

**Biosynthesis, transport and combinatorial
metabolic engineering of *Tanacetum parthenium*
(feverfew) and *Artemisia annua* (sweet
wormwood) sesquiterpene lactones**

Arman Beyraghdar Kashkooli

Thesis committee

Promotor

Prof. Dr. Harro J. Bouwmeester
Professor of Plant Physiology
Wageningen University & Research
Swammerdam Institute for Life Sciences
University of Amsterdam

Co-promotor

Dr. Sander van der Krol
Laboratory of Plant Physiology
Wageningen University & Research

Other members

Prof. Dr. S.C. de Vries - Wageningen University & Research
Prof. Dr. J. Memelink - Leiden University
Dr. Ir. R.C. Schuurink - University of Amsterdam
Dr. M.J. Beekwilder - Wageningen Plant Research

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

**Biosynthesis, transport and combinatorial
metabolic engineering of *Tanacetum parthenium*
(feverfew) and *Artemisia annua* (sweet
wormwood) sesquiterpene lactones**

Arman Beyraghdar Kashkooli

Thesis

submitted in fulfilment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public on

Monday 10 September 2018

at 4:00 p.m. in the Aula.

Arman Beyraghdar Kashkooli

Biosynthesis, transport and combinatorial metabolic engineering of *Tanacetum parthenium* (feverfew) and *Artemisia annua* (sweet wormwood) sesquiterpene lactones,

163 pages

PhD thesis Wageningen University, Wageningen, the Netherlands (2018)

With references, with summary in English

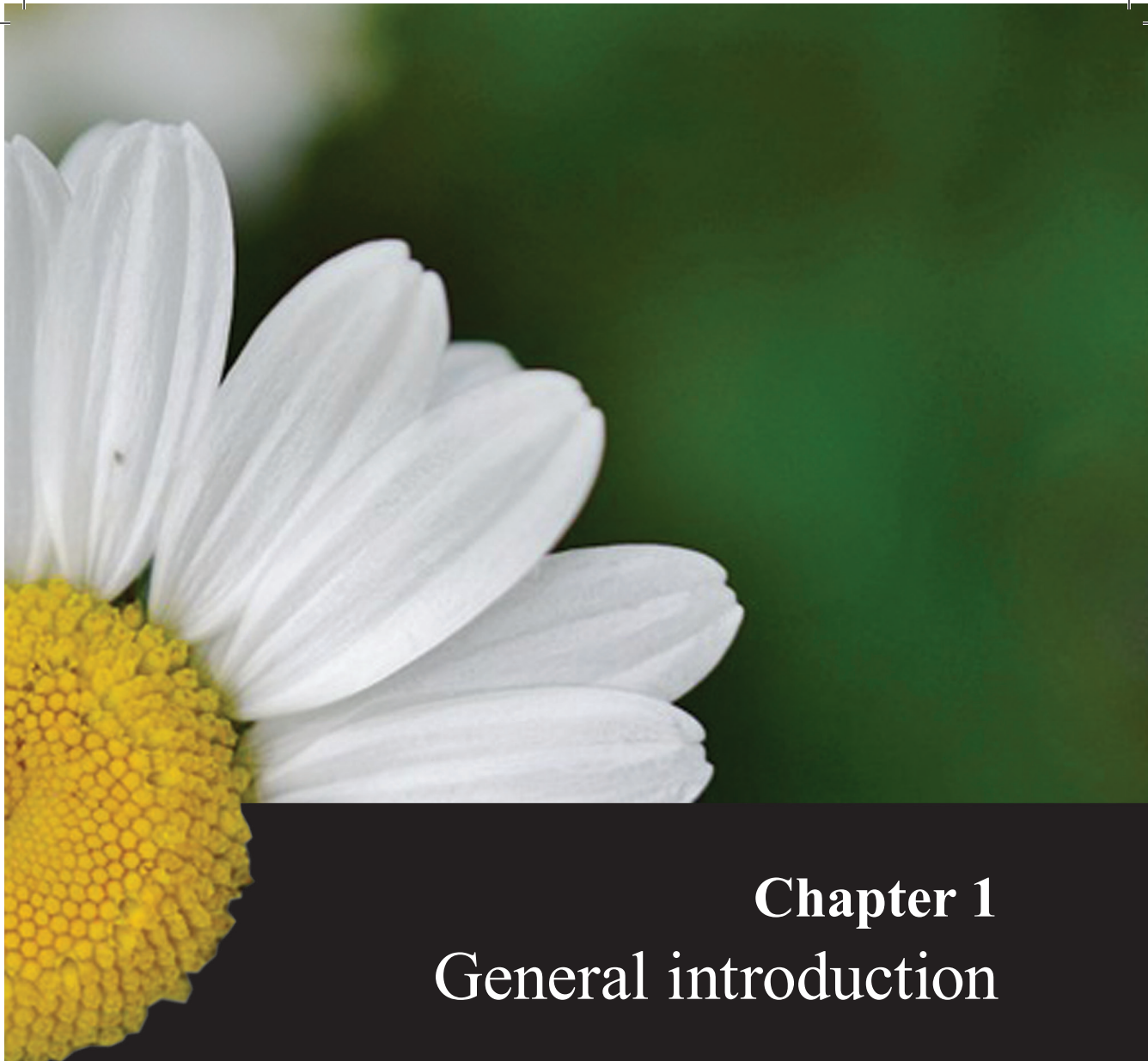
ISBN: 978-94-6343-476-8

DOI: 10.18174/455405

Contents

| | | |
|----------------------------|---|-----|
| Chapter 1 | General introduction | 7 |
| Chapter 2 | Kauniolide synthase, a P450 with unusual hydroxylation and cyclization-elimination activity | 29 |
| Chapter 3 | Substrate promiscuity of enzymes from the sesquiterpene biosynthetic pathways from <i>Artemisia annua</i> and <i>Tanacetum parthenium</i> allows for novel combinatorial sesquiterpene production | 63 |
| Chapter 4 | Transient production of artemisinin in <i>Nicotiana benthamiana</i> is boosted by a specific lipid transfer protein from <i>A. annua</i> | 83 |
| Chapter 5 | Non-specific Lipid Transfer Proteins exhibit specificity for feverfew sesquiterpene lactones | 111 |
| Chapter 6 | General discussion | 135 |
| Summary | | 153 |
| Acknowledgements | | 155 |
| About the author | | 157 |
| Education statement | | 158 |





Chapter 1

General introduction

A part of this chapter is accepted for publication in
Weurman Book with the series title *Flavour Science*.

Introduction

Plants produce a wide range of structurally diverse natural products. These natural products play key roles in the interaction of plants with organisms in their environment. They, for example, act as defence compounds against herbivores and pathogens, or as attractants of insects for pollination. They also provide a natural resource for humans and are used as medicine (e.g. artemisinin and parthenolide), pigments (e.g. β -carotene and lycopene), fragrance (e.g. limonene and linalool) and flavours (e.g. vanillin and menthol).

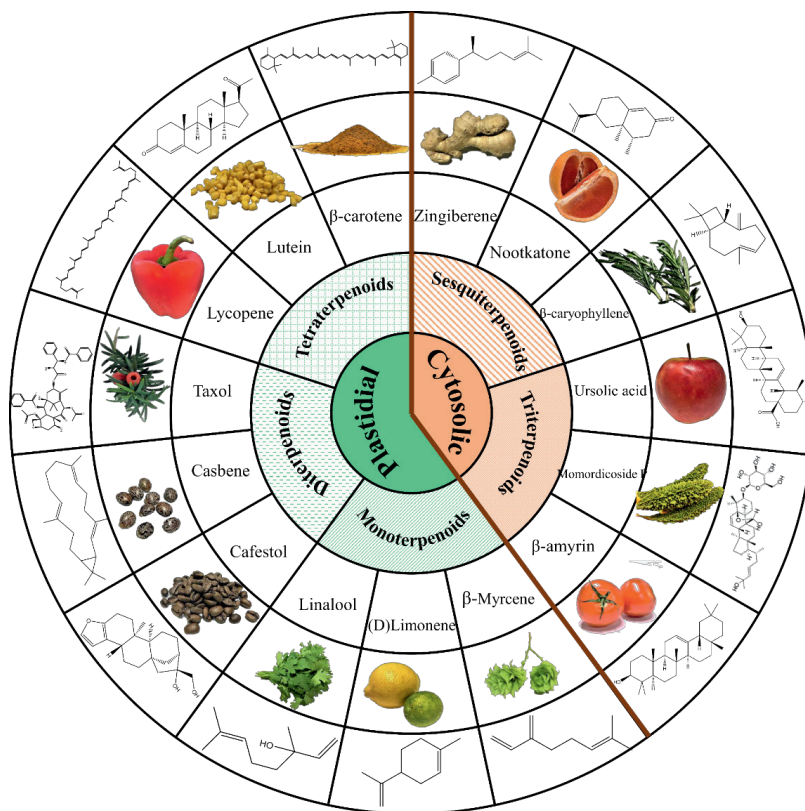


Figure 1. Schematic representation of plastidial and cytosolic terpenoids. Examples of different plant derived terpenoids and plant organs (seed, fruit, leaf or root) containing these terpenoids are shown.

Plant natural products can be divided into several classes such as nitrogen-containing (e.g. alkaloids, glucosinolates and cyanogenic glucosides) and nitrogen-free metabolites such as terpenoids, phenolics and flavonoids (Wink, 2010). Among these, the terpenoids are the most diverse class, constituting almost 12% of all known plant metabolites (Humphrey and Beale, 2006; Springob and Kutchan, 2009) and possessing many different functions in plants. Low molecular weight volatile terpenoids are involved in plant protection mechanisms during

biotic and abiotic stresses (Bleeker et al., 2011; Huang et al., 2012; Loreto et al., 2014) and when emitted from flowers can act as pollinator attracting signals (Byers et al., 2014). Terpenoids can be antifeedant compounds that protect the plant against insects, such as for example, geranylinalool in *Nicotiana obtusifolia* (Jassbi et al., 2010) and can be an activator signal for systemic acquired resistance (Chaturvedi et al., 2012; Kemen et al., 2014). Several terpenoids, such as gibberellins, abscisic acid and strigolactones are plant hormones and act as signalling molecules in physiological processes. Strigolactones, for example, are a regulator of plant axillary bud outgrowth and thus branching (Cardoso et al., 2014). In addition to their role as plant hormone, strigolactones are also secreted into the soil surrounding the plant's roots where they recruit the symbiotic arbuscular mycorrhizal fungi.

Biosynthesis of Terpenoids

Biosynthesis of the basic building blocks of terpenoids.

Terpenoids are produced from the universal building blocks, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP and DMAPP are synthesized through two different biosynthetic pathways (Figure 2). One of these occurs in the plastids and supplies mostly the substrates for the production of monoterpenoids (often present in essential oils), diterpenoids and tetraterpenoids (carotenoids) (Figure 3). The plastidial biosynthesis of IPP, also called the MEP pathway, starts with the condensation of pyruvate with glyceraldehyde 3-phosphate, yielding 1-deoxy-D-xylulose 5-phosphate (DXP). DXP is then rearranged and reduced by DXP reductoisomerase (DXR), which results in the formation of 2-C-methyl-D-erythritol 4-phosphate (MEP). MEP cytidyltransferase (MCT), 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (CMK), and 2-C-methyl-D-erythritol 2,4-cyclodiphosphate (MEcPP) synthase (MDS) convert MEP into MEcPP. Subsequently, 4-hydroxy-3-methylbut-2-enyl diphosphate (HMBPP) synthase (HDS) produces HMBPP. Formation of IPP and DMAPP is then catalysed by HDR (Herrera et al., 2006) (Figure 2). IPP isomerase finally (partially) converts IPP to DMAPP. This reaction is reversible, but with a preference towards formation of DMAPP. DXP synthase is regarded as the rate-limiting enzyme in the MEP biosynthetic pathway (Estévez et al., 2001).

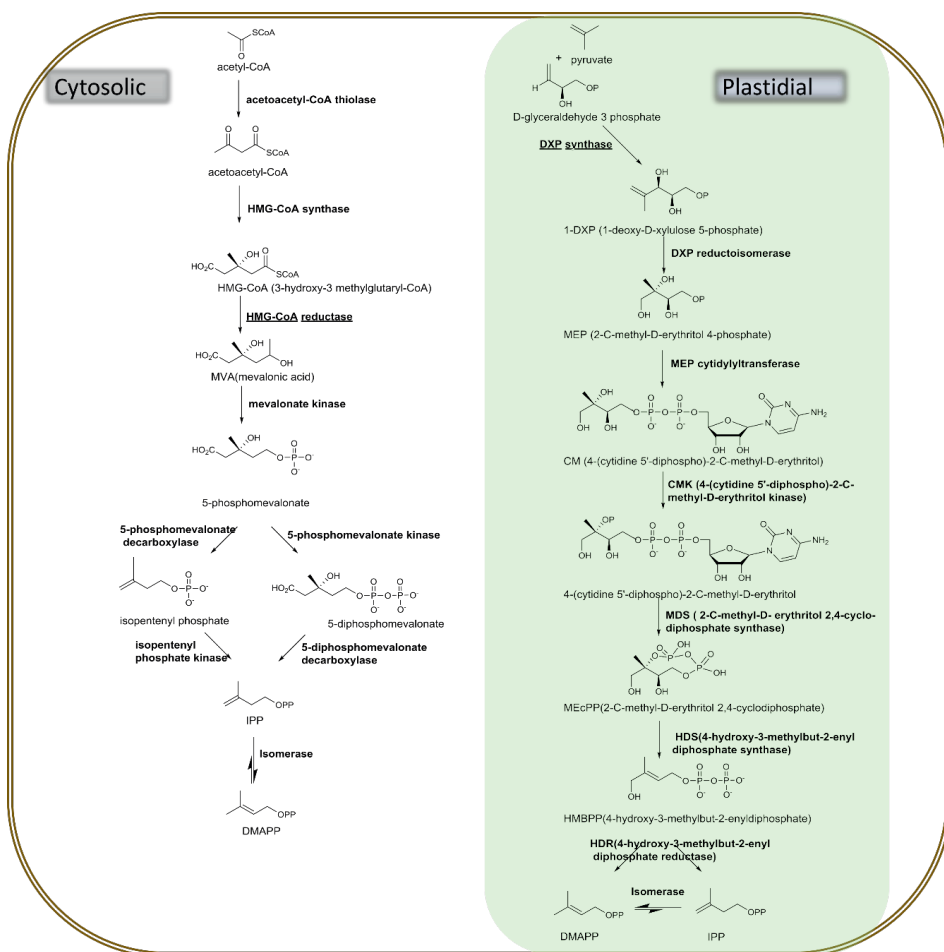


Figure 2. The MVA and MEP pathways. Enzymes and intermediates involved in the biosynthesis of IPP and DMAPP in the cytosolic pathway (MVA) (in white) and plastidial pathway (MEP) (in green).

The other pathway, known as the mevalonate (MVA) pathway takes place in the cytosol (Figure 2). The IPP and DMAPP derived from this pathway are mostly used as substrates in the production of sesquiterpenoids and triterpenoids. The cytosolic pathway encompasses two condensation events of three acetyl-CoA molecules. These reactions are catalysed by acetoacetyl-CoA thiolase and hydroxymethylglutaryl-CoA (HMG-CoA) synthase and result in the formation of HMG-CoA. HMG-CoA is subsequently reduced by HMG-CoA reductase (HMGR), which results in mevalonic acid (MVA) production. MVA phosphorylation followed by phosphorylation/elimination assisted decarboxylation results in IPP formation. Enzymes involved in these reactions are mevalonate kinase, 5-phosphomevalonate kinase, and 5-diphosphomevalonate decarboxylase, respectively (Tholl, 2015). Finally, just as in the

plastids an isomerase (partially) converts IPP to DMAPP. The rate-limiting step in the MVA pathway is catalysed by HMGR (Tholl and Lee, 2011).

The condensation of the IPP and DMAPP building blocks produced by the MEP and MVA pathways provides the prenyl diphosphate substrates such as C10 (geranyl diphosphate, GPP), formed by condensation of IPP and DMAPP through enzymatic activity of geranyl diphosphate synthase (GPS) or C15 (farnesyl diphosphate, FPP), formed from condensation of GPP and IPP by farnesyl diphosphate synthase (FPS) (Figure 3). GPP and FPP are the universal precursors of monoterpenoids (C10) and sesquiterpenoids (C15), respectively. Geranylgeranyl diphosphate synthase (GGPS) catalyses the condensation of FPP with IPP, which results in the formation of the C20 precursor of the diterpenes, geranylgeranyl diphosphate (GGPP), while dimerization of two FPP molecules and removal of the diphosphate groups through the activity of squalene synthase (SQS) results in biosynthesis of squalene (C30) (Blagg et al., 2002). Squalene mono-oxygenase or epoxidase adds an oxygen group to the squalene, resulting in production of 2,3-oxidosqualene, the precursor of triterpenoids (C30) as well as sterols and steroids in plants. Dimerization of two GGPP molecules and elimination of the two diphosphate groups by phytoene synthase (PS) results in the formation of a C40 compound, phytoene, the precursor of the tetraterpenoids or carotenoids (Figure 3).

Biosynthesis of monoterpenoids (C10). Monoterpenoids are C10 compounds derived from GPP. Monoterpenoids are well-known for their biological activity, but also for their strong odour and aromatic properties (Bakkali et al., 2008). These compounds are used for various applications such as fragrances, drinks, food additives, perfumes and cosmetics (Aharoni et al., 2005). Geraniol (isolated from rose flowers) (Mahboubi et al., 2011) and linalool (from coriander; Figure 1) (Mandal and Mandal, 2015) are two of most important monoterpenoids used in the flavour industry which is reaching to annual consumption of 5000 tons/year (Schwab et al., 2015). Monoterpenoids in plants are often induced upon biotic and/or abiotic stress conditions and they are supposed to possess properties to enable plants to deal with these stresses (Holopainen and Gershenzon, 2010). 1,8-Cineole, for example, is toxic to certain insects and is produced by *Artemisia annua* upon infestation by the root feeding insect, *Diuraphis noxia* (Tripathi et al., 2001). GPP synthase (GPS), the enzyme responsible for synthesis of the monoterpenoid precursor GPP, was first characterised from the essential oil glands of sage (Croteau and Purkett, 1989). Formation of monoterpenoids takes place through the activity of enzymes called monoterpene synthases (sometimes also called monoterpene cyclase if catalysing the formation of a cyclic monoterpene). Other functional groups may be added by the activity of tailoring enzymes such as oxygenases, methyltransferases, acetyltransferases and glycosyltransferases (Bian et al., 2017). In general, terpene synthases are closely related to the isoprenyl diphosphate synthases (IPS) enzymes (Humphrey and

Beale, 2006), however they are not able to bind the precursor IPP. The DDxxD motif for Mg^{2+} cation binding is conserved among all terpene synthases, which allows their identification (Bohlmann et al., 1998). A single monoterpene synthase often catalyses formation of several monoterpenoids from GPP. For example, a promiscuous monoterpene synthase enzyme, *CsTPS2FN*, isolated from *Cannabis sativa* encodes the formation of (+)- α -pinene, (+)- β -pinene, myrcene, (-)-limonene and β -phellandrene (Booth et al., 2017). An *Arabidopsis* terpene synthase, *AtTPS-Cin*, was shown to produce mainly 1,8-cineole, but also lower amounts of other monoterpenes such as (+)- α -thujene and (-)- α -pinene (Chen et al., 2004). As detailed above, it is generally assumed that the plastids are the major organelle for production of monoterpenoids and their substrate, GPP. Intriguingly, however, exchange of mitochondrial produced GDP to the plastids for production of monoterpenoids has been demonstrated (Dong et al., 2016).

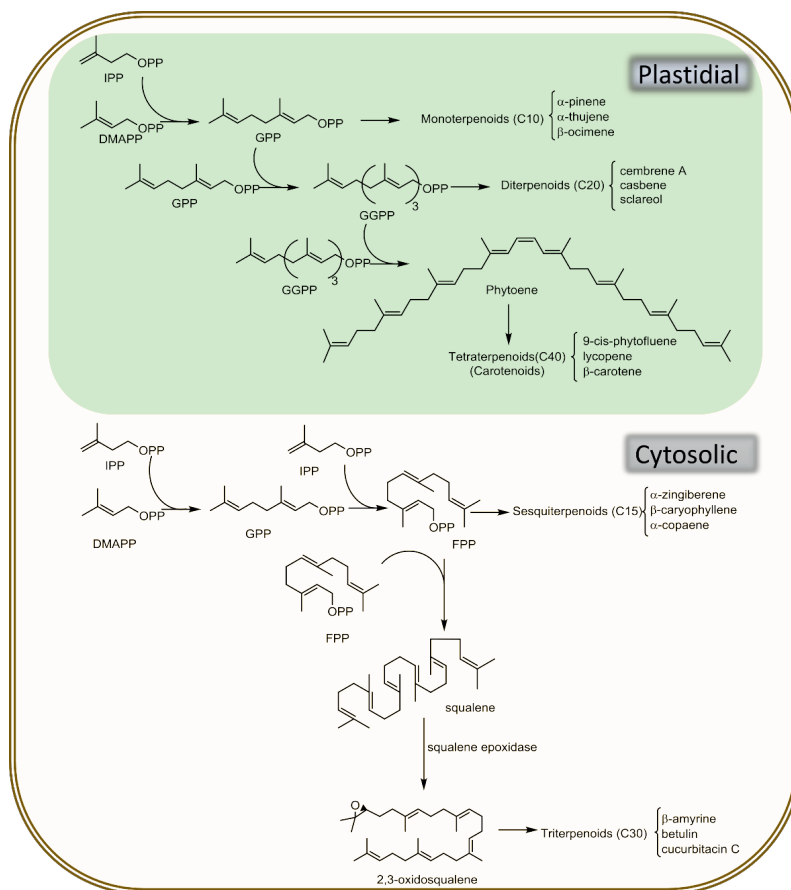


Figure 3. Terpenoid biosynthesis pathways. Biosynthesis of monoterpenoids (C10), diterpenoids (C20) and tetraterpenoids (C40) takes place in the plastids (in green) while the biosynthesis of sesquiterpenoids (C15) and triterpenoids (C30) is localized in the cytosol (in white).

Biosynthesis of sesquiterpenoids (C15). Sesquiterpenoids are produced from FPP, again through the catalytic activity of terpene synthases, in this case called sesquiterpene synthases. Sesquiterpenoids are often aromatic, and constituents of plant essential oils. The sesquiterpene β -caryophyllene (Figure 1) has been reported to be present in many plant species; it is the major essential oil component of basil (*Ocimum spp.*), oregano (*Origanum vulgare* L.) and rosemary (*Rosmarinus officinalis*) (Fidyt et al., 2016) (Figure 1) and, together with humulene, is the main sesquiterpene obtained from cannabis plants and responsible for its odor (Nissen et al., 2010). β -caryophyllene is widely used in frozen dairy, chewing gums and beverages (Flavoring and Association, 1997). Zingiberene, a sesquiterpene present in ginger (*Zingiber officinale*) is a spider mite repellent (Figure 1) (Maluf et al., 2001). Sesquiterpenes such as germacrene A, germacrene B, germacrene C, and germacrene D are found in plant species in the Asteracea such as feverfew (*Tanacetum parthenium*) and costus (*Saussurea lappa*) (Kraker et al., 2001). In chamomile (*Matricaria chamomila*) it was shown that sesquiterpene biosynthesis starts in the plastids with GPP that is exported to the cytosol where IPP is added (Adam et al., 1999). Interestingly, results from transient expression of a sesquiterpene synthase from feverfew (*Tanacetum parthenium*) show that addition of mitochondrial targeting to a sesquiterpene synthase will result in higher sesquiterpene biosynthesis, presumably because mitochondrial FPP is accessed (Kappers et al., 2005; van Herpen et al., 2010; Liu et al., 2011). Indeed, localization of one of the Arabidopsis FPP synthases in the mitochondria has been demonstrated (Cunillera et al., 1997). Protein localization studies using GFP fusions of *cis*-FPS, and santalene and bergamotene synthase (SBS) suggest that biosynthesis of these unusual sesquiterpenoids take place in the plastids using IPP and DMAPP from the MEP pathway (Sallaud et al., 2009).

Sesquiterpene lactones are a sub-class of the sesquiterpenoids with over 4000 different known structures. With a lactone functional group in its structure a sesquiterpene becomes a sesquiterpene lactone, which is reflected in the suffix 'olide' (e.g. costunolide, parthenolide, kauniolide, etc.). Sesquiterpene lactones are mainly colourless, bitter, compounds found mainly in plant species in the Asteracea (Picman, 1986). Their biological properties such as antibacterial (e.g. vernolide (Rabe et al., 2002)), antifungal (8 α -hydroxy-4-epi-sonchucarpolide (Skaltsa et al., 2000)), anticancer (e.g. parthenolide (Liu et al., 2014)) make them of interest for medical use. Lactucin and lactucopicrin, sesquiterpene lactones from chicory (*Cichorium intybus*) (Price et al., 1990; Kisiel and Zielińska, 2001), also show strong anti-cancer activity (Zhang et al., 2016) and sedative properties (Yuki et al., 2011). Sesquiterpene lactones are classified in six bicyclic or tricyclic classes named guaianolides, pseudoguaianolides, xanthanolides, eremophilanolides, eudesmanolides and germacranolides (Lepoittevin et al., 2009). Costunolide is the simplest germacranolide and is generally accepted to be the intermediate for biosynthesis of eudesmanolides and guaianolides (Kraker

et al., 2002). Formation of the lactone ring in costunolide is coupled with the activity of two cytochrome P450s, germacrene A oxidase and costunolide synthase (Nguyen et al., 2010; Liu et al., 2011). Costunolide may then serve as the precursor of the other germacranolides (e.g. parthenolide) and/or guaianolides (e.g. the main constituents of bitter compounds in chicory and endive). Further modification of sesquiterpene lactones is carried out by double bond reductases and glycosyl transferases (Zhang et al., 2008; Chadwick et al., 2013).

Biosynthesis of diterpenoids (C20). With more than 10,000 different natural plant derived structures, the diterpenoids are one of the most diverse classes of plant secondary metabolites (Zerbe et al., 2013). They are also part of plant primary metabolism as plant growth regulators such as the gibberellins are diterpenes (Martínez et al., 2016). Many diterpenoids have medicinal properties, such as taxol (Figure 1), which is isolated from the Pacific yew (*Taxus brevifolia*) (Goodman and Walsh, 2001), and is used for the treatment of ovarian and breast cancer (Suffness, 1995). Cafestol (Figure 1) and the structurally related kahweol are two diterpenes from *Coffea arabica* that induce apoptosis in malignant pleural mesothelioma (MPM) cancer cells (Lee et al., 2012). A valuable compound for the fragrance industry is *cis*-abienol which is an aromatic diterpene isolated from fir trees (*Abies balsamea*). *Cis*-abienol is an important oxygen containing diterpenoid serving as the precursor of Ambrox® in perfume formulations (Zerbe et al., 2012) and the major labdane type diterpenoid responsible for the fragrance of tobacco leaves. Biosynthesis of *cis*-abienol proceeds in two sequential steps. First a diterpene synthase converts GGPP to 8-hydroxy-copalyl diphosphate and then a kaurene synthase like enzyme converts the latter into *cis*-abienol by removing the diphosphate group (Sallaud et al., 2012).

Biosynthesis of triterpenoids (C30). This class of specialized metabolites constitutes more than 20,000 identified plant compounds so far (Thimmappa et al., 2014). Triterpenoids show a lot of diversity in plant families. Saponins, glycosylated triterpenoids, are, for example, found in *Quillaja saponaria* (a native Chilean tree) and *Camellia oleifera*. Saponins are used in detergents, shampoos and emulsifiers due to their foaming properties (Copaja et al., 2003; Chen et al., 2010). Many plants produce saponin type triterpenoids during normal growth (e.g. apple fruit peel, producing ursolic acid (Yamaguchi et al., 2008)), however their saponin levels strongly depends on plant species, organs (fruit, leaf etc.) and developmental stage (Moses et al., 2014). Butelin, isolated from the bark of *Butela* spp., is another natural triterpenoid which is used in cosmetic products such as hair conditioners (Patocka, 2003). Another important saponin type of triterpenoids found in *Glycyrrhiza glabra* (liquorice) is glycyrrhizin which is used as a natural sweetener. Steroids are also a class of triterpenoids which are harbouring an aliphatic side chain. Many triterpenoids are used to cure major diseases such as cancer and HIV. Celastrol, a triterpenoids isolated from *Tripterygium wilfordii* exhibits Tat inhibitory action (Narayan et al., 2011). 'Tat' is a virus encoded protein which is required for HIV

genome transcription. Cucurbitacin B is also a triterpenoid present in pumpkins, which is used for treatment of breast cancer (Dakeng et al., 2012). Triterpene synthases convert 2,3-oxidosqualene through a Chair-Boat-Chair (CBC) or the Chair-Chair-Chair (CCC) conformation into the different triterpene skeletons. An example of a triterpene synthase is β -amyrin synthase (Brendolise et al., 2011) responsible, for example, for β -amyrin biosynthesis in tomato (Figure 1). P450s and glycosyl transferases play an important role in further decoration of triterpenoids, for example, for the production of the triterpene glycoside glycyrrhizin.

Biosynthesis of tetraterpenoids (C40). The tetraterpenoids contain 750 different reported structures (Takaichi, 2011). The carotenoids (tetraterpenoids) are the most common natural pigments and also possess antioxidant properties. Carotenoids are industrially used as dyes and colorants, in the food industry (e.g. β -carotene, lutein and lycopene), as nutraceuticals and in the pharmaceutical industry, as well as in cosmetics (Zakynthinos and Varzakas, 2016) (Figure 1). Ten percent of the carotenoids display provitamin A activity (being a precursor for the formation of retinol, vitamin A) (Minguez-Mosquera and Hornero-Mendez, 1997; Xavier and Pérez-Gálvez, 2016). They are mostly present in photosynthetic organisms (Stange, 2016) and often are responsible for red, orange and yellow colours (Alcaíno et al., 2016). Carotenoids are essential and play a vital role in photosynthesis as they absorb the solar light and transfer that energy to the chlorophyll (Green and Parson, 2003). Carotenoid biosynthesis starts with the activity of phytoene synthase making pre-phytoene diphosphate (Dogbo et al., 1988). Phytoene synthase then converts pre-phytoene diphosphate to 15-*cis*-phytoene. Several other enzymes namely a desaturase and an isomerase are involved to produce *trans*-lycopene. Cyclisation is the next step; activity of an α -cyclase results in α -carotene biosynthesis, while a β -cyclase can convert *trans*-lycopene to β -carotene. Xanthophylls are oxygenated carotenoids involved in photoharvesting complexes in photosynthesis (Britton, 1995). Canthaxanthin is used as a food colouring agent while zeaxanthin was shown to have anticancer properties (Leyon et al., 1990; Yang et al., 1996). Another class of naturally occurring carotenoid-derived terpenoid type molecules are the strigolactones. Their biosynthesis starts with isomerization of β -carotene by D27 (Alder et al., 2012). Then a carotenoid cleavage (CCD7) cleaves the resulting 9-*cis*- β -carotene which then results in production of 9-*cis*- β -apo-10-carotenal and β -ionone (Bruno et al., 2014). Then, another carotenoid cleavage enzyme, CCD8, converts 9-*cis*- β -apo-10-carotenal into carlactone (Alder et al., 2012). This ubiquitous strigolactone precursor will be oxidised by a cytochrome P450, the MAX1 homologs, which results in the formation of carlactonoic acid or *ent*-2-*epi*-5-deoxystrigol (Abe et al., 2014; Zhang et al., 2014).

Heterologous production of terpenoids in plants and micro-organisms

As explained above, the terpenoids are very important compounds from medicinal, nutraceutical and nutritional point of view. However, commercialization of these compounds is often restricted due to their low concentrations in the plant and their high structural complexity which makes chemical synthesis approaches too costly (Misawa, 2011). In addition, some of the plant species that produce attractive molecules grow slowly, may have a low yield, are threatened by extinction, or are susceptible to environmental conditions. Several approaches have been followed in the last decades to overcome these limitations. In an approach called metabolic engineering, scientists use alternative organisms (expression platforms) to optimize production of these metabolites. Terpenoid production in microbial systems, for example, is an appealing approach. Rapid growth and regeneration (e.g. 1 to 3 days for *Escherichia coli* and *Saccharomyces cerevisiae*, respectively) and well established tools for transformation make them suitable organisms for metabolic engineering purposes. However, ectopic expression of plant derived genes (enzymes) in these microbial platforms comprise some limitations which needs to be solved for a successful engineering strategy. For example, neither *E. coli* nor *S. cerevisiae* contain plastids. Hence in order to prevent possible miss-folding of the enzymes in these platforms, removal of a possible plastid targeting signal is suggested (Schalk et al., 2012). The subcellular targeting strategy used by plants makes expressing cytochrome P450s in micro-organisms even more challenging. Expression of cytochrome P450s in *E. coli* generates huge amounts of reactive oxygen species (Zhou et al., 2015), hence disturbing the MEP pathway, unless a substrate later would be fed to the recombinant proteins. However, *S. cerevisiae* is a suitable expression platform for cytochrome P450s as it is a eukaryotic microorganism containing endoplasmic reticulum, the maturation and activity site of cytochrome P450s. Another advantage of yeast is the ability of *in vivo* recombination of DNA fragments, such that several DNA fragments (harbouring homologous flanking regions) can be recombined upon transformation into yeast in a so called transformation associated recombination (TAR) (Larionov et al., 1996; Siddiqui et al., 2012). Almost all required precursors for the biosynthesis of the different terpenoids are produced in yeast. Carotenoid and diterpenoid production in yeast is achieved often by overexpression of a GGPP synthase as yeast produces GGPP in small quantities. Carotenoids have been successfully expressed in yeast (Verwaal et al., 2007). Overexpression of genes such as HMGR, the rate limiting enzyme in the MVA pathway, has been shown to enhance the pool of precursor for the biosynthesis of, for example, sesquiterpenoids and triterpenoids. Alternatively, down regulation of competing pathways like sterol biosynthesis through down regulation of *ERG9* (squalene synthase) (Ro et al., 2006) are molecular strategies which are implemented for successful engineering programs. WAT11 yeast strain is an optimal yeast strain for expression of recombinant cytochrome P450s (Schoch et al.,

2003). Successful in-vivo production of artemisinic acid and costunolide in WAT11 yeast is reported by introduction of sesquiterpene synthases (amorphadiene synthase and germacrene A synthase) and P450s (amorphadiene oxidase (for artemisinic acid) and germacreneA oxidase plus costunolide synthase for costunolide production)) (Liu et al., 2011; Paddon et al., 2013).

Metabolic engineering can also be pursued in the plant species that is already making the attractive product by overexpression of biosynthetic pathway genes or downregulation of competing pathways. However, this homologous engineering - optimization and boosting of metabolic pathways in the original plant species - is sometimes difficult and time consuming. Hence, other *in planta* expression systems have been explored for heterologous expression of genes involved in the biosynthesis of secondary metabolites. Here we discuss a number of examples of such hosts that have been used for metabolic engineering and reconstruction of terpenoid biosynthesis pathways. Overexpression of taxadiene synthase, which converts GGDP to taxadiene - a precursor of the anti-cancer molecule taxol (Figure 1) - has been studied in *Nicotiana benthamiana*. Taxadiene was produced to an astonishing yield of 27 µg/g dry weight (Hasan et al., 2014). An example of successful reconstruction of a full biosynthetic pathway is the biosynthesis of parthenolide in *N. benthamiana*. The transient co-expression of germacrene A synthase (*GAS*), germacrene A oxidase (*GAO*), costunolide synthase (*COS*) and parthenolide synthase (*PTS*) yielded 1.4 mg/g fresh weight parthenolide in the leaves (Liu et al., 2014). Artemisinin was also successfully synthesized in *N. benthamiana*, by transient expression of five biosynthetic pathway genes, amorphadiene synthase (*ADS*), amorphadiene oxidase (*ADO*), alcohol dehydrogenase 1 (*ALDH1*), artemisinic aldehyde double-bond reductase (*DBR*) and aldehyde dehydrogenase 1 (*ALDH1*). Artemisinin production was enhanced upon co-expression of an ABCg-type transporter and lipid transfer proteins (Wang et al., 2016). *Physcomytrella patens* is another plant expression platform which recently has raised a lot of interest for metabolic engineering of valuable terpenoids. Novel and relatively easy transformation technology (King et al., 2016) has made this platform a suitable putative heterologous system for bulk production of terpenoids. Successful artemisinin production in *P. patens* was shown recently with a yield of 0.21 mg/g dry weight (Khairul Ikram et al., 2017). This yield was obtained upon co-expression of the same five biosynthesis pathway genes mentioned above, *ADS*, *ADO*, *ALDH1*, *DBR* and *ALDH1*. Yield in these heterologous production platforms is still quite low and could be due to toxicity due to inefficient efflux from the plant cells, also resulting in detoxification by conjugating activities, reducing the level of desired free products (Liu et al., 2011; Ting et al., 2013; Liu et al., 2014). However, some endogenous modifications by the heterologous hosts such as methylation, desaturation and esterification may result novel new-to-nature terpenoids which may enhance their bioactivities (Arendt et al., 2016). Such modifications

with beneficial effects may be used in combinatorial metabolic engineering for production of novel compounds. For example, coexpression of three carotenoid desaturases, a cyclase and one hydroxylase in *E. coli* in a combinatorial biosynthesis approach resulted in four novel carotenoids not identified in nature (Albrecht et al., 2000). A better knowledge of the natural site of biosynthesis and accumulation, the chemical properties of the terpenoids produced, and the mechanisms involved in their transport (from the biosynthesis site to the accumulation site) will provide novel solutions to be implemented in metabolic engineering programs, as we will discuss below.

Terpenoid transport and storage in plants

Many of the specialized metabolites discussed above are produced in plants to protect them from herbivores and pathogens. These compounds, due to their chemical reactivity, are often toxic for the plant itself as well, and often are translocated from their biosynthesis site to a storage compartment, where high levels can accumulate without harming the biosynthetic tissues. Plants have evolved several mechanisms to avoid this toxicity (Sirikantaramas et al., 2008). Transport to the vacuole or secretion to the extracellular space (trichomes, ducts, etc) are included in this. Trichomes are specialized plant structures differentiated from epidermal cells (Schellmann and Hulskamp, 2004). Sesquiterpene lactones in feverfew (e.g. costunolide and parthenolide) are biosynthesized in the glandular trichomes. However, these hydrophobic compounds are then secreted into the extracellular, subcuticular, space of the trichome. Besides trichomes, there are other intercellular secretory cavities for the storage of plant terpenoids. Secretory cavities of *Citrus* fruit peel are formed in early developmental stages and consist of many major essential oil constituents of *Citrus* fruits, such as, for example, limonene (Figure 1) (Voo et al., 2012). The other secretory structure found in more than 10% of flowering plants is the laticifer (Lange, 2015). These structures are elongated structures of which the cytoplasm changes to a substance called latex (Hagel et al., 2008). The sesquiterpene lactone taraxinic acid β -D-glucopyranosyl ester is one of the terpenoids stored in the laticifer (latex) of *Taraxacum officinale* (Huber et al., 2016). Other terpenoid molecules detected in laticifers are lactucin (*Lactuca sativa*), phorobol (a diterpenoid detected in *Euphorbia spp.*) and cardenolides in the *Apocynaceae*.

Many terpenoids (e.g. monoterpenoids and sesquiterpenoids) are classified as volatile organic compounds and it was suggested that they are emitted into the atmosphere through passive diffusion. However recently, through calculations it was concluded that extra agent(s) should play a role, as emission just by passive diffusion cannot explain the high rates of release into the atmosphere (Widhalm et al., 2015). Two major phenomena should be taken into account when trying to explain the mechanism of transport of terpenoids; (i) the chemical properties

of the terpenoids and (ii) the physio-chemical properties of the cell components over which the terpenoids need to be transported. The first barrier for secretion out of the cell is the plasma membrane. ATP-Binding Cassette type G (ABCG) transporters have been shown to facilitate transport of terpenoids over the cell membrane. Recently it was shown that emission of β -caryophyllene, a volatile sesquiterpenoid from *Artemisia annua* trichomes, is enhanced by an ABCG protein, called PDR3 (Pleiotropic Drug Resistance 3) (Fu et al., 2017). After passing the cell membrane, however, the hydrophobic β -caryophyllene must be transported over the cell wall, which is strongly hydrophilic, to allow emission into the atmosphere. For transport over the cell wall, it has been suggested that lipid transfer proteins (LTPs), which contain a hydrophobic pocket, play a role (Dudareva and Pichersky, 2006). Artemisinin, a sesquiterpene lactone present in the trichomes of *A. annua*, is a product of photo-oxidation of dihydroartemisinic acid. It was demonstrated that sequestration of artemisinic acid and dihydroartemisinic acid, both lipophilic molecules, to the extracellular space in the trichomes is enhanced by a putative interaction between an *A. annua* ABC transporter and LTPs (Wang et al., 2016). ABC transporters were also shown to be involved in diterpene transport in *Nicotiana plumbaginifolia* that produces the diterpenoid sclareol, which is secreted from the trichomes of this species and accumulates in crystalline epicuticular form on the leaf surface (Jasiński et al., 2001; Caissard et al., 2012).

Carotenoids, however, are biosynthesized and stored in the plastids. Plants have various types of plastids, classified according to their main function and morphology. Carotenoid biosynthesis takes place in almost all plastid types except the proplastids (Li et al., 2016). However, all different plastids are developed from proplastids. The main carotenoids involved in photosynthesis (lutein, β -carotene, violaxanthin, and neoxanthin) are stored in the thylakoids of chloroplasts (Domonkos et al., 2013). Often bright colours (red, yellow and orange) found in flowers, fruits and vegetables are due to the presence of carotenoids accumulated in chromoplasts (Li et al., 2016). The presence of xanthophylls neoxanthin and violaxanthin in tomato flowers (yellow colour) is due to their accumulation in chromoplasts (Galpaz et al., 2006). According to Yuan et al., chromoplasts can directly derive from proplastids during the ripening process of fruits (Yuan et al., 2015). However, chromoplasts can also derive from chloroplasts and amyloplasts. Chromoplasts are the accumulation site of the red coloured apocarotenoids in the stigma of saffron (Grilli Caiola and Canini, 2004). On the other hand, chromoplasts of carrot root (containing the orange β -carotene) are derived from leucoplasts (Kim et al., 2010). Size and type of plastids significantly affect accumulation and stability of carotenoids (Nisar et al., 2015). In this regard it was shown that carotenoids (e.g. lycopene) accumulation can be increased by increasing the number and/or volume of plastids in plants without changing the activity of the biosynthesis genes (Kolotilin et al., 2007).

Thesis outline

In my thesis I study the metabolic engineering of valuable pharmaceutical sesquiterpene lactones with several objectives: (i) to elucidate the biosynthetic pathway of sesquiterpene lactones and isolate the corresponding biosynthesis genes, (ii) to investigate the production of these compounds - as well as new compounds - through (combinatorial) metabolic engineering in heterologous hosts and (iii) to identify and characterise the molecular components involved in extracellular transport of sesquiterpene lactones.

For the selection of the relevant biosynthesis genes and transport related genes I used the *in-house* developed transcript databases of *A. annua* and feverfew. From these I selected candidate genes that I tested in two heterologous host organisms, *N. benthamiana* and *S. cerevisiae*. I used protein modelling to explain the activity of some of the encoded proteins, and used combinations of enzymes from different plant species to make chimeric pathways to produce novel sesquiterpenoids. For my studies on the involvement of LTP proteins in sesquiterpene transport I developed two novel assays to test LTP function: a sesquiterpene export assay and a sesquiterpene exclusion assay using transient expression of biosynthetic and (putative) transport genes in *N. benthamiana*.

In **Chapter 1** I describe the biosynthesis of different classes of terpenoids in plants. This chapter describes the background knowledge about the approaches by which metabolic engineering of plant derived terpenoids are accomplished. This chapter also introduces the limited knowledge on accumulation and transport of terpenoids at the onset of this thesis work and the questions related to extracellular transport of terpenoids that I addressed in my work. Feverfew contains both germacranolide and guanolide sesquiterpene lactones. The pathway towards the common germacranolide sesquiterpene lactones costunolide and parthenolide in feverfew was described before. In **Chapter 2** we describe the characterisation of an enzyme that is responsible for the conversion of the germacranolide costunolide to the guaianolide, kauniolide. The unusual mode of action of this enzyme, kauniolide synthase, is confirmed by modelling and in-vitro studies. In **Chapter 3** I describe a combinatorial metabolic engineering approach. In this chapter I took advantage of the substrate promiscuity of the enzyme Double Bond Reductase (DBR) from *A. annua*, to alter products from the feverfew sesquiterpene lactone biosynthesis pathway. This resulted in formation of several novel (for the reconstituted pathway in *N. benthamiana*) sesquiterpene lactones, with a reduced exocyclic double bond. Moreover, I demonstrated that different routes towards formation of these sesquiterpene lactones are possible and evaluated the efficiency of each of these pathways. In **Chapter 4** we studied the heterologous production of artemisinin in *N. benthamiana* and the need of transport to the apoplast to prevent reaction with conjugating enzymes. We describe the identification and molecular characterization of three LTPs and two PDR proteins from *A.*

annua that are involved in the accumulation of the artemisinin biosynthetic precursor, dihydroartemisinic acid, in the apoplast. In **Chapter 5** I continued the research on LTPs that was initiated with the *A. annua* LTPs, in Chapter 4, with LTPs from feverfew. The substrate specificity of these LTPs for sesquiterpene lactones of feverfew and their role in extracellular accumulation of costunolide and parthenolide is described. One of the functional and highly specific LTPs carries a GPI anchor attachment site and we demonstrate that this anchor site is essential for its function in substrate exclusion and substrate export from the cell. In **Chapter 6** I discuss my thesis and bring up a number of remaining questions. I describe the limitations and advantages of the metabolic engineering platforms that I used in my study, and discuss the limitations, advantages and possibilities of combinatorial metabolic engineering of feverfew dihydro-sesquiterpene lactones. Finally, I discuss the remaining questions related to the transport of (sesqui)terpenes and how the novel assays I developed for testing LTP activity can help to answer these.

References

- Abe S, Sado A, Tanaka K, Kisugi T, Asami K, Ota S, Kim HI, Yoneyama K, Xie X, Ohnishi T (2014) Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro. *Proceedings of the National Academy of Sciences* **111**: 18084-18089
- Adam K-P, Thiel R, Zapp J (1999) Incorporation of 1-[1-¹³C] deoxy-D-xylulose in chamomile sesquiterpenes. *Archives of biochemistry and biophysics* **369**: 127-132
- Aharoni A, Jongsma MA, Bouwmeester HJ (2005) Volatile science? Metabolic engineering of terpenoids in plants. *Trends in plant science* **10**: 594-602
- Albrecht M, Takaichi S, Steiger S, Wang Z-Y, Sandmann G (2000) Novel hydroxycarotenoids with improved antioxidative properties produced by gene combination in *Escherichia coli*. *Nature biotechnology* **18**: 843
- Alcaíno J, Baeza M, Cifuentes V (2016) Carotenoid Distribution in Nature. In *Carotenoids in Nature*. Springer, pp 3-33
- Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S (2012) The path from β -carotene to carlactone, a strigolactone-like plant hormone. *Science* **335**: 1348-1351
- Arendt P, Pollier J, Callewaert N, Goossens A (2016) Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. *The Plant Journal* **87**: 16-37
- Bakkali F, Averbeck S, Averbeck D, Idaomar M (2008) Biological effects of essential oils – A review. *Food and Chemical Toxicology* **46**: 446-475
- Bian G, Han Y, Hou A, Yuan Y, Liu X, Deng Z, Liu T (2017) Releasing the potential power of terpene synthases by a robust precursor supply platform. *Metabolic Engineering* **42**: 1-8
- Blagg BS, Jarstfer MB, Rogers DH, Poulter CD (2002) Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene. *J Am Chem Soc* **124**: 8846-8853
- Bleeker PM, Diergaarde PJ, Ament K, Schütz S, Johné B, Dijkink J, Hiemstra H, de Gelder R, de Both MT, Sabelis MW (2011) Tomato-produced 7-epizingiberene and R-curcumen act as repellents to whiteflies. *Phytochemistry* **72**: 68-73
- Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proceedings of the National Academy of Sciences* **95**: 4126-4133
- Booth JK, Page JE, Bohlmann J (2017) Terpene synthases from *Cannabis sativa*. *PLOS ONE* **12**: e0173911
- Brendolise C, Yauk YK, Eberhard ED, Wang M, Chagne D, Andre C, Greenwood DR, Beuning LL (2011) An unusual plant triterpene synthase with predominant alpha-amyrin-producing activity identified by characterizing oxidosqualene cyclases from *Malus x domestica*. *Febs j* **278**: 2485-2499
- Britton G (1995) Structure and properties of carotenoids in relation to function. *The FASEB Journal* **9**: 1551-1558

- Bruno M, Hofmann M, Vermathen M, Alder A, Beyer P, Al-Babili S** (2014) On the substrate- and stereospecificity of the plant carotenoid cleavage dioxygenase 7. *FEBS Letters* **588**: 1802-1807
- Byers KJ, Bradshaw H, Riffell JA** (2014) Three floral volatiles contribute to differential pollinator attraction in monkeyflowers (*Mimulus*). *Journal of Experimental Biology* **217**: 614-623
- Caissard J-C, Olivier T, Delbecq C, Palle S, Garry P-P, Audran A, Valot N, Moja S, Nicolé F, Magnard J-L, Legrand S, Baudino S, Jullien F** (2012) Extracellular Localization of the Diterpene Sclareol in Clary Sage (*Salvia sclarea* L., Lamiaceae). *PLOS ONE* **7**: e48253
- Cardoso C, Zhang Y, Jamil M, Hepworth J, Charnikhova T, Dimkpa SO, Meharg C, Wright MH, Liu J, Meng X** (2014) Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs. *Proceedings of the National Academy of Sciences* **111**: 2379-2384
- Chadwick M, Trewin H, Gawthrop F, Wagstaff C** (2013) Sesquiterpenoids Lactones: Benefits to Plants and People. *International Journal of Molecular Sciences* **14**: 12780-12805
- Chaturvedi R, Venables B, Petros RA, Nalam V, Li M, Wang X, Takemoto LJ, Shah J** (2012) An abietane diterpenoid is a potent activator of systemic acquired resistance. *The Plant Journal* **71**: 161-172
- Chen F, Ro DK, Petri J, Gershenzon J, Bohlmann J, Pichersky E, Tholl D** (2004) Characterization of a root-specific *Arabidopsis* terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. *Plant Physiol* **135**: 1956-1966
- Chen Y-F, Yang C-H, Chang M-S, Ciou Y-P, Huang Y-C** (2010) Foam properties and detergent abilities of the saponins from *Camellia oleifera*. *International journal of molecular sciences* **11**: 4417-4425
- Copaja SV, Blackburn C, Carmona R** (2003) Variation of saponin contents in *Quillaja saponica* Molina. *Wood Science and Technology* **37**: 103-108
- Croteau R, Purkett PT** (1989) Geranyl pyrophosphate synthase: characterization of the enzyme and evidence that this chain-length specific prenyltransferase is associated with monoterpene biosynthesis in sage (*Salvia officinalis*). *Archives of biochemistry and biophysics* **271**: 524-535
- Cunillera N, Boronat A, Ferrer A** (1997) The *Arabidopsis thaliana* FPS1 gene generates a novel mRNA that encodes a mitochondrial farnesyl-diphosphate synthase isoform. *Journal of Biological Chemistry* **272**: 15381-15388
- Dakeng S, Duangmano S, Jiratchariyakul W, Y UP, Bogler O, Patmasirawat P** (2012) Inhibition of Wnt signaling by cucurbitacin B in breast cancer cells: reduction of Wnt-associated proteins and reduced translocation of galectin-3-mediated beta-catenin to the nucleus. *J Cell Biochem* **113**: 49-60
- Dogbo O, Laferrière A, d'Harlingue A, Camara B** (1988) Carotenoid biosynthesis: Isolation and characterization of a bifunctional enzyme catalyzing the synthesis of phytoene. *Proceedings of the National Academy of Sciences* **85**: 7054-7058
- Domonkos I, Kis M, Gombos Z, Ughy B** (2013) Carotenoids, versatile components of oxygenic photosynthesis. *Progress in lipid research* **52**: 539-561
- Dong L, Jongedijk E, Bouwmeester H, Van Der Krol A** (2016) Monoterpene biosynthesis potential of plant subcellular compartments. *New Phytologist* **209**: 679-690
- Dudareva N, Pichersky E** (2006) *Biology of floral scent*. CRC Press
- Estévez JM, Cantero A, Reindl A, Reichler S, León P** (2001) 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *Journal of Biological Chemistry* **276**: 22901-22909
- Fidyt K, Fiedorowicz A, Strzadala L, Szumny A** (2016) β -caryophyllene and β -caryophyllene oxide—natural compounds of anticancer and analgesic properties. *Cancer Medicine* **5**: 3007-3017
- Flavoring, Association EM** (1997) FEMA Database: Beta-Caryophyllene (FEMA No. 2252). Washington, DC: Flavor and Extract Manufacturers Association
- Fu X, Shi P, He Q, Shen Q, Tang Y, Pan Q, Ma Y, Yan T, Chen M, Hao X, Liu P, Li L, Wang Y, Sun X, Tang K** (2017) AaPDR3, a PDR Transporter 3, Is Involved in Sesquiterpene beta-Caryophyllene Transport in *Artemisia annua*. *Front Plant Sci* **8**: 723
- Galpaz N, Ronen G, Khalfa Z, Zamir D, Hirschberg J** (2006) A Chromoplast-Specific Carotenoid Biosynthesis Pathway Is Revealed by Cloning of the Tomato *white-flower* Locus. *The Plant Cell* **18**: 1947-1960
- Goodman J, Walsh V** (2001) The story of Taxol. *Nature and Politics in the Pursuit of an Anti-Cancer Drug* **8**
- Green B, Parson WW** (2003) Light-harvesting antennas in photosynthesis, Vol 13. Springer Science & Business Media
- Grilli Caiola M, Canini A** (2004) Ultrastructure of chromoplasts and other plastids in *Crocus sativus* L.(Iridaceae). *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology* **138**: 43-52
- Hagel JM, Yeung EC, Facchini PJ** (2008) Got milk? The secret life of laticifers. *Trends in plant science* **13**: 631-639

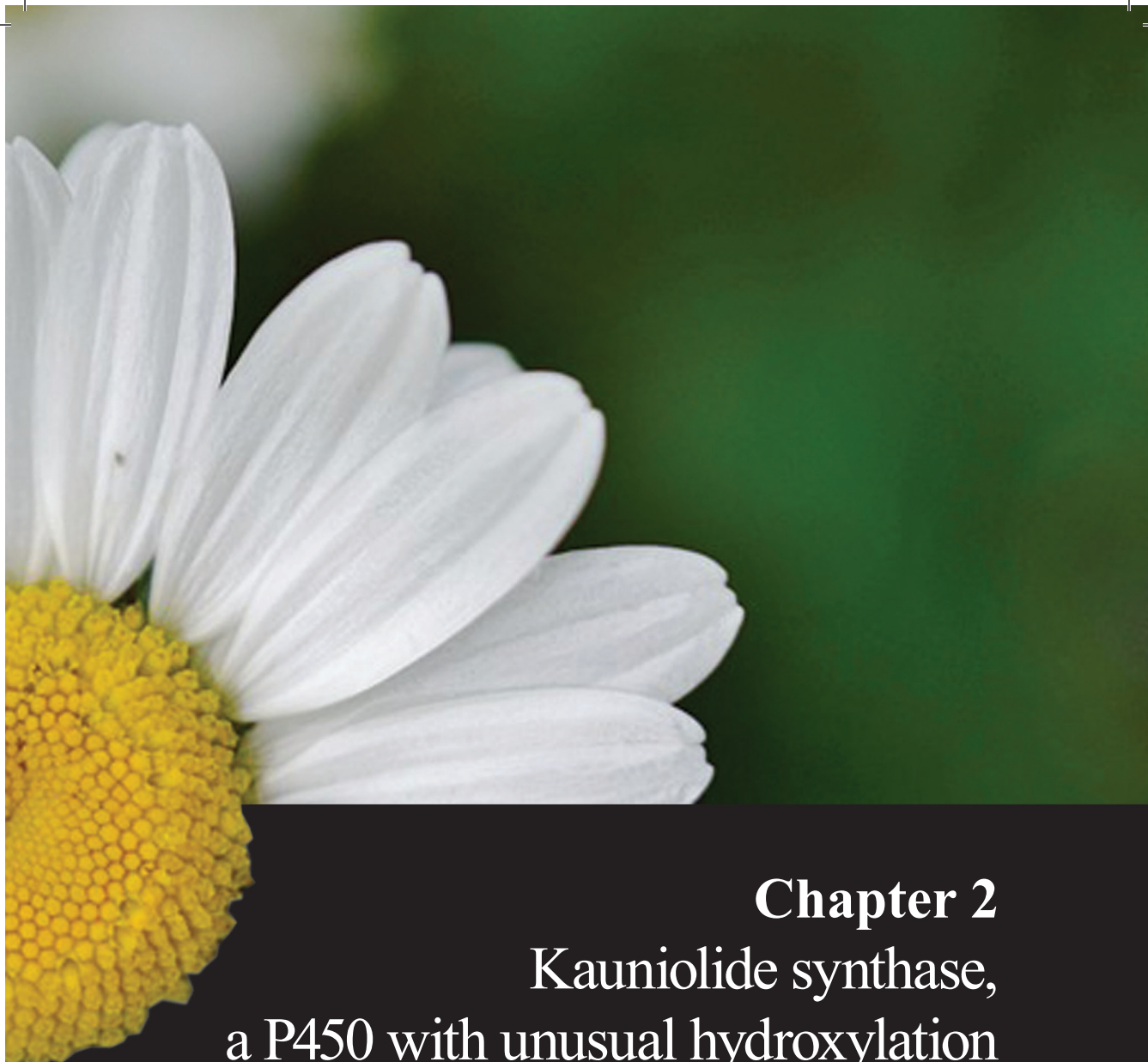
- Hasan MM, Kim H-S, Jeon J-H, Kim SH, Moon B, Song J-Y, Shim SH, Baek K-H** (2014) Metabolic engineering of *Nicotiana benthamiana* for the increased production of taxadiene. *Plant cell reports* **33**: 895-904
- Herrera MD, Rodriguez-Rodriguez R, Ruiz-Gutierrez V** (2006) Functional Properties of Pentacyclic Triterpenes Contained in. *Current Nutrition & Food Science* **2**: 45-49
- Holopainen JK, Gershenzon J** (2010) Multiple stress factors and the emission of plant VOCs. *Trends in Plant Science* **15**: 176-184
- Huang M, Sanchez-Moreiras AM, Abel C, Sohrabi R, Lee S, Gershenzon J, Tholl D** (2012) The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (E)- β -caryophyllene, is a defense against a bacterial pathogen. *New Phytologist* **193**: 997-1008
- Huber M, Epping J, Schulze Gronover C, Fricke J, Aziz Z, Brillatz T, Swyers M, Köllner TG, Vogel H, Hammerbacher A, Triebwasser-Freese D, Robert CAM, Verhoeven K, Preite V, Gershenzon J, Erb M** (2016) A Latex Metabolite Benefits Plant Fitness under Root Herbivore Attack. *PLOS Biology* **14**: e1002332
- Humphrey AJ, Beale MH** (2006) Terpenes. *Plant Secondary Metabolites: Occurrence, Structure and Role in the Human Diet*: 47-101
- 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 Online* **13**: 1095-1107
- Jassbi AR, Zamanizadehnajari S, Baldwin IT** (2010) 17-Hydroxygeranylinalool glycosides are major resistance traits of *Nicotiana obtusifolia* against attack from tobacco hornworm larvae. *Phytochemistry* **71**: 1115-1121
- Kappers IF, Aharoni A, van Herpen TWJM, Luckerhoff LLP, Dicke M, Bouwmeester HJ** (2005) Genetic engineering of terpenoid metabolism attracts bodyguards to *Arabidopsis*. *Science* **309**: 2070-2072
- Kemen AC, Honkanen S, Melton RE, Findlay KC, Mugford ST, Hayashi K, Haralampidis K, Rosser SJ, Osbourn A** (2014) Investigation of triterpene synthesis and regulation in oats reveals a role for β -amyrin in determining root epidermal cell patterning. *Proceedings of the National Academy of Sciences* **111**: 8679-8684
- Khairul Ikram NKB, Beyraghdar Kashkooli A, Peramuna AV, van der Krol AR, Bouwmeester H, Simonsen HT** (2017) Stable Production of the Antimalarial Drug Artemisinin in the Moss *Physcomitrella patens*. *Frontiers in Bioengineering and Biotechnology* **5**
- Kim JE, Rensing KH, Douglas CJ, Cheng KM** (2010) Chromoplasts ultrastructure and estimated carotene content in root secondary phloem of different carrot varieties. *Planta* **231**: 549-558
- King BC, Vavitsas K, Ikram NKBK, Schröder J, Scharff LB, Hamberger B, Jensen PE, Simonsen HT** (2016) In vivo assembly of DNA-fragments in the moss, *Physcomitrella patens*. *Scientific reports* **6**: 25030
- Kisiel W, Zielińska K** (2001) Guaianolides from *Cichorium intybus* and structure revision of *Cichorium* sesquiterpene lactones. *Phytochemistry* **57**: 523-527
- Kolotilin I, Koltai H, Tadmor Y, Bar-Or C, Reuveni M, Meir A, Nahon S, Shlomo H, Chen L, Levin I** (2007) Transcriptional profiling of high pigment-2dg tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiol* **145**: 389-401
- Kraker JWd, Franssen MCR, de Groot A, Shibata T, Bouwmeester HJ** (2001) Germacrenes from fresh costus roots. *Phytochemistry* **58**: 481-487
- Kraker JWd, Franssen MCR, Joerink M, de Groot A, Bouwmeester HJ** (2002) Biosynthesis of Costunolide, Dihydrocostunolide, and Leucodin. Demonstration of Cytochrome P450-Catalyzed Formation of the Lactone Ring Present in Sesquiterpene Lactones of Chicory. *Plant Physiology* **129**: 257-268
- Lange BM** (2015) The evolution of plant secretory structures and emergence of terpenoid chemical diversity. *Annual review of plant biology* **66**: 139-159
- Larionov V, Kouprina N, Graves J, Chen X, Korenberg JR, Resnick MA** (1996) Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proceedings of the National Academy of Sciences* **93**: 491-496
- Lee K-A, Chae J-I, Shim J-H** (2012) Natural diterpenes from coffee, cafestol and kahweol induce apoptosis through regulation of specificity protein 1 expression in human malignant pleural mesothelioma. *Journal of Biomedical Science* **19**: 60
- Lepoittevin JP, Berl V, Giménez-Arnau E** (2009) α -methylene- γ -butyrolactones: versatile skin bioactive natural products. *The Chemical Record* **9**: 258-270
- Leyon H, Ros AM, Nyberg S, Algvere P** (1990) Reversibility of canthaxanthin deposits within the retina. *Acta ophthalmologica* **68**: 607-611
- Li L, Yuan H, Zeng Y, Xu Q** (2016) Plastids and Carotenoid Accumulation. *In Carotenoids in Nature*. Springer, pp 273-293

- 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 Q, Manzano D, Tanic N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S, Bouwmeester H** (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metab Eng* **23**: 145-153
- Loreto F, Dicke M, SCHNITZLER JP, Turlings TC** (2014) Plant volatiles and the environment. *Plant, cell & environment* **37**: 1905-1908
- Mahboubi M, Kazempour N, Khamechian T, Fallah MH, Kermani MM** (2011) Chemical Composition and Antimicrobial Activity of *Rosa damascena* Mill Essential Oil. *Journal of Biologically Active Products from Nature* **1**: 19-26
- Maluf WR, Campos GA, das Graças Cardoso M** (2001) Relationships between trichome types and spider mite (*Tetranychus evansi*) repellence in tomatoes with respect to foliar zingiberene contents. *Euphytica* **121**: 73-80
- Mandal S, Mandal M** (2015) Coriander (*Coriandrum sativum* L.) essential oil: Chemistry and biological activity. *Asian Pacific Journal of Tropical Biomedicine* **5**: 421-428
- Martínez C, Espinosa-Ruiz A, Prat S** (2016) Gibberellins and plant vegetative growth. *Annual plant reviews* **49**: 285-322
- Minguez-Mosquera MI, Hornero-Mendez D** (1997) Changes in provitamin A during paprika processing. *Journal of food protection* **60**: 853-857
- Misawa N** (2011) Pathway engineering for functional isoprenoids. *Current opinion in biotechnology* **22**: 627-633
- Moses T, Papadopoulou KK, Osbourn A** (2014) Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Critical Reviews in Biochemistry and Molecular Biology* **49**: 439-462
- Narayan V, Kodihalli RC, Chiaro C, Cary D, Aggarwal BB, Henderson AJ, Prabhu KS** (2011) Celastrol Inhibits Tat-mediated Human Immunodeficiency Virus (HIV) Transcription and Replication. *Journal of molecular biology* **410**: 972-983
- Nguyen DT, Göpfert JC, Ikezawa N, MacNevin G, Kathiresan M, Conrad J, Spring O, Ro D-K** (2010) Biochemical Conservation and Evolution of Germacrene A Oxidase in Asteraceae. *The Journal of Biological Chemistry* **285**: 16588-16598
- Nisar N, Li L, Lu S, Khin Nay C, Pogson Barry J** (2015) Carotenoid Metabolism in Plants. *Molecular Plant* **8**: 68-82
- Nissen L, Zatta A, Stefanini I, Grandi S, Sgorbati B, Biavati B, Monti A** (2010) Characterization and antimicrobial activity of essential oils of industrial hemp varieties (*Cannabis sativa* L.). *Fitoterapia* **81**: 413-419
- Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell M, Tai A, Main A, Eng D** (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**: 528
- Patocka J** (2003) Biologically active pentacyclic triterpenes and their current medicine signification. *J Appl Biomed* **1**: 7-12
- Picman AK** (1986) Biological activities of sesquiterpene lactones. *Biochemical Systematics and Ecology* **14**: 255-281
- Price KR, Dupont MS, Shepherd R, Chan HWS, Fenwick GR** (1990) Relationship between the chemical and sensory properties of exotic salad crops—coloured lettuce (*Lactuca sativa*) and chicory (*Cichorium intybus*). *Journal of the Science of Food and Agriculture* **53**: 185-192
- Rabe T, Mullholland D, van Staden J** (2002) Isolation and identification of antibacterial compounds from *Vernonia colorata* leaves. *Journal of Ethnopharmacology* **80**: 91-94
- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J** (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940-943
- Sallaud C, Giacalone C, Töpfer R, Goepfert S, Bakaher N, Rösti S, Tissier A** (2012) Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. *The Plant Journal* **72**: 1-17
- Sallaud C, Rontein D, Onillon S, Jabès F, Duffé P, Giacalone C, Thoraval S, Escoffier C, Herbette G, Leonhardt N, Causse M, Tissier A** (2009) A Novel Pathway for Sesquiterpene Biosynthesis from Z,Z-Farnesyl Pyrophosphate in the Wild Tomato *Solanum habrochaites*. *The Plant Cell* **21**: 301-317
- Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L** (2012) Toward a biosynthetic route to sclareol and amber odorants. *Journal of the American Chemical Society* **134**: 18900-18903

- Schellmann S, Hulskamp M** (2004) Epidermal differentiation: trichomes in Arabidopsis as a model system. *International Journal of Developmental Biology* **49**: 579-584
- Schoch GA, Attias R, Belghazi M, Dansette PM, Werck-Reichhart D** (2003) Engineering of a Water-Soluble Plant Cytochrome P450, CYP73A1, and NMR-Based Orientation of Natural and Alternate Substrates in the Active Site. *Plant Physiology* **133**: 1198-1208
- Schwab W, Fischer TC, Giri A, Wüst M** (2015) Potential applications of glucosyltransferases in terpene glucoside production: impacts on the use of aroma and fragrance. *Applied microbiology and biotechnology* **99**: 165-174
- Siddiqui MS, Thodey K, Trenchard I, Smolke CD** (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS yeast research* **12**: 144-170
- Sirikantaramas S, Yamazaki M, Saito K** (2008) Mechanisms of resistance to self-produced toxic secondary metabolites in plants. *Phytochemistry Reviews* **7**: 467
- Skaltsa H, Lazari D, Panagouleas C, Georgiadou E, Garcia B, Sokovic M** (2000) Sesquiterpene lactones from *Centaurea thessala* and *Centaurea attica*. Antifungal activity. *Phytochemistry* **55**: 903-908
- Springob K, Kutchan TM** (2009) Introduction to the different classes of natural products. *In Plant-Derived Natural Products*. Springer, pp 3-50
- Stange C** (2016) Carotenoids in nature: biosynthesis, regulation and function, Vol 79. Springer
- Suffness M** (1995) *Taxol: science and applications*, Vol 22. CRC press
- Takaichi S** (2011) Carotenoids in algae: distributions, biosyntheses and functions. *Marine drugs* **9**: 1101-1118
- Thimmappa R, Geisler K, Louveau T, O'Maille P, Osbourn A** (2014) Triterpene biosynthesis in plants. *Annual Review of Plant Biology* **65**: 225-257
- Tholl D** (2015) Biosynthesis and biological functions of terpenoids in plants. *Adv Biochem Eng Biotechnol* **148**: 63-106
- Tholl D, Lee S** (2011) Terpene specialized metabolism in Arabidopsis thaliana. *The Arabidopsis Book*: e0143
- Ting HM, Wang B, Rydén AM, Woittiez L, Herpen T, Verstappen FW, Ruyter-Spira C, Beekwilder J, Bouwmeester HJ, 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
- Tripathi AK, Prajapati V, Aggarwal KK, Kumar S** (2001) Toxicity, Feeding Deterrence, and Effect of Activity of 1,8-Cineole from *Artemisia annua* on Progeny Production of *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Journal of Economic Entomology* **94**: 979-983
- van Herpen TWJM, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J** (2010) *Nicotiana benthamiana* as a Production Platform for Artemisinin Precursors. *PLOS ONE* **5**: e14222
- Verwaal R, Wang J, Meijnen J-P, Visser H, Sandmann G, van den Berg JA, van Ooyen AJJ** (2007) High-Level Production of Beta-Carotene in *Saccharomyces cerevisiae* by Successive Transformation with Carotenogenic Genes from *Xanthophyllomyces dendrorhous*. *Applied and Environmental Microbiology* **73**: 4342-4350
- Voo SS, Grimes HD, Lange BM** (2012) Assessing the Biosynthetic Capabilities of Secretory Glands in Citrus Peel. *Plant Physiology* **159**: 81-94
- Wang B, Kashkooli AB, Sallets A, Ting H-M, de Ruijter NCA, Olofsson L, Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR** (2016) Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metabolic Engineering* **38**: 159-169
- Widhalm JR, Jaini R, Morgan JA, Dudareva N** (2015) Rethinking how volatiles are released from plant cells. *Trends in Plant Science* **20**: 545-550
- Wink M** (2010) Introduction. *In Annual Plant Reviews Volume 39: Functions and Biotechnology of Plant Secondary Metabolites*. Wiley-Blackwell, pp 1-20
- Xavier AAO, Pérez-Gálvez A** (2016) Carotenoids as a Source of Antioxidants in the Diet. *In Carotenoids in Nature*. Springer, pp 359-375
- Yamaguchi H, Noshita T, Kidachi Y, Umetsu H, Hayashi M, Komiyama K, Funayama S, Ryoyama K** (2008) Isolation of Ursolic Acid from Apple Peels and Its Specific Efficacy as a Potent Antitumor Agent. *Journal of Health Science* **54**: 654-660
- Yang Y, Huang CY, Peng S, Li J** (1996) Carotenoid analysis of several dark-green leafy vegetables associated with a lower risk of cancers. *Biomedical and environmental sciences: BES* **9**: 386-392
- Yuan H, Zhang J, Divyashree Nageswaran LL** (2015) Carotenoid metabolism and regulation in horticultural crops. *Horticulture Research* **2**: 15036
- Yuki K, Soriano SG, Shimaoka M** (2011) Sedative drug modulates T-cell and lymphocyte function-associated antigen-1 function. *Anesth Analg* **112**: 830-838
- Zakynthinos G, Varzakas T** (2016) Carotenoids: From plants to food industry. *Current Research in Nutrition and Food Science* **4**: 38-51

- Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Britton R, Bohlmann J** (2012) Bifunctional cis-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *Journal of Biological Chemistry* **287**: 12121-12131
- Zerbe P, Hamberger B, Yuen MMS, Chiang A, Sandhu HK, Madilao LL, Nguyen A, Hamberger B, Bach SS, Bohlmann J** (2013) Gene Discovery of Modular Diterpene Metabolism in Nonmodel Systems. *Plant Physiology* **162**: 1073-1091
- Zhang F-H, Yan Y-L, Wang Y, Liu Z** (2016) Lactucin induces potent anti-cancer effects in HL-60 human leukemia cancer cells by inducing apoptosis and sub-G1 cell cycle arrest. *Bangladesh Journal of Pharmacology* **11**: 478-484
- Zhang Y, Teoh KH, Reed DW, Maes L, Goossens A, Olson DJ, Ross AR, Covello PS** (2008) The molecular cloning of artemisinic aldehyde Δ^{11} (13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *Journal of Biological Chemistry* **283**: 21501-21508
- Zhang Y, Van Dijk AD, Scaffidi A, Flematti GR, Hofmann M, Charnikhova T, Verstappen F, Hepworth J, Van Der Krol S, Leyser O** (2014) Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nature chemical biology* **10**: 1028-1033
- Zhou K, Qiao K, Edgar S, Stephanopoulos G** (2015) Distributing a metabolic pathway among a microbial consortium enhances production of natural products. **33**: 377-383





Chapter 2

Kauniolide synthase, a P450 with unusual hydroxylation and cyclization-elimination activity

Qing Liu[†], Arman Beyraghdar Kashkooli[†], David Manzano, Irini Pateraki,
Lea Richard, Pim Kolkman, Maria Fátima Lucas, Victor Guallar, Ric de Vos,
Maurice C.R. Franssen, Alexander van der Krol, Harro Bouwmeester

[†]: These authors contributed equally to this paper

Accepted in Nature Communications

Abstract

Guaianolides are an important class of sesquiterpene lactones with unique biological and pharmaceutical properties. Supposedly these compounds are derived from germacranolides, but for years no progress has been made in the elucidation of their biosynthesis pathway which would require some kind of cyclisation mechanism. Here we demonstrate the isolation and characterization of cytochrome P450 enzyme from *Tanacetum parthenium*, kauniolide synthase (TpKLS), which can perform this reaction by a sequence of hydroxylation, water elimination, cyclisation and regioselective deprotonation. TpKLS catalyses the formation of the guaianolide kauniolide from the germacranolide substrate costunolide. The unique mechanism of action of this special enzyme was elucidated through *in-silico* modelling and docking experiments. Unlike most cytochrome P450s, TpKLS combines stereoselective hydroxylation of costunolide at the C3 position, with three other reactions including cyclisation. Moreover, the full kauniolide biosynthesis pathway was reconstructed in the heterologous hosts *Nicotiana benthamiana* and yeast, paving the way for metabolic engineering and biochemical production of guaianolide-type sesquiterpene lactones.

Keywords: guaianolide biosynthesis, P450, stereoselective hydroxylation, *in-silico* modelling, pathway reconstitution

Introduction

Sesquiterpene lactones are C₁₅ terpenoids and constitute a major class of plant secondary metabolites with diverse chemical structures. They are present in plant species in the *Acanthaceae*, *Apiaceae* and *Asteraceae* (Chaturvedi, 2011) with over 4000 different structures reported so far (de Kraker et al., 2002). Sesquiterpene lactones are classified in six bicyclic or tricyclic classes named guaianolides, pseudoguaianolides, xanthanolides, eremophilanolides, eudesmanolides and germacranolides (Lepoittevin et al., 2009). All classes of natural occurring sesquiterpene lactones contain an α -methylene- γ -butyrolactone moiety which accounts for their biological activity against cancer and inflammation (Lepoittevin et al., 2009). Guaianolides are defined by their special 5-7-5 tricyclic structure (Fischer et al., 1979) and these compounds are of interest because of their biological activity against prostate cancer (Kim et al., 2012), anti-mycobacteria activity (Cantrell et al., 1998), inhibition of parasite *Trypanosoma cruzi* growth (Cogo et al., 2012) and human antioxidant response element activation at low concentrations (Fischedick et al., 2012). The biosynthetic pathways of a number of important germacranolide sesquiterpene lactones have been elucidated. For instance, the biosynthesis of costunolide and parthenolide has been elucidated by isolation and characterization of the germacrene A synthase (GAS) (Bouwmeester et al., 2002; Majdi et al., 2011), germacrene A oxidase (GAO) (Nguyen et al., 2010), costunolide synthase (COS) (Liu et al., 2011) and parthenolide synthase (PTS) (Liu et al., 2014). In contrast, at present enzymes involved in guaianolide formations have not been identified, preventing metabolic engineering of these classes of sesquiterpene lactones. There are some hints about the biosynthesis of guaianolides: in the family of *Asteraceae* and *Apiaceae* (e.g. feverfew (*Tanacetum parthenium*), chicory (*Cichorium intybus*), thapsia (*Thapsia garganica*)), two classes of sesquiterpene lactones, germacranolides and guaianolides, often occur together (Simonsen et al., 2013), suggesting that the biosynthesis of guaianolides may be related to that of germacranolides. Moreover, root extracts of chicory can convert costunolide to the guaianolide leucodin (de Kraker et al., 2002). This bioconversion could be inhibited by a range of chemical P450 enzyme activity inhibitors (de Kraker et al., 2002), suggesting that one or more P450 enzyme are involved in the conversion of costunolide to leucodin (de Kraker et al., 2002). On the other hand, chemical synthesis showed that one guaianolide, kauniolide, could be formed from parthenolide (Fig. 1a) (Zhai et al., 2012). In plants, parthenolide is produced from costunolide through the action of a P450 monooxygenase (Liu et al., 2014). This implies that biosynthesis of guaianolides could proceed through parthenolide or similar epoxides (Piet et al., 1995). However, according to Piet *et al.* (Piet et al., 1996) enzyme mediated cyclisation of germacranolides could also start with allylic C3 hydroxylation of costunolide, followed by protonation of the hydroxyl group, cyclisation and deprotonation which would result in the production of kauniolide, the basic guaianolide

backbone (Fig. 1b). Indeed, a cytochrome P450 from feverfew encoding an enzyme capable of C3- β hydroxylation of either costunolide or parthenolide has been previously reported (Liu et al., 2014). Neither 3 β -hydroxycostunolide nor 3 β -hydroxyparthenolide, however, could undergo a further cyclisation towards kauniolide, suggesting that a novel enzyme with cyclase activity is required to catalyse this step (Fig. 1a and 1b).

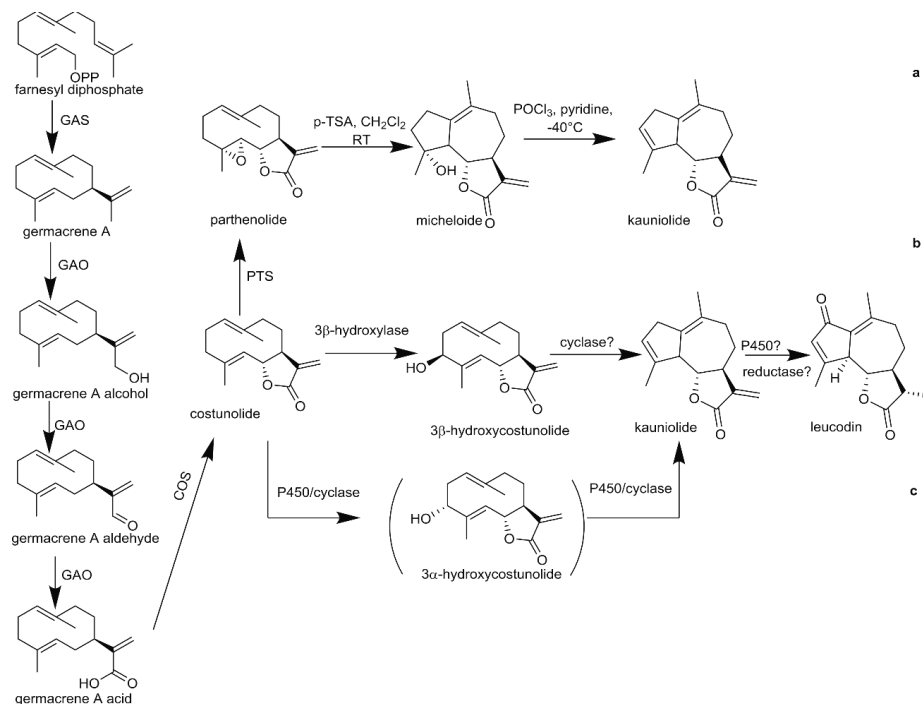


Figure 1. Proposed (bio)synthetic pathways for kauniolide. (a) As reported by chemical biomimetic synthesis (Piet et al., 1996); (b) as proposed as part of the biosynthetic pathway to leucodin in chicory (de Kraker et al., 2002); (c) as identified in the present study. GAS=germacrene A synthase, GAO= germacrene A oxidase, COS= costunolide synthase, PTS= parthenolide synthase. p-TSA= p-toluene sulfonic acid, CH₂Cl₂= dichloromethane, RT= room temperature, POCl₃= phosphoryl chloride.

Here we have identified a group of 10 co-expressed P450 genes of which 5 genes were characterized as part of the costunolide and/or parthenolide biosynthesis pathways (Liu et al., 2011; Majdi et al., 2011; Liu et al., 2014). Thus, the other 5 P450 genes of this group may also be part of this or related pathways. Therefore we expressed these genes in yeast to determine whether any of these P450 enzymes can act on costunolide or parthenolide. Unexpectedly, in these assays we identified a Kauniolide Synthase (KLS), a single P450 enzyme able not only to form hydroxylation activity but also water elimination coupled to cyclization and regioselective deprotonation leading to the synthesis of kauniolide. This shows that kauniolide biosynthesis does not require a cyclase as suggested before (Fig. 1a and 1b), but instead proceeds through a direct conversion of the germacranolide costunolide into the

guaianolide kauniolide (Fig. 1c). The full kauniolide biosynthesis pathway was successfully reconstructed in the heterologous hosts *Nicotiana benthamiana* and in yeast.

Modelling helped to elucidate the putative enzymatic steps catalysed by KLS. KLS thus establishes a novel enzyme class in the CYP71 family, having classical cytochrome P450 hydroxylation activity, combined with water elimination/cyclisation/deprotonation activity. Elucidation of this enzymatic step marks kauniolide as the first committed intermediate of the guaianolides and opens up opportunities for metabolic engineering of this pharmaceutically attractive class of sesquiterpene lactones.

Results

Screen for *Tanacetum parthenium* CYP71-family enzymes able to modifying costunolide and/or parthenolide

In feverfew extracts, multiple costunolide and parthenolide derived products are detected (Fischedick et al., 2012). Previously it was shown that feverfew P450s acting on germacrene A, germacrene A acid, costunolide and parthenolide all belong to the CYP71 family (Nguyen et al., 2010; Ikezawa et al., 2011; Liu et al., 2011; Majdi et al., 2011; Liu et al., 2014). Moreover, the genes encoding the enzymes of the parthenolide biosynthesis pathway (*TpGAS* (Liu et al., 2011), *TpGAO* (Liu et al., 2011), *TpCOS* (Liu et al., 2011) and *TpPTS* (Hullaert et al., 2014)) show a distinct expression profile during feverfew ovary development, matching with the accumulation profile of parthenolide (Liu et al., 2011; Majdi et al., 2011; Liu et al., 2014) (Supplementary Fig. 1a). The costunolide 3 β -hydroxylase (Hullaert et al., 2014) though, displays a slightly different expression profile (Supplementary Fig. 1a and 1b). For our screen we therefore targeted CYP71 P450s with an expression profile similar to those of the parthenolide biosynthesis pathway genes. P450 sequences were obtained by 454 sequencing of a trichome enriched cDNA library (www.terpmed.eu). Expression profiling was done by *Illumina* RNA sequencing of six different stages of feverfew ovary development. In total 59 P450s of the CYP71 family were detected of which 27 showed differential expression during ovary development (Supplementary Fig. 1a). Cluster analysis of these 27 candidate P450s with the parthenolide biosynthesis pathway genes resulted in a further selection of five uncharacterized P450s for which expression in ovaries shows clustering with those of the parthenolide biosynthesis pathway genes (Supplementary Fig. 1a). For four of these uncharacterized P450s (*Tp9025*, *Tp4149*, *Tp8879* and *Tp8886*), their expression profile was independently validated by RT-qPCR, confirming the RNA-sequencing data analysis in that they display an expression profile that is similar to the *bona fide* pathway genes (Supplementary Fig. 1b). For those candidates not represented in the transcriptome by full length cDNAs, RACE-PCR experiments were performed in order to obtain the full sequences.

Subsequently, the full-length coding sequences were cloned into yeast expression vectors for functional characterization of the encoding enzymes activity (see methods). In this paper we describe the further characterization of four P450s, *Tp8886*, *Tp8879*, *Tp4149* and *Tp9025*.

Functional characterization of selected Cyp71s in yeast

Candidate P450 cDNAs were expressed in yeast WAT11²⁰ strain and microsomes of the transgenic yeast were isolated for in vitro testing of the enzymatic activity towards costunolide or parthenolide. Enzymes encoded by *Tp4149* and *Tp9025* did not show any catalytic activity on parthenolide or costunolide and their substrate specificity could not be determined. In contrast, the enzyme encoded by *Tp8879* produced a product with mass $[M+H]^+=231.1378$ from costunolide as shown by LC-Orbitrap-FTMS analysis, although the peak in the chromatogram was very small. The mass reduction by 2.015 D of this novel product compared to the mass of costunolide ($[M+H]^+=233.1524$) signifies the loss of two protons, which suggests introduction of a C-double bond in costunolide or a ring closure resulting in a tricyclic sesquiterpene lactone. The detection of sesquiterpene lactones with an exocyclic double bond in LC-Orbitrap-FTMS can sometimes be improved by conjugation to glutathione or cysteine (Liu et al., 2011). This conjugation occurs non-enzymatically, is irreversible and in case of cysteine adds an exact mass of $[M]=121.01464$ and results in a shorter retention time (RT) (Hullaert et al., 2014). The reaction product of the enzyme encoded by *Tp8879* using costunolide as substrate was therefore incubated with cysteine. Analysis of the cysteine-conjugates revealed two new peaks at RT=10.07 min ($[M+H]^+=370.16827$) and RT=21.41 min ($[M+H]^+=352.1577$). The product eluting at RT=10.07 was identified as 3 β -hydroxycostunolide-cysteine (Fig. 2d and 2e) according to Qing et al., (2014) (Liu et al., 2014). More interestingly, the mass, fragmentation spectrum and RT of the product eluting at RT=21.41 min matched that of a kaunolide-cysteine standard (Fig. 2a and 2b). Indeed, co-injection of kaunolide-cysteine and the enzyme products incubated with cysteine showed co-elution, showing that the enzyme encoded by *Tp8879* catalyses ring closure resulting in kaunolide formation from costunolide (Supplementary Fig. 2). Moreover, the peak intensity of kaunolide-cysteine was 6.7 times higher than that of 3 β -hydroxycostunolide-cysteine, indicating that 3 β -hydroxycostunolide is a minor product (Fig. 2d and 2e). We therefore called *Tp8879* Kaunolide Synthase (*TpKLS*). When microsomes expressing *TpKLS* were incubated with parthenolide, novel peaks were not detected, not even after incubation of the enzyme products with cysteine. This suggests that costunolide is the natural substrate for *TpKLS* and that *TpKLS* is the first committed enzyme in the biosynthesis towards guaianolides, revealing for the first time that a guaianolide-type sesquiterpene lactone (like kaunolide) is derived from a germacranolide-type sesquiterpene lactones (i.e. costunolide).

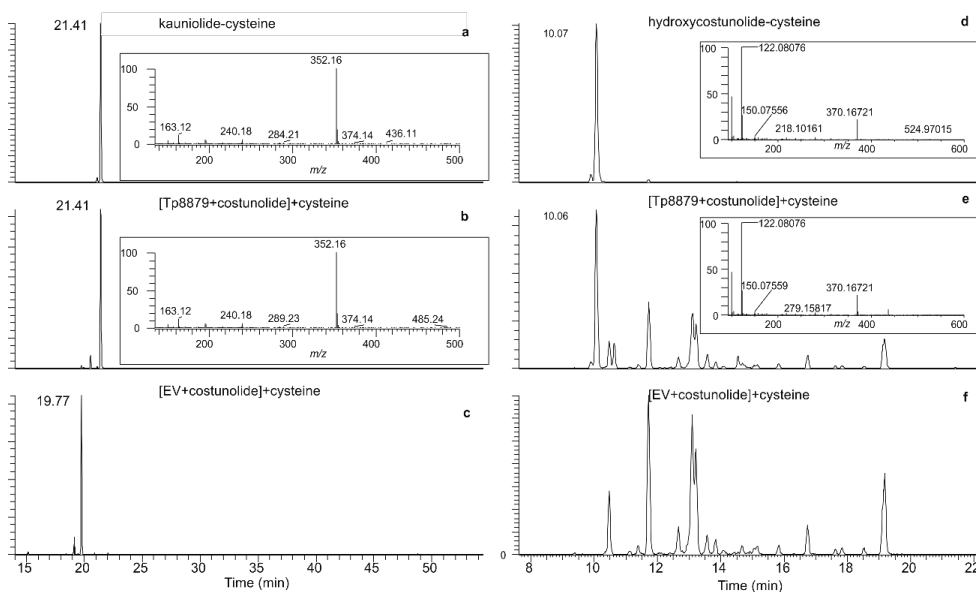


Figure 2. Identification of kauniolide-cysteine in an assay with yeast microsomes using LC-Orbitrap-FTMS. (a) Incubation of kauniolide with cysteine (b) Microsomes expressing *Tp8879* (*TpKLS*) fed with costunolide and incubated with cysteine. (c) Negative control where no kauniolide cysteine was formed by feeding costunolide to microsomes expressing an empty vector (EV) and incubation with cysteine. (d) Formation of 3 β -hydroxycostunolide-cysteine by incubating 3 β -hydroxycostunolide with cysteine. (e) LC-Orbitrap-FTMS chromatograms at $m/z=370.16827$ (10 ppm, positive ionization mode, mass for 3 β -hydroxycostunolide-cysteine) of 3 β -hydroxycostunolide-cysteine formed in the enzymatic assay of feeding costunolide to *TpKLS* and incubated with cysteine. (f) Negative control where no 3 β -hydroxycostunolide-cysteine was formed by feeding costunolide to EV and incubation with cysteine.

The production of 3 β -hydroxycostunolide, though as a minor product, from feeding costunolide to *TpKLS* suggests that this enzyme can perform hydroxylation of costunolide at the C3 position. Moreover, 3-hydroxycostunolide has been suggested to be an intermediate in kauniolide biosynthesis (Piet et al., 1996). We have previously identified a P450 in feverfew which can hydroxylate costunolide at C3 (*Tp3 β -hydroxylase* (Hullaert et al., 2014)), but does not produce kauniolide. To assess if 3 β -hydroxycostunolide is an intermediate in the production of kauniolide we incubated *TpKLS* with 3 β -hydroxycostunolide. This did not result in conversion into kauniolide, indicating that 3 β -hydroxycostunolide cannot serve as a substrate for *TpKLS* (Supplementary Fig. 3). The product of the enzyme encoded by *Tp8886* enzyme could not be identified due to its low abundance.

***In-silico* docking of costunolide in *TpKLS* favors costunolide C3 α -hydroxylation activity**

To get further insight into the mechanism by which kauniolide is formed by *TpKLS*, we used *in silico* modelling of *TpKLS*, followed by costunolide substrate docking experiments. As validation for the modelling approach we used two other P450 enzymes that use costunolide

as substrate and for which the site of costunolide oxidation is known, i.e. epoxidation of C4-C5 by TpPTS and hydroxylation of C3 by costunolide 3 β -hydroxylase (Liu et al., 2011; Liu et al., 2014). The protein structure models of TpPTS, costunolide 3 β -hydroxylase and TpKLS were generated using the crystal structure of *Homo sapiens* P450 2C9 (Reynald et al., 2012) (Protein Data Bank (PDB) code 4GQS), *Oryctolagus cuniculus* P450 2C5 (Wester et al., 2003) (PDB code 1NR6) and *Homo sapiens* dual specificity protein phosphatase 13 (PDB code 2PQ5_A), respectively (Supplementary Table 2). Substrate-docking of costunolide was simulated in PELE for each of the three modelled enzyme structures (Borrelli et al., 2005; Madadkar-Sobhani and Guallar, 2013). Two models were used: (i) a free simulation where the substrate has no constraints, and (ii) simulations with a defined constraint, where the substrate was not allowed to move beyond 15 Å from the heme⁺-Fe(IV)-O²⁻ (noted as heme-oxyanion). Following a Monte Carlo (MC) move where the substrate and the protein are first perturbed and then relaxed, the substrate binding in PELE is scored by the enzyme-substrate interaction energy, as described by the OPLS force field. PELE was ran on each system for maximally 5000 MC steps, after which the distance from the heme-oxyanion to the sesquiterpene lactone and docking orientation was scored using binding energy, thereby performing a population analysis (Fig. 3b and 3c).

The *in silico* docking studies of costunolide with TpKLS indicate that the C3 position of costunolide is highly favoured to interact with the heme oxy-anion in the TpKLS model structure (Fig. 3b), for both the constrained and unconstrained models (1379 and 1417 occasions, respectively). The calculated distance of costunolide C3 to the heme oxyanion in the model is 3.6 Å which translates into a hydrogen atom abstraction coordinate \sim 2.2 Å, sufficient to trigger hydroxylation at costunolide C3". Also in the *in silico* docking experiments of costunolide with costunolide-3 β -hydroxylase the PELE simulations indicated a preferred regioselective C3 orientation of costunolide towards the heme oxyanion, which is in agreement with its enzymatic activity (Figs. 3b,d; Supplementary Fig. 4). Moreover, the *in silico* docking of costunolide in TpKLS indicates that α -hydroxylation of C3 is favoured (Fig. 3d; Supplementary Fig. 4), while for costunolide-3 β -hydroxylase, in agreement with its enzymatic function, β -hydroxylation of C3 is preferred (Fig. 3c,e). Presumably, α -hydroxylation of costunolide at C3 is directly followed by protonation, dehydration, cyclisation and deprotonation in the enzymatic cavity, resulting in the formation of kauniolide. In contrast, the less preferred β -hydroxylation – which seems to occur, though at a much lower rate - at costunolide C3 apparently cannot be protonated and is removed from the enzymatic cavity as minor side-product.

Although for these two enzymes the predictions from the *in silico* modelling match with the actual enzyme activity, for TpPTS the modelling suggests epoxidation of C9 while the

enzyme epoxidises C4-C5 of costunolide (Supplementary Fig. 5). The number of predictions for this preferred costunolide C9 orientation towards the heme of *TpTPS* was much lower (~252) than the number of preferred orientations of costunolide C3 towards the heme of *TpKLS* (1417) and costunolide 3 β -hydroxylase (365), suggesting that the model of costunolide docking to *TpPTS* is less reliable than for the other two enzymes.

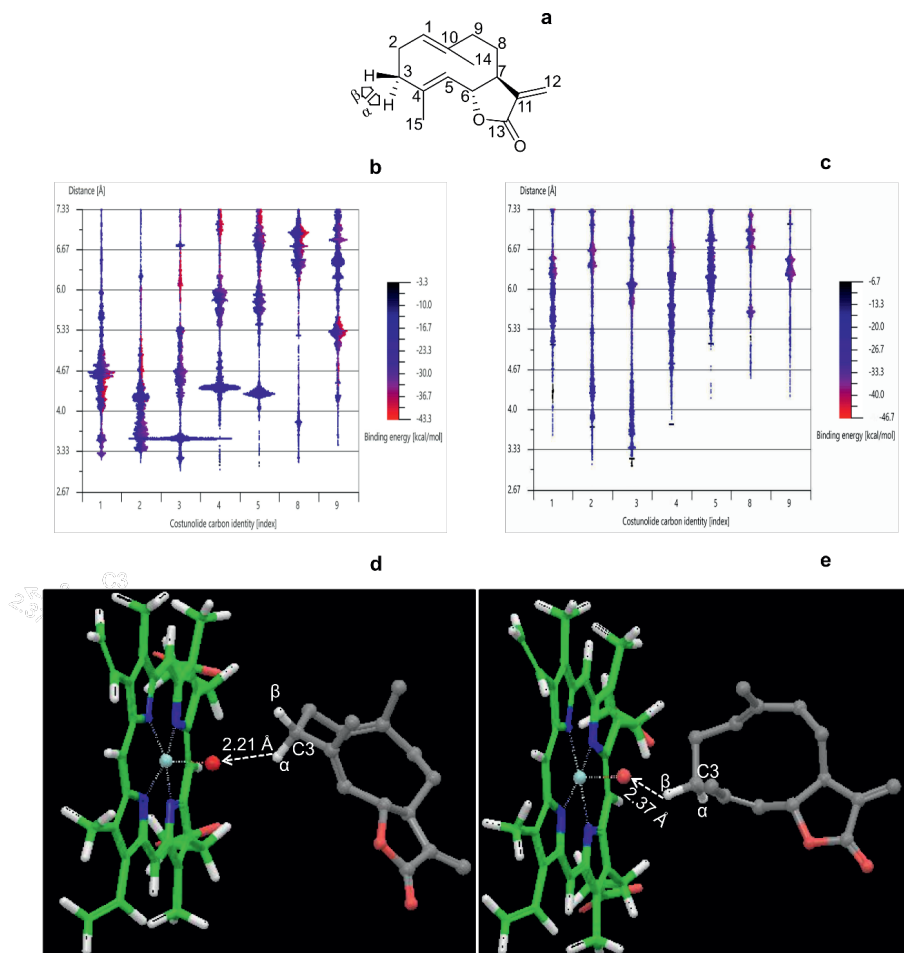


Figure 3. *In silico* docking of costunolide into the active site of the feverfew enzymes, kauniolide synthase and costunolide 3 β -hydroxylase. (a) Structure of costunolide and pro- α and pro- β hydroxylation orientations **(b)** Preferred docking orientations of costunolide as calculated using Protein Energy Landscape Exploration software. Costunolide carbon distance distribution relative to the heme-oxyanion in a KLS homology model (constrained model, at 15Å). **(c)** Costunolide carbon distance distribution relative to the heme-oxyanion in a costunolide 3 β -hydroxylase homology model (constrained model, at 15Å). *In silico* docking of costunolide into the active site of kauniolide synthase and costunolide 3 β -hydroxylase by Glide. **(d)** Preferred docking of costunolide in the active site of kauniolide synthase is in α -orientation. **(e)** Preferred docking of costunolide in the active site of costunolide 3 β -hydroxylase is in β -orientation.

3 α -hydroxycostunolide is the intermediate substrate for TpKLS

We therefore according to *in silico* docking studies hypothesized that 3 α -hydroxycostunolide serves as the intermediate substrate in kauniolide biosynthesis. Hence we chemically synthesized 3 α -hydroxycostunolide (Supplementary Fig. 6) and fed it to TpKLS. GC-MS analysis of extracts from feeding experiments showed conversion of 3 α -hydroxycostunolide to kauniolide by TpKLS (Fig. 4a and 4b). No kauniolide was detected in samples where 3 α -hydroxycostunolide was incubated with EV (Fig. 4c). We therefore suggest that conversion of costunolide into kauniolide starts with hydroxylation at the C3 position, predominantly in α -orientation and with β -orientation as a minor side reaction. This is immediately followed by a ring closure reaction for 3 α -hydroxycostunolide, which results in kauniolide formation, while the other intermediate, 3 β -hydroxycostunolide, leaves the enzymatic cavity as side product. The occasional release of the intermediate 3 β -hydroxycostunolide from the enzymatic cavity would explain the presence of 3 β -hydroxycostunolide as side product of *TpKLS*.

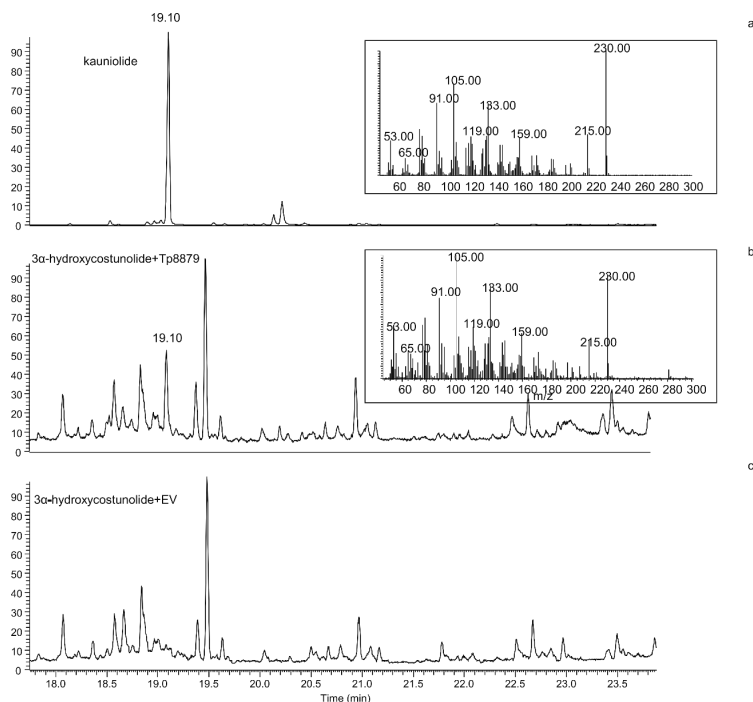


Figure 4. 3 α -hydroxycostunolide serves as the intermediate compound in kauniolide formation. (a). GC-MS chromatogram of kauniolide standard. **(b)** GC-MS chromatogram of product of assay with 3 α -hydroxycostunolide as substrate fed to yeast microsomes expressing *Tp8879*: production of kauniolide (nominal mass 230). **(c).** Negative control; feeding 3 α -hydroxycostunolide to microsomes containing EV does not lead to kauniolide formation.

Reconstruction of the kauniolide biosynthetic pathway in *Nicotiana benthamiana*

Previously we reconstituted the full biosynthesis pathway towards costunolide and parthenolide by transient gene expression in *N. benthamiana* (Liu et al., 2011; Liu et al., 2014). To obtain production of kauniolide in *N. benthamiana*, the *TpKLS* was cloned into a binary plant expression vector under control of the Rubisco small subunit promoter (RBC) which was used for agrobacterium transformation. All genes of the biosynthetic pathway towards kauniolide (*TpGAS*, *CiGAO*, *CiCOS* and *TpKLS*) were subsequently transiently expressed in *N. benthamiana* leaves by co-agroinfiltration. In addition, overexpression of the *Arabidopsis thaliana* 3-hydroxy-3-methylglutaryl-CoA reductase (*AtHMGR*) was used to boost farnesyl-diphosphate (FPP) production, which is the precursor of all sesquiterpenoids. *N. benthamiana* leaves were harvested four days after agro-infiltration (4 dpi), extracted with methanol:formic acid (0.1%) and extracts were measured by LC-Orbitrap-FTMS using targeted analysis for free and conjugated products (costunolide, kauniolide and costunolide/kauniolide-cysteine, and costunolide/kauniolide-glutathione conjugates). No free kauniolide was detected, but both costunolide-cysteine and a mass presumed to be kauniolide-cys ($[M+H]^+ = 352.1577$) were detected in the leaf extracts (Fig. 5b), indicating that kauniolide produced in *N. benthamiana* leaves is conjugated to cysteine.

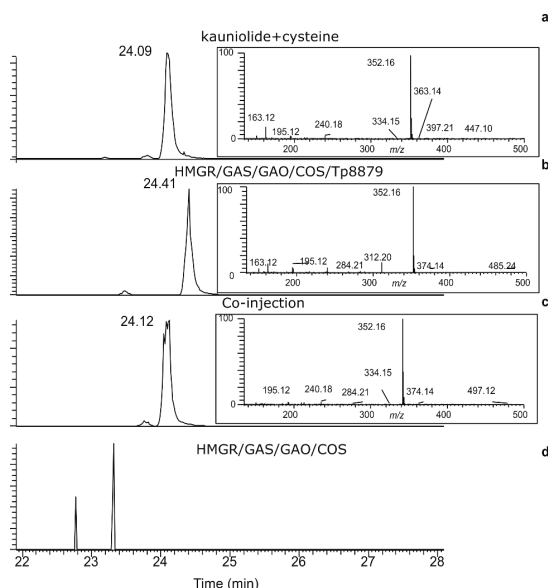


Figure 5. Characterization of *TpKLS* and reconstitution of kauniolide biosynthesis pathway in *Nicotiana benthamiana*. LC-Orbitrap-FTMS chromatograms at $[M+H]^+ = 352.1577$ of kauniolide-cysteine. **(a)** Kauniolide-cysteine standard formed from incubation of kauniolide and cysteine. **(b)** Demonstrates kauniolide-cysteine production by expressing the costunolide biosynthetic pathway (*AtHMGR*, *TpGAS*, *TpGAO* and *CiCOS*) in combination with *Tp8879* (*TpKLS*). **(c)** The chromatogram of co-injection of kauniolide-cysteine and the product of the costunolide biosynthetic pathway in combination with *Tp8879* (*TpKLS*). **(d)** Negative control. Expression of the costunolide biosynthetic pathway does not lead to formation of kauniolide-cysteine.

To confirm the identity of the compound with mass 352.1577, kauniolide was (non-enzymatically) conjugated *in vitro* to cysteine to produce kauniolide-cys. Indeed, retention time and molecular mass ($[M+H]^+ = 352.1577$) of kauniolide-cys were identical to the product obtained *in planta* upon expression of the kauniolide biosynthetic pathway (Fig. 5a-5c). Neither kauniolide-cysteine nor kauniolide-GSH were detected in control samples (costunolide pathway co-expressed with empty vector).

Reconstruction of the kauniolide biosynthetic pathway in yeast

Recently, reconstruction of the biosynthetic pathways of the sesquiterpene lactones, costunolide and artemisinin, in yeast was demonstrated (Paddon et al., 2013; Liu et al., 2014). To achieve the same for kauniolide, we cloned the four biosynthetic genes (*TpGAS*, *TpGAO*, *TpCOS* and *TpKLS*) into two dual yeast expression vectors and transformed yeast cells by plasmid transformation under the control of a galactose inducible promoter. Yeast cultures were induced by galactose and after 72 hrs cells plus medium were extracted with ethyl acetate. The extracts were concentrated, dried over a Na_2SO_4 column and the dehydrated sample was injected into GC-MS for analysis. Yeast cells expressing *TpGAS*, *TpGAO* and *TpCOS* produced costunolide, confirming previous results²⁰. The GC-MS chromatogram from the yeast cells expressing *TpGAS*, *TpGAO*, *TpCOS* and *TpKLS* showed a new peak at 19.10 min, with the same retention time and mass spectrum as the kauniolide standard (Fig. 6a and 6b). This indicates a successful production of free kauniolide in yeast cells. No conjugated kauniolide/costunolide were detected by GC-MS.

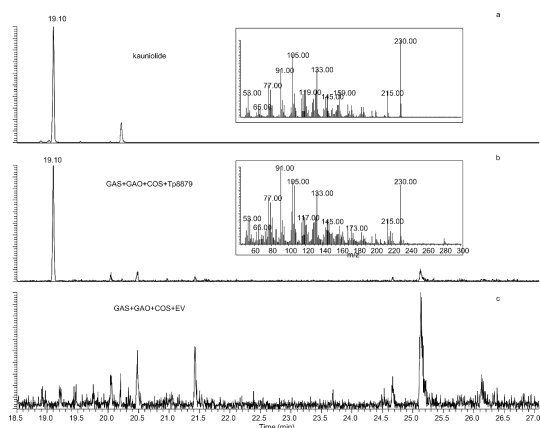


Figure 6. Characterization of *TpKLS* and reconstitution of the kauniolide biosynthesis pathway in yeast. GC-MS chromatograms at nominal mass 230 of kauniolide. (a) Kauniolide standard (b) Demonstrates kauniolide production by expressing the costunolide biosynthetic pathway (*TpGAS*, *TpGAO* and *CiCOS*) in combination with *Tp8879* (*TpKLS*). (c) Negative control. Expression of the costunolide biosynthetic pathway with EV does not lead to formation of kauniolide.

Discussion

Several sesquiterpene lactones from feverfew have been identified in feverfew (Fischedick et al., 2012) and several biosynthetic pathway enzymes and their corresponding genes have been characterized such as germacrene A synthase, costunolide synthase, parthenolide synthase, and costunolide 3 β -hydroxylase (Liu et al., 2011; Hullaert et al., 2014). Nevertheless the biosynthesis of several feverfew sesquiterpene lactones (e.g. santamarine, reynosin and artecamin) is still a mystery. Costunolide may play a central role in sesquiterpene lactones biosynthesis since several P450s have been shown to use this compound as substrate for oxidation (Ikezawa et al., 2011; Hullaert et al., 2014). Furthermore other researchers have proposed that costunolide might be the branching point for biosynthesis of other classes of sesquiterpene lactones such as guaianolides and eudesmanolides (Khasenov and Turdybekov, 2001; de Kraker et al., 2002; Ikezawa et al., 2011).

Here, for the first time we demonstrate the catalytic activity of a CYP71 (TpKLS) responsible for the production of a guaianolide type sesquiterpene lactone, kauniolide, from a germacranolide type sesquiterpene lactone, costunolide. Guaianolide biosynthesis has been suggested to start with C3 hydroxylation of costunolide, followed by cyclase activity by a separate enzyme (Piet et al., 1996). The cyclization would involve the following steps: i) protonation of the newly introduced hydroxyl group; ii) loss of water, producing an allylic carbocation C3⁺-C4=C5/C3=C4-C5⁺; iii) nucleophilic attack of the C1-C10 double bond to the C5 carbocation, forming the annellated 5-7 bicyclic system and a tertiary carbocation at C10; iv) regioselective deprotonation at C1 giving a C1-C10 double bond. It is likely that at least some of these steps occur in a concerted fashion (Piet et al., 1996). Here we show that all these reaction steps take place within a single P450. The production of 3 β -hydroxycostunolide as a minor side product of TpKLS enzyme activity, confirms that the enzyme hydroxylates costunolide at the C3 position. It has been shown that a different CYP71 from feverfew (costunolide 3 β -hydroxylase), stereospecifically adds a hydroxyl group in β -orientation to the C3 position of costunolide but does not convert it to kauniolide (Liu et al., 2014). Using yeast microsome feeding assays, we show that 3 β -hydroxycostunolide cannot be converted to kauniolide by TpKLS. Indeed, protein modelling suggests that C3- β -hydroxylation of costunolide does not allow the correct orientation in the active site of TpKLS for further postulated protonation (Fig. 3d). In contrast, α -orientation of the C3-hydroxyl group results in a suitable orientation of the substrate in the enzyme cavity which allows such a reaction to occur. Costunolide was indeed docked in a rather peculiar way in TpKLS. A conserved active site threonine in α -helix I appears to have different functions for different substrates in literature (Imai et al., 1989; Yoshigae et al., 2013) and may have an additional role in this system. It may well be that the conserved threonine of α -helix I protonates costunolide after

the hydroxylation step. Protonation by this threonine does not occur in the other discussed enzymes. After ring closure the conformation of the product is very different from the substrate, enabling it to move out of the active site, after which the threonine can be protonated again. Perhaps the threonine oxyanion is involved in the regioselective deprotonation of the intermediate carbocation, leading to kauniolide. We therefore suggest that TpKLS has costunolide C3-hydroxylase activity with a preference for α - over the β -orientation, while the 3 α -hydroxycostunolide is subsequently converted to kauniolide by the same enzyme. We note that modelling results are based on a heterologous P450 template model. Future acquisition of the TpKLS crystal structures may substantially improve the *in silico* simulation.

Other enzymes in the feverfew sesquiterpene lactone biosynthesis pathway also display a number of sequential activities, all initiated by oxidation. For instance, the P450 TpGAO converts germacrene A into germacrene A alcohol, germacrene A aldehyde and finally germacrene A acid (Liu et al., 2011). A similar type of single-enzyme multiple reactions also occurs in the closely related *Artemisia annua*, in which amorphadiene is converted by CYP71AV1 into amorphadiene alcohol, amorphadiene aldehyde and finally artemisinic acid (Ro et al., 2006). Also the hydroxylation of germacrene A acid by COS is followed by spontaneous lactone formation (de Kraker et al., 2002; Liu et al., 2011). However, in all these examples the sequential enzyme activities target the same position in the substrate. What makes TpKLS unique is that this enzyme not only carries out the conventional P450 catalysed hydroxylation, but also the protonation and dehydration, cyclisation and subsequent deprotonation (Fig. 7).

Phylogenetic analysis of the cytochrome P450s from feverfew and some other Asteraceae plants shows that TpKLS clusters more closely with COS proteins than with PTS or costunolide 3 β -hydroxylase and the GAOs (Supplementary Fig. 8). Analysis of more than 2.3 million expressed sequenced tags (ESTs) from different Asteraceae species (<http://compgenomics.ucdavis.edu>) identified contigs from these species having significant homology with TpKLS (but still with less than 65% homology) (Supplementary Fig. 7, Supplementary Table 1). Indeed, some of these Asteraceae species are known to produce guaianolide type sesquiterpene lactones such as *Cichorium* spp. (SETO et al., 1988) (lactucin, lactucopicrin and 11 β ,13-dihydrolactucin), *Cynara cardunculus* (Rial et al., 2016) (aguerin B, cynaropicrin). This suggests that a specific and quite unique protein with KLS activity has evolved in feverfew out of COS, but that putative KLSs in other Asteraceae may have evolved independently. This is in stark contrast with the strong sequence similarity of COS genes in Asteraceae species, resulting in a strong clustering in phylogenetic analysis (Supplementary Fig. 8).

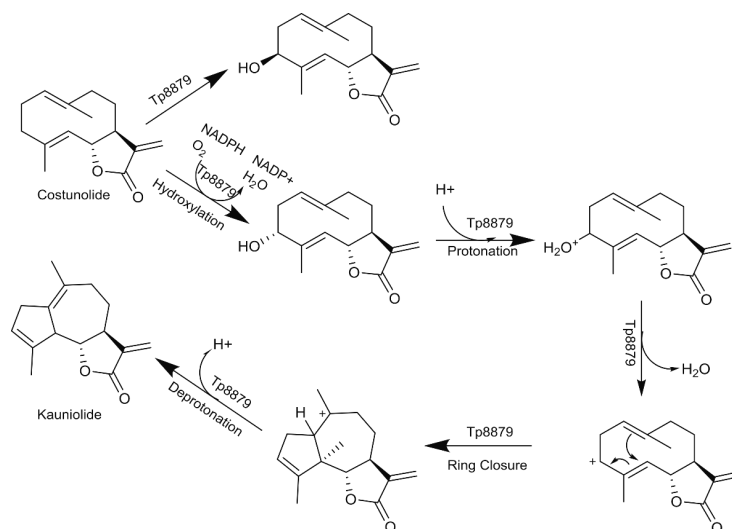


Figure 7. Mechanism of action of *TpKLS*. Costunolide is hydroxylated at the C3 position (in which the α -orientation is favoured). The C3 α -hydroxylation is followed by protonation and water abstraction resulting in an allylic carbocation. The resonance form with the C5-carbocation is attacked by the C1-C10 double bond which results in C1-C5 ring closure. Subsequent deprotonation results in kauniolide.

Feverfew indeed contains guaianolide sesquiterpene lactones, artecanin and tanaparthin- β -peroxide (Fischedick et al., 2012). Artecanin has anti-proliferative properties on leukemia cancer cells (Molnár et al., 2016). Artecanin biosynthesis requires multiple oxidation steps of kauniolide, likely catalysed by P450s, and these may be identified by feverfew P450 expression profiling and comparison with the accumulation pattern of artecanin (Supplementary Fig. 9). Artecanin shows a stable accumulation level during development of feverfew flowers, but suddenly increases late during development of the flowers. The increased production of artecanin in this late stage of flower development may be due to reduced expression of PTS (Liu et al., 2014) and therefore reduced competition with PTS for the costunolide substrate. Our elucidation of the biosynthesis of kauniolide, the most basic guaianolide, will provide new opportunities for bio-engineering of valuable guaianolides like artecanin.

Materials and Methods

RNA-sequencing and expression analysis

The *Tanacetum parthenium* EST library used in the present study was made from flower mRNA as reported (Majdi et al., 2011). Later Illumina-reads of ovary developmental stages were mapped (Langmead et al., 2009) against this assembly and eXpress (bio.math.berkeley.edu/eXpress) was accommodated for expression level estimation in different developmental stages. DEseq(Anders and Huber, 2010) was used to check

differential expression levels for three biological replicates. A negative binomial distribution test with the FDR threshold of 0.05 was used for Deseq. Mean expression values of CYP71 candidates together with feverfew sesquiterpene lactones biosynthesis pathway genes were directly imported to the software GeneMaths XT (2.12). This was followed by log transformation and normalization of the complete set with the average algorithm for data offset calculation. A complete linkage algorithm was used for cluster analysis (Fig. 2a). Focusing on genes showing a similar expression pattern as the pathway genes limited the number of candidates to six, which were clustering together with other known biosynthesis pathway genes.

Gene expression analysis

Relative gene expression of candidate P450s was measured by quantitative real time RT-PCR. Total RNA and cDNAs were obtained from ovary developmental stages as previously reported (Liu et al., 2011). Real time RT-PCR was performed using a LightCycler 480 (Roche Diagnostics). The LightCycler experimental run protocol used was: 95 °C for 10 min, 95 °C for 10s, 60 °C for 30s for 40 cycles and finally a cooling step to 40 °C. LightCycler Software 1.5.0 was used for data analysis. In order to determine the efficiency, a six serial dilution series point (from 200 to 6.25 ng) was used which was performed in triplicate. Primer pairs for *TpGAS* and *TpActin* were described previously (Majdi et al., 2011). The following primer pairs were used for *TpGAO*, *TpCOS* and *TpPTS* amplification:

| | | | |
|------------------------------|---------|---------------|-----|
| TGCAGCTCCCGCTTGCTAATATAC-3', | reverse | <i>TpGAO</i> | 5'- |
| AGTCTTTCTTTGAACCGTGGCTCC-3', | forward | <i>TpCOS</i> | 5'- |
| TAGCTTCATCCCGGAGCGATTGGA-3', | reverse | <i>TpCOS</i> | 5'- |
| AAATTCTTCGGCCCGCACCAAATG-3', | forward | <i>TpPTS</i> | 5'- |
| AGACATTACGTTTACACCCTCCCG-3', | reverse | <i>TpPTS</i> | 5'- |
| ATCACGACACAAGTCCCAGGGAAA-3', | forward | <i>Tp8879</i> | 5'- |
| AGTTCCTTACGGCCATTTCTGGGA-3', | reverse | | 5'- |

AGAAGATCGTCCTCAGTTGCTCCA-3'. Transcript levels were quantified in three independent biological replicates and three technical replicates for each biological replicate. ΔCT was calculated according to the following formula: $\Delta CT = CT (\text{Target}) - CT (\text{Actin})$. Fold change calculation was done by using the $2^{-\Delta CT}$ equation (Schmittgen and Livak, 2008).

Isolation and cloning of full length candidate genes from feverfew

Six candidate cytochrome P450 contigs were identified by sequence homology to known sesquiterpene monooxygenases. These candidates were selected for functional characterization. A RACE-PCR (Clontech) approach was used to obtain the complete coding sequence of the 5'- and/or 3'-regions. The coding sequence (CDS) was amplified by PCR and

at the same time introducing NotI and PacI restriction sites. The CDS was subsequently cloned into a yeast expression vector, pYEDP60 which was described before (Pompon et al., 1996). Both the TPS and P450 cDNA sequences used in this study were identified in the same glandular trichome enriched cDNA library and we assume that both genes are expressed in the same cell. The cDNA sequence of the candidate genes has been submitted in GeneBank under the accession numbers MF197558 (*Tp8879*) and MF197559 (*Tp8886*). The sequences were also deposited in David Nelson's cytochrome P450 database (<http://drnelson.uthsc.edu/cytochromeP450.html>), *Tp8879* and *Tp8886* were assigned the names CYP71BZ6X and CYP71DD5 respectively (Nelson, 2009).

Plasmid construction for gene expression in yeast

Candidate *P450s* were cloned into the pYED60 vector using NotI/PacI restriction sites. The obtained constructs were transformed into the WAT11 (Urban et al., 1997) yeast strain. After transformation yeast clones were selected on SD minimal medium supplemented with amino acids, but omitting uracil and adenine sulphate for auxotrophic selection of transformants.

Yeast microsome isolation and *in vitro* microsome assay

Microsome isolation of transformed yeasts was done according to (Pompon et al., 1996) with some modifications. Transformed yeast single colonies were pre-cultured in 50 mL SGI medium for 3 days at 30°C and shaking at 300 rpm. Then a 250 mL YPL medium, containing 2% galactose was added for induction of gene expression and was kept for 24 hours under the same conditions. Subsequently, the cells were harvested and chilled on ice for 20 min, followed by centrifugation at 4900×g for 10 min. Cell pellets were re-suspended in 100 mL extraction buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.6 M sorbitol and 10 mM β-mercaptoethanol) and kept for 10 min at room temperature. Again the cells were centrifuged (4900×g for 10 min) and cells were washed three times (3 mL) with extraction buffer, but omitting β-mercaptoethanol. The centrifuge tube was rinsed with 2 mL of the same buffer and the mixture was transferred to a 50 mL Falcon tube. Approximately 25 mL of glass beads (450 μm) were added to the tubes and cells were lysed by vigorous shaking in a cold room for 10 min. The cell lysate was transferred to a 25 mL centrifuge tube and were centrifuged at 10500×g for 10 min. later the supernatant was again centrifuged at 195000×g for 2 hours. The pellet (microsomal fractions) were re-suspended in a 4 mL solution of ice-chilled 50 mM Tris-HCl (pH 7.5) containing 1 mM EDTA and 20% (v:v) glycerol using a glass Tenbroek homogenizer. Aliquots of these microsomal fractions were kept at -80 °C until use.

Yeast microsome *in vitro* assays were done in a mixture of 28.8 μl isolated microsomal fractions, 4 μl substrate (10mM in DMSO), 40 μl NADPH (10 mM in 100 mM potassium buffer), 8 μl potassium buffer (1 M, pH7.5), and 90.4 μl of water which was incubated for 2.5

h at 25 °C with shaking (200 rpm). The obtained mixture was centrifuged at 12000 ×g (4 °C) and supernatant was passed through a 0.22 µm filter before injection into LC-MS.

LC-Orbitrap-MS analysis of yeast microsome assay mixture

In order to analyse the enzymatic mixture of microsome assays we used a LC-LTQ-Orbitrap FTMS system (Thermo Scientific) consisting of an Accela HPLC, an Accela photodiode array detector, connected to an LTQ/Orbitrap hybrid mass spectrometer equipped with an ESI source. Chromatographic separation was done on an analytical column (Luna 3 µ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA). Eluent A (degassed) [HPLC grade water: formic acid (1000:1, v/v)] and eluent B [acetonitrile: formic acid (1000:1, v/v)] were used. The flow rate was set to be 0.19 mL min⁻¹. The applied gradient in 45 min was from 5 to 75% acetonitrile, followed by a 15 min washing step and equilibration. Full scan mass analysis (FTMS) at the resolution of 60000 were recorded while the resolution for MSⁿ scans was set to 15000. Calibration of FTMS was done externally in negative ionisation mode by using sodium formate clusters in the range m/z 150–1200, and automatic tuning was performed on m/z 384.93. Injection volume for each sample was 5 µl.

Plasmid construction for expression in *Nicotiana benthamiana*

Heterologous transient expression of candidate P450s was done according to (Liu et al., 2014). Candidate P450 CDSs were cloned into Impactvector 1.1 under the control of Rubisco (RBC) promoter (Liu et al., 2014). Other pathway genes (*GAO*, *COS*, *PTS*) were also cloned into the same expression vector. *TpGAS* was also cloned into ImpactVector1.5 to fuse it with the RBC promoter and the CoxIV mitochondrial targeting sequence as it was shown before that mitochondrial targeting of sesquiterpene synthases results in improved sesquiterpene production (Liu et al., 2011). Later the CDS was transformed into the pBinPlus binary vector (Vanengelen et al., 1995) by an LR reaction (Gateway-LR Clonase TM II) to put the CDS between the right and left borders of T-DNA for plant transformation. All these constructs were finally transformed into and AGL-0 *Agrobacterium* strain. An LR reaction (Gateway-LR Clonase TM II) was carried out to clone each gene into the pBinPlus binary vector (Vanengelen et al., 1995) between the right and left borders of the T-DNA for plant transformation.

Transient expression in *Nicotiana benthamiana*

Transient heterologous expression of biosynthetic pathway genes was done by *Agrobacterium* mediated transformation (agro-infiltration) of *Nicotiana benthamiana* plants. Transformed *Agrobacteria* were grown at 28 °C at 250 rpm for 48 h in LB media with proper antibiotics. Then *Agrobacterium* cells were harvested at 4000×g for 20 min at room temperature. Cells

were then resuspended in agro-infiltration buffer (10mM MES, 10 mM MgCl₂ and 100 μ M acetosyringone). The final optical density (OD) in the 600nm wavelength absorbance was set to 0.5. Gene mixture dosage for control treatments were adjusted by addition of representative number of empty vector(s). Cultures were mixed on a roller-mixer for 2.5 h (50 rpm). 4 weeks old *N.benthamiana* plants were agroinfiltrated with a needleless 1 mL syringe injection to the abaxial side of the leaf. Transformed plants were kept for another four and half days and then harvested for metabolites analysis.

LC-QTOF-MS analysis of leaf extracts

Transformed leaves were harvested 4.5 days post infiltration (dpi) and snap-frozen in liquid N₂. Leaves were ground into fine powder and extracted with 300 methanol:formic acid (1000:1, v:v) extraction buffer. Short vortexing was followed by 15 minutes sonication. Then extracts were centrifuged at 13000 \times g for 10 min and 200 μ l of supernatant was passed through a 0.22 μ m inorganic membrane filter (RC4, Sartorius, Germany) and were kept in injection vials for LC-QTOF-MS analysis.

LC-QTOF-MS analysis of plant leaf extracts was done on a Waters Alliance 2795 HPLC connected to a Waters 2996 PDA detector and subsequently a QTOF Ultima V4.00.00 mass spectrometer (Waters, MS technologies, UK). Samples were measured in negative ionisation mode. The analytical column was Luna 3 μ C18/2 100A; 2.0 \times 150 mm (Phenomenex, USA) which was attached to a C18 pre-column (2.0 \times 4 mm; Phenomenex, USA). Eluent A ([HPLC-grade water: formic acid (1000:1, v/v)]) and Eluent B [acetonitril:formic acid (1000:1, v/v)] were run over the flow rate of 0.19 mL.min⁻¹. Masses were recorded between *m/z* 80 and *m/z* 1500. The leucine enkephalin ([M-H]⁻=554.2620) was used as a lock mass for on-line accurate mass correction. The HPLC gradient started with 5% eluent B and linearly increased to 75% at 45 min. Then the column was washed and equilibrated for 15 min before the next injection. The sample injection volume was set to be 5 μ l.

Cysteine conjugation

Cysteine conjugation was performed as described by Liu *et al.*(2011). In brief cysteine (150 mM) in 7 μ l potassium buffer (100 mM; pH 6.5), and standards (30 mM) in 7 μ l ethanol were added to 1000 μ l potassium buffer (100 mM; pH 6.5). The reaction was initiated by adding 7 μ l of GST (1g L⁻¹, in 100 mM potassium buffer; pH 6.5) into the mixture. Complete assay mixtures with(out) GST enzyme or either of the substrates were used as controls. Samples were incubated for 30 min at room temperature and were kept at -20 °C until analysis. Costunolide was purchased from TOCRIS Bioscience (United Kingdom). Parthenolide, 3 β -hydroxycostunolide, and 3 β -hydroxyparthenolide, isolated from dried aerial parts of feverfew plants, were provided by Dr. Justin T. Fishedick from PRISNA company(Fishedick et al.,

2012). Kauniolide, 1 β ,10 β -epoxy-kauniolide (arglabin), 1 α ,10 α -epoxy-kauniolide and 3 β ,4 β -epoxy- kauniolide were kindly provided by Prof. Yue Chen (State Key Laboratory of Medicinal Chemical Biology, Nankai University, China) (Zhai et al., 2012).

Chemical synthesis of 3 α -hydroxycostunolide

1 mg of 3 β -hydroxycostunolide was dissolved in 1 ml of pentane. Then 50 mg of manganese dioxide (MnO₂, Sigma-Aldrich) was added and shaken at room temperature for 40 hr. The supernatant was passed through a filter to remove MnO₂ and checked by GC-MS. No 3 β -hydroxycostunolide was detected in extracts and a new peak with the same mass as 3-oxocostunolide was detected. The solvent was evaporated to dryness by gently blowing a stream of N₂. In order to reduce 3-oxocostunolide, the sample was dissolved in 1 ml of ethanol and 10 mg of sodium borohydride (NaBH₄) was added. The mixture was shaken for 1 hr at room temperature. Then 1 drop of acetic acid was added to the mixture to destroy the complex of the product with borium and liberate the free alcohol. The solvent was then evaporated with a vacuum concentrator and the residue dissolved in 1 ml diethyl ether. Subsequently, 1 ml of saturated NaHCO₃ was added to the mixture to remove the acetic acid and the sample was washed several times with 1 ml of MQ water. The mixture was then extracted two more times with 1 ml diethyl ether, dried over Na₂SO₄, concentrated and injected into GC-MS.

GC-MS analysis

2000 mL of yeast culture (medium and cells) was extracted twice with 500 mL of ethyl acetate. Extracts were concentrated to 100 μ L using a rotary evaporator and a N₂ flow and used for GC-MS analysis. A gas chromatograph (7809A, Agilent, USA) equipped with a 30 m \times 0.25 mm \times 0.25 mm film thickness column (DB-5) was used. Helium was used as the carrier gas and the flow rate was adjusted to 1 mL min⁻¹ for GC-MS analysis. No splitting was used for injection and inlet temperature set to 250°C. The initial oven temperature was 45°C for 1 min, and increased to 300°C after 1 min at a rate of 10°C min⁻¹ which was held for 5 min at 300°C. The GC was coupled to a Triple-Axis detector (5975C, Agilent).

Homology modelling

To create homology models of kauniolide synthase, costunolide 3 β -hydroxylase and parthenolide synthase, tertiary structure templates were sought using the Protein-Protein Basic Local Alignment Tool (BLASTP) from NCBI, focussing on the Protein Data Bank (PDB) database (Altschul et al., 1997). Additional templates were retrieved using Protein-Protein HMMER (phmmmer), directed to all available databases (Finn et al., 2015). The resulting query-template alignments were compared in the Multiple Sequence Viewer of the software

package Prime (Jacobson et al., 2002; Jacobson et al., 2004) within Maestro (<https://www.schrodinger.com/maestro>). One template hit was chosen based on sequence identity, residue homology and the presence of a ligand. Substrate Recognition Sites (SRS) (which were identified using the Cytochrome P450 Engineering Database [CYPED,(Fischer et al., 2007)]) were checked thoroughly for homology. A few cycles of homology model building and manual realignment followed. During the realignments, gaps/insertions were shifted in regions of low homology to connect residues smoothly with regions of higher homology. The N-terminal α -helix is absent in the model as none of the templates covered this region in their X-ray structure. Supplementary Table 2 lists the final query-template alignments. The homology models were cleaned up: excess waters/ions from the template structure were removed, appropriate charges were applied to the residues and disulfide bonds of cysteines were created using the Protein Preparation Wizard of Prime(Jacobson et al., 2002; Jacobson et al., 2004). The heme was inherited from the template but was replaced by a quantum mechanically described heme from the Parthenolide synthase model. For this quantum mechanical model of the heme, Compound I was created (Fe^{2+} was replaced by Fe^{5+} and O^{2-} was bonded to Fe^{5+}) in Maestro. Costunolide was docked in this modified Parthenolide synthase before QSite ^{42,43} (<https://www.schrodinger.com/qsite>), a mixed quantum mechanics/molecular mechanics software (QM/MM), was used on this system. The molecular mechanic region was comprised of the protein plus the substrate without the cysteine that bonds with the iron center. The quantum mechanical region was comprised of Compound I and the iron-chelating cysteine. For the quantum mechanical description, the DFT-MO6 function set was chosen. The OPLS2005 *in vacuo* Force Field was set for the molecular mechanical calculations. QSite (Philipp and Friesner, 1999; Murphy et al., 2000) kept the substrate and residues beside the cysteine frozen during the calculations. After inserting the heme from Parthenolide synthase in Kaunilide synthase and Costunolide β -Hydroxylase, the homology model was minimized. 500 Cycles of OPLS2005 Force Field minimization were run using Impact under Maestro (<https://www.schrodinger.com/maestro>). Implicit water was included in the minimization acting as an Analytic Generalized Born solvent. The resulting structure was used in substrate – enzyme interaction modelling.

Substrate modelling

Costunolide was drawn in Chem3D (Version 14.0.0.117) and minimized with the OPLS2005 Force Field over 100 cycles in Impact from Maestro. Implicit water was modelled during the minimization by the Poisson Boltzmann Solver. After this molecular mechanics minimization, costunolide was described quantum mechanically with Jaguar ⁴⁴ (<https://www.schrodinger.com/jaguar>) using the B3LYP-DFT level of theory and the 6-31G**

basis set. Electrostatic potential surface atomic charges were then extracted and used for the following enzyme-substrate interactions.

Enzyme - substrate interaction modelling

In preparation of binding simulations costunolide was docked in KLS, costunolide β -hydroxylase and PTS. Hereto, Glide (Friesner et al., 2004; Halgren et al., 2004) was used to produce a receptor grid of the active site, using as center the position of the ligand in the homology template. Polar side chains were allowed to rotate in the grid. Glide was used to dock costunolide in extra precision mode, using the partial charges from the quantum mechanical description of costunolide. 5000 Initial docking poses were produced. The 100 best poses were then minimized. One final docking pose was returned by Glide. This system was used as input for a molecular dynamics simulation using Desmond 47-50. Explicit water molecules with 150 mM of sodium chloride were simulated in a box of 10Å surrounding the protein. A NPT system was maintained, providing a constant number of particles, pressure and temperature. Standard pressure (~1bar) and a temperature of 300K were applied. A 100 N/m constraint was set to Compound I and a 50 N/m constraint was set onto the protein backbone. The simulation duration was 5 ns maximally. System volume and energy were inspected afterwards using the Quality Analysis tool of Desmond (Guo et al., 2010; Shivakumar et al., 2010). At a potential energy minimum as close to the end of the simulation as possible one simulation frame was selected. Compound I from the first simulation frame was inserted in the selected final simulation frame. For kauniolide synthase and parthenolide synthase this simulation frame was used as input for a subsequent simulation with Protein Energy Landscape Exploration (PELE)^{22,23}, aiming at refining the binding mode of costunolide in the CYP450s. The costunolide β -hydroxylase with docked costunolide was used as input directly. A force constant of 400 N/m was set on the bond of 2.45Å between the iron-chelating cysteine sulphur and iron cation.

PELE generates a MC chain where each step uses a ligand and backbone perturbation followed by a relaxation step involving a side chain prediction and a full-system minimization. Each ligand perturbation involved 50 different tries with: (i) a rotational change set to 0.02 radian (70%)/0.25 radian (30%); (ii) a translational displacement of 1.0Å(70%)/0.5Å (30%). The backbone perturbation followed an anisotropic network model approach using the lowest 6 eigenvectors. All side chains within 6 Å from the substrate were included in the relaxation step. PELE ran on the system for maximally 5000 steps. Two different set of simulations were run: one without a constraint on the substrate and one with a restricted search region where the substrate was not allowed to move beyond 15Å from Compound I. To characterise the enzyme-substrate interactions, PELE used the 2005 OPLS-AA force field and a surface-SGB implicit model solvent.

References

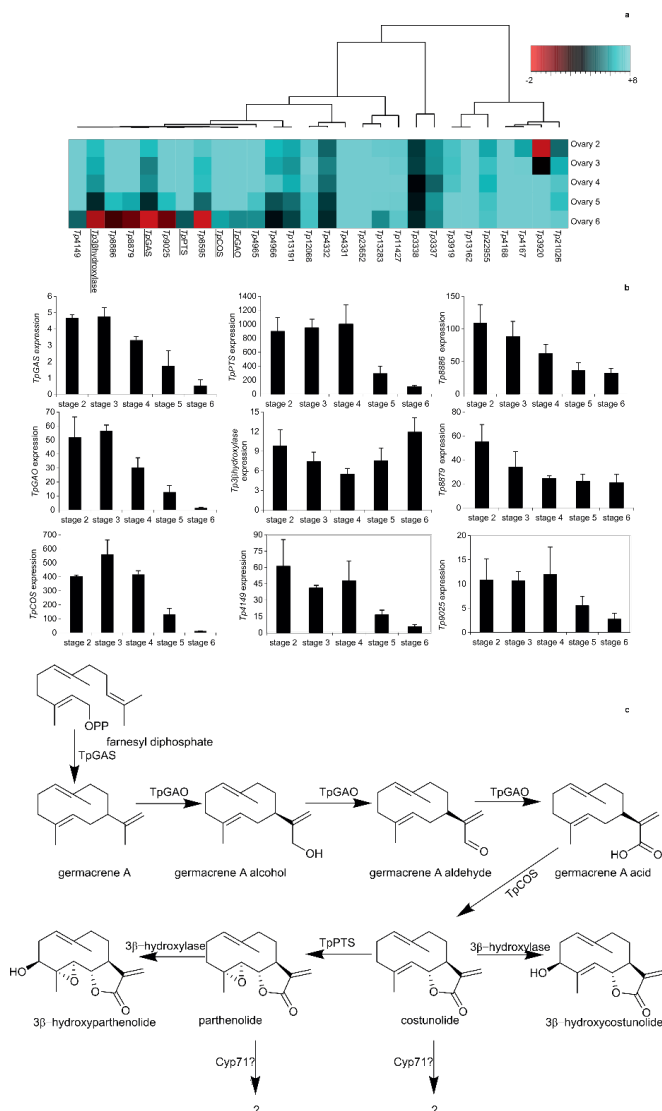
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**: 3389-3402
- Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome biology* **11**: R106
- Borrelli KW, Vitalis A, Alcantara R, Guallar V (2005) PELE: protein energy landscape exploration. A novel Monte Carlo based technique. *Journal of Chemical Theory and Computation* **1**: 1304-1311
- Bouwmeester HJ, Kodde J, Verstappen FWA, Altug IG, de Kraker J-W, Wallaart TE (2002) Isolation and Characterization of Two Germacrene A Synthase cDNA Clones from Chicory. *Plant Physiology* **129**: 134-144
- Cantrell CL, Nuñez IS, Castañeda-Acosta J, Foroozesh M, Fronczek FR, Fischer NH, Franzblau SG (1998) Antimycobacterial Activities of Dehydrocostus Lactone and Its Oxidation Products. *Journal of Natural Products* **61**: 1181-1186
- Chaturvedi D (2011) Sesquiterpene lactones: structural diversity and their biological activities, In-Opportunity, Challenges and Scope of Natural Products in Medicinal Chemistry. ISBN: 978-81-308-0448-4, Research Signpost, Trivandrum: 313-334
- Cogo J, Caleare AdO, Ueda-Nakamura T, Filho BPD, Ferreira ICP, Nakamura CV (2012) Trypanocidal activity of guaianolide obtained from *Tanacetum parthenium* (L.) Schultz-Bip. and its combinational effect with benznidazole. *Phytomedicine* **20**: 59-66
- Kraker J-W, Franssen MCR, Joerink M, de Groot A, Bouwmeester HJ (2002) Biosynthesis of Costunolide, Dihydrocostunolide, and Leucodin. Demonstration of Cytochrome P450-Catalyzed Formation of the Lactone Ring Present in Sesquiterpene Lactones of Chicory. *Plant Physiology* **129**: 257-268
- Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, Eddy SR (2015) HMMER web server: 2015 update. *Nucleic acids research*: gkv397
- Fischedick JT, Standiford M, Johnson DA, De Vos RCH, Todorović S, Banjanac T, Verpoorte R, Johnson JA (2012) Activation of antioxidant response element in mouse primary cortical cultures with sesquiterpene lactones isolated from *Tanacetum parthenium*. *Planta medica* **78**: 1725-1730
- Fischer M, Knoll M, Sirim D, Wagner F, Funke S, Pleiss J (2007) The Cytochrome P450 Engineering Database: a navigation and prediction tool for the cytochrome P450 protein family. *Bioinformatics* **23**: 2015-2017
- Fischer NH, Olivier EJ, Fischer HD (1979) The Biogenesis and Chemistry of Sesquiterpene Lactones. In Herz, H Grisebach, GW Kirby, eds, *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*. Springer Vienna, Vienna, pp 47-320
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **47**: 1739-1749
- Guo Z, Mohanty U, Noehre J, Sawyer TK, Sherman W, Krilov G (2010) Probing the α -helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. *Chemical biology & drug design* **75**: 348-359
- Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry* **47**: 1750-1759
- Hullaert J, Laplace DR, Winne JM (2014) A Three-Step Synthesis of the Guaianolide Ring System. *European Journal of Organic Chemistry* **2014**: 3097-3100
- Ikezawa N, Göpfert JC, Nguyen DT, Kim S-U, O'Maille PE, Spring O, Ro D-K (2011) Lettuce Costunolide Synthase (CYP71BL2) and Its Homolog (CYP71BL1) from Sunflower Catalyze Distinct Regio- and Stereoselective Hydroxylations in Sesquiterpene Lactone Metabolism. *The Journal of Biological Chemistry* **286**: 21601-21611
- Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishimura Y (1989) Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation. *Proceedings of the National Academy of Sciences* **86**: 7823-7827
- Jacobson MP, Friesner RA, Xiang Z, Honig B (2002) On the role of the crystal environment in determining protein side-chain conformations. *Journal of molecular biology* **320**: 597-608
- Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B, Shaw DE, Friesner RA (2004) A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics* **55**: 351-367
- Khasenov B, Turdybekov K (2001) Modelling transannular cyclization of 1 (10) E, 4E-germacranolide costunolide into eudesmane derivatives. *Chemistry of natural compounds* **37**: 451-454

- Kim EJ, Hong JE, Lim SS, Kwon GT, Kim J, Kim J-S, Lee KW, Park JHY (2012) The hexane extract of *Saussurea lappa* and its active principle, dehydrocostus lactone, inhibit prostate cancer cell migration. *Journal of medicinal food* **15**: 24-32
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**: R25
- Lepoittevin J-P, Berl V, Giménez-Arnau E (2009) α -methylene- γ -butyrolactones: versatile skin bioactive natural products. *The Chemical Record* **9**: 258-270
- 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 Q, Manzano D, Tanić N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metabolic engineering* **23**: 145-153
- Liu Q, Manzano D, Tanić N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S, Bouwmeester H (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metab Eng* **23**: 145-153
- Madadkar-Sobhani A, Guallar V (2013) PELE web server: atomistic study of biomolecular systems at your fingertips. *Nucleic acids research* **41**: W322-W328
- Majdi M, Liu Q, Karimzadeh G, Malboobi MA, Beekwilder J, Cankar K, de Vos R, Todorović S, Simonović A, Bouwmeester H (2011) Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.). *Phytochemistry* **72**: 1739-1750
- Molnár J, Szebeni GJ, Csupor-Löffler B, Hajdú Z, Szekeres T, Saiko P, Ocsóvszki I, Puskás LG, Hohmann J, Zupkó I (2016) Investigation of the Antiproliferative Properties of Natural Sesquiterpenes from *Artemisia asiatica* and *Onopordum acanthium* on HL-60 Cells in Vitro. *International journal of molecular sciences* **17**: 83
- Murphy RB, Philipp DM, Friesner RA (2000) A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *Journal of Computational Chemistry* **21**: 1442-1457
- Nelson DR (2009) The cytochrome P450 homepage. *Human Genomics* **4**: 59-65
- Nguyen DT, Göpfert JC, Ikezawa N, MacNevin G, Kathiresan M, Conrad J, Spring O, Ro D-K (2010) Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. *Journal of Biological Chemistry* **285**: 16588-16598
- 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, Lievens 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
- Philipp DM, Friesner RA (1999) Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *Journal of computational chemistry* **20**: 1468-1494
- Piet DP, Franssen MCR, de Groot A (1996) Biotransformation of allylically activated (E, E)-cyclodeca-1, 6-dienols by *Cichorium intybus*. *Tetrahedron* **52**: 11273-11280
- Piet DP, Schrijvers R, Franssen MCR, de Groot A (1995) Biotransformation of germacrene epoxides by *Cichorium intybus*. *Tetrahedron* **51**: 6303-6314
- Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Cytochrome P450, Pt B* **272**: 51-64
- Reynald RL, Sansen S, Stout CD, Johnson EF (2012) Structural Characterization of Human Cytochrome P450 2C19 active site differences between P450s 2C8, 2C9, and 2C19. *Journal of Biological Chemistry* **287**: 44581-44591
- Rial C, García BF, Varela RM, Torres A, Molinillo JM, Macías FA (2016) The Joint Action of Sesquiterpene Lactones from Leaves as an Explanation for the Activity of *Cynara cardunculus*. *Journal of Agricultural and Food Chemistry* **64**: 6416-6424
- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940-943
- Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **3**: 1101-1108

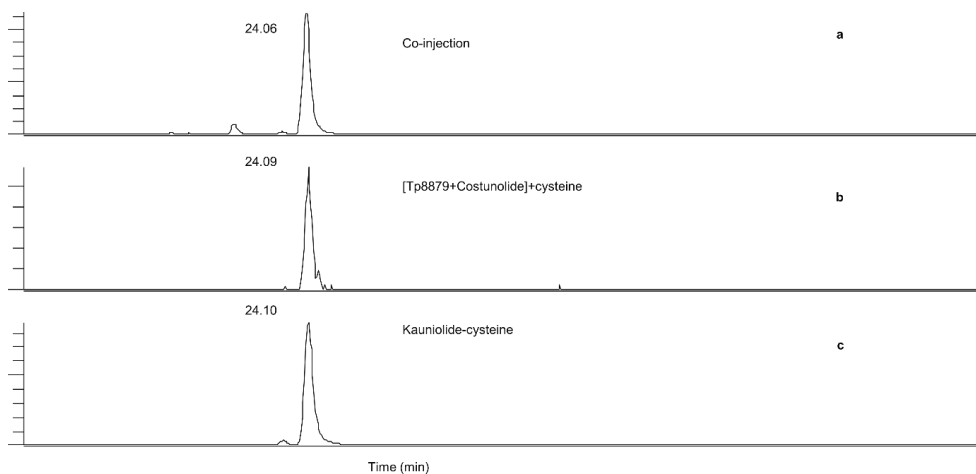
- SETO M, Miyase T, Umehara K, Ueno A, Hirano Y, Otani N** (1988) Sesquiterpene lactones from *Cichorium endivia* L. and *C. intybus* L. and cytotoxic activity. *Chemical and Pharmaceutical Bulletin* **36**: 2423-2429
- Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W** (2010) Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Journal of chemical theory and computation* **6**: 1509-1519
- Simonsen HT, Weitzel C, Christensen SB** (2013) Guaianolide Sesquiterpenoids: Pharmacology and Biosynthesis. In KG Ramawat, J-M Mérillon, eds, *Natural Products: Phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 3069-3098
- Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D** (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *Journal of Biological Chemistry* **272**: 19176-19186
- Vanengelen 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
- Wester MR, Johnson EF, Marques-Soares C, Dijols S, Dansette PM, Mansuy D, Stout CD** (2003) Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 Å resolution: evidence for an induced fit model of substrate binding. *Biochemistry* **42**: 9335-9345
- Yoshigae Y, Kent UM, Hollenberg PF** (2013) Role of the highly conserved threonine in cytochrome P450 2E1: prevention of H₂O₂-induced inactivation during electron transfer. *Biochemistry* **52**: 4636-4647
- Zhai J-D, Li D, Long J, Zhang H-L, Lin J-P, Qiu C-J, Zhang Q, Chen Y** (2012) Biomimetic Semisynthesis of Arglabin from Parthenolide. *The Journal of Organic Chemistry* **77**: 7103-7107

Supplementary information

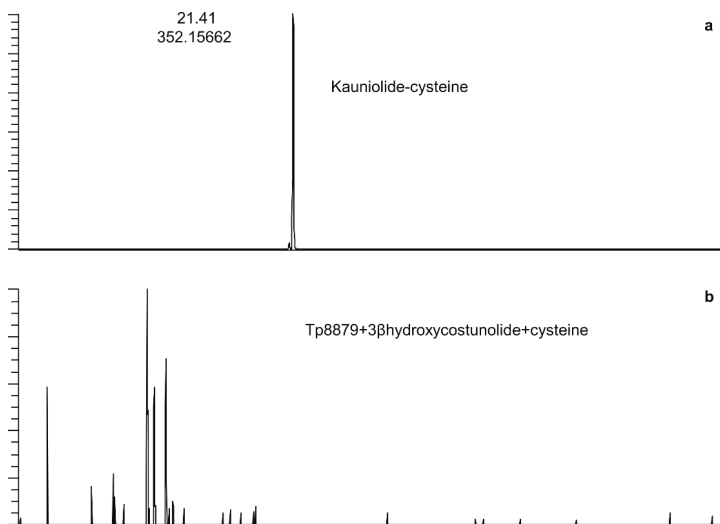
Supplementary Figures



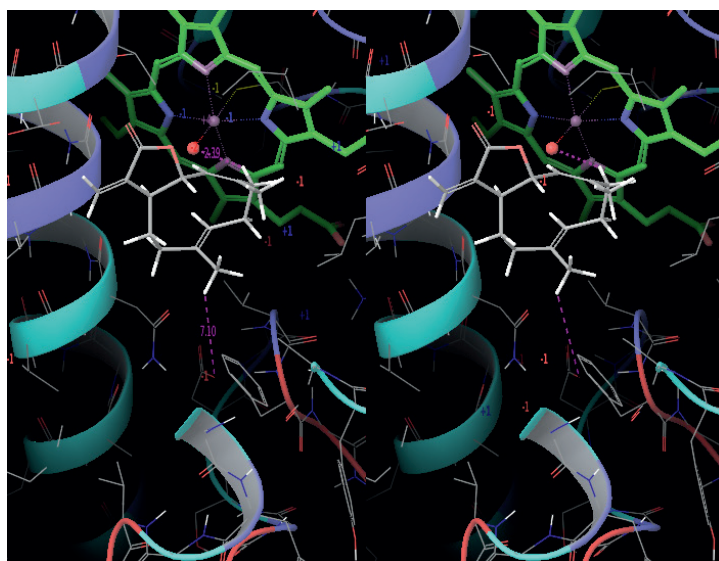
Supplementary Figure 1. Expression profile of candidate feverfew CYP71 genes grouped together with characterized feverfew CYP71 genes. (a) Complete linkage cluster analysis of differentially expressed feverfew CYP71 genes. Heat-map representation of the expression data of genes during five developmental stages of feverfew flowers. Each data point is the mean value of the three measured biological replicates. Underlined genes are the characterized TpCYP71 genes. (b) Validation of expression of selected obtained genes from RNA-sequencing. Expression profile of candidate feverfew CYP71 genes grouped together with characterized feverfew CYP71 genes (n=3) by Real-Time qPCR. (c) Proposed feverfew sesquiterpene lactone biosynthesis pathway extension by candidate CYP71 genes.



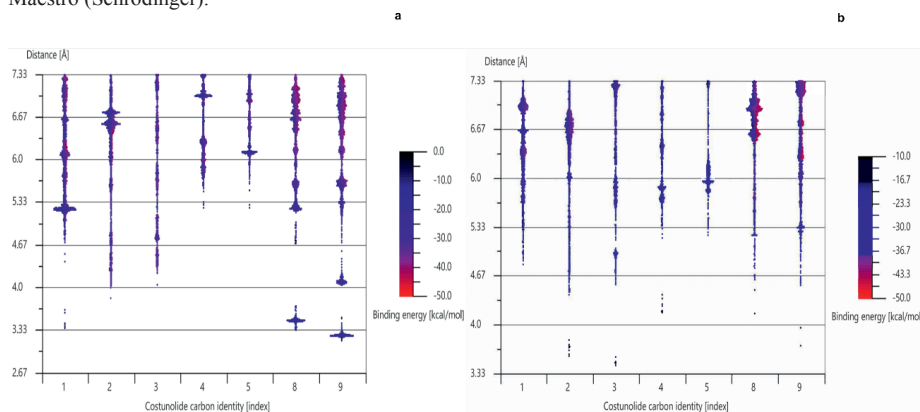
Supplementary Figure 2. Chromatogram of LC-Orbitrap-FTMS analysis of enzymatic reaction mixture. (a) co-injection of kauniolide-cysteine and product of feeding costunolide to *Tp8879* (*TpKLS*) incubated with cysteine. (b) kauniolide-cysteine production from feeding costunolide to *Tp8879* (*TpKLS*) and subsequent incubation with cysteine (c) kauniolide-cysteine.



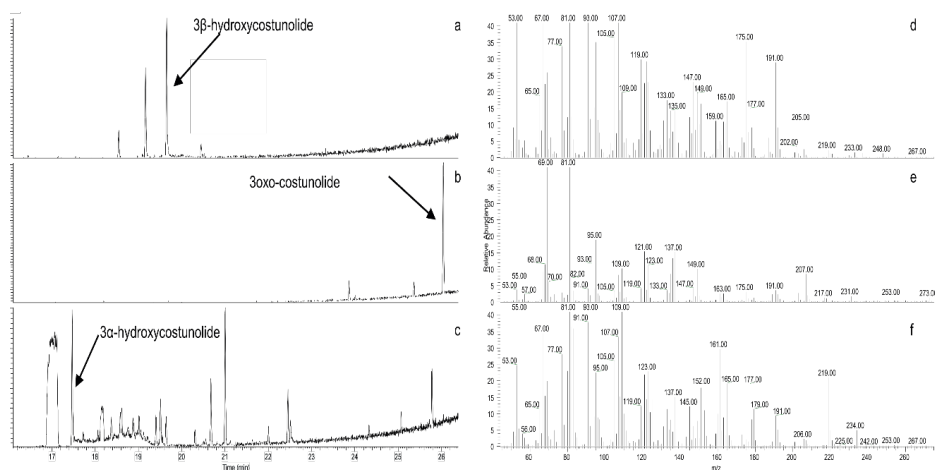
Supplementary Figure 3. Feeding 3βhydroxycostunolide to *Tp8879* (KLS) does not lead to kauniolide formation. (a) Chromatogram of enzymatic reaction mixture of kauniolide-cysteine and (b) feeding 3β-hydroxycostunolide to *Tp8879* (*TpKLS*) incubated with cysteine.



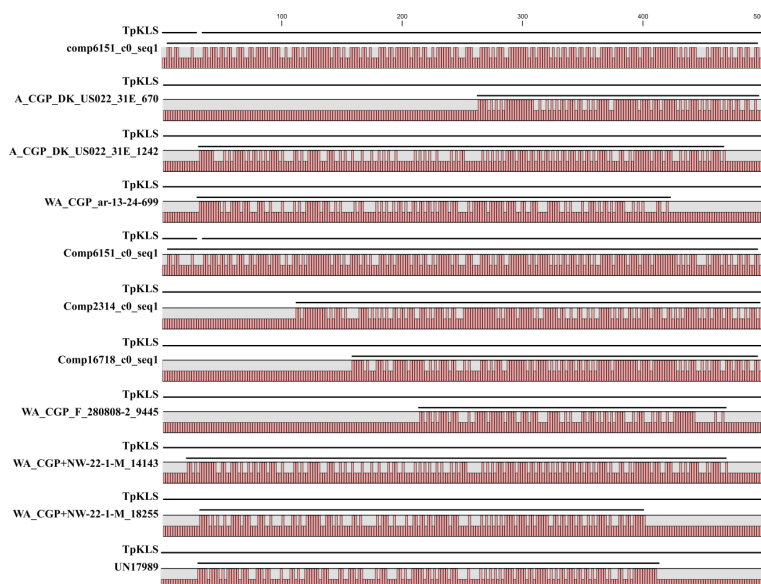
Supplementary Figure 4. A view at the active site of the kaunilide synthase model with a costunolide docking pose. C3 of costunolide approaches the iron-oxo species well. Aspartic acid 113 is located $\sim 7\text{\AA}$ from C14 of costunolide. A water bridge might be formed within the 7\AA . (Asparagine 283 in between aspartic acid and costunolide is also found in the other homology models.) The (walleyed stereo) image was taken in Maestro (Schrödinger).



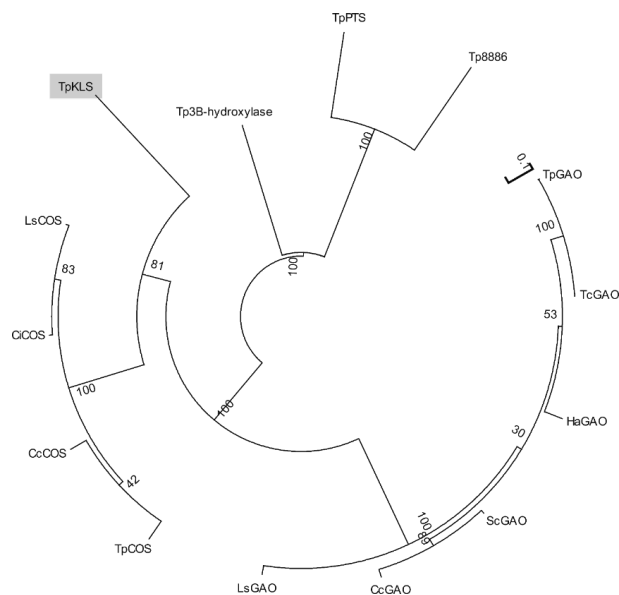
Supplementary Figure 5. Plotting preferred docking orientation of costunolide with PELE software. (a) Distance distribution of costunolide carbons relative to the heme-oxyanion in a *TpPTS* homology model (unconstrained). **(b)** Costunolide carbon distance distribution relative to the heme-oxyanion in a *TpPTS* homology model (constrained model (15\AA)).



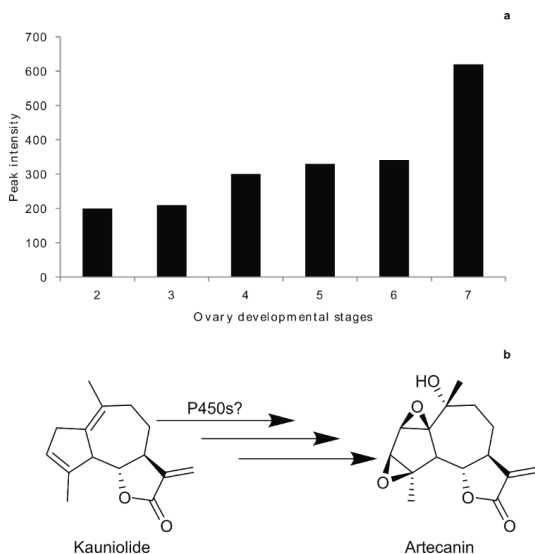
Supplementary Figure 6. Chemical synthesis of 3α-hydroxycostunolide. (a) GC-MS chromatogram of 3β-hydroxycostunolide. (b) GC-MS chromatogram of 3-oxocostunolide, obtained after oxidation with activated MnO₂. (c) GC-MS chromatogram of the reduction product of 3-oxocostunolide, 3α-hydroxycostunolide. The peak at RT = 17.80 is derived from 3α-hydroxycostunolide; note the very broad peak from RT = 17.50-19.50 min, characteristic for on-column Cope rearrangement of some germacrenes (e.g. see De Kraker, J.-W., Franssen, M.C.R., de Groot, Ae., König, W.A. & Bouwmeester, H.J. (+)-Germacrene A biosynthesis. The committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. *Plant Physiol.* **117**, 1381-1392 (1998).



Supplementary Figure 7. Homology analysis Asteraceae transcripts with *Tanacetum parthenium* kauniolide synthase. More than 2.3 million Asteraceae transcripts (<http://compgenomics.ucdavis.edu/>) analysed and highest ranking candidates with *TpKLS* were individually aligned with *TpKLS*.



Supplementary Figure 8. Phylogenetic tree of CYP71 genes from Asteraceae with known function. Tp: *Tanacetum parthenium*; Ci: *Chicorium intybus*; Cc: *Cynara cardunculus*; Tc: *Tanacetum cinerariifolium*; Ls: *Lactuca sativa*; Ha: *Helianthus annuus* and Sc: *Saussurea costus*. Bootstrap values (×10) are from 1000 replicates.



Supplementary Figure 9. *Tp8886* adds an oxygen group to kauniolide which might be an intermediate in artecanin biosynthetic pathway. (a) Accumulation pattern of artecanin during ovary developmental stages of feverfew flowers. **(b)** Proposed biosynthetic pathway of artecanin. Several P450 enzymes would be involved in biosynthesis of artecanin from kauniolide.

Supplementary Tables

Supplementary Table 1. Contig names and Asteraceae species from which the contigs were selected in Supplementary Figure 6.

| Contig name | Plant species |
|-------------------------|--|
| comp6151_c0_seq1 | <i>Cichorium endivia</i> Collection ID CHE-3178 |
| A_CGP_DK_US022_31E_670 | <i>Centaurea diffusa</i> Collection ID DK_US022-31E |
| A_CGP_DK_US022-31E_1242 | <i>Centaurea diffusa</i> Collection ID DK_US022-31E |
| WA_CGP_AR-13-24_669 | <i>Centaurea solstitialis</i> Collection ID AR-13-24 |
| Comp2314_c0_seq1 | <i>Leontodon</i> |
| Comp16718_c0_seq1 | <i>Leontodon</i> |
| WA_CGP_F_280808-2_9445 | <i>Cirsium arvense</i> Collection ID 280808-2 |
| WA_CGP+NW-22-1-M_14143 | <i>Cirsium arvense</i> Collection ID 280808-2 |
| WA_CGP+NW-22-1-M_18255 | <i>Cirsium arvense</i> Collection ID 280808-2 |
| UN17989 | <i>Cynara cardunculus</i> Collection ID ATL |

Supplementary Table 2. Query – template alignments where the homology models have been built from.

| Query (seq. identity; coverage) | Final query-template alignment |
|--------------------------------------|---|
| Kauniolide synthase (18% I; 95% C) | Kauniolide Synthase, 495 residues MALYITFLFIVSSLVLFYFFVLNQKPKGKLP PGP PPKLP IIGNIPQVAGKLP HHVLRDLARKYGP VMHLQLGHLSTIVVSSPRLAEHVLTNDLAVSNRPYSLVGDDVLYGGSDVVF~GNYGDYW RQMKKIMTTEALSA~~~~~KKVREFSGIRDHEINNMIEFIRSTLGKPFHLREGVMQRNNNIICK ALFGDHSKQQ~DLLIEIVEELVVLASGFQLADFFP~~~~~KLKFLT AISGMSKSLTKVHNELD NIFDELFRERKIKROT~NGATEDDLLD~VLFNIKER~~~~~GGLQFPEDNNIAIFVNMFIGGTD SVVTIEWTMTQMMRFPEVMKKAQAEVRRVFKGKQTITEKDLEQLVYLRCVVEALRLYAPI PILLPRESREKFQIDGYDIPVGTRVLVNAYACSTDPEYWDADSFKPERFEKSAVDFMGRNYE YLPFGTGRRICPGITFGLNVAEIIIAKLIYHFDWELPNGLSPKIDIDLSNFVGVVADKKVPLEIIP RYYPM |
| | 2PQ5_A, 471 residues ~~~~~SSKGKLP PGP TPLPFIGNYLQNT EQMYNSLMKISERYGPVFTI HLGPRRVVVL CGHDAVREALVDQAE EFSGRGEQATFDWV~FKGYGVVFSN~~GERAKQLR RFSIATLRDFGVGKRG~~~~~IEERIQEEAGFLIDALRGTTGGANIDPTFFLSRTVSNVISSIVFGDR FDYKDKE~~~~~FLSLRMMLGIFQFTSTSTGQLYEM~~~~~FSSVMKHLPGPQQQAFQLQGLE FIAKKVEHNQRTLDPN SPRDFIDSFLIRMQEEKNP~~~NTEFYLKNLVMTTLLQFIGGTETVST TLRYGFLLLMKHPEVEAKVHEIDRVIGKNRQPKFEDRAKMPYMEAVIHEIQRFQDVIPMSL ARRVKDKTKFRDFFLPKGTEVYPMLGSLVRDPSFFSNPQDFNPQHFLNEKGQF~KKSDAFVFP SIGKRNCFGEGLARMELFLFFTTVMQNFR LKSSQSPKIDIDVSPKHVGFATIPRNYTMSFLPRX HHH |
| | 3β-Hydroxylase, 499 residues MFSSFETLILSFVSLFFMMIFIHSK WISSYSKMAKNLPPSPFGLPIIGNLHQLGMTPY~NSLRTLA HKYGSLMLIHLGSPVPVIVASSAEAAQEIMKTHDQIFSTRPKMNIAIVSFDAKIVAFSPYGEHW RQSKSVYLLNLLSTKR~VQSF RHVREDETNLMLDVIENSCGSEIDLSNMIMSLTNDVVCRIAY GRKY~~~~~Y~EDWFKELMKEVMDVLGVFVSGVNYVPSLSWIDRLSGLEGRAKYAAKQLDAF LEGVVKQHETKS NESMRDQDVVDILLETQREQASAGTPFHRDTL KALMQEMFIAGTDTTSTA IEWEISEVIKHPRVMKKLQQLDEIAQGRQRITEEDLED TQHPYLEAILKESMRLHIPVPLLPR EATHDVKVMGYDIAAGTQVLINAWMIARDPTIWEDADEFKPERFLDTNIDYKGLNFELLFPF AGRRGCPGIQFAMSVNKLALANLVYKFDKLPNGLRLEQLDMT DSTGITVRRKYPLLVIPTA RF |
| 3β-Hydroxylase (95% I; 95% C) | 1NR6_A, 472 residues ~~~~~MAKKTSSKGKLP PGP TFPPIIGNILQIDAKDISKSLTKFSECY GPVFTVYLGMPKPTVVLHG YEAVKEALVDLGEEFAGRGSPVILEKVS~KGLGIAFS~NAKTWK EMRRFSLMTRLNFGMGKRSIEDRIQEEARCLVEELRKTNASPCDPTFILGCAPCNVICSVIFHN RFDYKDE~EFLKLMESLHENVELLGTPWLQVYNNFPALLDYFPGIHKTL LKNADYIKNFIMEK VKEHQKL~LDVNNPRDFIDCFLIKMEQENNL~~EFTLESLVIAVSDFLGAGTETTSTTLRYSL LLLKHPEVAARVQEEIERVI~GRHRSPCMQDRSRM~PYTDAVIHEIQRFIDLLPTNLPHAVTRD VRFRNYFIPKGTDIITSLTSVLHDEKAFNPVKVDPGHFLDESNGFKKSDY~FMPFSAGKRMCV GEG LARMELFLFLT SILQNFKLQSLVEPKDL DITAVVNGFVSVPPSYQLCFPIHHH |
| Parthenolide Synthase (26% I; 93% C) | Parthenolide Synthase, 506 residues MDTSTSFP SLFLPTLCTILISYIIHKYVLIWNRSSMAAFNLPPSPKLP IIGNIHVFSKNV NQTLW |

KLKKYGPVMLIDTGAKSFLVVSSQMAMEVLKTHQEILSTRPSNEGTKRLSYNFSIDITFSPH
GDHWDRMRKV~NEFLGPKRAGWVFNQVLRMEIKDVINNLSNPLNTSINLNEMLLSLVY
RVVCKFAFGKSYREEPFNGVTLKEMLDESMVVLGSSADMFTFGWILDKLYGWNDRLK
FGNLDGFFEMIINEHLQSASETSEDEKDFVHS~LVELSL~KDPQFTKDYIKALLNV~LLGA~I
DTTFTTIVWAMSEIVKNTQVMQKLQTEIRSCIGRKEEVDATDLTNMAYLKMVIKETLRLHPP
APLLFPRECPCHKIGGYDVFPGTCVVMNGWGIARDPNVWKEIPN~EFYPERFENFNIDFLG
NHCEMIPFGAGRRSCPGMKSATSTIEFTLVNLLYWFDWEVPSGMNNQDLME~EDGFLVIQ
KKSPLFLIPIKH

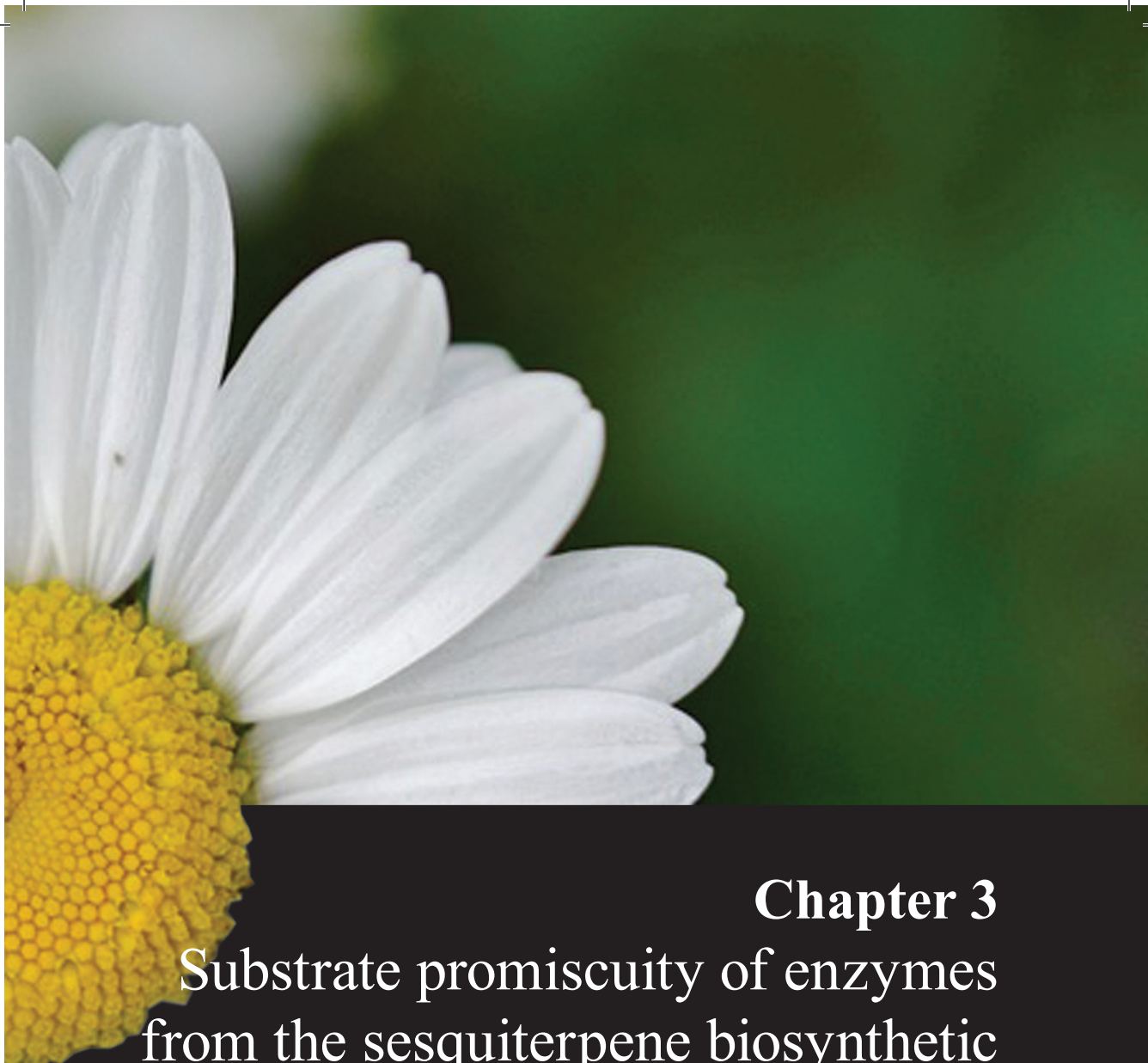
4GQS_A, 470 residues

~~~~~SSGRGKLPPGP7PLPVIGNILQIDIKDVSLSLTNLSK  
IYGPVFTLYFGLERMVVLHGYEVVKEALIDLGEFFSGRGHFPLAERANRGFG~IVFS~NGKRW  
KEIRRFSLMTLRNFGMGKRS~IEDRVQEEARCLVEELRKTAS~PCDPTFILGCAPCNVICSIF  
QKRFDYKDQQFLNLMKLNENIRIVS7PWIQICNNFPTIIDYFPGTHNLLKNLAFMESDILEK  
VKEHQESMD~INNPRDFIDCFLIKMEKEKQNNQSEFTIENLVITAADLLGAGTETTSTTLRYA  
LLLLLKHPEVTAKVQEEIERVVGRNRSPCMQDRGHMPYTDVVHEVQRYIDLPTSLPHAVT  
CDVKFRNYLIPKGTILTSLTSVLHDN~KEFPNPEMFDPRHFLDEGGNFKKSNY~FMPFSAG  
KRICVGEGLARMELFLFTFILQNF~NLKSLIDPKDLDT7PVVNGFASVPPFYQLCFPIHH~

---







## Chapter 3

Substrate promiscuity of enzymes from the sesquiterpene biosynthetic pathways from *Artemisia annua* and *Tanacetum parthenium* allows for novel combinatorial sesquiterpene production

Arman Beyraghdar Kashkooli, Alexander van der Krol, Harro Bouwmeester

In preparation for submission

## Abstract

The specific spatial configuration of a terpene may have a big effect on its therapeutic properties. However, stereo-specific chemical modifications of terpenes are very challenging and for stereo-specific modifications we mostly rely on enzymes. Therefore, isolation and characterisation of enzymes which are involved in (stereo)specific modification of medicinal terpenes is one of the goals of metabolic engineering. Here we exploit the stereo-specific modification by the *Artemisia annua* enzyme artemisinic aldehyde  $\Delta^{11}(13)$  double bond reductase2 (DBR2) on products of the feverfew sesquiterpene biosynthesis pathway. This allowed us to produce dihydroparthenolide, a compound that has been reported in feverfew extracts and displays anti-bacterial activity against the multidrug-resistant *Mycobacterium tuberculosis*. In addition, we demonstrate dihydroparthenolide biosynthesis *in planta* and *in vitro* and underpin the preferred order of biosynthesis, that is by reduction of the exocyclic methylene of parthenolide rather than through C4-C5 epoxidation of dihydrocostunolide. Moreover, the promiscuity of CYP71CB1 from feverfew for production of 3 $\beta$ -hydroxy-dihydrocostunolide or 3 $\beta$ -hydroxy-dihydroparthenolide is described, which offers new opportunities for engineering novel sesquiterpene lactones with potential improved medicinal value.

**Keywords:** Combinatorial metabolic engineering, sesquiterpene lactone, double bond reductase, feverfew, sweet wormwood, dihydroparthenolide.



## Introduction

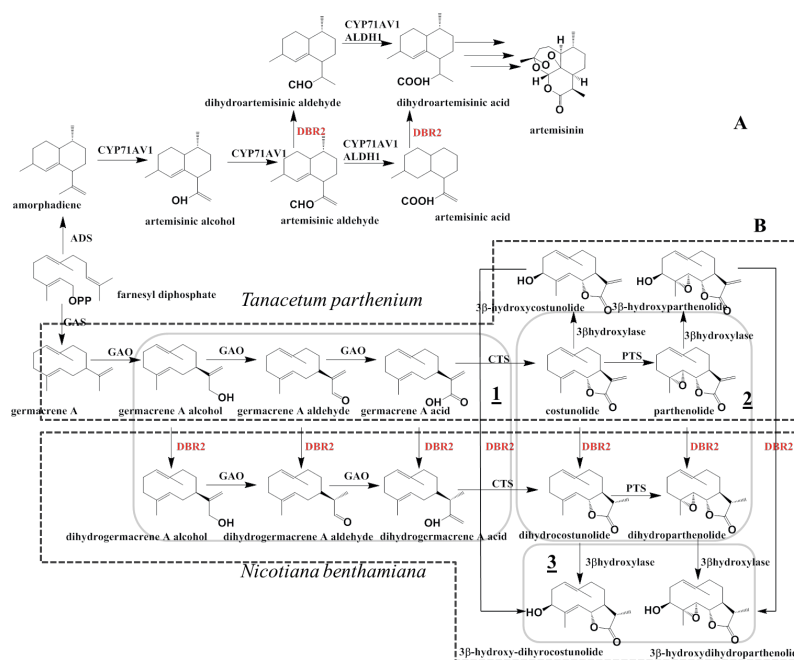
The terpenoids are a diverse and important group of secondary metabolites, which are composed of C5 isoprene units. One class of the terpenoids, the sesquiterpenoids (C15), is composed of 3 isoprene building blocks. Many of the sesquiterpenes contain a lactone group and are then called sesquiterpene lactones (STLs). STLs are mainly present in the *Asteraceae*. Feverfew, *Tanacetum parthenium*, is a well-known medicinal plant from this family, which has been used since ancient times for multiple afflictions from arthritis to insect bites, but mainly, as its name indicates, against fever. It is also used against migraines (Palevitch et al., 1997), and has been shown to have therapeutic effects on cancer (Guzman et al., 2005), by specifically inducing apoptotic cell death in cancer cells. Parthenolide (PT) is the main bioactive sesquiterpene lactone isolated from feverfew (Pareek et al., 2011). The chemical complexity of PT (C15 compound containing a macrocyclic 10C ring - with a C4-C5 epoxide - attached to an  $\alpha$ -methylene- $\gamma$ -lactone) and other STLs that have medicinal properties make them difficult targets for chemical synthesis. Most often, chemical synthesis is highly complex and expensive and it is even more difficult to get stereospecific compounds. Recent studies have demonstrated that several important bioactive compounds such as artemisinic acid (Ro et al., 2006), costunolide (CT) (Liu et al., 2011) and PT (Liu et al., 2014) can be produced in microbial and plant production platforms. Such pathway reconstruction in heterologous hosts has also opened up the possibility of combinatorial biosynthesis, in which compatible biosynthesis genes from different organisms are combined in novel biosynthesis pathways with the potential to produce novel bioactive secondary metabolites (Julsing, 2006).

Several examples of combinatorial biosynthesis of secondary metabolites are based on the substrate promiscuity of P450 enzymes, which may produce different products, depending on the substrate provided (O'Reilly et al., 2013; Eudes et al., 2016; Mafu et al., 2016). For example, the cytochrome P450, CYP71AV8, from chicory can catalyse a triple oxidation of C12 of both amorpho-4,11-diene and germacrene A, resulting in artemisinic acid and germacrene A acid, respectively (Nguyen et al., 2010), while the same enzyme catalyses a double oxidation at C2 if the precursor (+)-valencene is provided, such that nootkatone is produced (Cankar et al., 2011).

Feverfew contains several types of STLs (CT, PT, artemcanin and santamarin) (Fischedick et al., 2012). PT has already been shown to be a promising compound for medicinal applications. PT is synthesized through epoxidation of the C4-C5 double bond of the biosynthetic precursor CT (Liu et al., 2014). Although chicory, *Chicorium intybus*, does not contain PT, CT is also produced in chicory as an intermediate in the production of 11(*S*),13-dihydrocostunolide and leucodin (Kraker et al., 2002). Some STLs with reduced  $\alpha$ -methylene- $\gamma$ -lactone such as 11,13- $\beta$ -dihydrolactucopicrin are more cytotoxic than their non-reduced

forms (Ren et al., 2005). Kraker et al., (2002) reported the presence of an enzymatic activity in chicory root extracts that reduces the exocyclic double bond of CT resulting in dihydrocostunolide (DHCT). Interestingly, a reduced form of PT, dihydroparthenolide (DHPT), which shows bioactivity against multi-drug resistant *Mycobacterium tuberculosis* (causative agent of tuberculosis) and *M. avium* (Rugutt et al., 2001), is present in feverfew extracts (Fischedick et al., 2012), suggesting that both chicory and feverfew possess a so-called double bond reductase (DBR) that can reduce the exocyclic double bond of these STLs. However, the genes encoding these putative DBRs in chicory and feverfew have not been identified and characterized.

*Artemisia annua* L. is also a member of the *Asteraceae* and known for production of the antimalarial drug artemisinin. The direct precursor of artemisinin is dihydroartemisinic acid (Figure 1). An artemisinic aldehyde  $\Delta^{11}(13)$  double bond reductase<sup>2</sup> (DBR2) is responsible for reduction of the exocyclic double bond of artemisinic aldehyde to form dihydroartemisinic aldehyde, the precursor of dihydroartemisinic acid. In contrast to chicory and feverfew, DBR2 from *A. annua* has been characterized and the gene has been cloned (Zhang et al., 2008).



**Figure 1.** Biosynthesis pathway of artemisinin in *Artemisia annua* (A) and proposed combinatorial biosynthesis pathway of dihydrocostunolide and dihydroparthenolide in *Tanacetum parthenium* (B). ADS: amorphadiene synthase; CYP71AV1: amorphadiene oxidase; ALDH1: alcohol dehydrogenase; DBR2: double bond reductase; GAS: germacrene A synthase; GAO: germacrene A oxidase; CTS: costunolide synthase; PTS: parthenolide synthase.

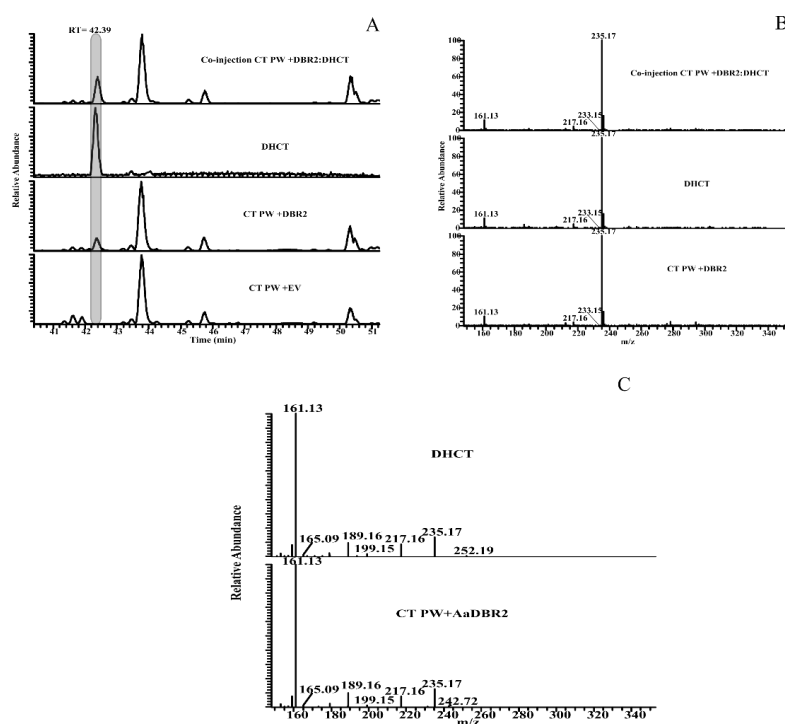
Here we tested whether DBR2 from *A. annua* can act on different STLs and pathway intermediates from feverfew to explore the options to produce novel STLs with potential medicinal value. DBRs have been shown to act on early (the aldehyde in *A. annua*), but also later intermediates (CT in chicory) of STL biosynthesis (Figure 1B). We studied whether DBR2 from *A. annua* (AaDBR2) can also reduce the exocyclic double bond of feverfew STLs and pathway intermediates, and if so whether this occurs in early or late pathway intermediates, and what the products formed are. We discuss how specific reduction of the methylene exocyclic double bond of STLs offers new opportunities for engineering novel compounds with potentially improved medicinal value.

## Results

### Functional Characterization of AaDBR2 in the CT pathway

Incubation of CT with a chicory root extract results in the formation of DHCT in an NADPH dependent manner (Kraker et al., 2002). We supposed that DHCT, a germacranolide type STL, is formed through the action of a reductase. DHCT and its C4-C5 epoxide (DHPT) are detected in chicory and feverfew extracts, respectively, but a reductase that is responsible for the exocyclic double bond reduction in these compounds has not been identified in these two species. Hence we tested DBR2 from *A. annua*, which is active in a closely related biosynthesis pathway (artemisinin production). AaDBR2 reduces the exocyclic double bond of artemisinic aldehyde, resulting in the formation of a precursor of artemisinin, dihydroartemisinic aldehyde (Zhang et al., 2008; Ting et al., 2013). Here we tested whether AaDBR2 can also reduce the exocyclic double bond of CT. For this the CT biosynthesis pathway genes (*GAS*, *GAO* and *CTS*) were co-expressed with *AaDBR2* using transient expression in *Nicotiana benthamiana* by agro-infiltration. The pathway activity was boosted by co-expression of the *Arabidopsis thaliana* 3-Hydroxy-3-Methyl-Glutaryl-coenzymeA Reductase (*AtHMGR*). AtHMGR can boost mevalonate production which leads to higher FPP precursor input for STL production in *N. benthamiana* (Liu et al., 2014). Leaves were harvested 5 days after agro-infiltration (5dpi) and metabolites were extracted with MeOH. A non-targeted metabolite profiling by LC/Q Exactive MS was used for product detection. Comparison of the chromatograms of samples that expressed the CT pathway genes with and without AaDBR2 revealed a new peak eluting at 42.32 min ( $[M+H]^+=235.169$ ) (Figure 2). This mass differs 2.01 Dalton from that of CT ( $[M+H]^+=233.153$ ), which fits with a double bond reduction in CT. In *N. benthamiana* CT is detoxified by addition of glutathione (GSH) resulting in CT<sup>GSH</sup>, which over time may breakdown to CT-cysteine conjugate (CT<sup>C</sup>). This GSH conjugation supposedly takes place at the methylene group adjacent to the carbonyl group of the lactone ring of CT (Liu et al., 2011). Expression of the CT pathway genes with *AaDBR2* in *N. benthamiana* resulted in a 70% reduction in CT<sup>GSH</sup> and CT<sup>C</sup> levels, providing

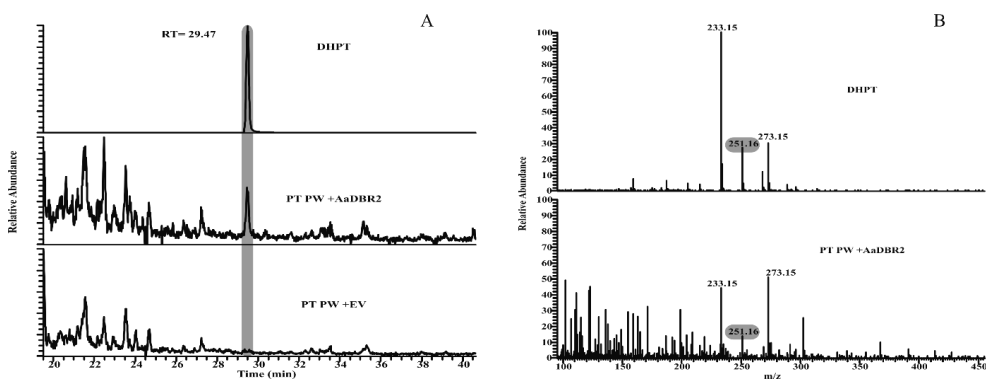
additional support for AaDBR2 activity on CT. Since a commercial standard of DHCT is not available, we used chemical synthesis to produce DHCT to confirm the identity and stereochemistry of the novel product in *N. benthamiana* (Kraker et al., 2002). Optimization of chemical synthesis required control of neutral pH and resulted in synthesis of 11(*S*),13-dihydrocostunolide. This compound has the same mass and mass spectrum and eluted at the same RT as the novel product produced in *N. benthamiana* (Figure 2). Co-injection of the synthetic standard and a *N. benthamiana* extract confirmed that they are the same (Figure 2A,B). These results show that AaDBR2 can indeed reduce the exocyclic double bond of CT to produce DHCT. Next, we investigated whether AaDBR2 can also be used for DHPT production in this transient expression platform.



**Figure 2.** Functional characterization of AaDBR2 in *N. benthamiana*. (A). Q-ExactivePlus Orbitrap LC-MS/MS chromatograms of *N. benthamiana* leaves agroinfiltrated with CT PW (AtHMGR, TpGAS, CiGAO, CiCTS) with or without AaDBR2 (bottom chromatograms), dihydrocostunolide (DHCT) and co-injection of the DHCT standard and CT PW+DBR2 (top chromatogram). DHCT was formed where CT PW was expressed together with AaDBR2. pBIN<sup>+</sup> vector was used instead of AaDBR2 as the control. (B) MS spectrum of eluting peak at 42.39 min for CT PW +DBR2, DHCT standard and co-injection (CT PW +DBR2:DHCT). (C) MS/MS fragmentation of peak eluting at 42.39 min for DHCT standard (top chromatogram) and CT PW+AaDBR2 (bottom) shows similar fragmentation pattern. CT PW: costunolide biosynthesis pathway, HMGR: 3-hydroxy-3-methylglutaryl coenzyme A synthase, GAS: germacrene A synthase, GAO: germacrene A oxidase, CTS: costunolide synthase, DBR2: double bond reductase.

## Reduction of PT by AaDBR2

In order to test whether AaDBR2 can act on PT, all the PT biosynthesis pathway genes (*GAS*, *GAO*, *CTS* and *PTS*) (Liu et al., 2014) were co-expressed with *AaDBR2*. Again, boosting of FPP production was achieved by co-expression of *AtHMGR1*. Agro-infiltrated leaves were harvested at 5dpi and methanolic extracts were injected into LC-Orbitrap-FTMS to analyse product formation. Comparison of the chromatograms of extracts from samples expressing the PT pathway with and without *AaDBR2* showed a new peak at RT = 29.47 min ( $[M+H]^+ = 251.16$ ). The mass of the new compound fits that of DHPT and the retention time shift between PT and the putative DHPT of about 0.4min is similar to the retention time shift between CT and DHCT (Supplementary Figure 1). The identity of the novel compound was confirmed by comparison with an authenticated DHPT standard, isolated from feverfew extracts (Fischedick et al., 2012) (Figure 3).

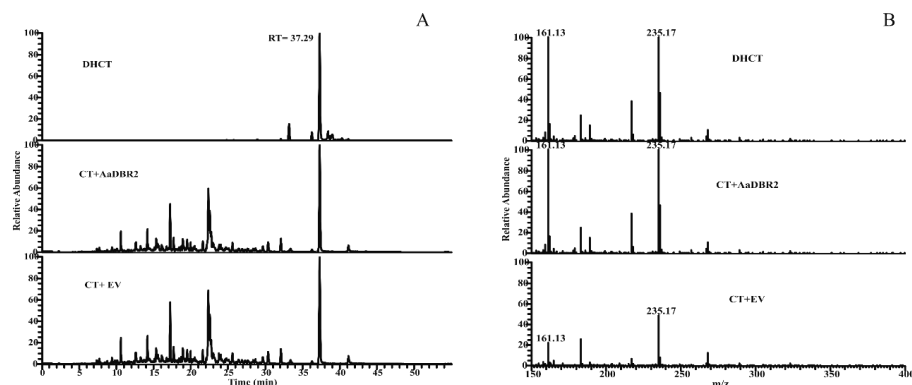


**Figure 3.** Functional characterization of AaDBR2 in the context of PT PW. (A) LC-Orbitrap-FTMS chromatograms of PT PW (*AtHMGR*, *TpGAS*, *CiGAO*, *CiCTS* and *TpPTS*) without AaDBR2 (bottom), PT PW with AaDBR2 (middle) and dihydroparthenolide (DHPT) standard (top). A similar peak to the DHPT standard was eluted at 29.47 min in the PT PW+AaDBR2 which was not in control samples. (B) in source spectrum of DHPT (top) and PT PW+ AaDBR2 of the eluting peak at 29.47 min. The grey highlighted box is indicating the parent ion  $[M+H]^+ = 251.16472$ . PT PW: parthenolide biosynthesis pathway, HMGR: 3-hydroxy-3-methylglutaryl coenzyme A synthase, GAS: germacrene A synthase, GAO: germacrene A oxidase, CTS: costunolide synthase, PTS: parthenolide synthase, DBR2: double bond reductase.

## AaDBR2 can act directly on CT

Formation of DHCT by co-expression of *AaDBR2* and the CT pathway genes in the experiment described above may occur through reduction of CT as suggested before (Kraker et al., 2002), but could also occur through reduction of CT pathway intermediates produced by GAS and/or GAO (Figure 1B-1), particularly considering that the native function of AaDBR2 is to reduce AAA. If AaDBR2 forms double-bond reduced CT pathway intermediates, such products may subsequently be converted by GAO and CTS to DHCT

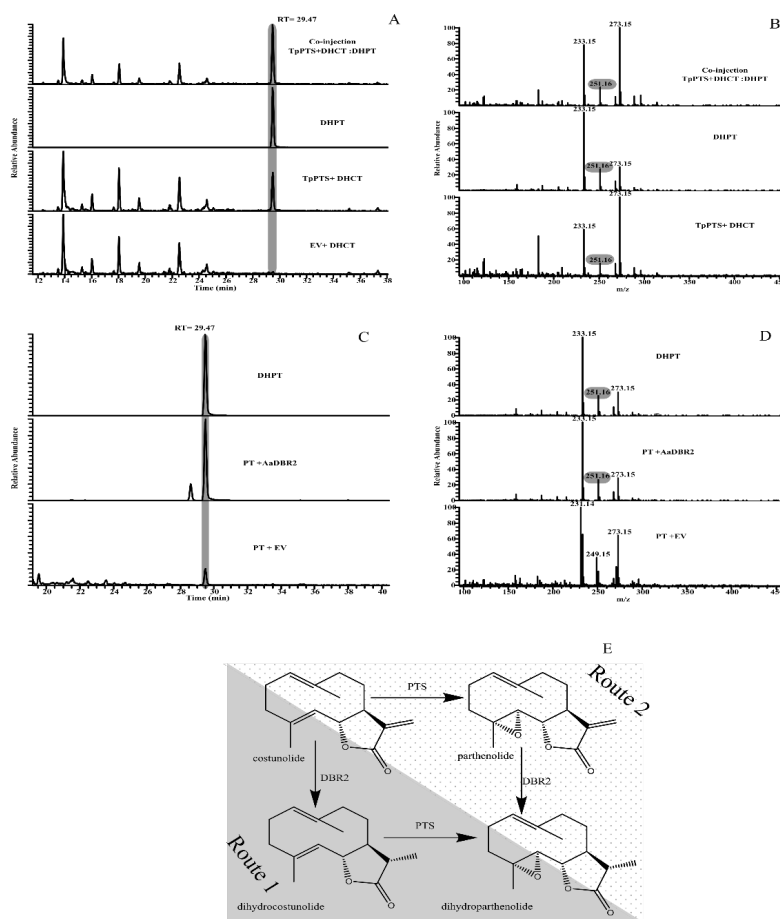
(Figure 1B-2), provided these enzymes are promiscuous for dihydro-substrates. To test whether AaDBR2 acts on CT we supplied CT to *N. benthamiana* leaf discs transiently expressing only AaDBR2. At three dpi, leaf discs were taken from agro-infiltrated leaves and vacuum infiltrated with water containing CT. The infiltrated leaf disks were harvested after 2h and extracted with methanol for product analysis by LC-Orbitrap-FTMS. Product analysis showed that DHCT was formed, indicating that AaDBR2 acts on CT or one of its precursors (Figure 4). Control experiments with leaf discs infiltrated with EV showed very low level conversion of CT to DHCT (~60% less active than AaDBR2), suggesting the presence of some endogenous reductase activity with low activity on CT in *N. benthamiana*.



**Figure 4.** Functional characterization of AaDBR2 to produce DHCT from costunolide (CT) in *N. benthamiana*. (A) LC-Orbitrap-FTMS chromatograms of *N. benthamiana* leaf extracts. *N. benthamiana* leaves were agroinfiltrated with AaDBR2 or pBIN+ (EV as the control). After 3 days (3dpi) a CT solution was vacuum infiltrated and incubated for 2h. chromatograms representing feeding CT to discs expressing EV (bottom), feeding CT to discs expressing AaDBR2 (middle) and DHCT standard. DHCT, regardless of presence of AaDBR2 was produced in *N. benthamiana* leaves. (B) DHCT standard and eluted peak at 37.29 min for CT feeding to AaDBR2 and EV demonstrating identical parent ions and fragmentation pattern at  $[M+H]^+ = 235.16981$ . DBR2: double bond reductase2, DHCT: dihydrocostunolide.

### AaDBR2 can act directly on PT

Formation of DHPT by co-expression of AaDBR2 with the PT pathway genes in *N. benthamiana* may be the result of promiscuous activity of TpPTS oxidising both CT and DHCT (Figure 5E, route 1) or by promiscuous activity of AaDBR2, reducing also the double bond of PT (Figure 5E, route 2). To test the promiscuity of TpTPS we used a yeast microsome feeding assay. Yeast WAT11 was transformed with TpPTS and used for isolation of the microsomal fraction which contains the TpTPS. The microsomal fractions were subsequently incubated with either DHCT or CT for 2.5 hours. As control, microsome fractions from yeast expressing EV were incubated with CT or DHCT. Product analysis of these *in vitro* enzyme assays by LC- Orbitrap-FTMS showed that PTS can use both CT and DHCT as substrate, resulting in PT and DHPT, respectively (Figure 5A and 5B).



**Figure 5.** Bidirectional characterization of DHPT biosynthesis in *N. benthamiana* and yeast. (A) LC-Orbitrap-FTMS (positive ionisation mode) chromatograms of yeast microsomes enzymatic mixture extracts to investigate dihydroparthenolide (DHPT) *in-vitro* biosynthesis through enzymatic epoxidation of dihydrocostunolide (DHCT) by the activity of TpPTS (parthenolide synthase) (Figure 5E, route 1). From bottom; feeding DHCT to yeast microsomes expressing EV, feeding DHCT to yeast microsomes expressing parthenolide synthase (TpPTS), DHPT standard and co-injection of TpPTS+DHCT: DHPT. A peak eluting at RT= 29.47 was present in mixtures of feeding assay of DHCT to TpPTS which was similar to DHPT standard. (B) Spectrum of produced compound from feeding assay of DHCT to TpPTS (bottom), DHPT standard (middle) and co-injection of TpPTS+DHCT:DHPT (top). (C) LC-Orbitrap-FTMS (positive ionisation mode) chromatograms of in-vivo synthesis of DHPT in *N. benthamiana* through enzymatic reduction PT by the activity of AaDBR2. A highly concentrated solution of PT was vacuum infiltrated to the leaf disks of *N. benthamiana* expressing AaDBR2 or EV (control) 3 days after agroinfiltration. A peak eluting at RT= 29.47 was present in samples with or without AaDBR2. Peak intensity of this peak was ~95% less than what was produced in the presence of AaDBR2. (D). Spectrum of produced compound eluting at 29.47 min from leaf feeding assay experiments. Spectrum of feeding PT to cells expressing AaDBR2 (middle) was identical to the DHPT standard (top). The fragmentation pattern of product of feeding PT to EV was not completely identical, however the representative mass to DHPT standard was still present. (E) proposed bidirectional biosynthesis of DHPT in feverfew plants. Route 1, suggests biosynthesis of DHPT, through reduction PT by AaDBR2 and route 2 proposes reduction of CT through the activity of AaDBR2 which later epoxidases (C4-C5) by the activity TpPTS.

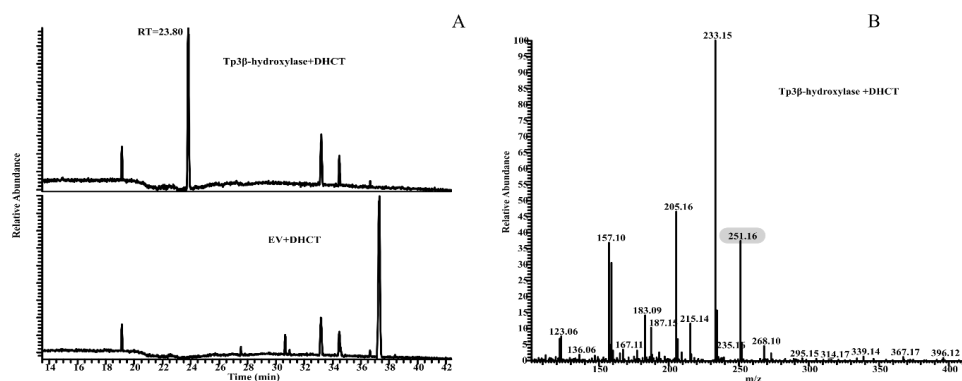


No PT or DHPT was detected in microsomes from yeast expressing EV, incubated with CT or DHCT, respectively. The conversion of DHCT to DHPT by PTS is not very efficient as only ~2.5% of the infiltrated DHCT was converted to DHPT.

To test whether AaDBR2 can use PT as substrate (Figure 5D, route 2) *N. benthamiana* leaves were infiltrated with *A. tumefaciens* harbouring the *AaDBR2* expression construct. Three dpi, discs were taken from agro-infiltrated leaves and infiltrated with water containing PT. Vacuum infiltration of PT in discs from leaves transformed with EV was used as negative control. Leaf discs were incubated for 2 hours after which they were extracted with MeOH and analysed by LC-Orbitrap-FTMS. Results showed that AaDBR2 is catalysing the exocyclic double bond reduction of PT (Figure 5C and 5D). A tiny peak, eluting at the same time as the authentic DHPT standard, was detected in leaves expressing EV (Figure 5C). However, the mass spectrum of this product contains fragment masses  $m/z=231.16$ ,  $m/z=249.16$  and  $m/z=273.15$  of which only one ( $m/z=273.15$ ) fits with fragment masses of authentic DHPT, suggesting that this is likely not DHPT (Figure 5D). More than 90% of infiltrated PT was converted by AaDBR2 compared to the activity in leaves expressing EV (Figure 5C). This shows that PT is a good substrate for AaDBR2.

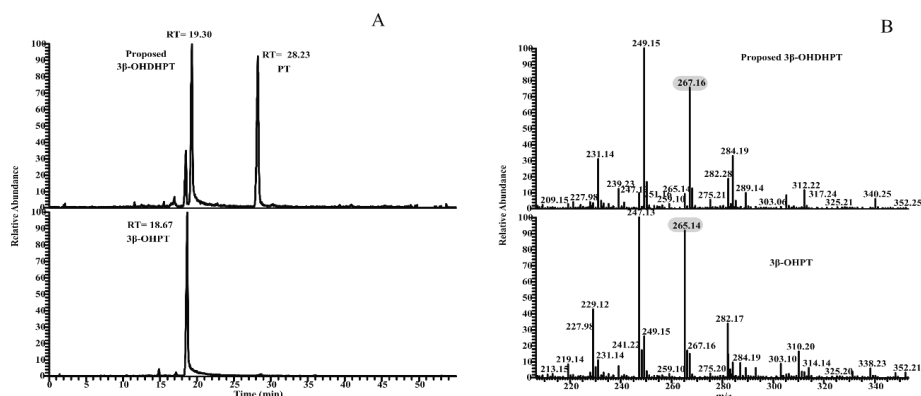
### Characterization of substrate promiscuity of Tp3 $\beta$ -hydroxylase

Feverfew also contains a Tp3 $\beta$ -hydroxylase which catalyses the stereo-specific 3 $\beta$ -hydroxylation of CT and PT to form 3 $\beta$ -hydroxycostunolide (3 $\beta$ -OHCT) and 3 $\beta$ -hydroxyparthenolide (3 $\beta$ -OHPT), respectively (Liu et al., 2014). We assessed whether Tp3 $\beta$ -hydroxylase also accepts DHCT as a substrate, just as TpPTS that can oxidise both CT and DHCT. Hereto, microsomal fractions isolated from yeast expressing Tp3 $\beta$ -hydroxylase were incubated for 2 hours with either DHCT or DHPT. Products of the *in vitro* enzyme assay were extracted and analysed by LC-Orbitrap-FTMS. The resulting chromatograms show a new peak eluting at 23.80 min ( $[M+H]^+= 251.16472$ ) in samples from yeast expressing Tp3 $\beta$ -hydroxylase enzyme incubated with DHCT. This peak was not present in the control samples where DHCT was fed to microsome fractions from yeast expressing EV (Figure 6). The mass of this novel compounds fits with the expected mass of 3 $\beta$ -hydroxy-dihydrocostunolide (3 $\beta$ -OH-DHCT). Also the retention shift of this new peak compared with DHCT is similar to the retention shift between CT and 3 $\beta$ -OHCT (~0.4 min, Supplementary Figure 1). A commercial standard of 3 $\beta$ -OH-DHCT is not available hampering unambiguous confirmation of the identity of this new compound as 3 $\beta$ -OH-DHCT.



**Figure 6.** Production of tentative 3 $\beta$ -hydroxy-dihydrocostunolide (3 $\beta$ -OH-DHCT) from DHCT by the activity of Tp3 $\beta$ -hydroxylase. (A) LC-Orbitrap-FTMS (positive ionisation mode) chromatograms of yeast microsomes expressing Tp3 $\beta$ -hydroxylase or EV. Chromatogram of enzymatic mixture extracts of feeding DHCT to Tp3 $\beta$ -hydroxylase showed a peak eluting at 23.80 min ( $[M+H]^+ = 251.16472$ ) which was 0.4 min delay from the produced 3 $\beta$ -hydroxycostunolide (3 $\beta$ -OHCT). (B) fragmentation spectrum of eluted peak at 23.80 min. grey highlighted box indicates the parent ion representing 3 $\beta$ -OH-DHCT. DHCT: dihydrocostunolide.

Tp3 $\beta$ -hydroxylase also acts on PT to form 3 $\beta$ -OHPT (Liu et al., 2014) and we tested whether DHPT is also substrate for this enzyme. Since the amount of DHPT we had available was low, we did not use the yeast microsome feeding assay, but produced DHPT *in planta* by feeding PT to *AaDDBR2* expressing *N. benthamiana* leaves. The Tp3 $\beta$ -hydroxylase was co-expressed with *AaDDBR2* or EV. At three dpi, PT was vacuum infiltrated into the leaf disks and discs were extracted after 2h incubation for product analysis by LC-Orbitrap-FTMS. Results showed that both in absence and presence of *AaDDBR2* a product was formed which elutes at 18.70 min ( $[M+H]^+ = 265.14399$ ), which was identified as 3 $\beta$ -OHPT, confirming the enzyme activity of Tp3 $\beta$ -hydroxylase on PT. A putative novel product formed by *AaDDBR2*, double-bond reduced 3 $\beta$ -OHPT, is expected to have mass  $[M+H]^+ = 267.15964$  and a slightly later retention time of approximately  $\sim 0.4$  min. Such product peak was indeed detected in samples expressing *AaDDBR2* (RT shift 0.65 min). A similar but smaller peak was identified in the samples only expressing EV (10-fold lower peak intensity), indicating that an endogenous *N. benthamiana* enzyme may also catalyse this double bond reduction (data not shown). Because we lack the standard for 3 $\beta$ -hydroxy-dihydroparthenolide (3 $\beta$ -OH-DHPT) we cannot unambiguously confirm product identity. However, the five main peaks in the mass spectrum of the novel product, were similar to the five main peaks in the mass spectrum of 3 $\beta$ -OHPT, with a 2 D reduction in the  $m/z$  for each of the five peaks, suggesting that the novel compound is indeed 3 $\beta$ -OH-DHPT (Figure 7).



**Figure 7.** Production of 3β-hydroxy-dihydroparthenolide (3β-OH-DHPT) from parthenolide (PT) by the activity of AaDBR2 and Tp3β-hydroxylase in *N. benthamiana*. (A) LC-Orbitrap-FTMS (positive ionisation mode) chromatograms of 3β-hydroxyparthenolide (3β-OHPT) (bottom) for selected mass  $[M+H]^+=265.14399$  (RT=18.67) and selected mass of  $[M+H]^+=267.15964$  (RT=19.30). The chromatogram of leaves expressing Tp3β-hydroxylase and AaDBR2 and incubated with PT showed a peak eluting at 19.30 min ( $[M+H]^+=267.15964$ ) which was 0.6 min delay from the produced 3β-OHPT. (B) fragmentation spectrum of the peak at 19.30 min (top) representative of putative 3β-OH-DHPT and 18.67 min (bottom) representing 3β-OHPT spectrum. grey highlighted box indicates the parent ion representing 3β-OH-DHPT (top) and 3β-OHPT (bottom). DBR2: double bond reductase2.

## Discussion

Combinatorial biosynthesis, by co-expression of compatible terpene biosynthesis genes from different species, provides a powerful strategy to produce novel terpene structures. Also, combinatorial biosynthesis can be used to explore the full range of substrate specificities of an enzyme, which may be hidden in the original host due to limited availability of compatible substrates. Finally, combinatorial biosynthesis can be used in instances where the full complement of biosynthesis genes is lacking, but through which a biosynthesis pathway may be extended by a compatible enzyme from a different plant species.

Although sesquiterpene lactones with reduced exocyclic methylene double bond have been reported both in feverfew and chicory, full reconstruction of their biosynthesis pathway has not been achieved yet, as from neither species the full complement of sesquiterpene lactone biosynthesis genes have been characterized. In this paper we took advantage of two sets of available enzymes from *A. annua* and feverfew in a combinatorial metabolic engineering approach for biosynthesis of sesquiterpene lactones with reduced exocyclic double bond. This was possible due to substrate promiscuity for a number of the enzymes involved in sesquiterpene lactone biosynthesis in these two species.

The artemisinin biosynthesis pathway in *A. annua* L. proceeds quite similar to the CT biosynthesis pathway in chicory and feverfew (Figure 1). Indeed, it was previously

demonstrated that germacrene A oxidase (GAO) from chicory can act on the non-native substrate amorphadiene from *A. annua*, although amorphadiene oxidase (CYP71AV1) from *A. annua* cannot oxidise germacrene A (Ro et al., 2006). Here we show that AaDBR2 is able to act on products of the sesquiterpene biosynthesis pathway of feverfew. AaDBR2 can reduce the exocyclic methylene group of both CT and PT, producing DHCT and DHPT, respectively (Figure 4 and 5). These products could subsequently further be modified by CYP71CB1 (Tp3 $\beta$ -hydroxylase) of feverfew into 3 $\beta$ -OH-DHCT or 3 $\beta$ -OH-DHPT (Figure 6 and 7). We cannot distinguish whether Tp3 $\beta$ -hydroxylase hydroxylates DHCT and DHPT or whether AaDBR2 reduces the double bond of 3 $\beta$ -OHCT and 3 $\beta$ -OHPT. Similar to DHPT biosynthesis (Figure 5E), a bidirectional biosynthesis pathway for 3 $\beta$ -OH-DHCT or 3 $\beta$ -OH-DHPT may be present where both routes are active, of which presumably one is more active.

The enzymatic combinatorial biosynthesis with AaDBR2 resulted in a stereo-specific bioconversion of CT and PT, producing only the *S*-enantiomers 11(*S*),13-dihydrocostunolide and 11(*S*),13-dihydroparthenolide. In contrast, conventional chemical reduction of the double bond by NaBH<sub>4</sub> results in both the *S* and *R* enantiomers (personal communications Prof. Giovanni Appendino, Università del Piemonte Orientale). Stereo-selective enzymatic reactions are more interesting from a pharmaceutical point of view, as often just one of specific stereoisomers of a metabolite may show medicinal properties (Julsing, 2006). Results from bidirectional biosynthesis of DHPT (through DHCT by PTS or through PT by DBR) suggests a preferred biosynthetic route (Figure 5E, route 2). Efficiency of this route 2 (exocyclic double bond reduction of PT) is ~90% while efficiency of route 1 (Figure 5E) is calculated to be only ~2.5%.

#### *Product biosynthesis potential versus actual product accumulation*

Our inventory of the capacity of sesquiterpene biosynthesis enzymes and the identification of sesquiterpene lactones in feverfew (Fischedick et al., 2012) provide insight in how competition between enzymes in feverfew may regulate accumulation of individual sesquiterpenes. For instance, DHPT is detected in feverfew extracts (Fischedick et al., 2012), suggesting that feverfew has an enzyme similar to AaDBR2, which we name here TpDBR. In contrast to DHPT, the compound DHCT is not detected in feverfew (Fischedick et al., 2012), while this is produced in our *N. benthamiana* assays (Figure 2). This could be due to efficient capturing in feverfew of CT by TpPTS, preventing CT to be converted to DHCT by the putative TpDBR. Alternatively, or in addition, any DHCT may be efficiently converted to DHPT by TpPTS in feverfew or PT may be the preferred substrate over CT for the putative TpDBR. Recently we also characterized a kauniolide synthase in feverfew (TpKLS), which converts CT into the guaianolide-type sesquiterpene lactone kauniolide (Chapter 2). Possibly, TpKLS could use DHCT as a substrate, resulting in the formation of dihydrokauniolide. We

incubated DHCT with microsomal fractions of yeast expressing KLS. However, there was no decrease in the amount of DHCT and no dihydrokauniolide was detected (data not shown) suggesting that the biosynthesis of dihydrokauniolide takes place by reduction of kauniolide itself and not its precursor, CT.

We have also shown that the methylene group of PT is efficiently reduced by artemisia DBR2 in the *in planta* (*N. benthamiana*) assay, and this pathway is more efficient than the one in which first CT is reduced to DHCT and subsequently converted by PTS to DHPT (route 1, Figure 5). It could be that the putative DBR from feverfew also has a bias towards PT compared to CT as substrate. Another explanation for the preferred production of DHPT is a close association of both CTS and PTS enzymes on the ER membrane, resulting in efficient substrate channelling. This would limit accessibility of CT for the putative feverfew DBR enzyme, which is likely located in the cytosol.

We recently showed that Lipid Transfer Proteins (LTPs) may exhibit high selectivity for the transport/sequestration of specific sesquiterpene lactones to the extracellular space, hence making them unavailable for enzymatic conversion (Wang et al., 2016). Competition for feverfew pathway intermediates may thus not only be at the enzyme level, but also at the level of transport to the extracellular space. We detected up to eight different LTPs in trichomes of feverfew where sesquiterpene biosynthesis takes place. For two of these (TpLTP1 and TpLTP2) we demonstrated involvement in extracellular accumulation of CT, while TpLTP3 was highly specific for PT. Action of these LTPs in feverfew would place CT and PT out of reach of intracellular (cytosolic) enzymes for further conversion. Characterization of the putative TpDBR and other enzymes in combination with further characterization of the different feverfew trichome LTPs on pathway products may eventually provide a full understanding of product accumulation in feverfew.

## Material and methods

### Plasmid construction for gene expression in *N. benthamiana*

TpGAO (KC964544), CiCTS (JF816041), TpPTS (KC954155), 3 $\beta$ -hydroxylase (KC954153) and AaDBR2 (JX898526) were previously cloned into pIV1A1.1. TpGAS (JF819848) was cloned into impact vector 1.5 ([www.impactvector.com](http://www.impactvector.com)) between the CaMV35S promoter and the RBCS1 terminator. All individual constructs were later cloned into the plant binary vector, pBINPlus. *Agrobacterium tumefaciens* strain AGL-0 was used for agroinfiltration. Agroinfiltration was performed according to the described method by (van Herpen et al., 2010). *Agrobacterium* constructs were grown at 28°C for 48h in a shaker incubator at 250 rpm in Luria broth medium supplied with 50mgL<sup>-1</sup> kanamycin and 37.5mg.L<sup>-1</sup> rifampicin. Cells were then harvested at 3500×g for 15 min at room temperature. 20 mL of 10 mM MES

buffer (2-(N-morpholino) ethanesulfonic acid), 20 mL of 10 mM MgCl<sub>2</sub> and 1 mL of 100 mM acetosyringone (4'-Hydroxy-3',5'-dimethoxyacetophenone) (to a final volume of 1L) was used as agroinfiltration buffer to resuspend cells. The OD<sub>600</sub> of these cells were adjusted to 0.5 by addition of agroinfiltration buffer. After mixing of the cultures for co-expression studies they were incubated on a roller mix for 120 min. The mixtures were infiltrated with a 1ml syringe to the abaxial side of *N.benthamiana* fully developed leaves. Plants were kept in a day-night temperature of 22-20°C for 3 or 5 days for substrate feeding experiments or metabolites extraction, respectively.

#### **Metabolite extraction from *N.benthamiana* leaves**

Agroinfiltrated leaves were harvested and snap-frozen in liquid N<sub>2</sub> for metabolites extraction. For substrate feeding, 2 leaf disks were punched from each leaf and vacuum infiltrated with 20µM parthenolide or costunolide and kept for 2h at room temperature. Then the leaves were ground into a fine powder using a tissuelyser and glass beads. 100 mg of fine ground material was extracted with 300 µl of MeOH:Formic acid (1000:1). Samples were sonicated for 15 min and centrifuged for 15 min at 13500xg (5°C). Later the supernatant was filtered through a 0.45 µm filter (Minisart® RC4, Sartorius, Germany) for metabolites analysis by LC-MS.

#### **Yeast microsome assay and metabolites extraction**

In order to perform yeast in-vitro assays microsomal fractions (72 µl), NADPH (100 µl of 10mM stock), potassium buffer (20 µl of 1mM, pH=7.5) and 288 µl of distilled water was mixed with substrate (10 µl of 10mM stock solution). The mixture was incubated for 150 min at 25°C in a shaker incubator (250 rpm). The enzymatic mixtures were then centrifuged for 15 min at 13500xg (5°C). The supernatant was filtered through a 0.45 µm filter (Minisart® RC4, Sartorius, Germany) for metabolites analysis by LC-MS.

#### **Metabolites analysis by LC-Orbitrap-FTMS**

In order to analyse the agro-infiltrated *N. benthamiana* leaf extracts or yeast microsome enzyme assays we used a LC-LTQ-Orbitrap-FTMS system (Thermo Scientific). This system a HPLC, an Accela photodiode array detector (PDA). This was connected to an LTQ/Orbitrap hybrid MS detector coupled with an ESI source. An analytical column of Luna 3 µm C18/2 100A; 2.0 × 150 mm (Phenomenex, USA) was used for chromatographic separation. We used HPLC H<sub>2</sub>O: formic acid (1000:1, v:v) as eluent A while eluent B was combination of acetonitrile: formic acid (1000:1, v:v). Flow rate was set to 0.19 mL min<sup>-1</sup> and the gradient was from 5 to 75% acetonitrile in a 45 min gradient. This was then followed by a washing step of 15 min and later equilibrated. Full scan mass analysis was done at a resolution equal to

60000. FTMS calibration was externally done in negative ionisation mode by CHNaO<sub>2</sub>. Injection volume for each sample was set to be 5 µL.

### Chemical synthesis of 11(*S*),13-dihydrocostunolide

In order to obtain 11(*S*),13-dihydrocostunolide, 2 mg of (+)-costunolide (TOCRIS bioscience) was dissolved in 1.5 mL of ethyl acetate. Then 1.5 mg of NaBH<sub>4</sub> was added on ice based on the protocol described for reduction of costunolide and exocyclic double bond reduction of other sesquiterpene lactones (Seto et al., 1988; Kraker et al., 2002). The reaction was stopped after 45 min by addition of 1% (w/w) acetic acid and 2 ml extra of ethyl acetate was added. The organic phase was then filtered through a glass-wool plugged Pasteur pipet filled with anhydrous Na<sub>2</sub>SO<sub>4</sub> to dehydrate the organic phase.

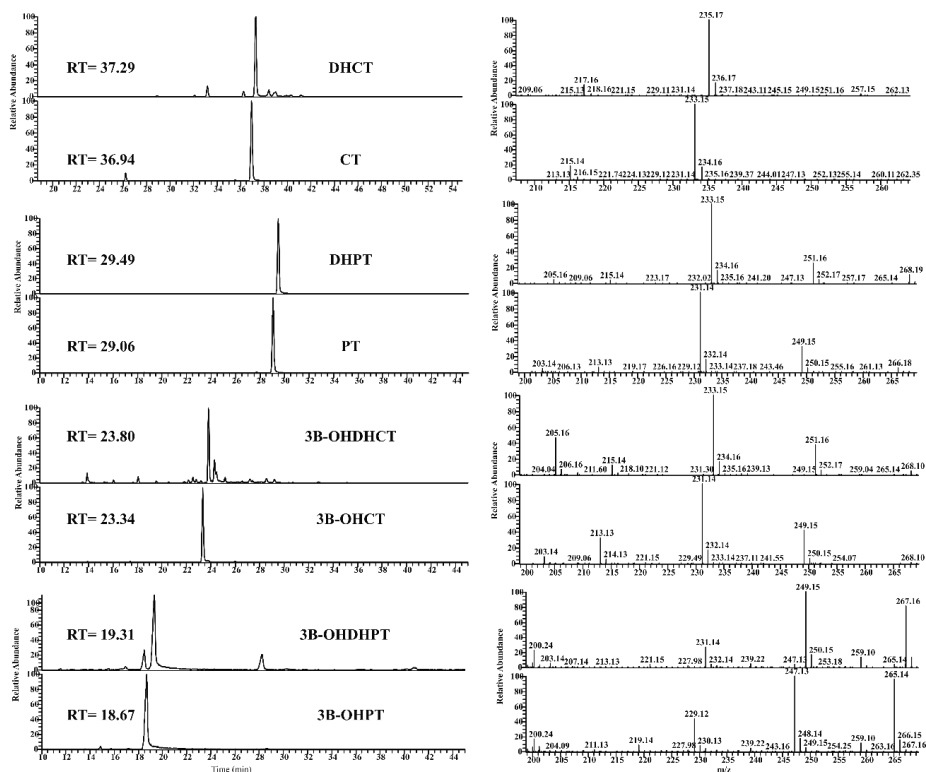
### References

- Cankar K, van Houwelingen A, Bosch D, Sonke T, Bouwmeester H, Beekwilder J (2011) A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS letters* **585**: 178-182
- Eudes A, Pereira JH, Yogiswara S, Wang G, Teixeira Benites V, Baidoo EEK, Lee TS, Adams PD, Keasling JD, Loqué D (2016) Exploiting the Substrate Promiscuity of Hydroxycinnamoyl-CoA:Shikimate Hydroxycinnamoyl Transferase to Reduce Lignin. *Plant and Cell Physiology* **57**: 568-579
- Fischedick JT, Standiford M, Johnson DA, De Vos RCH, Todorović S, Banjanac T, Verpoorte R, Johnson JA (2012) Activation of antioxidant response element in mouse primary cortical cultures with sesquiterpene lactones isolated from *Tanacetum parthenium*. *Planta medica* **78**: 1725-1730
- Guzman ML, Rossi RM, Karnischky L, Li X, Peterson DR, Howard DS, Jordan CT (2005) The sesquiterpene lactone parthenolide induces apoptosis of human acute myelogenous leukemia stem and progenitor cells. *Blood* **105**: 4163-4169
- Julsing MK (2006) Bioconversion and combinatorial biosynthesis of selected terpenoids and lignans. University of Groningen
- Kraker J-Wd, Franssen MCR, Joerink M, de Groot A, Bouwmeester HJ (2002) Biosynthesis of Costunolide, Dihydrocostunolide, and Leucodin. Demonstration of Cytochrome P450-Catalyzed Formation of the Lactone Ring Present in Sesquiterpene Lactones of Chicory. *Plant Physiology* **129**: 257-268
- 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 Q, Manzano D, Tanić N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metabolic engineering* **23**: 145-153
- Mafu S, Jia M, Zi J, Morrone D, Wu Y, Xu M, Hillwig ML, Peters RJ (2016) Probing the promiscuity of ent-kaurene oxidases via combinatorial biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 2526-2531
- Nguyen DT, Göpfert JC, Ikezawa N, MacNevin G, Kathiresan M, Conrad J, Spring O, Ro D-K (2010) Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. *Journal of Biological Chemistry* **285**: 16588-16598
- O'Reilly E, Corbett M, Hussain S, Kelly PP, Richardson D, Flitsch SL, Turner NJ (2013) Substrate promiscuity of cytochrome P450 RhF. *Catalysis Science & Technology* **3**: 1490-1492
- Palevitch D, Earon G, Carasso R (1997) Feverfew (*Tanacetum parthenium*) as a prophylactic treatment for migraine: a double-blind placebo-controlled study. *Phytotherapy Research* **11**: 508-511
- Pareek A, Suthar M, Rathore GS, Bansal V (2011) Feverfew (*Tanacetum parthenium* L.): A systematic review. *Pharmacognosy Reviews* **5**: 103-110
- Ren Y, Zhou Y, Chen X, Ye Y (2005) Discovery, Structural Determination and Anticancer Activities of Lactucinlike Guaianolides. *Letters in Drug Design & Discovery* **2**: 444-450



- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD** (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. **440**: 940-943
- Rugutt JK, Fronczek FR, Franzblau SG, Warner IM** (2001) Dihydroparthenolide diol, a novel sesquiterpene lactone. *Acta Crystallographica Section E: Structure Reports Online* **57**: o323-o325
- Seto M, Miyase T, Umehara K, Ueno A, Hirano Y, Otani N** (1988) Sesquiterpene lactones from *Cichorium endivia* L. and *C. intybus* L. and cytotoxic activity. *Chem Pharm Bull (Tokyo)* **36**: 2423-2429
- Ting H-M, Wang B, Rydén A-M, Woittiez L, van Herpen T, Verstappen FWA, 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
- 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
- Wang B, Kashkooli AB, Salles A, Ting H-M, de Ruijter NCA, Olofsson L, Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR** (2016) Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metabolic Engineering* **38**: 159-169
- Zhang Y, Teoh KH, Reed DW, Maes L, Goossens A, Olson DJ, Ross AR, Covello PS** (2008) The molecular cloning of artemisinic aldehyde Delta11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J Biol Chem* **283**: 21501-21508

## Supplementary information



**Supplementary Figure 1.** LC-Orbitrap-FTMS chromatograms of DHCT, CT, DHPT, PT, 3 $\beta$ -OHDHCT, 3 $\beta$ -OHCT, 3 $\beta$ -OHDHPT and 3 $\beta$ -OHPT (from top to bottom), positive ionisation mode. Opposite of each chromatogram is the representative spectrum of the selected RT in that chromatogram. Reduced forms of PT and CT according to authenticated injected standards were deviating  $\sim 0.4$  min from their non-reduced forms. The RT shift for 3 $\beta$ -OHDHCT and 3 $\beta$ -OHDHPT compared to their non-reduced forms were suggesting similar reduction of exocyclic double bond of their non-reduced forms. DHCT: dihydrocostunolide, CT: costunolide, DHPT: dihydroparthenolide, PT: parthenolide, 3 $\beta$ -OHDHCT: 3 $\beta$ -hydroxy-dihydrocostunolide, 3 $\beta$ -OHCT: 3 $\beta$ -hydroxycostunolide, 3 $\beta$ -OHDHPT: 3 $\beta$ -hydroxy-dihydroparthenolide and 3 $\beta$ -OHPT: 3 $\beta$ -hydroxyparthenolide.







## Chapter 4

# Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*

Bo Wang<sup>†</sup>, Arman Beyraghdar Kashkooli<sup>†</sup>, Adrienne Sallets, Hieng- Ming Ting,  
Norbert C.A. de Ruijter, Linda Olofsson, Peter Brodelius, Marc Boutry,  
Harro Bouwmeester, Alexander van der Krol

<sup>†</sup>: These authors contributed equally to this manuscript

Published in *Metabolic Engineering*

## Abstract

Our lack of full understanding of transport and sequestration of the heterologous products currently limit metabolic engineering in plants for the production of high value terpenes. For instance, although all genes of the artemisinin/artemisinin B (AN/AB) biosynthesis pathway (AN-PW) from *Artemisia annua* have been identified, ectopic expression of these genes in *Nicotiana benthamiana* yielded mostly glycosylated pathway intermediates and only very little free (dihydro)artemisinic acid [(DH)AA]. Here we demonstrate that Lipid Transfer Protein 3 (AaLTP3) and the transporter Pleiotropic Drug Resistance 2 (AaPDR2) from *A. annua* enhance accumulation of (DH)AA in the apoplast of *N. benthamiana* leaves. Analysis of apoplast and cell content and apoplast exclusion assays show that AaLTP3 and AaPDR2 prevent reflux of (DH)AA from the apoplast back into the cells and enhances overall flux through the pathway. Moreover, AaLTP3 is stabilized in the presence of AN-PW activity and co-expression of AN-PW+AaLTP3+AaPDR2 genes yielded AN and AB in necrotic *N. benthamiana* leaves at 13 days post-agroinfiltration. This newly discovered function of LTPs opens up new possibilities for the engineering of biosynthesis pathways of high value terpenes in heterologous expression systems.

**Keywords:** Lipid transfer proteins, ABC transporters, Pleiotropic Drug Resistance protein, artemisinin, *Artemisia annua*, *Nicotiana benthamiana*

## Introduction

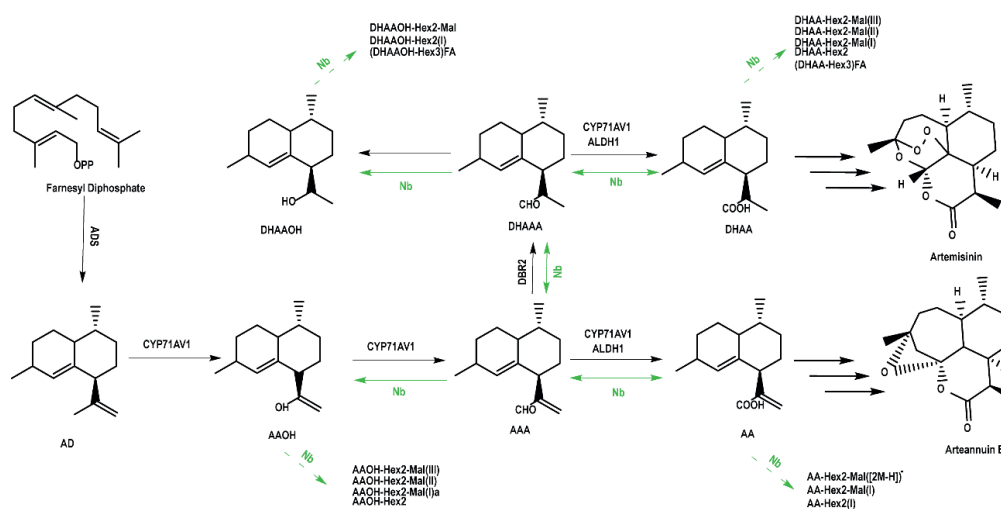
Artemisinin (AN) is a highly oxygenated sesquiterpene which is used to treat malaria. Both AN and arteannuin B (AB) are products of a branched biosynthesis pathway (AN-PW) in glandular trichomes on leaves and ovules from the plant *Artemisia annua*. After biosynthesis, the AB and AN precursors, artemisinic acid (AA) and dihydroartemisinic acid (DHAA), respectively, are transported to the subcuticular space of the trichomes, where (presumably through non-enzymatic oxidation) they are converted to AB and AN, respectively (Brown, 2010) (Fig. 1: AN pathway given with solid arrows).

All genes required for biosynthesis of DHAA and AA have been isolated (amorpha-4,11-diene synthase (ADS), CYP71AV1, artemisinic aldehyde  $\Delta$ 11(13) reductase (DBR2), aldehyde dehydrogenase 1 (ALDH1)) (Bouwmeester et al., 1999; Teoh et al., 2006; Covello et al., 2007; Zhang et al., 2008; Teoh et al., 2009). Reconstitution of the AN-PW in *Nicotiana benthamiana* results in mainly glycosylated PW intermediates and only low levels of free (non-glycosylated) intermediates (dihydro)artemisinic alcohol ((DH)AAOH), (dihydro)artemisinic aldehyde ((DH)AAA) or (DH)AA (Ting et al., 2013) (Fig. 1: glycosylation reactions from endogenous *N. benthamiana* enzymes indicated by dashed arrows). It is assumed that glycosylated AN-PW products are in the vacuole and cannot contribute to AN or AB accumulation. Therefore, sequestering of (DH)AA to the apoplast could potentially help increase the yield of AN and AB in heterologous plant production platforms, but little is known about the factors involved in transport and sequestering of terpenes. Terpene transport and sequestering activity could be by ATPBinding Cassette (ABC) transporters and Lipid Transfer Proteins (LTP), both of which have been detected in terpene producing glandular trichome cells. ABC transporters are involved in the active transport of various metabolites. More particularly, within the ABC family, Pleiotropic Drug Resistance (PDR) transporters are localized to the plasma membrane and involved in the transport of hormones and specialized metabolites. For example, in *Nicotiana tabacum* NtPDR1 was shown to be involved in the transport of antifungal diterpenes (Stukkens et al., 2005; Crouzet et al., 2013). In the same species, a Lipid Transfer Protein (NtLTP1) was shown to be involved in diterpene secretion. Indeed diterpenes in the trichome exudate were clearly increased in the NtLTP1-overexpressing lines and decreased in the NtLTP1-RNAi lines, compared with the wild-type line (Choi et al., 2012). LTPs are small, basic proteins with a hydrophobic cavity, which are characterized in in-vitro assays by their ability to transfer lipids between membranes. Most LTP proteins contain a signal peptide sequence for secretion to the apoplast and extracellular localization of LTP proteins has been confirmed (Kader, 1997; Yeats and Rose, 2008; Edstam et al., 2011; Huang et al., 2013). Plants have multiple *LTP* genes and individual LTPs have been implicated in diverse functions like plant



lipid metabolism (Debono et al., 2009), plant defence (Salcedo et al., 2007), seed development (Edstam and Edqvist, 2014; Wang et al., 2015), sexual reproduction (Zhang et al., 2010), and cell wall extension (Nieuwland et al., 2005). However, for most LTPs the function and mode of action have not been clarified. Indications for a putative role of LTPs in terpene accumulation come from the fact that *LTP* genes show very high transcriptional activity in glandular trichomes which produce terpenes (peppermint (Lange et al., 2000), alfalfa (Aziz et al., 2005), *A. annua* (Bertea et al., 2006), hop (Wang et al., 2008), *Salvia fruticosa* (Chatzopoulou et al., 2010), tomato (Schilmiller et al., 2010) and tobacco (Harada et al., 2010)).

We identified six LTPs from *A. annua* (AaLTPs), three of which are expressed in glandular trichomes. A putative role of these AaLTPs in accumulation of AN-PW products could be tested by silencing of AaLTP expression in *A. annua* using Virus Induced Gene Silencing (VIGS). However, VIGS induced silencing was not successful in our hands. We also identified two AaPDR genes, which are expressed in *A. annua* trichomes. Therefore, functional analysis of the AaLTPs and AaPDRs was done using the fully reconstituted AN-PW by transient co-expression in *N. benthamiana* leaves.



**Figure 1.** Schematic representation of the artemisinin and arteannuin B biosynthetic pathway in *Artemisia annua* (Black arrows). Green arrows indicate *N. benthamiana* enzyme activities within the AN-PW, dashed green arrows indicate reactions in *N. benthamiana* draining from AN-PW. ADS, amorphaadiene synthase; CYP71AV1, amorphaadiene oxidase; ALDH1, aldehyde dehydrogenase; DBR2, double bond reductase. AD, Amorpha-4,11-diene; AAOH, artemisinic alcohol, AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol, DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Hex, compound conjugated with hexose; Mal, compound conjugated with malonate; (I-III), different isobaric forms (i.e. identical accurate mass, but different retention times); FA, formic acid adduct.

## Results

### Isolation of *Artemisia annua* PDR genes expressed in glandular trichomes

Analysis of the EST databases for *A. annua* identified 15 clones expressed in trichomes. In order to identify the PDR transporters highly expressed in glandular trichomes, we designed degenerated primers (Table S2) corresponding to two conserved regions of the 3' coding sequence and surrounding a poorly conserved sequence so as to discriminate different *PDR* genes. Using these primers, cDNA clones were obtained from *A. annua* glandular trichome RNA and sequenced. Among 26 clones characterized, two genes were highly represented: AaPDR1 (NCBI EZ148819, 17 clones) and AaPDR2 (NCBI EZ144340, 8 clones).

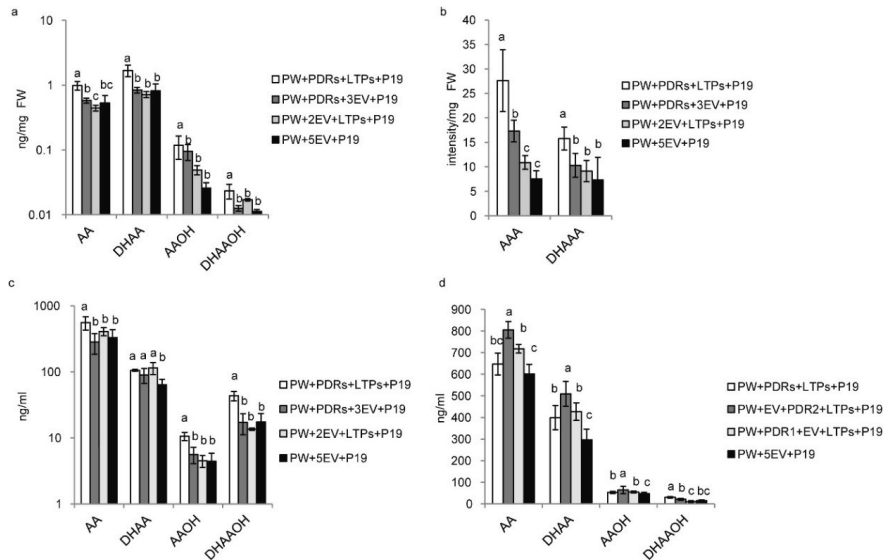
### Isolation of *Artemisia annua* LTP genes expressed in glandular trichomes

The six different LTP sequences were identified in databases resulting from *A. annua* trichome-enriched (glandular + filamentous) cDNA libraries (Bertea et al., 2006; Covello et al., 2007; Dai et al., 2010). Selection of AaLTP1–3 was based on detection of these LTP sequences in glandular trichome-enriched cDNA libraries (Ma et al., 2009; Soetaert et al., 2013) (Table S1). Subsequently three genes (AaLTP1, AaLTP2, AaLTP3) were selected for further functional analysis, based on co-expression with amorphaadiene synthase gene (*ADS*), representing the first step in the artemisinin biosynthesis pathway, or co-expression with AaPDR1 or AaPDR2 (Fig. S1). It was noted that the expression profile in *A. annua* of AaLTP3 and AaPDR2 was distinct from that of *ADS*/ *AaLTP1*/*AaLTP2*/*AaPDR1* and more closely matches the accumulation profile of (DH)AA over same leaf developmental stages (Zhang et al., 2012). The full length sequence of AaLTP1, AaLTP2 and AaLTP3 were obtained through PCR and cloned into binary expression vectors ([www.impactvector.com](http://www.impactvector.com)) for subcellular localization studies and for function in the AN-PW using transient expression assays.

All three glandular-expressed *AaLTP* genes contain an N-terminal signal peptide sequence, indicative for targeting to the ER and secretion to the apoplast. We tested the subcellular targeting of AaLTP-GFP fusion proteins in transient expression assays in *N. benthamiana*. Results show that at 4 days post-agroinfiltration (dpi) the GFP signal was detected both inside and outside the cells (Fig. S2a and b), but at 7 dpi the GFP signal was mainly outside cells (Fig. S2c), which in the spongy mesophyll is visible as bridges of fluorescent material between cells. Remarkably, for all three LTP-GFP fusion proteins, the extracellular accumulation seemed to be coordinated between cells around intracellular cavities (Fig. S2c). Combined, the localization studies confirm an extracellular localization of all three AaLTps.

### AaLTps+AaPDRs enhance freeform AN-PW products in *N. benthamiana* leaves

For functional analysis of the AaLTPs, the set of three *AaLTP* genes (*AaLTP1/2/3*) were co-infiltrated with the AN-PW genes (35S::ADS-tHMGFR-FPP, 35S::CYP71AV1, 35S::DBR2, 35S::ALDH1 (Ting et al., 2013)) and a set of two AaPDR genes (*AaPDR1/2*). In all these experiments a p19 construct was co-infiltrated to repress silencing (Voinnet et al., 2003) and relative dosage of agro-infiltrated AN-PW genes was kept constant by dilution with an *Agrobacterium tumefaciens* strain with empty vector (EV) where required. Free product accumulation (AAOH, AAA, AA, DHAAOH, DHAAA and DHAA) in leaves harvested at six dpi was measured by LC-trip-quad-MS (Ting et al., 2013). The co-expression of either three AaLTPs or two AaPDRs with the AN-PW genes has no effect on (DH)AA accumulation, and a small but significant effect on (DH) AAA and (DH)AAOH accumulation (Fig. 2). However, the combination of AN-PW+AaLTPs+AaPDRs results in a significant increase in accumulation of all free AN-PW products (Fig. 2), suggesting that *AaLTPs*+*AaPDRs*, can sequester these compounds away from competing sugar-conjugating or glutathione-conjugating enzyme activities in *N. benthamiana* leaves. Subsequently, we tested whether this relates to an accumulation of free AN-PW products in the apoplast of *N. benthamiana* leaves.



**Figure 2.** Artemisinin intermediates in *N. benthamiana* infiltrated artemisinin pathway genes with the *AaLTPs* and/or *AaPDRs*. Leaves were harvested at 6 dpi and leaf extracts were analyzed by LC-trip-quad-MS. (a) accumulation of DH(AA), DH(AAOH) in total leaf extract of *N. benthamiana* infiltrated with AN-PW or AN-PW in combination with *AaLTPs* and/or *AaPDRs*. The biggest effect on free product accumulation is found for AN-PW genes coexpressed with *AaPDRs*+*AaLTPs* (2.1-fold increase for AA). (b) relative levels of DH(AAA) in total leaf extract (based on peak intensity because of lack of standard for (DH)AAA) of *N. benthamiana* leaves infiltrated with AN-PW or AN-PW in combination with *AaLTPs* and/or *AaPDRs*. For (DH)AAA coexpression of AN-PW genes with either *PDRs* or *LTPs* both resulted in increased levels of (DH)AAA, but the effect of combined coexpression (*AaPDRs*+*AaLTPs*+AN-PW genes) was the strongest. (3.7-fold increase for AAA compared to levels in leaves infiltrated with only AN-PW genes +5×EV). (c) Concentration of AN-PW

products in the apoplast of *N. benthamiana* infiltrated with AN-PW or AN-PW in combination with *AaLTPs* and/or *AaPDRs*. Agroinfiltrated leaves were harvested at 6 dpi and leaf apoplast wash was analyzed by LC-trip-quad-MS. The biggest effect on free product accumulation in the apoplast is found for AN-PW genes coexpressed with *AaPDRs+AaLTPs* (2.4-fold increase for AAOH, 2.5-fold increase for DHAAOH). (d) concentration of AN-PW products in the apoplast of *N. benthamiana* infiltrated with AN-PW or AN-PW+AaLTPs with *AaPDR1* or *AaPDR2*. Agroinfiltrated leaves were harvested at 6 d.p.i. and leaf apoplast wash was analyzed by LC-trip-quad-MS. The biggest increase compared to product accumulation by AN-PW genes alone was for coexpression of AN-PW genes with *AaPDR2+AaLTPs* (1.3-fold increase AA; 1.7-fold increase for DHAA). The relative gene dosage of AN-PW was constant between the treatments of each experiment. Statistical significant differences ( $p \leq 0.05$ ) are indicated by letters. Error bar is SE ( $n=4$ ).; In (a) & (c) scale y axis is logarithmic.

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

For analysis of product accumulation in leaf apoplast, agro-in- filtrated leaves were infiltrated with water at 6 dpi and subsequently the infiltrated water was collected by mild centrifugation (Witzel et al., 2011). Analysis by UPLC-trip-quad-MS of such apoplast wash of *N. benthamiana* leaves expressing only AN-PW genes showed the presence of AN-PW products AAOH, AA, DHAAOH and DHAA in the apoplast (Fig. 2c). This indicates an intrinsic transport activity for these compounds in *N. benthamiana*. In contrast to total leaf extract, the intermediates (DH)AAA were not detected in the apoplast wash fluid (not shown), indicating little or no cell leakage during the apoplast wash procedure. As further control for cell leakage we measured the flavonoid rutin, which accumulates inside cells of *N. benthamiana* leaves and not in the apoplast (Marrs et al., 1995; Markham et al., 2001; Ökmen et al., 2013). No rutin was detected in apoplast fluid, indicating minimal contamination by broken cells (Fig. S3a). Analysis of apoplast samples by LC-QTOF-MS did occasionally show the presence of very low levels of DHAA-Hex3 (maximal 5% of total in whole leaf extract; Fig. S3b). Because no other glycosylated AN-PW products, which were present in the total extraction, were detected in the apoplast, we attribute DHAA-Hex3 in the apoplast to a low intrinsic transport activity for this compound by *N. benthamiana* cells, rather than cell leakage. Analysis of apoplast wash of leaves expressing either *AN-PW+AaLTPs* or *AN-PW+AaPDRs* showed no increase in AN-PW intermediates in the apoplast, confirming results from total leaf extracts (Fig. 2c). However, expression of *AN-PW+AaPDRs+AaLTP* genes resulted in a significant increase in apoplast levels of (DH)AAOH (2.4-fold increase for AAOH, 2.5-fold increase for DHAAOH) and (DH)AA (1.7-fold increase) (Fig. 2c), again matching results from total leaf extracts. Having confirmed that the set of *AaLTPs* with *AaPDRs* enhances accumulation of ANPW intermediates in the apoplast, we subsequently tested the specificity of the *AaLTP* gene set with either *AaPDR1* or *AaPDR2*.

### **AaLTPs are more effective with AaPDR2 than with AaPDR1**

We tested whether there is a difference in apoplast sequestering activity for the combination *AaLTPs+AaPDR1+AN-PW* and *AaLTPs+AaPDR2+AN-PW*. Leaves agro-infiltrated with the

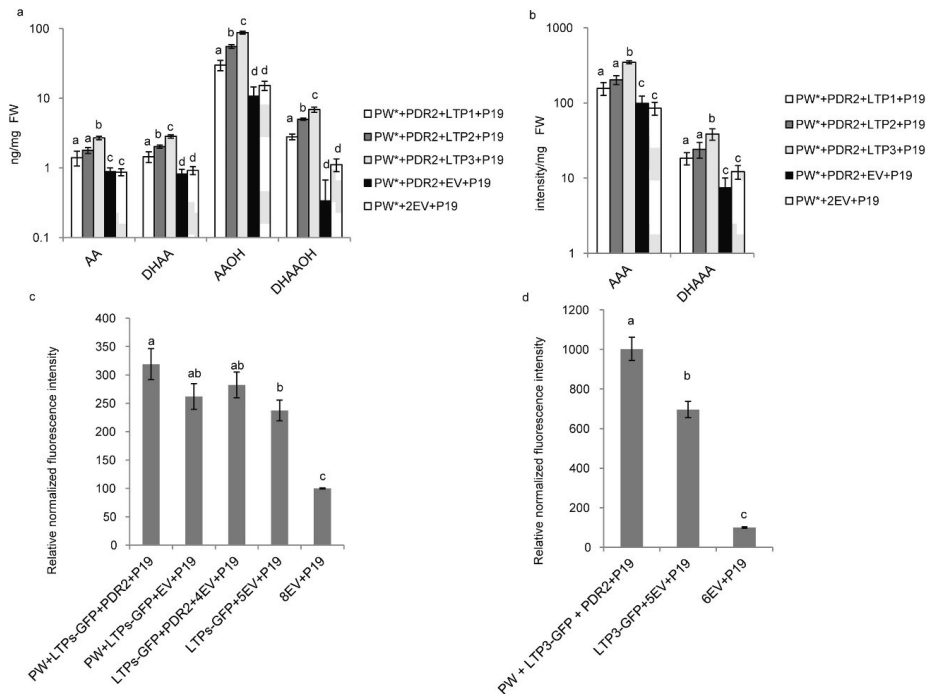
different gene sets were harvested at 7 dpi and the apoplast fluid was isolated and analyzed by UPLC-trip-quad-MS. Results show that AaLTPs were more effective with AaPDR2 than with AaPDR1 for accumulation of (DH)AA in the apoplast (Fig. 2d). Results show that AaPDR1 and AaPDR2 have little or no effect on accumulation of (DH)AAOH, while both proteins do give an increase in (DH)AA levels in the apoplast. For apoplast accumulation of (DH)AA the AaPDR2 has a stronger activity than AaPDR1. All the PDR transporters characterized so far have been localized to the plasma membrane. Nevertheless, we checked the plasma membrane localization of AaPDRs fused to GFP and expressed in *Nicotiana tabacum* suspension cells (Fig. S4). Results show that GFP-AaPDR1 is not detected in cells expressing the fusion protein (not-shown), even though AaPDR1 is functionally active in the transient expression assays (Fig. 2d). In contrast, GFP-AaPDR2 is clearly detected in the plasma-membrane of cells where the fusion protein clearly co-localized with FM4-64, a fluorescent dye used as a plasma membrane marker (Fig. S4). Combined the results indicate a higher transport activity for AaPDR2, which in part may be explained by higher intrinsic stability of the AaPDR2 protein compared to AaPDR1.

The activity of the set of AaPDRs for (DH)AA was lower than the activity of AaPDR2 for (DH)AA alone. This could be caused by protein crowding on the plasma membrane, resulting in a competition between the active AaPDR2 and less active AaPDR1 during the high expression levels of the PDRs during transient expression. Therefore, we subsequently tested the specificity of individual *AaLTPs* in combination with the *AN-PW+AaPDR2*.

### **AaLTP3+AaPDR2 are most active in enhancing (DH)AA levels in *N. benthamiana* and increasing flux through the AN-PW**

To determine whether there is a difference in activity of the individual AaLTPs, the individual *AaLTPs* were co-expressed with the *AN-PW\** (*PW\** refers to AN-PW with additional boosting by extra HMGR: HMGR+AN-PW)+AaPDR2 genes in *N. benthamiana* leaves. Agro-infiltrated leaves were harvested at 7 dpi and freeform AN-PW products were quantified in total leaf extracts. Since the (DH)AA are the direct precursors for the final products AN and AB, we focus on the accumulation of (DH)AA to characterize the activity of individual PDRs and LTPs. The results show that all three AaLTPs enhance the level of (DH)AA in total leaf extract, compared to only expression of *AN-PW+AaPDR2* (Fig. 3a and b). However, AaLTP3 has the biggest effect on enhancing accumulation of (DH) AA (Fig. 3a and b). Control experiment showed that the transcript levels of the individual AaLTPs in these transient assays were very similar (Fig. S5), indicating that the difference in AaLTP activity is either related to difference in individual AaLTP function and/or difference in individual LTP protein stability. Pilot studies showed that the fusion with GFP does not affect activity of the individual AaLTP proteins and therefore we used the fluorescence signal of AaLTP-GFP

fusion proteins to investigate possible differences in intrinsic stability of the individual AaLTPs. Results show that GFP fluorescence of leaves expressing AaLTP3-GFP is significantly higher than that of leaves expressing AaLTP1-GFP or AaLTP2-GFP (Fig. S6), while control experiments again show that transcriptional activity of these genes is very similar. The higher accumulation of (DH)AA in the leaf in the presence of AaLTP3 therefore seems to be directly related to a higher intrinsic stability of AaLTP3. Moreover, fluorescence of AaLTP3-GFP was increased by 69% when co-expressed with *AN-PW* and *AaPDR2* genes, suggesting that AaLTP3 interacts with the AN-PW (Fig. 3c, d).

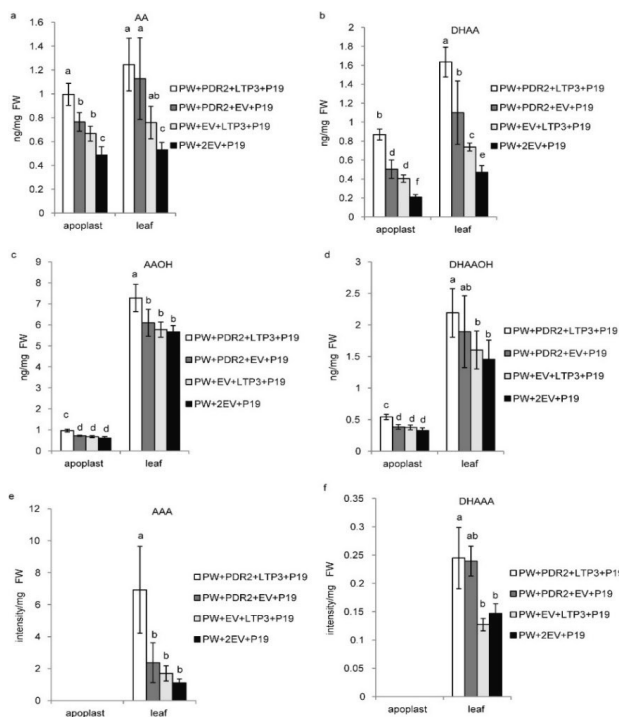


**Figure 3.** AN-PW intermediates in leaf extract of *N. benthamiana* agro-infiltrated with AN-PW\*+*AaPDR2* and individual *AaLTP* genes in (a) and (b). Agroinfiltrated leaves were harvested at 7 d.p.i. and leaf extract was analyzed by LC-trip-quad-MS. Error bar is SE (n=4); Note that (a) and (b) are in logarithmic scale. (a) accumulation of DH(AA), DH(AAOH); (b) relative levels of DH(AAA) (based on peak intensities due to lack standard (DH)AAA). The biggest effect of AN-PW+*PDR2* was with *AaLTP3* compared to leaves infiltrated with only AN-PW genes (+2×EV) (5.7-fold increase AAOH; 6.2-fold increase DHAAOH; 4.1-fold increase AAA). Statistical significant differences based on Student's *t*-test,  $p < 0.05$  are indicated by letters. Error bar is SE; n=4. Relative fluorescence intensity of LTP-GFP transiently expressed in *N. benthamiana* increases in presence of AN PW activity in (c) and (d). (c) Relative fluorescence of GFP-AaLTP1+GFP-AaLTP2+AaGFP-LTP3 with or without co-expressed AN-PW+*AaPDR2* gene expression. (d) Relative fluorescence of LTP3-GFP with or without AN-PW+*AaPDR2* gene expression. The relative gene dosage of *LTPs-GFP* was equal between all experiments by equilibrating with EV where needed. Results of two individual experiments combined (n≥54). PW: Artemisinin Pathway constructs V5, ALDH1, P450, DBR2; PW\* means AN-PW genes with an extra HMGR in the agroinfiltration; LTPs-GFP: AaLTP1-GFP+AaLTP2-GFP+AaLTP3-GFP; PDR2: AaPDR2; EV: Empty Vector.

The efficiency for apoplast accumulation of individual AN-PW compounds as determined by the ratio of product level in apoplast and product level in total leaf extract, was higher for acid than for the alcohol products (Fig. 4). Not all free AN-PW products are present in the apoplast fluid, indicating that either some products are still inside the cells, or extraction of the apoplast with water is not complete. Indeed, control experiments showed that the partitioning of (DH)AA between a polar water phase and an inorganic chloroform phase is very similar to the partitioning of (DH)AA over apoplast and total leaf (Fig. S7). In contrast, the partitioning of (DH)AAOH over apoplast and total leaf is actually lower than the partitioning of (DH)AAOH between water and chloroform. We determined whether the increased accumulation of free AN-PW products by AaLTP3 has an effect on total AN-PW activity by measuring (DH)AA glycosides and glutathione conjugations by the LC-QTOF-MS in leaves expressing the different combinations of AN-PW, AaLTP3 and AaPDR2 genes. These measurements indicate that glycosylated (DH)AAOH product levels are not much affected by AaLTP3, but that glycosylated (DH)AA product levels are significantly increased by AaLTP3 (Fig. S8). This is consistent with the higher specific transport activity of AaLTP3 for (DH)AA than for (DH)AAOH. The total flux through the AN-PW is also increased by AaPDR2 and even stronger by AaLTP3+AaPDR2 (Fig. S8).

**Figure 4.** AN-PW product levels in *N. benthamiana* leaf apoplast and leaf cells expressing AN-PW with or without AaLTP3 and AaPDR2 genes

Leaves were harvested at 6 dpi. The leaf weight before and after the apoplast extraction was used to calculate the apoplast fluid not extracted from leaf. AN-PW product level remaining in leaf was calculated and corrected for product level in non-extracted apoplast fluid. The biggest effect on product accumulation in the apoplast is by coexpression of AaPDR2 + AaLTP3 + AN-PW genes (2.6-fold increase AA; 5.3-fold increase DHAA) compared to levels infiltrated with AN-PW+2×EV. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE;  $p < 0.05$ ,  $n = 5$ .



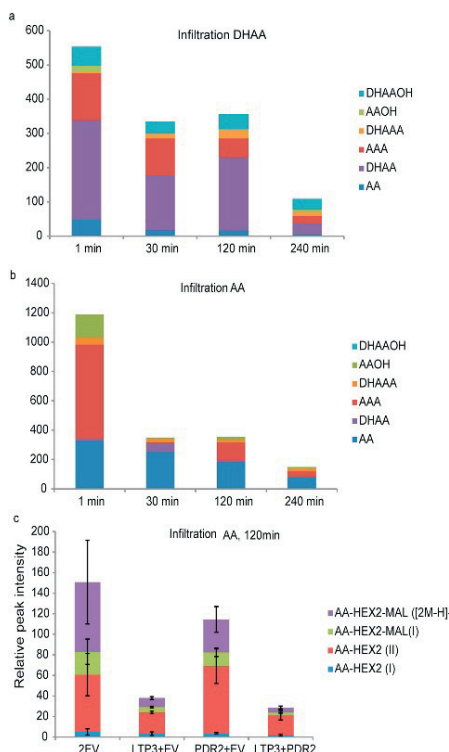


## AaLTP3 blocks reflux of (DH)AA from apoplast into *N. benthamiana* leaf cells

Since AaLTP3 enhances AN-PW product accumulation in the apoplast when co-expressed with *AN-PW* genes, we tested whether this is due to a retention of AN-PW products by AaLTP3. In- filtration of AA or DHAA in *N. benthamiana* leaves shows very rapid uptake by cells and conversion to aldehydes and alcohols (Fig. 5a and b), which we attribute to endogenous alcohol dehydrogenase (ADH) activity. We used the uptake of AA by cells and subsequent intracellular conversion to different glycosylated products, to study the effect of AaLTP3 on retaining AA in the apoplast. Leaves expressing only EV were not able to exclude AA from the cells, resulting in high levels of AA- and DHAA-glycosides at 2 h post-substrate infiltration. Expression of only AaPDR2 was able to prevent ~27% of product conversion to glycosides, however, AaLTP3 was able to prevent ~74% of product conversion to glycosides, while the combined action of AaLTP3+AaPDR2 prevented up to ~82% of AA conversion inside the cell to glycosides (Fig. 5c). These results indicate that export of AA by AaPDR2 is not very effective when this activity is not combined with a retention system in the apoplast and that AaLTP3 functions in retention of AA in the apoplast.

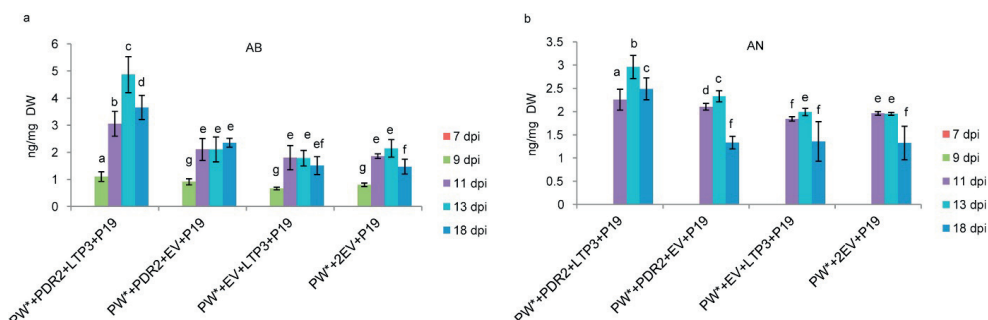
**Figure 5.** Bioconversion of DHAA or AA infiltrated in *N. benthamiana* leaves. (a) Apoplast infiltration of DHAA (500 ng/mg) and harvest whole leaf disks at 1–240 min; (b) Apoplast infiltration of AA (500 ng/mg) and harvest at 0–4 h; The bioconversion of infiltrated AA or DHAA was only done once to determine the best harvest time for the exclusion experiment of Fig. 4c (no error bars). Unit of AA, DHAA, AAOH, DHAAOH is ng/mg except for AAA and DHAAA (relative peak intensity) in (a) and (b). (c) Semi quantification of AA entering cells after apoplast infiltration of AA (2.67 mM) at 3 dpi of *N. benthamiana* leaves

expressing *AaPDR2*+EV, *AaLTP3*+EV or *AaPDR2*+*AaLTP3* at 3 dpi and harvest whole leaf disks 120 min after infiltration of AA into the apoplast. Inside the cells AA is rapidly glycosylated by addition of two hexose sugars (HEX2) which in some instances are capped by and additional malonyl group (HEX2-MAL). Different forms of AA glycosylated products elute at different times from the LC: (AA-HEX2-MAL (2M-H), AA-HEX2-MAL (I), AA-HEX2 (I) and A-HEX2 (II). Total AA levels entering the cell are therefore quantified by combining the different peak intensities of glycosylated AA products in the chromatograms. Units are based on peak intensities of the different AA glycosylated products.



### AaLTP3+AaPDR2 enhance accumulation of AN and AB in *N. benthamiana*

The accumulation of the different pathway intermediates (AAOH, AAA, AA, DHAAOH, DHAAA, DHAA) and the desired endproducts (AB and AN) in *N. benthamiana* leaves was measured at different times post-agroinfiltration (Fig. 6, S9). Results show that (DH)AA accumulates till 11 dpi, after which product levels decline again. At 7 dpi no AN or AB is detected in leaves, but these compounds start to accumulate from day 9 till 13, after which their levels start to decline. Because T-DNA from agro-infiltration may survive up to ~9 dpi (Kanagarajan et al., 2012; Kanagarajan et al., 2012), this suggests that proteins may be active till 11 dpi and that the end products AB and AN are not stable in *N. benthamiana*. In leaves expressing *AN-PW+AaPDR2+AaLTP3* we detect signs of necrosis at 9 dpi, which coincides with the first detection of AB, suggesting that necrosis may play a part in the conversion of AA to AB. Because of the leaf necrosis we did not perform apoplast measurements for samples harvested after 7 dpi.



**Figure 6.** Accumulation of AN and AB in leaf tissues of *N. benthamiana* agro-infiltrated with AN-PW\* genes with or without *AaLTP3* and/or *AaPDR2*. Leaves were harvested at times indicated and AN and AB levels were quantified by LC-trip-quad-MS. Because around 6 dpi necrosis starts in leaves for some of the treatments, all results are expressed as ng per mg Dry Weight (DW). Error bar is SE; n=4. PW\* means AN-PW genes with an extra HMGR in the agroinfiltration.

## Discussion

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

Here we have investigated the role for AaLTPs and AaPDRs in AN-PW product accumulation using transient expression in *N. benthamiana*. The AaPDRs+AaLTPs together have a limited effect on boosting transport of AN-PW intermediates to the apoplast of *N. benthamiana* leaves (Fig. 2). In contrast, the effect of coexpression of only *AaPDR2+AaLTP3* with the AN-PW genes is more effective in accumulation of free AN-PW products (Fig. 3). This difference could be explained by the different number of *Agrobacterium* strains carrying the independent expression constructs that are used for co-infiltration into the *N. benthamiana* leaves. In Fig. 2

the total PW plus transport activity depends on co-expression of 10 genes in each cell. In Fig. 3 the total PW plus transport activity is dependent on co-transformation with 8 gene constructs.

Indeed, control experiments have shown that the efficiency of cotransformation by independent agrobacterium strains that are coinfiltrated into *N. benthamiana* leaves drops off when more than 7 constructs are co-transfected. Therefore, in Fig. 2, more cells may only express (part of the) PW genes and no genes for transport activity compared to the experiment in Fig. 3. For the two AaPDR genes we show that AaPDR2 is more effective for AN PW products (Fig. 2d). Moreover, we have collected multiple lines of evidence that AaLTP3 functions with AaPDR2 in the sequestration of AN-PW products: (1) AaLTP3 is coexpressed with AN-PW genes and PDR transporter genes in glandular trichomes (Table S1); (2) the AaLTP3 expression profile is similar to that of AaPDR2 (Fig. S1) and AaLTP3 activity on free AN-PW compound accumulation is highest in combination with AaPDR2 (Fig. 3); (3) AaLTP3 protein is extracellular (Fig. S2), while AaLTP3 enhances AN-PW product accumulation in the apoplast; (4) AaLTP3 functions in retention of AA in the apoplast (Fig. 5); (5) AaLTP3 protein stability is increased by AN-PW activity (Fig. S6) suggesting some kind of functional interaction between AaLTP3 and AN-PW activity. For production of AN/AB in the apoplast only the transport of (DH)AA supposedly would be relevant. It could be that intermediates like (DH)AAOH, which apparently are transported to the apoplast (Figs. 2c, 4c and d), recycle back into the cell for further conversion and only for the end-product (DH)AA there is a requirement to keep the product in the apoplast for conversion to AN/AB. We note that PDRs+LTPs have an effect on accumulation of (DH)AAA, a product not found in the apoplast wash. This could indicate that also inside of the cells PDRs+LTPs affect sequestering of metabolites.

### **The need of AaLTP3 function in planta**

The experiments with apoplast infiltration of AA and DHAA have shown that both of these compounds are rapidly taken up by the cells of *N. benthamiana* and that endogenous enzyme activity in *N. benthamiana* causes a strong reverse flux of (DH)AA to (DH) AAA and (DH)AAOH. This is supposedly caused by endogenous alcohol dehydrogenases (ADHs) or oxidoreductases that reduce (DH)AA. Such activity may compromise accumulation of desired AN-PW end-products in *N. benthamiana*. A similar activity was detected in *A. annua* by Ryden et al., who characterized an oxidoreductase Red1, expressed in *A. annua* flowers that catalyses the reduction of DHAAA to DHAAOH (Rydén et al., 2010). This suggests that for heterologous expression systems and possibly also for *A. annua* the removal of (DH)AA from the cytosol through transport to and retention in the apoplast by AaLTP3 may be an important driver for the AN-PW flux towards (DH)AA. Indeed, the specific removal of (DH)AA from

the endogenous AN-PW in *N. benthamiana* increased the overall flux through the AN-PW, as both free (DH)AA and glycosylated (DH)AA product accumulation were increased by co-expression of PDR2 and LTP3 (Figs. 3, 4 and S8), while there was much less effect on free and glycosylated (DH) AAOH (Fig. S8). Removal of DHAA by increased transport to the apoplast may not necessarily lower the DHAA concentration, when there is product feedback inhibition in the pathway (accumulated DHAA inhibits pathway enzyme activity). In such case, increased removal of DHAA by increased transport to the apoplast will result in similar DHAA steady state level, but with increased flux. Both the branch of transport to the apoplast and the branch towards DHAA glycosylation may benefit from the increased flux through the pathway.

### **Specificity of AaLTPs for AN-PW products**

There are several lines of evidence that this effect of AaLTP3 is specific for (DH)AA and is not caused by generation of a general (lipid) sink for sesquiterpenes in the apoplast: (1) in contrast to AaLTP3, AaLTP1 and AaLTP2 have little or no effect on free (DH)AA accumulation in the apoplast. (2) AaLTP3 protein stability is increased by AN-PW activity. Finally, transient co-expression of AaLTP3 with the costunolide biosynthetic pathway from feverfew (a sesquiterpene PW very similar to the AN-PW; (Liu et al., 2011)) in *N. benthamiana* did not affect costunolide accumulation, suggesting specificity for (DH)AA and not any other lipid molecule.

### **Putative model of AaLTP3 function**

The retention activity of AaLTP3 for AA cannot be explained by a simple 1:1 binding to AA. If we assume an extreme high AaLTP3 production at 3 dpi of 1 mg/g FW in *N. benthamiana*, the estimated molar ratio of (AaLTP3): (infiltrated AA) is approximately 1:635. Although at present we can only speculate on how AaLTP3 fulfills the retention activity in the apoplast, we favor the model in which AaLTP3 transports AN-PW products from the space between the plasma-membrane and the cell wall to the other side of the cell wall at the site of inter cellular spaces. Once deposited at this site a reflux back into the cell is then largely prevented. The empty LTP may subsequently diffuse back to the plasma membrane to be reloaded by another (DH)AA molecule. The characterization of the AaLTP3 and AaPDR2 has here been done in the context of *N. benthamiana* leaf cells and not in the context of secretory trichome cells. Although by agroinfiltration of *N. benthamiana* leaves leaf epidermal cells are also subject to transformation, we did not find any evidence that export of AN-PW products was enhanced in the epidermal cell layer (e.g. rapid chloroform dipping, which is sufficient to release AN from glandular trichomes of *A. annua*, was not sufficient to extract AN-PW products from *N. benthamiana* leaves expressing AN-PW genes+AaPDR2+AaLTP3). Glandular trichomes of *A.*

*annua* are fully specialized in producing AN/AB and part of the secretion system active in such specialized cells may still be missing in our reconstituted system of AN-PW+PDR+LTP genes. Finally, the insight that LTPs may be needed to remove secreted products from the plasma membrane to prevent reflux opens up new possibilities in the engineering of biosynthesis pathways of high value terpenes in heterologous expression systems.

## Materials and methods

### Cloning procedures

The full-length *AaLTP1*, *AaLTP2* and *AaLTP3* were amplified from cDNA made from RNA isolated from *A. annua* flowers by RACE PCR (Clontech, USA) (Bertea et al., 2006). The primer pairs *AaLTP1*-F/ *AaLTP1*-R, *AaLTP2*-F/*AaLTP2*-R and *AaLTP3*-F/*AaLTP3*-R (Table S2) introduce a *NcoI* site at the start codon and a *NotI* restriction site after the stop codon of each LTP coding sequence. The PCR fragments were cloned into pIV1A2.1 vector ([www.impactvector.com](http://www.impactvector.com)) between the CaMV35S promoter and the RBCS1 terminator fragment. pIV1A2.1/*AaLTP1*, pIV1A2.1/*AaLTP2*, pIV1A2.1/*AaLTP3* were transferred into binary vector pBinPlus by LR reaction (Invitrogen) and the resulting pBIN/*AaLTP1*, pBIN/*AaLTP2*, and pBIN/*AaLTP3* were transformed into the *A. tumefaciens* AGL-0 strain which was used for the agroinfiltration experiments (Van Herpen et al., 2010; Ting et al., 2013). For the LTP-GFP fusion protein studies we first made an entry vector with enhanced green GFP (EGFP). The EGFP coding sequence was amplified by PCR using primers EGFP\_c-term-F/ EGFP\_c-term-R (Table S2) by using pBinEGFP as template (<http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriële-eiwitten-door-planten.htm>). After digesting with *NotI* and *SacI*, GFP was cloned into ImpactVectorpIV1A\_2.1 ([www.impactvector.com](http://www.impactvector.com)) under the control of CaMV35S promoter to generate an entry vector pIV1A\_2.1/EGFP. To remove the stop codon of *AaLTP1*, *AaLTP2* and *AaLTP3*, DNA fragments were amplified using the primer pairs *AaLTP1\_gfp*-F/*AaLTP1\_gfp*-R, *AaLTP2\_gfp*-F/ *AaLTP2\_gfp*-R and *AaLTP3\_gfp*-F/*AaLTP3\_gfp*-R (Table S2) by using the pIV1A\_2.1/*AaLTP1*, pIV1A\_2.1/*AaLTP2* and pIV1A\_2.1/*AaLTP3* vectors as templates, respectively. The PCR products were digested with *BamHI* and *NotI* and subcloned into vector pIV1A\_2.1/EGFP. The resulting pIV1A\_2.1/*AaLTP1*-GFP, pIV1A\_2.1/ *AaLTP2*-GFP and pIV1A\_2.1/*AaLTP3*-GFP containing CaMV35S promoter were cloned into the pBinPlus binary vector (Invitrogen) using LR recombination. The pBinPlus constructs containing the 35S: *AaLTP1*-EGFP, 35S: *AaLTP2*-EGFP and 35S: *AaLTP3*-EGFP were transferred to *A. tumefaciens* AGL-0 using electroporation. For the LTP-RFP fusion protein genes the LTP coding sequence without stop-codon was cloned in frame to the RFP coding sequenced in TOPO vector (Invitrogen). The resulting TOPO/*AaLTP1*-RFP, TOPO/*AaLTP2*-RFP and TOPO/*AaLTP3*-RFP were cloned under control of CaMV35S

promoter into the pK7RWG2 binary vector (<http://gateway.psb.ugent.be/>) using LR recombination (Invitrogen). Partial AaPDR cDNAs were amplified from glandular trichome cDNA (Olofsson et al., 2012) using two sets of forward (AaPDR-F) and reverse (AaPDR-R) degenerated primers and cloned in pGEMT-Easy (Promega). Full cDNAs were then obtained by RACE PCR. The PDR coding sequences were cloned between the pEn2PMA4 promoter and the tNOS terminator in the pAUX3131plasmid (De Muynck et al., 2009) and then transferred into the I-SceI site of the pPZP binary plasmid (Goderis et al., 2002). For subcellular localization, a fusion mGFP4::AaPDR2 was obtained in the pPZP vector, which already contained an mCherry gene (Mercx et al., 2016). *A. tumefaciens*-mediated transformation of *Nicotiana tabacum* BY-2 cells was performed as in Mercx et al. (2016).

### **Metabolite analysis**

#### **LC-QTOF-MS for free AN-PW products**

For the AN-PW free compound analysis, the agro-infiltrated leaves were harvested, ground in liquid nitrogen. Frozen grinded samples (100 mg) were extracted in 300 ml methanol using 10 min sonication, after which the samples were centrifuged at 18,625 g at 4 °C. The supernatant was collected and filtered through 0.45 mm PTEE-membrane filter and diluted 10-fold with 10% methanol for analysis by a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described (Ting et al., 2013). For the AN-PW glycosides and conjugations, 100 mg of infiltrated leaf material was extracted in 300 ml methanol/0.137% formic acid using 10 min sonication, after which the samples were centrifuged at 18,625g at 4 °C and filtered through a 0.45 mm PTEE-membrane filter. The supernatant was analyzed on a LC-QTOF-MS (Waters). Since leaves showed signs of necrosis after 7 dpi, samples collected from 7 dpi to 18 dpi were freeze-dried and yield was determined by dry-weight rather than by fresh weight. For analysis 100 mg of dried samples was used following the same extraction procedure as for fresh-weight samples.

#### **Isolation apoplast fluid**

Apoplastic fluids were collected using the infiltration centrifugation technique with modifications (Witzel et al., 2011; Joosten, 2012). For the vacuum infiltration, the fresh leaves were cut into 5 cm strips and stacked in a beaker filled in with deionized water. For the apoplast fluid collection, leaf strips were placed on a grid resting on the conical bottom part of 50 ml plastic centrifuge tubes (Greiner bio-one). Tubes were centrifuged at 400g for 15 min at 4 °C to collect the fluid in the bottom of the tubes. After collection of the released apoplast fluid the leaf strips were weighted again, and subsequently shock frozen in liquid nitrogen for the metabolites analysis. The weights were recorded before/ after the infiltration and

centrifugation, on average, 30–40% of the infiltrated water could be recovered by the centrifugation step. The apoplast wash fluid was filtered through a 0.45 mm PTEE membrane filter and the filtered fluid was dilute 5 times with 12.5% MeOH for analysis on Auity UPLC system (Waters). The concentration of the different products in the apoplast wash fluid was used to compare the different gene combinations in the transient expression assays. The transport activity in the leaf for different substrates was calculated: transport activity = mg substrate in total apoplast fluid volume / mg of product in total leaf minus apoplast content.

### **Expression profiling in *A. annua***

Shoot apex and leaves from top (Leaf 1) to bottom (Leaf 10) were collected from 2-month old *A. annua* plants and pooled per leaf number for RNA extraction. From each pooled and frozen grinded sample 100 mg was extracted for gene expression analysis. RNA was isolated from the frozen grinded samples and DNase I (Invitrogen) was used to remove remaining genomic DNA. Reverse transcription was carried out with the SuperScript First-Strand Synthesis System for the cDNA synthesis (Invitrogen). Real-time quantitative PCR was performed using the iQ5 RT-PCR (BioRad). Primers used are listed in Table S2. The *A. annua*  $\beta$ -actin gene was used for the normalization (Olofsson et al., 2011; Lu et al., 2013).  $\Delta\Delta CT$  was calculated as follows:  $\Delta CT = CT(\text{Target}) - CT(\text{Actin})$ ,  $\Delta\Delta CT$  was normalized using  $\Delta CT$ . The fold change value was calculated by  $2^{-\Delta\Delta CT}$ .

### **Transient expression in *N. benthamiana* leaves**

The following binary expression constructs were used in pathway reconstruction experiments by transient expression in *N. benthamiana*: PW: 35S::ADS-HMGR-FPS (A2 construct (Van Herpen et al., 2010)), 35S::CYP71AV1, 35S::DBR2, 35S::ALDH1. PW\*: PW+35S::HMGR. LTPs: 35S::AaLTP1, 35S::AaLTP2, 35S::AaLTP3. PDRs: 35S::AaPDR1, 35S::AaPDR2. The following binary expression constructs were used in subcellular localization studies by transient expression in *N. benthamiana*: AaLTPs and AaPDRs: 35S:: AaLTP1-GFP, 35S::AaLTP2-GFP, 35S::AaLTP3-GFP. Agroinfiltration of *N. benthamiana* leaves was done as described (Van Herpen et al., 2010). The relative dosage of each *A. tumefaciens* strain between treatments within the same experiment was kept constant by dilution with *A. tumefaciens* carrying an empty vector construct. Infiltrated leaves were harvested at the days indicated (varying from 3 to 18 days post agroinfiltration). In some experiments AN-PW activity was boosted by agroinfiltration of an additional truncated HMGR expression construct (Van Herpen et al., 2010; Ting et al., 2013), indicated by PW\*.



### **Semi-quantification of LTP-EGFP fusion protein levels in agroinfiltrated leaves**

Leaves expressing AaLTP-EGFP or empty vector (control) were harvested at 6 days post infiltration and four representative fully infiltrated flat regions of equal size were imaged per leaf under UV illumination at equal exposure times with a Leica MZIII fluorescence stereomicroscope. Care was taken that pixel saturation never occurred. Color images were split into individual R, G or B channels and the signal in the G channel was quantified in ImageJ 1.49 n. Four biological replicate leaves (16 individual areas) were used to determine the average fluorescence signal in the G channel for each treatment.

### **Subcellular localization of LTP-GFP and GFP-PDR**

The LTP-GFP and GFP-PDR proteins were transiently expressed in the *N. benthamiana* leaves. Expression and targeting analysis was carried out at 5–7 days post agroinfiltration (dpi) by confocal laser scanning microscopy. In order to remove air pockets from intercellular spaces in the spongy parenchyma, the leaf disks ( $r=0.5$  cm) were mounted in perfluorodecalin (Sigma-Aldrich, P9900) (Littlejohn and Love, 2012). A Zeiss 510-META inverted confocal laser scanning microscope (Carl Zeiss, Jena, Germany) equipped with a 63x (oil, 1.4 N. A.) Plan Apochromat DIC objective was used for image acquisition (Dong et al., 2013). Clear localizations were obtained with 1% of 488-nm Ar laser line using BP\_505–530 nm for GFP and LP\_560 nm to observe chlorophyll auto-fluorescence. Images were processed with Zeiss LSM image browser. For GFP::AaPDR2, observations were performed on *N. tabacum* BY-2 cells from a five day-old culture. Cells were incubated with 10 mM FM4–64 dye for 10 min before their analysis on a LSM 710 confocal microscope (Zeiss). Excitation wavelengths were 488 nm (GFP and FM4–64) and 558 nm (mCherry). Emission wavelengths were 495–535 nm (GFP), 640–740 nm (FM4–64), and 578–620 nm (mCherry).

### **Apoplastic exclusion assay**

Four-weeks old *N. benthamiana* leaves were infiltrated with *A. tumefaciens* harboring AaLTP3, AaPDR2, AaLTP3+AaPDR2 or EV as the control; each together with p19. In order to achieve the highest expressed protein levels, infiltrated leaves were harvested after five days. For each treatment, four leaf disks of 1 cm diameter were excised from four individual leaves. The leaf disks were put in 6 well plate (1 leaf disc/well) and vacuum infiltrated with 3500  $\mu$ l of 500  $\mu$ M Na<sub>2</sub>HPO<sub>4</sub> buffer to make a constant flow between the buffer and apoplastic region. Then 2500  $\mu$ l of solution was taken out and a solution of artemisinic acid was added into each well in a manner that each well got 0.625 mg of free compound. The 16-well plate was shaken gently for 2 h on an orbital shaker and then the leaf disks were extracted with 300  $\mu$ l of MeOH and 0.1% formic acid. Extracts were analyzed by LC-QTOF-MS.

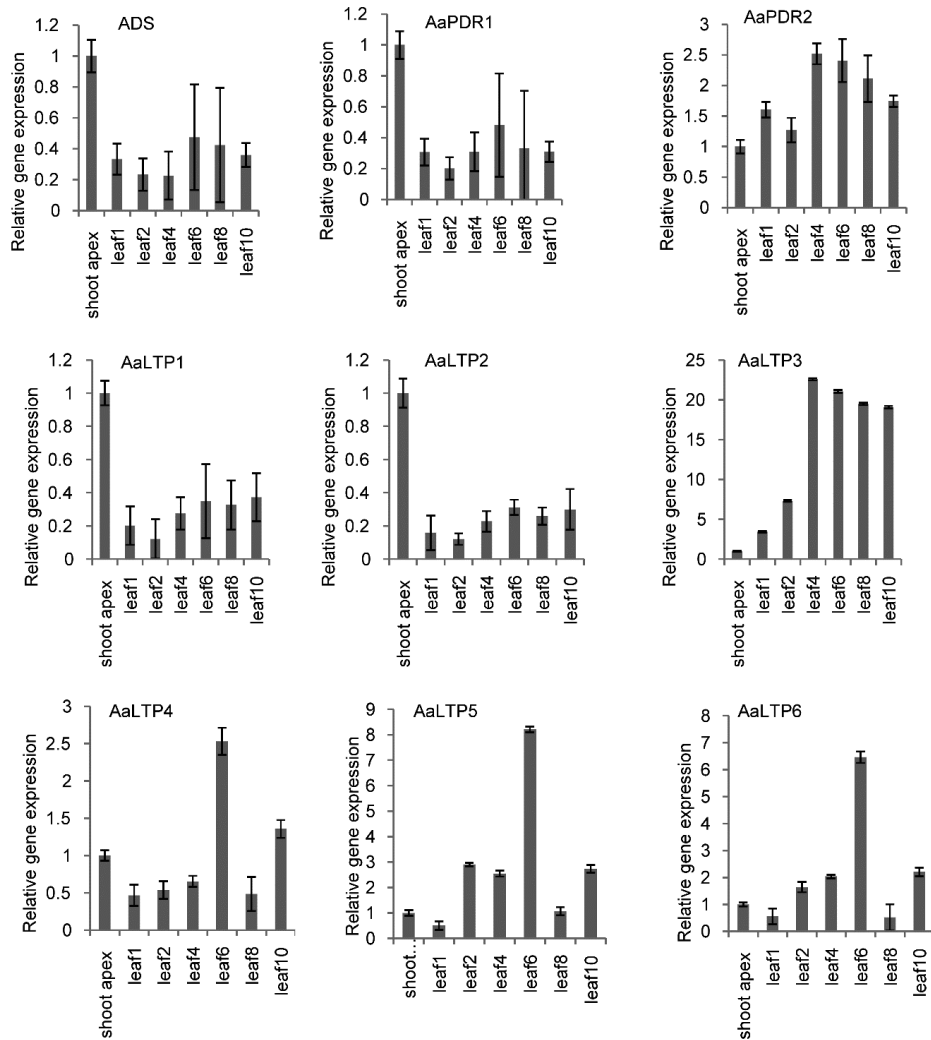
## References

- Aziz N, Paiva NL, May GD, Dixon RA (2005) Transcriptome analysis of alfalfa glandular trichomes. *Planta* **221**: 28-38
- 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. *Arch Biochem Biophys* **448**: 3-12
- Bertea CM, Voster A, Verstappen FWA, Maffei M, Beekwilder J, Bouwmeester HJ (2006) Isoprenoid biosynthesis in *Artemisia annua*: Cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Archives of Biochemistry and Biophysics* **448**: 3-12
- Bouwmeester HJ, Wallaart TE, Janssen MH, van Loo B, Jansen BJ, Posthumus MA, Schmidt CO, De Kraker JW, Konig WA, Franssen MC (1999) Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* **52**: 843-854
- Brown GD (2010) The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of *Artemisia annua* L. (Qinghao). *Molecules* **15**: 7603-7698
- Chatzopoulou FM, Makris AM, Argiriou A, Degenhardt J, Kanellis AK (2010) EST analysis and annotation of transcripts derived from a trichome-specific cDNA library from *Salvia fruticosa*. *Plant Cell Rep* **29**: 523-534
- 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. *Plant J* **70**: 480-491
- Covello PS, Teoh KH, Polichuk DR, Reed DW, Nowak G (2007) Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**: 1864-1871
- Covello PS, Teoh KH, Polichuk DR, Reed DW, Nowak G (2007) Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**: 1864-1871
- Crouzet J, Roland J, Peeters E, Trombik T, Ducos E, Nader J, Boutry M (2013) NtPDR1, a plasma membrane ABC transporter from *Nicotiana tabacum*, is involved in diterpene transport. *Plant Mol Biol* **82**: 181-192
- 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 Physiol* **152**: 44-54
- 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
- De Muynck B, Navarre C, Nizet Y, Stadlmann J, Boutry M (2009) Different subcellular localization and glycosylation for a functional antibody expressed in *Nicotiana tabacum* plants and suspension cells. *Transgenic Res* **18**: 467-482
- Debono A, Yeats TH, Rose JK, Bird D, Jetter R, Kunst L, Samuels L (2009) Arabidopsis LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *Plant Cell* **21**: 1230-1238
- Dong L, Miettinen K, Goedbloed M, Verstappen FW, Voster A, Jongsma MA, Memelink J, van der Krol S, Bouwmeester HJ (2013) Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: similar activity but difference in subcellular localization. *Metab Eng* **20**: 198-211
- Edstam MM, Edqvist J (2014) Involvement of GPI-anchored lipid transfer proteins in the development of seed coats and pollen in *Arabidopsis thaliana*. *Physiol Plant* **152**: 32-42
- Edstam MM, Viitanen L, Salminen TA, Edqvist J (2011) Evolutionary history of the non-specific lipid transfer proteins. *Mol Plant* **4**: 947-964
- Goderis IJ, De Bolle MF, François IE, Wouters PF, Broekaert WF, Cammue BP (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. *Plant molecular biology* **50**: 17-27
- 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 Physiol* **51**: 1627-1637
- Huang MD, Chen TL, 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 Physiol* **163**: 1218-1229
- Joosten MH (2012) Isolation of apoplastic fluid from leaf tissue by the vacuum infiltration-centrifugation technique. *Methods Mol Biol* **835**: 603-610
- Kader J-C (1997) Lipid-transfer proteins: a puzzling family of plant proteins. *Trends in Plant Science* **2**: 66-70
- Kanagarajan S, Muthusamy S, Gliszczynska A, Lundgren A, Brodelius PE (2012) Functional expression and characterization of sesquiterpene synthases from *Artemisia annua* L. using transient expression system in *Nicotiana benthamiana*. *Plant cell reports* **31**: 1309-1319

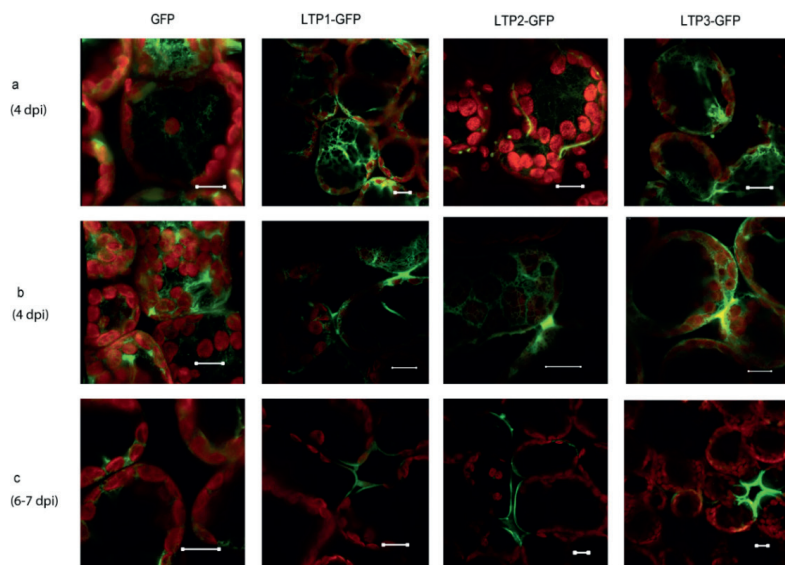
- Kanagarajan S, Tolf C, Lundgren A, Waldenström J, Brodelius PE** (2012) Transient expression of hemagglutinin antigen from low pathogenic avian influenza A (H7N7) in *Nicotiana benthamiana*. *PLoS one* **7**: e33010
- Lange BM, Wildung MR, Stauber EJ, Sanchez C, Pouchnik D, Croteau R** (2000) Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proceedings of the National Academy of Sciences* **97**: 2934-2939
- Lu X, Zhang L, Zhang F, Jiang W, Shen Q, Zhang L, Lv Z, Wang G, Tang K** (2013) AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. *New Phytologist* **198**: 1191-1202
- Ma C, Wang H, Lu X, Wang H, Xu G, 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
- Markham KR, Gould KS, Ryan KG** (2001) Cytoplasmic accumulation of flavonoids in flower petals and its relevance to yellow flower colouration. *Phytochemistry* **58**: 403-413
- Marrs KA, Alfenito MR, Lloyd AM, Walbot V** (1995) A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*.
- Mercx S, Tollet J, Magy B, Navarre C, Boutry M** (2016) Gene Inactivation by CRISPR-Cas9 in *Nicotiana tabacum* BY-2 Suspension Cells. *Frontiers in Plant Science* **7**: 40
- Nieuwland J, Feron R, Huisman BAH, Fasolino A, Hilbers CW, Derksen J, Mariani C** (2005) Lipid Transfer Proteins Enhance Cell Wall Extension in Tobacco. *The Plant Cell* **17**: 2009
- Ökmen B, Etalo DW, Joosten MH, Bouwmeester HJ, Vos RC, Collemare J, Wit PJ** (2013) Detoxification of  $\alpha$ -tomatine by *Cladosporium fulvum* is required for full virulence on tomato. *New Phytologist* **198**: 1203-1214
- Olofsson L, Engström A, Lundgren A, Brodelius PE** (2011) Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua* L. *BMC plant biology* **11**: 45
- Rydén A-M, 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 Med* **76**: 1778-1783
- Salcedo G, Sanchez-Monge R, Barber D, Diaz-Perales A** (2007) Plant non-specific lipid transfer proteins: an interface between plant defence and human allergy. *Biochim Biophys Acta* **1771**: 781-791
- Schillmiller 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
- Soetaert S, Van Neste C, Vandewoestyne M, Head S, Goossens A, Van Nieuwerburgh F, Deforce D** (2013) Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biology* **13**: 220
- Soetaert SS, Van Neste CM, Vandewoestyne ML, Head SR, Goossens A, Van Nieuwerburgh FC, Deforce DL** (2013) Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC plant biology* **13**: 220
- Stukkens Y, Bultreys A, Grec S, Trombik T, Vanham D, Boutry M** (2005) NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. *Plant Physiol* **139**: 341-352
- Teoh KH, Polichuk DR, Reed DW, Covello PS** (2009) Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua* This paper is one of a selection of papers published in a Special Issue from the National Research Council of Canada-Plant Biotechnology Institute. *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
- Ting HM, Wang B, Rydén AM, Woittiez L, Herpen T, Verstappen FW, Ruyter-Spira C, Beekwilder J, Bouwmeester HJ, Krol A** (2013) The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytologist* **199**: 352-366
- Van Herpen T, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J** (2010) *Nicotiana benthamiana* as a production platform for artemisinin precursors. *PLoS One* **5**: e14222
- Voinnet O, Rivas S, Mestre P, Baulcombe D** (2003) Retracted: an enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal* **33**: 949-956
- Wang G, Tian L, Aziz N, Broun P, Dai X, He J, King A, Zhao PX, Dixon RA** (2008) Terpene biosynthesis in glandular trichomes of hop. *Plant physiology* **148**: 1254-1266

- Wang W, Wang Y, Zhang Q, Qi Y, Guo D** (2009) Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. *BMC genomics* **10**: 465
- Wang X, Zhou W, Lu Z, Ouyang Y, O CS, Yao J** (2015) A lipid transfer protein, OsLTPL36, is essential for seed development and seed quality in rice. *Plant Sci* **239**: 200-208
- Witzel K, Shahzad M, Matros A, Mock H-P, Mühling KH** (2011) Comparative evaluation of extraction methods for apoplastic proteins from maize leaves. *Plant Methods* **7**: 1-11
- Yeats TH, Rose JK** (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Science* **17**: 191-198
- Zhang D, Liang W, Yin C, Zong J, Gu F, Zhang D** (2010) <em>OsC6</em>, Encoding a Lipid Transfer Protein, Is Required for Postmeiotic Anther Development In Rice. *Plant Physiology* **154**: 149
- 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 Y, Teoh KH, Reed DW, Maes L, Goossens A, Olson DJ, Ross AR, Covello PS** (2008) The molecular cloning of artemisinic aldehyde  $\Delta^{11}$  (13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *Journal of Biological Chemistry* **283**: 21501-21508

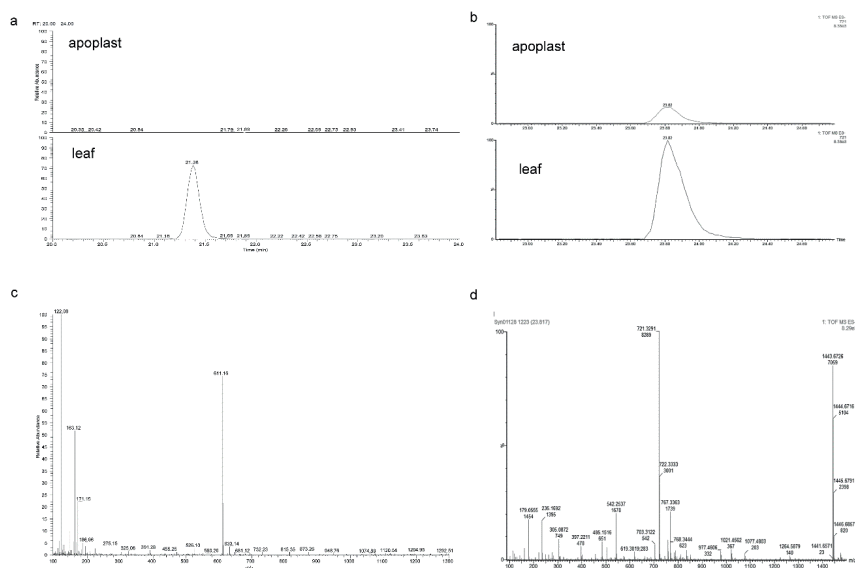
## Supplementary information



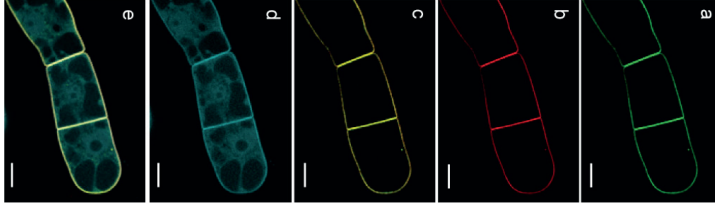
**Fig. S1.** Relative gene expression of *AaPDRs* and *AaLTPs* genes in *A. annua* shoot apex and leaves at different stages. Values indicate the mean fold relative to sample shoot apex, which was normalized to 1. Error bar is SE;  $p < 0.01$ ,  $n = 3$ .



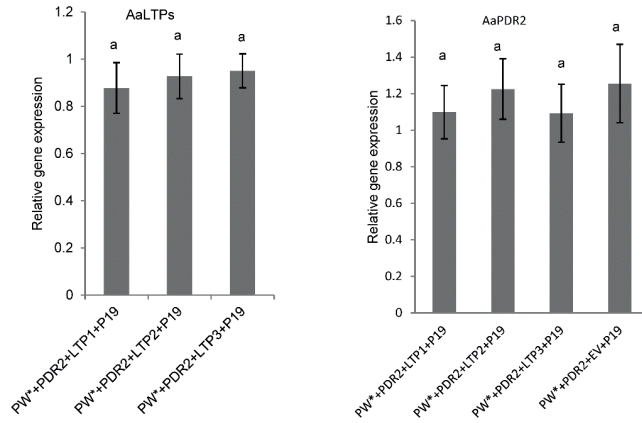
**Fig. S2.** Subcellular localization of AaLTP1-3. Confocal microscopy analysis of *N. benthamiana* leaf infiltrated with GFP, AaLTP1-GFP, AaLTP2-GFP and AaLTP3-GFP separately in (a) intercellular cytoplasmic signal at 4 dpi showing LTPs with accumulation at cell periphery (arrows); (b) excreted LTPs present in bridges between spongy mesophyll cells at 4 dpi; (c) extracellular LTPs present in cell walls of specific cell facets cavity between the mesophyll cells at 6-7 dpi. GFP fluorescence is displayed in green and auto fluorescence of chloroplasts in red. Scale bars represents 10 $\mu$ m.



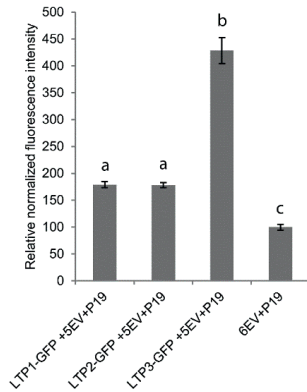
**Fig. S3.** Detection of rutin and (DHAA-Hex3)FA in the apoplast samples under the LC -MS. (a) LC-MS mass-traces for apoplast and leaf extraction samples infiltrated with AN-PW+AaPDRs+AaLTPs at m/z= 611 for rutin detection; (b) LC-MS mass-traces for apoplast and leaf extraction samples infiltrated with AN-PW+AaPDRs+AaLTPs at m/z=721 for (DHAA-Hex3)FA detection; (c) mass spectrum of rutin; (d), mass spectrum of (DHAA-Hex3)FA.



**Fig. S4.** GFP::AaPDR2 is targeted to the plasma membrane. Subcellular localization of GFP::AaPDR2 in *N. tabacum* BY-2 cells. Confocal cross sections show GFP::AaPDR2 (a), FM4-64, a plasma membrane marker (b), overlay between GFP::AaPDR2 and FM4-64 (c), cytosolic mCherry (d), and overlay between GFP::AaPDR2, FM4-64 and mCherry (e). Scale bar = 10  $\mu$ m.

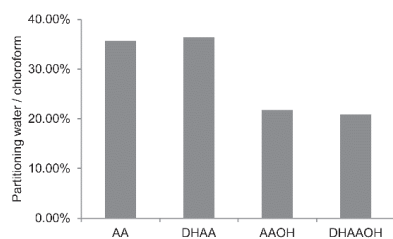


**Fig. S5.** The relative gene expression of *AaPDR2* and *AaLTP1/2/3* in the *N. benthamiana* leaves. There are no significant changes between treatment. Error bar is SE;  $p < 0.01$ ,  $n = 4$ .

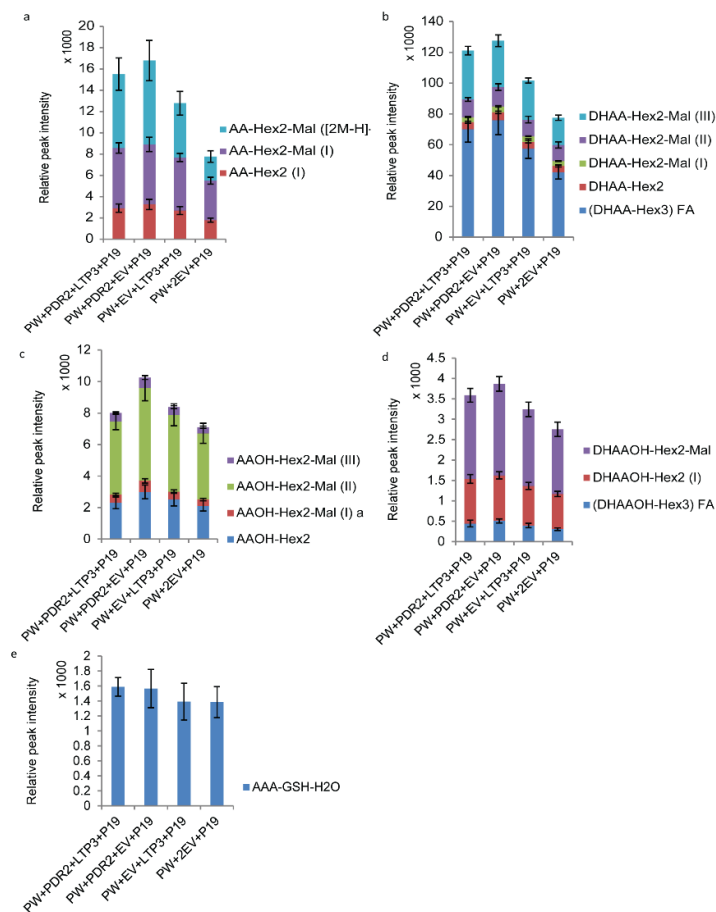


**Fig. S6.** Intrinsic fluorescence of individual LTP-GFP, transiently expressed in *N. benthamiana*. LTP1: AaLTP1-GFP. LTP2: AaLTP2-GFP. LTP3: AaLTP3-GFP. EV: Empty Vector. Relative gene dosage of LTPs-GFP was equal between all experiments by equilibrating with EV where needed. Error bar is SE;  $p < 0.05$ ,  $n = 12$ .

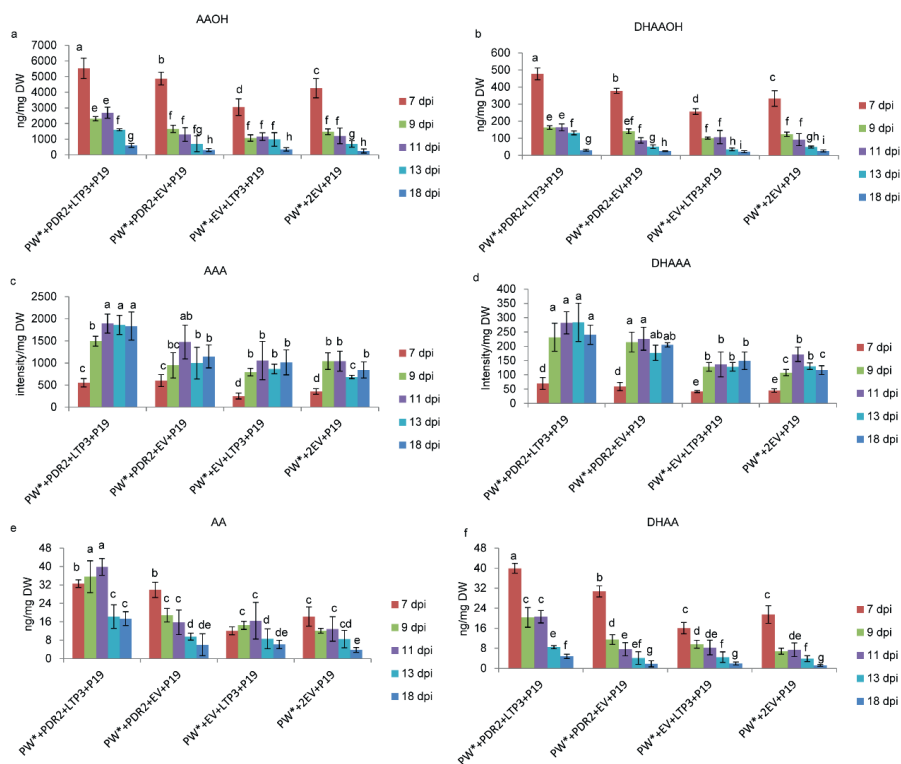




**Fig. S7.** Partitioning of artemisinin-intermediates standards in water vs. chloroform



**Fig. S8.** Artemisinin-related glycosides and glutathione conjugates in leaf extracts of *N. benthamiana* infiltrated with artemisinin pathway, AaLTP3 and AaPDR2 genes as identified by LC-QTOF-MS at 6 dpi. Error bar is SE; n=5.



**Fig. S9.** Production of AN-PW compounds in leaf of *N. benthamiana* infiltrated with artemisinin pathway, *AaLTP3* and *AaPDR2* genes as identified by LC-trip-quad-MS at different time points. Error bar is SE; n=4.

#### Supplementary Table


**Table S1.** Expression of *AaLTPs* and *AaPDRs* in different trichome databases. F: Filamentous trichomes; G: Glandular trichomes; ND: not detected; +: detected.

| Data type                        | Tissues | <i>AaLTP</i> 1 | <i>AaLTP</i> 2 | <i>AaLTP</i> 3 | <i>AaLTP</i> 4 | <i>AaLTP</i> 5 | <i>AaLTP</i> 6 | <i>AaPDR</i> 1 | <i>AaPDR</i> 2 | Reference                         |
|----------------------------------|---------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------------------------|
| RNAseq data                      | G vs. F | G              | G              | ND             | F              | F              | F              | G              | ND             | (Soetaert et al., 2013)           |
| EST                              | G       | +              | +              | +              |                |                |                | +              | ND             | (Covello et al., 2007)            |
| 454 pyrosequencing transcriptome | G       |                |                |                |                |                |                |                | ND             | (Wang et al., 2009)               |
| cDNA library                     | G + F   | +              | +              | +              |                |                |                | ND             | +              | (Bertea et al., 2006)             |
| Trichome database                | G + F   | +              | +              | +              |                |                |                | +              | +              | (Dai et al., 2010)                |
| RNAseq                           | G vs. F | +              | +              | +              |                |                |                | +              | +              | Tang, personal communication      |
| RNAseq                           | G vs. F |                |                |                |                |                |                | G              | G              | Brodelius, personal communication |

Table S2. Primers used in this manuscript

| <i>Primers</i>       | <i>Sequence(5' &gt; 3')</i>                 |
|----------------------|---------------------------------------------|
| <i>β-actin_Fw</i>    | CCAGGCTGTTTCAGTCTCTGTAT                     |
| <i>β-actin_Rev</i>   | CGCTCGGTAAGGATCTTCATCA                      |
| <i>AaADS_Fwd</i>     | GGAGTATGCCCAAACCTTGA                        |
| <i>AaADS_Rev</i>     | TCGTCTCCCATACGTGTGAA                        |
| <i>AaLTP1_Fwd</i>    | GCAGTAGAAGGCGAGGTGAC                        |
| <i>AaLTP1_Rev</i>    | AAGCAGCTTGTTGTGTCAGCA                       |
| <i>AaLTP2_Fwd</i>    | AAACTCCGCTGCTAAAACGA                        |
| <i>AaLTP2_Rev</i>    | TCGGTGCTAAGGCTGATCTT                        |
| <i>AaLTP3_Fwd</i>    | AGCACCTATGCTGAGGCTA                         |
| <i>AaLTP3_Rev</i>    | AGTTTGACGATCAGGGTTG                         |
| <i>AaPDR-F</i>       | TGATGACBRTYGCRYTYACC + GGTGGAGATGGTAYTAYTGG |
| <i>AaPDR-R</i>       | ATGACHRTYKCMRTYACDCC + GGTGGAKDTGGTAYTAYTGG |
| <i>AaPDR1_Fwd</i>    | AGAACGAGCTGCTGGAATGT                        |
| <i>AaPDR1_Rev</i>    | GAGCAACGAGCAAAACATGA                        |
| <i>AaPDR2_Fwd</i>    | GTTGGTCTGGATTGGTGCT                         |
| <i>AaPDR2_Rev</i>    | CCGAGTCTTGGTTGTTGGTT                        |
| <i>AaLTP1-F</i>      | TTCCATGGTTGGAAAGGTTGTGTTGG                  |
| <i>AaLTP1-R</i>      | GCGGCCGCTCATTGTATCGTTGAGCAAT                |
| <i>AaLTP2-F</i>      | TTCCATGGCAAGTATGACAATGAG                    |
| <i>AaLTP2-R</i>      | GCGGCCGCTCACTGACCTTGTGTCAGT                 |
| <i>AaLTP3-F</i>      | TTCCATGGCAAGGATGGCAATGATTGT                 |
| <i>AaLTP3-R</i>      | GCGGCCGCTCACTGCACCTTGGTGCAGT                |
| <i>EGFP_c-term-F</i> | GAAGGAGCGGCCGCGAGCAAGGGCGAGGAGCTG           |
| <i>EGFP_c-term-R</i> | CACAAAGAGCTCTTTATACAGCTCGTCCATGC            |
| <i>AaLTP1_gfp-F</i>  | AACACCGGATCCAATGGTTGGAAAGGTTGTGTTGGTC       |
| <i>AaLTP1_gfp-R</i>  | GACAAAGCGGCCGCTTGTATCGTTGAGCAATCAGTTGTGG    |
| <i>AaLTP2_gfp-F</i>  | AACACCGGATCCAATGGCAAGTATGACAATGAGGGTTTTATG  |
| <i>AaLTP2_gfp-R</i>  | GACAAAGCGGCCGCTTGACCTTGTGTCAGTCGGTG         |
| <i>AaLTP3_gfp-F</i>  | AACACCGGATCCAATGGCAAGGATGGCAATGATTGTTTC     |
| <i>AaLTP3_gfp-R</i>  | GACAAAGCGGCCGCTGCACCTTGGTGCAGT              |
| <i>AaLTP1_rfp-F</i>  | CACCATGGTTGGAAAGGTTGTGTTG                   |
| <i>AaLTP1_rfp-R</i>  | TTGTATCGTTGAGCAATCAGTTG                     |
| <i>AaLTP2_rfp-F</i>  | CACCATGGCAAGTATGACAATGAGG                   |
| <i>AaLTP2_rfp-R</i>  | CTTGACCTTGTGTCAGTCG                         |
| <i>AaLTP3_rfp-F</i>  | CACCATGGCAAGGATGGCAATG                      |
| <i>AaLTP3_rfp-R</i>  | CTGCACCTTGGTGCAGTC                          |





## **Chapter 5**

# **Non-specific Lipid Transfer Proteins exhibit specificity for feverfew sesquiterpene lactones**

Arman Beyraghdar Kashkooli, Aalt D.J. van Dijk,  
Harro Bouwmeester, Alexander van der Krol

**Under review in *Metabolic Engineering***

## Abstract

Hydrophobic specialized metabolites produced by plants often accumulate in specialized tissues, outside the cells in which they are synthesized. Transport over the plasma membrane (PM) is presumably carried out by specialized membrane transporters, but it is still not clear how these compounds are subsequently transported over the hydrophilic cell wall. Here we demonstrate that so-called Non-specific Lipid Transfer Proteins (NsLTPs) from feverfew (*Tanacetum parthenium*) have acquired a specialized function in the transport of specific lipophilic products (sesquiterpene lactones) produced in feverfew trichomes. Of the eight feverfew LTP candidate genes analysed in a sesquiterpene export assay using transient expression in *Nicotiana benthamiana*, two (TpLTP1 and TpLTP2) specifically improved costunolide export and one (TpLTP3) highly specifically improved parthenolide export. Subsequent characterization of these three LTPs in a substrate exclusion assay identified TpLTP3 as most effective in blocking influx of both costunolide and parthenolide when the substrate was infiltrated into the apoplast. TpLTP3 is special in having a GPI-anchor domain, which we showed is essential for the activity of TpLTP3, while addition of this domain to TpLTP1 resulted in loss of TpLTP1 activity.

**Keywords:** lipid transfer protein; transport; specialized metabolism; heterologous expression; sesquiterpenoids.

## Introduction

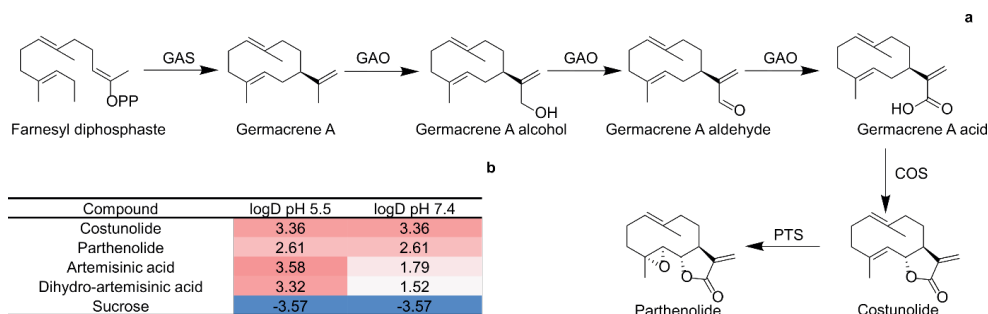
In many plants specialized metabolites are synthesised in specialised cells and subsequently transported over the PM and cell wall for accumulation in the extracellular space. The bioactivity of many of these metabolites, such as for example sesquiterpenoids, is based on their chemical reactivity with proteins and other cellular components, and transport into the extracellular space is likely required to avoid that these compounds react with the plant's own cellular components. The medicinal sesquiterpenoids and their precursors produced by (for instance) feverfew and *Artemisia annua* are lipophilic as indicated by the chemical distribution coefficient (LogD (Leo et al., 1971)) (Fig. 1b). Such lipophilic compounds may therefore show low spontaneous diffusion over the hydrophilic cell wall once transported over the PM of the glandular trichome cells, especially since the pH of the cell wall is lower than that of the cytosol. Here we studied the extracellular transport of two medicinal sesquiterpenoids produced by feverfew: costunolide and parthenolide, both of which are active against ovary, prostate, breast and oral cancer as well as leukaemia (Choi et al., 2005; Zunino et al., 2007; Hsu et al., 2011; Yang et al., 2011; Yu et al., 2015). In feverfew these compounds are produced in the glandular trichomes on the ray florets and ovaries of the flowers and are secreted and accumulated in the subcuticular space of these trichomes (Majdi et al., 2011).

Biosynthesis of costunolide and parthenolide from the precursor germacrene A (GA) takes place in the cytosol and on the ER membrane by a sesquiterpene synthase and cytochrome P450s (Fig. 1a). The mechanism responsible for the transport of costunolide and parthenolide into the extracellular space of the trichomes is not known. Transport of these compounds within the cell could be by spontaneous or vesicular transport (Moreau and Cassagne, 1994; Ting et al., 2015). Subsequent transport over the PM may occur through the action of different membrane transporters. For instance, two *A. annua* ABC type G PM transporters have been identified that are involved in the transport of (dihydro)artemisinic acid ((DH)AA) over the PM (Personal communication Prof. Marc Boutry). We postulate that ABC type G transporters are also involved in the secretion of costunolide and parthenolide in feverfew. However, once on the outside of the PM the hydrophobic sesquiterpenoids are confronted with the hydrated polysaccharides of the cell wall which forms a hydrophilic environment suggesting that transport of sesquiterpenoids over the cell wall must somehow be facilitated. Indeed, also modelling suggests that the actual observed dynamics of (hydrophobic) volatiles emitted by plants requires the involvement of something that facilitates passage over the hydrophilic cell wall (Widhalm et al., 2015). For the product of the artemisinin biosynthesis pathway (DH)AA it was shown that an *A. annua* NsLTP (AaLTP3) enhances sequestration of (DH)AA into the apoplast (Wang et al., 2016). Here we describe the discovery of a highly specialised function of feverfew LTPs in the transport of two products of the feverfew sesquiterpenoid



biosynthetic pathway. Lipid Transfer Proteins (LTPs) are small proteins usually with a size of 7-9 KDa, which contain a highly conserved 8 cysteine motif (Kader, 1996). These cysteines connect to each other in four disulphide bridges resulting in four  $\alpha$ -helices which together form a hydrophobic cavity (Shin et al., 1995) which in *in vitro* assays can bind a range of lipid cargo molecules, hence their name: Non-specific LTPs (de Oliveira Carvalho and Gomes, 2007). NsLTPs contain a secretion signal and are predicted to be secreted into the cell wall of plants. Indeed, extracellular localisation has been confirmed for multiple LTPs in several different plant species (Sterk et al., 1991; Coutos-Thevenot et al., 1993; Choi et al., 2005; Wang et al., 2016).

The genes for parthenolide biosynthesis (*TpGAS*, *TpGAO*, *TpCOS* and *TpPTS*) in feverfew have been identified (Fig. 1a) and were characterized by expression in yeast and *Nicotiana benthamiana* (Liu et al., 2011; Majdi et al., 2011; Liu et al., 2014). Reconstitution of the parthenolide biosynthesis pathway in *N. benthamiana* – using transient expression – results in the accumulation of conjugates of costunolide and parthenolide with glutathione (GSH) or cysteine (Cys), while only very low levels of free costunolide and parthenolide are detected (Liu et al., 2011; Majdi et al., 2011; Liu et al., 2014). Here we tested whether the presence of free costunolide and parthenolide in such transient expression assays is due to active transport to the apoplast, thus escaping from intracellular conjugation. In the background of this low intrinsic export activity for costunolide and parthenolide in *N. benthamiana* leaves, we tested eight feverfew Lipid Transfer Proteins (TpLTPs) for their effect on the accumulation of free costunolide and parthenolide. The *TpLTP* genes were selected based on their co-expression pattern with costunolide/parthenolide biosynthetic pathway genes over six developmental stages of feverfew flowers (Majdi et al., 2011). The activity of each TpLTP was tested in combination with the costunolide and parthenolide biosynthesis pathway by transient co-expression of the *TpLTPs* and costunolide/parthenolide biosynthesis genes in *N. benthamiana*. Product yield analysis of extracts of the leaf and the apoplast indicated that most LTPs have no, or only a small effect on costunolide export, while TpLTP1 and TpLTP2 cause a ~2.5-fold increase in costunolide export. The co-expression of TpLTP3 with the parthenolide pathway resulted in a ~23-fold increase in free parthenolide production, while the other seven TpLTPs did not have any affect. In an alternative assay, based on preventing bioconversion of costunolide and parthenolide infiltrated into *N. benthamiana* leaves to their conjugates, we show that TpLTP3 is most effective in blocking bioconversion of both costunolide and parthenolide. Combined, the results suggest that in feverfew glandular trichomes, so-called Non-specific Lipid Transfer Proteins (NsLTPs) have acquired a specialized function by aiding sesquiterpene lactone transport over the cell wall. Implications for metabolic engineering of (sesqui)terpenes and potential wider implications for the role of NsLTPs in plants are discussed.



**Figure 1. Biosynthetic pathway of parthenolide and partitioning coefficient of some sesquiterpene lactones and sucrose.** (a) Biosynthesis pathway of parthenolide (Liu et al., 2011; Liu et al., 2014). *GAS*: germacrene A synthase; *GAO*: germacrene A oxidase; *COS*: costunolide synthase and *PTS*: parthenolide synthase. (b) log D (partitioning of a chemical compound between an organic versus aqueous phases; corrected for the pH) of costunolide, parthenolide, (dihydro)artemisinic acid and sucrose. The lower and/or more negative the number, the higher the partitioning in the aqueous phase (= better water solubility).

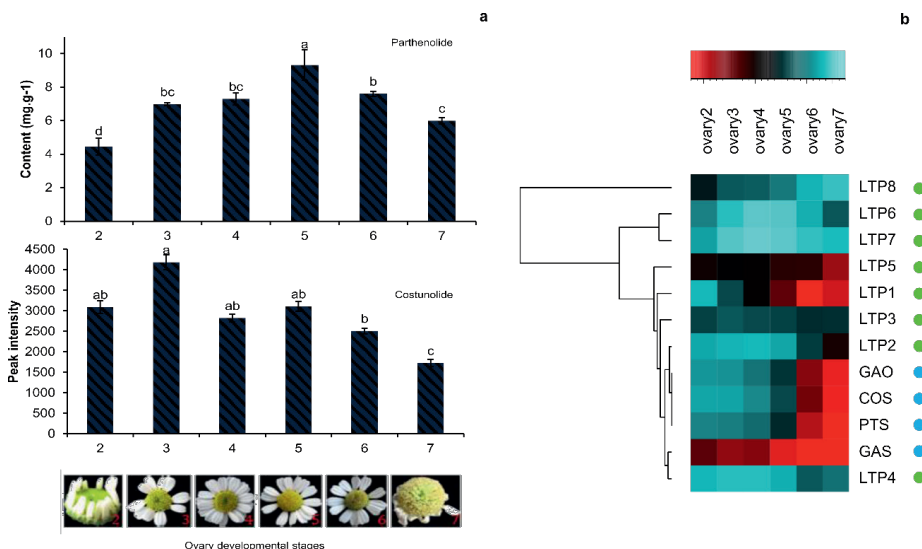
## Results

### Identification of candidate Lipid Transfer Protein genes from co-expression analysis

Based on our experience with the LTPs that play a role in artemisinin biosynthesis in *A. annua* (Wang et al., 2016), we searched for candidate feverfew LTPs that may be active in the costunolide/parthenolide biosynthesis pathway. First, a feverfew trichome enriched reference transcriptome assembly was obtained. The number of contigs longer than 500 bp was 29,799 and the total contig length 27 Mb. Subsequently, the expression profile over six ovary developmental stages was obtained from triplicate cDNA samples per developmental stage. RNA-seq reads were mapped to the reference transcriptome assembly. On average, 35M $\pm$ 10M reads per developmental stage were obtained, and the average alignment rate was 79%  $\pm$  4%. Analysis of the expression profiles indicated that 5,414 contigs were differentially expressed between at least two subsequent stages. To further narrow down this set to the most relevant genes, BLAST was used to identify the known costunolide/parthenolide pathway genes as well as sequences with similarity to LTPs, P450s and ATP binding cassette membrane transporters. LTP sequences were selected based on GO terms associated with the BLAST hits and the conserved 8 cysteine motif of NsLTPs (Kader, 1996; Supplementary Fig. 4a). This resulted in a set consisting of 13 putative LTP, 75 putative P450, and 35 putative ABC transporter sequences (Supplementary Fig. 1).

The expression profile of the known costunolide/parthenolide biosynthesis genes (*GAS*, *GAO*, *COS*, and *PTS*) was used to identify the *TpLTP* genes that show a similar expression profile, resulting in selection of eight candidate LTP sequences (Fig. 2). *LTP6* and 7 were clustered

together while *LTP1* to *LTP5* were closely clustered together with pathway genes. *LTP8* was sitting in a separate branch in the cluster analysis. For these eight candidate TpLTPs, the coding sequence was cloned into a binary expression vector under the control of the CamV 35S promoter. The LTPs were functionally analysed by transient co-expression with the costunolide or parthenolide biosynthesis pathway genes in *N. benthamiana*.



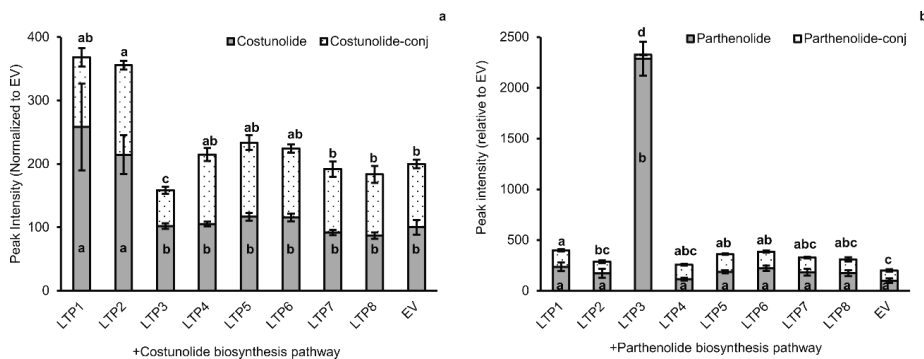
**Figure 2. Accumulation of costunolide and parthenolide and expression pattern of feverfew parthenolide biosynthesis pathway genes and Lipid Transfer Proteins (LTPs)** (a) Parthenolide and costunolide content of feverfew flowers in different developmental stages (Majdi et al., 2011; Liu et al., 2014). (b) Complete linkage cluster analysis of feverfew sesquiterpene lactone biosynthesis pathway genes and LTPs. Heatmap representation of log<sub>2</sub> transformed expression data of genes during six developmental stages of feverfew flowers. Each data point is the mean value of the three measured biological replicates. Green circles represent LTPs whereas blue circles represent feverfew sesquiterpene lactones biosynthesis pathway genes.

### The effect of TpLTPs on costunolide and parthenolide levels

In order to test whether any of the eight TpLTPs has an effect on costunolide production, the costunolide pathway genes (*AtHMGR*, *TpGAS1.5*, *CiGAO*, *CiCOS*) were co-expressed with each individual TpLTP in 4 weeks old *N. benthamiana* leaves by co-infiltration of *Agrobacterium* batches carrying the different expression constructs. The relative dosage of the costunolide pathway genes was kept equal between experiments by using empty vectors (EV) where needed and a *p19* construct was used to repress the silencing mechanisms of *N. benthamiana* (Voinnet et al., 2003). Ectopic expression of the LTPs together with the costunolide and parthenolide pathway caused necrotic lesions in leaves at 5 days post-infiltration (dpi). Therefore, leaves were harvested at 3 and 4 dpi for the costunolide and parthenolide pathway, respectively. After harvest, metabolites were extracted for targeted

metabolite analysis using UPLC-triple quadrupole-MS. The levels of free costunolide (non-conjugated) and costunolide conjugates (costunolide-conj) in leaves expressing the costunolide pathway+EV were set at 100%. Figure 3a shows that only co-expression of TpLTP1 and TpLTP2 results in a substantial increase in both free and conjugated costunolide (2.4 and 2.65 times, respectively; Fig. 3a). Of the other LTPs, only TpLTP3 had a significant effect, slightly decreasing the level of conjugated costunolide (Fig. 3a).

Subsequently, the eight TpLTPs were tested for their effect on parthenolide pathway product accumulation. For this, the parthenolide pathway genes (*AtHMGR*, *TpGAS1.5*, *CiGAO*, *CiCOS*, *TpPTS*) were co-expressed with each of the eight individual *TpLTPs* in 4 weeks old *N. benthamiana* leaves by agro-infiltration of the expression constructs. The level of free parthenolide and parthenolide conjugates in leaves expressing the parthenolide pathway+EV were again set at 100%. In contrast to their effect on costunolide, neither TpLTP1 nor TpLTP2 had an effect on free or conjugated parthenolide levels (Fig. 3b). However, expression of TpLTP3 with the parthenolide biosynthesis pathway genes resulted in a remarkable 23-fold increase in free parthenolide, while the level of conjugated parthenolide decreased (Fig. 3b). Thus, three of the eight TpLTPs that showed co-expression with the biosynthetic pathway genes (Fig. 2) have an effect on the accumulation of sesquiterpenes, with TpLTP1 and TpLTP2 showing specificity for costunolide and TpLTP3 for parthenolide.



**Figure 3. Quantification of the effect of individual feverfew Lipid Transfer Proteins (LTPs) on metabolite accumulation as analysed by UPLC-MRM-MS. (a)** TpLTP effect on the accumulation of costunolide biosynthesis pathway products at 3 dpi. **(b)** TpLTP effect on the accumulation of parthenolide biosynthesis pathway products at 4 dpi. All values are normalized to the levels produced by the costunolide and parthenolide biosynthesis pathways in combination with an empty vector (EV) (n=4). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

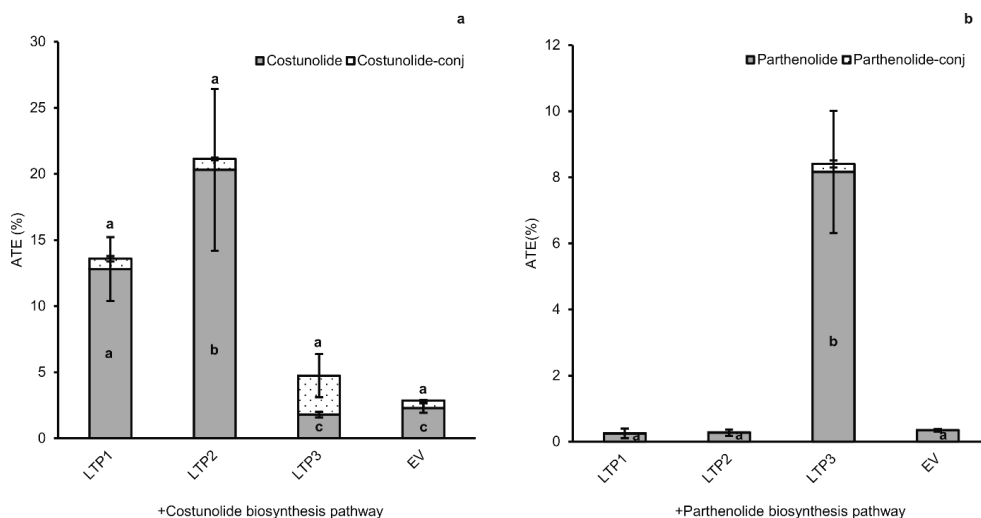
Analysis of the free and conjugated costunolide levels in the parthenolide-pathway experiment, shows that TpLTP1 and TpLTP2 are still active in enhancing accumulation of free costunolide, but they do not contribute to accumulation of free or conjugated parthenolide (Supplementary Fig. 2). This indicates that the enhanced costunolide accumulation in the presence of TpLTP1 or TpLTP2 is not available for TpPTS activity. This would be consistent

with the LTPs mainly having an effect on product accumulation in the apoplast of the *N. benthamiana* leaves, where costunolide is no longer available for TpPTS. Therefore, we next analysed the effect of TpLTP1 to 3 on costunolide and parthenolide pathway product accumulation in the apoplast.

### **Expression of *TpLTP1* to 3 results in increased levels of costunolide and parthenolide in the apoplast**

In feverfew flowers, the costunolide level peaks at stage 3 of flower development, while the parthenolide level peaks at stage 5 of flower development (Fig. 2a). Also the expression profile of *TpLTP1* and *TpLTP2* shows an earlier peak (stage 2) than *TpLTP3* (stage 3) (Fig. 2b). Moreover, more than 90% of the parthenolide and costunolide in feverfew is present in the extracellular space of the glandular trichomes (Liu et al., 2014), indicating an efficient export of these sesquiterpenes from the cell into the apoplast. To test whether the three active TpLTPs boost pathway product accumulation in the apoplast, the apoplast of *N. benthamiana* leaves transiently expressing the pathway-LTP combinations was extracted and analysed. Leaves with transient expression of either the costunolide or parthenolide pathway in absence or presence of the individual TpLTPs were harvested at 3 or 4 dpi (for costunolide or parthenolide pathway, respectively), briefly vacuum infiltrated with water, followed by slow centrifugation for non-destructive extraction of infiltrated water from the extracellular space. In general, approximately 75% of the infiltrated water was retrieved after centrifugation. After extraction of the apoplast wash fluid, leaves were extracted with MeOH for analysis of the remaining, non-apoplastic, products. Apoplastic wash fluid and leaf extracts were analysed by UPLC-MRM-MS for quantification of free and conjugated compounds. From these data the ratio of compound level in the apoplast of one leaf and the level in the remaining leaf material (with correction for non-extracted apoplast fluid) was calculated. We consider the ratio of ‘apoplast content/leaf content’ to be a measure for overall Apoplast Transport Efficiency [ATE = 100×(apoplast content/leaf content)]. The data show that also without co-expression of TpLTPs, both free and conjugated costunolide are detected in the apoplast, indicating an intrinsic transport activity for costunolide in *N. benthamiana*. The presence of conjugated costunolide in the apoplast can either be from export of conjugated costunolide or from costunolide conjugating activity (to GSH) in the apoplast of *N. benthamiana* leaves. Without TpLTPs approximately 2% of free and conjugated forms of costunolide are in the apoplast ( $ATE_{\text{costunolide}} + ATE_{\text{costunolide-conj}} = 4$ ), while with *TpLTP1* ~12% and with *TpLTP2* ~20% of free costunolide is in the apoplast (Fig. 4a). The effect of *TpLTP1* and 2 on free costunolide in the apoplast seems to be specific, because they had no significant effect on transport of costunolide conjugates into the apoplast ( $P < 0.05$ ) (Fig. 4a), while *TpLTP3* has no or even inhibitory effect on the accumulation of free and conjugated costunolide (Fig. 3a), *TpLTP3*

seems to enhance accumulation of conjugated costunolide in the apoplast ( $ATE_{\text{costunolide}}$  increases from  $\sim 0.5$  to  $\sim 3$ ) which is not significantly different to the EV (Fig. 4a).



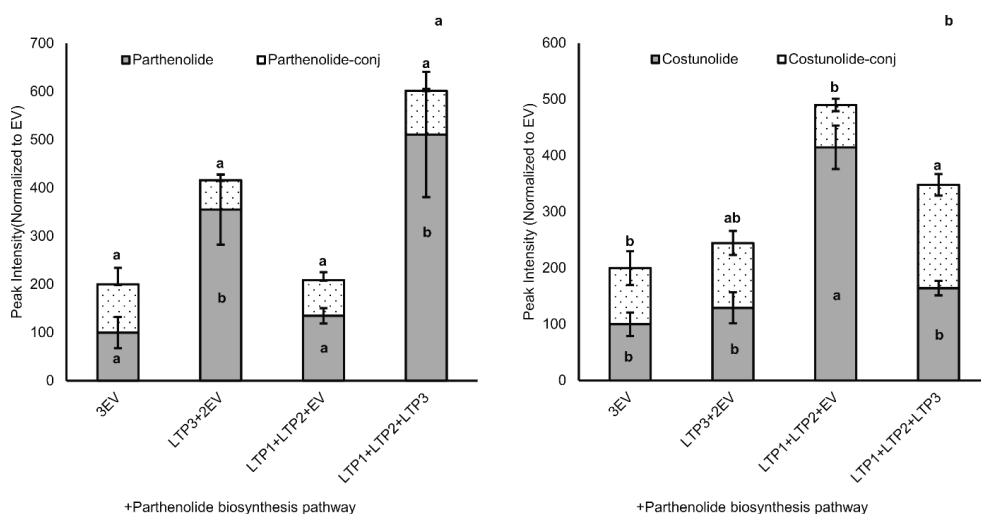
**Figure 4. Quantification of the effect of TpLTPs on metabolite accumulation in the apoplast as analysed by UPLC-MRM-MS. (a)** Costunolide ( $ATE_{\text{costunolide}}$ ) and costunolide-conjugate apoplast transport efficiency; grey bars represent free costunolide and black-dotted bars represent sum of costunolide-conj. **(b)**  $ATE_{\text{parthenolide}}$  and  $ATE_{\text{parthenolide-conj}}$ ; grey bars represent free parthenolide and black-dotted bars represent the sum of parthenolide-conj. All values are normalized to yield in (costunolide/parthenolide) pathway +EV (empty vector) ( $n=4$ ). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

Similarly, the  $ATE_{\text{parthenolide}}$  was determined for leaves expressing the parthenolide pathway genes with or without the TpLTPs. Results show that without TpLTPs,  $\sim 0.2\%$  of free parthenolide is in the apoplast ( $ATE_{\text{parthenolide}}=0.2$ ). TpLTP1 and TpLTP2 had no effect on  $ATE_{\text{parthenolide}}$ , but *TpLTP3* increased  $ATE_{\text{parthenolide}}$  from 0.2 to 8%. Again this effect seems to be specific as *TpLTP3* did not enhance the level of conjugated parthenolide in the apoplast (Fig. 4b). When co-expressed with the parthenolide pathway, both *TpLTP1* and 2 enhanced the level of free costunolide in the apoplast ( $ATE_{\text{costunolide}}$  increases from 0.2 to 12 and 14, respectively) (Supplementary Fig. 3). Overall, the results show that TpLTP1 and TpLTP2 are more active in apoplast accumulation of costunolide, while TpLTP3 is more active in apoplast accumulation of parthenolide.

### Parthenolide accumulation not significantly increased by TpLTP1/2, but costunolide accumulation reduced upon co-expression of TpLTP1/2/3

In feverfew the three LTPs are presumably expressed together at the same time and result in extracellular accumulation of costunolide, but mostly parthenolide. Therefore, we performed an experiment in *N. benthamiana* where all three functional LTPs were coexpressed with the parthenolide pathway. The data show that there is no significant difference between

TpLTP3+2EV and TpLTP3+TpLTP2+TpLTP1 for free parthenolide production, although the yield of parthenolide and parthenolide conjugates tended to be highest when all three TpLTPs were coexpressed (Fig. 5a). Analysis of costunolide and its conjugates showed that TpLTP3 did not contribute to free costunolide accumulation, while co-expression of TpLTP1 and 2 significantly enhanced free costunolide accumulation. Co-expression of TpLTP3 with TpLTP1 and 2 ruled out the positive effect of the latter two LTPs (Fig. 5b). We note that accumulation of parthenolide in these experiments did not reach to the same level as for the experiment shown in Figure 3. This may be explained by a dilution effect on the parthenolide biosynthesis pathway gene dosage. In the experiment of Figure 3 the *PTS* expression construct was co-infiltrated with six other expression constructs, while in the experiment of Figure 5, the *PTS* expression construct was co-infiltrated with eight other constructs. At a high number of co-infiltrated expression constructs not all cells of the infiltrated leaf may be transformed by all gene constructs, which would reduce overall parthenolide biosynthesis and LTP activity.



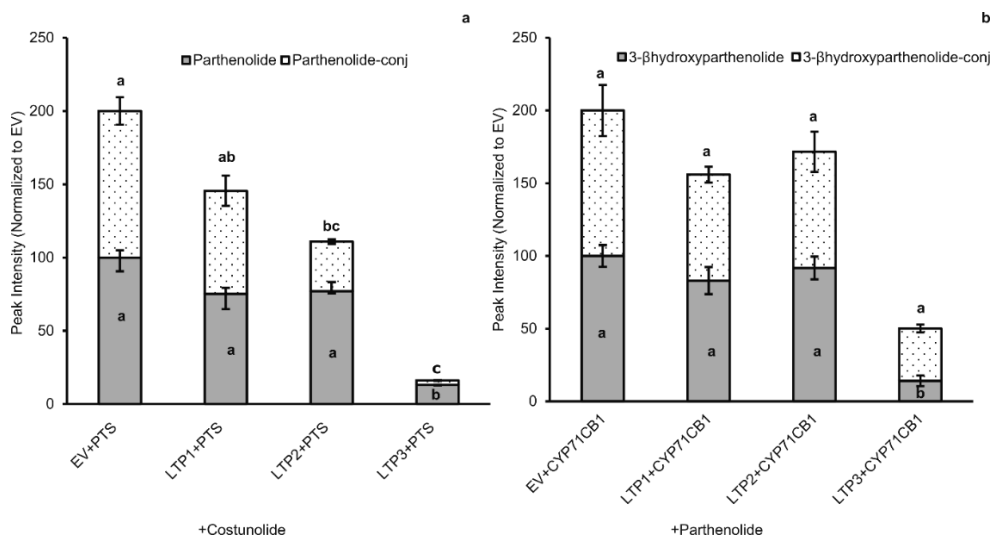
**Figure 5. UPLC-MRM-MS quantification of products of the parthenolide pathway transiently expressed in combination with TpLTPs in *N. benthamiana*.** (a) level of parthenolide (grey bars) and parthenolide conjugates (black-dotted bars) in leaves expressing the parthenolide biosynthesis pathway genes (*HMGR*, *GAS*, *GAO*, *COS*, *PTS*) co-expressed with different TpLTPs (b) level of costunolide and costunolide conjugates in leaves expressing the parthenolide biosynthesis pathway with different TpLTPs. (n=4). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

### TpLTPs prevent influx of costunolide and parthenolide into leaf cells

An alternative way to test for LTP activity and specificity *in planta* is by a substrate exclusion assay (Wang et al., 2016). In such an assay, a substrate is infiltrated into the apoplast of leaves and the uptake of the substrate by the cells is estimated by analysis of the conversion of the



substrate inside the cells. If ectopic expression of an LTP (partly) prevents this bioconversion, that suggests that the LTP helps to retain the compound in the apoplastic space outside the cellular PM. For monitoring the influx of infiltrated costunolide into cells we expressed TpPTS in leaves of *N. benthamiana* and used parthenolide production as a measure for the influx of costunolide. *PTS*, in combination with an empty vector or a *TpLTP* were expressed in *N. benthamiana* leaves and at 3 dpi leaf discs were taken from the agro-infiltrated leaves. The leaf discs were vacuum infiltrated with the substrate costunolide, were harvested and rinsed at 2 hr post substrate infiltration and were subsequently analysed for costunolide and parthenolide. In leaves expressing *PTS* and *EV* and infiltrated with costunolide, both parthenolide and parthenolide conjugates were detected, indicating efficient influx of costunolide into the cells and bioconversion by the ectopically expressed parthenolide synthase (Fig. 6a). Co-expression of *TpLTP1* or *TpLTP2* with *PTS* reduced parthenolide production by approximately 20% (Fig. 6a). In contrast, co-expression of *TpLTP3* with *PTS* blocked bioconversion of costunolide to parthenolide by almost 90% (Fig. 6a), indicating efficient costunolide exclusion from the cells by *TpLTP3*. In a similar exclusion assay with parthenolide as infiltrated substrate, the ectopic expression of a parthenolide 3 $\beta$ -parthenolide hydroxylase (CYP71CB1 (Liu et al., 2014)) was used to monitor bioconversion of parthenolide to 3 $\beta$ -hydroxyparthenolide and 3 $\beta$ -hydroxyparthenolide conjugates (Fig. 6b). In this exclusion assay, the co-expression of either *TpLTP1* or *TpLTP2* with CYP71CB1 could block parthenolide influx by 18% or 9%, respectively (Fig. 6b), while co-expression of *TpLTP3* with CYP71CB1 blocks influx of parthenolide by 86%. Thus, in the substrate exclusion assays with the three LTPs, only *TpLTP3* efficiently blocks both costunolide and parthenolide influx into the cell (Fig. 6a and 6b).

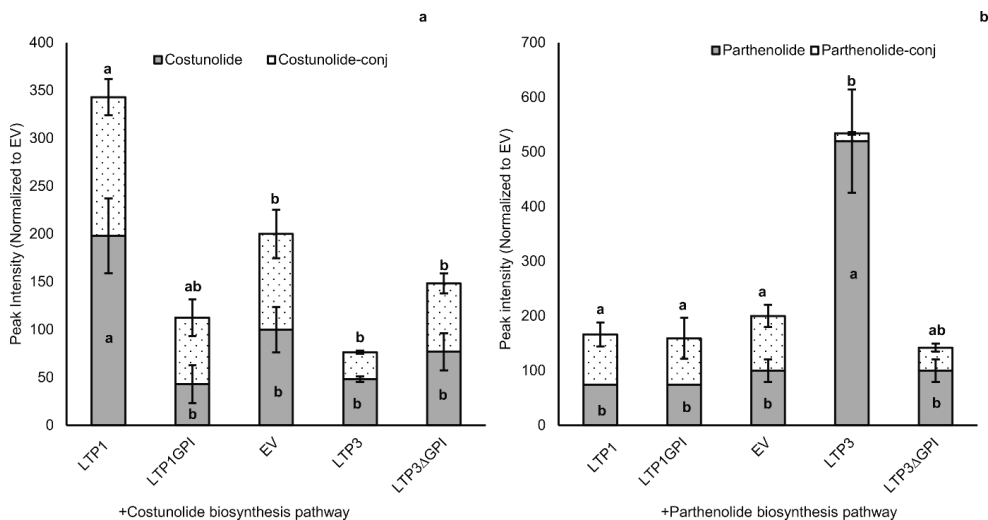


**Figure 6. Intercellular bioconversion of costunolide to parthenolide by parthenolide synthase and parthenolide to 3β-hydroxy parthenolide by 3β-parthenolide hydroxylase in *N. benthamiana* leaves.** Parthenolide synthase or 3β-parthenolide hydroxylase were infiltrated into *N. benthamiana* leaves together with different LTPs. After 3 days, substrates (either costunolide or parthenolide (1 mg ml<sup>-1</sup>) were vacuum infiltrated into leaf disks of 1 cm diameter. EV was used as a control. Metabolites were extracted with methanol and analysed with LC-MS. **(a)** UPLC-MRM-MS analysis of extracted metabolites from leaves expressing parthenolide synthase, vacuum infiltrated with costunolide and incubated for 120 minutes **(b)** Bioconversion of parthenolide to 3β-hydroxy parthenolide. Apoplast vacuum infiltration of parthenolide and extraction of leaf metabolites at 120 min (analysed by LC-Orbitrap-FTMS). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

### TpLTP3 has a GPI anchor domain which is essential for activity

The AA coding sequence of TpLTP3 indicates that it contains an extended domain with a structural motif that is associated with addition of a GPI-anchor. TpLTP3 therefore belongs to the LTPG (GPI-anchored) class of LTPs. In Arabidopsis and cotton, LTPG proteins are involved in the deposition of leaf cuticular waxes and transport of phosphatidylinositol monophosphates in fibres, respectively (Kim et al., 2012; Deng et al., 2016). The role of the GPI anchor domain in these processes is unknown. In order to identify the role of the GPI-anchor domain of TpLTP3 we produced a truncated version of TpLTP3 by removing the GPI anchor site (TpLTP3<sup>ΔGPI</sup>) (Supplementary Fig. 4b). In addition, we tested whether the GPI anchor domain of TpLTP3 enhances the functionality of TpLTP1 by linking it to the C terminus of TpLTP1 (TpLTP1<sup>GPI</sup>). Co-expression of TpLTP3<sup>ΔGPI</sup> with either the parthenolide or costunolide biosynthesis pathway showed that the truncated protein no longer increases parthenolide accumulation (Fig. 7a). The deletion of the anchor motif did not affect costunolide production. The addition of the GPI anchor motif to TpLTP1 resulted in a loss of TpLTP1 activity in costunolide accumulation (Fig. 7a). Addition of the GPI anchor to TpLTP1 did not have a positive effect on accumulation of parthenolide (Fig. 7b). The GPI

anchor domain therefore seems to be essential for the function of TpLTP3, but does not provide parthenolide transport activity to TpLTP1. Moreover, adding the anchor compromises the activity of *TpLTP1* in costunolide production.



**Figure 7.** Effect of addition of a GPI anchor to LTP1 and truncation of the GPI anchor from LTP3 on metabolite accumulation in *N. benthamiana*, as analysed by UPLC-MRM-MS. (a) effect of the addition of a GPI anchor to LTP1 and removal of the GPI anchor from LTP3 on the production of costunolide (grey bars) and costunolide conjugates (black-dotted bars) when transiently coexpressed with the costunolide biosynthesis pathway genes and (b) effect of addition of GPI anchor to LTP1 and removal of LTP3 GPI anchor on the production of parthenolide (grey bars) and parthenolide conjugates (black-dotted bars) upon transient coexpression with the parthenolide biosynthesis pathway genes (n=4). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

## Discussion

### Specific boosting of sesquiterpene accumulation by individual feverfew LTPs

Here we discovered that three feverfew trichome LTPs are specific for different sesquiterpene lactones stored in the trichomes. This was assessed by characterizing the efficiency of the three feverfew LTPs in increasing product formation of a transiently reconstituted costunolide and parthenolide biosynthesis pathway in *N. benthamiana* leaves. The increase in free costunolide production by coexpression of TpLTP1 and TpLTP2 with the costunolide pathway genes (Fig. 3a) seems to be caused by the enhanced accumulation of free costunolide in the apoplast (Fig. 4a). In contrast, coexpression of TpLTP3 had no effect on costunolide accumulation when co-expressed with either the costunolide or parthenolide pathway genes (Fig. 3b and 4b). However, TpLTP3 had a big effect on free parthenolide production, which, again, can be ascribed to enhanced accumulation into the apoplast (Fig. 4b). Combined, these results show that these LTPs play a role in transport of costunolide (TpLTP1/2) and

parthenolide (TpLTP3) to the apoplast. We interpret these results as a transport activity over the cell wall, although we do not have detailed spatial resolution of the extracellular product accumulation.

The substrate exclusion assays reveal another activity of these LTPs. The results show that the exclusion activity by TpLTP1/2 is limited, while TpLTP3 can effectively exclude influx of both costunolide and parthenolide from the apoplast into the cell. We interpret this result as a transport of general lipid molecules (e.g. waxes, fats, etc.) by these LTPs to the outer cell wall of mesophyll cells of the *N.benthamiana* leaves, thus forming a lipid layer that blocks influx of either costunolide or parthenolide. If this is true, then all three TpLTPs are non-specific LTPs as they act on the sesquiterpene lactones as well as other lipids (Liu et al., 2015; Salminen et al., 2016). We think that the exclusion effect of LTPs is not related to their action on the terpene but on their action on lipids. The fact that TpLTP3 is more efficient in excluding both costunolide and parthenolide we interpret as TpLTP3 being more efficient in creating a lipid layer on the cell wall of mesophyll cells which prevents influx of the substrate. We did not identify the lipid cargo of these LTPs in the absence of a co-expressed sesquiterpene biosynthesis pathway. However, if the substrate exclusion by TpLTP1/2 and TpLTP3 is indeed caused by deposition of a lipophilic layer on the outer-face of the cell wall of *N. benthamiana* mesophyll cells, the prediction is that desiccation of leaves expressing these LTPs is slower than leaves expressing only empty vector, while desiccation of leaves expressing TpLTP3 is slower than that of leaves expressing TpLTP1/2.

### **Feverfew TpLTP activity co-dependent on membrane transporters?**

For the AaLTP3 from *Artemisia annua*, the activity on (DH)AA depends on co-expression with the membrane transporter AaPDR2 which transports (DH)AA over the PM<sup>10</sup>. In the experiments described in the present study, the activities of TpLTP1/2 and TpLTP3, are found in a background of an existing, low, intrinsic export activity by *N. benthamiana* cells (see Fig. 4a and 4b EV controls). It could be that the action of the five feverfew LTPs that show little or no activity on costunolide or parthenolide have affinity of feverfew pathway products not tested here. The sequestration activity of individual LTPs may also depend on their intrinsic stability. Moreover, previously we have shown that AaLTPs3 protein stability is increased in the presence of its putative substrate as was shown for AaLTP3 (Wang et al., 2016). Extracts from feverfew trichomes contain an array of different derivatives from costunolide and parthenolide (e.g. costunolide di-epoxide, 3 $\beta$ -hydroxycostunolide, 3 $\beta$ -hydroxyparthenolide, artecamin, artemorin, etc. (Fischedick et al., 2012)) and it will be of interest to determine the activity of the different TpLTPs, also in combination with membrane transporters expressed in feverfew trichomes, for these other feverfew pathway products.

### **GPI anchor required for TpLTP3 activity**

The mechanism by which LTPs (TpLTP1/2, AtLTP1/3, AaLTP3) enhance the transport of sesquiterpenoids into the apoplast is easiest explained when we assume a carrier function over the cell wall. It is therefore surprising that the most active TpLTP in sesquiterpene transport belongs to the subclass of GPI-anchored LTPs. The group of LTPs with GPI anchor (NsLTPGs) is a subfamily of the Non-specific LTPs (NsLTPs), and have been shown to function in sporopollenin, suberin, and wax deposition and transport of phosphatidylinositol mono-phosphates (PtdIns3P, PtdIns4P and PtdIns5P) (Edstam et al., 2013; Deng et al., 2016). TpLTP3 is therefore the first member of the NsLTPG subfamily that seems to function in sesquiterpenoid transport. In many instances NsLTPGs function in deposition of lipophilic material on the outside of the cell wall, but in none of these cases it is clear how the supposedly restricted movement of these LTPGs can cause transport over the cell wall. One option may be loss of the GPI anchor through phospholipase action. However, for LTPGs this seems to be more an exception than the rule (Ferguson et al., 2009).

Moreover, we show that TpLTP3 loses its bioactivity without GPI anchor domain, while addition of this domain to TpLTP1 destroys TPLTP1 function. Another option is that GPI-anchored LTPs function as shuttle between membrane transporters and other LTPs that can move freely over the cell wall. However, there is no experimental evidence for such cooperation between different types of LTPs. In theory, a GPI anchor on TpLTP3 only allows for lateral movement in the membrane. However, studies on an LTPG from *Arabidopsis thaliana* show that LTPGs may accumulate in the apoplastic space when expressed early during development (DeBono et al., 2009). The GPI anchor prevents free movement over the cell wall, except near structures that penetrate the cell wall like the plasmodesmata (PD). It is therefore intriguing that recently it was shown that GPI anchors are involved in targeting proteins to PD (Zavaliev et al., 2016). Furthermore, we show that the loss of the GPI anchor results in a loss in export activity (Figure 7b), but also the loss of exclusion activity (Supplementary Fig. 5). Interestingly TpLTP3 displays relatively early expression. Although parthenolide accumulation and TpPTS expression reach to the maximum level at a later stage, parthenolide biosynthesis and accumulation start early in flower development, matching the relative early expression of TpLTP3.

### **Wider implications for role of LTPs in cell-to-cell transport of lipophilic compounds?**

It has now been demonstrated that for hydrophobic sesquiterpenoids, specific LTPs may enhance their transfer over a hydrophilic plant cell wall. We speculate that facilitated transfer over the cell wall may be relevant for other lipophilic compounds as well. Table S1 lists the predicted LogD (Leo et al., 1971) of different plant hormones and other sesquiterpene

lactones. It has been shown that in *Nicotiana plumbaginifolia* ABC1 (NpABC1), expressed in glandular trichomes, plays an important role in sclareolide excretion to the leaf surface (Jasiński et al., 2001; Stukkens et al., 2005; Bultreys et al., 2009). Similar to what is discussed in the present paper and in accordance with to our previous findings (Wang et al., 2016), *N. plumbaginifolia* LTPs may play a similar role as feverfew and artemisia LTPs in combination with the reported NpABC1 membrane transporter. Table S1 shows that the LogD of some strigolactones is close to that of the sesquiterpene lactones in the present study. In petunia it has been shown that an ABC type G membrane transporters (PhPDR1) is involved in secretion of strigolactones. Moreover, the apical subcellular localisation of PhPDR in root cells suggests that strigolactones may be subject to polar apical transport in the root tip (Sasse et al., 2015). Such polar transport would require transfer of the strigolactones over two cell walls and this raises the question whether the transfer over the cell walls is facilitated by LTPs.

The specific activity of individual TpLTPs for different sesquiterpenoid pathway products may be useful for pathway engineering in feverfew or in heterologous host plants, to prevent endogenous detoxification by glycosylation or glutathione conjugation. It seems that in feverfew multiple LTPs are involved in directing different pathway intermediates and/or products to the extracellular space. Thus, it would be of interest to test whether elimination of TpLTP1/2 activity in feverfew would result in a higher yield of parthenolide and parthenolide derivatives.

Future research will have to show if optimal combinations between specific membrane transporters and LTPs can enhance transfer of specific lipophilic compounds over plant cell walls, whether this can be used to engineer extracellular accumulation of desired secondary metabolites and whether this has a role in homeostasis of strigolactones and other plant hormones.

## References

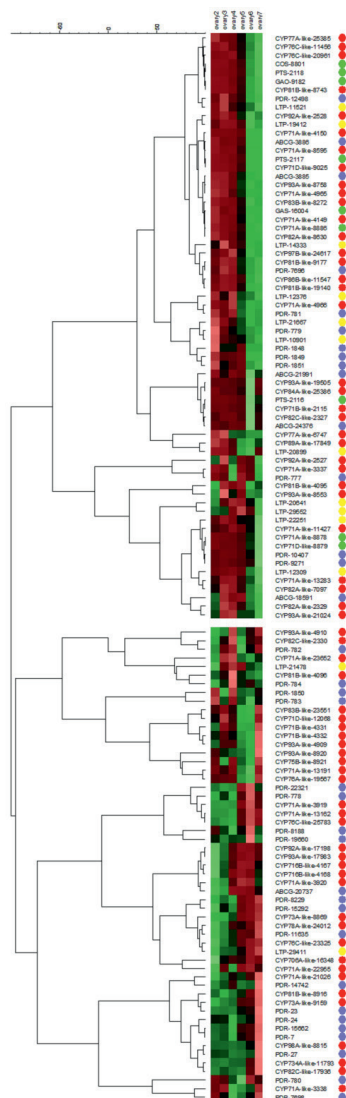
- Bultreys A, Trombik T, Drozak A, Boutry M** (2009) *Nicotiana plumbaginifolia* plants silenced for the ATP-binding cassette transporter gene NpPDR1 show increased susceptibility to a group of fungal and oomycete pathogens. *Mol Plant Pathol* **10**: 651-663
- Choi S-H, Im E, Kang HK, Lee J-H, Kwak H-S, Bae Y-T, Park H-J, Kim ND** (2005) Inhibitory effects of costunolide on the telomerase activity in human breast carcinoma cells. *Cancer letters* **227**: 153-162
- Coutos-Thevenot P, Jouenne T, Maes O, Guerbette F, Grosbois M, Le Caer JP, Boulay M, Deloire A, Kader JC, Guern J** (1993) Four 9-kDa proteins excreted by somatic embryos of grapevine are isoforms of lipid-transfer proteins. *Eur J Biochem* **217**: 885-889
- de Oliveira Carvalho A, Gomes VM** (2007) Role of plant lipid transfer proteins in plant cell physiology—a concise review. *Peptides* **28**: 1144-1153
- DeBono A, Yeats TH, Rose JKC, Bird D, Jetter R, Kunst L, Samuels L** (2009) Arabidopsis LTPG Is a Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein Required for Export of Lipids to the Plant Surface. *The Plant Cell* **21**: 1230-1238

- Deng T, Yao H, Wang J, Wang J, Xue H, Zuo K (2016) GhLTPG1, a cotton GPI-anchored lipid transfer protein, regulates the transport of phosphatidylinositol monophosphates and cotton fiber elongation. *Scientific Reports* **6**: 26829
- Edstam MM, Blomqvist K, Eklof A, Wennergren U, Edqvist J (2013) Coexpression patterns indicate that GPI-anchored non-specific lipid transfer proteins are involved in accumulation of cuticular wax, suberin and sporopollenin. *Plant Mol Biol* **83**: 625-649
- Ferguson MA, Kinoshita T, Hart GW (2009) Glycosylphosphatidylinositol anchors.
- Fischedick JT, Standiford M, Johnson DA, De Vos RC, Todorović S, Banjanac T, Verpoorte R, Johnson JA (2012) Activation of antioxidant response element in mouse primary cortical cultures with sesquiterpene lactones isolated from *Tanacetum parthenium*. *Planta medica* **78**: 1725-1730
- Hsu J-L, Pan S-L, Ho Y-F, Hwang T-L, Kung F-L, Guh J-H (2011) Costunolide induces apoptosis through nuclear calcium  $2^{+}$  overload and DNA damage response in human prostate cancer. *The Journal of urology* **185**: 1967-1974
- 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
- Kader J-C (1996) Lipid-transfer proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 627-654
- Kim H, Lee SB, Kim HJ, Min MK, Hwang I, Suh MC (2012) Characterization of glycosylphosphatidylinositol-anchored lipid transfer protein 2 (LTPG2) and overlapping function between LTPG/LTPG1 and LTPG2 in cuticular wax export or accumulation in *Arabidopsis thaliana*. *Plant Cell Physiol* **53**: 1391-1403
- Leo A, Hansch C, Elkins D (1971) Partition coefficients and their uses. *Chemical reviews* **71**: 525-616
- Liu F, Zhang X, Lu C, Zeng X, Li Y, Fu D, Wu G (2015) Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *Journal of Experimental Botany* **66**: 5663-5681
- Liu Q, Majdi M, Cankar K, Goedbloed M, Charnikhova T, Verstappen FW, De Vos RC, 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 Q, Manzano D, Tanić N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van de Krol S (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metabolic engineering* **23**: 145-153
- Majdi M, Liu Q, Karimzadeh G, Malboobi MA, Beekwilder J, Cankar K, Vos Rd, Todorović S, Simonović A, Bouwmeester H (2011) Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.). *Phytochemistry* **72**: 1739-1750
- Moreau P, Cassagne C (1994) Phospholipid trafficking and membrane biogenesis. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1197**: 257-290
- Salminen TA, Blomqvist K, Edqvist J (2016) Lipid transfer proteins: classification, nomenclature, structure, and function. *Planta* **244**: 971-997
- Sasse J, Simon S, Gubeli C, Liu GW, Cheng X, Friml J, Bouwmeester H, Martinoia E, Borghi L (2015) Asymmetric localizations of the ABC transporter PaPDR1 trace paths of directional strigolactone transport. *Curr Biol* **25**: 647-655
- Shin DH, Lee JY, Hwang KY, Kim KK, Suh SW (1995) High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure* **3**: 189-199
- Sterk P, Booij H, Schellekens GA, Van Kammen A, De Vries SC (1991) Cell-specific expression of the carrot EP2 lipid transfer protein gene. *The Plant Cell* **3**: 907-921
- Stukkens Y, Bultreys A, Grec S, Trombik T, Vanham D, Boutry M (2005) NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. *Plant Physiology* **139**: 341-352
- Ting H-M, Delatte TL, Kolkman P, Misas-Villamil JC, Van Der Hoorn RA, Bouwmeester HJ, van der Krol AR (2015) SNARE-RNAi results in higher terpene emission from ectopically expressed caryophyllene synthase in *Nicotiana benthamiana*. *Molecular plant* **8**: 454-466
- Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) Retracted: An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal* **33**: 949-956
- Wang B, Kashkooli AB, Sallets A, Ting H-M, de Ruijter NCA, Olofsson L, Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR (2016) Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metabolic Engineering* **38**: 159-169

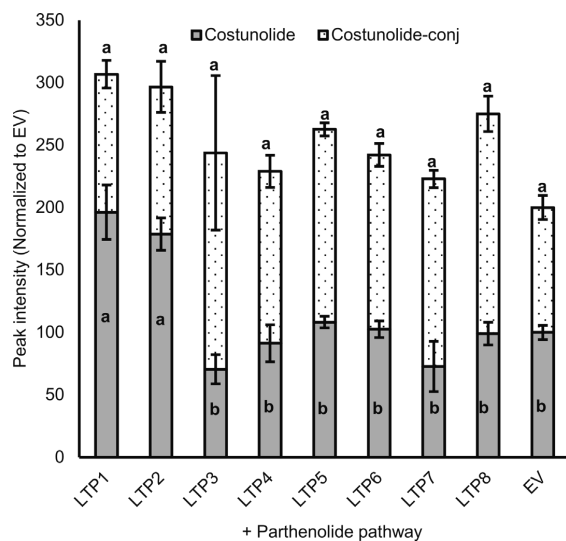


- Widhalm JR, Jaini R, Morgan JA, Dudareva N** (2015) Rethinking how volatiles are released from plant cells. *Trends in Plant Science* **20**: 545-550
- Yang Y-I, Kim J-H, Lee K-T, Choi J-H** (2011) Costunolide induces apoptosis in platinum-resistant human ovarian cancer cells by generating reactive oxygen species. *Gynecologic oncology* **123**: 588-596
- Yu H-J, Jung J-Y, Jeong JH, Cho S-D, Lee J-S** (2015) Induction of apoptosis by parthenolide in human oral cancer cell lines and tumor xenografts. *Oral Oncology* **51**: 602-609
- Zunino SJ, Ducore JM, Storms DH** (2007) Parthenolide induces significant apoptosis and production of reactive oxygen species in high-risk pre-B leukemia cells. *Cancer Letters* **254**: 119-127

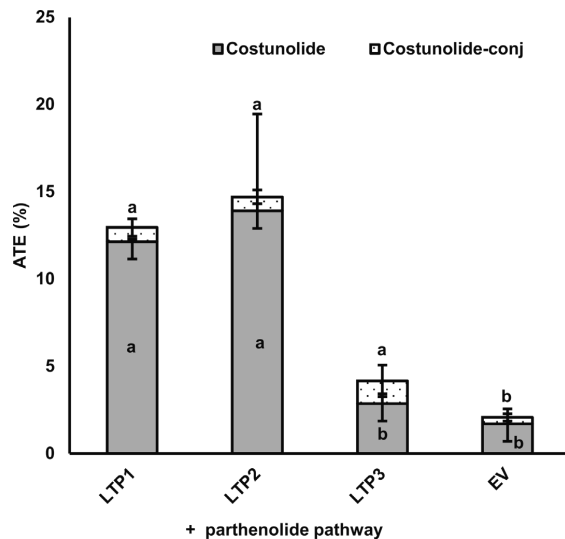
## Supplementary information



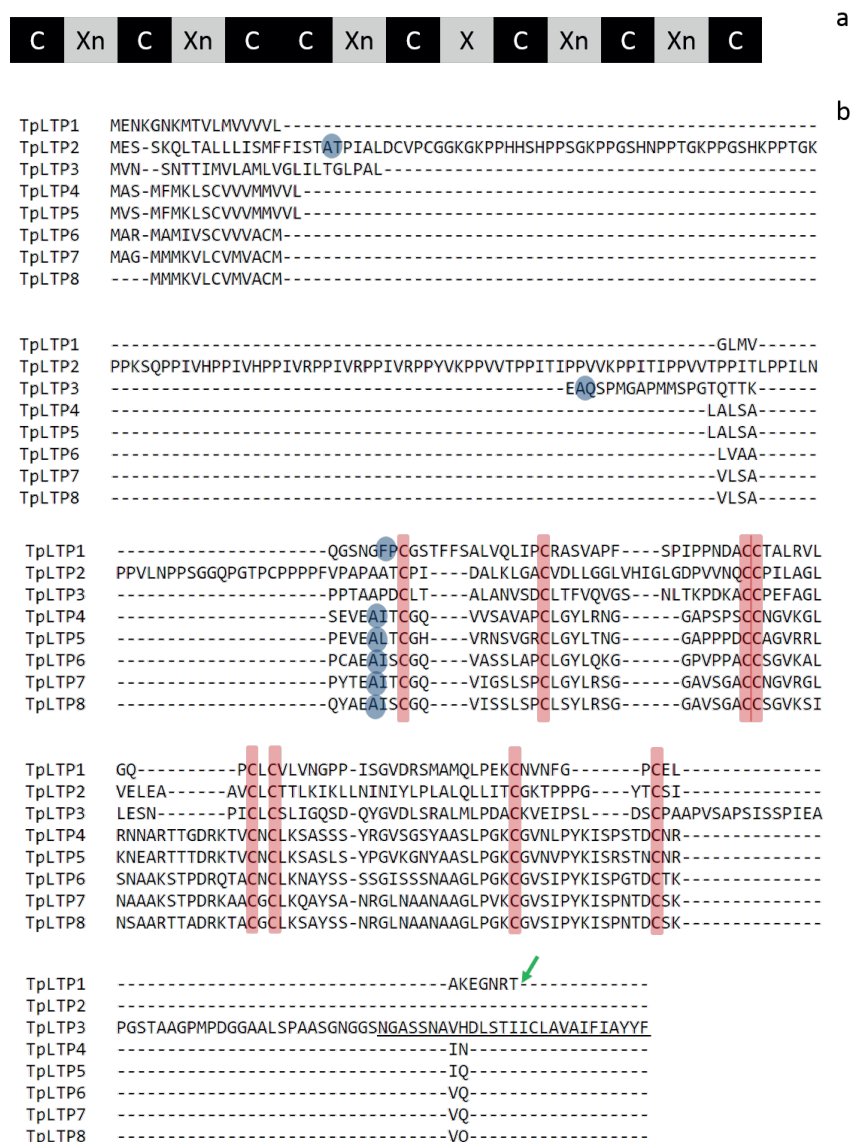
**Supplementary Figure 1- Cluster analysis of differentially expressed genes in 6 developmental stages of feverfew flowers.** Feverfew sesquiterpene lactone biosynthesis pathway genes (green circles), unknown function cytochrome P450s (red), ABC transporters (purple) and Lipid transfer proteins (yellow). 2 LTPs are not co-clustered in the same cluster of biosynthesis pathway genes. In order to simplify this figure, unknown function (based on GO terms) and very short contigs were not included.



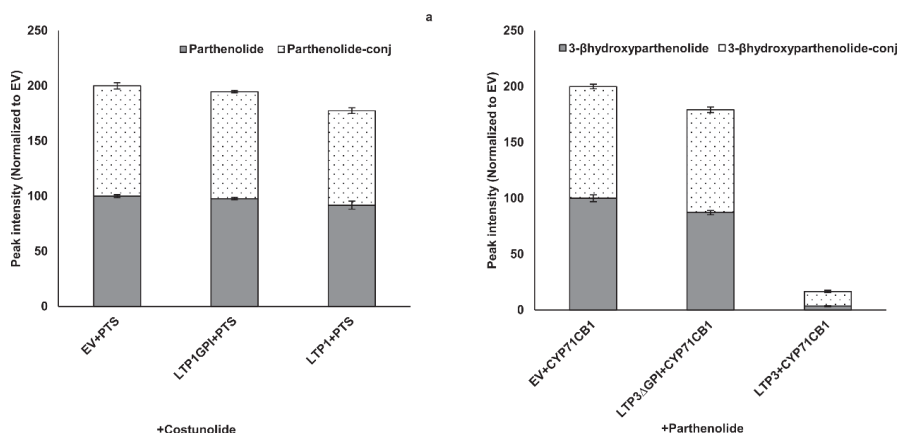
**Supplementary Figure 2. Effect of individual *TpLTPs* on total costunolide and costunolide-conj accumulation upon transient co-expression with parthenolide pathway genes in *N. benthamiana*.** All values are normalized to EV (empty vector) and EV values for each compound is set at 100; costunolide (grey bars) and costunolide-conj (black-dotted bars). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).



**Supplementary Figure 3. Costunolide apoplast transport efficiency (ATE) upon transient co-expression with the parthenolide pathway genes in *N. benthamiana*.** Only free costunolide accumulates in the apoplast upon co-expression of *TpLTP1* and *TpLTP2*. Means followed by a different letter are significantly different ( $P \leq 0.05$ ).



**Supplementary Fig. 4. Alignment of 8 *T. parthenium* lipid transfer proteins. (a) Characteristic structure of LTPs.** Conserved 8 cysteine residues (C). X represents amino acid residue and Xn represents a number of amino acid residues. **(b)** Multiple sequence alignment was done by Promals3d (<http://prodata.swmed.edu/promals3d/promals3d.php>). Blue indicates the secretion signal peptide, predicted by SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>). 8 Cysteine motifs are indicated in red. The GPI anchor of TpLTP3 is underlined, as predicted by PredGPI (<http://gpcr.biocomp.unibo.it/predgpi/pred.htm>). Green arrow represents the insertion site of TpLTP3 GPI anchor. Truncation in LTP3<sup>AGPI</sup> was done by deletion of the underlined GPI motif.



**Supplementary Fig. 5. Effect of addition of a GPI anchor to LTP1 and truncation of the GPI anchor from LTP3 on intercellular bioconversion of costunolide to parthenolide by parthenolide synthase and parthenolide to 3β-hydroxy parthenolide by 3β-parthenolide hydroxylase in *N. benthamiana* leaves.** After 3 days, substrates (either costunolide or parthenolide (1mg ml<sup>-1</sup>) were vacuum infiltrated into leaf disks of 1cm diameter. EV was used as a control. Metabolites were extracted with methanol and analysed with LC-MS. **(a)** UPLC-MRM-MS analysis of extracted metabolites from leaves expressing parthenolide synthase, vacuum infiltrated with costunolide and incubated for 120 minutes **(b)** Bioconversion of parthenolide to 3β-hydroxy parthenolide. Apoplast vacuum infiltration of parthenolide and extraction of leaf metabolites at 120 min (analysed by LC-Orbitrap-FTMS). Means followed by a different letter are significantly different ( $P \leq 0.05$ ).

**Supplementary Table 1. Predicted log D values (corrected for pH) of studied compounds in this study.** artemisinic acid, dihydroartemisinic acid and some known phytohormones. Lower and negative values represent higher affinity for partitioning in the aqueous phase.

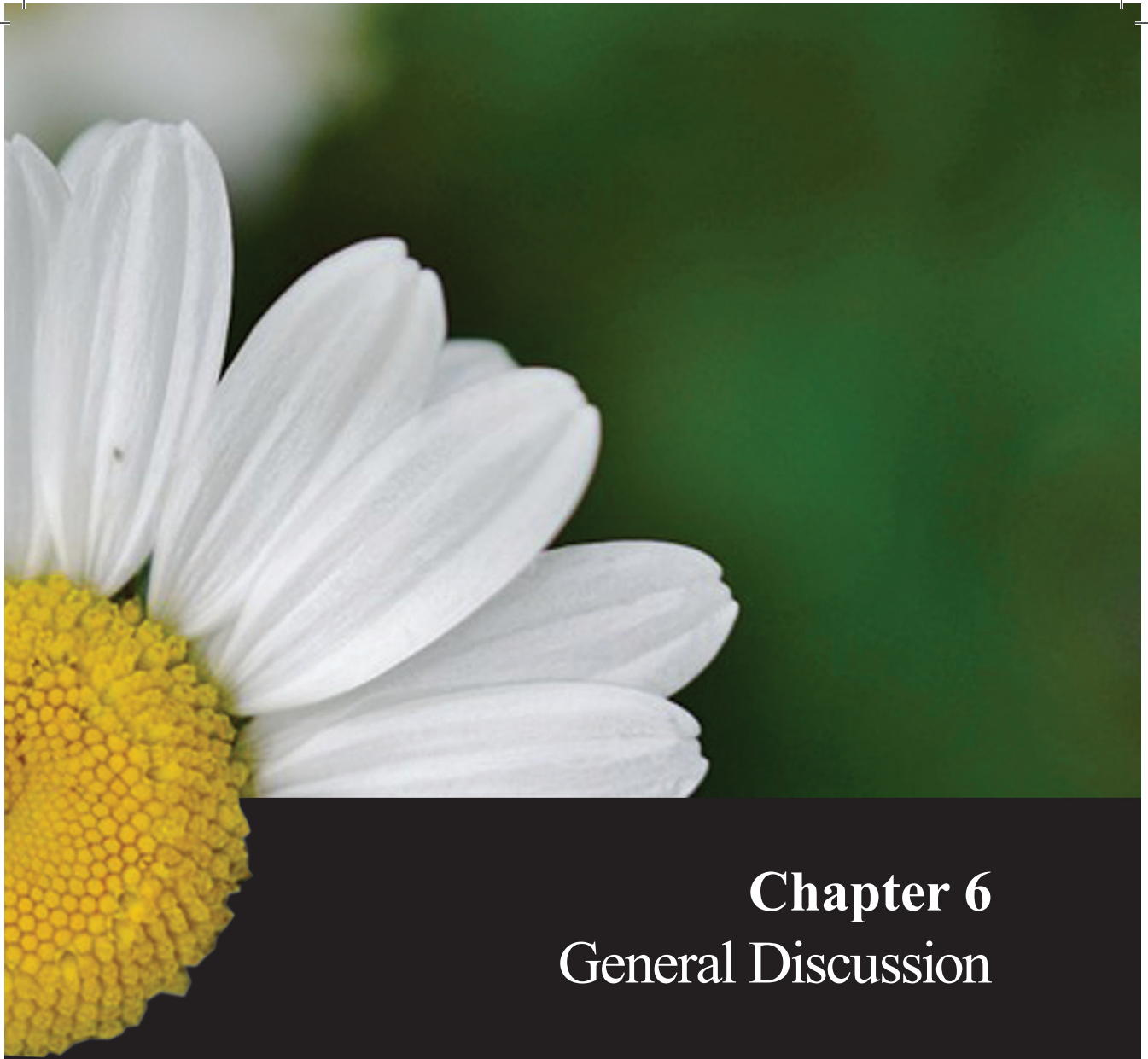
| Compound                 | logD pH5.5 | logD pH 7.4 |
|--------------------------|------------|-------------|
| Carlactone               | 4.66       | 4.66        |
| Sclareolide              | 4.12       | 4.12        |
| Caryophyllene            | 6.47       | 6.47        |
| Artemisinic acid         | 3.58       | 1.79        |
| <u>Costunolide</u>       | 3.36       | 3.36        |
| Dihydroartemisinic acid  | 3.32       | 1.52        |
| Brassinolide             | 3.18       | 3.18        |
| 5-Deoxystrigol           | 2.64       | 2.64        |
| <u>Parthenolide</u>      | 2.61       | 2.61        |
| 1-Naphthaleneacetic acid | 1.73       | -0.06       |
| Orobanchol               | 1.41       | 1.41        |
| Abcisic Acid             | 0.93       | -0.87       |
| Jasmonic Acid            | 0.72       | -1.08       |
| 3-Indole acetic acid     | 0.68       | -1.12       |
| Kinetin                  | -0.33      | 0.41        |
| Zeatin                   | -0.99      | -0.21       |
| Sucrose                  | -3.57      | -3.57       |

**Supplementary Table 2. Primers used in this study.**

| <b>Name primer</b>     | <b>5' to 3' sequence</b>                    |
|------------------------|---------------------------------------------|
| <i>TpLTP1</i> -Forward | AAACTAGATCTGAATAATGGAAAACAA                 |
| <i>TpLTP1</i> -Reverse | GACaCaACAATGCTTAACAATAATTC                  |
| <i>TpLTP2</i> -Forward | GACACCAATACACATTGAGTGACA                    |
| <i>TpLTP2</i> -Reverse | GGAAAACCTATATCCTTAAACACTCCA                 |
| <i>TpLTP3</i> -Forward | CACCAAATATGGTTAATTCAAA                      |
| <i>TpLTP3</i> -Reverse | TTAGAAGTAGTATGCTATGAAG                      |
| <i>TpLTP4</i> -Forward | CACCAATGGCATCCATGTTT                        |
| <i>TpLTP4</i> -Reverse | TCAATTAATCCTGTTGCAGTCAGT                    |
| <i>TpLTP5</i> -Forward | TGTCCATGTTTCATGAAGTTATC                     |
| <i>TpLTP5</i> -Reverse | TTAGTGCTATAGTTTGCTTCCTTA                    |
| <i>TpLTP6</i> -Forward | CAGCCCCCATGCATAAGTATCACA                    |
| <i>TpLTP6</i> -Reverse | TCATGTGTAGTCGTGGCTTGCATGTTGG                |
| <i>TpLTP7</i> -Forward | CACACGCCAAGTGAATATCCATGC                    |
| <i>TpLTP7</i> -Reverse | ATGTGTTATGGTGGCGTGCATGGTGTTG                |
| <i>TpLTP8</i> -Forward | ATTCGAACATTGGCAGGAATGA                      |
| <i>TpLTP8</i> -Reverse | ACATTGAAGAGGCCAACCTCAC                      |
| 22251 3' RACE          | ATCAGCCGATCTACCAACTGCAACAGGA                |
| 22251 5' RACE          | TTACGGTCCGTAGTGGTTCGTGCTTCAT                |
| <i>TpLTP2</i> 3' RACE  | GATTACGCCAAGCTTTGGCAAGACTCCACCTCCTGGTTACACT |
| <i>TpLTP2</i> 5' RACE  | GATTACGCCAAGCTTGCATGTGGCTGCTGGAGCTGGCACAA   |
| <i>TpLTP8</i> 3' RACE  | GGCGTGTTGCAGCGGCGTTAAGAGTATT                |







## **Chapter 6**

### General Discussion

Plant secondary metabolites constitute one of the main components of plant defence against a range of organisms. For humans, however, plant secondary metabolites are a source of flavouring agents, fragrances, herbicides, but more importantly, they are a rich source of pharmaceutically active molecules. Plants produce different classes of secondary metabolites such as terpenoids, alkaloids and phenolic compounds. Terpenoids are the largest and most widespread class of secondary metabolite in nature with over 40,000 known examples (Connolly and Hill, 1991; Aharoni et al., 2005). An important class of the terpenoids are the sesquiterpene lactones, which are promising as drugs against cancer (e.g. parthenolide and costunolide). Artemisinin is another example of a plant-derived sesquiterpene lactone, which is isolated from the Chinese herb, *Artemisia annua*, and is used for treatment of malaria.

Despite the low molecular weight of the sesquiterpenoid backbone, the complicated, stereo-specific, functionalization of the basic skeleton of these natural products renders them difficult targets for chemical synthesis. Indeed, biosynthesis of these medicinal sesquiterpenes is generally far more efficient than chemical synthesis. Therefore, discovery of the enzymes involved in the biosynthesis of sesquiterpene lactones was an important part of this thesis research. Although many studies have been conducted on metabolic engineering of terpenoids, there are still numerous questions regarding biosynthesis and especially inter/(intra)-cellular transport. Understanding the molecular mechanism by which these secondary metabolites are synthesised, transported and/or accumulated in plants will pave the way towards more successful metabolic engineering approaches.

The aims of this study were to **(1)** characterise additional enzymes which can modify parthenolide, costunolide and/or their derivatives to other valuable sesquiterpene lactones; **(2)** testing novel combinations of enzymes, with the aim of producing novel sesquiterpenoids (combinatorial biosynthesis) and **(3)** revealing the molecular components involved in extracellular transport and accumulation of parthenolide, costunolide and artemisinin.

During this research we have identified an enzyme (kauniolide synthase) catalysing the first committed step in the biosynthesis of guaianolide sesquiterpene lactones. This discovery has opened up a whole new class of sesquiterpenes with important medicinal functions for metabolic engineering. Furthermore, by combining sesquiterpene lactone biosynthesis pathway genes from *Tanacetum parthenium* (feverfew) and a double bond reductase enzyme from *A. annua* we produced new sesquiterpene lactones with reduced exocyclic double bond. These novel compounds were produced both *in vitro* (yeast enzymatic assays) and *in planta*. Unfortunately, time did not permit identification of the bioactivity of these novel compounds. To achieve this, first larger scale production of these compounds would be needed. This could be achieved by either feeding dihydrocostunolide to isolated yeast microsomes or *Nicotiana benthamiana* leaves expressing *Tp3 $\beta$ -hydroxylase*. Since the exocyclic double bond of the

substrate and product sesquiterpene lactones is reduced, no conjugation (to cysteine/glutathione) will take place making isolation of the products easier. In a next step the bioactivity could be assessed by examining the effect of the novel compounds on for example both sensitive and multi-drug resistant carcinoma cells and identifying parameters such as minimum inhibitory concentration (MIC). Finally, in this thesis we identified for the first time a role for lipid transfer proteins (LTPs) in the transport of sesquiterpene lactones (costunolide, parthenolide and artemisinin precursors), and we established a novel bio-assay for pleiotropic drug resistance proteins (PDRs). Moreover, we demonstrated for the first time the successful production of artemisinin in the heterologous host plant *N. benthamiana*.

My research also left us with some questions, which will need to be addressed in the future to make metabolic engineering of medicinal sesquiterpenes to a full success story. Some of these remaining questions I will address in this final discussion chapter such as:

- Choice of host organism for metabolic engineering
- Additional options for combinatorial biochemistry
- Transport and sequestration of sesquiterpenes not completely solved

## **Metabolic engineering of sesquiterpene lactones**

### **Manipulation host plant**

Biosynthetic pathway genes/enzymes identification and characterization through transcriptomics, genetical and biochemical approaches provides novel tools and possibilities for engineering secondary metabolite (for example sesquiterpene lactones) production. The plant species, from which these biosynthetic pathway genes are isolated, could serve as the host for metabolic engineering. Stable transformation methods for some of these plant species such as *A. annua* have recently been developed (Fu et al., 2017). Genetic engineering of the gene donor plant species, however, is often difficult due to the difficulties/absence of transformation methods. Furthermore, many plant derived terpenoids are synthesized in specialized tissues such as trichomes or other secretory tissues. Hence metabolic engineering would entail very tightly regulated tissue specific expression of the pathway genes (Jindal et al., 2015; Liu et al., 2016). In case of feverfew no transformation protocol is available so far. Hence, alternative metabolic engineering platforms and approaches are required for identification and characterization of biosynthetic pathway genes and pathway engineering.

## Ectopic expression in heterologous host species

Metabolic engineering of sesquiterpene lactones can be done in other organisms than the donor plant species, such as other host plant species and yeast. Each of these platforms has advantages and disadvantages, and needs to be optimised for commercial production. In this study we have used the transient expression plant platform, *N. benthamiana*, for production of costunolide, parthenolide and artemisinin. In order to obtain higher yields in this system we have co-expressed *HMGR* (*3-Hydroxy-3-Methyl-Glutaryl-coenzyme A Reductase*). *HMGR* is the rate limiting enzyme in the mevalonate pathway, which provides the precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), required for the production of sesquiterpenoids and triterpenoids. It was previously reported that co-expression of *HMGR* results in increased yield, from transiently reconstructed parthenolide and artemisinin pathways (van Herpen et al., 2010; Liu et al., 2014). Ectopic expression of a terpenoid pathways *in planta* often does not result in desired production levels compared with the endogenous production in the donor plant species that is naturally producing these compounds (Wallaart et al., 2001). This production can be enhanced by downregulation of competing endogenous pathways. For example, in some cases where endogenous pathways in *N. benthamiana* are competing with transiently expressed genes for the FPP pool for sesquiterpenoid production, downregulation of squalene synthase and 5-*epi*-aristolochene synthase enhanced production of (+)-valencene, by transiently expressed (+)-valencene synthase (Cankar et al., 2015). It is also noteworthy to mention that many researchers showed that directing and targeting genes to special compartments can enhance ectopic production of metabolites in heterologous plant hosts (Wu et al., 2006; Liu et al., 2011). Some of these also offer perspective to enhance sesquiterpenoid production in *N. benthamiana*. For example, *HMGR* that I co-expressed together with other biosynthetic pathway genes, was used to increase the FPP substrate pool. It was also previously reported that mitochondrial targeting of the germacrene A synthase results in increased production of germacrene A (Liu et al., 2014). I used the same germacrene A synthase (containing a mitochondrial targeting signal) to increase germacrene A production and subsequently final product yield in our *in planta* experiments. However, in my thesis I did not use downregulation of competing pathways with targeted RNAi constructs, with as most important reason that we then would have to omit P19, the silencing inhibitor that we use to avoid silencing of the feverfew pathway genes.

One of the major differences of *N. benthamiana* as a heterologous expression system compared with the original donor plant (for example feverfew) is that the heterologously produced costunolide and/or parthenolide do not accumulate in the *N. benthamiana* trichomes, but in the apoplast of the leaf mesophyll cell layers. This was confirmed by lack of costunolide or parthenolide in chloroform dip extracts of *N. benthamiana* leaves expressing

the costunolide or parthenolide biosynthetic pathway genes with or without LTPs (data not shown), but presence of these compounds in leaf apoplast extracts.

The other disadvantage of plant expression systems such as *N. benthamiana* is the endogenous detoxification mechanisms such as glycosylation and cysteine/glutathione conjugation. This conjugation alters the chemical properties of the produced compounds and may channel the metabolite towards the vacuoles. This could on the one hand be an advantage (reducing toxicity), but it has been shown that the bioactivity of some sesquiterpene lactones such as costunolide and parthenolide is reduced upon conjugation to cysteine and glutathione (Liu et al., 2014). Extracellular transport and storage of these metabolites could be a solution to this conjugation problem. We have first demonstrated this in **chapter 4** where more free product (non-conjugated artemisinin) was detected in the leaf apoplast. A much stronger effect of extracellular storage and prevention from detoxifying mechanisms of the cell was observed in **chapter 5** where the extracellular product accumulation profile (ratio non-conjugated to conjugated) was dramatically increased.

Also in artemisinin metabolic engineering, many glycoside and glutathione conjugates of artemisinin precursors were produced when the artemisinin biosynthesis pathway genes were co-expressed in *N. benthamiana* (Ting et al., 2013). In theory this could be prevented by downregulation of endogenous glycosyltransferases, hence obtaining more of the final product, artemisinin of which so far we produce about 3 µg/g dry weight in *N. benthamiana* (**chapter 4**, (Wang et al., 2016)). The moss, *Physcomitrella patens*, a non-vascular plant expression system, was recently also used for the biotechnological production of artemisinin (Khairul Ikram et al., 2017). In this system, the amount of artemisinin was about 70-fold higher at 210 µg/g dry weight without any optimization such as increasing substrate availability or reducing competing activities (Khairul Ikram et al., 2017). Interestingly no conjugated products were detected in extracts of the engineered moss, which might be due to the low number of glycosyltransferases 1 (GT1) genes, which in higher plants are involved in detoxification/deactivation processes of secondary metabolites, in *P. patens*. This suggests that moss is a suitable platform for metabolic engineering of sesquiterpene lactones.

### **Ectopic expression of pathway genes in yeast**

Yeast is another possible expression platform for identification and characterization of genes involved in biosynthesis of sesquiterpenes and/or for metabolic engineering/production. Takahashi et al., (2007), for example, used yeast to express *Hyoscyamus muticus* premnaspirodiene synthase (*HPS*). Co-expression of *HMGR* with *HPS* resulted in accumulation of approximately 170 mg.L<sup>-1</sup> of the sesquiterpene premnaspirodiene. The endogenous FPP pool of yeast together with a relatively straight forward transformation

procedures have made it a suitable host for *in vivo* reconstruction of sesquiterpenes biosynthetic pathways, of for example costunolide (Liu et al., 2011), artemisinic acid (Ro et al., 2006) and, as shown in this thesis, kauniolide (**Chapter 2**). However, yeast requires some optimization in order to be able to produce significant levels of sesquiterpenes. Many of the modifications (like oxidation, epoxidation or reduction) that are required to produce the final product from the sesquiterpene backbone produced by a sesquiterpene synthase from FPP, are catalysed by cytochrome P450 enzymes that are anchored to the endoplasmic reticulum (ER) membrane. The WAT11 yeast strain, which I used in my work has been previously engineered to express NADPH-P450 reductase (CPR) from *Arabidopsis thaliana*, hence making it a suitable platform for the expression of cytochrome P450 enzymes. Further optimisation of the WAT11 strain can be achieved by downregulation of endogenous competing pathways, for example by reducing sterol biosynthesis by knocking out ERG9 (encoding farnesyl-diphosphate farnesyltransferase, aka squalene synthase). This strategy was used for the optimisation of artemisinic acid production in yeast (Paddon et al., 2013). Alternatively, providing additional copies of the gene of interest in yeast may also enhance the final yield (Ignea et al., 2014). Compounds such as artemisinin, costunolide and parthenolide are quite hydrophobic and this might limit their export from the cell to the medium, thus reducing yield. Enhancement of lipid bodies/droplets by improving triacylglycerol storage (Winichayakul et al., 2013) might provide a sink for terpenes storage in the yeast culture medium. The other approach to increase accumulation of lipophilic compounds is the application of a two-phase reactor, in which a mild organic solvent can trap hydrophobic compounds, but does not inhibit growth of the yeast cells. The produced lipophilic compounds will accumulate in the organic solvent, which will stimulate production and avoid toxicity. Two-phase extractive fermentation was used to reduce monoterpenes toxicity for bioproduction of jet fuel in yeast (Brennan et al., 2012). When comparing yeast and plant (*N. benthamiana*) organic solvent extracts we note a big difference in terms of extract purity. Plant extracts contain more undesired metabolites than the yeast extracts. This is a disadvantage for obtaining pure compounds from the plant expression systems.

### **Opportunities for combinatorial metabolic engineering**

Combinatorial metabolic engineering has also been used to a certain extent by others (O'Reilly et al., 2013; Eudes et al., 2016; Mafu et al., 2016). The principle of combinatorial biochemistry was first defined at the metabolic level. Different precursors were provided to a single enzyme, resulting in different products (Julsing et al., 2006). However, nowadays approaches are more towards reconstruction of biosynthetic pathways in other organisms such as plants and microbial hosts. This approach allows researchers to establish a basis for complex pathways reconstruction in which many biosynthetic steps are involved. With this



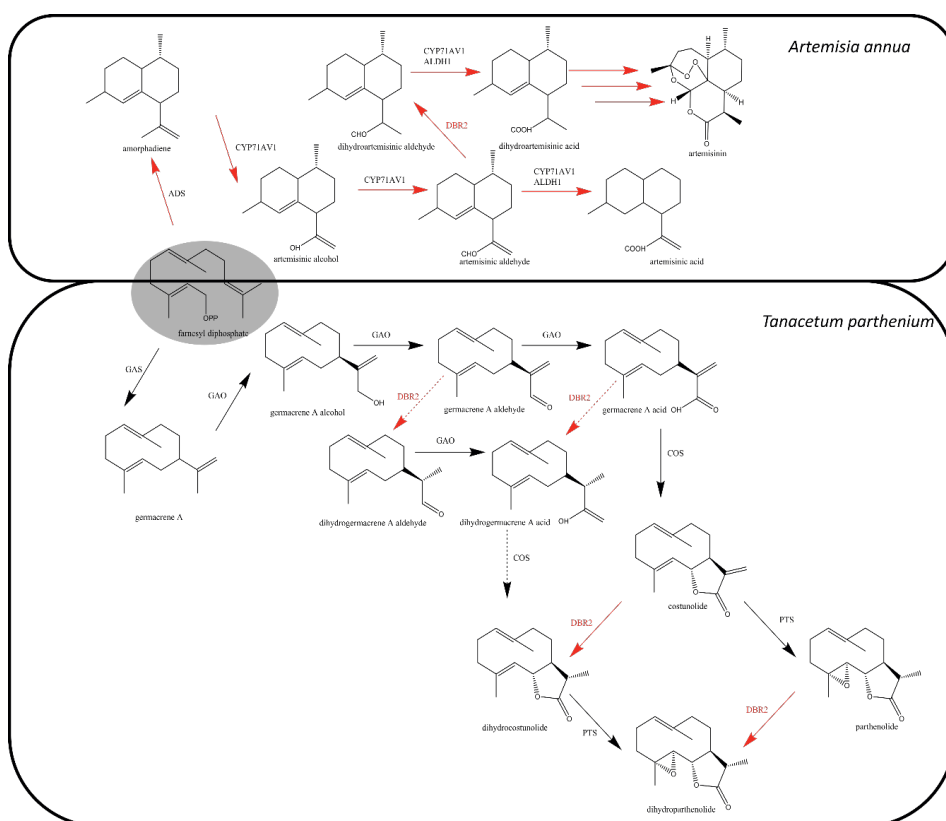
approach scientists first define the required enzymes for pathway reconstruction. Such an approach also includes characterization of the most active homologous or heterologous enzyme. An example of such an approach is taxol biosynthesis. Paclitaxel (Taxol®) is a diterpenoid isolated from the bark and needles of the pacific yew tree which shows strong activity against breast and ovary cancer cells (Caldas and McGuire, 1993; Rao, 1993). Biosynthesis of paclitaxel requires 19 enzymatic steps, starting from cyclisation of geranylgeranyl diphosphate to taxadiene. Taxadiene biosynthesis ( $1.3 \text{ mgL}^{-1}$ ) was achieved through a combinatorial metabolic engineering approach in which genes from *Taxus brevifolia* (encoding a taxadiene synthase, with increased solubility), *Erwinia herbicola* (encoding a geranylgeranyl diphosphate (GGPP) synthase (*E. coli* produces low levels of GGPP)), *Schizosaccharomyces pombe* (encoding an isopentenyl diphosphate (IPP) isomerase (to increase intracellular concentration of IPP) and *E. coli* (encoding a 1-deoxy-D-xylulose 5-phosphate (DXP) synthase, to increase isoprenoids levels through the non-mevalonate pathway) were co-expressed in *E. coli* (Huang et al., 2001). In this thesis, for the reconstruction of the costunolide/parthenolide biosynthetic pathway we used the germacrene A synthase from feverfew, which was approximately three times more active than the germacrene A synthases (short and long) from chicory (Liu et al., 2011).

Another aspect of combinatorial metabolic engineering consists of the mixing of enzymes from different pathways, assuming that these enzymes will also be active on other substrates (substrate promiscuity). A great example of this approach is the promiscuous CYP71AV8, isolated from chicory which can catalyse C12 oxidation of germacrene A and amorpha-4,11-diene and C2 oxidation of (+)-valencene (Nguyen et al., 2010; Cankar et al., 2011). In **chapter 3** by using yeast and plant-based platforms we explored the promiscuity of double bond reductase 2 (DBR2) from *A. annua* with sesquiterpene lactones from feverfew (Figure 1). Although AaDBR2 reduces the double bond of artemisinic aldehyde, it did not show that activity for germacrene A alcohol, aldehyde or acid.

NMR studies and analysis of sesquiterpene lactones from feverfew confirms the presence of the reduced (dihydro) forms of sesquiterpene lactones (e.g. dihydroparthenolide) at about 10% of the non-reduced form (Fischedick et al., 2012). This lower amount of dihydroparthenolide might be due to a lag phase/delayed expression of the DBR2-like enzyme in feverfew (compared with parthenolide synthase expression). Parthenolide is likely transported and accumulated in the extracellular space of the trichomes (through the activity of LTP3 (**chapter 5**), where it is not accessible anymore to the putatively later expressed feverfew DBR2-like enzyme. In **chapter 3**, two putative biosynthetic pathways of dihydroparthenolide were assessed. Parthenolide was efficiently reduced by the AaDBR2 in an *in planta* expression platform. In contrast, the route in which first costunolide was reduced (producing



dihydrocostunolide) and then being used by the parthenolide synthase was not so efficient (**Chapter 3**). The co-localization of costunolide synthase and parthenolide synthase on the ER membrane could support a preferred further oxidation of costunolide to parthenolide and then reduction by the DBR enzyme rather than a reduction of costunolide by the DBR and then oxidation of dihydrocostunolide by the parthenolide synthase. However, this may be clarified by isolation and characterization of the feverfew DBR2 homolog(s).



**Figure 1.** Biosynthesis pathway of artemisinin in *Artemisia annua* (top) and proposed combinatorial biosynthesis pathway of dihydrocostunolide and dihydroparthenolide in *Tanacetum parthenium* (bottom). Red arrows: enzymes from *Artemisia annua*. ADS: amorpha-4,11-diene synthase, CYP71AV1: amorpha-4,11-diene oxidase, ALDH1: aldehyde dehydrogenase 1, DBR2: double bond reductase 2, GAS: germacrene A synthase, GAO: germacrene A oxidase, COS: costunolide synthase, PTS: parthenolide synthase.

## Remaining questions related to transport of (sesqui)terpenes

### Transport over plasma membrane?

Both parthenolide and artemisinin are biosynthesized in trichomes (Duke et al., 1994; Majdi et al., 2011). Trichomes are specialized plant structures differentiating from epidermal cells and many different types of trichomes can exist within one plant (Schellmann and Hulskamp, 2004). Sesquiterpene lactones in feverfew (e.g. costunolide and parthenolide) and in *A. annua* ((dihydro)artemisinic acid) are biosynthesized in so-called glandular trichomes, after which these products are transported across the plasma membrane and stored in the extracellular space of the trichome cells. Presumably this transport prevents these reactive compounds to react with cellular proteins. The compounds are mainly hydrophobic and therefore face a hydrophilic barrier of the cell cytosol, the hydrophilic ‘head’ region of phospholipids of the plasma-membrane and the hydrophilic environment of the plant cell wall. In my thesis I have tested pleiotropic drug resistance (PDR) proteins for their involvement in sesquiterpene transport. PDRs are classified as ATP-binding cassette (ABC) type ‘G’ membrane transporters. ABC transporters are involved in active transport of several metabolites (like diterpenoids) and have indeed been detected in plant trichomes (using immunolocalization studies) that are producing terpenes (Stukkens et al., 2005; Crouzet et al., 2013). We obtained two artemisia PDR genes which were characterized to be involved in transport of dihydroartemisinic acid (personal communication, Prof. Marc Boutry), using BY2-cell exclusion assays. In an *in planta* export assay we could show that the artemisia PDR2 has a synergistic effect with artemisia LTP3 on transport of artemisinin intermediates to the apoplast in a transient assay. At present it is not clear whether this is due to higher intrinsic protein stability of AaPDR2 compared to AaPDR1(**chapter 4**). While for the Artemisia LTP3 no activity could be detected without AaPDR2, I was lucky that for the feverfew LTPs I was not dependent on co-expression of a PDR (at least not for LTP 1-3). I assume that in this case some endogenous *N. benthamiana* PDR activity could already accommodate transport over the plasma-membrane. Our feverfew transcriptomic data does contain a number of PDR sequences which show co-expression with parthenolide biosynthetic pathway genes (**chapter 5**). These would be good candidates to clone and to characterize in combination with the feverfew LTPs in transient assays. As PDRs from glandular trichomes may have evolved for efficient transport of feverfew sesquiterpenes, it is postulated that the feverfew PDRs in combination with the feverfew LTPs show more efficient transport of feverfew sesquiterpene products. Clearly the role of PDRs in sesquiterpene transport needs further investigation. Recently the involvement of a PDR in extracellular transport of the volatile sesquiterpene  $\beta$ -caryophyllene was reported (Fu et al., 2017), raising the question whether initially PDRs in ancestors of feverfew were active on the volatile germacrene A (the product of the first step in

feverfew sesquiterpene biosynthesis) and only later evolved to act on derivatives of germacrene A as the pathway in feverfew evolved with more P450 modifying activities.

### **Transport over the cell wall**

The lipophilic sesquiterpene lactones need PDR proteins for transport over the plasma membrane (PM). Once on the other side of the PM, they need to be transported over the hydrophilic cell wall to be stored in the sub-cuticular space of the trichome. In my research I speculated that the LTP proteins are involved in this latter step. Indeed, I showed that LTPs enhance accumulation of non-conjugated sesquiterpene lactones, which suggest they are in the apoplast (**chapters 4 and 5**). LTPs are small low molecular weight proteins, harbouring a signal peptide and no ER retention signal. LTPs are therefore secreted into the ER and move via the Golgi apparatus to the apoplast through the cell secretory pathway. Indeed, we could demonstrate the extra-cellular localisation of the *Artemisia* LTPs. We assume that the feverfew LTP1 and LTP2 show similar extracellular localisation, and this may be further confirmed in the future using GFP-fusion proteins. LTPs contain a hydrophobic pocket and show relatively high abundance in glandular trichomes of terpene producing plants (Bertea et al., 2006; Chatzopoulou et al., 2010). Candidate LTPs were selected from the trichome enriched cDNA library based on their expression profile. Interestingly the expression pattern of the most active LTPs in feverfew and *A. annua* suggests a delayed (late) expression, compared with the biosynthetic pathway genes. This delayed expression might be explained by the combined time required for biosynthesis gene expression, protein synthesis and biosynthesis of sesquiterpene lactones by the enzymes (expression of four and three biosynthetic pathway genes in feverfew and *A. annua*, respectively). Not all LTPs tested showed activity in our transport assays, raising the question whether these LTPs are involved in the transport of other sesquiterpenes, which were not produced in our pathway reconstructions in *N. benthamiana*. This hypothesis could be tested by isolation and characterization of other feverfew sesquiterpene lactone biosynthetic pathway genes and co-expression with other uncharacterized feverfew LTPs. A possible candidate for LTP transport is artemcanin. Artemcanin is a guaianolide sesquiterpene lactone detected in feverfew extracts (Fischedick et al., 2012). The postulated biosynthetic pathway of artemcanin consist of several oxidation steps beyond kauniolide (or costunolide). Furthermore, artemcanin shows a delayed accumulation pattern in feverfew ovary developmental stages, with the maximum accumulation occurring at the very last stage (**Chapter 2**), which would fit with the delayed expression pattern of feverfew LTPs.

## Novel assays for testing LTP activity

### Substrate export assays

LTPs are notoriously enigmatic because no good *in vivo* gain of function assays exist for their biological function. Enhanced secretion of both lipids and terpenes was shown by overexpression of tobacco LTP1 by the group of Choi et al. (2012), but in this case the specificity of the LTP was not tested. In my research I provided two important novel bioassays for LTP function: enhancing free sesquiterpene accumulation in a terpene export assay and blocking import of free sesquiterpenes from the apoplast into the cell in a terpene exclusion assay. We expressed biosynthetic pathway genes of the artemisinin and parthenolide pathways in *N. benthamiana* and co-expressed individual LTPs with the biosynthetic pathway genes. The *N. benthamiana* mesophyll leaf cells act as compound (sesquiterpene) production factories and the extracellular space is a proxy for the sub-cuticular space of the glandular trichome. With regard to the sesquiterpenes tested the LTPs showed remarkable specificity: AaLTP3 did not enhance costunolide/parthenolide accumulation when transiently co-expressed with the feverfew biosynthetic pathway genes in *N. benthamiana* and feverfew LTPs did not enhance accumulation of artemisinin pathway products. However, also within one plant species the LTPs show individual specificity as TpLTP1 and 2 improve extracellular accumulation of costunolide and TpLTP3 of parthenolide. Both the specificity of PDR transporters for sesquiterpenes and the specificity of LTPs for certain sesquiterpenes raises the question how these complicated multi-step biosynthesis pathways evolved: the evolution of a new P450 activity acting on a sesquiterpene product could result in the loss of the existing LTP activity on the novel product and a new LTP has to evolve to deal with the new product.

In feverfew itself the majority of parthenolide (~90%) is detected in its free form (Liu et al., 2014). When expressing the parthenolide biosynthetic pathway genes in *N. benthamiana* without LTPs the majority of detected parthenolide (~95%) is in the (glutathione, cysteine) conjugated form. However, when the parthenolide biosynthetic pathway is expressed together with TpLTP3, only the amount of free parthenolide in both total extract and apoplastic wash is increased (**chapter 5**). This shows that LTPs play a role in getting the sesquiterpenes out of reach of detoxification enzymes in the cell, by facilitating their transport to the apoplast.

### Substrate exclusion assays

Another way by which I characterised LTPs is by the substrate exclusion assay. This assay provides a measure on sesquiterpenes influx blockage into the cell. Influx was quantified by measuring intracellular conjugation (to sugar and glutathione conjugates) (**chapter 4**) or intracellular conversion of sesquiterpene lactones by the activity of a transiently co-expressed

cytochrome P450 (**chapter 5**). The substrate exclusion assays showed us that there is a very active influx of sesquiterpenes into the cell. This was obvious when artemisinic acid was infiltrated into the leaf, which was then directly extracted (1 min post infiltration). Analysis by UPLC-MRM-MS showed that artemisinic acid was rapidly converted to other artemisinin intermediates by *N. benthamiana* enzymes (**chapter 4**). The rapid uptake actually argues against the hydrophilic cell wall being a barrier for sesquiterpene transport. As the influx was (partly) blocked by LTPs we postulate that LTPs are also involved in lipid transport and create a lipophilic layer on the outside of the cell wall. Such lipid layer may also provide a sink for the sesquiterpenes and may prevent the reflux of exported sesquiterpenes. To test whether enhanced lipid metabolism in combination with the LTPs create a larger sink for sesquiterpenes we tested the effect of the transcription factors MIXTA and SHINE, which regulate cuticle development (Oshima et al., 2013). Hereto, MIXTA and SHINE were co-expressed with the parthenolide or costunolide biosynthetic pathway genes, either with or without TpLTP3 (parthenolide specific) or TpLTP1 (costunolide specific). MIXTA and SHINE did not affect the accumulation of free parthenolide or costunolide to a significant level but in the future it would be of interest to test whether MIXTA and SHINE do function in the sesquiterpene exclusion assay. Membranes and lipid layers may be labelled with different fluorescent probes. Unfortunately, the resolution of available techniques (confocal microscopy) is not high enough to distinguish between PM and a possible lipid layer on the outer cell wall. However, in the future a comprehensive lipidomics approach of the *N. benthamiana* plants expressing biosynthetic pathway genes with and without LTPs may provide additional information on whether the LTPs active on sesquiterpenes also affect the lipid profile upon ectopic expression in plants.

### **LTP function is sensitive to tagging**

All so far studied lipid transfer proteins contain a secretion targeting signal on their N-terminus. Therefore, where addition of a tag for localization/purification was required, the amino acid residues of the tag protein (GFP or His) were added to the C-terminus to prevent interference with the secretion peptide. In the case of the *A. annua* LTP3 the C-terminal GFP tag did not interfere with functionality and allowed for the observation that the LTP was stabilized in the presences of pathway activity. For the isolation of the feverfew TpLTP3 and TpLTP1 we fused a His-tag to the C-terminus. Expression of these constructs in *N. benthamiana* with the costunolide or parthenolide biosynthetic pathway genes showed that the His-tag interfered with LTP function. Such an interference has not been reported for LTPs before, however a hexahistidine affinity tag was shown to interfere with the catalytic properties of short chain dehydrogenases isolated from *Solanum dulcamara* (Freydank et al., 2008). Interestingly this effect is apparent when the His-tag was attached to the C-terminus of

the enzyme. Modelling studies showed that the hexahistidine tag on the C-terminus can interfere with the enzyme active site, possibly by steric or electrostatic interference (Freydank et al., 2008). It could be that the HIS tag on the LTPs give a similar steric hinder of the binding pocket and thus block their carrier activity. This also means that purification of LTPs using a HIS tag may not be the way to do binding studies with potential substrates.

### **Expression of LTPs in yeast**

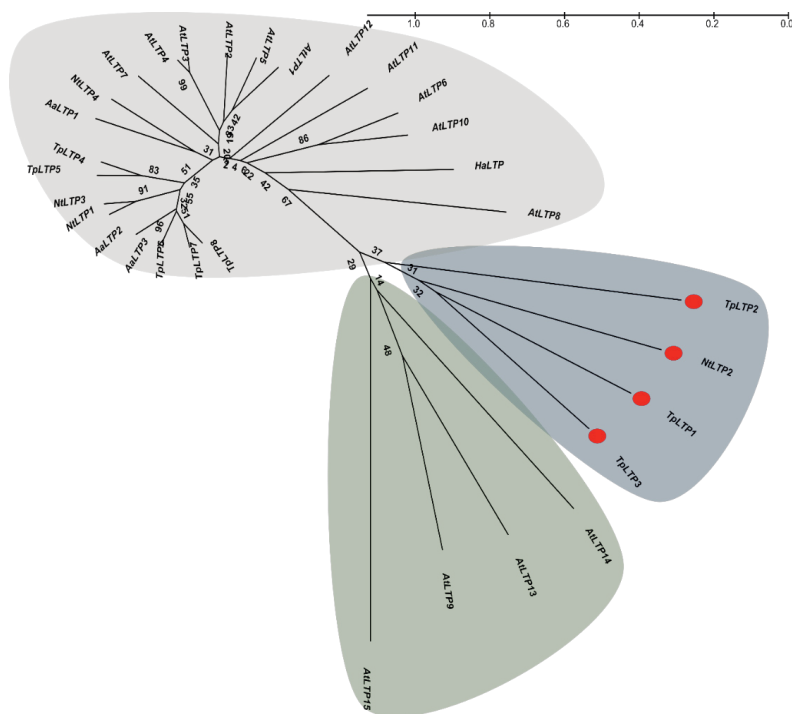
As an alternative platform for the production of medicinal sesquiterpenes, we have also used yeast as an expression system. Yeast has a different type of cell wall than plant cells and is cultured in liquid medium. We tested whether pathway activity in combination with LTPs also resulted in greater sesquiterpene production in yeast. This was done by expression of AaLTP3 together with artemisinic acid biosynthetic pathway genes (ADS and CYP71AV1) or TpLTP1/TpLTP2 with costunolide biosynthetic pathway genes (GAS, GAO and COS), in the WAT11 yeast strain. Although we could demonstrate successful production of (dihydro)artemisinic acid and costunolide, respectively, no significant positive effect was detected for the LTPs (data not shown). The accumulation pattern over 120 hrs also did not show any significant difference. To identify where the sesquiterpenes accumulated we separately analysed the medium and the yeast pellet. Surprisingly the majority of the produced sesquiterpene lactones were present in the medium with only 5% of the produced sesquiterpene lactones present in the yeast extract. Accumulation of sesquiterpene lactones in the yeast growth medium was not expected, due to their hydrophobicity. It could be that these compounds are suspended in microscopic lipid droplets. This could be tested by ultra-centrifugation of the yeast medium to determine if the sesquiterpenes can be concentrated by centrifugation. It has been demonstrated that extracellular lipid vesicles in yeast are involved in the molecular traffic across the cell wall, where such vesicles are implicated in transport of proteins, lipids, polysaccharide and pigments to the extracellular space (Oliveira et al., 2010; Rodrigues et al., 2015). Alternatively, the lack of effect of LTPs in yeast could be due to misfolding of the heterologously expressed LTPs. It has been reported that heterologous proteins might be retained in the endoplasmic reticulum of yeast because of its fairly strict biosynthetic quality control machinery. When the heterologously expressed LTPs show delayed or less precise folding they may never reach an active state (Ton and Rao, 2004).

### **Role of LTPs in transporting other biologically active molecules**

We have isolated eight feverfew LTPs (from which three were showing activity on parthenolide or costunolide) and three *A. annua* LTPs (showing activity on (dihydro)artemisinic acid) inter- and intracellular accumulation. Parthenolide and to a lesser extent costunolide are the main sesquiterpene lactones present in feverfew, however there are

many more sesquiterpene lactones detected in feverfew (e.g. 3 $\beta$ -hydroxycostunolide, 3 $\beta$ -hydroxyparthenolide, artemcanin, santamarine, etc.) (Fischedick et al., 2012). Similar to parthenolide, these sesquiterpene lactones are chemically reactive and thus require transport to the extracellular space of the trichome in order to prevent reaction with endogenous compounds of the cell. Therefore, we may expect similar transport mechanism to be involved in extracellular sequestration of these other sesquiterpene lactones. The other feverfew LTPs (not showing transport activity for costunolide and parthenolide) may thus be involved in extracellular transport of other sesquiterpene pathway products.

Sequence homology analysis of LTPs (Figure 2) showed that functional feverfew LTPs are not clustered with *Nt*LTP1, but with *Nicotiana tabacum* LTP2. *N. tabacum* LTP1, which is a glandular trichome specific lipid transfer protein, was shown to be involved in diterpenoid sequestration to the plant surface (Choi et al., 2012). *Nt*LTP2 is expressed in old leaves, petioles, roots and stems of *N. tabacum*. Although *Nt*LTP2 is not yet functionally characterized, our homology based cluster analysis results suggest a role of *Nt*LTP2 in transport of small lipophilic molecules in *N. tabacum*.



**Figure 2.** Phylogenetic tree of 15 *Arabiopsis thaliana*, 8 *Tanacetum parthenium*, 3 *Artemisia annua* and 4 *Nicotiana tabacum* LTP genes. Bootstrap values are shown in frequency values from 1000 replicates. There are 3 major clusters. The three functional TpLTPs cluster together with *Nt*LTP2.



As presented in **chapter 2**, some plant hormones have LogD values similar to the studied sesquiterpene lactones in this research. Moreover, the terpenoid-like plant hormones strigolactones, also require an ABCG-type plasma membrane transporter (PhDPR1) to be secreted (Sasse et al., 2015), but the subsequent transport over the cell wall of strigolactones has not been investigated. In *Petunia hybrida* PhDPR1 is apically localised in cells of the root tip, suggesting a polar cell-to-cell transport of strigolactones (Sasse et al., 2015). Such polar strigolactone transport would also require a transport over the cell wall, which may be facilitated by LTPs. A recent study showed altered levels of ABA and salicylic acid when *Arabidopsis* LTP3 was overexpressed (Gao et al., 2016), but whether this is due to an effect on hormone transport or an effect on cuticle deposition (or both) is at present not clear. In *Arabidopsis* more than 70 LTPs have been annotated (Beisson et al., 2003). A combination of transcriptome analysis, LTP knock-down/knock-out lines, complementation of LTP mutants and metabolomics studies may provide more insight in a possible role for LTPs in plant hormone transport.

## Prospects

My thesis project provides insight into the biosynthesis and transport of sesquiterpene lactones. The functional characterization of the novel cytochrome P450, kauniolide synthase, was the first step towards elucidation of guaianolide sesquiterpene lactone biosynthesis. In the future, the challenge is to identify the additional enzymes that modify kauniolide to other biologically active guaianolides, such as artecannabin. For metabolic engineering of such compounds, a bottleneck is that many enzymes are involved in their biosynthesis (for kauniolide for example: germacrene A synthase, germacrene A oxidase, costunolide synthase and kauniolide synthase) which will result in a stepwise decrease in product level and hence substrate for the next biosynthetic step and therefore low final yield. Metabolic engineering approaches such as boosting the precursor, utilization of the most efficient enzymes, down regulation of competing pathways (for example downregulation of 5-*epi*-aristolochene synthase and squalene synthase in *N. benthamiana*; (Cankar et al., 2015) or restriction of the squalene formation by downregulation of *ERG9* expression in yeast (Paddon et al., 2013)) and implementation of knowledge obtained in this thesis on transport of sesquiterpene lactones should be used to optimise the production, on a large scale, of kauniolide and sesquiterpene lactones derived from it by even more additional enzymatic steps. With regard to transport and sequestration, the characterization of LTPs and PDRs in this thesis provide a first basis of molecular knowledge about extracellular transport and storage of sesquiterpene lactones. This knowledge offers opportunities to optimise biotechnological production in combination with improved sequestration, of the medicinally so important sesquiterpene lactones.

## References

- Aharoni A, Jongsma MA, Bouwmeester HJ (2005) Volatile science? Metabolic engineering of terpenoids in plants. *Trends in plant science* **10**: 594-602
- Beisson F, Koo AJ, Ruuska S, Schwender J, Pollard M, Thelen JJ, Paddock T, Salas JJ, Savage L, Milcamps A (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, 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
- Brennan TC, Turner CD, Kromer JO, Nielsen LK (2012) Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **109**: 2513-2522
- Caldas C, McGuire WP, 3rd (1993) Taxol in epithelial ovarian cancer. *J Natl Cancer Inst Monogr*: 155-159
- Cankar K, Jongedijk E, Klompmaier M, Majdic T, Mumm R, Bouwmeester H, Bosch D, Beekwilder J (2015) (+)-Valencene production in *Nicotiana benthamiana* is increased by down-regulation of competing pathways. *Biotechnology journal* **10**: 180-189
- Cankar K, van Houwelingen A, Bosch D, Sonke T, Bouwmeester H, Beekwilder J (2011) A chicory cytochrome P450 mono-oxygenase CYP71AV8 for the oxidation of (+)-valencene. *FEBS letters* **585**: 178-182
- Chatzopoulou FM, Makris AM, Argiriou A, Degenhardt J, Kanellis AK (2010) EST analysis and annotation of transcripts derived from a trichome-specific cDNA library from *Salvia fruticosa*. *Plant cell reports* **29**: 523-534
- 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
- Connolly J, Hill R (1991) Dictionary of Terpenoids. Vol. 1: Mono- and Sesquiterpenoids, Vol. 2: Diterpene and higher Terpenoids, Vol. 3: Indexes. In: New York, Tokyo, Melbourne, Madras: Chapman & Hall London
- 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
- 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* **155**: 365-372
- Eudes A, Pereira JH, Yogiswara S, Wang G, Teixeira Benites V, Baidoo EE, Lee TS, Adams PD, Keasling JD, Loqué D (2016) Exploiting the substrate promiscuity of hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase to reduce lignin. *Plant and Cell Physiology* **57**: 568-579
- Fischedick JT, Standiford M, Johnson DA, De Vos RCH, Todorović S, Banjanac T, Verpoorte R, Johnson JA (2012) Activation of antioxidant response element in mouse primary cortical cultures with sesquiterpene lactones isolated from *Tanacetum parthenium*. *Planta medica* **78**: 1725-1730
- Freydank AC, Brandt W, Dräger B (2008) Protein structure modeling indicates hexahistidine-tag interference with enzyme activity. *Proteins* **72**: 173-183
- Fu X, Shi P, He Q, Shen Q, Tang Y, Pan Q, Ma Y, Yan T, Chen M, Hao X, Liu P, Li L, Wang Y, Sun X, Tang K (2017) AaPDR3, a PDR Transporter 3, Is Involved in Sesquiterpene  $\beta$ -Caryophyllene Transport in *Artemisia annua*. *Frontiers in Plant Science* **8**: 723
- Gao S, Guo W, Feng W, Liu L, Song X, Chen J, Hou W, Zhu H, Tang S, Hu J (2016) LTP3 contributes to disease susceptibility in *Arabidopsis* by enhancing abscisic acid (ABA) biosynthesis. *Mol Plant Pathol* **17**: 412-426
- Huang Q, Roessner CA, Croteau R, Scott AI (2001) Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorganic & medicinal chemistry* **9**: 2237-2242
- Ignea C, Pontini M, Maffei ME, Makris AM, Kampranis SC (2014) Engineering Monoterpene Production in Yeast Using a Synthetic Dominant Negative Geranyl Diphosphate Synthase. *ACS Synthetic Biology* **3**: 298-306
- Jindal S, Longchar B, Singh A, Gupta V (2015) Promoters of AaGL2 and AaMIXTA-Like1 genes of *Artemisia annua* direct reporter gene expression in glandular and non-glandular trichomes. *Plant Signaling & Behavior* **10**: e1087629
- Julsing MK, Koulman A, Woerdenbag HJ, Quax WJ, Kayser O (2006) Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomolecular engineering* **23**: 265-279

- Khairul Ikram NKB, Beyraghdar Kashkooli A, Peramuna AV, van der Krol AR, Bouwmeester H, Simonsen HT (2017) Stable Production of the Antimalarial Drug Artemisinin in the Moss *Physcomitrella patens*. *Frontiers in Bioengineering and Biotechnology* 5
- Liu M, Shi P, Fu X, Brodelius PE, Shen Q, Jiang W, He Q, Tang K (2016) Characterization of a trichome-specific promoter of the aldehyde dehydrogenase 1 (ALDH1) gene in *Artemisia annua*. *Plant Cell, Tissue and Organ Culture (PCTOC)* 126: 469-480
- 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 Q, Manzano D, Tanic N, Pesic M, Bankovic J, Pateraki I, Ricard L, Ferrer A, de Vos R, van der Krol S, Bouwmeester H (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metab Eng* 23: 145-153
- Mafu S, Jia M, Zi J, Morrone D, Wu Y, Xu M, Hillwig ML, Peters RJ (2016) Probing the promiscuity of ent-kaurene oxidases via combinatorial biosynthesis. *Proceedings of the National Academy of Sciences* 113: 2526-2531
- Majdi M, Liu Q, Karimzadeh G, Malboobi MA, Beekwilder J, Cankar K, Vos R, Todorovic S, Simonovic A, Bouwmeester H (2011) Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.). *Phytochemistry* 72: 1739-1750
- Nguyen DT, Göpfert JC, Ikezawa N, MacNevin G, Kathiresan M, Conrad J, Spring O, Ro D-K (2010) Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. *Journal of Biological Chemistry* 285: 16588-16598
- O'Reilly E, Corbett M, Hussain S, Kelly PP, Richardson D, Flitsch SL, Turner NJ (2013) Substrate promiscuity of cytochrome P450 RhF. *Catalysis Science & Technology* 3: 1490-1492
- Oliveira DL, Nakayasu ES, Joffe LS, Guimarães AJ, Sobreira TJP, Nosanchuk JD, Cordero RJB, Frases S, Casadevall A, Almeida IC, Nimrichter L, Rodrigues ML (2010) Characterization of Yeast Extracellular Vesicles: Evidence for the Participation of Different Pathways of Cellular Traffic in Vesicle Biogenesis. *PLoS ONE* 5: e11113
- Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M (2013) MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in *Arabidopsis* and *Torenia fournieri*. *The Plant Cell* 25: 1609
- Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell M, Tai A, Main A, Eng D (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496: 528
- Rao KV (1993) Taxol and related taxanes. I. Taxanes of *Taxus brevifolia* bark. *Pharm Res* 10: 521-524
- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440: 940
- Rodrigues ML, Godinho RMC, Zamith-Miranda D, Nimrichter L (2015) Traveling into Outer Space: Unanswered Questions about Fungal Extracellular Vesicles. *PLoS Pathogens* 11: e1005240
- Sasse J, Simon S, Gübeli C, Liu G-W, Cheng X, Friml J, Bouwmeester H, Martinoia E, Borghi L (2015) Asymmetric localizations of the ABC transporter PaPDR1 trace paths of directional strigolactone transport. *Current Biology* 25: 647-655
- Schellmann S, Hulskamp M (2004) Epidermal differentiation: trichomes in *Arabidopsis* as a model system. *International Journal of Developmental Biology* 49: 579-584
- Stukkens Y, Bultreys A, Grec S, Trombik T, Vanham D, Boutry M (2005) NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. *Plant Physiology* 139: 341-352
- 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 Phytol* 199: 352-366
- Ton V-K, Rao R (2004) Functional expression of heterologous proteins in yeast: insights into Ca<sup>2+</sup> signaling and Ca<sup>2+</sup>-transporting ATPases. *American Journal of Physiology-Cell Physiology* 287: C580-C589
- van Herpen TWJM, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J (2010) *Nicotiana benthamiana* as a Production Platform for Artemisinin Precursors. *PLOS ONE* 5: e14222
- Wallaart TE, Bouwmeester HJ, Hille J, Poppinga L, Majers NC (2001) Amorpha-4, 11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. *Planta* 212: 460-465
- Wang B, Kashkooli AB, Sallets A, Ting H-M, de Ruijter NCA, Olofsson L, Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR (2016) Transient production of artemisinin in *Nicotiana*

- benthamiana is boosted by a specific lipid transfer protein from *A. annua*. *Metabolic Engineering* **38**: 159-169
- Winichayakul S, Scott RW, Roldan M, Hatier J-HB, Livingston S, Cookson R, Curran AC, Roberts NJ** (2013) In vivo packaging of triacylglycerols enhances *Arabidopsis* leaf biomass and energy density. *Plant Physiology* **162**: 626-639
- 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

## Summary

This thesis mainly focuses on sesquiterpene lactone biosynthesis, their (combinatorial) metabolic engineering and extracellular transport and accumulation. Terpenoids are the most diverse and largest class of natural products, many of which have important pharmaceutical, biological, nutraceutical and/or other beneficial properties. **Chapter 1** of this thesis provides the background on the knowledge of biosynthesis of different terpenoid classes in plants at the onset of this thesis work. Also, a background is given of metabolic engineering approaches and different expression platforms (organisms) to overcome limitations of production of these attractive molecules. Finally the topic ‘terpenoid transport’ in plants is introduced, a research area which was largely underdeveloped at the start of my thesis.

There are different classes of sesquiterpene lactones from which particularly the biosynthesis of germacranolides has been studied, mostly in plant species belonging to the Asteraceae. In contrast, how members of the guaianolide class of sesquiterpene lactones are synthesized has been a mystery for a very long time. In **chapter 2**, I report the first characterization of an enzyme from feverfew responsible for the first committed step branching the guaianolide pathway off from the germacranolides. This enzyme, which converts costunolide into kauniolide, is very special as it performs several sequential reactions, from hydroxylation, water elimination and cyclisation to regio-selective deprotonation. The mechanism of action of this enzyme was elucidated by testing different putative substrates and intermediates (obtained through chemical synthesis) and through *in-silico* substrate docking studies.

Recent progress in metabolic engineering and pathway reconstruction in microbes and plants has resulted in the production of the bioactive costunolide and parthenolide from feverfew as well as dihydroartemisinic acid, the precursor of the antimalarial artemisinin, from *Artemisia annua* in heterologous expression platforms. In *A. annua*, a double bond reductase (AaDBR) enzyme catalyses a branch point in the pathway towards dihydroartemisinic acid. Because artemisinin biosynthesis in *A. annua* proceeds quite similar to costunolide biosynthesis in feverfew, we suspected that AaDBR may act on products from the feverfew pathway. This was confirmed in **chapter 3** in a combinatorial metabolic engineering approach, in which I produced the novel dihydro-sesquiterpene lactones 3 $\beta$ -hydroxy-dihydroparthenolide and 3 $\beta$ -hydroxy-dihydrocostunolide.

In **Chapter 4** I switch from biosynthesis of sesquiterpene lactones to their transport. In this chapter I describe the involvement of Lipid Transfer Proteins (LTPs) and Pleiotropic Drug Resistance (PDR) membrane transporters in extracellular accumulation of (dihydro)artemisinic acid [(DH)AA], using transient expression assays in the heterologous host *Nicotiana benthamiana*. We show that AaLTP3 and AaPDR2 enhance (DH)AA accumulation in the *N. benthamiana* leaf apoplast as well as the overall flux through the (DH)AA pathway. For functional analysis of the LTPs I developed two novel assays: an *in planta* substrate export assay and an *in planta* substrate exclusion assay. Using these assays, we show that AaLTP3 is more effective than AaPDR2 in preventing influx of (DH)AA from the apoplast into the cells. This chapter also describes the first documented case of artemisinin and arteannuin B production in *N. benthamiana*.

In **chapter 5**, I continue my research on sesquiterpene transport by functional characterization of Lipid Transfer Protein genes from feverfew trichomes. I demonstrate that some of these LTPs have acquired a specialized function in the transport of specific lipophilic products (sesquiterpene lactones) produced in feverfew trichomes. I show that TpLTP1 and TpLTP2 specifically improve costunolide export and that TpLTP3 highly specifically improves parthenolide export. Moreover, the substrate exclusion assays revealed TpLTP3 as the most effective in blocking influx of both costunolide and parthenolide. TpLTP3 has a GPI anchor domain and I show that this GPI-anchor domain is essential for the activity of TpLTP3, while addition of this domain to TpLTP1 resulted in loss of TpLTP1 activity.

In **chapter 6** I discuss the biosynthesis and transport results obtained during my thesis research. I address remaining questions related to biosynthesis of sesquiterpene lactones in yeast and discuss opportunities for combinatorial metabolic engineering, e.g. what are putative applications of the enzyme-substrate promiscuity concept for the double bond reductase (DBR)? In this chapter I also discuss the many remaining questions related to transport of sesquiterpene lactones, as many details about transport over the plasma membrane and over the cell wall are still missing. Finally, I provide a wider perspective on the role of LTPs, based on clues from the exclusion assays and the chemical properties of other biologically active molecules in plants.

In summary, this thesis provides several novel discoveries in the field of sesquiterpene lactone metabolic engineering by identification of the biosynthetic branch point towards guaianolides, by combinatorial metabolic engineering to produce novel dihydro-sesquiterpene lactones and by revealing an important and very specific role of LTPs in extracellular transport of sesquiterpene lactones.

## Acknowledgments

This thesis would have not been reached to this status without the help and support of many people. First of all I would like to thank the almighty God for all the blessings. I also would like to thank the Iranian Ministry of Science, Research and Technology for the financial support of my PhD study at Wageningen University and Research.

My sincere thanks go to my daily supervisor and co-promotor, Dr. Sander van der Krol, and my promotor Prof. Harro Bouwmeester. Harro and Sander thanks for accepting me in the Terpene group.

Sander, I am pretty much sure that I cannot show my gratitude within these lines. You were not only a great daily scientific supervisor, but also were full of understanding. Your scientific and personal advices, learnt me how to self-criticize, but still being self-motivated. Sander, you have spent much time to supervise me and provided me chances to improve my skills in a continuous learning process. This, I wish to be continued...!

Harro, having you as the promotor was fantastic. I have learnt a lot from you. You have opened up many international opportunities for me by getting me involved in side projects, which resulted in two publications on which I am the co-author. The multidisciplinary nature of metabolic engineering requires these collaborations; you made it easy!

I would also like to appreciate all kind helps and supports from all Plant Physiology staff members, from whom I always received precious feedbacks. Mariëlle, Francel and Jacqueline, you were all great technicians with a great sense of responsibility. You have learnt me a lot and also provided an enjoyable atmosphere in the lab for which I am very thankful. I also would like to thank Leo, Juriaan and Lidiya for their helps during the past years. Iris, our collaboration lasted for more than one year which I enjoyed very much. Thanks for allowing me being a part of your very much multidisciplinary bio-interaction group. I also would like to express my appreciation to three of my best friends in Wageningen: Jules, Katja and Thierry. We spent a lot of time in the lab, during conferences, meetings and other side activities together, thanks for all those fantastic moments. I also would like to extend my appreciation to all my colleagues: Esmer, Mariana, Beatriz, Mahdere, Giovanni, Gonda, Alexandre, Bing, Deborah, Elise, Natalia, Emilie, Krystina, Qing, Bo, Lemeng, Rymyana, Melissa, Diaan, Caroline, Nikita, Karen, Neli, Desalegn, Margaret, Jimmy, Yuanyuan, Paolo, Alice, Carmen, Jun, Farzaneh, Mark vH., Umidjun, Bas, Sangsoek, Rina, Renake, Wei, Rummyana, Shuang, Farshid, my students (Joris, Mara and Sandra), my dear paranymphs Mark Levisson and Jalal Samia and others whom I might have forgotten. Saadat, we have met 9 years ago and became best friends; I don't forget all these moments and your support and help with every single problem that I had. I am happy that life path brought you to Wageningen and thanks to the kind invitation from Dr. Wilco Ligterink you did your sabbatical at Plant Physiology so we could spend more time together.

Being abroad, far away your family was difficult. Jalal, Shaghayegh and Taher we shared many beautiful moments together and became like a family. I believe our friendship will be for the lifetime. Narges and Mehdi, I wish you all the best in your personal life and career and



believe our friendship will stay for ever. I would like to thank all Iranian families (Sasan and Maryam, Saeid and Maryam, Mansoor and Farzaneh, Omid and Atousa, Mohammad and Rojin, Siavash and Samin). Also many thanks to my friends (Arash, Naser, Ehsan, Ahmad, Amir, Mahyar, Ayoub, Afshin, Abuzar); hope we can soon play football again, all together. Talking about all my Iranian friends, I could not forget the great supports from staff members of the Department of Horticulture of Tarbiat Modares University. Thanks a lot for all supports from the very first moments I have started my PhD. My special thanks to my supervisors during my Master program, Prof. R. Omidbaigi (passed away) and Prof. M.J. Saharkhiz. I also would like extend my gratitude to Dr. Alireza Babaei, as my study coordinator, who has supported me every single moment throughout the whole PhD program.

I am very much grateful to my dearest family members. Baba Ali and Maman Mahrokh, thousand words cannot explain my love and passion to you. Your patience, love and support resulted to this very specific moment. Very proud of having you as parents. My next thank goes to my dear parents in law, Madar Shakiba and Pedar Siavash. Thank you very much; your kindness and love during the past years means a lot to me. My dear brother, Omid, we have spent our beautiful childhood together. My sincere thanks to you and your wife Rezvan for being here and supporting me.

Last but not least I would like to express my deepest gratitude to my beloved wife, Nafiseh. Nafis, I still remember the sentence from my PhD coordinator 'Your wife is your best friend, ever'. You were, are and will be my closest person for ever. We have spent all our time together even at work, by sitting next to each other, in the same office. Nafis, I never forget your patience and love during the past years. This thesis compilation would have not been possible without your supports. My dear Nafiseh, we started the PhD journey together and I wish you all the best in finalizing your PhD thesis.

## About the author



Arman Beyraghdar Kashkooli was born on 13 June 1987 in Shiraz, Iran. He had started his bachelor on September 2005 and was graduated on September 2009 in the field of Horticultural Sciences from Shiraz University, Shiraz, Iran. Arman has been accepted to the Master's program with the excellent rank of two (among more than 4000 participants, nationwide). He did his research in the field of Physiology and Breeding of Medicinal Plants with the focus on ecology and the allelopathic properties of endemic Iranian *Thymus daenensis* Celak in Tarbiat Modares University (Tehran, Iran) under supervision of Prof. R. Omidbaigi (passed away) and Prof. M.J. Saharkhiz and graduated in 2011. Arman has been awarded a scholarship from Iranian Ministry of Science, Research and Technology and accepted to the PhD program in the laboratory of Plant Physiology of Wageningen University and Research (The Netherlands) under supervision of Prof. H.J. Bouwmeester and Dr. Sander van der Krol. He did his research in the field of Metabolic Engineering and his doctorate thesis is entitled "Biosynthesis, transport and combinatorial metabolic engineering of *Tanacetum parthenium* (feverfew) and *Artemisia annua* (sweet wormwood) sesquiterpene lactones".

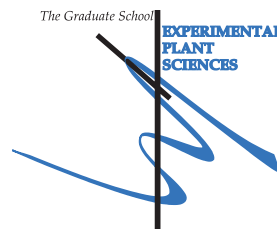
## List of Publications

- Dickschat JS, Rinkel J, Rabe P, **Beyraghdar Kashkooli A**, Bouwmeester HJ. (2017) 18-Hydroxydolabella-3, 7-diene synthase—a diterpene synthase from *Chitinophaga pinensis*. Beilstein journal of organic chemistry. 13, 1770.
- Binti Khairul Ikram NK, **Beyraghdar Kashkooli A**, Vithakshana Peramuna A, van der Krol AR, Bouwmeester H, Toft Simonsen H. (2017) Stable production of the antimalarial drug artemisinin in the moss *Physcomitrella patens*. Frontiers in Bioengineering and Biotechnology. 5: 1-8.
- Wang B, **Beyraghdar Kashkooli A**, Sallets A, Ting HM, de Ruijter NC, Olofsson L, Brodelius P, Pottier M, Boutry M, Bouwmeester H, van der Krol AR. (2016) Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. Metabolic engineering. 38:159-169.
- **Beyraghdar Kashkooli A**, Saharkhiz MJ. (2014) Essential Oil Compositions and Natural Herbicide Activity of Four Denaei Thyme (*Thymus daenensis* Celak.) Ecotypes. Journal of Essential Oil Bearing Plants. 5:859-874.

## Education Statement of the Graduate School

### Experimental Plant Sciences

**Issued to:** Arman Beyraghdar Kashkooli  
**Date:** 10 September 2018  
**Group:** Laboratory of Plant Physiology  
**University:** Wageningen University & Research



| 1) Start-up phase                                                                                                                                                | <u>date</u> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| ► <b>First presentation of your project</b><br>Metabolite engineering of feverfew parthenolide pathway                                                           | 21 Sep 2013 |
| ► <b>Writing or rewriting a project proposal</b><br>Metabolite engineering of feverfew parthenolide pathway                                                      | 29 Jun 2013 |
| ► <b>Writing a review or book chapter</b><br>Terpene Biosynthesis in Plants in 'Flavour Science',<br>Proceedings from 15th Weurman Flavour Research<br>Symposium | Jul 2018    |
| ► <b>MSc courses</b>                                                                                                                                             |             |
| ► <b>Laboratory use of isotopes</b>                                                                                                                              |             |

*Subtotal Start-up Phase*

*13.5 credits \**

| 2) Scientific Exposure                                                                                                                | <u>date</u>    |
|---------------------------------------------------------------------------------------------------------------------------------------|----------------|
| ► <b>EPS PhD student days</b><br>EPS PhD student day, Amsterdam, NL                                                                   | 29 Nov 2013    |
| Wageningen PhD symposium 'Healthy Food and Living<br>Environment', Wageningen, NL                                                     | 10 Dec 2013    |
| EPS PhD student day 'Get2Gether', Soest, NL                                                                                           | 28-29 Jan 2016 |
| ► <b>EPS theme symposia</b><br>EPS theme 4 symposium 'Genome Biology', Wageningen,<br>NL                                              | 03 Dec 2014    |
| EPS theme 3 symposium 'Metabolism and Adaptation',<br>Utrecht, NL                                                                     | 10 Feb 2015    |
| EPS theme 3 symposium 'Metabolism and Adaptation',<br>Amsterdam, NL                                                                   | 23 Feb 2016    |
| EPS theme 3 symposium 'Metabolism and Adaptation',<br>Wageningen, NL                                                                  | 14 Mar 2017    |
| <b>National meetings (e.g. Lunteren days) and other national<br/>platforms</b>                                                        |                |
| Annual meeting 'Experimental Plant Sciences', Lunteren, NL                                                                            | 22-23 Apr 2013 |
| Annual meeting 'Experimental Plant Sciences', Lunteren, NL                                                                            | 11-12 Apr 2016 |
| Annual meeting 'Experimental Plant Sciences', Lunteren, NL                                                                            | 10-11 Apr 2017 |
| ► <b>Seminars (series), workshops and symposia</b><br><i>Mini-symposium:</i> 'Arabidopsis in Wageningen' by Prof.<br>Marteen Koorneef | 11 Apr 2013    |
| <i>Symposium:</i> Start symposium Plant Developmental Biology,<br>Wageningen, NL                                                      | 14 Oct 2013    |

|                                                                                                                                                          |                       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| <i>Symposium: WURomics: Technology- Driven Innovation for Plant Breeding, Wageningen, NL</i>                                                             | 15 Dec 2016           |
| <i>Seminar: 'Glandular trichomes of tomato: from terpene biosynthesis to trichome differentiation' by Prof. Alain Tissier</i>                            | 10 Apr 2013           |
| <i>Seminar: 'Iridoid glycosides: biochemistry and role in biotic interactions in ribwort plantain' by Dr. Arjen Biere</i>                                | 03 May 2013           |
| <i>Seminar: 'Elucidating the Biosynthetic Pathway for Vibrallactone in the Basidiomycete Fungus <i>Boreostereum vibrans</i>' by Ying Zeng</i>            | 24 Oct 2013           |
| <i>Seminar: 'Terpenes from Microorganisms - Isotopic Labellings for Structure Elucidation and Biosynthetic Studies' by Prof. Dr. Jeroen S. Dickschat</i> | 12 Oct 2015           |
| <i>Seminar: 'Harnessing plant metabolic diversity' by Prof. Anne Osbourn</i>                                                                             | 31 Aug 2016           |
| <i>Seminar: 'The Arabidopsis BIR family – negative regulators of BAK1 receptor complexes and more' by Dr. Birgit Kemmerling</i>                          | 25 Nov 2016           |
| <i>Seminar: 'Pelargonidin in flowers - why not? Gerbera and petunia flowers block pelargonidin biosynthesis in a different way' by Teemu Teeri</i>       | 14 Mar 2018           |
| <i>Workshop: Advanced Light Microscopy facilities of Wageningen UR, Wageningen, NL</i>                                                                   | 13 Jun 2013           |
| <i>Workshop: CLC Bio - Analysis, interpretation, and reporting in Next Generation Sequencing</i>                                                         | 09 Dec 2015           |
| ► <b>Seminar plus</b>                                                                                                                                    |                       |
| ► <b>International symposia and congresses</b>                                                                                                           |                       |
| Terpnet 2015, Vancouver, Canada                                                                                                                          | 01-05 Jun 2015        |
| Gordon Research Conference -Plant Volatiles Across Multiple Scales: From Molecular Mechanisms to Ecological Functions, Lucca (Barga), Italy              | 03-04 Feb 2018        |
| ► <b>Presentations</b>                                                                                                                                   |                       |
| <i>Poster: Terpnet 2015, Vancouver, Canada</i>                                                                                                           | 01-03 Jun 2015        |
| <i>Poster: Annual meeting 'Experimental Plant Sciences', Lunteren, NL</i>                                                                                | 11-12 Apr 2016        |
| <i>Poster: Annual meeting 'Experimental Plant Sciences', Lunteren, NL</i>                                                                                | 10-11 Apr 2017        |
| <i>Poster: Gordon Research Conference, Lucca (Barga), Italy</i>                                                                                          | 03-04 Feb 2018        |
| <i>Talk: Plant Physiology PhD Trip, Max Planck Institute, Cologne, Germany</i>                                                                           | 22 Apr-01 May 2015    |
| <i>Talk: EPS theme 3 symposium 'Metabolism and Adaptation', Amsterdam, NL</i>                                                                            | 23 Feb 2016           |
| ► <b>IAB interview</b>                                                                                                                                   |                       |
| ► <b>Excursions</b>                                                                                                                                      |                       |
| Plant Physiology PhD Trip, the Netherlands, Germany and Switzerland                                                                                      | 22 Apr-01 May 2015    |
| <i>Subtotal Scientific Exposure</i>                                                                                                                      | <i>14.9 credits *</i> |

|                                                                       |                |
|-----------------------------------------------------------------------|----------------|
| <b>3) In-Depth Studies</b>                                            | <u>date</u>    |
| ▶ <b>EPS courses or other PhD courses</b>                             |                |
| Course: An Introduction to Mass Spectrometry-based Plant Metabolomics | 09-13 Dec 2013 |
| Course: Bioinformatics- A User's Approach                             | 25-29 Aug 2014 |
| Course: Data Analyses and Visualizations in R, Wageningen, NL         | 11-12 May 2017 |
| ▶ <b>Journal club</b>                                                 |                |
| Plant Physiology literature discussion group                          | 2013-2017      |
| ▶ <b>Individual research training</b>                                 |                |

*Subtotal In-Depth Studies*

6.5 credits \*

|                                                                                 |                |
|---------------------------------------------------------------------------------|----------------|
| <b>4) Personal development</b>                                                  | <u>date</u>    |
| ▶ <b>Skill training courses</b>                                                 |                |
| Mini-symposium: How to Write a World-class Paper, Wageningen, NL                | 17 Oct 2013    |
| Workshops: Wageningen Graduate School PhD Workshop Carousel, Wageningen, NL     | 02 Jun 2014    |
| Workshops: Wageningen Graduate School PhD Workshop Carousel, Wageningen, NL     | 08 Apr 2016    |
| Course: Scientific Writing, Wageningen, NL                                      | Sep-Nov 2016   |
| Course: Information Literacy PhD including Endnote Introduction, Wageningen, NL | 06-07 Dec 2016 |
| ▶ <b>Organisation of PhD students day, course or conference</b>                 |                |
| Discussion leader Gordon Research Conference, Lugga (Barga), Italy              | 03-04 Feb 2018 |
| ▶ <b>Membership of Board, Committee or PhD council</b>                          |                |

*Subtotal Personal Development*

3.2 credits \*

|                                                                                                                                                                                                           |                       |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| <b>TOTAL NUMBER OF CREDIT POINTS</b>                                                                                                                                                                      | <b>38.1 credits *</b> |
| 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 Work has been conducted at the Laboratory of Plant Physiology of Wageningen University and Research, Wageningen, Netherlands, and was financially supported by the Iranian Ministry of Science, Research and Technology (MSRT).

Thesis layout: Arman Beyraghdar Kashkooli

Cover design: Samad Bahrami Kashkooli, [kashkoolisamad@gmail.com](mailto:kashkoolisamad@gmail.com)

Printed by: ProefschriftMaken, || [www.proefschriftmaken.nl](http://www.proefschriftmaken.nl)

