane and/or vacuole (Pathway 3 and 4 in Fig. 1), while the *VAMP71-RNAi* is expected to target endocytotic vesicle trafficking to the vacuole (Pathway 4 in Fig. 1) (Carter *et al.*, 2004). The overlap in traffic pathway between VAMP72 type and VAMP71 type SNAREs is in pathway 3, so maybe that pathway is most relevant to the observed effect on caryophyllene emissions.

Table 2. SNARE proteins on different plant membrane compartments.

Membrane compartment	t-SNARE	v-SNARE
ER/Golgi	SYP81, SEC20, USE1	SEC22
cis-Golgi	SYP31/32, MEMB1, BET1	SEC22
trans-Golgi	SYP31/32, GOS1, SFT1	SEC22
TGN / EE	SYP41/42, SYP61, VTI12	YKT61, VAMP727
LE / Vacuole	SYP21/22, SYP51/52, VTI11	VAMP711, VAMP727
PM	SYP121/122, SYP131/132, SNAP33	VAMP721/722, VAMP724
Cell Plate	SYP111/112, SYP71, NPSN1	VAMP721/722

TGN: trans-Golgi network; EE: early endosome; LE: late endosome (also described as multivesicular body, or prevacuolar compartments); PM: Plasma membrane. [Modified from (Ivanov, 2012, Kim&Brandizzi, 2012)]

However, I did not test the specificity of the silencing by the Medicago *VAMP711-RNAi* construct, so it could be that this also targets *VAMP72* genes in *N. benthemiana*, but with reduced efficiency. Because the inhibition of SNARE function in combination with caryophyllene synthesis resulted in unexpected protein stability effects (**Chapter 5**), this confuses the issue of a role in terpene transport. The same *VAMP72-RNAi* construct were also used in combination with overexpression of a linalool synthase. The *VAMP72-RNAi* resulted in an increase in terpene emission, similar to what happened in the combination of *CST+VAMP72-RNAi*. It would be of interest to assess the effect of this *VAMP72-RNAi* with other terpene synthases, for instance with geraniol synthase (monoterpene) (Dong *et al.*, 2013) or the costunolide pathway (sesquiterpene) from feverfew (Liu *et al.*, 2011) to test the specificity of the terpene on the increased protein stability phenotype. Moreover, it still needs to be verified whether the transport of membrane transporters (e.g. PDRs) and LTPs to the apoplast are affected by the Medicago *VAMP72-RNAi* in the transient assays in *N. benthamiana* leaves.

As alternative to *SNARE-RNAi* in the study of a role for vesicle fusion in terpene transport, chemical transport inhibitors could be used. There are a number of chemical inhibitors that target single or multiple endosomal transport pathways in the cell, such as brefeldin A (BFA) which inhibits vesicle transport from TGN to PM and from TGN to EE/LE (reviewed in (Hachez *et al.*, 2013)). Recently, several other small molecules (Endosidin 1, Endosidin 5, Sortin 1, Sortin 2 and LG8) have been reported that can affect endomembrane trafficking at different positions in the plant cell (Mishev *et al.*, 2013). Such

chemicals may be used to test specific endocytotic transport pathways in relation to terpene emission.

Some terpenes act as proteasome inhibitor?

Analysis of the RNAseq data indicated that the increased caryophyllene production in the transient expression assay with *CST+VAMP72-RNAi* may be due to protein stabilisation. This was not directly confirmed for the CST protein but rather inferred from the higher emission. The DsRED protein, however, was clearly stabilized by this treatment (**Chapter 5**, Fig. 9), while the up-regulation of proteasome genes can also be considered as indication of non-functional proteasomes due to the feedback up-regulation in proteasome gene expression (Kobayashi&Yamamoto, 2005, Meiners *et al.*, 2003). Intriguingly, the combination of *VAMP72-RNAi* with a linalool synthase (monoterpene synthase; *FaNES*) from strawberry gave a similar result: when expressed in *N. benthamiana* leaves, this enzyme produces the volatile linalool and within the leaf linalool-derived glycosides accumulate. When expression of *FaNES* was combined with *VAMP72-RNAi*, DsRed protein stability as well as linalool emission was increased (**Chapter 5,** Figs. 9, 10). So far, the results indicate that monoterpenes (linalool) and sesquiterpenes (caryophyllene) have an inhibitory effect on proteasome function when production in leaves is under conditions where *VAMP72s* are inhibited. It will be interesting to see whether other terpene synthases have a similar effect. This can, for instance, be tested by infiltrating *N. benthamiana* leaves with *VAMP72-RNAi* in combination with other available terpene synthases, such as geraniol synthase (Dong *et al.*, 2013) or the costunolide pathway (Liu *et al.*, 2011).

Although there are several indirect indications that the proteasome function is inhibited, this can be verified with a number of assays. For instance, an inactive proteasome would lead to a higher level of ubiquitinated proteins, which could be assessed on a western blot with ubiquitin antibodies. Also the activity of the proteases of the proteasome could be assessed directly, and whether caryophyllene and linalool directly inhibit these proteases. Another possible explanation for the effect of caryophyllene and linalool, in combination with the SNARE-RNAi, on proteasome function is that this combination results in oxidative stress which damages proteasome components. Oxidative stress is caused by an excess of free radicals that can occur because of increased oxidant levels, decreased anti-oxidant amounts, or failure to repair oxidative damage induced by reactive oxygen species (ROS) (Apel&Hirt,

2004). ROS molecules can modify proteins, lipids, and nucleic acids and thus potentially impair their natural biological function (Apel&Hirt, 2004, Cabiscol *et al.*, 2010). Oxidative damage to proteins can be monitored because the damaged proteins have amino acids with carbonyl groups which can be detected using carbonyl specific antibodies (Oxiblot detection kit) (Lee *et al.*, 2000).

So what is now our interpretation of what happens when vesicle transport is inhibited under high production of caryophyllene or linalool? Apparently, without inhibition of vesicle trafficking the cells are able to get rid of these potential toxic molecules, either by emission to the headspace (caryophyllene and linalool) or by conversion to soluble glycosides (linalool glycosides) and (presumably) sequestering into the vacuole. When vesicle transport is inhibited by the *VAMP72-RNAi* construct the assumption is that initially the terpene concentration builds up in the cell, which then somehow causes the inhibition of the proteasome. Because of the reduced protein turnover, terpene production increases and eventually an alternative emission pathway, which does not require vesicle fusion (e.g. Pathway 1, 5 in Fig. 1) is resulting in higher terpene emission. To test this interpretation of the results it will be necessary in the future to determine the time course of for instance caryophyllene emission before, and directly after induction of an inducible *VAMP72-RNAi* construct. The prediction is that caryophyllene emission changes from normal to low and subsequently to higher than normal.

Preventing caryophyllene emission affects cell biology: Implications for cancer treatments?

Many terpenes have been shown to have specific biological activity against cancer (Demain&Vaishnav, 2011, Gershenzon&Dudareva, 2007) and indeed inhibition of the proteasome is one of the targets to screen for potential activity of metabolites that could be used in cancer treatment (Orlowski&Kuhn, 2008). Several studies have suggested that inhibition of the proteasome may trigger apoptosis in tumor cells (Adams, 2004, Almond&Cohen, 2002, Kupperman *et al.*, 2010). While several different types of terpenoids isolated from different natural sources have been shown to induce apoptosis, only for a few of these has this action been shown to be related to a specific inhibition of the proteasome. These include heretonemin, a sesterterpene isolated from a sponge (Schumacher *et al.*, 2010), petrosaspongiolide M, a sesterterpene isolated from another marine organism (Margarucci *et*

al., 2010) and celastrol, a triterpene isolated from *Tripterygium wilfordii* Hook F (Yang *et al.*, 2006)). Our results with the transient expression in *N. benthamiana* now indicate that the effectiveness of a biologically active terpene may be increased when the mechanism to remove the biologically active compound from the cell is impaired. In other words, a combination of low dosage of caryophyllene combined with a low dosage of a chemical interfering with vesicle transport (e.g. brefeldin A) might be more effective than a high dosage of either chemical alone.

A long road with distractions but with future perspectives

In this thesis, I tried to address the mechanism(s) of terpene transport in different ways. In the search for the role of LTPs, unfortunately initially I was misled by unexpected pollen and ovule phenotypes, while the experiments with SNARE inhibition resulted in unexpected effects on protein stability. Results have been difficult to interpret at times, most likely because we are dealing with multiple transport pathways, and affecting a single pathway may be compensated for by alternative pathways. Only very shortly after the experimental thesis work was finished, the role of LTPs in terpene transport became clearer as now there is evidence that LTPs can enhance the transport of terpenes over the plasma membrane by ABC transporter proteins (PDR) (Boutry, Wang, vdKrol, unpublished data) in a transient assay in *N. benthamiana.* Indeed this now does explain why overexpression of the *AaLTPs* alone with the *A. annua* biosynthesis pathway genes was not sufficient to have an effect on the product profile (**Chapter 3**). Currently, the characterisation of the specificity of the different LTPs and PDR proteins in relation to transport of specific terpenes is in progress. It now also is important to check whether expression of *VAMP72-RNAi* (**Chapter 5**) affects the export of PDR and LTP proteins in *N. benthamiana* or whether this treatment affects terpene production in a different way.

The complexity of terpene transport in and between the different cell compartments and to the extracellular matrix seems at present to be the bottleneck for efficient terpene production in different heterologous plant hosts. For instance, although supposedly all the relevant biosynthesis genes of the artemisinin pathway from *A. annua* have been cloned and this whole pathway can be expressed in tobacco plants, the yield of the ectopic expression of this set of genes is still very limited, suggesting that still crucial steps are missing (Farhi *et al.*, 2011). Especially the high number of side reactions seems to limit the yield for many desired

terpene products in heterologous plant production systems. In both stable as well as transient expression, an ectopically expressed pathway may result in the production of modified products in the form of further hydroxylated, reduced, glycosylated, and malonylated products of which the formation is catalysed by endogenous enzymes of the engineered host (**Chapter 2;** (Aharoni *et al.*, 2006, Yang *et al.*, 2011)). Particularly the conjugated compounds are considered to be storage forms that usually accumulate in specific cell compartments such as the vacuole, and are produced by the plant to reduce the phytotoxic effects of ectopically produced lipophilic terpenoids (Hatzios, 1997, Ikan, 1999). With the recent results, which were based on the investigations and tools produced in this thesis work, we now have new means for the manipulation of the flux of terpenes towards the apoplast in plants and away from non-desired side reactions.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Adams J. 2004. The development of proteasome inhibitors as anticancer drugs. *Cancer cell* **5**: 417-421.

Aharoni A, Jongsma MA, Bouwmeester HJ. 2005. Volatile science? Metabolic engineering of terpenoids in plants. *Trends in Plant Science* **10**: 594-602.

Aharoni A, Jongsma MA, Kim TY, Ri MB, Giri AP, Verstappen FWA, Schwab W, Bouwmeester HJ. 2006. Metabolic engineering of terpenoid biosynthesis in plants. *Phytochemistry Reviews* **5**: 49-58.

Ajikumar PK, Xiao W-H, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. 2010. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* **330**: 70-74.

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.

Almond J, Cohen G. 2002. The proteasome: a novel target for cancer chemotherapy. *Leukemia* **16**: 433443.

Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annuual Reviews of Plant Biology* **55**: 373-399.

Aquil S, Husaini AM, Abdin MZ, Rather GM. 2009. Overexpression of the HMG-CoA reductase gene leads to enhanced artemisinin biosynthesis in transgenic *Artemisia annua* plants. *Planta Medica* **75**: 1453-1458.

Arondel V, Vergnolle C, Cantrel C, Kader JC. 2000. Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*. *Plant Science* **157**: 1-12.

Bassham DC, Blatt MR. 2008. SNAREs: cogs and coordinators in signaling and development. *Plant Physiology* **147**: 1504-1515.

Beisson F, Koo AJ, Ruuska S, Schwender J, Pollard M, Thelen JJ, Paddock T, Salas JJ, Savage L, Milcamps A, Mhaske VB, Cho Y, Ohlrogge JB. 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.

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.

Block M, Debrouwer D. 1992. In-situ enzyme histochemistry on plastic-embedded plant material. The development of an artefact-free β-glucuronidase assay. *The Plant Journal* **2**: 261-266.

Bohlmann J, Keeling CI. 2008. Terpenoid biomaterials. *The Plant Journal* **54**: 656-669.

Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA, Martin GB. 2012. A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. *Molecular Plant-Microbe Interactions* **25**: 1523-1530.

Bouvier F, Rahier A, Camara B. 2005. Biogenesis, molecular regulation and function of plant isoprenoids. *Progress in Lipid Research* **44**: 357-429.

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.

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

Budziszewski GJ, Lewis SP, Glover LW, Reineke J, Jones G, Ziemnik LS, Lonowski J, Nyfeler B, Aux G, Zhou Q. 2001. Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning. *Genetics* **159**: 1765-1778.

Cabiscol E, Tamarit J, Ros J. 2010. Oxidative stress in bacteria and protein damage by reactive oxygen species. *International Microbiology* **3**: 3-8.

Caissard J-C, Joly C, Bergougnoux V, Hugueney P, Mauriat M, Baudino S. 2004. Secretion mechanisms of volatile organic compounds in specialized cells of aromatic plants. *Recent Research Developments in Cell Biology* **2**: 1-15.

Cameron KD, Teece MA, Smart LB. 2006. Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. *Plant Physiology* **140**: 176-183.

Campbell EJ, Schenk PM, Kazan K, Penninckx IA, Anderson JP, Maclean DJ, Cammue BP, Ebert PR, Manners JM. 2003. Pathogen-responsive expression of a putative ATP-binding cassette transporter gene conferring resistance to the diterpenoid sclareol is regulated by multiple defense signaling pathways in Arabidopsis. *Plant Physiology* **133**: 1272-1284.

Carter C, Pan S, Zouhar J, Avila EL, Girke T, Raikhel NV. 2004. The vegetative vacuole proteome of *Arabidopsis thaliana* reveals predicted and unexpected proteins. *The Plant Cell* **16**: 3285-3303.

Carvalho DO, Teodoro DS, Eduardo C, Da Cunha M, Okorokova-Façanha AL, Okorokov LA, Fernandes KV, Gomes VM. 2004. Intracellular localization of a lipid transfer protein in *Vigna unguiculata* seeds. *Physiologia Plantarum* **122**: 328-336.

Castle LA, Errampalli D, Atherton TL, Franzmann LH, Yoon ES, Meinke DW. 1993. Genetic and molecular characterization of embryonic mutants identified following seed transformation in Arabidopsis. *Molecular and General Genetics* **241**: 504-514.

Chae K, Gonong BJ, Kim SC, Kieslich CA, Morikis D, Balasubramanian S, Lord EM. 2010. A multifaceted study of stigma/style cysteine-rich adhesin (SCA)-like *Arabidopsis* lipid transfer proteins (LTPs) suggests diversified roles for these LTPs in plant growth and reproduction. *Journal of Experimental Botany* **61**: 4277-4290.

Chae K, Kieslich CA, Morikis D, Kim SC, Lord EM. 2009. A gain-of-function mutation of Arabidopsis lipid transfer protein 5 disturbs pollen tube tip growth and fertilization. *The Plant Cell* **21**: 3902-3914.

Chandran SS, Kealey JT, Reeves CD. 2011. Microbial production of isoprenoids. *Process Biochemistry* **46**: 1703-1710.

Charvolin D, Douliez JP, Marion D, Cohen-Addad C, Pebay-Peyroula E. 1999. The crystal structure of a wheat nonspecific lipid transfer protein (ns-LTP1) complexed with two molecules of phospholipid at 2.1 Å resolution. *European Journal of Biochemistry* **264**: 562-568.

Chen C, Chen G, Hao X, Cao B, Chen Q, Liu S, Lei J. 2011. *CaMF2*, an anther-specific lipid transfer protein (LTP) gene, affects pollen development in *Capsicum annuum* L. *Plant Science* **181**: 439-448.

Chen D-H, Ye H-C, Li G-F. 2000. Expression of a chimeric farnesyl diphosphate synthase gene in *Artemisia annua* L. transgenic plants via *Agrobacterium tumefaciens*-mediated transformation. *Plant Science* **155**: 179-185.

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

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

Chen HJ, Wen IC, Huang GJ, Hou WC, Lin YH. 2008. Expression of sweet potato asparaginyl endopeptidase caused altered phenotypic characteristics in transgenic Arabidopsis. *Botanical Studies* **49**: 109-117.

Chen Y, Shen Q, Wang Y, Wang T, Wu S, Zhang L, Lu X, Zhang F, Jiang W, Qiu B. 2013. The stacked over-expression of *FPS*, *CYP71AV1* and *CPR* genes leads to the increase of artemisinin level in *Artemisia annua* L. *Plant Biotechnology Reports* **7**: 287-295.

Choi YE, Lim S, Kim HJ, Han JY, Lee MH, Yang Y, Kim JA, Kim YS. 2012. Tobacco *NtLTP1*, a glandular-specific lipid transfer protein, is required for lipid secretion from glandular trichomes. *The Plant Journal* **70**: 480-491.

Chowhan N, Singh HP, Batish DR, Kaur S, Ahuja N, Kohli RK. 2013. β-Pinene inhibited germination and early growth involves membrane peroxidation. *Protoplasma* **250**: 691–700.

Chuang C-F, Meyerowitz EM. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **97**: 4985-4990.

Clark KA, Krysan PJ. 2010. Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *The Plant Journal* **64**: 990-1001.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676.

Covello PS. 2008. Making artemisinin. *Phytochemistry* **69**: 2881-2885.

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

Curtis MJ, Belcram K, Bollmann SR, Tominey CM, Hoffman PD, Mercier R, Hays JB. 2009. Reciprocal chromosome translocation associated with TDNA-insertion mutation in Arabidopsis: genetic and cytological analyses of consequences for gametophyte development and for construction of doubly mutant lines. *Planta* **229**: 731-745.

Custers JBM. 2003. Microspore culture in rapeseed (*Brassica napus* L.). In: *Doubled haploid production in crop plants: a manual*, edited by Maluszynski M, Kasha KJ, Forster B and Szarejko I: Kluwer Academic Publishers, p. 185–193.

Dai X, Wang G, Yang DS, Tang Y, Broun P, Marks MD, Sumner LW, Dixon RA, Zhao PX. 2010. TrichOME: a comparative omics database for plant trichomes. *Plant Physiology* **152**: 44-54.

Dayan FE, Hernández A, Allen SN, Moraes RM, Vroman JA, Avery MA, Duke SO. 1999. Comparative phytotoxicity of artemisinin and several sesquiterpene analogues. *Phytochemistry* **50**: 607-614.

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.

Dean M. 2009. ABC transporters, drug resistance, and cancer stem cells. *Journal of Mammary Gland Biology and Neoplasia* **14**: 3-9.

Demain AL, Vaishnav P. 2011. Natural products for cancer chemotherapy. *Microbial Biotechnology* **4**: 687-699.

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.

Dong L, Miettinen K, Goedbloed M, Verstappen FW, Voster A, Jongsma MA, Memelink J, Krol Svd, Bouwmeester HJ. 2013. Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis* : Similar activity but difference in subcellular localization. *Metabolic Engineering* **20**: 198-211.

Douliez J-P, Michon T, Elmorjani K, Marion D. 2000. Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. *Journal of Cereal Science* **32**: 1-20.

Drakakaki G, Robert S, Szatmari A-M, Brown MQ, Nagawa S, Van Damme D, Leonard M, Yang Z, Girke T, Schmid SL. 2011. Clusters of bioactive compounds target dynamic endomembrane networks in vivo. *Proceedings of the National Academy of Sciences, USA* **108**: 17850-17855.

Dreger H, Westphal K, Wilck N, Baumann G, Stangl V, Stangl K, Meiners S. 2010. Protection of vascular cells from oxidative stress by proteasome inhibition depends on Nrf2. *Cardiovascular Research* **85**: 395-403.

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

Duke MV, Paul RN, Elsohly HN, Sturtz G, Duke SO. 1994. Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L. *International Journal of Plant Sciences*: 365-372.

Duke SO, Paul RN. 1993. Development and fine structure of the glandular trichomes of *Artemisia annua* L. *International Journal of Plant Sciences*: 107-118.

Eckstein-Ludwig U, Webb RJ, Van Goethem ID, East JM, Lee AG, Kimura M, O'Neill PM, Bray PG, Ward SA, Krishna S. 2003. Artemisinins target the SERCA of *Plasmodium falciparum. Nature* **424**: 957-961.

Edwards K, Johnstone C, Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**: 1349.

Effmert U, Grosse J, Rose US, Ehrig F, Kagi R, Piechulla B. 2005. Volatile composition, emission pattern, and localization of floral scent emission in *Mirabilis jalapa* (Nyctaginaceae). *American Journal of Botany* **92**: 2-12.

Farhi M, Kozin M, Duchin S, Vainstein A. 2013. Metabolic engineering of plants for artemisinin synthesis. *Biotechnology and Genetic Engineering Reviews* **29**: 135-148.

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

Federico ML, Kaeppler HF, Skadsen RW. 2005. The complex developmental expression of a novel stress-responsive barley Ltp gene is determined by a shortened promoter sequence. *Plant Molecular Biology* **57**: 35-51.

Filippini F, Rossi V, Galli T, Budillon A, D'Urso M, D'Esposito M. 2001. Longins: a new evolutionary conserved VAMP family sharing a novel SNARE domain. *Trends in Biochemical Sciences* **26**: 407-409.

Friedmann M, Ralph SG, Aeschliman D, Zhuang J, Ritland K, Ellis BE, Bohlmann J, Douglas CJ. 2007. Microarray gene expression profiling of developmental transitions in Sitka spruce (*Picea sitchensis*) apical shoots. *Journal of Experimental Botany* **58**: 593-614.

Fujimoto M, Ueda T. 2012. Conserved and plant-unique mechanisms regulating plant post-Golgi traffic. *Frontiers in Plant Science* **3**.

Gersbach P. 2002. The essential oil secretory structures of *Prostanthera ovalifolia* (Lamiaceae). *Annals of Botany* **89**: 255-260.

Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nature Chemical Biology* **3**: 408-414.

Gerst JE. 1999. SNAREs and SNARE regulators in membrane fusion and exocytosis. *Cellular and Molecular Life Sciences* **55**: 707-734.

Graña E, Sotelo T, Díaz-Tielas C, Reigosa MJ, Sánchez-Moreiras AM. 2013. The phytotoxic potential of the terpenoid citral on seedlings and sdult plants. *Weed Science* **61**: 469-481.

Guo J, Yuan Y, Liu Z, Zhu J. 2013. Development and structure of internal glands and external glandular trichomes in *Pogostemon cablin*. *PLoS ONE* **8**: e77862.

Guo L, Yang H, Zhang X, Yang S. 2013. *Lipid transfer protein 3* as a target of MYB96 mediates freezing and drought stress in Arabidopsis. *Journal of Experimental Botany* **64**: 1755-1767.

Hachez C, Besserer A, Chevalier AS, Chaumont F. 2013. Insights into plant plasma membrane aquaporin trafficking. *Trends in Plant Science* **18**: 344-352.

Harada E, Kim JA, Meyer AJ, Hell R, Clemens S, Choi YE. 2010. Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. *Plant & Cell Physiology* **51**: 1627-1637.

Hatzios KK. 1997. *Regulation of enzymatic systems detoxifying xenobiotics in plants*: Springer Netherlands.

Heinemann B, Andersen KV, Nielsen PR, Bech LM, Poulsen FM. 1996. Structure in solution of a four-helix lipid binding protein. *Protein Science* **5**: 13-23.

Heinig U, Gutensohn M, Dudareva N, Aharoni A. 2013. The challenges of cellular compartmentalization in plant metabolic engineering. *Current Opinion in Biotechnology* **24**: 239-246.

Hollenbach B, Schreiber L, Hartung W, Dietz K-J. 1997. Cadmium leads to stimulated expression of the lipid transfer protein genes in barley: implications for the involvement of lipid transfer proteins in wax assembly. *Planta* **203**: 9-19.

Holopainen JK, Gershenzon J. 2010. Multiple stress factors and the emission of plant VOCs. *Trends in Plant Science* **15**: 176-184.

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*. *Metabolic Engineering* **15**: 88-97.

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)-beta-caryophyllene, is a defense against a bacterial pathogen. *New Phytologist* **193**: 997-1008.

Huang M-D, Chen T-LL, Huang AH. 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.

Ikan R. 1999. *Naturally occurring glycosides*: John Wiley & Sons.

Ivanov S, Fedorova EE, Limpens E, De Mita S, Genre A, Bonfante P, Bisseling T. 2012. Rhizobium-legume symbiosis shares an exocytotic pathway required for arbuscule formation. *Proceedings of the National Academy of Sciences, USA* **109**: 8316-8321.

Ivanov S. 2012. *The formation of endosymbiotic membrane compartments: membrane identity markers and the regulation of vesicle trafficking* (Doctoral thesis, Wageningen University, Wageningen, The Netherlands).

Jasiński M, Stukkens Y, Degand H, Purnelle B, Marchand-Brynaert J, Boutry M. 2001. A plant plasma membrane ATP binding cassette–type transporter is involved in antifungal terpenoid secretion. *The Plant Cell* **13**: 1095-1107.

Jing F-y, Zhang L, Li M-y, Tang K-x. 2008. Over-expressing *CYP71AV1* and *CPR* genes enhances artemisinin content in *Artemisia annua* L. *Journal of Agricultural Science Technology* **10**: 64-70.

Jürgens G. 2004. Membrane trafficking in plants. *Annual Review of Cell and Developmental Biology* **20**: 481-504.

Jung HW, Kim W, Hwang BK. 2003. Three pathogen-inducible genes encoding lipid transfer protein from pepper are differentially activated by pathogens, abiotic, and environmental stresses. *Plant, Cell & Environment* **26**: 915-928.

Kader JC. 1996. Lipid-transfer proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 627-654.

Kang J, Hwang J-U, Lee M, Kim Y-Y, Assmann SM, Martinoia E, Lee Y. 2010. PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proceedings of the National Academy of Sciences, USA* **107**: 2355-2360.

Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, Martinoia E. 2011. Plant ABC transporters. *The Arabidopsis Book* **9**: e0153.

Kim D-Y, Bovet L, Maeshima M, Martinoia E, Lee Y. 2007. The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance. *The Plant Journal* **50**: 207-218.

Kim D-Y, Jin J-Y, Alejandro S, Martinoia E, Lee Y. 2010. Overexpression of AtABCG36 improves drought and salt stress resistance in Arabidopsis. *Physiologia Plantarum* **139**: 170-180.

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.

Kim S-H, Chang Y-J, Kim S-U. 2008. Tissue specificity and developmental pattern of amorpha-4, 11-diene synthase (ADS) proved by ADS promoter-driven GUS expression in the heterologous plant, *Arabidopsis thaliana*. *Planta Medica* **74**: 188-193.

Kim S-J, Brandizzi F. 2012. News and views into the SNARE complexity in Arabidopsis. *Frontiers in Plant Science* **3**.

Knudsen JT, Eriksson R, Gershenzon J, Ståhl B. 2006. Diversity and distribution of floral scent. *The Botanical Review* **72**: 1-120.

Kobayashi M, Yamamoto M. 2005. Molecular mechanisms activating the Nrf2-Keap1 pathway of antioxidant gene regulation. *Antioxidants & Redox Signaling* **7**: 385-394.

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.

Kopetzki D, Levesque F, Seeberger PH. 2013. A continuous-flow process for the synthesis of artemisinin. *Chemistry* **19**: 5450-5456.

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

Kupperman E, Lee EC, Cao Y, Bannerman B, Fitzgerald M, Berger A, Yu J, Yang Y, Hales P, Bruzzese F. 2010. Evaluation of the proteasome inhibitor MLN9708 in preclinical models of human cancer. *Cancer Research* **70**: 1970-1980.

Kutchan TM. 2005. A role for intra- and intercellular translocation in natural product biosynthesis. *Current Opinion in Plant Biology* **8**: 292-300.

Kwon C, Bednarek P, Schulze-Lefert P. 2008. Secretory pathways in plant immune responses. *Plant Physiology* **147**: 1575-1583.

Lang T, Jahn R. 2008. Core proteins of the secretory machinery. *Handbook of Experimental Pharmacology*: 107-127.

Lange BM, Wildung MR, Stauber EJ, Sanchez C, Pouchnik D, Croteau R. 2000. Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proceedings of the National Academy of Sciences, USA* **97**: 2934-2939.

Lee KO, Jang HH, Jung BG, Chi YH, Lee JY, Choi YO, Lee JR, Lim CO, Cho MJ, Lee SY. 2000. Rice 1Cys-peroxiredoxin over-expressed in transgenic tobacco does not maintain dormancy but enhances antioxidant activity. *FEBS Letters* **486**: 103-106.

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.

Lenihan JR, Tsuruta H, Diola D, Renninger NS, Regentin R. 2008. Developing an industrial artemisinic acid fermentation process to support the cost-effective production of antimalarial artemisinin-based combination therapies. *Biotechnology Progress* **24**: 1026-1032.

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

Li R, Yu C, Li Y, Lam T-W, Yiu S-M, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**: 1966-1967.

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.

Liu X, Chen C-Y, Wang K-C, Luo M, Tai R, Yuan L, Zhao M, Yang S, Tian G, Cui Y. 2013. PHYTOCHROME INTERACTING FACTOR3 associates with the histone deacetylase HDA15 in repression of chlorophyll biosynthesis and photosynthesis in etiolated Arabidopsis seedlings. *The Plant Cell* **25**: 1258-1273.

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

Lücker J, Schwab W, van Hautum B, Blaas J, van der Plas LH, Bouwmeester HJ, Verhoeven HA. 2004. Increased and altered fragrance of tobacco plants after metabolic engineering using three monoterpene synthases from lemon. *Plant Physiology* **134**: 510-519.

Ma C, Wang H, Lu X, Wang H, Xu G, 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.

Maes L, Van Nieuwerburgh FCW, Zhang Y, Reed DW, Pollier J, Vande Casteele SRF, Inzé D, Covello PS, Deforce DLD, 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 SS, Croteau RB. 2001. Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. *Proceedings of the National Academy of Sciences, USA* **98**: 8915-8920.

Maldonado AM, Doerner P, Dixon RA, Lamb CJ, Cameron RK. 2002. A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature* **419**: 399-403.

Manford AG, Stefan CJ, Yuan HL, MacGurn JA, Emr SD. 2012. ER-to-plasma membrane tethering proteins regulate cell signaling and ER morphology. *Developmental Cell* **23**: 1129-1140.

Marchant R, Robards A. 1968. Membrane systems associated with the plasmalemma of plant cells. *Annals of Botany* **32**: 457-471.

Margarucci L, Monti MC, Tosco A, Riccio R, Casapullo A. 2010. Chemical proteomics discloses petrosapongiolide M, an antiinflammatory marine sesterterpene, as a proteasome inhibitor. *Angewandte Chemie International Edition* **49**: 3960-3963.

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

Martin DM, Toub O, Chiang A, Lo BC, Ohse S, Lund ST, Bohlmann J. 2009. The bouquet of grapevine (*Vitis vinifera* L. cv. Cabernet Sauvignon) flowers arises from the biosynthesis of

sesquiterpene volatiles in pollen grains. *Proceedings of the National Academy of Sciences, USA* **106**: 7245-7250.

Martinoia E, Klein M, Geisler M, Bovet L, Forestier C, Kolukisaoglu Ü, Müller-Röber B, Schulz B. 2002. Multifunctionality of plant ABC transporters–more than just detoxifiers. *Planta* **214**: 345-355.

Meiners S, Heyken D, Weller A, Ludwig A, Stangl K, Kloetzel PM, Kruger E. 2003. Inhibition of proteasome activity induces concerted expression of proteasome genes and *de novo* formation of mammalian proteasomes. *Journal of Biological Chemistry* **278**: 21517-21525.

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.

Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Woittiez L, van der Krol S, Lugan R, Ilc T, Verpoorte R, Oksman-Caldentey K-M, Martinoia E, Bouwmeester HJ, Goossens A, Memelink J, Werck-Reichhart D. 2014. Discovery and reconstitution of the secoiridoid pathway from *Catharanthus roseus*. *Nature Communications* : Accepted.

Miki D, Itoh R, Shimamoto K. 2005. RNA silencing of single and multiple members in a gene family of rice. *Plant Physiology* **138**: 1903-1913.

Mishev K, Dejonghe W, Russinova E. 2013. Small molecules for dissecting endomembrane trafficking: a cross-systems view. *Chemistry & Biology* **20**: 475-486.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**: 621-628.

Muñoz-Bertomeu J, Arrillaga I, Ros R, Segura J. 2006. Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant Physiology* **142**: 890-900.

Nagegowda DA. 2010. Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. *FEBS Letters* **584**: 2965-2973.

Ng TB, Cheung RC, Wong JH, Ye X. 2012. Lipid-transfer proteins. *Biopolymers* **98**: 268-279.

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.

Nieuwland J, Feron R, Huisman BA, Fasolino A, Hilbers CW, Derksen J, Mariani C. 2005. Lipid transfer proteins enhance cell wall extension in tobacco. *The Plant Cell* **17**: 2009-2019.

Niinemets U, Loreto F, Reichstein M. 2004. Physiological and physicochemical controls on foliar volatile organic compound emissions. *Trends in Plant Science* **9**: 180-186.

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **27**: 29-34.

Olofsson L, Lundgren A, Brodelius PE. 2012. Trichome isolation with and without fixation using laser microdissection and pressure catapulting followed by RNA amplification: Expression of genes of terpene metabolism in apical and sub-apical trichome cells of *Artemisia annua* L. *Plant Science* **183**: 9-13.

Orlowski RZ, Kuhn DJ. 2008. Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clinical Cancer Research* **14**: 1649-1657.

Ossowski S, Schwab R, Weigel D. 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal* **53**: 674-690.

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD. 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**: 528-532.

Pagnussat L, Burbach C, Baluška F, de la Canal L. 2012. An extracellular lipid transfer protein is relocalized intracellularly during seed germination. *Journal of Experimental Botany* **63**: 6555-6563.

Pagnussat LA, Lombardo C, Regente M, Pinedo M, Martín M, de la Canal L. 2009. Unexpected localization of a lipid transfer protein in germinating sunflower seeds. *Journal of Plant Physiology* **166**: 797-806.

Park K-R, Nam D, Yun H-M, Lee S-G, Jang H-J, Sethi G, Cho SK, Ahn KS. 2011. β-Caryophyllene oxide inhibits growth and induces apoptosis through the suppression of PI3K/AKT/mTOR/S6K1 pathways and ROS-mediated MAPKs activation. *Cancer Letters* **312**: 178-188.

Pichersky E, Gershenzon J. 2002. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Current Opinion in Plant Biology* **5**: 237-243.

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.

Potocka I, Baldwin TC, Kurczynska EU. 2012. Distribution of lipid transfer protein 1 (LTP1) epitopes associated with morphogenic events during somatic embryogenesis of *Arabidopsis thaliana*. *Plant Cell Reports* **31**: 2031-2045.

Reyes FC, Buono R, Otegui MS. 2011. Plant endosomal trafficking pathways. *Current Opinion in Plant Biology* **14**: 666-673.

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.

Růžička K, Strader LC, Bailly A, Yang H, Blakeslee J, Łangowski Ł, Nejedlá E, Fujita H, Itoh H, Syōno K. 2010. Arabidopsis PIS1 encodes the ABCG37 transporter of auxinic compounds including the auxin precursor indole-3-butyric acid. *Proceedings of the National Academy of Sciences, USA* **107**: 10749-10753.

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.

Sacchetti G, Romagnoli C, Nicoletti M, Di Fabio A, Bruni A, Poli F. 1999. Glandular trichomes of *Calceolaria adscendens* Lidl.(Scrophulariaceae): histochemistry, development and ultrastructure. *Annals of Botany* **83**: 87-92.

Sanderfoot A. 2007. Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants. *Plant Physiology* **144**: 6-17.

Schilmiller AL, Miner DP, Larson M, McDowell E, Gang DR, Wilkerson C, Last RL. 2010. Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. *Plant Physiology* **153**: 1212-1223.

Schumacher M, Cerella C, Eifes S, Chateauvieux S, Morceau F, Jaspars M, Dicato M, Diederich M. 2010. Heteronemin, a spongean sesterterpene, inhibits TNF alpha-induced NF-kappa B activation through proteasome inhibition and induces apoptotic cell death. *Biochemical Pharmacology* **79**: 610-622.

Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. 2006. Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *The Plant Cell* **18**: 1121-1133.

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 Science, USA* **96**: 1863-1868.

Shin DH, Lee JY, Hwang KY, Kyu Kim K, Suh SW. 1995. High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure* **3**: 189-199.

Skubatz H, Kunkel DD, Patt JM, Howald WN, Hartman TG, Meeuse B. 1995. Pathway of terpene excretion by the appendix of *Sauromatum guttatum*. *Proceedings of the National Academy of Sciences, USA* **92**: 10084-10088.

Smyth DR, Bowman JL, Meyerowitz EM. 1990. Early flower development in Arabidopsis. *The Plant Cell* **2**: 755-767.

Staniek A, Bouwmeester H, Fraser PD, Kayser O, Martens S, Tissier A, van der Krol S, Wessjohann L, Warzecha H. 2013. Natural products–modifying metabolite pathways in plants. *Biotechnology Journal* **8**: doi: 10.1002/biot.201300224.

Stein M, Dittgen J, Sánchez-Rodríguez C, Hou B-H, Molina A, Schulze-Lefert P, Lipka V, Somerville S. 2006. Arabidopsis PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration. *The Plant Cell* **18**: 731-746.

Stein W. 1986. *Transport and diffusion across cell membranes*: Elsevier.

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

Tellez MR, Canel C, Rimando AM, Duke SO. 1999. Differential accumulation of isoprenoids in glanded and glandless *Artemisia annua* L. *Phytochemistry* **52**: 1035-1040.

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.

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* **37**: 914-939.

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. *The Plant Journal* **42**: 757-771.

Tholl D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Current Opinion in Plant Biology* **9**: 297-304.

Thoma S, Hecht U, Kippers A, Botella J, De Vries S, Somerville C. 1994. Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from Arabidopsis. *Plant Physiology* **105**: 35-45.

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.

Ting HM, Wang B, Ryden AM, Woittiez L, van Herpen T, Verstappen FW, Ruyter-Spira C, Beekwilder J, Bouwmeester HJ, 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.

Tissier A. 2012. Glandular trichomes: what comes after expressed sequence tags? *The Plant Journal* **70**: 51-68.

Torres-Barceló C, Martín S, Daròs JA, Elena SF. 2008. From hypo- to hypersuppression: effect of amino acid substitutions on the RNA-silencing suppressor activity of the *Tobacco etch potyvirus* HC-Pro. *Genetics* **180**: 1039-1049.

Trapp SC, Croteau RB. 2001. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **158**: 811-832.

Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD. 2009. High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS ONE* **4**: e4489.

Van Den Brûle S, Müller A, Fleming AJ, Smart CC. 2002. The ABC transporter SpTUR2 confers resistance to the antifungal diterpene sclareol. *The Plant Journal* **30**: 649-662.

van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ. 1995. pBINPLUS: 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.

Verkhusha VV, Kuznetsova IM, Stepanenko OV, Zaraisky AG, Shavlovsky MM, Turoverov KK, Uversky VN. 2003. High stability of *Discosoma* DsRed as compared to *Aequorea* EGFP. *Biochemistry* **42**: 7879-7884.

Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, Klein M, Kolukisaoglu Ü, Lee Y, Martinoia E. 2008. Plant ABC proteins–a unified nomenclature and updated inventory. *Trends in Plant Science* **13**: 151-159.

Vickers CE, Gershenzon J, Lerdau MT, Loreto F. 2009. A unified mechanism of action for volatile isoprenoids in plant abiotic stress. *Nature Chemical Biology* **5**: 283-291.

Villeneuve NF, Lau A, Zhang DD. 2010. Regulation of the Nrf2-Keap1 antioxidant response by the ubiquitin proteasome system: an insight into cullin-ring ubiquitin ligases. *Antioxidants & Redox Signaling* **13**: 1699-1712.

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 H, Olofsson L, Lundgren A, Brodelius PE. 2011. Trichome-specific expression of amorpha-4, 11-diene synthase, a key enzyme of artemisinin biosynthesis in Artemisia annua L., as reported by a promoter-GUS fusion. *American Journal of Plant Sciences* **2**: 619-628.

Wang NJ, Lee CC, Cheng CS, Lo WC, Yang YF, Chen MN, Lyu PC. 2012. Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB). *BMC Genomics* **13 Suppl 1**: S9.

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.

Weitzel C, Simonsen HT. 2013. Cytochrome P450-enzymes involved in the biosynthesis of mono-and sesquiterpenes. *Phytochemistry Reviews*: 1-18.

Westfall PJ, Pitera DJ, Lenihan JR, Eng D, Woolard FX, Regentin R, Horning T, Tsuruta H, Melis DJ, Owens A. 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proceedings of the National Academy of Sciences, USA* **109**: E111-E118.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**: e718.

Wu S, Schalk M, Clark A, Miles RB, Coates R, Chappell J. 2006. Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nature Biotechnology* **24**: 1441-1447.

Xiong L, Lee M-W, Qi M, Yang Y. 2001. Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. *Molecular Plant-Microbe Interactions* **14**: 685-692.

Yang H, Chen D, Cui QC, Yuan X, Dou QP. 2006. Celastrol, a triterpene extracted from the Chinese "Thunder of God Vine," is a potent proteasome inhibitor and suppresses human prostate cancer growth in nude mice. *Cancer Research* **66**: 4758-4765.

Yang L-m, Mercke P, van Loon JJ, Fang Z-y, Dicke M, Jongsma MA. 2008. Expression in *Arabidopsis* of a strawberry linalool synthase gene under the control of the inducible potato PI2 promoter. *Agricultural Sciences in China* **7**: 521-534.

Yang T, Stoopen G, Yalpani N, Vervoort J, de Vos R, Voster A, Verstappen FW, Bouwmeester HJ, Jongsma MA. 2011. Metabolic engineering of geranic acid in maize to achieve fungal resistance is compromised by novel glycosylation patterns. *Metabolic Engineering* **13**: 414-425.

Yeats TH, Rose JK. 2008. The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Science* **17**: 191-198.

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

Yu ZX, Li JX, Yang CQ, Hu WL, Wang LJ, Chen XY. 2012. The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. *Molecular Plant* **5**: 353-365.

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

Zhang L, Lu X, Shen Q, Chen Y, Wang T, Zhang F, Wu S, Jiang W, Liu P, Zhang L. 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 Molecular Biology Reporter* **30**: 838-847.

Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H. 2006. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols* **1**: 641-646.

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 Δ11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *Journal of Biological Chemistry* **283**: 21501-21508.

Summary

Terpenoids represent the largest diverse group of natural products that is produced by plants. Over the years, many studies aimed at understanding the biosynthesis of terpenes in order to improve their production through different engineering strategies. Mostly these studies focussed on identification, isolation and characterization of the biosynthesis genes. They focussed much less on how these terpenes are transported within the cell and from within the cell to the apoplast/storage organ, although this may be important steps in obtaining an efficient flux through the pathway.

To study the role of transport between enzymes within the cell, in **Chapter 2**, we analysed the branched multi-enzyme pathway for artemisinin biosynthesis from *Artemisia annua*. In this pathway, the intermediate artemisinic aldehyde sits at a branchpoint in the pathway. It can either be converted to artemisinic acid by CYP71AV1 or to dihydroartemisinic aldehyde by DBR2. We showed that CYP71AV1 from a high artemisinin producing (HAP) chemotype is less efficient with the substrate artemisinic aldehyde than the CYP71AV1 from a low artemisinin producing (LAP) chemotype, resulting in a preference of the DBR2 branch of the pathway in the HAP chemotype. Although DBR2 and ALDH1 from the HAP and LAP chemotypes have the same effciency, the relative dosage of these genes also could contribute to determining the metabolite chemotype.

Lipid Transfer Proteins (LTPs) are often expressed to a high level in tissues or under conditions where terpenes are produced, but it is so far unclear whether their high expression has any function in terpene production and/or transport. To study this, we focused on the putative role of *A. annua* LTPs. In **Chapter 3**, three *A. annua LTPs* were cloned and subcellular localisation studies of *AaLTPs:GFP* showed that these proteins are secreted and accumulate in the apoplast. However, three different functional assays (transient expression in combination with the artemisinin biosynthesis pathway, stable transformation of *A. annua* with *AaLTP-RNAi* and stable overexpression of *AaLTP* in Arabidopsis) did not result in affirm conclusion about the function of AaLTPs in sesquiterpene production/transport. Subsequently we studied the role of LTPs in sesquiterpene emission from Arabidopsis flowers. Two Arabidopsis T-DNA insertion LTP mutants were obtained (*ltp1* and *ltp3*) which

each showed ~20% reduction in emission of the sesquiterpene caryophyllene by flowers. However, the role of these LTPs cannot solely be in terpene emission as detailed analysis of GUS reporter lines only shows limited overlap in the expression profiles of LTP1 and LTP3 on the one hand and the Arabidopsis sesquiterpene synthases TPS21 and TPS11 on the other.

In **Chapter 4**, we failed to obtain the *ltp1 ltp3* and *tps11 tps21* homozygous double mutants, which initially suggested a role for LTPs and/or sesquiterpenes in gametophyte development. However, control experiments showed this to be caused by chromosomal translocations in the *ltp1* and *tps21* mutants. The apparent chromosomal translocation in the *ltp1* and *tps21* lines caused aberrant pollen and ovule phenotypes, of which the latter has not been reported before, which were studied in detail and discussed.

To study whether vesicle transport plays a role in terpene transport, in **Chapter 5**, we studied the effect of inhibition of vesicle transport (by *VAMP72-RNAi*) on terpene emission. Surprisingly, transient expression of a sesquiterpene synthase (caryophyllene synthase; CST) or a monoterpene synthase (linalool synthase; FaNES) in combination with a *VAMP72-RNAi*, resulted in an increase in terpene emission. At the same time, a co-expressed DsRed protein showed higher levels of accumulation, suggesting increased protein stability. RNAseq analysis of the leaves indicated an up-regulation of proteasome-related genes. Because proteasome genes are under feedback regulation, this suggests that the transient expression of the terpene synthases in combination with *VAMP72-RNAi* results in an inhibition of proteasome function. While the effect of expressing only a terpene synthase or only *VAMP72-RNAi* was limited, the combination also resulted in strong down-regulation of photosynthesis genes in the agro-infiltrated *N. benthamiana* leaves.

In the final discussion in **Chapter 6**, we discuss the research reported in this thesis. Recent reports suggest that there is indeed a role for LTPs in terpene transport, as we postulated. I discuss how the role of the LTPs that I studied could be characterised in more detail in the future. Also I discuss the future perspectives of some of my findings. Although the roles of LTPs and vesicle transport need to be explored further, the results from this thesis can already contribute to improving terpene production by metabolic engineering in heterologous plant hosts.

Samenvatting

Terpenen zijn de grootste groep van natuurlijke producten die door planten worden gemaakt. Al sinds veel jaren wordt onderzoek gedaan aan de synthese van terpenen om de productie te kunnen verbeteren met verschillende plant-engineering technieken. Dit onderzoek richtte zich voornamelijk op de identificatie, isolatie en karakterisering van biosynthese genen en veel minder op hoe deze terpenen worden getransporteerd binnen de cel of vanuit de cel naar de apoplast, terwijl dit een belangrijke stap kan zijn voor het verkrijgen van een efficiënte flux door de biosynthese route.

Om de rol van transport tussen enzymen binnen de cel te bestuderen hebben we in **Hoofdstuk 2** de vertakte biosynthese route van artemisinine van *Artemisia annua* bestudeerd. In de biosynthese zit het intermediair artemisinine aldehyde op een splitsing in de route; het kan of worden geconverteerd naar artemisinine zuur door CYP71AV1 of naar dihydroartemisinine aldehyde door DBR2. We laten zien dat CYP71AV1 van een hoog artemisinine producerend (HAP) chemotype van *A. annua* minder efficiënt is in de omzetting van het substraat artemisinine aldehyde dan de CYP71AV1 van een laag artemisinine producerend (LAP) chemotype. Dit resulteert in een preferentiële omzetting van artemisinine aldehyde via de DBR2 tak van de biosynthese route in het HAP chemotype. Hoewel DBR2 en ALDH1 van de HAP en LAP chemotypes dezelfde efficiëntie hebben, zou de relatieve activiteit van deze twee genen ook nog kunnen bijdragen aan het bepalen van het metabole chemotype.

Lipid Transfer Proteins (LTPs) komen vaak sterk tot expressie in weefsels waar terpenen worden geproduceerd en worden geïnduceerd door dezelfde condities die terpeen biosynthese induceert. Het is echter niet duidelijk of de LTP expressie iets te maken heeft met de terpeen biosynthese dan wel hun transport. Om te bestuderen of LTPs betrokken zijn bij het transport van terpenen, hebben we eerst naar de *A. annua* LTPs gekeken. Drie AaLTP genen werden gekloneerd, maar in transiente expressie hadden deze geen effect op het product profiel van de artemisinine biosynthese route. Vervolgens hebben we in **Hoofdstuk 3** de rol van LTPs in Arabidopsis sequiterpeen emissie bestudeerd. Twee Arabidopsis T-DNA

insertie LTP mutanten werden geïdentificeerd (*ltp1* en *ltp3*) die beiden een reductie van ongeveer 20% in emissie van het sesquiterpeen caryophylleen in bloemen lieten zien.

De rol van deze LTPs kan niet beperkt zijn tot de emissie van terpenen omdat gedetailleerde analyse van GUS reporter lijnen maar een beperkte overlap liet zien tussen het expressie profiel van LTP1 en LTP3 en de Arabidopsis sesquiterpeen synthases TPS21 en TPS11.

Voor het werk aan de Arabidopsis LTPs wilde ik graag dubbelmutanten maken, maar het lukte niet om *ltp1 ltp3* en *tps11 tps21* dubbel mutanten te maken, wat aanvankelijk suggereerde dat LTPs en/of sesquiterpenen betrokken zijn bij de gametofyt ontwikkeling. Echter, controle experimenten lieten uiteindelijk zien dat dit veroorzaakt werd door chromosomale translocaties in de *ltp1* en *tps21* mutant. De analyse van de afwijkende pollen en ovules die onstaan bij dit soort chromosomale translocaties is beschreven in **Hoofdstuk 4**.

Om te onderzoeken of vesicle transport betrokken is bij het transport van terpenen, hebben we in **Hoofdstuk 5** het effect van remming van vesicle transport (door *VAMP72-RNAi*) op terpeen emissie bestudeerd. Verrassend genoeg blijk dat bij transinte expressie van een sesquiterpeen synthase (caryophylleen synthase; CST) of een monoterpene synthase (linalool synthase; FaNES) in combinatie met *VAMP72-RNAi*, de terpeen emissie hoger werd. Dit ging samen met een toename in het DsRed eiwit dat tegelijkertijd tot expressie werd gebracht, wat duidt op een verhoogde eiwit stabiliteit. RNAseq analyse van de bladeren liet zien dat proteasoom genen een hogere expressie vertoonden. Omdat proteasome genen onder feedback regulatie staan, suggereert dit dat de transiente expressie van de terpeen synthases samen met *VAMP72-RNAi* resulteert in een remming van de proteasoom functie. Het effect van alleen terpeen synthase expressie of alleen *VAMP72-RNAi* expressie was beperkt. De combinatie van de twee resulteerde ook in sterke suppressie van de fotosynthese genen in agro-geïnfiltreerde *N. benthamiana* bladeren.

In het laatste **Hoofdstuk 6** bediscussieer ik de resultaten van de verschillende hoofdstukken. Ook bediscussieer ik recente bevindingen die er op wijzen dat er inderdaad een rol is voor LTPs in terpenen transport en bespreek ik een aantal strategieën om beter aan te tonen wat de rol is van de *A. annua* en Arabidopsis LTPs in terpenen transport. Tenslotte bediscussieer ik een aantal van de toekomst perspectieven van mijn werk aan terpenen transport. Hoewel de rol van LTPs en vesicle transport nog verder onderzocht moet worden, kunnen de resultaten die zijn behaald in dit onderzoek al worden toegepast in de metabole engineering van terpeen productie in heterologe host planten.

Acknowledgments

Finally, it took me almost 5.5 years to finish my PhD thesis. I felt like I am on some kind of roller coaster ride during this PhD period: very often, I got a very “good result” that drove me straight up and made me hopeful; soon after, it also drove me down so quickly. I always tried to detach myself from those “negative” results and keep on going to finish my thesis. Thank you so much for so many people who either directly or indirectly helped me during this process. My apologies and thanks to all people I might have forgotten to mention in these acknowledgements.

To begin with, I would like to express my sincerest gratitude to my promoter, Prof. Dr. Harro Bouwmeester and daily supervisor, Dr. Sander van der Krol for giving me the opportunity to join the Laboratory of Plant Physiology and to start my PhD in Wageningen. I really appreciated both of you guiding me towards the completion of this challenging, risky project until the very last moments. Thank you Harro to give me comments and feedbacks on my research and papers. You helped me to find collaborators to do different experiments and ask materials for my experiments. You always tried your best to deal with my requests even if you were super busy.

I am really amazed, Sander, about you, for the billion ideas in your mind and the ability to connect different subjects together. From that, I learned how to arrange my time to do different kinds of experiments, to negotiate, to focus, and to be a more independent researcher. Thank you for giving me many advices on my presentations, reports, papers and helping me to improve my writing and presentation skills. Moreover, thanks for arranging writing weeks to force us to write as much as possible. I also enjoyed a lot the boat trips and dinner at your nice house!

I would also like to thank my external supervisor, Prof. Dr. Dirk Bosch to give me suggestions at the beginning of my PhD study. I extend my gratitude to Dr. Jules Beekwilder for guiding me to analyze the LC-MS/MS data and giving me a lot of feedback and suggestions on my paper. I would like to acknowledge Dr. Kim Boutilier for her knowledge and her contribution to the LTP work. I would like to express my appreciation to Prof. Dr. Richard Immink, Prof. Dr. Maarten Koornneef and Dr. Mark Aarts, who are willing to spend

their time discussing and giving comments on my research. I am very grateful to Dr. Maarten Jongsma, who often joined the terpene group meetings and gave me useful advices and critical comments. Thanks Dr. Ric de Vos for the helpful discussion on LC-MS/MS data. Also, I thank Prof. Dr. Ton Bisseling, Prof. Dr. Marc Boutry, Dr. Maarten Jongsma and Dr. Matthieu Joosten for sparing time and giving comments on this thesis.

I appreciate the kind support on GC, LC, SPME measurements and corresponding data processing from Francel, Bert, Roland, Geert, Harry, Yury and Berhane. Also I would like to acknowledge the technical support in the lab from Mariëlle, Miriam and Jacqueline. Thanks Casper, Taede and Gerrit from Unifarm who took care of my plants in the green house.

I would like to thank all my fellow PhD students and the post-docs in the terpene group: Marina, Liping, Jing Mao, Teun, Benyamin, Mohammad, Mahmood, Tila, Yi Shang, Danijela, Rashmi, Iris, Thierry, Katja, Maurice, Aldana, Neli, Qing Liu, Ting Yang, Léa, Bo, Andreas, Lemeng, Jun He, Yuanyuan, Esmer, Arman and Pim, for all your support and the happy times in the lab. Moreover, I express my heartiest thanks to my student Rutger, Seifu and Tzu-Ying who were involved in the work described in this thesis. Roddy, although you are not my student, thanks for your kind support, warm concern and happy time together.

I also would like to thank all my fellow PhD students and other colleagues from PPH/ PRI for the pleasant memories during my stay in Wageningen: Anke and Desalegn (we started our PhD on the same day), Natalia Carreno Quintero, Victoria, Anna, Catarina, Ralph, Manickam, Wei Song, Carolien, Emilie, Henk, Leonie, Wilco, Bas ,Ronny, Leo, Lidiya, Dick, Diaan, Bing, Hanzi, Xi, Yanxia, Yumeng, Julio, Rik, Imran, Phuong, Fe, Rashid, Noorullah, Myriam, Wouter, Cimille, Aaron, Karen, Rene, Johanna, Natalia Moreno Pachon, Beatriz, Cecilia, Giovanni, Anderson, Deborah, Paulo, Nasr, Farzaneh, Kerstin, Melissa, Manus and Sang. I would like to express my gratitude to Rina and Magaret for always helping me with administrative issues. Also a thanks to colleagues from the Laboratory of Genetics: Pádraic, Corrie, Wytske, Ana Carolina, Aina, Charles, Frank, Mina, Mohamed, Nihal, Roxanne and Zeshan. Thanks to colleagues from other group: Juliane, Anneke, Peter Dinh, Nasim, Arwa and Nikolay.

To my dear Taiwanese friends, Chang-Lin Chen, Chenling Huang (Jenny), Che-Yang Liao, Chia-chi Chang, Chia-Kai Kang, Chia-Yi Liu, Chia-Ying Lin, Chih-Chieh Chuang, Ellen Lin, Han-Ni Tang, Hsiao-Li Hsu, Hsi-Yu Chen, Hsuan Chen, I-Chiao Lee, Jacourr Liu, Jen-Jung Liu, Jia-Yuh Lee (Eunice), Jie-Hong Huang (Red), Jo-Chien Liao (Wendy), Jou-Fu Wang, Lilly Hsiao,Michelle Yeh, Meng-Yuan Jen, Wei-Shan Chen (MoMo), Shang-Jung Lin,

Shu-Chen Liu (Phabbie), TzeYi Huang, Wen-Kai Feng, Ya-Ping Chang, Yu-Fang Chen, Yun-Shan Liu, Yun-Yueh Lu, Yu-wei Chang, I want to express my appreciation for your company in Wageningen and for those great & funny times we had. My sincere thanks to Pengying Piena-Yang for kindly taking care of us as your children and always preparing very nice food for us; to Jhan-Ye Wang for your kind help with the cover design; and to Song-Bin Chang for your kind help with our application for a PhD at Wageningen University.

To my Chinese friends, Wei Liu, Na Li, Ying Zhang, Chunzhao Zhao, Guodong Wang, Xiaomin Tang, Chenlei Hua, Lisha Zhang, Ke Lin, Ningwen Zhang, Chunxu Song, Wei Qin, Weicong Qi, Xi Chen, Xu Cheng, Yanru Song, Junwei Liu, Feng Zhu, Huchen Li, Guiling Ren, Hui Li, Jianhua Zhang, Jimeng Li, Suxian Zhu, Tingting Xiao, Wei Gao, Xianwen Ji; Yan Wang, Yanli Wang, Xiaoqian Shi, Xuan Xu, Shuhang Wang, Tao Li, Yuchen Long, Zhen Wei, Tao Zhao, I am thankful for all your support and kindness.

To my friends at the Student Chaplaincy, Rev. Josine, Fr. Wiel, Ingeborg Brouwer, Alexandre Villela, Danielle Lucas Barbosa, Robin Liu, and other friends at Chaplaincy, I'm so grateful for the time we had in Taizé, for the singing at a prison, for the great food at the global dinner, for the good wishes you delivered to us, and for the peace we prayed for. Thank you very much from the very bottom of my heart.

Thanks to the Malaysian community here in Wageningen, Kak Mai, Ching, Yana, Asyraf, Ilah, Yani, Ameen, Izan, Naim, Fatimah, Razak, Nozie, Huda, Termizi, Su, Aniey, Uncle Alan, Siew Ling, Apple Teo, Edmund, Hong Ling, Jack, Cheon, Loo Wee, Shi Pey and Freddy for sharing pleasant moments on various occasions.

I would like to express my heartfelt gratitude to my most loving parents (Hock Ing & Ah Moi) and siblings (Hieng Hock, Hieng Huat and Ley Yiew) in Malaysia; parents in law, brother and sisters in law in Taiwan; thank you for your full support, understanding and encouragement.

Most importantly, I deeply thank my beloved wife, Ya-Fen, for your love, patience and support. Lastly, I would like thank my lovely son, Yew-Yew, for being with us and giving us a lot of joy!

Thank you!!!

Hieng-Ming Ting
Wageningen, 24 March 2014

Curriculum Vitae

Hieng-Ming Ting was born on August 13, 1979 in Sibu, Sarawak, Malaysia. After obtaining his high school degree in Malaysia, he got a scholarship for his Bachelor study in the Department of Biology of National Taiwan Normal University in Taiwan (1998~2002). In 2001, he joined Prof. Shong Huang's lab to study plant population ecology and genetics, in a project entitled "Study on the genetic variation among populations of *Kandelia candel* (L.) Druce in East Asia". After obtaining his B.Sc. degree, he joined Dr. Chiu-Ping Cheng's lab in the Agricultural Biotechnology Research Centre (ABRC), Academia Sinica, Taiwan from August 2002 to August 2004. He worked on a project called "Functional genomics of plant root systems responding to stresses and secondary metabolic pathways", in which plant-pathogen interaction was studied by means of tomato cDNA microarrays. In addition, he worked on the project "Functional assays and application of plant pathogen-inducible promoters in transgenic tomato and tobacco". These two research experiences gave him a good training as a researcher and a broad view on the complexity of plant gene responses to biotic stresses. Subsequently, he pursued an M.Sc. degree in the Graduate Institute of Life Science of National Central University, Taiwan under the supervision of Dr. Ching-Hui Yeh and Prof. Tuan-Hua David Ho. Here his interest was towards plants responses to abiotic stresses in a project entitled "Functional analysis of group 3 LEA proteins". After his graduation in June 2006, he worked as a research assistant in the same lab continuing his research on the LEA proteins. In September 2007, he started his PhD study at the Institute of Plant Biology of National Taiwan University in Taipei, where his first rotation lab was Dr. Paul E. Verslues's lab and his second was Dr. Tien-Shin Yu's lab. In November 2008, Hieng-Ming started working as a PhD student at the Laboratory of Plant Physiology of Wageningen University under daily supervision of Dr. Sander van der Krol and promoter Prof. Harro Bouwmeester on the thesis entitled "Biosynthesis and transport of terpenes". Currently he is applying for a postdoctoral position in the field of secondary metabolism in plants.

Publications

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

● Na Tian, Shuoqian Liu, **Hieng-Ming Ting**, Jianan Huang, Sander van der Krol, Harro Bouwmeester and Zhonghua Liu (2013) An improved *Agrobacterium tumefaciens* mediated transformation of *Artemisia annua* L. by using stem internodes as explants.
Czech J. Genet. Plant Breed., 49(3): 123–129.

● **Hieng-Ming Ting**, Kim Boutilier, Francel W.A. Verstappen, Mariëlle Schreuder, Harro J. Bouwmeester and Alexander van der Krol (2014) LTP1 and LTP3 are involved in sesquiterpenes emission in Arabidopsis.
(In preparation for PLoS ONE)

● **Hieng-Ming Ting**, Thierry L. Delatte, Pim Kolkman, Alexander van der Krol and Harro J. Bouwmeester (2014) Inhibition of vesicle transport during terpene biosynthesis causes proteasome malfunction.
(In preparation for The Plant Journal)

● **Hieng-Ming Ting**, Ching-Hui Yeh, and Tuan-Hua David Ho (2006) Functional Analysis of Group 3 LEA Proteins. M.Sc. thesis of Institute of Life Science, National Central University, Taiwan.

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to: Hieng-Ming Ting
Date: 24 March 2014
Group: Plant Physiology, Wageningen University & Research Centre

1) Start-up phase	date
► **First presentation of your project**	
Subcellular transport in terpene biosynthesis	Apr 06, 2009
► **Writing or rewriting a project proposal**	
Subcellular transport in terpene biosynthesis	Sep 29, 2009
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	date
► **EPS PhD student days**	
EPS PhD Student Day, Leiden University, Leiden, NL	Feb 26, 2009
EPS PhD Student Day, Utrecht University, Utrecht, NL	Jun 01, 2010
2nd European Retreat of PhD Students in Plant Sciences, Cologne, DE	Apr 15-17, 2010
EPS PhD Student Day, Wageningen University, Wageningen, NL	May 20, 2011
ExPectationS Day 2011: Career Day	Nov 18, 2011
ExPectationS Day 2013: Creative and Inspiration in Science	Feb 01, 2013
► **EPS theme symposia**	
EPS theme 2 'Interactions between Plants and Biotic Agents', Utrecht University	Jan 22, 2009
EPS theme 3 'Metabolism and Adaptation', University of Amsterdam	Feb 18, 2009
EPS theme 1 'Developmental Biology of Plants', Wageningen University	Jan 28, 2010
EPS theme 3 'Metabolism and Adaptation', Leiden University	Feb 19, 2010
EPS theme 3 'Metabolism and Adaptation', Utrecht University	Apr 26, 2012
► **NWO Lunteren days and other National Platforms**	
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 06-07, 2009
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 19-20, 2010
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 04-05, 2011
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 02-03, 2012
► **Seminars (series), workshops and symposia**	
Invited seminar Sjef Smeekens	Nov 27, 2008
Invited seminars Adam Price, Ottoline Lyser	Jun 19, Aug 28, 2009
KEYS2009 seminar "Phenomics & The Green Gene Revolution"	Sep 21, 2009
Invited seminars Nick Panopoulos, Laurent Zimmerli, Chris Hawes, John Yoder, David Baulcombe, Kirsten Bomblies, Jose Jimenez-Gomez, Ales Pecinka	Jan 11- Nov 19, 2010
One day symposium on plant endomembrane system	Jul 02, 2010
CBSG Technology Symposium	Nov 25, 2010
Mini symposium: How to Write a World-class Paper	Oct 26, 2010
Invited seminar Javier Palatnik	Aug 25, 2011
Symposium: Parasitic plants and strigolactones. More than rhizosphere communication	Oct 07, 2011
Invited seminars Paul Fraser, Cornelia Spetea Wiklund, Rainer Melzer, Inez Slamet-Loedin, Doris Wagner, David G. Robinson, Ruth Finkelstein, Larry Griffing, Aaron Fait	Feb 15-Dec 04, 2012
COST Action Workshop "Data handling and interpretation in PNP biosynthesis"	Feb 17-18, 2012
Workshop: New developments in plant endomembrane biology	Jun 15, 2012
NCSB Symposium_Systems Biology: From models and data to real life applications	Nov 01-02, 2012
Invited seminars Alain Tissier, Howard S. Judelson, Rays H.Y. Jiang	May 03,07, 2013
Symposium: Systems Biology for Food, Feed, and Health	Jun 27, 2013
Plant Science seminars (2009)	2009
Plant Science seminars (2010)	2010
Plant Science seminars (2011)	2011
► **Seminar plus**	
Patrick Hussey (EPS Theme 1)	Jan 28, 2010
► **International symposia and congresses**	
FESPB 2010 (Valencia, Spain)	Jul 04-09, 2010
TERPNET 2011 (Kalmar, Sweden)	May 22-26, 2011
ASPB 2012 (Austin, Texas)	Jul 20-24, 2012
TERPNET 2013 (Crete, Greece)	Jun 01-05, 2013
► **Presentations**	
Poster: FESPB 2010	Jul 04-09, 2010
Poster: TERPNET 2011	May 22-26, 2011
Oral: EPS Theme 3	Apr 26, 2012
Poster: ASPB 2012	Jul 20-24, 2012
Poster: TERPNET 2013	Jun 01-05, 2013
► **IAB interview**	
► **Excursions**	
Green Life Science Company Visit at Seed Valley	Jun 23, 2011
KeyGene excursion	Jan 26, 2012
Subtotal Scientific Exposure	*24.6 credits**

CONTINUED ON NEXT PAGE

3) In-Depth Studies		*date*
► **EPS courses or other PhD courses**		
Postgraduate course 'Gateway to gateway technology'		Nov 17-21, 2008
EPS PhD summer school "Environmental Signaling", Utrecht		Aug 24-26, 2009
Advanced PhD Course Comparative Proteomics		April 21-23, 2010
Postgraduate course 'System Biology: omics data analysis'		Dec 13-17, 2011
► **Journal club**		
Literature discussions and work discussions at PRI and Plant Physiology		2008-2012
► **Individual research training**		
	Subtotal In-Depth Studies	*7.5 credits**

4) Personal development		*date*
► **Skill training courses**		
Project and Time Management		Sep 07,21, Oct 19, 2010
Techniques for writing and presenting a scientific paper		Dec 14-17, 2010
Systems Biology "Basic of Parameter Estimation"		Feb 10, 2012
Stress identification and Management workshop		Apr 05, 2012
Systems Biology Tutorial: Statistics with R		Apr 20, 2012
► **Organisation of PhD students day, course or conference**		
► **Membership of Board, Committee or PhD council**		
Activity committe of Laboratory of Plant Physiology		2010
	Subtotal Personal Development	*3.6 credits**

TOTAL NUMBER OF CREDIT POINTS*	**43.2**

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 credits

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

This project was financially supported by the Graduate School of Experimental Plant Sciences (EPS), Wageningen University, The Netherlands.

Layout: Hieng-Ming Ting

Cover design: Jhan-Ye Wang & Hieng-Ming Ting
"Terpene" letters on *Nicotiana benthamiana* leaves were agroinfiltrated with *CST+SNARE-RNAi* (**Chapter 5**).
Pictures were taken under normal and UV light.

Printed by Uitgeverij BOXPress || Proefschriftmaken.nl ||