Elucidation of the sesquiterpene lactone biosynthetic pathway in feverfew (Tanacetum parthenium)

Qing Liu
Thesis committee

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

Co-promotor
Dr. S.V.D. Krol
Associate professor, Laboratory of Plant Physiology
Wageningen University

Other members
Prof. dr. A. Ferrer, Centre for Research in Agricultural Genomics, Bellaterra, Spain
Dr. M.C.R. Franssen, Wageningen University
Prof. dr. M.E. Schranz, Wageningen University
Dhr. dr. ir. R.C. Schuurink, University of Amsterdam

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.
Elucidation of the sesquiterpene lactone biosynthetic pathway in feverfew (Tanacetum parthenium)

Qing Liu
Qing Liu

Elucidation of the sesquiterpene lactone biosynthetic pathway in feverfew (Tanacetum parthenium)
134 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2013)
With summaries in English and Dutch

ISBN 978-94-6173-756-4
Contents

**Chapter 1**  General introduction 7

**Chapter 2**  Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.) 19

**Chapter 3**  Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana* 39

**Chapter 4**  *In planta* reconstitution of the parthenolide biosynthetic pathway 59

**Chapter 5**  Kauniolide synthase, a cytochrome P450 terpenoid cyclase that catalyses the first step in guaianolide sesquiterpene lactone biosynthesis 87

**Chapter 6**  General discussion 105

Summary 121

Samenvatting 123

Acknowledgements 125

*Curriculum vitae* 128

Publication list 129

Education statement 131

Funding 133
Chapter 1

General introduction

Qing Liu
Chapter 1

Plant secondary metabolites and terpenoids

Plants produce a diverse group of secondary metabolites that are required for adaptation to their environment. The majority of plant secondary metabolites include alkaloids, cyanogenic glycosides, saponins, tannins, flavonoids, anthocyanins and terpenoids. These compounds are involved in many aspects of plants during development, such as defense against herbivores and pathogens, attraction of pollinators, regulation of symbiosis, control of seed germination, and chemical inhibition of competing plant species (Makkar et al., 2007).

Terpenoids are the most numerous and structurally diverse class of secondary metabolites. As built of basic five-carbon units, they can be categorized as mono- (C10), sesqui- (C15), di- (C20), tri- (C30), tetra- (C40) and poly- (>C40) terpenoids. Terpenoids play diverse function roles in plants as hormones (strigolactones, gibberellins and abscisic acid), electron carriers (plastoquinone and ubiquinone), mediators of polysaccharide assembly (polyprenyl phosphates), structural components of membranes (phytosterols), and photosynthetic pigments (carotenoids) (Theis & Lerdau, 2003). Besides the importance for the plant itself, terpenoids are also important to humans as many of them can be used as commercial flavourings, fragrances, antimalarial drugs and anticancer drugs (Jirschitzka et al., 2012). They are usually the major constituents of essential oils (commonly C5, C10, C15 and C20 terpenoids) of most plants offering a wide variety of pleasant scents, such as flowery, fruity, woody or balsamic notes, which makes them a very important class of compounds for flavour and fragrance industries, e.g. menthol from wild mint (Mentha arvensis); D-carvone from caraway (Carum carvi); D-limonene from citrus species; citral from lemongrass (Cymbopogon citratus); 1,8-cineole from eucalyptus (Eucalyptus globulus) (Caputi & Aprea, 2011). Some terpenoids are widely applied either directly as drugs, or as lead compound for the synthesis of drugs that mimic terpenoids found in nature (Feher & Schmidt, 2003). A sesquiterpene lactone artemisinin, an antimalarial drug, was isolated from a herb Artemisia annua which has been used for many centuries in Chinese tradition medicine as a treatment for fever and malaria (Klayman, 1985). Another terpenoid paclitaxel (taxol), first isolated form the pacific yew tree (Taxus brevifolia), has been considered as the most successful anticancer drugs (Heinig et al., 2013).

Sesquiterpenoids

Sesquiterpenoids are terpenoids with a skeleton of 15 carbons. They occur as hydrocarbons or in oxygenated forms, such as ketones, alcohols, aldehydes, acids, and lactones. Often these oxygen containing sesquiterpenoids are further modified, for example by double bond reduction or to form glycosides and acetyl esters. When a lactone is contained in the sesquiterpene skeleton, the compound's name is given the suffix 'olide'. Sesquiterpene lactones (STLs) are a major class of plant secondary metabolites, which are found in most species of the Asteraceae. Many of these colourless, bitter tasting, lipophilic molecules are the active constituents of a variety of medicinal plants used in traditional medicine (Rodriguez et al., 1976; Zhang et al., 2005). Over 4000 different STLs have been identified so far (de Kraker et al., 2002). Although sesquiterpene lactones generally are present throughout the plant, they are most commonly
General introduction

General introduction

located in glandular trichomes on flowers, leaves and stems (Kelsey & Shafizadeh, 1980; Siegler, 1998). There are five subclasses of STLs categorised according to their carbon skeleton: germacrane, eudesmanolide, elemanolides, guaianolide, and pseudoguaianolide (Rodriguez et al., 1976; Picman, 1986; Fischer, 1990). Germacrane are the most common type. Generally, they contain a trans,trans-cyclodecadiene system, but also a number of cis,cis-germacranolides have been reported. All the sesquiterpene lactones are considered to be derived from a germacrane precursor, likely germacrene A (Figure 1) (de Kraker et al., 1998). The germacrane precursor undergoes cyclisation, ring fusion, and sometimes methyl migration to yield the other skeletal types of STLs.

Sesquiterpene lactones in feverfew

Feverfew (Tanacetum parthenium) is one of the most prominent medicinal species in the Asteraceae family (Bedoya et al., 2008). Feverfew is a daisy-like annual or perennial herb (Figure 2). The term parthenium has been considered to originate from the Greek word parthenios.
meaning ‘virgin’, probably because the herb has been used as an antidote for women’s ailments (Jackson & McDonald, 1986). An alternative explanation comes from a Greek story about an accident where feverfew was used to save the life of someone who had fallen from the Parthenon during its construction in the 5th century BC (Johnson, 1984). The herb has been used since ancient times as a general febrifuge, which means ‘fever reducer’, hence its common name.

Feverfew is native in southern Europe, but it has become naturalized throughout Europe, Asia, North Africa, Australia and North America. During summer, the plant flowers and most of the aerial parts are harvested when they are required for use in herbal medicine. Feverfew has been used for at least two millennia for the treatment of fever, as well as headache, menstrual irregularities, stomach-ache and as an aid for those suffering from arthritis and inflammation (Pareek et al., 2011). Its extracts have been approved in Europe as a herbal drug for the treatment of migraine without prescription. The clinical safety of these extracts has been verified.

More than 30 STLs have been identified in feverfew, including parthenolide, costunolide, 3β-hydroxyparthenolide, 3β-hydroxycostunolide, artecanin, artemorin, balchanin, canin, 10-epicanin, epoxyartemorin, 1β-hydroxyarbusculin, 8α-hydroxyestagiatin, 8β-hydroxy-reynosin, manolialide, reynosin, santamarine, epoxysantamarine, secotanaparthenolide A, secotanaparthenolide B, tanaparthin-α-peroxide, and 3,4β-epoxy-8-deoxycumambrin B (Pareek et al., 2011; Fischedick et al., 2012). These STLs belong to three structural types: eudesmanolides, germacranolides, and guaianolides. The structures of some representative STLs are shown in Figure 3. Parthenolide is the principal bioactive STL component in feverfew (0.1~0.9 %, dry weight) (Bork et al., 1997).

Why do sesquiterpene lactones have such exceptional bioactivities?
Parthenolide exhibits exceptional anti-cancer, anti-inflammatory, and anti-microbial activity (Mathema et al., 2012). It has been reported to induce reactive oxygen species (ROS) exclusively in tumor cells (Zhang et al., 2004; Zunino et al., 2007; Juliana et al., 2010), enhance platelet production and attenuate platelet activation through the inhibitory activity of NF-κB signalling pathway (Sahler et al., 2011), and interfere with microtubule formation by reducing
impaired control of spindle positioning (Fonrose et al., 2007). It has also been reported to induce apoptosis of human acute myelogenous leukemia stem and progenitor cells (Guzman et al., 2005), act as inhibitor of inflammasomes (Juliana et al., 2010), and prevent human lens epithelial cells from oxidative stress-induced apoptosis through inhibition of the activation of inflammation-related genes (Li-Weber et al., 2005). However, despite promising activity, this potent natural product has one major limitation which precludes its further development as a therapeutic agent and that is its poor water-solubility (Sweeney et al., 2005). Hence, there is a great need for derivatives with similar bioactivity but improved water solubility. An analogue of parthenolide, dimethylamino-parthenolide (DMAPT), was shown to retain the biological activity of parthenolide, while exhibiting oral-bioavailability (Guzman et al., 2007). Eight additional, water-soluble, parthenolide analogues were found to exhibit good anti-leukemic activity (Neelakantan et al., 2009).

The biological activity of most STLs is due to alkylation of biological macromolecules by Michael-type additions. The cytotoxic activity of STLs has been attributed to the reaction of the α,β-unsaturated lactone moiety with thiols, such as cysteine residues in proteins, which leads to the disruption of various cysteine-dependent structural enzymes and proteins. In addition, the redox balance in biological tissues is altered by the reaction of the lactone moiety with free intracellular reduced glutathione (Rodriguez et al., 1976).

The α-methylene-γ-lactone group and α,β-unsaturated carbonyl (Figure 1b), as well as conjugated aldehyde groups, are considered to be the reactive groups (Lee et al., 1971; Merfort, 2011). The cytotoxic activity of STLs decreases strongly through reduction of the double bond of the α,β-unsaturated ketone (Lee, 2010). When the α-methylene-γ-lactone moiety is the only alkylating centre, a conjugated ester side chain increases cytotoxicity regardless of its lipophilic properties (Ghantous et al., 2010). These alkylating STLs can react with nucleophiles, which consequently react reversibly with sulphydryl groups in the cell, including those on free cysteine residues (Ghantous et al., 2010). Guaianolide and pseudoguaianolide skeletal types are found to be among the most active STLs through in vitro cytotoxicity assays (Fernandes et al., 2008).

Biosynthesis of sesquiterpene lactones

**MEP and Mevalonate pathway**

The STLs in feverfew are thought to be located in trichomes on the abaxial side of the leaves as well as in the flowers and seeds (Blakeman & Atkinson, 1979). Parthenolide is found in highest concentration in the flowers and fruits. The amounts present in leaves are normally well in excess of 2 mg g⁻¹. Other sesquiterpene lactones are only present in mg kg⁻¹ quantities (Hewlett et al., 1996). Sesquiterpene lactones are derived from the universal C5 precursor isopentenyl diphosphate (IPP) which can be synthesized via two different pathways: the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Figure 4). Both pathways exist in plants but have different subcellular localisation. The MEP pathway enzymes are located in plastids, while the MVA enzymes are found in the cytosol.
In the MEP pathway, glyceraldehyde-3-phosphate and pyruvic acid are transformed to IPP and DMAPP in a ratio of 5:1 through seven enzymes. In the MVA pathway, acetyl-CoA is transformed to IPP through 6 steps, followed by an IPP isomerase (IPPI) that maintains a balance between IPP and DMAPP. Two IPPs and one DMAPP are then converted to farnesyl diphosphate (FPP), the precursor for all sesquiterpenes, through farnesyl diphosphate synthase (FPS). FPP can be converted to sesquiterpenes by sesquiterpene synthases. The sesquiterpenes are often further modified by hydroxylation and/or other oxidation reactions, usually

(Vranová et al., 2013).
mediated by cytochrome P450 enzymes. Cytochrome P450 monooxygenases are heme protein–dependent mixed-function oxidases that utilize NADPH and/or NADH to reductively cleave atmospheric dioxygen to produce a functionalized organic substrate and a molecule of water (Schuler & Werck-Reichhart, 2003). In many cases, the result of a cytochrome P450 catalysed reaction is insertion of oxygen (hydroxylation), but in a number of cases, P450s mediate dehydrogenation, isomerization, dimerization, carbon–carbon bond cleavage, reduction, as well as N-, O- and S-dealkylations, sulphoxidations, epoxidations, deaminations, and desulphurations (Weitzel & Simonsen, 2013). The term ‘P450’ originated from their shared ability to display a typical absorption peak at 450 nm when carbon monoxide is bound to the reduced form of the enzyme (Omura & Sato, 1964).

Several P450s have been reported to be involved in sesquiterpene biosynthesis. CYP71AV1, amorpha-4,11-diene oxidase of *Artemisia annua*, was found to catalyse the conversion of amorpha-4,11-diene into artemisinic acid in three consecutive hydroxylation steps (Ro et al., 2006). By co-expressing amorpha-4,11-diene synthase (ADS) and CYP71AV1 in *Nicotiana benthamiana*, artemisinic acid was produced, which was partially conjugated to form a di-glucoside (van Herpen et al., 2010). A cytochrome P450 enzyme, CYP71D20, was reported to catalyse the regio- and stereospecific insertion of two hydroxyl moieties into the bicyclic sesquiterpene 5-epiaristolochene (Takahashi et al., 2005). Another P450 CYP71D55, premnaspirodiene oxygenase from *Hyoscyamus muticus* (Solanaceae), was found to be capable of hydroxylating valencene, 5-*epi*-eremophilene, and 5-*epi*-aristolochene in addition to its native substrate premnaspirodiene (Takahashi et al., 2007).

**Biosynthesis of sesquiterpene lactones in feverfew**

The biosynthesis pathway of parthenolide, the most prominent compound in feverfew, was postulated to proceed as shown in Figure 5 when the work of this PhD thesis was initiated.

![Diagram of biosynthetic pathway for parthenolide](image)
Chapter 1

First, germacrene A synthase (GAS) catalyses the cyclization of FPP to germacrene A (de Kraker et al., 1998). GAS has been cloned from chicory (Bouwmeester et al., 2002), lettuce (Bennett et al., 2002) and A. annua (Bertea et al., 2006). In a number of additional steps, germacrene A is then oxidized to germacranoic carboxylic acid (de Kraker et al., 2001). This sequence of reactions is very similar to the biosynthesis of artemisinic acid, hydroxylation of amorphadiene to artemisinic alcohol, artemisinic aldehyde, and then artemisinic acid by CYP71AV1 (Ro et al., 2006; Teoh et al., 2006; Zhang et al., 2008). Thus, it is likely that the oxidation of germacrene A to germacranoic acid in Tanacetum parthenium is performed by a similar P450 monooxygenase. The subsequent hydroxylation of germacranoic acid, required to form costunolide (Figure 5), is likely also catalysed by a P450 monooxygenase. Parthenolide is produced through epoxidation of costunolide at the C4-C5 double bond of costunolide, possibly by another P450 monooxygenase.

Thesis outline

This thesis is part of a project within the “Food, Agriculture and Fisheries, and Biotechnology” theme of the Seventh Framework Programme of the European Commission project entitled ‘Plant Terpenoids for Human Health: a chemical and genomic approach to identify and produce bioactive compounds’ (acronym: TERPMED).

As part of the TERPMED project, this thesis aimed to elucidate the biosynthetic pathway of parthenolide and other sesquiterpene lactones in feverfew (Tanacetum parthenium), and to improve parthenolide production - and produce novel STLs through combinatorial biochemistry - through genetic engineering of plants. Enzymes involved in the different steps of the biosynthesis of these bioactive STLs are characterized at molecular and biochemical level and the corresponding genes used for metabolic engineering.

The present chapter, Chapter 1, gives an introduction to the sesquiterpene lactones in general and parthenolide and feverfew in particular, the medicinal properties of the STLs and what is known and postulated about their biosynthesis.

Chapter 2 describes work on the localisation of parthenolide biosynthesis to provide the basis for the subsequent characterisation of the pathway. The work shows that parthenolide is produced and localized mainly in trichomes on feverfew flowers. Isolated trichomes are subsequently used to isolate the sesquiterpene synthase encoding the first dedicated step in parthenolide biosynthesis, germacrene A synthase (TpGAS).

During my work on TpGAS, germacrene A oxidase (GAO), the gene that likely encodes the next step in the pathway, was reported from lettuce (Nguyen et al., 2010) and chicory (Cichorium intybus L.) (Cankar et al., 2011). As the feverfew trichome transcriptomics data were not yet available, in Chapter 3 I used a chicory root 454 cDNA library to identify the gene that is responsible for the next step of the pathway, costunolide synthase, CiCOS. I used this gene, in combination with TpGAS and CiGAO for reconstitution of the costunolide biosynthesis pathway through transient expression in N. benthamiana. This resulted in formation of the expected product, costunolide, but also new costunolide conjugates, which have never been...
reported before in plants.

Chapter 4 describes the isolation and characterization of TpGAO, TpCOS and TpPTS using feverfew trichome cDNA sequences. Together with TpGAS that I characterised in Chapter 2, these are all the pathways genes required for biosynthesis of parthenolide from farnesyl diphosphate. I used these genes to reconstitute the entire parthenolide biosynthesis pathway in N. benthamiana. Parthenolide was formed, mostly as conjugates, and these are tested for bioactivity against cancer cells.

The feverfew trichome EST data yielded a wealth of other interesting gene candidates and Chapter 5 describes the isolation and characterization of a kauniolide synthase (TpKS). TpKS is also a cytochrome P450 and catalyses ring closure in the germacranolide costunolide to form a guaianolide STL, kauniolide.

Chapter 6 integrates and discusses all the results from this thesis in a broader perspective. Some future perspectives are also discussed.

References


Chapter 1

43(10): 2197-2205.


Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J,


Chapter 2

Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.)

Mohammad Majdi¹, Qing Liu¹, Ghasem Karimzadeh, Mohammad Ali Malboobi, Jules Beekwilder, Katarina Canka, Ric de Vos, Sladjana Todorović, Ana Simonović and Harro Bouwmeester

¹ These two authors contributed equally to this work.

Feverfew (*Tanacetum parthenium*) is a perennial medicinal herb and is a rich source of sesquiterpene lactones. Parthenolide is the main sesquiterpene lactone in feverfew and has attracted attention because of its medicinal potential for treatment of migraine and cancer. In the present work the parthenolide content in different tissues and developmental stages of feverfew was analyzed to study the timing and localization of parthenolide biosynthesis. The strongest accumulating tissue was subsequently used to isolate sesquiterpene synthases with the goal to isolate the gene encoding the first dedicated step in parthenolide biosynthesis. This led to the isolation and characterization of a germacrene A synthase (TpGAS) and an (E)-β-caryophyllene synthase (TpCarS). Transcript level patterns of both sesquiterpene synthases were analyzed in different tissues and glandular trichomes. Although TpGAS was expressed in all aerial tissues, it was expressed higher in tissues with higher concentrations of parthenolide and particularly in the biosynthetically active stages of flower development. The high expression of TpGAS in glandular trichomes which also contain the highest concentration of parthenolide, suggests that glandular trichomes are the organ where parthenolide biosynthesis and accumulation occur.
Introduction
Terpenoids are the largest class of plant secondary metabolites with over 20,000 compounds known (Davis & Croteau, 2000). Sesquiterpene lactones are a major class of terpenoids that arise from the assembly of a 15-carbon skeleton into bisabolane, cuparane, cadinane, humulane and germacrane backbones and the addition of functional groups to these backbones (Chappell et al., 2010; Chappell, 2010). They are particularly abundant in the Asteracea family and are biologically significant because of key roles in plant environment interaction, for example as chemical defense compounds against pathogens and insects (Cheng et al., 2007; Hristozov et al., 2007; Chappell et al., 2010; Chappell, 2010). Sesquiterpenes have various biological activities and uses for humans, including antimicrobial, anticancer and anti-inflammatory properties (Min Li-Weber, 2002); (Elisa Saranitzky, 2009; Trusheva et al., 2010).

Feverfew (2n=2x=18) is a perennial herb belonging to the Asteracea that has been used traditionally as a fever-reducer (Palevitch et al., 1997). The most important constituent (up to 85% of total sesquiterpenes) among the 30 sesquiterpene lactones identified in feverfew is parthenolide (Williams et al., 1995; Brown et al., 1997; Cretnik et al., 2005). Parthenolide (Fig.1) is a germacronolide lactone which recently raised quite a lot of attention because clinical trials showed medicinal value and pharmacological activities especially as a migraine prophylaxis agent and also for treatment of cancer (Knight, 1995; Palevitch et al., 1997; Vogler et al., 1998; Min Li-Weber, 2002; Pfaffenrath et al., 2002; Lesiak K, 2010).

Considering these promising applications, the need for parthenolide may be increasing in the near future. In order to increase the availability of parthenolide several approaches can be envisaged, such as heterologous engineering into other plant species or microbial platforms. In order to be able to do so knowledge about the biosynthetic pathway of parthenolide is mandatory.

Sesquiterpene lactones are mostly derived from the mevalonic acid (MVA) pathway (van Klink et al., 2003). They are classified on the basis of their carbon skeletons as germacrano-lides, guaianolides, pseudoguanolides and eudesmanolides (de Kraker et al., 1998). Parthenolide has a germacrano-lide backbone structure and costunolide has been proposed to be the common precursor of all germacrone-derived sesquiterpene lactones (de Kraker et al., 2002) (Fig. 1). The first committed step in the biosynthesis of costunolide is the cyclization of farnesyl diphosphate (FDP) to germacrene A by a germacrene A synthase (de Kraker et al., 1998; Bouwmeester et al., 2002). Recently two research groups showed that the cytochrome P450, germacrene A oxidase, is responsible for the conversion of germacrene A to germacrenoic acid (germacra-1(10),4,11(13)-trien-12oic acid) (Fig. 1) (Cankar et al.; Nguyen et al., 2010).

To verify this pathway in feverfew and to identify the genes missing in the parthenolide pathway, it is crucial to study the spatial and temporal regulation of the pathway. The biosynthesis and accumulation of secondary metabolites in plants is closely associated with the transcription level of the relevant genes in organs, tissues, as well as in different developmental stages. Biosynthesis and accumulation of metabolites may occur at the same place
(cells, tissues) or be spatially separated. For example, biosynthesis as well as accumulation of artemisinin have been observed in the glandular trichomes of Artemisia annua (Bertea et al., 2006; Teoh et al., 2006). In the Asteraceae, several different specialized accumulation/biosynthesis structures for secondary metabolites have been reported e.g cavities in Solidago canadensis (Curtis & Lersten, 1990; Cury & Appezzo-da-Gloria, 2009), ducts in Ambrosia trifida (Curtis, 1988), laticifers in Lactuca sativa (Esau, 1965; Sessa et al., 2000) and glandular trichomes in Artemisia annua (Kelsey, 1980).

Glandular trichomes are specialized structures consisting of usually 6 to 10 cells, that produce secondary metabolites which are stored in a sub-cuticular cavity on top of the trichome, such that the phytotoxic secondary metabolites are stored away from primary metabolism (Wagner, 1991). Glandular trichomes have been used in several species of the Asteracea to elucidate the biosynthesis of terpenes, using the fact that they are the most active or even exclusive organ for biosynthesis e.g in A. annua (Bertea et al., 2006; Covello et al., 2007; Olsson et al., 2009; Lies Maes, 2010) and Helianthus annuus (Gopfert et al., 2009).

The objectives of the present investigation were to determine the developmental and spatial regulation of parthenolide biosynthesis in feverfew. This detailed analysis should provide the crucial knowledge necessary for further characterization of the parthenolide biosynthetic pathway, for example to identify the right tissue and developmental stage to create an EST library that can be used for gene discovery. In addition, we set out to identify and characterize the sesquiterpene synthase that catalyses the first dedicated step in parthenolide biosynthesis.
Results

Parthenolide concentrations in feverfew

To see which organs of feverfew contain parthenolide, several different tissues including flower, leave, stem and root were analyzed by LC-QTOF-MS. There were significant differences between different tissues in term of parthenolide concentration (P < 0.01) (Fig. 2A). The highest amount of parthenolide was observed in flowers, followed by leaves and stems. No parthenolide was detected in feverfew roots.

Considering the high parthenolide concentration in flowers compared to the other tissues we further investigated which part(s) of the flower contains the highest amount of parthenolide. The parthenolide concentration differed between the two main parts of the flower, with the highest amount of parthenolide being observed in the disc florets when compared with the ray florets (P < 0.01) (Fig. 2B). Scanning electron microscopy (SEM) showed that...
the ray florets do not contain trichomes (Fig. 3B) in contrast to the disc florets (Fig. 3C). Disc florets consist of two parts, the upper (corolla) and lower part (ovary), of which the latter contained more parthenolide (Fig. 2B). SEM showed that the trichome density on disc florets is higher on the ovary than the corolla (Fig. 2C). When comparing the trichome density in several other plant organs it showed to be highest on the disc florets (Fig. 3C), followed by leaves (Fig. 3A) while stems contain an even lower density of trichomes (data not shown). To elucidate whether parthenolide is present in the glandular trichomes, chloroform dipping was used, which extracts the glandular trichomes but not - or to a much lesser extent - the remainder of the plant tissue (Duke et al., 1994). Indeed, scanning electron microscopy showed that a chloroform dip of 30 sec extracts most of the trichomes on a disc floret without inducing visible damage to the epidermal cells (Fig. 4). Chloroform-dipped disc florets contained significantly (P<0.01) less parthenolide than non-dipped disc florets. More than 80% of the parthenolide was extracted from disc florets by the 30-sec chloroform dip.

Isolation of sesquiterpene synthases

The degenerate primer approach followed by 3’- and 5’-rapid amplification of cDNA ends (RACE), resulted in the isolation of two full-length feverfew sesquiterpene synthase cDNAs, TpGAS and TpCarS. The TpGAS ORF showed a length of 1,677 bp encoding 559 amino acids. A Blast search in GenBank revealed that TpGAS has high homology with other GASs of the Asteraceae, such as germacrene A synthase short form from C. intybus (highest identity, 87%), germacrene A synthase LTC1 from Lactuca sativa, germacrene A synthase 1 from H. annuus, germacrene A synthase from Crepideastrum sonchifolium, germacrene A synthase from A. annua and germacrene A synthase long form from C. intybus (lowest identity, 73%). The molecular weight of the TpGAS protein was calculated to be 64.5 kDa with an isoelectric point (pI) of 5.03. The deduced TpGAS protein alignment with other germacreneA synthases is shown in Fig. 5A. Phylogenetic analysis using ClustalW (http://www.ebi.ac.uk/Tools/clusterw2/index.html) showed that TpGAS and A. annua GAS are in the same cluster and have the shortest evolutionary distance from each other. In spite of the higher similarity between TpGAS and short form GAS from chicory they have been categorized in different clusters (Fig. 5B). The TpCarS ORF has a length of 1,647 bp encoding for 549 amino acids. Blast search in GenBank showed high identity (93%) with (E)-β-caryophyllene synthase from A. annua and 61% with GAS from Solidago canadensis. The molecular weight of the predicted TpCarS protein was calculated to be 63.5 kDa with isoelectric point (pI) of 5.48. A protein sequence alignment and a phylogenetic tree of TpCarS with a number of other sesquiterpene synthases are shown in Fig. 5B.
Fig. 5. (A) Multiple sequence alignment of TpGAS with other germacrene A synthases (GAS) using ClustalW (Bio Edit software) and (B) Phylogenetic tree analysis of TpGAS with other GASs from Asteraceae family; Aa [Artemisia annua] ABE03980.1; Ci.SF [Cichorium intybus short form] AF498000.1; Ci.LF [Cichorium intybus long form] AF497999.1; Ha [Helianthus annuus] AAY41421.2; Ls, [Lactuca sativa] AF489964.1; Sc [Solidago canadensis] CAC36896.1; Cs [Crepidiastrum sonchifolium] ABB00361.1; Pc [Pogostemon cablin] AAS86321.1; 5-Epi-aristolochene synthase [Nicotiana tabacum], AAA19216.
are shown in Fig. S2 and S3.

**Functional characterization of TpGAS and TpCarS**

TpGAS and TpCarS were functionally characterized by cloning of the ORF in pACYCDuet™-1 vector and heterologous expression in E. coli. An enzyme assay with the recombinant TpGAS in the presence of FDP, followed by GC–MS analysis showed a single peak which was
identified as β-elemene by comparing its spectrum and retention index with library data (Fig. 6 A,B). β-Elemene is the Cope rearrangement product of germacrene A that is formed upon injection into the hot injection port of the GC-MS (de Kraker et al., 1998). A comparable enzyme assay with TpCarS generated mostly β-caryophyllene and a trace of α-humulene (also called α-caryophyllene) (Fig. 6 C, D and E).

Expression of TpGAS and TpCarS

In order to assess the expression of TpGAS and TpCarS in different organs of feverfew, qRT-PCR experiments were carried out using cDNA templates from whole flowers, leaves, stems and roots as well as glandular tichomes. Comparison of relative gene expression (RGE) patterns for TpGAS and TpCarS revealed that the expression of TpCarS was higher in leaves than in flowers and stems whereas the expression of TpGAS in flowers was more than 4-fold higher than in leaves and 25-fold higher than in stems (Fig. 7). Both genes were not expressed in roots. The highest expression of TpGAS was observed in trichomes (RGE of more than 100), which is at least 10–fold higher than the RGE of TpCarS in trichomes (Fig. 8).

Analysis of variance showed significant differences in parthenolide concentration of ovaries of different developmental stages of flowers (P<0.001) (Fig. 9A). Mean comparison with Duncan’s test showed that the parthenolide concentration significantly increased with flower developmental stage, reaching a peak at stage 5 (2-fold higher than in stage 2) and subsequently declining gradually during stage 6 and 7 (Fig. 9A). The parthenolide concentration in stage 7 was still higher than in stage 2 (Fig. 9A). Analysis of variance showed that also for TpGAS expression significant differences existed in the RGE in ovaries of different developmental stages (P<0.001) (Fig. 9B). RGE increased significantly with flower developmental stage, reaching a peak at stage 3 and 4 (about 2-fold increase compared with stage 2) then strongly decreasing in stage 5 (>4-fold lower than in stages 3 and 4). The lowest RGEs were observed for stage 6 and 7 ovaries. TpCarS displayed a stable, low expression pattern (RGE < 1) in all different developmental stages (Fig. 9B).
Chapter 2

Discussion

Parthenolide is mainly accumulating in glandular trichomes

Our work show that the parthenolide present in feverfew is particularly concentrated in the flowers, which is in accordance with the results of others (Awang, 1991; Majdi et al., 2010). Within the flowers, the disc florets contain a higher concentration of parthenolide than the ray florets which coincide with a higher glandular trichome density on the disc florets. Within the disc florets the ovary contains a higher concentration of parthenolide than the corolla and also here this coincides with a higher glandular trichome density on the former (Fig. 2B-C). All this suggests that the glandular trichomes are likely the place where parthenolide accumulation occurs. The mild chloroform dipping extraction method, which extracted about 80% of parthenolide from the disc florets without visible damage to the epidermis, further proved the presence of parthenolide in the glandular trichomes (Fig.4). The presence of sesquiterpene lactones has usually been associated with other structures such as cavities, ducts and laticifers (Esau, 1965; Kelsey, 1980; Curtis, 1988; Curtis & Lersten, 1990; Cury & Appelazzi-da-Gloria, 2009). On the other hand, the presence of sesquiterpenes, including oxidized derivatives, in glandular trichomes has been reported in other plants as well such as A. annua (Bertea et al., 2006), Solanum habrochaties (Gianfagna Tj, 1992), Mentha x piperita (McCaskill & Croteau, 1995) and Solanum lycopersicum (Schilmiller et al., 2010).

TpGAS expression reveals glandular trichomes as the site of parthenolide biosynthesis

As discussed above, the parthenolide concentration is associated with the distribution of glandular trichomes in several different organs. The highest density of glandular trichomes occurred in flowers (ovaries) followed by leaves and stems and the same pattern was observed for the parthenolide concentration (Fig. 2 and 3) as well as the expression of TpGAS (Fig. 7). TpGAS is encoding the enzyme that highly likely catalyses the first step in parthenolide biosynthesis, germacrene A synthase (Fig. 1). The expression of TpCarS - another feverfew sesquiterpene synthase – showed a very different profile and was for example highest in leaves (Fig. 7). Also parthenolide concentration and TpGAS expression in trichomes - isolated from different developmental stages of ovaries - correlated, in contrast to TpCarS of which expression did not correlate with parthenolide concentration (Fig. 9). The association of parthenolide with glandular trichomes is further substantiated by the fact that parthenolide was not detected in roots which do not have glandular trichomes (Fig. 2 and 7). Also closely related plant species belonging to the Asteraceae such as A. annua make and store sesquiterpene lactones in the glandular trichomes (Bertea et al., 2006; Covello et al., 2007). Consistent with our results, the relationship between the density of glandular trichomes and the essential oil content has also been reported in other plant species (Maffei et al., 1989; McCaskill & Croteau, 1999; Bertea et al., 2006). In feverfew, TpGAS is only expressed in the aerial parts in which also parthenolide exists, while the expression of GAS in roots of other Asteracea – that produce sesquiterpene lactones in their roots - has been reported e.g C. intybus and H. annuus (Bouwmeester et al., 2002; Gopfert et al., 2009). In the latter case, not only has different
Biosynthesis and localization of parthenolide

regulation of transcription of biosynthetic genes allowed for biosynthesis of sesquiterpene lactones in the roots but also storage of the sesquiterpene lactones was adapted to accommodate for the absence of trichomes on the roots (Gopfert et al., 2009).

The much higher expression of \( TpGAS \) in trichomes when compared with other tissues (Fig. 7) shows that \( TpGAS \) is a glandular trichome specific gene. In consistence with our work there have been many reports on different plant species which have demonstrated a strong correlation between terpene amount (or emission) and the level of the corresponding mRNA, indicating that terpenoid biosynthesis is mainly regulated at the transcript level (Nagegowda, 2010). Also gene expression analysis in different cell types of glandular trichomes of \( A. \ \textit{annua} \) including apical, sub-apical and mesophyl cells showed that the expression of three enzymes specific to the artemisinin biosynthetic pathway are active only in the apical cells (Olsson et al., 2009). These apical cells are likely the ones in which artemisinin biosynthesis occurs, close to the subcuticular cavity where artemisinin and/or its precursors are stored. In contrast, FDP synthase which is not specific only for artemisinin biosynthesis was expressed in all the glandular trichomes cell types (Olsson et al., 2009). All this suggests that also for uncovering the regulation and site of parthenolide biosynthesis, analysis of \( TpGAS \) expression and its localisation is a reliable approach.

**Developmental changes in \( TpGAS \) expression and parthenolide concentration**

Analysis of parthenolide concentration and \( TpGAS \) expression in different developmental stages of flowers showed that parthenolide accumulation and \( TpGAS \) expression closely correlate. Although \( TpCarS \) gene expression was observed in the trichomes in several flower developmental stages, both the level of transcript and the pattern of transcript change do not show any relationship with parthenolide biosynthesis, excluding that \( TpCarS \) is involved in parthenolide biosynthesis. Parthenolide accumulation is developmentally regulated and displays three distinct phases (Fig. 9). The first phase (stage 1 to stage 4) can be considered as “parthenolide accumulation phase” in which there is a direct relationship between the increase in \( TpGAS \) expression and the increase in parthenolide concentration (Fig. 9). The second phase (stage 5) can be considered as “parthenolide saturation phase” in which the parthenolide concentration reaches its maximum while the expression of \( TpGAS \) is down-regulated compared with the previous phase. This may indicate a feed back regulatory mechanism or a developmental program controlling parthenolide accumulation by down-regulating the expression of the gene(s) in the biosynthetic pathway of parthenolide, in this case \( TpGAS \). Also feedback inhibition of artemisinin biosynthesis by repression of ADS and CYP71AV1 expression by artemisinic acid and artemisinin has recently been suggested (Arsenault et al., 2010). Considering the relationship between flower development and glandular trichome development (Werker, 2000; Gopfert et al., 2005), this phase may represent physiological maturity of the glandular trichomes which coincides with the end of cell expansion in floret development. The third phase (stage 6-7) can be considered as “parthenolide reduction phase”, in which a decrease in parthenolide concentration was observed, along with the strongly reduced expression of \( TpGAS \) (Fig. 9). The \( TpGAS \) expression pattern resembles that of the \( H. \ \textit{annuus} \)
germacrene A synthase. In this species, up-regulation of the germacrene A synthase during the secretory stage and down-regulation during the post-active secretory phase has been reported confirming developmental stage-specific expression in glandular trichomes (Gopfert et al., 2009). The high expression of $TpGAS$ during the parthenolide accumulation phase and continued expression of $TpGAS$ during the parthenolide decrease phase show the necessity of active and continuous transcription of $TpGAS$ (and likely other genes downstream of $TpGAS$) for parthenolide accumulation. This is consistent with the high positive correlation between the expression of $ADS$ and $CYP71AV1$ and the accumulation of the artemisinin precursor dihydroartemisinic acid in A. annua (Arsenault et al., 2010). Also in peppermint a high rate of monoterpenic biosynthesis and high enzyme activities were observed during early leaf development (Gershenzon, 1994). The strong developmental changes in parthenolide concentration that we find (Fig. 9) are supported by the literature. For example, Omidbaigi et al. (2007) reported developmental changes in feverfew essential oil content and composition in flowers at various harvesting times. Developmental regulation appears to be an important factor in parthenolide biosynthesis and this is consistent with other reports on terpene biosynthesis (McConkey et al., 2000; Lommen et al., 2006; Kim et al., 2008; Arsenault et al., 2010). The decrease in parthenolide concentration during the reduction phase is more difficult to explain. Volatilization and leaching have been postulated as possible mechanisms that cause a loss of sesquiterpene lactones (Gershenzon, 1994). However, parthenolide is not volatile and leaching is not likely to occur in a greenhouse or climate room. Possibly, parthenolide decrease in the reduction phase is caused by a biodegradation mechanism or further conversion to another compound(s). Further conversion of parthenolide to epoxyparthenolide and hydroxyparthenolide in Anvillea radiata a member of the Asteracea has been reported (Hassany et al., 2004). Biodegradation or further conversion could also explain the diurnal changes in parthenolide content (decrease during the night) in feverfew leaves (Fonseca et al., 2005; Fonseca et al., 2006). The decrease in parthenolide concentration in phase 3 could also indicate trichome collapse, as has been reported for A. annua (Lommen et al., 2006; Arsenault et al., 2010). Lommen et al. (2006) have proposed that trichome collapse in A. annua is responsible for the decrease in artemisinin content during the later stages of leaf development.

Conclusion

Cloning of $TpGAS$ and the study of its transcript patterns in different tissues and developmental stages of flowers in feverfew shows that the expression of $TpGAS$ is closely associated with parthenolide biosynthesis. The high expression of $TpGAS$ in glandular trichomes and the localization of parthenolide in the trichomes show that the trichomes are the specific site of parthenolide biosynthesis and accumulation. Hence, isolation of feverfew glandular trichomes will be a powerful tool to generate an EST library, which can be used for high throughput transcript sequencing and identification of additional genes involved in the biosynthetic pathway of parthenolide. This approach has been successful in other plant species and has lead to the identification of genes involved in the biosynthesis of several secondary metabolites (Bertea et al., 2006; Covello et al., 2007; Nagel et al., 2008). Transcript profiling
in combination with metabolite profiling (untargeted large scale metabolomics) of glandular trichomes isolated from flowers in different developmental stages could also provide insight into the regulation of accumulation and loss of parthenolide in *Tanacetum parthenium*.

**Materials and methods**

**Plant growth condition**

*Tanacetum parthenium* L. plants were grown under controlled conditions in a climate room with a 16h day length supplied with a photosynthetic photon flux density of 320 μmol m⁻² s⁻¹ with day/night temperature of 24/18 °C. Samples for phytochemical and gene expression analysis were immediately frozen in liquid nitrogen and then stored at -80 °C until further use.

Seven developmental stages of feverfew flower development were defined based on a scale developed for *Chrysanthemum cinerariaefolium* (Head, 1966). Stage 1 (well developed closed buds), stage 2 (ray florets vertical), stage 3 (ray florets horizontal and first row of disc florets open), stage 4 (approximately three rows of disc florets open), stage 5 (all disc florets open, fully mature), stage 6 (early senescence; color of disc florets diminishing but ray florets still intact), stage 7 (late senescence; disc florets dried out) (Head, 1966). Ovaries were isolated from all these stages except stage 1.

**Parthenolide measurement**

Plant materials (flower, leave, stem, root and ovary) were ground to a fine powder using a mortar and pestle in liquid nitrogen. Non-volatile compounds were analyzed using a protocol for untargeted metabolomics of plant tissues (De Vos et al., 2007).

Three biological replicates were used for each sample. In brief, 100 mg of sample for each replicate (leaves from one month old plants in the vegetative stage, whole stage 3 flowers, ovaries from different flower developmental stages, stems and roots) were ground in liquid nitrogen and extracted with 0.3 mL methanol:formic acid (1000:1, v/v). The extracts were briefly vortexed and then sonicated for 15 min. Subsequently, the extracts were centrifuged and filtered through a 0.2 μm inorganic membrane filter (RC4, Sartorius, Germany). LC-PDA-MS analysis was performed using a Waters Alliance 2795 HPLC connected to a Waters 2996 PDA detector and a QTOF Ultima V4.00.00 mass spectrometer (Waters, MS technologies, UK) operating in positive ionization mode. The column used was an analytical column (2.0 × 150 mm; Phenomenex, USA) with a C18 pre-column (2.0 × 4 mm; Phenomenex, USA). Degassed eluent A (ultra pure water:formic acid; 1000:1, v/v) and eluent B (acetonitril:formic acid; 1000:1, v/v) were pumped at 0.19 mL min⁻¹ into the HPLC system. The gradient started at 5% B and increased linearly to 35% B in 45 min. Then the column was washed and equilibrated for 15 min before the next injection.

An authentic standard of parthenolide (Sigma-Aldrich, USA) was used to make a calibration curve by injecting different concentrations of parthenolide for quantification.

**Chloroform dipping**
To 100 mg of ovaries from feverfew in a 2 mL Eppendorf vial, 1 mL of chloroform was added followed by vortexing for 30 seconds and short centrifugation. The chloroform phase was removed and the chloroform evaporated in a fume hood for 3 minutes. Then the parthenolide content of dipped and non-dipped ovaries was measured by LC-MS as described above.

Isolation of glandular trichomes and RNA extraction

Ovaries (about 300 flowers) were separated from the rest of the flower for trichome isolation. About 3 gram of ovaries were placed in a 50 mL pre-cooled Greiner tube and 10 mL liquid nitrogen was added. The tube was covered using a 20 μm mesh followed by vortexing for 1 min and addition of another 10 mL liquid nitrogen. The vortexing was repeated 6 to 10 times. The isolated trichomes were separated from the other cells and debris by passing through a 120 μm mesh and collection on a 20 μm mesh. RNA was isolated from the glandular trichomes using the Spectrum™ Plant Total RNA isolation kit (Sigma, USA).

RNA extraction and cDNA synthesis for sesquiterpene synthases

Flowers of Tanacetum parthenium were used for RNA extraction. Tripure isolation reagent (Roche, Mannheim, Germany) was used for extraction of total RNA according to the manufacturer’s instructions with minor modifications. Isolated RNA was subsequently treated with DNase I (Invitrogen, USA) and then purified through a silica column using the RNeasy RNA clean up kit (Qiagen #74104, USA). The quantity and quality of RNA was determined using a NanoDrop (NanoDrop Technologies, USA) and agarose gel electrophoresis. The reverse transcription reaction was carried out with the Taqman Reverse Transcription Reagent kit (Applied Biosystems, USA) using 1μg of total RNA and 50μM oligo dT according to the manufacturer’s instructions.

To clone sesquiterpene synthases, a homology-based PCR cloning strategy was used. Sequence alignments of published Asteracea sesquiterpene synthase cDNAs in public databases were used to design degenerate oligonucleotide primers. Primer pairs were designed corresponding to the conserved regions, avoiding degeneracy at the 3’ end: forward primer 5’-TATWCNGTNCAYCGTCTYGG 3’ and reverse primer 5’- AYYKCATATCRTTCCACCAC 3’. The temperature program for PCR was initial denaturation at 94 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min and the final elongation at 72 °C for 5 min.

A second set of degenerate primers (set 2 degenerate primers) consisted of: forward primer 5’-GAY GAR AAY GGI AAR TTY AAR GA-3’ and reverse primer 5’- CCR TAI GCR TCR AAI GTR TCR TC-3’ (Wallaart et al., 2001). The temperature program for PCR was initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 48 °C for 1 min, 72 °C for 1 min and the final elongation at 72 °C for 5 min. The resulting purified fragments were cloned into pGEM-T Easy Vector (Promega, USA) and sequenced.

cDNA extension using 5’ and 3’-RACE

5’-RACE and 3’-RACE were performed using the SMART-RACE cDNA amplification kit (Clontech, USA) according to the manufacturer’s protocol. The first strand cDNA for 5’ and
3’ RACE were synthesized from 1 µg of total RNA extracted from ovaries according to the protocol (Clontech, USA). According to the partial sequence of the TpGAS fragment (550 bp) amplified using degenerate primers set 1, high annealing temperature (70 °C) primers were designed for both 5’ and 3’ RACE. Nested primers were designed according to the protocol suggestions. The Gsp1-GAS (Gene specific primer 5’) and NGsp1-GAS (Nested gene specific primer 5’) for 5’-RACE were 5’ GCATCCCTTGATGGAATGGTCTCATT -3’ and 5’ CCCAACATACCCCTCACAATCG -3’, respectively. The Gsp2-GAS (Gene specific primer 3’) and NGsp2-GAS (Nested gene specific primer 3’) for 3’-RACE were 5’-GTGATGTGAGGGGTATGTTGGGCTTA-3’ and 5’ TGCAACAGTTGAGAATAAGAGG -3’, respectively.

Two additional sesquiterpene synthase like sequences were generated using the second set of degenerate primers. For one of them the 5’ and 3’ RACE did not produce the expected cDNA fragments, while the other did. The Gsp1-CarS (Gene specific primer 5’) and NGsp1-CarS (Nested gene specific primer 5’) for 5’-RACE were 5’CTTCGCCCTTGACCCCCCAAGTATG-3’ and 5’ CAAGCAGGGCCTTTCTACATCGTTGGT-3’, respectively. The Gsp2-CarS (Gene specific primer 3’) and NGsp2-CarS (Nested gene specific primer 3’) for 3’-RACE were 5’ GGCCTGCTTGAGCTGTATGAGGCAAC -3’ and 5’ TTCGCTGGCCACTGTTCTTGACGATA -3’, respectively. 5’-RACE and 3’-RACE products were cloned into the pGEM-T easy vector (Promega) and sequenced. Assembly of the sequences from the 3’ RACE and 5’ RACE products using SeqMan software yielded the full-length cDNA sequences of TpGAS and TpCarS.

**Heterologous expression of TpGAS and TpCarS in Escherichia coli**

According to the 5’ and 3’ ends new primers were designed introducing BamHI and NotI restriction sites to clone the fragment into pACYC-DUET-1 (Novagen). A full length of TpGAS was amplified using 5’ Race cDNA as a template, using primers forward 5’-ACTACGGATCCGGCAGCGGTTAACAGCTTAC-3’ and reverse 5’-ACCACGCGGCCGCTTACAGCTTACGG-GTAGAGAATCCCAA-3’ (restriction sites for BamHI and NotI underlined). cDNA amplification for the construction of expression plasmids was performed using the Phusion DNA polymerase. The full length of TpCarS was also cloned into pACYC-DUET-1 (Novagen) using primers forward 5’ ACTACGGATCCGCTGCTAAGAGAAAGAAAGTA-3’ and reverse 5’ ACCACGCGGCCCCTTATATAGGTATAGGATGAACGAG-3’. For functional characterization the two sesquiterpene synthases ORFs (cloned into the pACYC-DUET-1 (Novagen) expression vector with an aminoterminal histidine tag) were expressed in Escherichia coli BL-21 under an isopropyl-thio-β-galactoside and arabinose inducible promoter. For the functional assay, 1mL of transformed bacteria was cultured overnight at 250 rpm at 37 °C in LB supplemented with chloramphenicol. Subsequently, 500 µL of the bacteria were added to 50 ml 2xYT supplemented with chloramphenicol for 1.5 h on shaker incubator at 250 rpm and 37 °C. Then 50 µL 20% L-arabinose was added to induce the bacteria and the culture was incubated overnight on a shaker at 18 °C at 250 rpm. The bacteria were centrifuged for 10 min at 2500 rpm. The resulting pellet was resuspend in 1 mL buffer A (50 mM Tris, pH 8; 300 mM
NaCl, 1.4 mM β-mercaptoethanol) followed by sonication on ice. The sample was centrifuged for 10 min at 13,000xg at 4 °C and the supernatant transferred to a Qiagen Nickel column which was then centrifuged for 2 min at 3000 rpm. The column was washed two times with 600 μL buffer A and then eluted with 200 μL buffer A containing 175 mM imidazole. 100 μL of this purified enzyme was diluted into 800 μL assay buffer (15 mM MOPSO, pH 7.5; 12.5% (v/v) glycerol, 1 mM ascorbic acid, 0.001% (v/v) Tween-20, 1 mM MgCl2, 2 mM DTT). To the assay a final concentration of 50 μM farnesyl diphosphate was added after which the assay was overlayed with 1mL of pentane and incubated at 30 °C with mild agitation for 2h. Subsequently, the pentane phase was collected and the assay extracted with 1 mL ethylacetate. The ethylacetate and pentane phases were combined, centrifuged at 1200xg and then dried using anhydrous Na2SO4 packed in a small glass column and used for GC-MS analysis.

**GC-MS analysis**

Volatile produced in the enzyme assays were analyzed by an Agilent Technologies 7890A GC, equipped with a 5975C inert MSD with Triple Axis Detector using helium as carrier gas at a flow rate of 1ml.min⁻¹. The injector was used in splitless mode with the inlet temperature set to 250 °C. The initial temperature of 45 °C was gradually increased after 1 min to 300 °C by a ramp of 10°C min⁻¹ held for 5 min at 300 °C. Peaks were identified by comparison of mass spectra to the Wiley mass spectra library. Kovats Indices (KI) were calculated for each peak according to the retention time relative to alkane standards (Sekiwa-Iijima, 2001).

**Gene expression analysis**

First-strand cDNA was synthesized using the iScript cDNA Synthesis kit (BioRad) according to the manufacturer’s instructions. Gene expression analysis was done by real-time quantitative PCR (qPCR) with the iCycler iQ5 system (BioRad) (Spinsanti G, 2006) using the iQ™ SYBR® Green Supermix master mix (Biorad) in three independent biological and two technical replicates. Actin was used as a housekeeping gene. The ΔCt was calculated as follows: ΔCt = Ct(TpGAS) – Ct(TpActin) and the RGE as: RGE= POWER(2,-ΔCt) (Livak & Schmittgen, 2001). A two step program was used as follows: 3 min at 95 °C; 40 cycles of 10 s at 95 °C, 30 s at 55 °C followed by melting curve analysis. The following primer pairs were used: forward actin 5’-CCTCTTTAATCCTAAGGCTAATC-3’ ; reverse actin 5’-CCAGGAATCCAGCA-CAATACC-3’; forward TpGAS 5’-TTCTCCTCTTTATTTCAACTGTG-3; reverse TpGAS 5’-TGCTATCTCGGTGACTTTCAAGG-3’; forward TpCarS 5’-GCATCCAGTTGGAAAA-GTTAC-3’; reverse TpCarS 5’-GGTCATCAGAGGCCATCGG -3’.

**Data analysis**

Sequence comparison was performed by Blast Search in GenBank (http://www.ncbi.nih.gov). Phylogenetic tree analysis and sesquiterpene alignment analysis were performed using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and Bio-edit software, respectively.

A completely randomized design with 3 replications was used for the comparison of the parthenolide concentration and RGE in different tissues, within flower parts and in different developmental stages of the flower. Mean comparison was conducted using Duncan's test.
Student's t-test was conducted to compare parthenolide concentration in disc florets and chloroform dipped disc florets and for the comparison of the RGE for TpGAS and TpCarS in glandular trichomes. All the data were statistically analysed using SPSS statistical software.

Acknowledgements

MM acknowledges the Ministry of Sciences, Research and Technology of Iran for awarding a fellowship. LQ, RdV, ST, AS and HB acknowledge funding by the European Commission (EU-project TerpMed, 227448). JB and KC were funded by IBOS-ACTS 053.63.305. We acknowledge Francel Verstappen, Benyamin Houshyani, Aldana Ramirez, Miriam Goedbloed, Liping Gao, and Jos Moltthoff for their helpful suggestions.

References


Chapter 2


Biosynthesis and localization of parthenolide


Chapter 3

Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*

Qing Liu¹, Mohammad Majdi¹, Katarina Cankar, Miriam Goedbloed, Tatsiana Charnikhova, Francel W.A. Verstappen, Ric C.H. de Vos, Jules Beekwilder, Sander van der Krol, Harro J. Bouwmeester

¹These two authors contributed equally to this work.

Chapter 3

Abstract

The sesquiterpene costunolide has a broad range of biological activities and is the parent compound for many other biologically active sesquiterpenes such as parthenolide. Two enzymes of the pathway leading to costunolide have been previously characterized: germacrene A synthase (GAS) and germacrene A oxidase (GAO) which together catalyse the biosynthesis of germacra-1(10),4,11(13)-trien-12-oic acid. However, the gene responsible for the last step towards costunolide has not been characterized until now. Here we show that chicory costunolide synthase (CiCOS), CYP71BL3, can catalyse the oxidation of germacra-1(10),4,11(13)-trien-12-oic acid to yield costunolide. Co-expression of feverfew GAS (TpGAS), chicory GAO (CiGAO), and chicory COS (CiCOS) in yeast resulted in the biosynthesis of costunolide. The catalytic activity of TpGAS, CiGAO and CiCOS was also verified in planta by transient expression in Nicotiana benthamiana. Mitochondrial targeting of TpGAS resulted in a significant increase in the production of germacrene A compared with the native cytosolic targeting. When the N. benthamiana leaves were co-infiltrated with TpGAS and CiGAO, germacrene A almost completely disappeared as a result of the presence of CiGAO. Transient expression of TpGAS, CiGAO and CiCOS in N. benthamiana leaves resulted in costunolide production of up to 60 ng.g\(^{-1}\) FW. In addition, two new compounds were formed that were identified as costunolide-glutathione and costunolide-cysteine conjugates.
Introduction

Sesquiterpene lactones (SLs) are a major class of plant secondary metabolites. These bitter tasting, lipophilic molecules form the active constituents of a variety of medicinal plants used in traditional medicine (Rodriguez et al., 1976; Zhang et al., 2005). Some SLs show bioactivities which are beneficial to human health, such as anti-inflammatory (e.g. helenalin) (Lyss et al., 1998), anti-cancer (e.g. costunolide) (Koo et al., 2001), and anti-malarial (artemisinin) (Klayman, 1985). The majority of SLs have been reported from the Asteraceae family, with over 4000 different SLs that have been identified (de Kraker et al., 2002). While the detailed structure of those SLs varies, their backbones are constrained to a limited set of core skeletons, such as germacranolide, eudesmanolide and guaianolide (Seto et al., 1988; Fischer, 1990; Van Beek et al., 1990). For all these three types of sesquiterpene lactones costunolide is generally considered the common precursor (de Kraker et al., 2002). Costunolide has been detected in many medicinal plants and several biological activities were ascribed to it including anti-carcinogenic, anti-viral, anti-fungal, and immunosuppressive activities (Mori et al., 1994; Chen et al., 1995; Taniguchi et al., 1995; Barrero et al., 2000; Wedge et al., 2000). Synthetic derivatives of costunolide such as 13-amino costunolide derivatives have anti-cancer activity (Srivastava et al., 2006) and also biosynthetic downstream products derived from costunolide have been reported to have interesting biological properties. For example, parthenolide has been reported to have anti-inflammatory and anti-cancer activity (Bedoya et al., 2008; Zhang et al., 2009).

Despite the importance of costunolide-derived SLs, the biosynthesis pathway of costunolide has not been fully elucidated. The pathway from FPP to costunolide was first proposed by de Kraker et al. based on the presence of enzymes in chicory roots that convert FPP to costunolide (de Kraker et al., 1998; de Kraker, JW et al., 2001; de Kraker et al., 2002) (Figure 1). First, farnesyl diphosphate is converted to germacrene A by germacrene A synthase (GAS) (de Kraker et al., 1998). GAS genes have been isolated and characterized from several members of the Asteraceae family, such as chicory (Bouwmeester et al., 2002), lettuce (Bennett et al., 2002), Artemisia annua (Bertea et al., 2006), and feverfew (Majdi et al., 2011).

In the next step of the pathway, germacrene A is oxidized at its C13 methyl by germacrene A oxidase (GAO) to form germacr-1(10),4,11(13)-trien-12-ol, which is then further oxidised to germacr-1(10),4,11(13)-trien-12-al and germacr-1(10),4,11(13)-trien-12-oic acid (de Kraker, J-W et al., 2001; de Kraker et al., 2002). The C6 position of germacr-1(10),4,11(13)-trien-12-oic acid is subsequently hydroxylated by a putative cytochrome P450 mono-oxidase, after which presumably spontaneous cyclization of the C6 hydroxyl and C12 carboxylic group leads to the formation of costunolide (de Kraker et al., 2002).

Although biosynthesis of costunolide from germacr-1(10),4,11(13)-trien-12-oic acid has been demonstrated in chicory biochemically (de Kraker et al., 2002), the corresponding gene responsible for this step has not been identified to date. It was shown that both germacrene A oxidase and costunolide synthase are cytochrome P450 enzymes. Recently, genes that encode germacrene A oxidase were cloned from a number of Asteraceae species (Nguyen et al.,
A valencene oxidase gene (CYP71AV8) was also reported to have the germacrene A oxidase activity (Cankar et al., 2011). All these genes belong to the CYP71 group of cytochrome P450s. In the present study, we investigated 5 candidate CYP71 P450 genes from a chicory cDNA library for costunolide synthase activity. The putative CiCOS gene was characterised by reconstitution of the costunolide biosynthetic pathway in yeast as well as in Nicotiana benthamiana, and the products formed were analysed using GC-MS and LC-MS metabolic profiling.

**Results**

**Optimizing germacrene A production in planta**

To produce germacrene A we used the GAS gene isolated from feverfew (Tanacetum parthenium), TpGAS (Majdi et al., 2011). After cloning of the full length coding sequence into a yeast expression vector, the TpGAS activity was compared with the previously characterized GAS genes from chicory (CiGAS-l and CiGAS-s) (Bouwmeester et al., 2002). Results showed that yeast culture expressing TpGAS had an approximately three fold higher activity than that of the CiGAS gene(s) (Figure 2).

Subsequently, the TpGAS cDNA - using its native targeting to the cytosol (cTpGAS) or equipped with a mitochondrial targeting
Biosynthetic pathway reconstitution of costunolide

A biosynthetic pathway reconstitution of costunolide signal (mTpGAS) - was cloned into a binary expression vector under control of the Rubisco promoter and introduced into Agrobacterium tumefaciens. For analysis of in planta activity, N. benthamiana leaves were agro-infiltrated with the indicated genes, and were analysed after 3 days. In the headspace of cTpGAS agro-infiltrated N. benthamiana, germacrene A was detected while no germacrene A was detected in leaves infiltrated with the empty vector (Figure 3A, line a and b). It has been shown for several terpene synthases that targeting to the mitochondria rather than to the cytosol which is the native compartment for sesquiterpene synthases results in higher production, presumably because of higher substrate availability in the mitochondria (Kappers et al., 2005; van Herpen et al., 2010). The TpGAS coding sequences was therefore fused to the CoxIV mitochondrial targeting sequence (mTpGAS). In N. benthamiana leaves infiltrated with A. tumefaciens carrying mTpGAS, the germacrene A production was approximately 15-fold higher than obtained by expression of cTpGAS (Figure 3A, line b and line c). Therefore, mTpGAS was used for the reconstruction of the costunolide pathway in N. benthamiana.

**Functional characterization of GAO in planta**

Previously the chicory germacrene A oxidase (CiGAO, GenBank: GU256644) was character-

---

**Figure 3.** Headspace analysis of volatiles emitted from agro-infiltrated Nicotiana benthamiana leaves. A, GC-MS chromatograms are shown for the volatiles emitted from N.benthamiana leaves infiltrated with the indicated genes. Line a is a negative control, line b and c display the different amount of compound 1 (germacrene A) produced by N. benthamiana leaves infiltrated with TpGAS with different targeting signals: mTpGAS, mitochondrial targeting; cTpGAS, cytosolic targeting. Line d shows that compound 1 which is produced upon mTpGAS agro-infiltration disappears upon agro-infiltration with CiGAO. Agro-infiltration with CiGAO alone does not induce any volatile formation (Line e). B, the mass fragmentation patterns of compound 1 (a) and a β-elemene from the Wiley library (b). C, cope rearrangement of germacrene A to β-elemene by heat.
Chapter 3

We amplified the same gene (CiGAO) from a chicory cDNA library and the enzymatic activity was confirmed in our yeast system by co-expression of TpGAS and CiGAO (Figure 4A, line b). To test the activity of CiGAO in planta, an expression vector containing CiGAO was co-infiltrated with the mTpGAS expression vector into N. benthamiana leaves. In the headspace of mTpGAS+CiGAO agro-infiltrated N. benthamiana leaves, the germacrene A peak was no longer detected (compare Figure 3A line c with line d), suggesting that CiGAO can efficiently catalyse the conversion of the product of mTpGAS, germacrene A, into one or more other products. However, new peaks were visible neither in the headspace (Figure 3), nor in dichloromethane (DCM) extracts (data not shown) of mTpGAS+CiGAO agro-infiltrated N. benthamiana leaves compared with those of mTpGAS. Agro-infiltrated with CiGAO alone in N. benthamiana leaves, used as negative control, does not induce any volatile formation (compare Figure 3A line e and line d).

To investigate whether any non-volatile products were formed, the infiltrated leaves were extracted with aqueous methanol and the extracts analysed by accurate mass LC-QTOF-MS analysis. Metabolite profiles of the pBIN (empty vector), mTpGAS, and mTpGAS+CiGAO samples were recorded and the mass signals extracted in an untargeted manner using Metalign software (www.metAlign.nl), followed by clustering of extracted mass features into recon-
Biosynthetic pathway reconstitution of costunolide

structured metabolites (Tikunov et al., 2005). The comparison of the mTpGAS+CiGAO co-infiltrated leaves and the mTpGAS infiltrated leaves revealed that, of the 2023 individual mass peaks with a signal-to-noise ratio higher than 3, none differed more than 2-fold (p<0.01, n=3; student T-test) between the two sample groups. This result could be explained by conjugation of the expected products of GAO oxidation of germacrene A (germacra-1(10),4,11(13)-trien-12-ol, germacra-1(10),4,11(13)-trien-12-al and germacra-1(10),4,11(13)-trien-12-oic acid) to multiple compounds, resulting in a distribution of the product signal over multiple masses that apparently remain below the 2-fold threshold or the level of detection.

Characterisation of a costunolide synthase gene from chicory

Chicory (Cichorium intybus L.) accumulates costunolide in roots (Van Beek et al., 1990). We assumed costunolide synthase to be a cytochrome P450 enzyme (as demonstrated by de Kraker et al. (2002)) which evolved from the GAO gene and therefore should show close homology to the CiGAO amino acid sequences. A root-specific cDNA library from chicory was available (Cankar et al., 2011), the sequences from the library were combined with chicory ESTs from GenBank (http://www.ncbi.nlm.nih.gov) and UC Davis database (http://compgenomics.ucdavis.edu/compositae_index.php), and these were searched for sequences with homology to the cytochrome P450 sequences of germacrene A oxidase from Cichorium intybus (Nguyen et al., 2010; Cankar et al., 2011) and Lactuca sativa (Nguyen et al., 2010). Five P450 sequences were identified which clustered into class CYP71 and had high similarity to the GAO genes mentioned above. Figure 5 shows the phylogenetic relationship of the candidate chicory CYP71 P450 sequences and GAO genes from different plant species. Each of these candidate cDNA sequences was then cloned into a yeast expression vector and tested in the yeast expression system by co-transformation with TpGAS and CiGAO. One of the isolated cDNAs (3368) encodes an enzymatic activity which was able to produce costunolide (Figure 4) in the presence of TpGAS and CiGAO, and therefore was designated as costunolide synthase (CiCOS) (CYP71BL3).

As the conversion of the germacra-1(10),4,11(13)-trien-12-oic acid to costunolide was quite low in HEPES buffer (Figure 4), we tested the effect of another buffer, MOPS, on cos-
tunolide production in yeast, using UPLC-MRM-MS to quantify costunolide production. For $\text{TpGAS}+\text{CiGAO}+\text{CiCOS}$ transformed yeast cultured in HEPES buffer (pH 7.5) the costunolide production was 9 μg mL$^{-1}$ culture, while in MOPS buffer (pH 7.5) production was about 3-fold higher (28 μg mL$^{-1}$ culture).

To test the activity of the newly identified $\text{CiCOS}$ in planta, the cDNA was cloned into a binary expression vector under control of the Rubisco promoter. $N.\ benthamiana$ leaves were co-infiltrated with agrobacterium cultures with $\text{RBC::mTpGAS}, \text{RBC::CiGAO}$ and $\text{RBC::CiCOS}$ and after 3 days leaves were extracted with methanol and extracts were analysed by UPLC-MRM-MS for quantification of free costunolide. Results show that average production of costunolide from eight infiltration experiments was 48.6 ± 13.4 ng g$^{-1}$ FW. No costunolide was detected in extracts from leaves infiltrated with either $\text{pBIN}$ (empty vector), $\text{RBC::mTpGAS}, \text{RBC::mTpGAS+RBC::CiGAO}$ or $\text{RBC::mTpGAS+RBC::CiCOS}$ (data not shown), indicating that the production of costunolide by $\text{CiCOS}$ in $N.\ benthamiana$ leaves is dependent on the presence of both $\text{TpGAS}$ and $\text{CiGAO}$.

To investigate whether there were any other metabolic changes caused by co-infiltration of $\text{RBC::mTpGAS}, \text{RBC::CiGAO}$ and $\text{RBC::CiCOS}$, an untargeted LC-QTOF-MS analysis of aqueous methanol extracts from leaves was carried out. Comparison of the chromatograms...
of extracts from co-infiltrated leaves showed two new compounds, eluting at 22.30 and 22.48 min, in the leaves infiltrated with mTpGAS+CiGAO+CiCOS compared to leaves infiltrated with mTpGAS+CiGAO (Figure 6)

**Identification of costunolide conjugates**

In order to identify the two new compounds in the leaves infiltrated with mTpGAS+CiGAO+CiCOS, the apparent parent masses of the peaks at 22.30 and 22.48 min were fragmented by LC-MS/MS in negative mode. Within the MS/MS fragments of 352.1615 (parent ion of peak at 22.30 min, a 9.2 ppm deviation from the elemental formula C_{18}H_{27}NO_4S), we detected an ion with mass 120.0142, a 19 ppm deviation from the elemental formula of cysteine (C_{3}H_{7}NO_2S). This MS/MS experiment therefore suggests that the peak at 22.30 min is a costunolide (C_{15}H_{20}O_2)-cysteine (C_{3}H_{7}NO_2S) conjugate. Within the MS/MS fragments of 538.2206 (parent ion of peak at 22.48 min, a -19.9 ppm deviation from the elemental formula C_{25}H_{37}N_3O_8S), we detected an ion with mass 306.0760, a -2.9 ppm deviation from the elemental formula of glutathione (C_{10}H_{17}N_3O_6S). This MS/MS experiment therefore suggests that the peak at 22.48 min is a costunolide (C_{15}H_{20}O_2)-glutathione (C_{10}H_{17}N_3O_6S) conjugate (Figure 6).

To further confirm the identity of these putative costunolide glutathione and cysteine conjugates, we tested the activity of a glutathione-S-transferase (GST) enzyme on costunolide in an in vitro enzyme assay. Analysis of the reaction mix of costunolide with glutathione and GST by LC-QTOF-MS showed that costunolide was efficiently converted into a costunolide-glutathione conjugate (Figure 7) that had the same retention time and exact mass and MS fragments as the postulated costunolide conjugate detected in the extract of the mTpGAS+CiGAO+CiCOS agro-infiltrated leaf sample (Figure 7A and 7C). When costunolide and glutathione were incubated without GST enzyme, the same costunolide-glutathione conjugate was formed indicating that the conjugation of costunolide and glutathione can occur spontaneously. Similarly, when costunolide was incubated with cysteine a costunolide-cysteine conjugate was spontaneously formed (Figure 7B). The cysteine conjugate that was spontaneously formed by the in vitro reaction had the same mass spectrum and retention time as the compound produced in mTpGAS+CiGAO+CiCOS agro-infiltrated N. benthamiana leaves (Figure 7B and 7D). Thus the two new peaks in mTpGAS+CiGAO+CiCOS agro-infiltrated N. benthamiana leaf were confirmed to be a costunolide-glutathione and a costunolide-cysteine conjugate. The presumed structure of these two compounds is shown in Figure 7E.

**Discussion**

**Reconstitution of the costunolide biosynthesis pathway in yeast**

Costunolide is the precursor of many biologically active SLs, and reconstitution of its biosynthetic pathway in heterologous hosts could form an attractive option for commercial production of these compounds. Partial reconstruction of the sesquiterpene biosynthesis pathways in yeast has been demonstrated for the antimalarial drug artemisinin (Ro et al., 2006) and for germacr-1(10),4,11(13)-trien-12-oic acid (Nguyen et al., 2010; Cankar et al., 2011). Here we
Figure 7. Costunolide-glutathione and costunolide-cysteine conjugate identification by GST enzyme assay and LC-MS analysis. A. LC-MS chromatograms of \( m/z \) =538 of extracts of *N. benthamiana* leaves agro-infiltrated with TpGAS+CiGAO+CiCOS, costunolide-GSH conjugate formed in an enzyme assay of costunolide and GSH with GST, costunolide-GSH conjugate formed by non-enzymatic conjugation of costunolide and GSH. B. LC-MS chromatograms of \( m/z \) =352 of extracts of *N. benthamiana* leaves agro-infiltrated with TpGAS+CiGAO+CiCOS, costunolide-cysteine (Cys) conjugate formed in an enzyme assay of costunolide and Cys with GST, costunolide-cysteine (Cys) conjugate formed non enzymatically from costunolide and Cys. C. \( m/z \) spectrum of peak 22.48 and costunolide-GSH conjugate (RT=22.52). Arrows indicate parent ions of GSH-glutathione. D. \( m/z \) spectrum of peak 22.30 and costunolide-Cys conjugate (RT=22.30). Arrows indicate parent ions of GSH-cysteine. E. Presumed molecular structure of costunolide-GSH (a) and costunolide-Cys (b) conjugates. GSH, glutathione; Cys, cysteine; GST, glutathione S-transferase. RT, retention time. Y-axis scale is identical in all chromatograms.
reconstituted the biosynthetic pathway of costunolide in yeast and in planta. To achieve this, we screened five candidate CYP71 P450 genes from chicory for costunolide synthase activity, which yielded one gene which had this activity, CiCOS. This novel gene was combined with a new GAS gene from feverfew (TpGAS) and the previously identified chicory GAO (Nguyen et al., 2010; Cankar et al., 2011). Co-expression of TpGAS together with CiGAO and CiCOS in yeast yielded just low levels of costunolide, so it was of interest to see if production in yeast could be boosted. We showed that using GAS from different sources may have a strong effect on germacrene A production (Figure 2).

Also the culture buffer conditions strongly affect costunolide production. We showed that germacr-1(10),4,11(13)-tri-en-12-oic acid can be more efficiently converted by CiCOS into costunolide when the yeast is cultured in MOPS (pH 7.5) buffer instead of HEPES buffer (pH 7.5). The pH of the buffered yeast culture decreased from 7.5 to 6.8 for both MOPS and HEPES buffered culture after 48 hours of cultivation. Therefore, the difference in costunolide production was not due to the buffering capacity of the buffer. We presume that the increased costunolide production results from improved growth of yeast in the presence of MOPS, compared to the HEPES-buffered yeast culture.

Reconstitution of the costunolide biosynthesis pathway in N. benthamiana

Reconstruction of pathways in the transient Nicotiana spp. plant expression system was demonstrated for many medically relevant proteins (Tremblay et al., 2010) and has been shown to be a good model to study the production of sesquiterpenoid pharmaceutical compounds (van Herpen et al., 2010). A marked peak of germacrene A was detected in the headspace of mTpGAS agro-infiltrated N. benthamiana leaves. The targeting of sesquiterpene synthases in metabolic engineering seems to have a great effect on its performance in plants. It has been shown in Arabidopsis that fusion of nerolidol synthase from strawberry to a mitochondrial targeting sequence leads to the biosynthesis of nerolidol, whereas this was not the case when using a cytosol targeting sequence (Kappers et al., 2005). This is in line with our observation that the germacrene A emitted from N. benthamiana leaves infiltrated with mitochondrial targeted germacrene A synthase is much higher than that from leaves infiltrated with the cytosolic germacrene A synthase (Figure 3).

The disappearance of germacrene A observed upon co-expression of mitochondrial TpGAS and CiGAO suggests that germacrene A is efficiently transferred from the mitochondrial compartment to the cytosol/ER, where it is presumably converted into germacrene A acid by CiGAO. The activity of mitochondrial TpGAS might produce a concentration gradient sufficient to drive diffusion of germacrene A into the cytosol. Or, as suggested by Turner and Croteau (2004), some type of terpenoid carrier protein, mitochondrial membrane pump, or transient contacts between ER and mitochondrial membranes might facilitate germacrene A movement from mitochondria to cytosol/ER. Germacrene A could not be detected in the headspace or solvent extracts of mTpGAS+CiGAO agro-infiltrated N. benthamiana leaves. We suppose that this is due to an efficient conversion of germacrene A into germacr-1(10),4,11(13)-tri-en-12-oic acid by the expressed CiGAO enzyme. The fact that we were
not able to detect germacr-1(10),4,11(13)-trien-12-oic acid may be explained by assuming that the compounds produced by CiGAO subsequently react with endogenous metabolites to form conjugates that were either not detected by our GC-MS and LC-QTOF-MS (ESI negative mode) profiling approaches, or were converted and diluted out into many new metabolites of which the signals were below our detection levels. The disappearance of germacrene A after introduction of CiGAO suggests an efficient transfer of substrate between GAS and GAO. Similar results have been reported by van Herpen et al. (2010): co-infiltration of N. benthamiana with the cDNA encoding amorphadiene synthase plus CYP71AV1, also a sesquiterpene oxidizing P450, leads to an almost complete conversion of amorphadiene. Also in the latter paper, the product of this conversion, artemisinic acid was efficiently further converted to a diglucose conjugate (2010).

Co-expression of TpGAS together with CiGAO and CiCOS in N. benthamiana leaves yielded up to 60 ng g\(^{-1}\) FW of costunolide. In addition, two novel costunolide conjugates were detected.

**Costunolide glutathione conjugates are present in planta but not in yeast**

Most glutathione conjugations are catalysed by glutathione S-transferase (GST), which may be constitutively active or be induced upon oxidative stress or exogenous heterocyclic compounds, such as herbicides (Marrs, 1996). The expression of two glutathione S-transferases (NbGSTU1 and NbGSTU3) in N. benthamiana was up-regulated progressively during infection by the fungus Colletotrichum destructivum (Dean et al., 2005). Here we show that glutathione conjugates may be formed spontaneously from costunolide and GSH in an in vitro enzyme assay (Figure 7). S-glutathionylated metabolites are likely tagged for vacuolar import by ATP binding cassette (ABC) transporters, which selectively transport GSH conjugates, as has been shown for other glutathione S-conjugates (Rea, 1999). Storage of target metabolites as glutathione S-conjugates may have the advantage that the storage capacity of the vacuole is used and that high concentrations can be reached without phytotoxic effects. Marrs et al. (1995) reported that anthocyanin pigments require GSH conjugation for transport into the vacuole. If conjugation is inhibited, this leads to inappropriate cytoplasmic retention of the pigment which is toxic for the cells.

In addition to costunolide-GSH, costunolide-cysteine was also found to accumulate in agro-infiltrated N. benthamiana leaves. This cysteine conjugate may be a breakdown product of costunolide-glutathione (Ohkama-Ohtsu et al., 2007). However, we showed that the costunolide-cysteine conjugate may also be formed spontaneously from costunolide and cysteine (Figure 7).

Remarkably, no costunolide-glutathione or costunolide-cysteine conjugates were detected in medium of yeast transformed by TpGAS+CiGAO+CiCOS (data not shown), even though GSH is present in yeast (Shimizu et al., 1991) and some transporter genes, such as Bpt1p (Klein et al., 2002), have been reported to mediate vacuolar sequestration of glutathione conjugates in yeast. It could be that free costunolide is excreted out of the yeast cells and that any costunolide-glutathione conjugate formed inside the cell has a short half-life.
**Spontaneous conjugation to glutathione related to bioactivity of costunolide?**

It has been shown that the effect of costunolide treatment of cancer cells is based on a rapid depletion of the intracellular reduced glutathione and protein thiols, which precedes apoptosis. Indeed, the effect of costunolide can be blocked by pretreatment with sulfhydryl compounds such as GSH, N-acetyl-L-cysteine, dithiothreitol and 2-mercaptoethanol (Choi et al., 2002). The apoptosis-inducing activity of costunolide likely depends on the exomethylene moiety because derivatives in which this group was reduced, such as dihydrocostunolide and saussurea lactone, did not deplete the cellular thiols and showed no apoptotic activity (Choi et al., 2002; Park et al., 2006). If the biological activity of costunolide depends on the ability to conjugate glutathione and thiols, then the costunolide-glutathione conjugate produced in *N. benthamiana* may not exhibit biological activity. On the other hand, the poor water-solubility of costunolide may also limit its potential as a promising clinical agent (Ma et al., 2007) and conjugation could improve this property. Regulation of conjugation in heterologous plant hosts and secretion into cell compartments that allow accumulation of free costunolide could therefore be an important target for further optimization of a costunolide production platform.

**In conclusion**

We describe here the discovery of a new gene, *CiCOS*, and its functional characterization in yeast as well as *in planta*. *CiCOS* encodes the enzyme catalyzing the formation of a sesquiterpene lactone, costunolide, a promising anti-cancer medicine, and a crucial intermediate in the biosynthesis of many other sesquiterpene lactones with important biological activities, such as parthenolide. The cloning of this gene allows for the development of platforms – in microbial as well as in plant hosts – for the production of sesquiterpene lactones that can potentially be developed into new drugs. The conjugation of costunolide to glutathione and cysteine detected upon *in planta* expression has never been reported before in plants and could present new opportunities for high production because of better storability as well as for the development of drugs with better water solubility.

**Materials and methods**

**Isolation and cloning of costunolide synthase candidate gene from chicory**

A previously reported chicory cDNA taproot library (Cankar et al., 2011) was used for gene isolation. Five candidate cytochrome P450 contigs belonging to the CYP71 family were identified by sequence homology to known sesquiterpene monooxygenases. RACE PCR (Clontech) was used to obtain the sequence of the 5’- and 3’-region of the candidate contigs. Full length cDNAs of candidate genes were amplified from chicory cDNA with the addition of NotI/ PacI restriction sites. They were subsequently cloned to the yeast expression vector pYEDP60 (Pompon et al., 1996), which was modified to contain PacI/ NotI sites at the polylinker, and sequenced. The DNA sequence for the chicory costunolide synthase (*CiCOS*), has been deposited in GenBank under the accession number JF816041. The sequence was also submitted.
to David Nelson’s cytochrome P450 homepage (http://drnelson.uthsc.edu/cytochromeP450.html) and was assigned the name CYP71BL3 (Nelson, 2009).

**Plasmid construction for gene expression in yeast**

For the production of germacra-1(10),4,11(13)-trien-12-oic acid in yeast, CiGAO and TpGAS genes were both cloned into the pESC-Trp yeast expression vector (Agilent technologies) with the TRP1 auxotrophic selection marker. CiGAO (Cankar et al., 2011) was subcloned from the yeast vector PYEDP60 (Pompon et al., 1996) to the pESC-Trp vector using NotI/PacI restriction sites. The obtained construct was named CiGAO pESC-Trp. Subsequently, TpGAS was amplified from the pACYCDeut™-1 vector (Majdi et al., 2011) using high fidelity Phusion polymerase (Finnzymes) with the addition of BamHI/KpnI restriction sites. The amplified product was digested by BamHI/KpnI and ligated into the CiGAO pESC-Trp plasmid, yielding the final plasmid TpGAS+CiGAO pESC-Trp. No terminal tags were added in these constructs. This plasmid was transformed into the WAT11 (Urban et al., 1997) yeast strain and the clones were selected on Synthetic Dextrose (SD) minimal medium (0.67% Difco yeast nitrogen base medium without amino acids, 2% D-glucose, 2% agar) supplemented with amino acids, but omitting L-tryptophane for auxotrophic selection of transformants. TpGAS+CiGAO pESC-Trp and pYEDP60 plasmids containing costunolide synthase candidates were co-transformed into the WAT11 yeast strain. After transformation yeast clones containing both plasmids were selected on SD minimal medium supplemented with amino acids, but omitting uracil, adenine sulphate and L-tryptophane for auxotrophic selection of transformants.

**Gene induction in yeast and metabolite extraction**

For the induction of gene expression in yeast, the transformed WAT11 yeast strain with Tp-GAS+CiGAO pESC-Trp or co-transformed with TpGAS+CiGAO pESC-Trp and costunolide synthase candidate-PYED60-Ura-Ade were inoculated in 3 mL Synthetic Galactose (SG) minimal medium (0.67% Difco yeast nitrogen base medium without amino acids, 2% D-galactose, 2% agar) but omitting TRP or Trp-Ura-Ade amino acids, respectively. The yeast was cultured overnight at 30 °C and 300 rpm. The start culture was diluted to OD 0.05 in SG minimal medium omitting Trp or Trp-Ura-Ade amino acids, respectively. All yeast induction experiments were performed in triplicates in 50 mL of culture volume. Cultures were buffered at pH 7.5 using 100 mM HEPES or 100 mM MOPS. After fermentation for 48h at 30 °C and 300 rpm, the medium was extracted with 20 mL ethyl acetate. From this, a 10 mL sample was taken and the ethyl acetate evaporated with a stream of N2 to a final volume of 1 mL which was analyzed by GC-MS. For UPLC-MRM-MS analysis the ethyl acetate in a 10 mL subsample was completely evaporated and the residue redissolved in 300 μl of 25% acetonitrile in water.

**Plasmid construction for expression in Nicotiana benthamiana**

For expression in N. benthamiana, TpGAS, CiGAO and CiCOS were cloned into ImpactVector1.1 (http://www.impactvector.com/) to express them under the control of the Rubisco (RBC) promoter (Outchkourov et al., 2003). TpGAS was also cloned into ImpactVector1.5
to fuse it with the RBC promoter and the CoxIV mitochondrial targeting sequence. An LR reaction (Gateway-LR Clonase TM II) was carried out to clone each gene into pBinPlus binary (Vanengelen et al., 1995) vector between the right and left borders of the T-DNA for plant transformation.

**Transient expression in *N. benthamiana***

*A. tumefaciens* infiltration (agro-infiltration) was performed according to the description of van Herpen et al. (2010). *A. tumefaciens* batches were grown at 28 °C at 220 rpm for 24 hours in YEP media with kanamycin (50 mg L−1) and rifampicillin (34 mg L−1). Cells were harvested by centrifugation for 20 min at 4000xg and 20 °C and then resuspended in 10 mM MES buffer containing 10 mM MgCl2 and 100 μM acetosyringone (4’-hydroxy-3’,5’-dimethoxyacetophenone, Sigma) to a final OD600 of ~ 0.5, followed by incubation at room temperature under gentle shaking at 50 rpm for 150 min. For co-infiltration, equal volumes of the Agrobacterium batches were mixed. Batch mixtures were infiltrated into leaves of three-week-old *N. benthamiana* plants by pressing a 1 mL syringe without metal needle against the abaxial side of the leaf and slowly injecting the bacterium suspension into the leaf. *N. benthamiana* plants were grown from seeds on soil in the greenhouse with a minimum of 16 hour light. Day temperatures were approximately 28 °C, night temperatures 25 °C. After agro-infiltration the plants were grown under greenhouse conditions for another 3 days and then harvested for analysis.

**Headspace analysis and GC-MS thermodesorption**

Volatile collection from agro-infiltrated *N. benthamiana* leaves and GC-MS analysis were performed according to van Herpen et al. (2010). Steel sorbent cartridges (89 mm × 6.4 mm O.D.; Markes) containing Tenax were used for volatile collection. Cartridges were conditioned at 280 °C for 40 min under a nitrogen flow of 20 psi in a TC-20 multi-tube conditioner and were capped airtight until use. *N. benthamiana* leaves were sampled and placed on water in a small vial and were enclosed in a glass container. To trap the leaf-produced volatiles, air was sucked through the containers with a flow rate of 90 mL min−1 for 24 hours and released through one cartridge. A second cartridge was used to purify the incoming air. Sample cartridges were dried for 15 min at room temperature with a nitrogen flow of 20 psi before GC-MS analysis on a Thermo Trace GC Ultra connected to a Thermo Trace DSQ quadruple mass spectrometer (Thermo Fisher Scientific, USA).

Cartridges were placed in an automated thermodesorption unit (Ultra; Markes, Llantrisant) in which they were flushed with helium at 50 mL min−1 for 2 min to remove moisture and oxygen just before thermodesorption. The volatiles were desorbed by heating of the cartridges at 220 °C for 5 min with a helium flow of 50 mL min−1. The compounds released were trapped on an electrically cooled sorbent trap (Unity; Markes, Llantrisant) at a temperature of 5 °C. Subsequently, the trapped volatiles were injected on the analytical column (ZB-5MSI, 30 m × 0.25 mm ID, 1.0 μm – film thickness, Zebron, Phenomenex) in splitless mode by ballistic heating of the cold trap to 250 °C for 3 min. The temperature program of the GC started at 40 °C (3 min hold) and rose 10°C min−1 to 280 °C (2 min hold). The column effluent was ionised by electron impact (EI) ionisation at 70 eV. Mass scanning was done from 33 to 280 m/z with
a scan time of 4.2 scans s⁻¹. Xcalibur software (Thermo, USA) was used to identify the eluted compounds by comparing the mass spectra with those of authentic reference standards.

**GC-MS analysis of solvent extracts**

Seven mL yeast culture was extracted three times with 2 mL ethyl acetate, which was concentrated, dried using anhydrous Na₂SO₄ and used for GC-MS analysis. Agro-infiltrated leaves (100 mg) were ground in liquid nitrogen and extracted with 800 μl dichloromethane. The extracts were prepared by brief vortexing and sonication for 10 min. Then the extracts were centrifuged for 15 min at 3000 rpm, dehydrated using Na₂SO₄, and then used for GC-MS analysis. A gas chromatograph (7890A, Agilent, USA) equipped with a 30 m × 0.25 mm, 0.25 mm film thickness column (ZB-5, Phenomenex) using helium as carrier gas at flow rate of 1 mL min⁻¹ was used for GC-MS analysis. Splitless mode was used for the injector with the inlet temperature set to 250°C. The initial oven temperature was 45 °C for 1 min, and was increased to 300 °C after 1 min at a rate of 10 °C min⁻¹ and held for 5 min at 300 °C. The GC was coupled to a Triple-Axis detector (5975C, Agilent). Compounds were identified by comparison of mass spectra and retention times (RT) with those of the following authentic standards: germacrene A, germacra-1(10),4,11(13)-trien-12-ol, germacra-1(10),4,11 (13)-trien-12-al (Cankar et al., 2011) and costunolide (TOCRIS bioscience). Quantification of sesquiterpenoids was conducted by determination of total ion count (TIC) peak area of the sesquiterpenoid peaks from three independent fermentation experiments. An absolute concentration of sesquiterpenoids was calculated from the peak area by comparison with calibration curves of the authentic standards. At the routine injection port temperature of 250 °C germacrene A, germacra-1(10),4,11(13)-trien-12-oic acid and costunolide are thermally converted into β-elemene, elematrien-12-oic acid, and saussurea lactone, respectively as discussed by de Kraker et al. (de Kraker et al., 1998; de Kraker et al., 2002; de Kraker et al., 2003; Yang et al., 2011). We also regularly injected samples with an injection port temperature of 150 °C to confirm the presence of non-rearranged germacrene A, germacra-1(10),4,11(13)-trien-12-oic acid and costunolide.

**LC-QTOF MS and MS/MS analysis**

Non-volatile metabolites were analysed by LC–QTOF-MS (liquid chromatography, coupled to quadrupole time-of-flight mass spectrometry) according to a protocol for untargeted metabolomics of plant tissues (De Vos et al., 2007). 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) operating in negative ionization mode was used. An analytical column (Luna 3 μ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA) attached to a C18 pre-column (2.0 × 4 mm; Phenomenex, USA) was used. Degassed eluent A [ultra-pure water: formic acid (1000:1, v/v)] and eluent B [acetonitril:formic acid (1000:1, v/v)] were used at a flow rate of 0.19 mL min⁻¹. Masses were recorded between m/z X and m/z Y; leucine enkaphalin ([M-H]⁻ = 554.2620) was used as a lock mass for on-line accurate mass correction.

For agro-infiltrated N. benthamiana, 100 mg infiltrated leaf from each treatment was ground in liquid nitrogen and extracted with 300 μl methanol:formic acid (1000:1, v/v). After
brief vortexing and sonication for 15 min, the extracts were centrifuged for 5 min at 13,000 rpm and filtered through a 0.2 μm inorganic membrane filter (RC4, Sartorius, Germany). The gradient of the HPLC started at 5% eluent B and increased linearly to 75% eluent B in 45 min, after which the column was washed and equilibrated for 15 min before the next injection. The injection volume was 5 μl. Data-directed MS-MS measurements were done at collision energies of 10, 15, 25, 35 and 50 eV.

**Costunolide detection and quantification by UPLC- MRM- MS**

Targeted analysis of costunolide in yeast extract and agro-infiltrated *N. benthamiana* leaves was performed with a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described by Kohlen et al. (2011) with some modifications. Chromatographic separation was obtained on an Acquity UPLC BEH C18 column (100 × 2.1 mm, 1.7 μm; Waters) by applying a water/acetonitrile gradient to the column, starting from 5% (v/v) acetonitrile in water for 1.25 min and rising to 50% (v/v) acetonitrile in water in 2.35 min, followed by an increase to 90% (v/v) acetonitrile in water in 3.65 min, which was maintained for 0.75 min before returning to 5% acetonitrile in water using a 0.15 min gradient. Finally, the column was equilibrated for 1.85 min using this solvent composition. Operation temperature and flow rate of the column were 50°C and 0.5mL min⁻¹, respectively. Injection volume was 5 μL. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L h⁻¹, respectively. The capillary voltage was set at 3.0 kV, the source temperature at 150°C, and the desolvation temperature at 650°C. The cone voltage was optimized for costunolide using the Waters IntelliStart MS Console. Argon was used for fragmentation by collision-induced dissociation in the ScanWave collision cell. MRM method was used for identification of costunolide in yeast extract and agro-infiltrated *N. benthamiana* leaves by comparing retention times and MRM mass transitions with that of a costunolide standard. MRM transitions for costunolide m/z 233.16>131.01 and 233.16>187.23 were optimized using the Waters IntelliStart MS Console.

**GC–MS and LC–MS data processing**

GC-MS and LC-MS data analysis was done according to the description by Yang et al. (2011) with minor modifications. GC–MS data were acquired using Xcalibur 1.4 (Thermo Electron Corporation) and LC–MS data using MassLynx 4.0 (Waters). The data were processed using MetAlign version 1.0 (www.metAlign.nl) for baseline correction, noise elimination and subsequent spectral data alignment (De Vos et al., 2007). The processing parameters of MetAlign for GC–MS data were set to analyse scan numbers 1,340–16,000 (corresponding to retention times 2.32 to 28.05 min) with maximum amplitude of 1.4×108. The processing parameters for LC–MS data were set to analyse scan numbers 60–2300 (corresponding to retention time 1.4 to 49.73 min) with a signal-to-noise ratio higher than 3.

To combine mass signals belonging to the same metabolite, all the detected masses were clustered by an in-house developed script called Multivariate Mass Spectra Reconstruction (MMSR) (Tikunov et al., 2005). The mass signal intensities (expressed as peak height using
MetAlign) obtained from agro-infiltrated plants and empty vector control plants were compared using the Student's t-test. Masses with a significant (p<0.05) intensity change of at least 2-fold were verified manually in the original chromatograms.

To annotate significantly different compounds in LC-QTOF-MS, accurate masses were manually calculated, taking into account only those scans with the proper intensity ratios of analyte and lock mass (between 0.25- and 2 (Moco et al., 2006)) and elemental formulae generated within 5 ppm deviation from the observed mass. In addition, data-directed LC-MS/MS experiments were performed on differential compounds. To obtain proper MS/MS spectra only molecular ions with signal intensities higher than 500 ion counts per scan were selected.

**Costunolide conjugation enzyme assay**

This enzyme assay was performed according to the method of Habig et al. (1974) with modifications. In brief, glutathione (150 mM) or cysteine (150 mM) in 7μl potassium buffer (100 mM; pH 6.5), and costunolide (30 mM) in 7μl ethanol were added to 200 μl potassium buffer (100 mM; pH 6.5). The reaction was initiated by adding 7 μl of GST (1g L⁻¹, in 100mM potassium buffer; pH 6.5) into the mixture. Complete assay mixtures without GST enzyme or either of the substrates were used as controls. After incubation for 15 min at room temperature, samples were kept at -20°C until analysis by LC-QTOF-MS.

**Acknowledgements**

We thank Bert Schipper for assistance in LC-MS analysis, Ting Yang for her help in MS data analysis, and Benyamin Houshyani for technical support in headspace analysis.

Note added in submission: During submission of this manuscript, a paper by Ikezawa et al. (2011) was published online (22 April 2011), in which a lettuce costunolide synthase (*CYP71BL2*) is reported. The amino acid sequence of chicory costunolide synthase (*CYP71BL3*) is 97% identical to that of *CYP71BL2* according to the alignment made by David Nelson (http://drnelson.uthsc.edu/cytochromeP450.html).

**References**


*Biosynthetic pathway reconstitution of costunolide.*
Chapter 3


Chapter 4

In planta reconstitution of the parthenolide biosynthetic pathway

Qing Liu, David Manzano, Nikola Tanić, Milica Pesic, Jasna Bankovic, Irini Pateraki, Lea Ricard, Albert Ferrer, Ric de Vos, Sander van de Krol, Harro Bouwmeester

Submitted
Abstract
Parthenolide, the main bioactive compound of the medicinal plant feverfew (*Tanacetum parthenium*), is a promising anti-cancer drug. Here we report on the isolation and characterization of all the genes from feverfew that are required for the biosynthesis of parthenolide, using a combination of 454 sequencing of a feverfew glandular trichome cDNA library, co-expression analysis and metabolomics. When parthenolide biosynthesis was reconstituted by transient co-expression of all pathway genes in *Nicotiana benthamiana*, up to 1.4 mg.g⁻¹ parthenolide was produced, mostly present as cysteine and glutathione conjugates. These relatively polar conjugates were highly active against colon cancer cells, with only slightly lower activity than free parthenolide. In addition to these biosynthetic genes, another gene encoding a costunolide and parthenolide 3β-hydroxylase was identified opening up further options to improve the water solubility of parthenolide and therefore its potential as a drug.
Introduction

Sesquiterpene lactones are a major class of plant secondary metabolites and over 4000 different structures have been elucidated (de Kraker et al., 2002). Many of these colourless, frequently bitter tasting, semi-polar molecules are the bioactive constituents of a variety of medicinal plants used in (traditional) medicine (Rodriguez et al., 1976; Zhang et al., 2005). Feverfew (Tanacetum parthenium) is one of the most prominent medicinal species in the Asteraceae family and a well-known remedy for the treatment of various diseases (Bedoya et al., 2008). It has been used for at least two millennia for the treatment of fever, as well as headache, menstrual irregularities, stomach-ache and to relieve arthritis and inflammation (Pareek et al., 2011). Parthenolide is the principal bioactive sesquiterpene lactone component in feverfew (Bork et al., 1997). The nucleophilic nature of the methylene-γ-lactone ring and epoxide group of parthenolide enables rapid interactions with different biological targets (Mathema et al., 2012). For instance, parthenolide can promote apoptosis by inhibiting the activity of the NF-κB transcription factor complex, and thereby down-regulating anti-apoptotic genes under NF-κB control (Bork et al., 1997; Wen, J. et al., 2002; Kishida et al., 2007; Parada-Turska et al., 2007; Zhang et al., 2009). Parthenolide has been reported to selectively target human leukaemia stem cells, while sparing normal stem or progenitor cells (Guzman et al., 2005). Despite these promising activities, application of this potent natural product is limited by its poor water-solubility (Sweeney et al., 2005). A number of chemically synthesized parthenolide derivatives with increased water solubility –hence allowing oral application - have been shown to retain bioactivity (Guzman et al., 2007; Neelakantan et al., 2009). Very recently, parthenolide and its cyclopropyl analogue have been synthesized chemically from costunolide (Long et al., 2013).

Parthenolide, a germacranolide type sesquiterpene lactone, is presumably derived from costunolide, in line with the proposed precursor role of costunolide for germacranolide-, eudesmanolide- and guaianolide-type sesquiterpene lactones (de Kraker et al., 2002). The initial committed step towards the formation of costunolide is the formation of germacrene A from farnesyl diphosphate (FDP), catalysed by the enzyme (+)-germacrene A synthase (GAS, Fig. 1) (de Kraker et al., 1998). Genes encoding GAS have been cloned from chicory (Cichorium intybus) (Bouwmeester et al., 2002), lettuce (Lactuca sativa) (Bennett et al., 2002), artemisia (Artemisia annua) (Bertea et al., 2006) and feverfew (Liu et al., 2011). In a number of oxidation steps, germacrene A is subsequently converted into germacr-1(10),4,11(13)-trien-12-oic acid by a cytochrome P450 enzyme, germacrone A oxidase (GAO) (de Kraker et al., 2001). Genes encoding GAO have previously been isolated from several Asteraceae species (Nguyen et al., 2010; Cankar et al., 2011). Germacr-1(10),4,11(13)-trien-12-oic acid is subsequently oxidised by costunolide synthase (COS) to 6α-hydroxy-germacr-1(10),4,11(13)-trien-12-oic acid, which undergoes spontaneous lactone ring formation to yield costunolide (de Kraker et al., 2002; Ikezawa et al., 2011; Liu et al., 2011). Finally, presumably a P450 monooxygenase catalyses the epoxidation of the C4-C5 double bond of costunolide, yielding parthenolide (Liu et al., 2011). The gene encoding the enzyme responsible for that epoxidation, parthenolide
synthase (PTS), has not been reported yet.

Elucidation of all parthenolide biosynthetic pathway and cloning of the structural genes may enable the production of parthenolide in heterologous systems. Therefore, we set out to identify and isolate all the genes of the parthenolide biosynthetic pathway in feverfew. We have previously reported on the TpGAS gene from feverfew (Majdi et al., 2011). Here we report on the isolation and characterisation of the remaining genes required for parthenolide biosynthesis in feverfew (TpGAO, TpCOS and TpPTS). Enzyme activities were characterized by expression of genes in yeast and subsequently the complete parthenolide biosynthetic pathway was reconstituted in Nicotiana benthamiana, by co-expression of TpGAS with the newly identified TpGAO, TpCOS and TpPTS. Extracts of N. benthamiana leaves co-expressing these genes contained free parthenolide, as well as cysteine and glutathione (GSH) conjugates of parthenolide. Because the conjugation to cysteine and GSH affects water solubility, we assessed the biological activity of these parthenolide derivatives in cancer cell lines. The parthenolide conjugates were less effective than free parthenolide but still displayed considerable anti-cancer activity, particularly in colon cancer cells. In addition to the parthenolide biosynthetic genes, another candidate gene was identified to encode a 3β-hydroxylase that uses costunolide as well as parthenolide as substrate. This may give additional possibilities to improve the water solubility of parthenolide. The production of parthenolide and more water-soluble conjugates through metabolic engineering of heterologous hosts may provide a sustainable alternative source for the further development of parthenolide as an anti-cancer drug.

Results

Identification of parthenolide biosynthesis candidate genes

Previously we have shown that parthenolide accumulates to high levels in floral glandular trichomes of feverfew (Tanacetum parthenium) (Majdi et al., 2011). To identify the genes in-
Biosynthetic pathway reconstitution of parthenolide involved in this parthenolide biosynthesis, mRNA was extracted from isolated flower trichomes and used for deep-sequencing to obtain a feverfew trichome EST database (Majdi et al., 2011). Subsequently, sequences of reported Asteraceae GAOs and COSs were used to blast against the feverfew EST database. Two sets of EST sequences with the highest homology to chicory GAO and chicory COS were assemble into two contigs from which full length open reading frames (ORF) were obtained. The expression of \( TpGAS \), putative \( TpGAO \), and putative \( TpCOS \) was profiled in feverfew during ovary development with real time RT-PCR. Those three genes showed similar patterns of expression, which was highest in stage 2 and stage 3 ovaries, and then decreased from stage 4 until virtually zero in stage 6 (Fig. 2B). Moreover, the expression pattern of \( TpGAS \), the putative \( TpGAO \), and the putative \( TpCOS \) is consistent

![Diagram](image-url)
with the accumulation profile of parthenolide in ovaries during flower development (Fig. 2): the parthenolide content increased from stage 2 to stage 5, and then decreased slightly. The content of its precursor, costunolide, increased in stage 2 and 3 and then decreased.

Most sesquiterpene-modifying P450s belong to the CYP71 subfamily (Ikezawa et al., 2011). Indeed, the putative feverfew TpGAO and TpCOS belong to this CYP71 subfamily. Identification of the expected parthenolide synthase (TpPTS) gene therefore was focussed on P450 sequences showing closest homology to the CYP71 class. Screening of the feverfew EST sequence database for putative TpPTS candidates identified twenty eight P450 contigs that belong to the CYP71 family and all have relatively high amino acid sequence similarity with TpCOS. To limit the number of candidate genes to be characterized for enzymatic activity, we compared the expression profiles of the candidate genes with TpGAS, and the putative TpGAO and TpCOS, assuming that TpPTS will have a similar expression pattern as the upstream genes. Three out of the twenty eight candidate genes - Tp2116, Tp4149, and Tp9025 - showed maximum expression in ovary development stage 2-4, similar as TpGAS, putative TpGAO and TpCOS and were therefore considered as TpPTS candidate genes (Fig. 2D). Costunolide and parthenolide levels decreased slightly after stage 4 (Fig. 2A), which suggests further metabolism of costunolide and parthenolide in these late stages. Indeed, one of the candidate genes (Tp8878) displayed increased expression after ovary development stage 4 (Fig. 2D) and was therefore considered as pathway-side branch candidate gene for costunolide and/or parthenolide conversion. The ORFs of putative TpGAO and TpCOS, three TpPTS candidates (Tp2116, Tp4149, Tp9025), and one pathway side branch candiate (Tp8878) were cloned into yeast expression vector pYED60 for expression and characterization.

**Functional characterization of parthenolide biosynthesis genes in yeast**

**TpGAO and TpCOS candidates:** To test the enzymatic function of the putative TpGAO, this gene was expressed in yeast together with the previously characterized gene TpGAS (Majdi et al., 2011), and crude yeast extracts were subsequently prepared and analyzed by GC-MS.
Figure 4. Functional identification of Tp8878 using a yeast microsome assay. (A) LC-Orbitrap-FTMS chromatogram at m/z=265.14344 (10 ppm, positive ionization mode) demonstrating that the compound produced by Tp8878 from parthenolide has identical retention time as a 3β-hydroxyparthenolide standard. EV, empty vector. (B) A 3β-hydroxyparthenolide standard and Tp8878 product have an identical fragmentation pattern. Grey boxes indicate the parent ion at [M+H]$^+ = 265.14344$. (C) LC-Orbitrap-FTMS chromatogram at m/z=249.14852 (10 ppm, positive ionization mode) demonstrating that the compound produced by Tp8878 from costunolide has identical retention time as a 3β-hydroxycostunolide standard. EV, empty vector. (D) A 3β-hydroxycostunolide standard and the Tp8878 product have an identical fragmentation pattern. Grey boxes indicate the parent ion at [M+H]$^+ = 249.14852$. (E) Molecular structures of costunolide, 3β-hydroxycostunolide, parthenolide, and 3β-hydroxyparthenolide.
for the presence of sesquiterpene lactones. Cells expressing TpGAS+TpGAO showed a clear GC-MS peak of elemantrien-12-oic acid, which is missing in cells expressing TpGAS alone. This compound is a cope-rearrangement product of germacr-1(10),4,11(13)-trien-12-oic acid (Supplementary Fig. 3), showing that the protein encoded by TpGAO is able to catalyze oxidation of germacrene A to germacr-1(10),4,11(13)-trien-12-oic acid (Fig. 1). To test the catalytic function of the putative TpCOS, the gene was co-expressed with TpGAS and TpGAO in yeast. Compared to the products produced by yeast cells expressing both TpGAS and CiGAO, the extracts of yeast cells expressing TpGAS+TpGAO+TpCOS showed a new GC-MS peak which was identified as costunolide, while the peak for germacr-1(10),4,11(13)-trien-12-oic acid was strongly reduced (Supplementary Fig. 3). Thus, it is confirmed that the protein encoded by the putative TpCOS gene is able to catalyze the conversion of germacr-1(10),4,11(13)-trien-12-oic acid to costunolide.

**Characterization of parthenolide synthase candidates**

To test the catalytic activity of TpPTS candidates, microsomes of yeast expressing Tp2116, Tp4149, and Tp9025 were isolated and incubated with costunolide. Compared to the microsomes from yeast transformed with the control construct, the microsomes from yeast expressing Tp2116 induced a new LC-MS peak that was unambiguously identified as parthenolide ([M+H]+ = 249, retention time and mass spectrum match with that of the parthenolide standard) (Fig. 3). An official name CYP71CA1 was assigned to this parthenolide synthase (TpPTS).

No new peaks were detected in the assays with microsomes isolated from yeast expressing Tp4149 or Tp9025.

To test the catalytic activity of Tp8878, a candidate assumed to be involved in a side branch of parthenolide biosynthesis, microsomes of yeast transformed with Tp8878 were isolated and incubated with costunolide. With parthenolide as a substrate, a new LC-MS peak was detected which was identified as 3β-hydroxyparthenolide ([M+H]+ = 265, retention time and mass spectrum matches that of the 3β-hydroxyparthenolide standard) (Fig. 4A and 4B). With costunolide as a substrate, also a new product peak was detected, which was identified as 3β-hydroxycostunolide ([M+H]+ = 249, retention time and mass spectrum matches that of a 3β-hydroxycostunolide standard) (Fig. 4C and 4D). An official name CYP71CB1 was assigned to this 3β-hydroxylase.

**Reconstitution of the parthenolide biosynthetic pathway in Nicotiana benthamiana**

With all the genes for the production of parthenolide available, we aimed to reconstitute the parthenolide biosynthetic pathway in the plant host N. benthamiana through transient heterologous gene expression. In addition, we tested the effect on product accumulation of the co-expression of the pathway genes together with a soluble Arabidopsis thaliana HMG-CoA reductase (AtHMGR) which can increase FDP substrate availability needed for the pathway. Hereto, the AtHMGR, TpGAS, TpGAO, TpCOS and TpPTS (CYP71CA1) coding sequences were cloned into the binary expression vector pBIN under the control of the Rubisco promoter (RBC). Agrobacterium tumefaciens was transformed with the various binary expression vectors and leaves were co-infiltrated with different combinations of the transformed A. tu-
Biosynthetic pathway reconstitution of parthenolide mefaciens strains to reconstitute the parthenolide biosynthetic pathway in *N. benthamiana* step by step.

**TpGAS and TpGAS+TpGAO:** Four days after infiltration with the *A. tumefaciens* strain carrying the RBC:TpGAS construct, the *N. benthamiana* leaves emitted the volatile compound germacrene A into their headspace (Table 1 and Supplementary Fig. 4). When leaves were co-infiltrated with two *A. tumefaciens* strains carrying the TpGAS and TpGAO constructs, respectively, germacrene A levels in the headspace were reduced by 90% compared to infiltration with TpGAS alone, suggesting that TpGAO can efficiently utilise germacrene A. However, no new product peaks were detected neither in the headspace nor in dichloromethane (DCM) extracts of the infiltrated leaves, indicating that the expected products of the TpGAO enzyme, germacr-1(10),4,11(13)-trien-12-oic acid, germacr-1(10),4,11(13)-trien-12-al and germacr-1(10),4,11(13)-trien-12-ol, are not stable *in planta* or are further metabolized into other products, analogous to our previous results obtained with the heterologous expressed CiGAO gene (Liu et al., 2011). Using high mass resolution LC-QTOF-MS in negative ionization mode, we checked whether any germacr-1(10),4,11(13)-trien-12-oic acid produced was possibly glycosylated within the *N. benthamiana* leaves, but accurate mass signals corresponding to the elemental formulae of the acid conjugated to either a hexose, a deoxyhexose or a pentose, or to combinations thereof, were not detected. We nevertheless decided to infiltrate the next gene of the pathway to see if the anticipated product and hence the substrate of the next enzyme was produced or not.

**TpGAS +TpGAO+TpCOS:** Co-infiltration of *A. tumefaciens* strains carrying the TpGAS, TpGAO and TpCOS expression constructs did result in the production of costunolide at four days post-infiltration (Table 1 and Supplementary Fig. 5A). The average production of costunolide was 9.6 ± 0.8 μg.g⁻¹ FW (n=8). No costunolide was detected in extracts from leaves upon transient expression of the empty vector (pBIN), neither in those of TpGAS alone, TpGAS+TpGAO or TpGAS+TpCOS, indicating that the production of costunolide in *N. benth-
amiana leaves is dependent on the presence of three genes: \( TpGAS, \) \( TpGAO \) and \( TpCOS \).

To investigate whether there were any other unexpected products formed in the infiltrated leaves, we performed untargeted LC-QTOF-MS analysis of leaf extracts. This resulted in the detection of two chromatographic peaks eluting at 22.24 and 22.48 min in leaves infiltrated with \( TpGAS+TpGAO+TpCOS \) that were absent in leaves infiltrated with \( TpGAS+Tp-\)GAO (Supplementary Fig. 5B-E). These two \( TpCOS \)-induced products were identified as the cysteine and glutathione (GSH) conjugates of costunolide, respectively.

\( TpGAS +TpGAO+TpCOS+TpPTS \) with boosting by \( AtHMGR \): No free parthenolide was detected in leaf extracts transiently expressing the four genes \( TpGAS+TpGAO+TpCOS+TpPTS \) (Table 1). In an attempt to boost the availability of substrate for the parthenolide pathway, we also co-expressed \( AtHMGR \). In combination with \( TpGAS+TpGAO+TpCOS \), this resulted in 4-fold increased costunolide production (Supplementary Fig. 6B). When \( AtHMGR \) was co-expressed with \( TpGAS+TpGAO+TpCOS+TpPTS \), free parthenolide was detected (2.05 ng.g\(^{-1}\) FW) four days after infiltration by MRM-LC-MS (Supplementary Fig. 6A). Moreover, two new LC-QTOF-MS peaks eluting at 17.74 and 18.53 min were detected, which were absent in leaves infiltrated with \( TpGAS+TpGAO+TpCOS+pBIN \) as control. The exact mass and comparison with standards identified these products as the cysteine and glutathione (GSH) conjugates of costunolide, respectively.

\( TpGAS +TpGAO+TpCOS+Tp8878 \) with boosting by \( AtHMGR \): To verify the function of \( Tp8878 \) (\( CYP71CB1 \)) in planta, \( A. \) \( tumefaciens \) strains with the parthenolide pathway constructs plus \( AtHMGR \) were infiltrated into \( N. \) \( benthamiana \) together with \( A. \) \( tumefaciens \) with the \( Tp8878 \) expression construct. Compared with the control (parthenolide pathway without \( Tp8878 \)) one new peak (RT=14.68 min) was detected that was identified as 3\( \beta \)-hydroxycostunolide-GSH (both retention time and mass spectrum match that of a 3\( \beta \)-hydroxycostunolide-GSH standard) (Table 1 and Supplementary Fig. 7A-B). A second new peak was identified as 3\( \beta \)-hydroxycostunolide-cysteine. GSH or cysteine conjugates of 3\( \beta \)-hydroxyparthenolide was not detectable in these samples. Compared with leaves infiltrated with only the costunolide pathway, about 50% of the original costunolide conjugates were converted into

### Table 2. IC 50 values (\( \mu \)M) of parthenolide and its conjugates acquired by sulforhodamine B viability test

<table>
<thead>
<tr>
<th></th>
<th>NCI-H460(^a)</th>
<th>NCI-H460/Rb(^b)</th>
<th>U87(^c)</th>
<th>U87-TxR(^d)</th>
<th>DLD1(^e)</th>
<th>DLD1-TxR(^f)</th>
<th>HT-29(^g)</th>
<th>HaCaT(^h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parthenolide</td>
<td>2.1</td>
<td>6.8</td>
<td>40.5</td>
<td>26.7</td>
<td>2.1</td>
<td>1.8</td>
<td>3.6</td>
<td>10.6</td>
</tr>
<tr>
<td>Parthenolide-Cysteine</td>
<td>76.8</td>
<td>46.9</td>
<td>107.7</td>
<td>105.0</td>
<td>37.0</td>
<td>33.1</td>
<td>17.3</td>
<td>65.7</td>
</tr>
<tr>
<td>Parthenolide-Glutathione</td>
<td>71.4</td>
<td>44.7</td>
<td>120.8</td>
<td>83.3</td>
<td>29.2</td>
<td>23.1</td>
<td>10.7</td>
<td>57.9</td>
</tr>
</tbody>
</table>

\(^a\) Sensitive non-small cell lung carcinoma cell line  
\(^b\) Multi-drug resistant non-small cell lung carcinoma cell line derived from its sensitive counterpart  
\(^c\) Sensitive glioblastoma cell line  
\(^d\) Multi-drug resistant glioblastoma cell line derived from its sensitive counterpart  
\(^e\) Sensitive colon carcinoma cell lines  
\(^f\) Multi-drug resistant colon carcinoma cell line derived from its sensitive counterpart DLD1  
\(^g\) Normal human keratinocyte
parthenolide when *TpPTS* was added to the infiltration mix. When *Tp8878* was co-infiltrated with the costunolide pathway together with *TpPTS*, about 40% of the costunolide conjugates were converted, yet the amount of parthenolide conjugates was decreased by 71% (Supplementary Fig. 7C). This is possibly the result of 3-hydroxylation of parthenolide by *Tp8878*, as we showed to occur in yeast microsomes (Fig. 4), but we were unable to detect any parthenolide-derived compounds.

**Anti-cancer activity of parthenolide conjugates in cell lines**

The anti-cancer effect of parthenolide GSH and cysteine conjugates was examined in 8 different human cell lines: both sensitive and multi-drug resistant lines of non-small cell lung carcinoma, glioblastoma and colon carcinoma cells as well as normal human keratinocytes (Table 2). Parthenolide-cysteine and parthenolide-GSH conjugates were less potent than free parthenolide: the concentrations necessary to inhibit cell growth by 50% (IC50 values) for conjugates were significantly higher than for free parthenolide in all tested cancer cell lines and normal human keratinocytes. Nevertheless, IC50 values of the conjugates for colon cancer cells are substantially lower than those for normal cells (HaCaT), indicating selectivity of both parthenolide conjugates towards colon carcinoma cells. The parthenolide-cysteine and parthenolide-GSH conjugates exerted the highest bioactivity in HT-29 cells (colon adenocarcinoma) with IC50s of 17.3 and 10.7 μM, respectively. The sensitivity to free or conjugated parthenolide was not affected by multi-drug resistance as the inhibitory profiles of the compounds were similar in both sensitive (DLD1) and resistant (DLD1-TxR) colon carcinoma cell lines (Table 2, Supplementary Fig. 8). Cysteine and GSH, when applied alone, had no influence on cell growth (Supplementary Fig. 8).

**Discussion**

The sesquiterpene lactone parthenolide from feverfew is a promising anti-cancer drug. The identification of feverfew *parthenolide synthase* (*TpPTS*) which uses costunolide as substrate confirms the hypothesis that parthenolide is derived from costunolide through epoxidation of the C4-C5 double bond (Liu et al., 2011). With the identified *germacrene A synthase* (*TpGAS*) (Majdi et al., 2011), *germacrene A oxidase* (*TpGAO*), costunolide synthase (*TpCOS*) and *TpPTS* we have isolated all structural genes of the biosynthetic pathway from the universal sesquiterpene precursor farnesyl diphosphate (FDP) (Majdi et al., 2011) up to parthenolide. Expression of these genes in the heterologous hosts *Nicotiana benthamiana* results in the formation of parthenolide plus a number of parthenolide conjugates, which may provide an attractive option for a more efficient and controlled production of this compound. The successful identification of the *TpPTS* gene shows that gene mining based on sequence similarity to related enzymes in combination with gene expression profiling is a good strategy to identify candidate genes involved in plant secondary metabolite pathways.

In the present study we showed that the production of parthenolide in a heterologous host plant species is feasible. No free parthenolide was detected in *N. benthamiana* leaves infiltrated with *TpGAS*+*TpGAO*+*TpCOS*+*TpPTS* by sensitive UPLC-MRM-MS, but when
AtHMGR was added to boost the supply of the precursor FDP, indeed a trace amount of free parthenolide (2.05 ng.g\(^{-1}\) FW) was detected. This low amount of free parthenolide was caused by the conjugation of the parthenolide produced towards both parthenolide-cysteine (1368.4 ng.g\(^{-1}\) FW) and parthenolide-GSH (87.5 ng.g\(^{-1}\) FW) conjugates. As costunolide, parthenolide and the hydroxylated products are cytotoxic, conjugation to GSH or cysteine may be part of a detoxification reaction of the N. benthamiana host cells. The cysteine-conjugates may be produced from the GSH-conjugate through the actions of peptidases (Marrs, 1996). As the conjugation to GSH is reversible (Heilmann et al., 2001) at physiological pH and the conjugation to cysteine is not, this would explain the relatively high levels of cysteine-conjugated products.

More than 90% of the total parthenolide produced in N. benthamiana was conjugated to either cysteine or GSH, while more than 95% of the parthenolide detected in the trichomes of feverfew was present as free parthenolide. TpGAS was found to be expressed much higher in the trichomes compared in the other tissues (Majdi et al., 2011). Trichome specific expression of a diterpene synthase in transgenic tobacco was recently reported (Ennajdaoui et al., 2010). To obtain higher production of free parthenolide in heterologous plants host, it would be a good option to try tissue specific expression in trichomes to prevent conjugation. An alternative host could be lettuce (Lactuca sativa) or chicory (Cichorium intybus). Both can produce costunolide and its derivatives, that are accumulating in specialized structures called laticifers throughout the plant (Hagel et al., 2008). Thus lettuce and chicory could potentially be used as a production platform for the heterologous production of parthenolide.

As water solubility is one of the major limiting factors for parthenolide being used as an...
anti-cancer drug (Shanmugam et al., 2006), obtaining more water-soluble parthenolide derivatives or analogues can be of interest. We have isolated Tp8878 and showed that it encodes a cytochrome P450 enzyme that can oxidise both costunolide and parthenolide to produce the more polar derivatives 3β-hydroxycostunolide and 3β-hydroxyparthenolide respectively (Fig. 4). Indeed, both compounds have also been detected in feverfew extracts (Fischedick et al., 2012). Hydroxylation makes these compounds more polar and may also allow additional enzymatic or chemical modifications to further improve water solubility. 3β-Hydroxyparthenolide has been shown to be active in the treatment of neurodegenerative disease (Fischedick et al., 2012), suggesting the additional hydroxyl group does not compromise its biological activity.

Previous studies have demonstrated the anti-cancer property of parthenolide in vitro, through induction of apoptotic cell death in a number of human cancer cell lines (Mathema et al., 2012). The depletion of intracellular GSH by parthenolide probably contributes to its apoptotic activity (Wen, Jing et al., 2002; Zhang et al., 2004), indicating that the anti-cancer effect of parthenolide involves interaction with GSH. Indeed, in our study, the parthenolide-GSH and parthenolide-cysteine conjugates showed less biological activity than free parthenolide in the cancer cell lines investigated. However, even though less effective in most cell lines, these conjugates showed quite high and selective activity against colon carcinoma cells and this feature could be an advantage in colon cancer treatment. Perhaps they act as a pro-drug in these cells, requiring biotransformation into free parthenolide to exert the anti-cancer effect.

The relative polarity (hydrophilicity) of the sesquiterpene lactones identified in this study can be deduced from their relative retention times in the C18 reverse phase LC-MS chromatograms (Fig. 5). Considering that poor water-solubility of parthenolide (and its oxidised derivatives) is a significant limitation for its application in cancer treatment (Sweeney et al., 2005) and that parthenolide-conjugates are selectively active against colon cancer cells, the conjugation of parthenolide and its oxidised derivatives could be a new strategic tool in drug development for cancer treatment.

In conclusion, the isolation of the genes encoding the entire parthenolide biosynthetic pathway will enable the industrial scale production of parthenolide in heterologous systems such as plants, yeast or other micro-organisms. The success of that will improve the availability of parthenolide - and parthenolide derivatives with improved chemical properties - and hence speed up the development of parthenolide-based anti-cancer drugs.

**Experimental Procedures**

Detailed description of the Gene expression analysis, Headspace analysis and Thermodesorption GC-MS, LC-QTOF-MS LC-Orbitrap-FTMS analysis of leaf extracts, parthenolide detection and quantification by LC-MRM-MS and Cysteine and glutathione (GSH) conjugation can be found in the supplementary data.

**Isolation and cloning of full length candidate genes from feverfew**
An EST library constructed from mRNA isolated from feverfew (Tanacetum parthenium) trichomes as reported before (Majdi et al., 2011) was used for gene isolation. For TpGAO and TpCOS candidates, one contig was identified for each gene by sequence homology to known GAOs (LsGAO, GU198171; CiGAO, GU256644; SlGAO, GU256646; HaGAO, GU256645; BsGAO, GU256647) and COSs (LsCOS, AEI59780; and CiCOS, AEG79727). Twenty-eight parthenolide synthase candidate cytochrome P450 contigs were identified by sequence homology to known sesquiterpene monoxygenases of the CYP71 subfamily. Four of them, Tp2116, Tp4149, Tp9025 and Tp8878, were selected for functional characterization based on their expression profile during ovary development. RACE PCR (Clontech) was used to obtain the sequence of the 5'- and 3'-region of all candidate contigs. Full length cDNAs of candidate genes were amplified from feverfew cDNA with the addition of NotI/PacI restriction sites. The cDNAs were subsequently cloned into the yeast expression vector pYEDP60 (Pompon et al., 1996) and sequenced. The cDNA sequences of all candidates genes have been deposited in GenBank: TpGAO, KC964544; TpCOS, KC964545, Tp8878, KC954153; Tp9025, KC954154; Tp2116 (TpPTS), KC954155; Tp4149, KC954156. The sequences were also submitted to David Nelson's cytochrome P450 homepage (http://drnelson.uthsc.edu/cytochromeP450.html), Tp2116 and Tp8878 were assigned the name as CYP71CA1 and CYP71CB1, respectively (Nelson, 2009).

**Plasmid construction for gene expression in yeast**

TpGAO, TpCOS and three parthenolide synthase candidates (Tp2116, Tp4149, Tp9025) were cloned into pYED60 vector using NotI/PacI restriction sites. The obtained constructs were named TpGAO::pYED60, TpCOS::pYED60, Tp2116::pYED60, Tp4149::pYED60, and Tp9025::pYED60. TpGAS::pYES3 (Liu et al., 2011) plus TpGAO::pYED60 and TpGAS/TpGAO::pESC-Trp plus TpCOS::pYEDP60 were co-transformed into the WAT11 (Urban et al., 1997) yeast strain. After transformation yeast clones containing both plasmids were selected on SD minimal medium supplemented with amino acids, but omitting uracil, adenine sulphate and L-tryptophane for auxotrophic selection of transformants.

**Yeast in vitro microsome assay**

The procedure of the yeast microsome isolation is described in detail in the supplementary data. For in vitro microsome assays, 72 μl isolated microsomal fractions, 10 μl substrate (of a 10mM stock in DMSO), 100 μl NADPH (of a 10mM stock in 100mM potassium buffer), 20 μl potassium buffer (1M, pH7.5), and 288 μl water were mixed and incubated for 2.5 h at 25 °C with shaking (200 rpm). Then the mixture was centrifuged at 12000 rpm for 10 min. The supernatant was filtered through an 0.22 μm filter before analysis of the products by LC-Orbitrap-FTMS (for details see supplementary data).

**Plasmid construction and transient expression in N. benthamiana**

For transient expression in N. benthamiana, TpGAS, TpGAO, TpCOS, Tp8878, and 3 parthenolide synthase candidates (Tp2116, Tp4149, Tp9025) were cloned into ImpactVector1.1 (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwit-
Biosynthetic pathway reconstitution of parthenolide

ten-door-planten.htm) to express them under the control of the Rubisco (RBC) promoter (Outchkourov et al., 2003). TpGAS was also cloned into ImpactVector1.5 to fuse it with the RBC promoter and the CoxIV mitochondrial targeting sequence as we have demonstrated before that mitochondrial targeting of sesquiterpene synthases results in improved sesquiterpene production (2011). 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. *A. tumefaciens* infiltration (agro-infiltration) for transient expression in N. benthamiana was performed as described by Liu et al. (2011). After infiltration the plants were grown for another four and half days and then harvested for analysis.

Acknowledgements

We thank Francel Verstappen for assistance with MRM-LC-MS analysis, Bert Schipper for assistance with LC-QTOF-MS and LC-Orbitrap-MS analysis, Miriam Goedbloed for assistance with making expression constructs, Justin T. Fischedick for providing standards isolated from feverfew. This project was funded by the European Commission (EU-project TerpMed, Plant Terpenoids for Human Health: a chemical and genomic approach to identify and produce bioactive compounds, Grant agreement no.: 227448). Ric C.H. de Vos acknowledges additional funding by the Netherlands Metabolomics Centre and the Centre for BioSystems Genomics, both of which are part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

References


Biosynthetic pathway reconstitution of parthenolide


SUPPLEMENTARY INFORMATION

Supplementary methods

Gene expression analysis
mRNA levels of candidates genes were measured by quantitative real time RT-PCR. Total RNA was extracted from ovaries of feverfew flowers collected at developmental stages 2 to 6 and the corresponding cDNAs were obtained as previously reported (Majdi et al., 2011). Real time RT-PCR was performed using a LightCycler 480 (Roche Diagnostics). All real time RT-PCR reactions were carried out in a total volume of 20 μl mastermix containing 10 μl LightCycler 480 SYBR Green I Master (Roche Diagnostics), 0.6 μl forward primer (0.3 μM), 0.6 μl reverse primer (0.3 μM), 6.8 μl water and 2 μl cDNA (50 ng). The LightCycler experimental run protocol used was: 95ºC for 10 min, 95ºC for 10s, 60ºc for 30s for 40 cycles and finally a cooling step to 40ºC. LightCycler Software 1.5.0 was used for data analysis. For efficiency determination, a standard curve of six serial dilution points (ranging from 200 to 6.25 ng) was made 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: forward TpGAO 5’-TGCAGCTCCCGCTTGCTAATATAC-3’, reverse TpGAO 5’-AGTCTTTTCTTTGAACCGTGCTCC-3’, forward TpCOS 5’-TAGCTTCATCCCGGAGCGATTTGA-3’, reverse TpCOS 5’-AAATTCTTTCGGCCCGCACAAATG-3’, forward TpPTS 5’-AGACATTACGTGTACCCCTCCCG-3’, reverse TpPTS 5’-ATCACGACACAAAGTCCCGAGGAAA-3’. Quantification of transcript levels was done in three independent biological replicates and for each biological replicate three technical replicates were run. Actin was used as a housekeeping gene. The ΔCT was calculated as follows: ΔCT= CT (Target)- CT (Actin).The fold change value was calculated using the expression $2^{-\Delta CT}$ (Schmittgen & Livak, 2008).

Headspace analysis and thermodesorption GC-MS
Volatile collection and GC-MS analysis were performed as described before (2011). Steel sorbent cartridges (89 mm × 6.4 mm O.D.; Markes) containing Tenax were used for volatile collection. N. benthamiana leaves were cut from the plant and their petioles placed in water in a small vial and were enclosed in a glass cuvette. To trap the leaf-produced volatiles, air was sucked through the cuvettes at a flow rate of 90 mL min$^{-1}$ for 2 h through a Tenax containing cartridge. A second cartridge was used to purify the incoming air. Sample cartridges were dried for 15 min at room temperature with a nitrogen flow of 20 psi before GC-MS analysis.

For GC-MS analysis, cartridges were placed in an automated thermodesorption unit (Ultra; Markes, Llantrisant, UK) and were flushed with helium at 50 mL min$^{-1}$ for 2 min to remove moisture and oxygen before thermodesorption. Volatiles in the cartridges were desorbed by heating at 220 °C for 5 min with a helium flow of 50 mL min$^{-1}$. The compounds released were trapped on an electrically cooled sorbent trap (Unity; Markes, Llantrisant) at a temperature of 5 °C. The trapped volatiles were injected on the analytical column (ZB-5MSI,
30 m × 0.25 mm ID, 1.0 μm – film thickness, Zebron, Phenomenex) by ballistic heating of the cold trap to 250 °C for 3 min. The temperature program of the GC started at 40 °C (3 min hold) and rose 10 °C min⁻¹ to 280 °C (2 min hold). The column effluent was ionised by electron impact (EI) ionisation at 70 eV. Mass scanning was done from 33 to 280 m/z with a scan time of 4.2 scans s⁻¹. Data analysis was done using Xcalibur (Thermo, USA) to identify compounds by comparing mass spectra with those of MS libraries such as NICT and the mass spectra and retention times of authentic reference standards.

LC-QTOF-MS analysis of leaf extracts

For LC-QTOF-MS analysis of leaf extracts, 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) operating in negative ionization mode was used. An analytical column (Luna 3 μ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA) attached to a C18 pre-column (2.0 × 4 mm; Phenomenex, USA) was used. Degassed eluent A [ultra-pure water: formic acid (1000:1, v/v)] and eluent B [acetonitril:formic acid (1000:1, v/v)] were used at a flow rate of 0.19 mL min⁻¹. Masses were recorded between m/z 80 and m/z 1500; leucine enkaphalin ([M-H]⁻=554.2620) was used as a lock mass for on-line accurate mass correction. The gradient of the HPLC started at 5% eluent B and increased linearly to 75% eluent B in 45 min, after which the column was washed and equilibrated for 15 min before next injection. Injection volume was 5 μl.

LC-Orbitrap-FTMS analysis of yeast microsome assay products

To analyse microsome assay mixtures, 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 was used. Chromatographic separation took place on an analytical column (Luna 3 μ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA). Degassed eluent A [ultra-pure water: formic acid (1000:1, v/v)] and eluent B [acetonitril:formic acid (1000:1, v/v)] were used at a flow rate of 0.19 mL min⁻¹. A linear gradient from 5 to 75% acetonitrile in 45 min was applied, which was followed by 15 min of washing and equilibration. FTMS full scans (m/z 100–1200) were recorded with a resolution of 60000, whereas for MSn scans a resolution of 15000 was used. The FTMS was externally calibrated in negative mode using sodium formate clusters in the range m/z 150–1200, and automatic tuning was performed on m/z 384.93. Injection volume was 5 μl.

Parthenolide detection and quantification by LC-MRM-MS

Targeted analysis of parthenolide and costunolide was performed with a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described before(2011). Chromatographic separation was obtained on an Acquity UPLC BEH C18 column (100 × 2.1 mm, 1.7 μm; Waters) by applying a water/acetonitrile gradient to the column, starting from 5% (v/v) acetonitrile in water for 1.25 min and rising to 50% (v/v) acetonitrile in water in 2.35 min, followed by an increase to 90% (v/v) acetonitrile in water in 3.65 min, which was maintained for 0.75 min
before returning to 5% acetonitrile in water using a 0.15 min gradient. Finally, the column was equilibrated for 1.85 min using this solvent composition. Operation temperature and flow rate of the column were 50 °C and 0.5 mL min⁻¹, respectively. Injection volume was 5 μL. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L h⁻¹, respectively. The capillary voltage was 3.0 kV, the source temperature 150 °C, and the desolvation temperature 650 °C. The cone voltage was optimized for parthenolide or costunolide using the Waters IntelliStart MS Console. Argon was used for fragmentation by collision-induced dissociation in the ScanWave collision cell. Multiple Reaction Monitoring (MRM) was used for identification and quantification of costunolide in yeast extract and agro-infiltrated N. benthamiana leaves by comparing retention times and MRM mass transitions with that of a costunolide standard. MRM transitions for costunolide [M+H]⁺=233.16>131.01 and [M+H]⁺=233.16>187.23, and for parthenolide [M+H]⁺=249.16>231.229 and [M+H]⁺= 249.16>185.224 were optimized using the Waters IntelliStart MS Console.

Yeast microsome isolation

Microsomes of yeast transformed with parthenolide synthase candidates were isolated as described by Pompon et al.(1996) with modifications. Yeast cell cultures were grown in 50mL SGI medium for 36 hours at 30 °C. After adding 250 mL YPL medium containing 2% galactose, induction was allowed to proceed for 24 hours at 30 °C. Cells were collected and chilled on ice for 20 min. After centrifugation at 4,900 x g for 10 min, 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 incubated for 10 min at room temperature. Following centrifugation at 4,900 x g for 10 min, the pellet was washed three times with extraction buffer (without resuspending the pellet). Cells were re-suspended in 3 mL extraction buffer (without β-mercaptoethanol) and transferred into a 50 mL Falcon tube. About 25 mL of glass beads (450-500 μm) were used to lyse the cells by shaking for 10 min in a cold room. The lysed cells were transferred to a 25 mL centrifuge tube and centrifuged at 10,500 x g for 10 min. The supernatant was centrifuged at 195,000 x g for 2 h. The pelleted microsomal fractions were re-suspended in 4 mL chilled 50 mM Tris-HCl buffer pH 7.5 containing 1 mM EDTA and 20% (v/v) glycerol using a chilled glass Tenbroeck homogenizer. Yeast microsomes were then aliquoted in pre-cooled 1.5 mL eppendorf tubes and stored at -80 °C until use.

Cysteine and glutathione (GSH) conjugation

Cysteine and GSH conjugation was performed as described by Liu et al.(2011). The reaction mixture contains GSH (1 mM) or cysteine (1 mM), substrates (0.2 mM), and GST (0.7 mg L⁻¹; Sigma, USA) in 1000 μl potassium buffer (100 mM; pH 6.5). Complete assay mixtures without GST enzyme or either of the substrates were used as controls. After incubation for 30 min at room temperature, samples were kept at -20 °C until analysis. Costunolide was purchased from TOCRIS Bioscience (United Kingdom). Parthenolide, 3β-hydroxycostunolide, and 3β-hydroxyparthenolide, isolated from dried aerial parts of feverfew plants, were provided by Justin T. Fischedick of PRISNA (Leiden, the Netherlands)(Fischedick et al., 2012).
Cytotoxicity assay

The NCI-H460, U87, DLD1 and HT-29 cell lines were purchased from the American Type Culture Collection (ATCC). Human normal keratinocytes (HaCaT) were obtained from Cell Lines Service (CLS). NCI-H460/R cells were selected originally from NCI-H460 cells and cultured in a medium containing 100 nM doxorubicin. U87-TxR and DLD1-T/R cells were selected from U87 and DLD1 cells, respectively, and cultured in a medium containing 300 nM paclitaxel (Urban et al., 1997). All cell lines were sub-cultured at 72 h intervals using 0.25% trypsin/EDTA and seeded into a fresh medium at the following densities: 8,000 cells/cm² for NCI-H460, DLD1, DLD1-T/R and HT-29, 16,000 cells/cm² for U87 and NCI-H460xR, and 32,000 cells/cm² for U87-T/R and HaCaT.

Cells grown in 25 cm² tissue flasks were trypsinized, seeded into flat-bottomed 96-well tissue culture plates, and incubated overnight. Treatment with all compounds (1-100 μM) lasted 72 h. The cellular proteins were stained with Sulforhodamine B (SRB), following a slightly modified protocol (Outchkourov et al., 2003). The absorbance after SRB staining was measured at 540 nm using an automatic microplate reader (LKB5060-006-Micro-PlateReader-Vienna-Austria).
Supplementary Figure 1. Phylogenetic tree of TpGAO, TpCOS, parthenolide synthase candidates (Tp2116, Tp4149, Tp9025, Tp8878) and other GAOs and COSs of the Asteraceae. Bootstrap values are shown as percentage from 1,000 replicates.

Supplementary Figure 2. Alignment of deduced amino acid sequence of TpGAO (germacrene A oxidase), TpCOS (costunolide synthase), and other GAOs and COSs from Asteraceae species. Amino acid sequences were obtained from cDNAs deposited at NCBI. Germacrene A oxidase from Tanacetum parthenium (TpGAO, KC964544), from Lactuca sativa (LsGAO; GU198171), Cichorium intybus (CiGAO; GU256644), Sussurea lappa (SlGAO; GU256646), Helianthus annuus (HaGAO; GU256645), and Barnadesia spinosa (BsGAO; GU256647); costunolide synthase from Tanacetum parthenium (TpCOS, KC964544), from Lactuca sativa (LsCOS; AEI59780), and Cichorium intybus (CiCOS; AEG79727).
Supplementary Figure 3. Functional characterisation of TpGAO and TpCOS in yeast. (A) GC-MS chromatograms showing the metabolites produced by yeast transformed with the indicated genes. Line a, negative control; line b, the yeast transformed with TpGAS+TpGAO; line c, yeast transformed with TpGAS+TpGAO+TpCOS. 1 = germacra-1(10),4,11(13)-trien-12-oic acid; 2 = costunolide. (B) The mass spectra of compound 1 (a) and compound 2 (b) produced by yeast, and elematrien-12-oic acid (c) and costunolide standards (d). (C) Cope rearrangement of germacra-1(10),4,11(13)-trien-12-oic acid to elematrien-12-oic acid.

Supplementary Figure 4. Germacrene A production in Nicotiana benthamiana infiltrated by pBIN (empty vector), TpGAS+pBIN, and TpGAS+TpGAO. Volatiles emitted from N. benthamiana leaves infiltrated with the indicated genes were measured by GC-MS. TpGAS, Tanacetum parthenium germacrene A synthase; TpGAO, Tanacetum parthenium germacrene A oxidase.
Supplementary Figure 5. Pathway reconstitution of costunolide biosynthesis in *N. benthamiana*. (A) UPLC-MRM-MS analysis of free costunolide (MRM transitions 233.16>187.232 and 233.16>131.005) in *N. benthamiana* leaves infiltrated with TpGAS+TpGAO+TpCOS and TpGAS+TpGAO+pBIN. (B) LC-QTOF-MS chromatograms (ion selection at m/z 306) of the costunolide-glutathione conjugate formed in an enzyme assay of costunolide and glutathione with glutathione-S transferase, and of extracts of *N. benthamiana* leaves agro-infiltrated with TpGAS+TpGAO+TpCOS and TpGAS+TpGAO+pBIN (control). (C) In source MS spectrum of peak 22.48 and costunolide-glutathione conjugate (RT=22.52 min). Grey boxes indicate the parent ion at [M-H]=538.20. (D) LC-QTOF-MS chromatograms (ion selection at m/z 120) of the costunolide-cysteine conjugate formed in an enzyme assay of costunolide and cysteine with glutathione-S transferase, and of extracts of *N. benthamiana* leaves agro-infiltrated with TpGAS+TpGAO+TpCOS and TpGAS+TpGAO+pBIN (control). (E) In source MS spectrum of peak 22.24 and costunolide-cysteine conjugate (RT=22.30 min). Grey boxes indicate the parental ion at [M-H]=352.14.
Supplementary Figure 6. Pathway reconstitution of parthenolide biosynthesis in *N. benthamiana*. (A) UPLC-MRM-MS analysis of parthenolide (MRM transitions 249.16>231.229 and 249.16>185.224) in *N. benthamiana* leaves infiltrated with *AtHMGR+P19+TpGAS+TpGAO+TpCOS+TpPTS*. (B) Costunolide concentration in *N. benthamiana* leaves infiltrated with *P19+TpGAS+TpGAO+TpCOS*, *AtHMGR+P19+TpGAS+TpGAO+TpCOS*, *P19+TpGAS+TpGAO+TpCOS+TpPTS* and *AtHMGR+P19+TpGAS+TpGAO+TpCOS+TpPTS*. (C) LC-QTOF-MS chromatograms (ion selection at m/z 306) of *N. benthamiana* leaves agro-infiltrated with *AtHMGR+p19+TpGAS+CiGAO+CiCOS+TpPTS*, the parthenolide-glutathione conjugate formed in an enzyme assay of parthenolide and glutathione with glutathione-S transferase, and of an extract of *N. benthamiana* leaves agro-infiltrated with *AtHMGR+p19+TpGAS1.5+CiGAO+CiCOS+pBIN*. (D) In source MS spectrum of peak 18.53 and the parthenolide-glutathione conjugate (RT=18.53 min). Grey boxes indicate the parent ion at [M-H]-=554.20. (E) LC-QTOF-MS chromatograms (ion selection at m/z 120) of *N. benthamiana* leaves agro-infiltrated with *AtHMGR+p19+TpGAS+CiGAO+CiCOS+TpPTS*, the parthenolide-cysteine conjugate formed in an enzyme assay of parthenolide and glutathione with glutathione-S transferase and of an extract of *N. benthamiana* leaves agro-infiltrated with *AtHMGR+p19+TpGAS1.5+CiGAO+CiCOS+pBIN*. (F) In source MS spectrum of peak 18.53 and parthenolide-cysteine conjugate (RT=18.53 min). Grey boxes indicate the parent ion at [M-H]-=368.15.
Supplementary Figure 7. Functional characterisation of Tp8878 in *N. benthamiana*. (A) LC-QTOF-MS chromatograms (ion selection at \textit{m/z} 306) of *N. benthamiana* leaves agro-infiltrated with *AtHMGR*p19+TpGAS+CiGAO+CiCOS+TpPTS+Tp8878, the 3β-hydroxycostunolide-glutathione conjugate formed in an enzyme assay of 3β-hydroxycostunolide and glutathione with glutathione-S transferase, and of an extract of *N. benthamiana* leaves agro-infiltrated with *AtHMGR*p19+TpGAS1.5+CiGAO+CiCOS+TpPTS or with *AtHMGR*p19+TpGAS1.5+CiGAO+CiCOS+pBIN as controls. (B) In source MS spectrum of peak 14.68 and 3β-hydroxycostunolide-glutathione conjugate (RT=14.72 min). Grey boxes indicate the parent ion at [M-H]−=554.20. (C) Relative content of costunolide conjugates (glutathione and cysteine) in *N. benthamiana* agro-infiltrated with *AtHMGR*p19+TpGAS+CiGAO+CiCOS+TpPTS, *AtHMGR*p19+TpGAS+CiGAO+CiCOS, and *AtHMGR*p19+TpGAS+CiGAO+CiCOS+TpPTS+Tp8878 analysed by LC-QTOF-MS.
Supplementary Figure 8. Anti-cancer activity of parthenolide, parthenolide-cysteine and parthenolide-GSH. Displayed is the cell growth inhibition by parthenolide, cysteine (CYS), glutathione (GSH), parthenolide- CYS, and parthenolide-GSH in (A) DLD1 colorectal carcinoma cell line; (B) DLD1-TxR multi-drug resistant cancer cell line derived from DLD1; (C) HT-29 colon carcinoma cell line and (D) in HaCaT normal human keratinocytes. The data are averages of five independent experiments (n=5) ± S.D.

References


Chapter 5

Kauniolide synthase, a cytochrome P450 terpenoid cyclase that catalyses the first step in guaianolide sesquiterpene lactone biosynthesis

Qing Liu, David Manzano Irini Pateraki, Lea Ricard, Ric de Vos, Sander van de Krol, Harro Bouwmeester

In preparation for submission.
Abstract

Sesquiterpene lactones are terpenoids with a skeleton of 15 carbons, which are important for human as flavourings and fragrances and possess a wide range of pharmacological activities, including anticancer, anti-inflammatory, antifungal and anti-bacterial activities. The core skeletons of sesquiterpene lactones are classified as guaianolide, pseudoguaianolide, germacranolide, eudesmanolide or helenanolides. However, how these sesquiterpene lactone are made remains unknown. The putative first committed step in guaianolide biosynthesis has been hypothesised to be the epoxidation or hydroxylation of the branch point germacranolide sesquiterpene lactone, costunolide, followed by a cyclisation reaction. Here we used metabolic profiling, 454 sequencing of a glandular trichome cDNA library and expression profiling of P450 genes to identify the enzyme catalysing this reaction in feverfew, *Tanacetum parthenium*. The cytochrome P450, kauniolide synthase (TpKS), catalyses the formation of the guaianolide sesquiterpene lactone, kauniolide from costunolide, through the intermediate 3β-hydroxycostunolide. We hypothesise that the enzyme catalyses the formation of kauniolide from costunolide via hydroxylation at C₃. Subsequent dehydration at C₃ induces double bond migration from C₄-C₅ to C₃-C₄. Attack of the resulting C₅ carbocation by the C₁-C₁₀ double bond subsequently results in cyclisation. The full biosynthetic pathway of kauniolide from the general sesquiterpene precursor FPP was reconstituted in *Nicotiana benthamiana* by transient co-expression of all the biosynthetic genes. In contrast to most P450s that only oxidise their substrates, TpKS oxidises its substrate and cyclises it, making it a special type of cytochrome P450.
Introduction

Sesquiterpene lactones (STLs) are a major class of plant secondary metabolites, characterized by their α-methylene γ-lactone moiety on the 15-carbon core backbone (Picman, 1986). Many of these colourless, bitter tasting, lipophilic molecules form the active constituents of extracts from a variety of medicinal plants used in traditional medicine (Rodriguez et al., 1976; Zhang et al., 2005). Bioactivity of STLs varies from anti-inflammatory (Lyss et al., 1998) and anti-cancer (Koo et al., 2001) to anti-malarial (Klayman, 1985), all of which are beneficial to human health. The majority of STLs have been reported in the Asteraceae, with over 4000 different STLs that have been identified (de Kraker et al., 2002). Although all these STLs differ in the details of their structure, the backbone of all of them consists of a limited set of core skeletons that are classified as guaianolide, pseudoguaianolide, germacranolide, eudesmanolide or helenanolide (Picman, 1986; Fischer, 1990; Neerman, 2003) (Fig. 1a).

Guaianolides consist of a bicyclic 5,7-ring and have been reported in a variety of plants and other organisms, but mainly in the Asteraceae and Apiaceae (Simonsen et al., 2013). Guaianolides represent a large number of naturally occurring STLs that have been used in traditional medicine throughout the history of mankind for treating various diseases (Schall & Reiser, 2008). The guaianolides exhibit anti-tumor, anti-schistosomal, anthelmintic, antimicrobial, contraceptive, root-growth stimulatory, anti-feedant, and germination inhibitory activities (Rodriguez et al., 1976; Drew et al., 2009; K & Merillon, 2013). Thapsigargin, for example, induces apoptosis in mammalian cells and eventually leads to the death of the cell (Søhoel et al., 2006). A thapsigargin derived drug is currently undergoing clinical trials (phase II) for the treatment of breast, kidney and prostate cancer (Denmeade et al., 2003; Janssen et al., 2006). Another guaianolide, 11,13-dehydro compressanolide from feverfew (Tanacetum parthenium), has been reported to have activity against leishmaniasis, which is one of the major infectious diseases affecting the poorest regions of the world (da Silva et al., 2010). Several other bioactive guaianolides have also been reported from feverfew and chicory (Cichorium intybus), such as artecanin, tanaparthin-β-peroxide, leucodin and dehydrocostus lactone (de Kraker et al., 2002; Fischedick et al., 2012).

Despite the importance of guaianolides, their biosynthesis has not yet been elucidated. The biogenesis of guaianolides presumably proceeds through costunolide, as a guaianolide, leucodin, was found to be formed by the incubation of costunolide with chicory root enzyme extract (de Kraker et al., 2002). The most basic guaianolide and putative precursor of leucodin, likely is kauniolide (Fig. 1b). However, it is not clear whether kauniolide is directly derived from costunolide or through parthenolide and two alternative kauniolide biosynthesis pathways have been suggested. One proceeds via parthenolide through an annular cyclization of the ring by a cyclase, followed by an elimination reaction by a guaianolide dehydroxylase (Qi et al., 1995; Zhai et al., 2012) (Figure 1b, presumed route 1). Alternatively, the pathway proceeds through direct enzymatic hydroxylation at C3 of costunolide, resulting in the intermediate 3β-hydroxycostunolide, while the guaianolide skeleton is subsequently formed via C1,C5-cyclization by a germacrene cyclase (de Kraker et al., 1998) (Figure 1b, presumed route 1).
Figure 1. (a), Core skeleton of guaianolides, pseudoguaianolide, germacranolide, eudesmanolide and helenanolide sesquiterpene lactones (STLs), and examples of some members of these STL classes. (b), Presumed biosynthetic pathway of guaianolide type sesquiterpene lactones in plants. GAS, germacrene A synthase; COS, costunolide synthase; PTS, parthenolide synthase. Costunolide is formed by conversion of farnesyl diphosphate (FDP) to germacrene A, catalysed by (+)-germacrene A synthase (GAS) (de Kraker et al., 1998). Subsequently, germacrene A is converted to germacra-1(10),4,11(13)-trien-12-oic acid by the cytochrome P450 enzyme, germacrene A oxidase (GAO) (de Kraker et al., 2001). Genes encoding GAS and GAO have been isolated from several Asteraceae species (Bennett et al., 2002; Bouwmeester et al., 2002; Bertea et al., 2006; Nguyen et al., 2010; Cankar et al., 2011; Liu et al., 2011). Germacra-1(10),4,11(13)-trien-12-oic acid is subsequently oxidised by costunolide synthase (COS) to 6α-hydroxy-germacra-1(10),4,11(13)-trien-12-oic acid, which undergoes spontaneous lactone ring formation to yield costunolide (Ikezawa et al., 2011; Liu et al., 2011). Costunolide is thought to be precursor for both guaianolide and eudesmanolide types of sesquiterpene lactones. Two alternative pathways have been suggested for the formation of the first basic guaianolide skeleton, kauniolide (Piet, 1996).
Kauniolide synthase, a cytochrome P450 cyclase

2). We have recently demonstrated that feverfew has a cytochrome P450 enzyme that can catalyse this C3-hydroxylation (Chapter 4).

Here we set out to determine whether kauniolide is derived from parthenolide or costunolide by identifying and isolating the gene(s) involved in biosynthesis of the guaianolide skeleton from feverfew (Tanacetum parthenium). We previously described the characterisation of the feverfew genes leading to the production of costunolide (Chapter 3) and parthenolide (Chapter 4), both of which are germacraneolide STLs (de Kraker et al., 2002). As several guaianolides have been reported in feverfew (Bork et al., 1997; Pareek et al., 2011; Fischedick et al., 2012) we assumed that a guaianolide synthase should be present in feverfew. We assumed that this reaction should be catalysed by a cytochrome P450. To identify this P450, we determined the expression profile of a series of P450 gene candidates - that we retrieved from our feverfew trichome cDNA library - and compared their expression profile with the STL accumulation profile in developing ovaries. The enzyme encoded by one of these candidates, Tp8879, catalysed the conversion of costunolide, but not parthenolide, into the guaianolide kauniolide. This shows that the branch point in the biosynthesis of guaianolide and germacraneolide type sesquiterpene lactones is the ubiquitous STL costunolide. The full kauniolide pathway encoded by farnesyl diphosphate synthase (FPS), germacrene A synthase (GAS), germacrene A oxidase (GAO), costunolide synthase (COS) and kauniolide synthase was successfully reconstituted by transient expression in Nicotiana benthamiana. As kauniolide forms the basic skeleton from which other guaianolides are likely derived, the identification of kauniolide synthase provides the basis for metabolic engineering approaches for the production of a range of different medicinal guaianolides.

Results

Identification and cloning of guaianolide skeleton biosynthesis candidate genes

Previously we have shown that the genes encoding the enzymes of the costunolide and parthenolide biosynthesis pathway, TpGAS, TpGAO, TpCOS and TpPTS, all have a similar expression pattern during ovary development, and that this expression pattern closely follows the accumulation of both costunolide and parthenolide in ovaries respectively (Figure 2a,b,c; Chapter 4). The genes involved in biosynthesis of germacrene A acid (Nguyen et al., 2010), costunolide (Ikezawa et al., 2011; Liu et al., 2011) and parthenolide (Chapter 4) all encode P450 enzymes (TpGAO, TpCOS and TpPTS) of class CYP71 and share a relatively high amino acid sequence similarity in feverfew (Chapter 4). We assumed that the enzyme acting on either parthenolide or costunolide to form kauniolide would also be a P450 of class CYP71 and that the expression pattern of this gene may show similar profile as TpGAO, TpCOS and TpPTS (Figure 2a). From the 59 class CYP71 P450 sequences in the feverfew trichome cDNA sequence database (Majdi et al., 2011) the expression profile over different ovary developmental stages was assessed (Chapter 4). The basic guaianolide kauniolide was not detected in the ovary extracts, presumably because it is rapidly converted to artecanin, which is detected in feverfew ovaries (Figure 2c). The accumulation profile in developing ovaries of artecanin
differs from that of costunolide and parthenolide: both costunolide and parthenolide levels decline at later stages of ovule development, which may be a reflection of the reduced gene expression level at later stages of ovule development. In contrast, the artecanin levels remain increasing throughout ovule development suggesting that expression of genes involved in the biosynthesis of guaianolide-like compounds persists throughout ovule development. Such expression profile was identified in candidate gene *Tp8879* (Figure 2d). The full sequence of this gene was retrieved and used to design primers to isolated the full length ORF for cloning into yeast expression vector pYED60 for enzymatic characterization.

**Functional characterization of candidate *Tp8879* in yeast**

To test the catalytic activity of candidate *Tp8879*, microsomes of yeast expressing the *Tp8879* ORF were isolated and incubated with costunolide or parthenolide. Analysis of the products formed from costunolide by LC-Orbitrap-FTMS showed a small new peak ([M+H]+=231.13780) not present in control samples. The mass of this new compound is 2 D
Kauniolide synthase, a cytochrome P450 cyclase

less than that of costunolide ([M+H]+=233.15244), suggesting introduction of a double bond or ring closure in costunolide. From previous experience we know that some compounds may escape detection by LC-Orbitrap-FTMS, because of low ionisation. Ionisation may be improved by conjugation of the compound to be detected to cysteine ([M]=121.01464), which for many sesquiterpene lactones can be achieved non-enzymatically and irreversibly by add-

Figure 3. Identification of kauniolide in yeast microsome assays. (a) LC-Orbitrap-FTMS chromatograms at \(m/z=352.15771\) (10 ppm, positive ionization mode, mass for kauniolide-cysteine conjugate) of extracts of microsomes expressing empty vector fed with costunolide and incubated afterwards with cysteine (control) (top panel), extracts of microsomes expressing Tp8879 fed with costunolide and incubated afterwards with cysteine (2\textsuperscript{nd} panel), kauniolide-cysteine conjugate formed in an enzyme assay of kauniolide and cysteine with glutathione-S transferase (3\textsuperscript{rd} panel), co-injection of the latter two (bottom panel). (b) MS spectrum of peak RT=24.45 min for kauniolide-cysteine conjugate (top panel), cysteine conjugated compounds in microsomes expressing Tp8879 fed with costunolide (middle panel), and co-injection of the former two (bottom panel). Grey boxes indicate the parental ion at [M+H]+=352.16. (c) LC-Orbitrap-FTMS chromatograms at \(m/z=370.16827\) (10 ppm, positive ionization mode, mass for 3β-hydroxycostunolide-cysteine conjugate) of 3β-hydroxycostunolide-cysteine conjugate formed in an enzyme assay of 3β-hydroxycostunolide and cysteine with glutathione-S transferase (top panel), cysteine conjugated products formed by microsomes expressing Tp8879 fed with costunolide (2\textsuperscript{nd} panel), and the co-injection of those two extracts (3\textsuperscript{rd} panel). Microsomes expressing empty vector fed with costunolide and treated with cysteine was used as control. (d) MS spectrum of peak RT=10.07 min for 3β-hydroxycostunolide-cysteine conjugate (top panel) and cysteine conjugated compounds in microsomes expressing Tp8879 fed with costunolide (bottom panel). Grey boxes indicate the parental ion at [M+H]+=370.17.
ing cysteine to the reaction mixture (Liu et al., 2011). To test whether products are better detected after conjugation to cysteine, the Tp08879 assay product formed from costunolide was incubated for 1 hr with cysteine. Two compounds conjugated to cysteine were identified with mass [M+H]+=352.15771 and [M+H]+=370.16827. Using several sesquiterpene lactone standards conjugated to cysteine we could establish that the product from Tp8879 conjugated to cysteine with the mass [M+H]+=352.15771 is kauniolide-cysteine: its mass and mass spectrum matched that of a kauniolide-cysteine standard and the peak co-eluted with a kauniolide-cysteine standard after co-injection (Figure 3a and b). The product with mass [M+H]+=370.16827 is 3β-hydroxycostunolide-cysteine: its mass and mass spectrum matched that of a 3β-hydroxycostunolide-cysteine standard (Figure 3c and d). The peak intensity ratio of kauniolide-cysteine / 3β-hydroxycostunolide-cysteine was 6.7:1. No product peaks were detected in the microsome fractions incubated with parthenolide, also not when the assay mix was incubated with cysteine to enhance detection of putative products (data not shown). The results of the incubation with costunolide suggest that Tp8879 is a kauniolide synthase which may produces 3β-hydroxycostunolide as intermediate. However, when microsomes of yeast cells expressing Tp8879 were provided directly with this putative intermediate 3β-hydroxycostunolide (or 3β-hydroxyparthenolide), no new peaks were detected, suggesting that the enzyme encoded by Tp8879 can not further convert this putative enzymatic intermediate.

Reconstitution of the kauniolide biosynthetic pathway in Nicotiana benthamiana

To verify the function of kauniolide synthase (TpKS) in planta, TpKS was cloned into a binary expression vector under control of the 35S-promoter and the full biosynthetic pathway of kauniolide was reconstituted in N. benthamiana by agro-infiltration. A. tumefaciens lines

---

Figure 4. Characterisation of kauniolide biosynthesis in N. benthamiana. (a) LC-Orbitrap-FTMS chromatograms of [M+H]^+=352.15771 (kauniolide-cysteine conjugate). Top panel: Extract of N. benthamiana leaves agro-infiltrated with AtHMGR+p19+TpGAS+CiGAO+CiCOS. 2nd panel: Extracts of N. benthamiana leaves agro-infiltrated with AtHMGR+p19+TpGAS+CiGAO+CiCOS+Tp8879. 3rd panel: extraction of kauniolide-cysteine conjugate formed in an enzyme assay of kauniolide and cysteine with glutathione-S transferase. Bottom panel: co-injection of (2) and (3). (b) MS spectrum of peak RT=24.41 min for N. benthamiana leaves agro-infiltrated with AtHMGR+p19+TpGAS+CiGAO+CiCOS+Tp8879, MS spectrum of peak RT=24.10 for kauniolide-cysteine conjugate and co-injection of the latter two. Grey boxes indicate the parental ion at [M+H]^+=352.16.
Kauniolide synthase, a cytochrome P450 cyclase

with the costunolide pathway constructs, TpGAS, CiGAO and CiCOS, as well as AtHMGR (to boost substrate availability) were co-infiltrated into N. benthamiana with and without the A. tumefaciens strain containing Tp8879. Leaves were harvested 4 days post-agroinfiltration and extracted with methanol for analysis by LC-Orbitrap-FTMS. Compared with the control (costunolide pathway without Tp8879) two new peaks were detected at RT=24.41 min and RT=24.66 min. The compound eluting at RT=24.41 min was identified as kauniolide-cysteine (Figure 4) and the peak at RT=24.66 min was identified as a kaunolide-glutathione conjugate (data not shown). No free kauniolide was detected.

Discussion

Here we show that cyclisation of the macrocyclic germacranolide to the bicyclic guaianolide sesquiterpene lactones is catalysed by a cytochrome P450 monooxygenase. This enzyme, kauniolide synthase, catalyses the formation of kauniolide from costunolide. The enzymatic reaction likely involves hydroxylation at C3 of costunolide to form 3β-hydroxycostunolide (1), which is a NADPH/O2 dependent step. Then the protonation of the 3β-hydroxycostunolide gives cation 3β-hydroxyxystunolide (2). After losing one H2O, 2 undergoes cation transition (3) and 1-5 cyclisation, which gives cation kauniolide (4). Finally, a selective deprotonation towards the bridgehead carbon atom of 4 gives kauniolide.

Figure 5. Proposed mechanism for the enzyme catalysed formation of kauniolide from costunolide. First, costunolide is hydroxylated at C3 to form 3β-hydroxycostunolide (1), which is a NADPH/O2 dependent step. Then the protonation of the 3β-hydroxycostunolide gives cation 3β-hydroxyxystunolide (2). After losing one H2O, 2 undergoes cation transition (3) and 1-5 cyclisation, which gives cation kauniolide (4). Finally, a selective deprotonation towards the bridgehead carbon atom of 4 gives kauniolide.
P450 CYP71 sequences that were identified in a feverfew trichome cDNA library (Chapter 4). Accumulation levels of artecanin increased throughout ovule development unlike those of costunolide and parthenolide (Figure 2). We therefore selected a candidate gene, Tp8879, that showed continued expression throughout ovule development for testing as kauniolide synthase. Subsequently, Tp8879 was indeed shown to encode a kauniolide synthase (TpKS).

Enzyme assays showed that TpKS uses costunolide, the simplest germacranolide, as substrate and not parthenolide. This places costunolide at the branch point of the guaianolide and germacranolide STL biosynthesis pathways. A side product of the TpKS enzyme assay with costunolide as substrate was 3β-hydroxycostunolide, although this product was not detected in the *in planta* pathway reconstitution. Although this may be an indication that kauniolide is formed from costunolide by TpKS through the intermediate 3β-hydroxycostunolide, TpKS did not accept 3β-hydroxycostunolide as substrate. The cyclisation of 3β-hydroxycostunolide is not spontaneous, as another P450 (Tp8878) oxidises costunolide at the same carbon position (C3) to form a stable 3β-hydroxycostunolide product (Chapter 4). Results thus suggest that TpKS has additional catalytic capability compared with Tp8878, enabling an additional protonation reaction as indicated in Figure 5. Apparently this protonation reactions needs to follow the hydroxylation reaction while the substrate is still bound to the enzyme. Once intermediate 3β-hydroxycostunolide is released it can not be used as substrate anymore as was evident from the lack of activity in enzyme assays with kauniolide synthase using 3β-hydroxycostunolide as substrate. The proposed multiple step enzymatic mechanism of TpKS is not without precedent. Several plant P450s have been reported to catalyse multiple oxidations of their substrate. For example, amorphadiene synthase, CYP71AV1, from *Artemisia annua* converts amorphadiene to the corresponding alcohol, aldehyde and acid sequentially (Teoh et al., 2006) and germacrene A oxidases from chicory, lettuce and feverfew have been reported to convert germacrene A to the corresponding alcohol, aldehyde and acid form (Nguyen et al., 2010; Cankar et al., 2011). However these P450s oxidise their substrate at the same carbon position. TpKS is able to oxidise its substrate followed by a protonation, double bond migration and carbocation-driven cyclization (Figure 5), which makes it a special type of P450.
Costunolide synthase, a cytochrome P450 cyclase

Costunolide synthase thus seems to form a central intermediate both for the different types of germacranolides and the different types of guaianolides in the majority of Asteraceae plants. Costunolide synthase has been hypothesized to have emerged relatively early in Asteraceae evolution (Ikezawa et al., 2011) from germacrene A oxidase (GAO), and is present in several Asteraceae species, as shown in the phylogenetic tree in Figure 6. Several ESTs from Asteraceae in the NCBI database (ncbi.nlm.nih.gov) show high (over 80%) amino acid sequence identity to TpKS (Figure 7) and these genes may therefore be good candidates for kauniolide synthase in these plant species. Although TpGAO, TpCOS, TpPTS, Tp8878 and TpKS all belong to the same CYP71 family, TpKS is more similar to TpCOS (~48% amino acid sequence identity) than to TpPTS and Tp8878 (~32% and 37% amino sequence identity, respectively), suggesting that the evolution of TpKS is more constrained than that of TpPTS and Tp8878 (Figure 7).

As thousands of guaianolides have been reported in plants and other organisms (Simonsen...
et al., 2013), the identification of TpKS, which catalyses the formation of the simplest guai-
nolide, kauniolide, makes the biosynthesis of guaianolides derived from it feasible. Indeed,
co-expression of all kauniolide pathway genes in a heterologous host, *N. benthamiana*, result-
ed in the production of kauniolide, showing the potential of engineering of this pathway into
alternative hosts. For the production of further functionalised guaianolide STLs additional
enzymes will be required. The expression profile of TpKS differs from that of earlier enzymat-
ic steps in the guaianolide biosynthesis pathway and may thus be used to select other P450
genes with similar expression pattern, for example the ones responsible for the formation of
artecanin (Figure 1).

**Materials and methods**

**Isolation and cloning of full length candidate genes from feverfew**

A previously reported EST library constructed from mRNA isolated from feverfew (*Tanace-
tum parthenium*) trichomes was used for gene isolation (Chapter 2)(Majdi et al., 2011). Fifty
nine candidate cytochrome P450 contigs of the CYP71 group were identified by sequence
homology to known sesquiterpene monooxygenases. One of them was selected for func-
tional characterization based on its expression profile during ovary development. RACE PCR
(Clontech) was used to obtain the 5'- and 3'-sequence of the candidate contigs. The full length
cDNA of the candidate gene was amplified from feverfew cDNA with the addition of NotI/
PacI restriction sites. The cDNA was subsequently cloned into the yeast expression vector
pYEDP60 (Pompon et al., 1996) and sequenced.

**Gene expression analysis**

Gene expression levels of candidates genes were measured by quantitative real time RT-PCR.
Total RNA and cDNA were obtained from ovaries of feverfew flowers collected at develop-
mental stages 2 to 6 as previously reported (Majdi et al., 2011). Real time RT-PCR was
performed using a LightCycler 480 (Roche Di agnostics). The LightCycler experimental run
protocol used was: 95°C for 10 min, 95°C for 10 s, 60°C for 30 s for 40 cycles and finally
a cooling step to 40°C. LightCycler Software 1.5.0 was used for data analysis. For efficien-
cy determination, a standard curve of six serial dilution points (ranging from 200 to 6.25
ng) was made 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: forward *TpGAO* 5’-TGCAGCTCCCGCTTGCTAATATAC-3’, reverse *TpGAO*
5’-AGTCTTTTCTTTGAACCGTGCTCC-3’, forward *TpCOS* 5’-TAGCTTCATCCCG-
GAGCGATTTGA-3’, reverse *TpCOS* 5’-AAATTCTTCGGCCGACCAAATG-3’, forward
*TpPTS* 5’-AGACATTACGTTACACCCTCCCG-3’, reverse *TpPTS* 5’-ATCACGACACAA-
GTCCCAGGGAAA-3’, forward Tp8879 5’-AGTTCCTTACGGCCATTTCTGGGA-3’, re-
verse 5’-AGAAGATCGTCTCAGTTGCTCCA-3’. Quantification of transcript levels was
done in three independent biological replicates and for each biological replicate three tech-
nical replicates were used. Actin was used as a housekeeping gene. The ΔCT was calculated
as follows: \( \Delta CT = CT (\text{Target}) - CT (\text{Actin}) \). The fold change value was calculated using the expression \( 2^{-\Delta CT} \) (Schmittgen & Livak, 2008).

**Plasmid construction for gene expression in yeast**

Candidates \( Tp8879 \) was cloned into pYED60 using NotI/PacI restriction sites. The obtained construct, named \( Tp8879::pYED60 \), was 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**

Microsomes of yeast transformed with candidate \( Tp8879 \) were isolated as described by Pompon et al. (1996) with modifications. Transformed yeast cell cultures were grown in 50 mL SGI medium for 36 hours at 30°C. By adding 250 mL YPL medium containing 2% galactose, the induction was performed for 24 hours at 30°C. Then cells were collected and chilled on ice for 20 min. After centrifugation at 4,900 x \( g \) for 10 min, 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 \( \beta \)-mercaptoethanol) and incubated for 10 min at room temperature. Following centrifugation at 4,900 x \( g \) for 10 min, the pellet was washed three times with extraction buffer (without resuspending the pellet). Cells were re-suspended in 3 mL extraction buffer (without \( \beta \)-mercaptoethanol) and transferred to a 50 mL Falcon tube. The centrifuge bottle was washed with another 2 mL of extraction buffer (without \( \beta \)-mercaptoethanol) and this was also transferred to the same 50 mL Falcon tube. About 25 mL of glass beads (450-500 μm) were used to lyse the cells by shaking for 10 min in a cold room. The lysed cells were transferred to a 25 mL centrifuge tube and centrifuged at 10,500 x \( g \) for 10 min. The supernatant was centrifuged at 195,000 x \( g \) for 2 h. The pelleted microsomal fraction was re-suspended in 4 mL chilled 50 mM Tris-HCl buffer pH 7.5 containing 1 mM EDTA and 20% (v/v) glycerol using a chilled glass Tenbroeck homogenizer. Yeast microsomes were then aliquoted in pre-cooled 1.5 mL eppendorf tubes and stored at -80°C until use.

All enzyme assays were carried out using 40 mM potassium buffer pH7.5. Reaction mixture that contains 14.4% (v/v) isolated microsomes, 0.2 mM substrate (10 mM stock in DMSO), 2mM NADPH (10 mM stock in 100 mM potassium buffer), 40mM potassium buffer (1 M, pH 7.5), and 45.2% (v/v) was incubated for 2.5 h at 25 °C with shaking (200 rpm). Then the mixture was centrifuged at 12000 rpm for 10 min. The supernatant was filtered through 0.22 μm filter before injection into LC-MS.

**LC-Orbitrap-MS analysis of yeast microsome assay mixture**

To analyse microsome assay mixtures, 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 was used. Chromatographic separation took place on an analytical column (Luna 3 μ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA). Degassed eluent A [ultra-pure water: formic acid (1000:1, v/v)] and eluent
B [acetonitril:formic acid (1000:1, v/v)] were used at a flow rate of 0.19 mL min$^{-1}$. A linear gradient from 5 to 75% acetonitrile in 45 min was applied, which was followed by 15 min of washing and equilibration. FTMS full scans ($m/z$ 100–1200) were recorded with a resolution of 60000, whereas for MS$n$ scans a resolution of 15000 was used. The FTMS was externally calibrated in negative mode using sodium formate clusters in the range $m/z$ 150–1200, and automatic tuning was performed on $m/z$ 384.93. Injection volume was 5 μl.

**Plasmid construction for expression in *Nicotiana benthamiana***

For transient expression in *N. benthamiana*, *TpGAS*, *TpGAO*, *TpCOS*, *TpPTS* and *Tp8879* were cloned into ImpactVector1.1 (http://www.impactvector.com/) to express them under the control of the Rubisco (RBC) promoter (Outchkourov et al., 2003). *TpGAS* was also cloned into ImpactVector1.5 to fuse it with the RBC promoter and the CoxIV mitochondrial targeting sequence as we have demonstrated before that mitochondrial targeting of sesquiterpene synthases results in improved sesquiterpene production (Liu et al., 2011). 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***

*A. tumefaciens* infiltration (agro-infiltration) for transient expression in *Nicotiana benthamiana* was performed as described by Liu et al. (2011). *A. tumefaciens* batches were grown at 28 °C at 220 rpm for 48 h in LB media with proper antibiotics. Cells were harvested by centrifugation at 4000 $x_g$ for 20 min and then resuspended in MES buffer (10 mM MES, 10 mM MgCl$_2$ and 100 μM acetosyringone to a final OD$_{600}$ of ~ 1, followed by incubation at room temperature under shaking at 50 rpm for 150 min. For co-expression, equal volumes of the *Agrobacterium* batches were mixed. Batch mixtures were infiltrated into leaves of three-week-old *N. benthamiana* plants by pressing a 1 mL syringe against the abaxial side of the leaf and slowly injecting into the leaf. After infiltration the plants were grown for another four and half days and then harvested for analysis.

**LC-QTOF-MS analysis of leaf extracts**

Extraction of agro-infiltrated leaves and LC-QTOF-MS (liquid chromatography, coupled to quadrupole time-of-flight mass spectrometry) analysis were performed as described by Liu et al. (2011). For extraction, 100 mg of fresh leaves were ground in liquid nitrogen and extracted with 300 μl methanol:formic acid (1000:1, v/v). After brief vortexing and sonication for 15 min, the extracts were centrifuged at 13,000 rpm for 5 min and filtered through a 0.22 μm inorganic membrane filter (RC4, Sartorius, Germany), and later injected into LC-QTOF-MS for analysis. 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) operating in negative ionization mode was used. An analytical column (Luna 3 μ C18/2 100A; 2.0 × 150 mm; Phenomenex, USA) attached to a C18 pre-column (2.0 × 4 mm; Phenomenex, USA) was used. Degassed eluent A [ultra-pure water: formic acid (1000:1, v/v)] and eluent B
[acetonitril:formic acid (1000:1, v/v)] were used at a flow rate of 0.19 mL min⁻¹. Masses were recorded between m/z 80 and m/z 1500; leucine enkaphalin ([M-H]⁻=554.2620) was used as a lock mass for on-line accurate mass correction. The gradient of the HPLC started at 5% eluent B and increased linearly to 75% eluent B in 45 min, after which the column was washed and equilibrated for 15 min before the next injection. Injection volume was 5 μl.

**Cysteine conjugation**

Cysteine conjugation was performed as described by Liu et al. (2011). In brief, glutathione (GSH) (150 mM) or cysteine (150 mM) in 7 μl potassium buffer (100 mM; pH 6.5), and standards (30 mM) in 7 μl ethanol were added to 1000 μl potassium buffer (100 mM; pH 6.5). The reaction was initiated by adding 7 μl of GST (1g L⁻¹, in 100 mM potassium buffer; pH 6.5) into the mixture. Complete assay mixtures without GST enzyme or either of the substrates were used as controls. After incubation for 30 min at room temperature, samples were kept at -20 °C until analysis. Costunolide was purchased from TOCRIS Bioscience (United Kingdom). Parthenolide, 3β-hydroxycostunolide, and 3β-hydroxyparthenolide, isolated from dried aerial parts of feverfew plants, were provided by Dr. Justin T. Fischedick of PRISNA (Fischedick et al., 2012). Kauniolide was provided by Prof. Yue Chen, Nankai University, China (Zhai et al., 2012).

**References**


Chapter 5


Chapter 6

General discussion

Qing Liu
Plant secondary metabolites are an important source for the development of new drugs. For example, antimalarial drug artemisinin, was isolated from a herb *Artemisia annua* which has been used for many centuries in Chinese tradition medicine as a treatment for fever and malaria (Klayman, 1985). Another terpenoid paclitaxel (taxol), first isolated form the pacific yew tree (*Taxus brevifolia*), has been considered as the most successful anticancer drugs (Heinig *et al.*, 2013). Parthenolide, a sesquiterpene lactone present in feverfew, is a promising lead for the development of an anti-cancer drug and has attracted a lot of attention from medical institutes and companies. The elucidation of the biosynthetic pathway of parthenolide should improve the availability of parthenolide, and possibly parthenolide derivatives with improved chemical properties, and hence speed up the development of parthenolide-based anti-cancer drugs. The aim of the study presented in this thesis was to elucidate the biosynthetic pathway of parthenolide and related sesquiterpene lactones and to deliver the proof of concept for the heterologous production of these compounds in a heterologous host plant. With the cDNA library of feverfew trichomes at hand, the parthenolide biosynthetic pathway was elucidated by isolating all the structural genes (Figure 1). Moreover, the whole pathway was reconstituted in a heterologous plant species, *Nicotiana benthamiana*, through transient expression (agro-infiltration). Besides the genes involved in the parthenolide biosynthetic pathway, we also isolated P450s that catalyse branches of the main pathway (Figure 1): Tp8878, a 3β-hydroxylase, which can hydroxylate both costunolide and parthenolide at C3 and Tp8879, kauniolide synthase, which can cyclises costunolide to form kauniolide. Another P450, Tp8886, can oxidise kauniolide to an alcohol or epoxide, but the exact identity of the product has not be determined yet.

**Strategies to enhance medicinal compound yield from medicinal plants**

One of the drawbacks of using plants as a source for the production of drugs is that many plant species produce the bioactive terpenoids in only low quantities. Sometimes this is because the biosynthesis of those terpenoids is restricted to specialized cells (e.g. trichome, laticifer) which form only a small portion of the cells in a certain tissue, e.g. leaf, root, ovary. To increase the production of bioactive terpenoids in medicinal plants where these compounds naturally occur, the options are limited, as most medicinal plants are difficult to cultivate, to multiply or to transform.

However, there are exceptions to this rule. For the anti-malarial drug artemisinin, for example, transformation of the host plant *Artemisia annua* has been established (Weathers *et al.*, 1994; Vergauwe *et al.*, 1996; Han *et al.*, 2005). Overexpression of genes involved in the biosynthetic pathway or transport of artemisinin in *A. annua* increased the yield of artemisinin by more than two-fold (Jing *et al.*, 2008; Wang *et al.*, 2011; Zhang *et al.*, 2012). In addition to this transgenic approach, the yield of artemisinin in *A. annua* could also be promoted through cultivation, such as exposure to salinity stress (Qian *et al.*, 2010) or through abscisic acid treatment (Jing *et al.*, 2009). For feverfew, as transformation has not yet been established, the promotion of parthenolide yield is limited to agricultural practices, such as harvesting time
or subjecting plants to a single water stress event (Fonseca et al., 2005).

**Strategies to enhance medicinal compound yield through metabolic engineering**

With the elucidation of the full pathway of some important terpenoids (e.g. all biosynthesis genes of the parthenolide biosynthesis pathway in this thesis), production of medicinal compounds in heterologous plant species that are more easy to manipulate and cultivate will open up new possibilities to produce valuable compounds in a cost-efficient way. Key components in the manipulation of medicinal compounds in heterologous plants include identification and isolation of the relevant pathway genes, gene characterisation, product detection and identification, and whole pathway reconstitution.

**Identification and isolation of the relevant pathway genes**

The identification and characterization of isoprenoid biosynthetic genes have mostly used similar strategies. First the presence of a bioactive terpenoid compound in a certain plant species, and its tissue-specific accumulation, are determined. After that, product accumulation in specific tissues or cells over plant development is determined. In this thesis, parthenolide was shown to be mainly present in trichomes of feverfew flowers (Chapter 2). The isolation
of specific cells that synthesize the product can greatly enhance the chances of identifying the relevant genes from sequencing of mRNA isolated from these specific cell types. For the identification of genes involved in parthenolide biosynthesis, the trichomes were isolated from ovaries and used for cDNA library construction and subsequent 454 sequencing. The isolation of specific metabolic pathway genes from mRNA isolated from glandular trichomes has been reported for several plant species, including mint (Mentha piperita) (Alonso et al., 1992; Rajaonarivony et al., 1992; Lange et al., 2000), basil (Ocimum basilicum) (Gang et al., 2002; Iijima et al., 2004; Xie et al., 2008), A. annua (Teoh et al., 2006; Zhang et al., 2008; Ting et al., 2013), tomato (Solanum lycopersicum) (Fridman et al., 2005; Besser et al., 2009; Schilmiller et al., 2009), and hop (Humulus lupulus) (Wang et al., 2002; Nagel et al., 2008; Xu et al., 2013).

With the EST sequences in hand, the next challenge is to recognize the potential candidate genes, which usually starts by comparing the sequences to those available in public databases such as NIST. Vice versa, known terpene synthase gene sequences are used to blast against the specific sequence database to identify terpene synthases. Terpene synthases can be recognized as they share a common aspartate-rich DDxxD motif which is thought to be involved in the coordination of divalent metal ions for substrate binding (Lesburg et al., 1997). Similar strategies may be used to identify cytochrome P450s which may be involved in the further modification of the terpenoid products. Especially for the P450 enzymes, this approach will renders multiple candidate genes from which the right one still needs to be selected. For instance, in this thesis, fifty nine candidate P450s genes were identified for the parthenolide biosynthetic pathway in the feverfew trichome cDNA database. An efficient way to select the most likely candidate gene from such a large group of candidates is to use a combination of expression and metabolite profiling. Expression profiling does not require full length sequence information and when different tissues or developmental stages of a tissue are available that display a trend in metabolite accumulation, the metabolite profile can be compared to the expression profile of the candidate genes over the same tissue samples. In this way the list of 58 candidate P450 genes could be reduced to nine top candidates.

In this thesis, we noticed that all the P450s in the same biosynthetic pathway belonged to the same group, e.g. both GAO and COS belong to the CYP71 group, which helped us to identify P450s involved in the following steps, e.g. PTS (CYP71CA1, Chapter 4). Actually the identification of GAO was based on the discovery of amorphadiene oxidase (CYP71AV1), a gene that can oxidise amorphadiene to artemisinic acid (Teoh et al., 2006). As the structure of germacrene A is close to that of amorphadiene, GAO was assumed to share high amino acid sequence similarity to that of CYP71AV1, leading to the finding of GAO (CYP71AV8) (Nguyen et al., 2010; Cankar et al., 2011). Later on, the identification of GAO (CYP71AV8) inspired the identification of COS from lettuce (CYP71BL2) (Ikezawa et al., 2011), chicory (CYP71BL3, Chapter 3) (Liu, Q et al., 2011) and feverfew (Chapter 4), as well as parthenolide synthase, kauniolide synthase and kauniolide oxidase from feverfew (Chapter 4 and 5).

**Candidate gene characterisation**

Once the candidate sequences have been identified, their full length sequences need to be ob-
General discussion

Terpene synthase genes can be expressed in E. coli for functional characterisation while yeast should be used for expressing proteins that are targeted to microsomal membranes, such as P450s. Plants offer the advantage that they carry out more complex post-translational modifications than the other two expression hosts. In this thesis, the function of the terpene synthase gene, TpGAS, was identified by expression in E. coli, yeast and N. benthamiana (Chapter 2). The function of P450 candidates was identified by expression in yeast and N. benthamiana (Chapter 3, 4 and 5).

For candidate genes that operate downstream in the terpenoid biosynthetic pathway, the characterisation by expression in yeast is difficult when specific substrates are not readily available for feeding assays. This problem can be overcome by co-expression of the genes that encode enzymes making the putative substrate of the candidate enzyme to be tested. For the identification of LsCOS, for example, candidate genes were co-transformed in yeast with GAS+GAO that together makes the substrate for LsCOS, Germacra-1(10),4,11(13)-trien-12-oic acid (Ikezawa et al., 2011). Another option would be to combine microsome extracts from multiple independent yeast cultures expressing single genes in one combined in vitro assay. In this thesis, we mixed the microsomes of yeast expressing kauniolide synthase (TpKS) and another P450 candidate, Tp8886, together. After feeding the mixture with the substrate of KS, costunolide, a new peak of oxidised kauniolide (based on molecular mass spectrum) was detected by LC-MS. The oxidised kauniolide is not from an epoxidation as the retention of this new peak is not the same as either of its kauniolide-epoxidized compounds (data not shown). We were unable to identify this new product yet, and it is also not present in feverfew! This demonstrates the power of combinatorial biochemistry where we combine enzymes into new pathways, producing compounds that may even be new to nature and could potentially have novel biological activity.

Agrobacterium-mediated transient gene expression is often used for identification of new viral suppressors of RNA silencing (Thomas et al., 2003) and screening and functional analysis of unidentified genes (Hashimoto et al., 2012; Kanagarajan et al., 2012). The system has been further used to test whole groups of genes in pathway reconstitution experiments through agro-co-infiltration in plant, such as N. benthamiana (van Herpen et al., 2010), lettuce (Joh et al., 2005), and tobacco (Yang et al., 2000). When expression constructs are available for all known steps of the pathway (starting with the specific TPS) and all candidate genes are cloned into individual expression vectors, such agro-infiltration assays may be used to test activity of the transiently expressed novel genes in N. benthamiana. By using this way, genes from the artemisinin biosynthesis pathway in Artemisia annua have been identified and characterized (Ting et al., 2013), and so does the genes of the costunolide and parthenolide biosynthetic pathway of feverfew in this thesis (Chapter 4).

Product detection and identification

Targeted detection
Different techniques could be used for the detection of products. When the standard of a dedicated product is available, targeted analysis could be applied by GC-MS, LC-QTOF-MS, LC-Orbitrap-FTMS, and LC-MRM-MS/MS. In this thesis, the content of costunolide and parthenolide in transformed yeast extract or agro-infiltrated *N. benthamiana* leaves was qualified and quantified by standards (Chapter 3 and 4).

**Chemical modification:** Some compounds are difficult to detected either because they are not well extracted by the solvent used or the show poor ionisation during analysis by QTOF. Such compound may be easier to detect after a simple modification. In this thesis, the product of *TpKS* was hardly detectable when expressed in yeast. Cysteine was added to the reaction mixture as we have known from previous work that the product could be conjugated with cysteine through Michael-reaction. A new obvious peak was detected and later on identified as kauniolide-cysteine (Chapter 5).

**Standards through bulk isolation:** To identify the products of candidates, standards need to be available. But for the complicated sesquiterpene biosynthesis pathways most substrates and products are unknown and are not commercially available. One option to obtained specific standard compounds is through bulk isolation and purification from host plant tissue. For example, several sesquiterpene lactones have been isolated from bulk feverfew plant tissue material for which the structure of the purified compound was elucidated by NMR (Fischedick *et al.*, 2012). Those compound were later crucial to identify two novel compounds produced by parthenolide candidates, 3β-hydroxycostunolide and 3β-hydroxyparthenolide (Chapter 4).

**Standards through chemical synthesis:** Another option to obtain standards would be chemical synthesis of the predicted compound. In this thesis, an unknown product of one of the candidate genes was identified as kauniolide based on references chemically synthesized (Zhai *et al.*, 2012) (Chapter 5). A third option would be using *in vitro* synthesized compounds as reference. In this thesis, two novel conjugates of costunolide were found in agro-infiltrated *N. benthamiana* leaves. Based on their molecular masses, they are assumed to be cysteine and glutathione conjugated of costunolide. Through *in vitro* enzyme assay, costunolide-glutathione and costunolide-cysteine were synthesized and used to identify those two novel compound (Chapter 3). Similar strategy was used to identify parthenolide-glutathione and parthenolide-cysteine later (Chapter 4). We have to admit the limitation of this option, which is better used when other options are not available.

**NMR elucidation:** When no standards are available to identify the product, NMR may be applied to elucidate the structure, as NMR is in principle the most uniform detection technique and is essential for the unequivocal identification of unknown compounds. This strategy has been successfully applied to the structure of two novel compound, artemisinic acid-12-β-diglucoside (van Herpen *et al.*, 2010) and geranoyl-6-O-malonyl-β-D-glucopyranoside (Yang *et al.*, 2011). However, structure elucidation by NMR can be very time and energy consuming, as the sensitivity of NMR is relatively low compared with MS.

**Untargeted detection**
When searching for unknown/novel compound is required, unbiased metabolomics detec-
General discussion

- A targeted LC-MS profiling of extract from maize leaves overexpressing geraniol synthase was performed to find the major product of the overexpressed gene as geranic acid glycoside (Yang et al., 2011). In this thesis, similar untargeted approach was applied and two novel conjugates costunolide-cysteine and costunolide-glutathione were identified in agro-infiltrated N. benthamiana leaves (Chapter 3).

- Functional characterisation of candidate genes can be done by monitoring product accumulation, or substrate consumption when products of genes are not detectable. In this thesis, the identification of TpGAO was conducted by co-expressing it with TpGAS in N. benthamiana. Although no new product was detected, we noticed that the production of germacrene A (substrate) was decreased dramatically in N. benthamiana expressing TpGAS/TpGAO compared to the one expressing TpGAS only. By expressing TpGAO in yeast, we were able to detect its products, germacrene A alcohol, germacrene A aldehyde and germacrene A acid (Chapter 4).

Whole pathway reconstitution

- Many plant-derived terpenoids have pharmaceutical and industrial applications, such as parthenolide, but their natural resources for extraction are often limited and, in many cases, their chemical synthetic routes are not commercially available. Metabolic engineering, either in the native producer or a heterologous host, is the only realistic alternative to improve yield and accessibility. The reconstruction of a biosynthesis pathway that involve multiple genes in stable transformed heterologous plants can be complicated as introduction of multiple genes under the same promoter may lead to silencing problems. Step wise introduction of pathways in stable transformed plants may also be hampered by accumulation of toxic intermediates which may cause selection of low expressing transformants. Transient gene expression system, mediated by Agrobacterium, in intact plant leaves is a rapid and useful method for metabolic engineering. By this method many different heterologous proteins can be produced without the need to generate transgenic plants, which is tedious, time consuming and sometimes difficult for many plant species (Fischer et al., 1999; Fischer et al., 2004).

- Several platforms have been reported to be suitable for transient expression, e.g. N. benthamiana (Nafisi et al., 2007; van Herpen et al., 2010), Nicotiana tabacum (tobacco) (Yang et al., 2000), and Lactuca sativa (lettuce) (Joh et al., 2005). Transient gene expression has several advantages over stable expression. First, the transient gene expression technique is simple and easy to perform and the results can be assayed one week post-agroinfiltration (Kapila et al., 1997), compared to months (to years) for stable transformation. Second, several genes can be introduced into plants at the same time through transient expression. (ref). In my thesis I introduced up to 15 genes have been successfully introduced into N. benthamiana plants through transient expressing and used that for the discovery of the P450 genes involved in STL biosynthesis in feverfew (Chapter 4 and 5). Finally, the relative expression level of the genes may also be manipulated by changing the relative amount of Agrobacterium carrying...
each of the genes, which is difficult to achieve by stable transformation.

Using transient expression in *N. benthamiana*, only about 50% of costunolide was converted to parthenolide when *TpPTS* was added to the pathway (Chapter 4). This is likely caused by the co-infiltration of too many genes at the same time, such that not all cells are transformed with the whole pathway. The more multiple independent agrobacterium strains are co-infiltrated, the higher the chance that some cells are not being transfected with all the pathway genes. When *TpGAS* was transiently expressed in *N. benthamiana*, a huge peak of peak of germacrene A could be detected by GC-MS. When *TpGAO* was co-expressed with *TpGAS*, the peak area of germacrene A decreased by more than 90%. However, we were not able to detect any new peaks produced by *TpGAO*. Nevertheless when we co-expressed *TpCOS* or *CiCOS* with *TpGAS* and *TpGAO* or *CiGAO* we could detect costunolide and its conjugates, showing that Germacr-1(10),4,11(13)-trien-12-oic acid is produced upon expression of GAS and GAO (and is available to costunolide synthase when that is co-expressed), but is apparently conjugated to something that we cannot detect with our mass spectrometer.

**Strategies for boosting flux through the pathway**

To increase productivity during pathway reconstitution, it is important to boost the flux through the pathway or make use of pathway precursors more efficiently. Several strategies can be used to achieve that, including targeted expression, eliminating competing endogenous enzyme activity, Boosting precursor availability, and construction of an artificial metabolon.

**Targeted expression**

As the biosynthesis of different terpenoids is highly compartmentalized, enzymes need to be targeted to the appropriate location to achieve better production of terpenoids. For example, by expressing a limonene synthase gene from *Perilla frutescens*, with different targeting signal (plastid, cytosolic, endoplasmic reticulum), in tobacco (*Nicotiana tabacum*), Ohara *et al.* (2003) found that the specific activity of limonene synthase was generally higher in transgensics expressing the plastid localization construct compared to those containing the cytosolic enzyme. In this thesis, by targeting *TpGAS* to the mitochondria, the production of germacrene A in *N. benthamiana* was increased 15-fold compared with cytosolic localisation (Chapter 3).

**Eliminating competing endogenous enzyme activity**

When reconstituting biosynthetic pathways by transiently expressing multiple genes in a heterologous host plant, losses may occur if intermediate products are recognized by endogenous enzymes which divert carbon flux from the desired product route. For example, when GAS+GAO+COS were co-infiltrated in *N. benthamiana*, more than 90% of the costunolide produced was conjugated either to cysteine or to glutathione. When PTS was co-infiltrated with GAS+GAO+COS, the concentration of costunolide conjugate decreased by 50%, showing that the heterologous PTS competes with the endogenous glutathione S transferase conjugating costunolide. Elimination or restriction of endogenous competing reactions has been
shown to increase production of target compounds in yeast (Ro et al., 2006; Paddon et al., 2013). The advantage of a transient expression/production system would be that inhibition of competing reactions in mature leaves would be of little consequence during the 4-5 days of transient parthenolide production, while permanent elimination of such competing reactions in stable transformants could potentially lead to reduced plant fitness. On the other hand, storage of target metabolites as conjugates may have the advantage that the storage capacity of the vacuole is used and that high concentrations can be reached without phytotoxic effects (1995).

**Boosting precursor availability:** The overexpression of genes encoding enzymes such as 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), deoxyxylulose 5-phosphate synthase (DXS) and prenyltransferases, has been used to elevate terpenoid levels in plants (Degenhardt et al., 2003; Wu et al., 2006). Overexpression of HMGR and farnesyl diphosphate synthase (FPS; farnesyl diphosphate is the immediate C15 precursor for sesquiterpene biosynthesis) together with amorphadiene synthase resulted in an 80-fold increase in the production of the artemisinin precursor, amorphadiene (van Herpen et al., 2010). In my thesis work, I used the same AtHMGR together with AtFPS to boost the production of costunolide 7-fold (free form) and 20-fold (conjugated form) (Figure 2). Later we found that the boosting activity was due
to AtHMGR, while AtFPS did not contribute (Figure 3). A possible explanation for that could be that the product of AtFPS is not available to the downstream GAS. In subsequent experiments (Chapter 4 and 5), only AtHMGR was therefore used to boost production in agro-infiltration.

Construction of an artificial metabolon:
Metabolons are multienzyme complexes which allow the direct passage of product from one enzymatic reaction to a consecutive enzyme in a metabolic pathway (Møller, 2010). The advantage of a metabolon is that it can limit the diffusion of intermediates into the surrounding medium, facilitates fast turnover of labile or toxic intermediates, and may prevent undesired crosstalk between different metabolic pathways (Møller, 2010). As the expression of parthenolide biosynthetic pathway genes in feverfew is highly coordinated and there is competition between heterologous genes and endogenous N. benthamiana genes, construction of an artificial metabolon using a common scaffold protein could coordinate the flux of heterologous enzyme products/substrates and prevent the competition for heterologous metabolites by endogenous enzymes.

**Comparing feverfew and N. benthamiana yields of parthenolide**

When the whole biosynthetic pathway of parthenolide was reconstituted in N. benthamiana in combination with boosting by co-expression of AtHMGR, the production of free and conjugated parthenolide in the leaves of N. benthamiana was 1.4 mg/g FW (Chapter 4). This is higher than the (free) parthenolide content in feverfew leaves, but lower than the content in feverfew flowers (Figure 4). Nevertheless, there are several advantages of producing parthenolide through agro-infiltration in N. benthamiana over production with feverfew. First, it takes 4 weeks for N. benthamiana plants to be ready for agro-infiltration, and leaves could be harvested 4-5 days later, compared to 3 months for feverfew plants to be ready for harvest. Second, the industrial culturing of N. benthamiana - including large-scale agro-infiltration - is available, which can ensure a stable plant supply, thus stable supply of parthenolide.

At present the parthenolide amount produced in agro-infiltrated N. benthamiana leaves is probably not yet high enough for industrial-scale production when compared with feverfew. However, in the evaluation of production capacity of feverfew plants and an industrial scale agroinfiltration of pre-grown tobacco plants the time aspects should also be taken into account. The yield from N. benthamiana is obtained 4-5 days after agro-infiltration, while that
from feverfew requires at least 10 weeks of growth followed by a second harvest from the same plants after an additional 27 weeks (http://www2.ca.uky.edu/ktrdc/Bulletin%20KTRDC-IB-3.pdf). In 2009, feverfew seed was sold for $424/kg and the buyer price for dried feverfew was about $3.20/kg. Premiums paid for organically produced product in Canada have fluctuated from year to year and are buyer dependent. Yields of feverfew in Saskatchewan can vary from 2,000 kg/ha to 5,000 kg/ha dry herb (about 20,000-50,000 kg FW/ha). (http://www.agriculture.gov.sk.ca/Default.aspx?DN=5a05a8da-ff3e-490b-b280-5f7b22b803b). Assuming that the yield from feverfew is per year and mainly consists of leaf material (with similar content per gr/FW as agroinfiltrated N. benthamiana plants), one would need to infiltrate 200,000-500,000 N. benthamiana plants (at about 100 gr FW/plant) per year to reach similar yields as with a hectare of feverfew. This mounts to about 4000 plant infiltrations per week.

**Options for host plant engineering**

Instead of using a heterologous plant host, there are also advantages to engineering the homologous host, if that host can be transformed. Several strategies for metabolic engineering of medicinal plants have been used to boost the production of terpenoids, such as overexpressing enzymes and boosting transcriptional regulation.

**Engineering enzymes**

The production of terpenoids can be boosted through overexpression of biosynthetic genes. The artemisinin content in A. annua was enhanced 2.4-fold through over-expression of two of its biosynthetic genes, CYP71AV1 and cytochrome P450 reductase (CPR), in the host plant (Jing et al., 2008). When FPS was co-overexpressed with these two biosynthetic genes, artemisinin production was enhanced 3.6-fold (Chen et al., 2012). A similar strategy could be applied to feverfew when its transformation is established.

**Boosting transcriptional regulation**

The use of transcription factors to up-regulate an entire terpenoid pathway (or large parts of it) has so far not been reported (Lange & Ahkami, 2012). However, there are transcription factors, such as ERF/Ap2 type (van der Fits & Memelink, 2000; Yu et al., 2012), WRKY (Xu et al., 2004) and MYC (Hong et al., 2012), that have been reported to be involved in transcriptional regulation of terpenoid biosynthesis in several different plant species. The understanding of transcriptional regulation of terpenoid biosynthesis may enable us to manipulate the biosynthesis pathway of terpenoids in cell or tissue cultures of medicinal plants. In Catharanthus roseus, overexpression of AP2/ERF transcription factor ORCA3 resulted in enhanced expression of several terpene indole-alkaloid biosynthetic genes and, consequently, in increased accumulation of the terpenoid indole alkaloid products (van der Fits & Memelink, 2000). Overexpression of ORCA2 lead to enhanced accumulation of catharanthine and vindoline in hairy roots (Liu, D-H et al., 2011). In Artemisia annua, overexpression of AP2/ERF transcription factors AaERF1 and AaERF2 resulted in elevated transcript levels of both ADS and CYP71AV1 and consequently in increased accumulation of artemisinin and artemisinic acid (Yu et al., 2012). Overexpression of AaORA - the expression of which could directly be
linked to expression of ADS, CYP71AV1 and DBR2 that encode enzymes of the artemisinin pathway - in transgenic A. annua resulted in an about 50% increase in artemisinin and 35% in DHAA content (Lu et al., 2013). In my thesis I show that the expression of the genes involved in the biosynthetic of parthenolide all show similar profiles, suggesting that these genes are under the regulation of a single transcription factor. Selection of transcription factors that show similar expression profiles as the pathway genes can help us to narrow down the number of candidates. If this transcription factor can be identified, overexpression in feverfew could probably increase the production of parthenolide.

**Increasing the polarity of terpenoids**

Water solubility is a crucial part of the terpenoid biological activity as poor solubility will make pharmacological use of the compound difficult. This poor water solubility is for example a limitation for the use of parthenolide as a drug. Solubility of a compound may be improved by, among others, amino acid conjugation, hydroxylation and glycosylation. Parthenolide amino analogs could be converted into water-soluble organic acid salts through Michael addition reactions, and this increased their water solubility and bioavailability (Guzman et al., 2007; Neelakantan et al., 2009). In my thesis, cysteine and glutathione conjugates of parthenolide were produced upon pathway reconstitution in N. benthamiana. These conjugates are more water-soluble than parthenolide and showed high activity against colon cancer cells, with only slightly lower activity than parthenolide (Chapter 4). Glycosylation has been reported to often change the biological activity of terpenoids as well as improve their pharmacokinetic parameters (Rivas et al., 2013). Hydroxylation, many of which are performed by P450 oxidases, may cause new bioactivity to the compounds as well as improve their water solubility (McGarvey & Croteau, 1995). In this thesis, through the hydroxylation at C3 position of costunolide and parthenolide, their water-solubility was increased while their bioactivity was not affected (Chapter 4).

**Future challenges**

In my thesis work I have elucidated the entire biosynthetic pathway of parthenolide and have reconstituted this pathway in planta. The future challenge lies in how to bring the production of parthenolide to a level high enough for industrial production. Solutions for that challenge include optimization of the enzymatic activity of the pathway genes, boosting of precursor availability, and restriction of competing pathways, like what has been done for the production of artemisinin precursors in yeast (Paddon et al., 2013). Besides those, solutions may also include targeting the expression of the pathway to other sub-cellular compartments, coordination of each gene’s expression level and/or artificial metabolon formation, to reach an optimum flux through the pathway. Finally, as only a few P450 candidates out of the 59 within the CYP71 group have been identified in this thesis, some of the remaining P450s may be able to convert parthenolide to more water-soluble derivatives while keeping its activity. This would possibly make parthenolide interesting enough to warrant large-scale efforts to use
the knowledge generated in my thesis for the large-scale biotechnological production of this natural compound.

References:


Chapter 6


Summary

Parthenolide is the major bioactive compound of feverfew and has anti-inflammatory and anti-cancer activity. Chapter 1 gives an overview of the history and current status of research on parthenolide in feverfew. As a promising anti-cancer drug, parthenolide has attracted a lot of attention from medical institutes and companies. A search with ‘parthenolide’ in Google patents yields more than 2000 hits on extraction of parthenolide or its use in treating cancer or other diseases. However, information on the parthenolide biosynthetic pathway is scarce. Elucidation of the full pathway to parthenolide would open up new opportunities for production of this compound in heterologous, more efficient production platforms.

To elucidate the biosynthetic pathway of parthenolide, knowledge on the tissue(s) in which parthenolide is produced and stored is important. In Chapter 2, parthenolide was found to highly accumulate particularly in floral trichomes, suggesting that this is also the preferred site of biosynthesis. These floral trichomes were subsequently used to isolate germacrene A synthase (TpGAS), the gene encoding the first dedicated step in parthenolide biosynthesis, using a degenerate primer PCR approach. The transcript level of TpGAS was indeed highest in glandular trichomes. The high expression of TpGAS in glandular trichomes which also contain the highest concentration of parthenolide, supports the assumption that glandular trichomes are the organ where parthenolide biosynthesis and accumulation occur.

During my work on Chapter 2, a Canadian group reported a germacrene A oxidase (GAO) from lettuce. As the 454 cDNA library of feverfew trichomes was not available yet, we decided to use a 454 cDNA library of chicory (which also produces costunolide) to continue screening candidate genes involved in the next step of the parthenolide biosynthetic pathway, costunolide synthase (COS). In Chapter 3, four P450s (belonging to the CYP71 group) were selected from the chicory cDNA library for functional characterisation in yeast. One of them, named CYP71BL3, was found to be costunolide synthase, and can catalyse the oxidation of germacr-1(10),4,11(13)-trien-12-oic acid to yield costunolide. The biosynthetic pathway of costunolide was reconstituted in Nicotiana benthamiana by transient expression (agro-infiltration) of TpGAS, CiGAO (which we also identified in the chicory library) and CiCOS, which resulted in costunolide production of up to 60 ng g⁻¹ FW. In addition, two new compounds were formed that were identified as costunolide-glutathione and costunolide-cysteine conjugates.

When the 454 sequences of the feverfew trichome library became available, we continued to identify additional genes involved in the biosynthetic pathway of parthenolide. In Chapter 4, the parthenolide biosynthetic pathway was elucidated by isolating all the structural genes from feverfew, TpGAS, TpGAO, TpCOS and TpPTS. Moreover, the whole pathway was reconstituted in N. benthamiana, through transient expression. In the agro-infiltrated plants, parthenolide as well as a number of conjugates (with cysteine and glutathione) were produced. In an anti-cancer bioassay, these relatively polar conjugates were highly active against colon cancer cells, with only slightly lower activity than free parthenolide. Finally, also a gene encoding
a costunolide and parthenolide 3β-hydroxylase was identified, which could potentially be used in biotechnological applications to produce hydroxylated parthenolide. The conjugation and hydroxylation of parthenolide open up new options to improve the water solubility of parthenolide and therefore its potential as a drug.

Besides genes involved in the biosynthetic pathway of parthenolide, we also identified two other P450 genes that can utilize costunolide as substrate. In Chapter 5, Tp8879 is identified. Tp8879 can cyclise the monocyclic germacranolide sesquiterpene lactone costunolide to form the bicyclic guaianolide sesquiterpene lactone kauniolide, and is hence called kauniolide synthase. The biosynthetic pathway of kauniolide was reconstituted in *N. benthamiana*, through transient expression.

This thesis combines a series of existing and new technologies for gene discovery – transcriptomics and metabolomics - as well as optimisation of plant metabolic engineering – using transient expression in *N. benthamiana* - and reports on novel combinatorial biochemistry occurring in metabolic engineering of heterologous plant hosts, resulting in novel sesquiterpene lactone derivatives with the potential to be new drug leads. The use of transient expression and metabolomics for unexpected product identification are technologies that will be of great value to others working in the field of metabolic engineering. The strategies for identification and characterization of candidate genes, the strategies and tools for metabolic engineering and the possibilities to further improve pathway metabolic engineering are discussed in Chapter 6.
Samenvatting

Parthenolide is de belangrijkste biologisch actieve stof in moederkruid (*Tanacetum parthenium*) en heeft zowel anti-ontstekings als anti-kanker activiteit. In **hoofdstuk 1** wordt een historisch overzicht gegeven van het onderzoek aan parthenolide in moederkruid. Als anti-kanker medicijn heeft parthenolide de belangstelling van medische instituten en de farmaceutische industrie. Een zoekopdracht in GOOGLE patenten geeft meer dan 2000 hits die variëren van extractie protocollen voor parthenolide tot het gebruik van parthenolide als medicijn tegen kanker. Echter, information over de biosynthese van parthenolide is in deze database niet te vinden. De opheldering van de parthenolide biosynthese route zou nieuwe mogelijkheden kunnen geven voor de productie van deze stof in efficiëntere productie systemen. Om de biosynthese route van parthenolide te ontrafelen, is kennis van de weefsels waar parthenolide wordt gemaakt en opgeslagen in de plant van belang. In **hoofdstuk 2** is vastgesteld dat parthenolide vooral ophoopt in de zogenaamde trichomen, die zich met name bevinden op de bloemen. Dit duidt er op dat de biosynthese ook voornamelijk in deze cellen actief is. Deze trichomen zijn vervolgens gebruikt om met behulp van gedegenereerde primers en PCR het germacreen A synthase gen (*TpGAS*) te isoleren dat codeert voor het eerste enzym in de parthenolide biosynthese route. De transcriptie van *TpGAS* was inderdaad het hoogst in de trichomen, waar ook de hoogste concentratie parthenolide wordt gevonden. Dit bevestigt dat de trichomen op de bloemen de weefsels zijn waar de parthenolide biosynthese plaatsvindt.

Tijdens mijn werk aan hoofdstuk 2 publiceerde een Canadese groep over een germacreen A oxidase (GAO) uit sla. Omdat in die tijd de sequentie informatie over de genen die in de moederkruid trichomen tot expressie komen nog niet bekend was hebben we besloten om gebruik te maken van een sequentie database van witlof (witlof maakt net als moederkruid ook costunolide, de waarschijnlijke precursor is van parthenolide). In **hoofdstuk 3** zijn vier cytochrome P450s (van de CYP71 klasse) geselecteerd uit de witlof sequentie database voor functionele karakterisering in gist. Eén van deze cDNAs (CYP71BL3) bleek inderdaad te coderen voor het costunolide synthase enzym (COS) dat de oxidatie van germacra-1(10),4,11(13)-trien-12-oic zuur naar costunolide katalyseert. De biosynthese route tot aan costunolide kon transient tot expressie gebracht worden in *Nicotiana benthamiana* middels agroinfiltratie van *TpGAS*, *CiGAO* en *CiCOS*, wat resulteerde in een costunolide productie van 60 ng.g⁻¹ FW. Bovendien werden twee nieuwe componenten gevormd, die werden geïdentificeerd als costunolide-glutathion en costunolide-cysteine conjugaten.

Toen de 454 sequenties van de moederkruid trichoom cDNA bank beschikbaar kwamen zijn we doorgegaan met de identificatie van additionele genen van de biosynthese route van parthenolide. In **hoofdstuk 4** is de hele parthenolide biosynthese route opgehelderd door isolatie en karakterisering van de moederkruid *TpGAS, TpGAO, TpCOS* en het parthenolide synthase, *TpPTS*. Ook werd nog een enzym geïdentificeerd dat parthenolide kan hydroxyleren. De parthenolide biosynthese route werd in zijn geheel tot expressie gebracht in *N. benthamiana*. In de agro-geïnfiltreerde planten werden parthenolide en een aantal conjugaten
daarvan (met cysteine of glutathion) geproduceerd. In een anti-kanker assay bleek dat deze relatief polaire conjugaten zeer actief waren tegen darmkanker cellen, met maar iets minder activiteit dan vrij parthenolide. Omdat de conjugatie en hydroxylatie van parthenolide de wateroplosbaarheid verbeteren, opent dit nieuwe mogelijkheden voor de toepassing van parthenolide als medicijn. Naast genen van de biosyntheseroute van parthenolide hebben we ook twee andere P450 genen geïdentificeerd die coderen voor enzymen die costunolide als substraat gebruiken. In hoofdstuk 5 is Tp8879 geïdentificeerd als een enzym dat het monocyclische germacranolide sesquiterpeen lacton costunolide omzet in het bicyclische guaianolide sesquiterpeen lacton, kauniolide, en dit enzym is daarom kauniolide synthase genoemd. De biosynthese route van kauniolide werd gereconstitueerd in *N. benthamiana* door middel van transiente expressie. In mijn proefschrift heb ik een aantal bestaande en nieuwe technieken gecombineerd voor ‘gene-discovery’, te weten transcriptomics en metabolomics, en heb ik gewerkt aan de optimalisatie van de metabole engineering van planten, gebruikmakend van transiente expressie in *N. benthamiana*. Bovendien heb ik nieuwe ‘combinatorial biochemistry’ aangetoond, die optreedt tijdens de metabole engineering die ik heb uitgevoerd in de heterologe gastplant, *N. benthamiana*, wat resulteerde in nieuwe sesquiterpeen lacton derivaten met potentieel voor de ontwikkeling van nieuwe medicijnen. Het gebruik van transiente expressie en metabolomics voor de identificatie van onverwachte nieuwe producten zijn technologieën die van grote waarde zijn voor anderen die werken in het veld van de metabole engineering. De strategie voor het identificeren en karakteriseren van kandidaat genen en de strategie en mogelijkheden van metabole engineering om biosynthese en productie van belangrijke metabollen verder te verbeteren worden bediscussieerd in hoofdstuk 6.
I’m feeling excited to finish my PhD, although it’s my second PhD degree. During the past 4 years, I became a confident and independent researcher through the PhD training in Wageningen University. This could not have been achieved without the help from my promoters, supervisors, colleagues, co-operates, friends, and families. Here I would like to take this opportunity to express my thanks and appreciations to all of them.

First of all, I would like to thank my promoter Prof. dr. Harro Bouwmeester for giving me the opportunity to join his research group. Harro, I still remember the time when we first met. You went to China for a meeting, and I picked you up from the airport. During the way from the airport to your hotel, we discussed a lot of scientific issues. I was impressed by your contagious scientific curiosity and your quick response. After I started working in your group, I didn’t do a very good job at the very beginning. I am grateful for your support. It is your continuous support that encouraged me to propose new ideas and new solution to overcome every difficulty. Your research vision and scientific thinking enlightened me to think and act like a scientist. Thanks a lot for your effort and help through every step of my PhD.

Special thanks go to my supervisor, Sander van de Krol. It is from you that I see one can enjoy playing with scientific research so well. Because of that, I begin to enjoy doing my research during the second half of my PhD study. One activity I would like to mention specially is the writing week (hard part) combined with boat trip (relaxed part) afterwards. It forced and helped us at the same time to start writing about whatever results we’ve got at that moment, which is really really helpful for me. Thank you for your elaborated instructions on my papers, PPT slides and this thesis, which helped me to improve my writing and presenting skills a lot.

My sincere thanks go to Carolien Ruyter-Spira. Thank you for your guidance during the first 8 months of my research in Plant Physiology. Your kindness and patience helped me to adjust to the new environment. I really appreciate that. By the way, it is with your help that I get used to planning my experiment for the coming month. Thank you for being supportive all the time.

I would like to thank the technician of my PhD project, Miriam Goedbloed. Thank you for your help with the cloning of most of the candidate genes in
I would like to thank the people in the TERPMED project, which my PhD project is a part of. Ric de Vos, thank you for being always available to answer my questions and help me out with experimental difficulties, especially LC-MS data analysis. Angelos K. Kanellis, thank you for your trust and inviting me to present my work in the TERPNET 2013 conference as an invited speaker. People from PRISNA, Justin Fischedick, Karel Lund, Annelies Schulte, thank you for all the standards you sent me. Albert Ferrer, David Manzano, Irene Pateraki, thank you for the nice cooperation. Alain Tissier, Kathleen Brückner, Ana Simonovic, Sladana Todorović, Dragana Bozic and Aalt-Jan van Dijk, thank you for all the discussions and suggestions during the TERPMED project meetings.

I would like to express my appreciation to my colleagues and friends in the Terpene lab 1.056. Francel, Bert and Jos, thank you for your help with LC-MS, GC-MS and qPCR analysis. I’ve learned a lot about how to operate those machines and analyse the data. Katja, thanks a lot for your help in and outside the lab. Rina, thanks for the help with all the administrative things. And all my PhD fellows in the lab, Neli, Aldana, Jimmy, Benyamin, Thierry, Lemeng, Jun, Bo, Yuanyuan, Mohammad, thanks you for creating such a pleasant working atmosphere. My friend Danijela, thank you for your help with the Inula work. I miss our discussions during the coffee break in the afternoon. My students Namraj, Rashmi, and Lea, thank you for your valued contribution to my PhD work.

To my friends in Wageningen, Ningwen Zhang and Ke Lin, Wei Song, Xi Cheng, Chunxu Song and Wei Qin, Yanfen Lin and Hieng Ming (Jimmy) Ting, Yanru Song and Xu Cheng, Na Li and Wei Liu, Hanzi He, Hui Li, Yanxia Zhang, Tingting Xiao, Lisha Zhang and Chenlei Hua, Xi Chen and Weicong Qi, Ying Li, Songlin Xie, Wei Gao, Suxian Zhu, Feng Zhu, Liping Gao, Jifeng Tang and Jinbo Wan, Xi Wan, Ke Peng, Xianwen Ji, Huchan Li, Bing Bai, Chunting Lang, Yan Wang, Margriet Roelse and Martijn Luking, Benyamin, Houshyani Hassanzadeh, I would like to thank you for your great company and support. With all of you, Ting and I had a very good time in the Netherlands.

I am very grateful to my promoter during my first PhD study in Huazhong Agricultural University in China, Prof. Xiuxin Deng. You opened the door of
scientific research to me and showed me how to become a qualified researcher. I thank my loving parents Xianqi Wang and Xinzhen Liu. Thank you for being supportive all the time and giving me the opportunity to make my own decisions and follow my own path. Special thanks for my brother Xiaowei Wang. You have no idea how much I admired you when we were growing up. To my parents-in-laws, thank you for raising such an amazing daughter.

My daughter, Yiyi. Thank you for bringing so much happiness and joyfulness to the family. You are my angel. For sure, I will love you forever!

Finally, my wife and soul mate, Ting. I feel like the luckiest guy in the world to have spent the past 11 years and going to spend the rest of my life with you!
Curriculum Vitae

Qing Liu was born on 21 January 1982 in Songzi City, Hubei province, China. Qing performed his bachelor study between 1999 and 2003 with a major in Horticulture in Huazhong Agricultural University and with an additional major in English in Huazhong University of Science and Technology. From September 2003 to October 2008, he performed his first PhD study under the supervision of Professor Xiuxin Deng in Huazhong Agricultural University, focusing on ‘Revealing the molecular mechanism of a red-flesh citrus mutant’. After that, he came to the Netherlands and temporarily worked as a research assistant in the group of Harro Bouwmeester at Wageningen UR. Started from August 2009, he became a PhD student in the same group under the supervision of Harro Bouwmeester and Sander van de Krol. This thesis summarizes the research of his study in elucidation of the sesquiterpene lactone biosynthetic pathway in feverfew, a medicinal plant.
Publication list

Publications related to this thesis:


Other publications:


**Liu Y, Liu Q**, Xiong J, Deng X. (2007). Difference of a citrus late-ripening mutant (*Citrus sinensis*) from its parental line in sugar and acid metabolism at the fruit ripening stage.


* authors contributed equally to this publication.
# Education Statement of the Graduate School

## Experimental Plant Sciences

**Issued to:** Qing Liu  
**Date:** 2 December 2013  
**Group:** Plant Physiology, Wageningen University & Research Centre

### 1) Start-up phase

<table>
<thead>
<tr>
<th>Activity</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>First presentation of your project</td>
<td>Feb 08, 2010</td>
</tr>
<tr>
<td>Writing or rewriting a project proposal</td>
<td>2009</td>
</tr>
<tr>
<td>Writing a review or book chapter</td>
<td></td>
</tr>
<tr>
<td>MSc courses</td>
<td></td>
</tr>
<tr>
<td>Laboratory use of isotopes</td>
<td></td>
</tr>
</tbody>
</table>

**Subtotal Start-up Phase** 7.5 credits*

### 2) Scientific Exposure

<table>
<thead>
<tr>
<th>Activity</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS PhD student days</td>
<td></td>
</tr>
<tr>
<td>EPS PhD student day, Wageningen University</td>
<td>May 20, 2011</td>
</tr>
<tr>
<td>EPS ExPerceptionS Day (EPS Career Day), Wageningen University</td>
<td>Feb 01, 2013</td>
</tr>
<tr>
<td>EPS theme symposia</td>
<td></td>
</tr>
<tr>
<td>EPS Theme 3 Symposium ‘Metabolism and Adaptation’, Leiden University</td>
<td>Feb 19, 2010</td>
</tr>
<tr>
<td>EPS Theme 3 Symposium ‘Metabolism and Adaptation’, Wageningen University</td>
<td>Feb 10, 2011</td>
</tr>
<tr>
<td>NWO Lunteren days and other National Platforms</td>
<td></td>
</tr>
<tr>
<td>ALW meeting ‘Experimental Plant Sciences’, Lunteren</td>
<td>Apr 06-07, 2009</td>
</tr>
<tr>
<td>ALW meeting ‘Experimental Plant Sciences’, Lunteren</td>
<td>Apr 19-20, 2010</td>
</tr>
<tr>
<td>ALW meeting ‘Experimental Plant Sciences’, Lunteren</td>
<td>Apr 04-05, 2011</td>
</tr>
<tr>
<td>ALW meeting ‘Experimental Plant Sciences’, Lunteren</td>
<td>Apr 02-03, 2012</td>
</tr>
<tr>
<td>Seminars (series), workshops and symposia</td>
<td></td>
</tr>
<tr>
<td>MIni symposium: How to write a world-class paper?</td>
<td>Oct 26, 2010</td>
</tr>
<tr>
<td>Seminar: ‘Metabolomics-based functional genomics - from Arabidopsis to crops and medicinal plants’ by Professor Kazuki Saito</td>
<td>Apr 08, 2013</td>
</tr>
<tr>
<td>Seminar: ‘Glandular trichomes of tomato: from terpene biosynthesis to trichome differentiation’ by prof. Alain</td>
<td>May 03, 2013</td>
</tr>
<tr>
<td>Seminar: ‘Using the Nicotiana-TMV system to study resistance gene evolution and plant genome stability by Professor Hanhui Kuang’</td>
<td>Sep 11, 2013</td>
</tr>
<tr>
<td>Seminar series: Plant Physiology-Genetics, Wageningen University</td>
<td>2009-2013</td>
</tr>
<tr>
<td>Seminar plus</td>
<td></td>
</tr>
<tr>
<td>International symposia and congresses</td>
<td></td>
</tr>
<tr>
<td>EU meeting TERPMED project (Baeclona, Spain)</td>
<td>Jul 03, 2009</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Chania, Greece)</td>
<td>May 14, 2010</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Wageningen, The Netherlands)</td>
<td>Oct 28, 2010</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Kalmar, Sweden)</td>
<td>May 22, 2011</td>
</tr>
<tr>
<td>TERPNET 2011 (Kalmar, Sweden)</td>
<td>Nov 03-04, 2011</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Halle, Germany)</td>
<td>Apr 23-24, 2012</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Leiden, The Netherlands)</td>
<td>Oct 25-26, 2012</td>
</tr>
<tr>
<td>EU meeting TERPMED project (Barcelona, Spain)</td>
<td>May 30-31, 2013</td>
</tr>
<tr>
<td>International Scientific Workshop on Plant Secondary Metabolism (Wuhan, China)</td>
<td>Jul 05-07, 2013</td>
</tr>
<tr>
<td>Presentations</td>
<td></td>
</tr>
<tr>
<td>Second Joint Retreat of Phd Students in Plant Sciences, Cologn, Germany (poster)</td>
<td>Apr 15-17, 2010</td>
</tr>
<tr>
<td>Summer school ‘Rhizosphere Signaling’ Wageningen (oral)</td>
<td>Aug 25, 2010</td>
</tr>
<tr>
<td>EU meeting TERPMED project Chania, Greece (oral)</td>
<td>May 14, 2010</td>
</tr>
<tr>
<td>EU meeting TERPMED project Wageningen (oral)</td>
<td>Oct 28, 2010</td>
</tr>
<tr>
<td>EPS Theme 3 Symposium Wageningen (oral)</td>
<td>Feb 10, 2011</td>
</tr>
<tr>
<td>TERPNET 2011 Kalmar, Sweden (poster)</td>
<td>May 23-26, 2011</td>
</tr>
<tr>
<td>EU meeting TERPMED project Kalmar, Sweden (oral)</td>
<td>May 22, 2011</td>
</tr>
<tr>
<td>EU meeting TERPMED project Halle, Germany (oral)</td>
<td>Nov 03-04, 2011</td>
</tr>
<tr>
<td>EU meeting TERPMED project Leiden (oral)</td>
<td>Apr 23-24, 2012</td>
</tr>
<tr>
<td>EU meeting TERPMED project Barcelona, Spain (oral)</td>
<td>Oct 25-26, 2012</td>
</tr>
<tr>
<td>EU meeting TERPMED project Chania, Greece (oral)</td>
<td>May 30, 2013</td>
</tr>
<tr>
<td>TERPNET 2013 Chania, Greece (oral)</td>
<td>Jun 02, 2013</td>
</tr>
<tr>
<td>International Scientific Workshop on Plant Secondary Metabolism Wuhan, China (oral)</td>
<td>Jul 05-07, 2013</td>
</tr>
<tr>
<td>IAB interview</td>
<td>Nov 05, 2012</td>
</tr>
<tr>
<td>Excursions</td>
<td></td>
</tr>
</tbody>
</table>

**Subtotal Scientific Exposure** 27.6 credits*

---

*Credits calculated based on the duration and significance of each activity as outlined in the document. The numbers may vary slightly due to the nature of the subject matter and the specific context provided in the statements.
3) In-Depth Studies

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS courses or other PhD courses</td>
<td>Aug 23–25, 2010</td>
</tr>
<tr>
<td>Summer school ‘Rhizosphere Signaling’</td>
<td>Aug 30–Sep 03, 2010</td>
</tr>
<tr>
<td>Bioinformatics- A user's Approach</td>
<td>Apr 26–28, 2011</td>
</tr>
<tr>
<td>Plant Metabolomics</td>
<td></td>
</tr>
<tr>
<td>Journal club</td>
<td></td>
</tr>
<tr>
<td>Plant Research International, Bouwmeester &amp; Jongsmi associate group</td>
<td>2009-2012</td>
</tr>
<tr>
<td>Literature discussion; Plant physiology group, Wageningen University</td>
<td>2009-2013</td>
</tr>
<tr>
<td>Individual research training</td>
<td></td>
</tr>
<tr>
<td>Subtotal In-Depth Studies</td>
<td>6.3 credits*</td>
</tr>
</tbody>
</table>

4) Personal Development

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill training courses</td>
<td></td>
</tr>
<tr>
<td>PhD Competence Assessment, Wageningen University</td>
<td>Mar 16, 2010</td>
</tr>
<tr>
<td>Techniques for writing and presenting a scientific paper</td>
<td>Feb 14-17, 2012</td>
</tr>
<tr>
<td>Advanced Course Guide to Scientific Artwork</td>
<td>May 06-07, 2013</td>
</tr>
<tr>
<td>Adobe InDesign</td>
<td>May 08, 2013</td>
</tr>
<tr>
<td>How to write a convincing Research Proposal</td>
<td>May 27, 2013</td>
</tr>
<tr>
<td>Organisation of PhD students day, course or conference</td>
<td></td>
</tr>
<tr>
<td>Summer school ‘Rhizosphere Signaling’ session organiser</td>
<td>Aug 25, 2010</td>
</tr>
<tr>
<td>Membership of Board, Committee or PhD council</td>
<td></td>
</tr>
<tr>
<td>Subtotal Personal Development</td>
<td>4.2 credits*</td>
</tr>
</tbody>
</table>

TOTAL NUMBER OF CREDIT POINTS*: 45.6

Here 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.
Funding

The research of this thesis was funded by the European Commission (EU-project TerpMed, Plant Terpenoids for Human Health: a chemical and genomic approach to identify and produce bioactive compounds, Grant agreement no.: 227448).

Financial support from Wageningen University and Research centre for printing this thesis is gratefully acknowledged.
A lot of science is detective work: observing, collecting evidence, finding clues, reconstructing events, establishing theories and drawing conclusions. Detective work with techniques straight out of Crime Scene Investigation (CSI) makes science almost romantic.

Adapted from Roelof Kleis & Bob Ramaker in RESOURCE magazine 2013.