Localization and in-vivo characterization of thapsia garganica CYP76AE2 indicates a role in thapsigargin biosynthesis

Andersen, Trine Bundgaard; Martinez-Swatson, Karen Agatha; Rasmussen, Silas Anselm; Boughton, Berin Alain; Jørgensen, Kirsten; Andersen-Ranberg, Johan; Nyberg, Nils; Christensen, Søren Brøgger; Simonsen, Henrik Toft

Published in:
Plant Physiology

Link to article, DOI:
10.1104/pp.16.00055

Publication date:
2017

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Localization and characterization of CYP76AE2 part of thapsigargin biosynthesis in Thapsia garganica

Andersen, Trine Bundgaard; Martinez-Swatson, Karen Agatha; Rasmussen, Silas Anselm; Boughton, Berin Alain; Jørgensen, Kirsten; Andersen-Ranberg, Johan; Nyberg, Nils; Christensen, Søren Brøgger; Simonsen, Henrik Toft

Published in:
Plant Physiology

Link to article, DOI:
10.1104/pp.16.00055

Publication date:
2017

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Localization of thapsigargin biosynthesis

Corresponding author(s):
Henrik Toft Simonsen
Department of Biotechnology and Biomedicine, Technical University of Denmark Søltofts Plads, 2800 Kgs. Lyngby, Denmark
+45 26986684
hets@dtu.dk
Localization and in-vivo characterization of *Thapsia garganica* CYP76AE2 indicates a role in thapsigargin biosynthesis

Trine Bundgaard Andersen, Karen Agatha Martinez-Swatson, Silas Anselm Rasmussen, Berin Alain Boughton, Kirsten Jørgensen, Johan Andersen-Ranberg, Nils Nyberg, Søren Brøgger Christensen, Henrik Toft Simonsen

Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark (T.B.A., K.J., J.A-R.); Natural History Museum of Denmark, University of Copenhagen, Østervoldgade 5-7, DK-1350 Copenhagen K, Denmark (K.A.M.); Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, 2800 Kgs. Lyngby, Denmark (K.A.M., S.A.R., H.T.S.); Metabolomics Australia, School of BioSciences, University of Melbourne, VIC, Australia, 3010 (B.A.B.); Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark (N.N., S.B.C.)

One sentence Summary

The secretory ducts in the root of *Thapsia garganica* harbor the cytotoxin thapsigargin and the cells lining these ducts express the first enzymes in the biosynthesis of thapsigargin.

---

1 This work was supported by SpotLight, a grant from the Danish Council for Strategic Research (T.B.A., H.T.S., S.B.C.). The work was supported by the Center for Synthetic Biology “bioSYNergy”, UCPH Excellence Program for Interdisciplinary Research (J.A-R.) and by MEDPLANT, a Marie Curie Actions Initial Training Network (K.A.M.). NMR equipment used here was purchased via grant #10-085264 from The Danish Research Council for Independent Research | Nature and Universe.

2 Present address: 441 Koshland Hall, Berkeley, CA 94720-3102, USA, joar@berkeley.edu

3 Present address: Vallgatan 5, Bruker BioSpin Scandinavia AB, SE-170 67 Solna, Sweden

* Address correspondence to hets@dtu.dk

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Henrik Toft Simonsen (hets@dtu.dk).

T.B.A. established the major part of the results and prepared the major part of the paper. K.A.M. prepared samples for MALD, performed the in situ PCR experiment and contributed to the manuscript. S.A.R. purified epidihydrocostunolide. S.B.C., S.A.R. and N.N. performed the NMR. B.A.B. performed the MALDI analysis. J.A.R. did part of the purification of compound 2 and 3, K.J. contributed to the in situ PCR experiment. H.T.S. initiated, directed and supported the research and writing of the manuscript. All authors edited and approved the final manuscript.
Abstract

The Mediterranean plant *Thapsia garganica* (dicot, Apiaceae), also known as Deadly carrot, produces the highly toxic compound thapsigargin. This compound is a potent inhibitor of the SERCA calcium pump in mammals, and is of industrial importance as the active moiety of the anticancer drug Mipsagargin, currently in clinical trials. Knowledge of thapsigargin *in planta* storage and biosynthesis has so far been limited. Here we present the putative second step in thapsigargin biosynthesis, by showing that the cytochrome P450 TgCYP76AE2, transiently expressed in *Nicotiana benthamiana*, converts epikunzeaol into epidihydrocostunolide. Furthermore, we show that thapsigargin is likely to be stored in secretory ducts in the roots. Transcripts from TgTPS2 (epikunzeaol synthase) and TgCYP76AE2 in roots were only found in the epithelial cells lining these secretory ducts. This emphasizes the involvement of these cells in the biosynthesis of thapsigargin. This study paves the way for the further studies of thapsigargin biosynthesis.
Introduction

Sesquiterpenoids are widely distributed across the plant kingdom and are recognized for their pharmacological properties and commercial value (Simonsen et al., 2013). Artemisinin, which today is the cornerstone for treatment of malaria, is an outstanding example (Wiesner et al., 2003; Tu, 2011), along with fragrances such as patchoulol and santalol (Zhan et al., 2014). The genus *Thapsia* L. ( Apiaceae) produces a variety of sesquiterpenoids including sesquiterpene lactones (Christensen et al., 1997; Drew et al., 2009). The most studied sesquiterpene from the genus is the sesquiterpene lactone thapsigargin, which is the most predominant sesquiterpene lactone in *Thapsia garganica* L. and *Thapsia gymnesica* Rouy (Christensen et al., 1997). In the Mediterranean area *T. garganica* has been used in traditional medicine for over 2000 years for the treatment of pulmonary diseases, catarrh, fever, pneumonia and as a counter irritant for the relief of rheumatic pains (Andersen et al., 2015). The pharmacological effect of thapsigargin has been studied thoroughly and it has been established that thapsigargin is an inhibitor of the sarco-endoplasmic reticulum Ca$^{2+}$-ATPase (SERCA) that leads to cell apoptosis (Thastrup et al., 1990). A pro-drug (Mipsagargin®) based on thapsigargin towards solid cancer tumors is currently in clinical trials (Doan et al., 2015; Mahalingam et al., 2016).

An unusual feature of thapsigargin and related guaianolides is the presence of a β-disposed C-6-O and an α-disposed C-7-C-11 bond. In the majority of guaianolides from other plant families, the C-6-O bond is α-disposed and the C-7-C-11 bond β-disposed (Christensen et al., 1997; Drew et al., 2009; Simonsen et al., 2013).

Thapsigargin is found in most parts of the plant *T. garganica*. Ripe fruits contain the highest amount of thapsigargin with 0.7-1.5% of the dry weight followed by roots (0.2-1.2% d.w.) and leaves (0.1% d.w.) (Smitt et al., 1995). It is well established that many Apiaceae species store lipophilic compounds such as phenyl propanoids and terpenoids in secretory ducts and vittae (Corsi et al., 1988; Poli et al., 1995; Maggi et al., 2015). We chose to examine the localization of sesquiterpenoids and their biosynthesis in *T. garganica* roots that have high amounts of thapsigargin. The roots could also be obtained commercially from plants grown in greenhouses, in contrast to the seasonal dependent harvest of fruits from the natural population. Fruits have so far not been obtained from greenhouse plants, not even from plants more than four years old. By histochemical staining, we show that *T. garganica* roots contain secretory ducts in parenchymatic tissue and that these may harbor terpenoids. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry Imaging (MALDI-MSI) of the roots was used to show that thapsigargin is present in
specific locations in the root and furthermore in a pattern likely to coincide with the location of
secretory ducts.
The site and route of thapsigargin biosynthesis have not yet been established. The first specific step
in the biosynthesis of most sesquiterpenoids is catalyzed by sesquiterpene synthases (Bohlmann et
al., 1998). Two sesquiterpene synthases have previously been described from T. garganica roots.
These were expressed in Saccharomyces cerevisiae and biochemically characterized. It was shown
that the sesquiterpenoid synthase TgTPS2 was of particular interest to thapsigargin biosynthesis.
The major product of TgTPS2 is epikunzeaol, a germacrenol with a hydroxyl group at C-6 (Figure
1) (Pickel et al., 2012).
Generally, the diversity of sesquiterpenoids is obtained by the catalytic activity of the sesquiterpene
synthases and followed by modifications to the C₁₅ backbone by cytochromes P450 (P450s), acyl
transferases and dehydrogenase amongst others (Weitzel and Simonsen, 2015). Complexity of
sesquiterpenoid structures, including chiral centers and regio- and stereospecific oxidations, makes
chemical synthesis difficult and synthesis approaches often result in low yields (Andrews et al.,
2007; Ball et al., 2007; Chu et al., 2016). Attempts to find alternative sustainable production
methods for sesquiterpenoids such as artemisinin by biological synthesis are actively pursued
(Paddon et al., 2013). Sesquiterpenoid lactones are a subgroup of sesquiterpenoids, and compounds
in this subgroup have been shown to have a potential for treatment of various cancers (Curry et al.,
2004; Simonsen et al., 2013). Despite recent advances, only a few steps in the complex biosynthetic
routes of plant sesquiterpenoid lactones have been characterized. These include enzymes involved
in the biosynthesis of artemisinin (Artemisia annua L.) and (+)-costunolide, a precursor for a range
of other sesquiterpene lactones (in Cichorium intybus L. and Tanacetum parthenium L.). These
plants are all from the family Asteraceae (Yu and Wen, 2011; Liu et al., 2014).
P450s are common participants in sesquiterpenoid biosynthesis and especially P450s from the
CYP71 clan have been shown to be involved in the biosynthesis of sesquiterpenoids (Luo et al.,
2001; Diaz-Chavez et al., 2013; Liu et al., 2014; Takase et al., 2015; Weitzel and Simonsen, 2015;
Yang et al., 2015). CYP71AV1, which is involved in artemisinin biosynthesis, is among the most
well studied P450s in sesquiterpenoid biosynthesis (Teoh et al., 2006). Here we functionally
characterize TgCYP76AE2, a P450 from the CYP71 clan, which was found in the root
transcriptome of T. garganica. Through the transient co-expression in Nicotiana benthamiana of
TgTPS2 and TgCYP76AE2, epikunzeaol is converted to epidihydrocostunolide, a likely precursor
for more complex sesquiterpenoid lactones including thapsigargin. Within the Apiaceae, only P450s
involved in the biosynthesis of the phenylpropanoid coumarine have been described (Larbat et al., 2007; Larbat et al., 2009; Drew et al., 2013; Dueholm et al., 2015). This work presents the functional characterization of the first sesquiterpene specific P450 found in the Apiaceae. 

*In situ* PCR was performed to investigate the specific site for expression of *TgTPS2* and *TgCYP76AE2*. We were thereby able to verify that *TgTPS2* and *TgCYP76AE2* are expressed in epithelial cells lining the secretory ducts containing thapsigargin.

*Figure 1. The structure of epikunzeaol and epidihydrocostunolide (1). Both metabolites are suggested intermediates in the thapsigargin pathway in *T. gorganica.*
Results

Identification of cytochromes P450 from transcriptome data

The root transcriptome data from *T. garganica* were mined to enable the discovery of possible sequential biosynthesis steps from epikunzeaol toward thapsigargin. The search was limited to P450s from the CYP71 clan due to the previous findings in this clan of P450s involved in sesquiterpenoid biosynthesis. We investigated the occurrence of orthologous genes to CYP71’s in the *T. garganica* transcriptome using BLAST searches. The eighteen full-length P450 genes were found to be distributed with twelve belonging to the CYP71 family, one to the CYP706 family and five to the CYP76 family. The P450s were named by David Nelson according to the current annotation system (Nelson, 2009); TgCYP71AH8 (KX826939), TgCYP71AJ5 (KP191555), TgCYP71AJ14 (KP191558), TgCYP71AS14 (KX845553), TgCYP71AT12 (KX826940), TgCYP71AU89 (KX845548), TgCYP71AU90 (KX845552), TgCYP71BK1 (KX826941), TgCYP71BK6 (KX845546), TgCYP71D183 (KX845554), TgCYP71D311 (KX845555), TgCYP71D319_ortholog (KX845550), TgCYP76AE1 (KX826942), TgCYP76AE2 (KX826943), TgCYP76AE8 (KX845545), TgCYP76AF7 (KX845549), TgCYP76B79 (KX845547) and TgCYP706C30_ortholog (KX845551).

Phylogeny

Phylogenetic analyses of 35 full-length genes from the CYP71 clade involved in sesquiterpenoid biosynthesis from several plant species and the 18 enzymes from *T. garganica* revealed that these enzymes are grouped in several subclades (Figure 2). The analysis shows that there are blooms of genes within species and families (Hamberger and Bak, 2013), like the CYP76F bloom in *Santalum album*. Although, the phylogeny only included P450s related to sesquiterpenoid metabolism, the sequences are from most families and subfamilies in the CYP71 clade. This shows that phylogenetic analysis are not a useful to predict specific functionality of P450s, as also shown previously (Dueholm et al., 2015). The analysis merely serve to indicate what range of enzymes that have to be examined biochemically; in this case 18 sequences.
GC-MS and LC-MS analysis of extracts from *N. benthamiana* expressing TgTPS2 and cytochromes P450, including TgCYP76AE2

The cyclodeca-1(10),4-diene ring of epikunzeaol (Figure 1) is susceptible to thermal Cope rearrangement at the high temperatures in the GC-MS injection port. Extracts of *N. benthamiana* expressing *At*HMGR and TgTPS2 were therefore analyzed with the two injection port temperatures, 250°C and 160°C. At 250°C, the TgTPS2 product epikunzeaol is detected along with degradation compounds whereas at 160°C, only epikunzeaol is detected (Figure 3A). Similar rearrangement have been identified for other cyclodeca-1(10),4-diene products of sesquiterpene synthases (Andersen et al., 2015). To enhance the level of precursors available for TgTPS2 in *N.*
benthamiana, a truncated version of the Arabidopsis thaliana HMGR, AtHMGR, was transiently co-expressed. AtHMGR has previously been shown to enhance production levels of sesquiterpenoids (Cankar et al., 2015).

In order to discover, which P450(s) that could utilize epikunzeaol as substrate the eighteen P450s from the CYP71 clan were transiently co-expressed with TgTPS2 and AtHMGR in N. benthamiana.

Of the tested P450s, only TgCYP76AE2 was evidently able to utilize epikunzeaol as a substrate. To test for further downstream pathway steps the remaining P450s were co-expressed with AtHMGR, TgTPS2 and TgCYP76AE2. No new products or decline in substrate was detected. At this stage, it cannot be excluded that new products were not detected due to low expression or lack of expression of the seventeen P450s.

GC-MS analysis of the hexane extracts of N. benthamiana leaves expressing AtHMGR, TgTPS2 and TgCYP76AE2 is shown in Figure 3B, where two new products, 2 and 3, are observed. It was not possible to detect these products with the injection port at 160°C, which can be explained by a lower volatility of the new product(s) in comparison with epikunzeaol. The co-expression of AtHMGR, TgTPS2 and TgCYP76AE2 in N. benthamiana resulted in a complete conversion of epikunzeaol (Figure 3B).
To expand the search for TgCYP76AE2 products or derivatives hereof, not detectable by GC-MS, the N. benthamiana extracts was analyzed by analytical LC-MS. In contrast to the GC-MS analysis, epikunzeaol was not detected in free form, but as the aglycon in a glycoside of a disaccharide, and only one TgCYP76AE2 product was detected (Figure 4). In the LC-MS analysis product 1 was detected as the protonated molecular ion (m/z 235.23, [M+H]^+), the sodium adduct (m/z 257.23, [M+Na]^+), and the base peak equal to m/z 491.3126 ([2M+Na]^+), corresponding to the sodiated dimer adduct.

Isolation of epidihydrocostunolide (1) and the two 1,3-elemandien-12,6-olides (2 and 3) by HPLC and preparative GC-MS for NMR analysis.

To determine the structure of the three new compounds (1, 2, and 3) these were isolated from the hexane extract of N. benthamiana leaves expressing ArtHMGR, TgTPS2 and TgCYP76AE2. Compound 1 was isolated by semi-preparative normal-phase HPLC and the purity was confirmed by LC-HRMS. Compounds 2 and 3 were isolated using preparative GC-MS. To confirm that 2 and 3 are thermal Cope rearrangement products of 1, pure 1 was injected into the GC-MS (Figure 5). No traces of compound 1 was seen in this GC-MS chromatogram, whereas the mass spectra of the peaks originating in compound 2 and 3 were identical to those obtained from the compounds isolated by preparative GC-MS (Figure S2). This confirm the thermal Cope rearrangement of 1 into 2 and 3.

Structure elucidation of product 1, 2 and 3

The structures of 1, 2 and 3 were elucidated by the interpretation of ^1H, ^13C and 2D COSY, HSQC, HMBC and ROESY spectra. The molecular formula of 1 was established using High Resolution Mass Spectrometry (HRMS) to be C_{15}H_{22}O_{2} (observed m/z 235.1692, calculated for C_{15}H_{22}O_{2} [M+H]^+ m/z 235.1693, 0.2 ppm error). Comparison of the ^1H and ^13C NMR spectra of 1 with that of dihydrocostunolide and in particular the ^13C NMR spectrum (Sanz et al., 1990; Barrero et al., 2002) revealed significant similarities. The structure elucidation was complicated as all nuclei exhibited one major signal and a minor signal, as is clearly observed in the ^13C NMR spectrum (Figure S1). This is most likely an effect caused by the presence of two slowly interconverting conformers of the decadiene ring. However, closer inspection showed that the isolated germacranolide had to be a stereoisomer of the previously proposed structure of dihydrocostunolide. The ROE interactions between H-7, H-11 and H-6 indicated that these protons were all cis-disposed, thus supporting the proposed structure of 1.
From the preparative GC-MS isolation, two products (2 and 3) were obtained (Figure 3B). Compound 2 had the molecular formula C_{15}H_{22}O_{2} as revealed by MS. Inspection of the $^{13}$C NMR spectrum indicated the presence of two signals for methylene groups at 110 and 116, a signal for a quaternary carbon at 144 (Table 1). A signal for a methine group at 149 ppm suggested the presence of a mono substituted and a germinal di-substituted double bond, respectively. A signal for a carbonyl group at 179 ppm indicated the presence of a $\gamma$-lactone. These chemical shift values were

Figure 4: A: Extracted Ion Chromatograms of epikunzeao [epikunzeao disaccharide+Na]+ $m/z$ 569.74 (red) and epidiodyrostomolide [2M+Na]+ $m/z$ 491.32 (blue) of the LC-MS analysis of methanol extract of *AnthOMGR* alone. *AnthOMGR* plus *TPS2*. B: The MS/MS spectrum of the epikunzeao peak at 18 minutes detected as the epikunzeao disaccharide. The MS spectrum of epidiodyrostomolide at 27 minutes shows a $m/z$ of 491.325 corresponding to a sodium adduct and a dimer of $m/z$ 235.1741. C: The structure of epikunzeao disaccharide including the fragmentation pattern.
very similar to those reported for the saussurea lactone (Barrero et al., 2002). Inspection of the \(^1\)H NMR spectrum of \(2\) (Table 2) and the reported spectrum of the saussurea lactone (Ando et al., 1983) however, revealed significant differences between the sizes of the 3-bond coupling constants in \(2\) and in the saussurea lactone. In the saussurea lactone, the lactone ring is trans-fused with the cyclohexane ring, enabling an axial location of H-6 and H-7 and consequently coupling constants of approximately 10 Hz are expected. Assuming that the lactone ring and the cyclohexane ring of \(2\) are cis-fused this would prevent an axial-axial coupling and consequently smaller \(^3J_{HH}\) couplings are expected. Inspection of the ROESY spectrum also revealed ROE-correlation between H-6 and H-7 confirming the cis-fusion. An additional ROE-correlation between H-7 and H-11 revealed that these protons also are cis-disposed. ROE-correlations between CH\(_3\)-15, CH\(_3\)-14 and CH\(_3\)-13 confirm that all of these methyl groups are β-disposed. Based on the data the stereochemistry of \(2\) is suggested as shown in Figure 5 to be (3\(R\),3\(a\)S,6\(S\),7\(S\),7\(a\)R)-3,6-dimethyl-7-(prop-1-en-2-yl)-6-vinylhexahydrobenzofuran-2(3H)-one. The spectra of \(3\) were very similar to those of \(2\). In \(3\) however, a similar ROE-correlation as described above for compound \(1\) was present except for ROE-correlation between CH\(_3\)-15 and CH\(_3\)-13. Combined with similar \(^3J_{HH}\) coupling constants this suggest the stereochemistry of \(3\) shown in Figure 5 to be (3\(R\),3\(a\)S,6\(R\),7\(R\),7\(a\)R)-3,6-dimethyl-7-(prop-1-en-2-yl)-6-vinylhexahydrobenzofuran-2(3H)-one. \(1\) is a new compound for which, we suggest the name epidihydrocostunolide. The stable structure of \(2\) and \(3\) led to the full structure elucidation of \(1\).

**Tissue localization of thapsigargin biosynthesis in roots**

Histochemical staining and MALDI-MS imaging (MALDI-MSI) was performed to determine the localization of thapsigargin and its biosynthesis. Histochemical staining was used to indicate the presence of terpenoids in secretory ducts in root tissue. The location of thapsigargin was analyzed by MALDI-MSI. Based on these results further investigations by in situ PCR was performed to observe if the expression of \(TgTPS2\) and \(TgCYP76AE2\) involved in the biosynthesis of thapsigargin showed co-localization with thapsigargin.
Histochemical analysis of *T. garganica* roots

The presence of special storage structures in the root possibly containing terpenoids was investigated by a histochemical analysis using NADI staining (David and Carde, 1964; Caissard et al., 2004; Jezler et al., 2013; Kromer et al., 2016; Muravnik et al., 2016; Stešević et al., 2016; Stojičić et al., 2016). The NADI reagent has been reported to give rise to a blue/purple color in the presence of oxygenated or lipophilic compounds (e.g. terpenoids) in oil secretory cells such as ducts, trichomes and other specialized tissue. NADI is a two component system consisting of dimethyl-p-phenylenediamine and α–naphathol. The mechanism is suggested to be a non-enzymatic reduction of the oxidized target and an oxidation of dimethyl-p-phenylenediamine followed by the formation of a radical, which react with α–naphathol to produce indophenol blue (Harwig, 1967; Takamatsu and Hirai, 1968).

Prior to staining, *T. garganica* roots were cut in 60 µm cross-sections. As in other *Daucus* species the roots showed a clear secondary growth (Havis, 1939; Korolev et al., 2000). White resin was observed oozing out of the root upon cutting. Figure 6 indicates the presence of oxygenated or lipophilic compounds (e.g. terpenoids) in the sections after staining with NADI and here secretory ducts are clearly visible as blue spots. These were located in concentric rings within the parenchymatic tissue radiating out from the pith until just below the periderm as also shown in *Daucus* (Deutschmann, 1969). At higher magnification, it is evident that both the ducts themselves and the epithelial cells surrounding these were stained blue indicating the presence of oxygenated or lipophilic compounds (e.g. terpenoids). Apart from the ducts in the parenchymatic tissue the periderm also showed blue staining, which could be due to tannins in this tissue. In general, ducts were not found in cross-sections from the crown or the very bottom of the roots, but otherwise generally distributed throughout the root.

MALDI-MSI of *T. garganica* roots

The spatial distribution of thapsigargin, the intermediates epikunzeaol and epihydrocostunolide; including related guaianolides were examined using High Resolution Matrix Assisted Laser Desorption Ionization-Fourier Transform-Ion Cyclotron Resonance-Mass Spectrometry Imaging (HR MALDI-FT-ICR-MSI). MALDI-MSI is able to measure ionized chemicals as their mass-to-charge ratio (*m/z*) in a highly localized manner (Boughton et al., 2016). Root sections of *T. garganica*, that had previously been confirmed to contain thapsigargin by HPLC, were prepared by initial cryo-sectioning, mounting to slides on double sided tape, freeze drying and application of DHB matrix by sublimation (Jarvis et al., 2017). Prepared sections were analyzed by MALDI-MSI.
in the positive ionization mode across the mass range m/z 200-3000. Results demonstrated complex mass spectra containing numerous ions that could be tentatively assigned as small molecule metabolites, sugars and lipids (Figure S3). Specific ions were found to localize to the epidermis,
parenchyma, stele and a series of concentric spots correlating with the observed distribution of
secretory ducts (Figure 7). The acquired spectra were screened for thapsigargin, epikunzeaol and
epidihydrocostunolide by searching for the calculated \( m/z \) of proton, sodium and potassium adducts.
Ions corresponding to thapsigargin, \([\text{M+Na}]^+\) ion \( m/z \) 673.3186 (calc. 673.31945, 1.26 ppm error) and a \([\text{M+K}]^+\) ion \( m/z \) 689.2924 (calc. 689.29339, 1.43 ppm error) were found. Proton adducts of
thapsigargin were barely observed and in general the K adduct showed a much higher signal
intensity relative to the Na adduct (signal intensity ratio of 2.5-3:1). The spatial distribution of both
Na and K adducts showed a distinct distribution within the parenchymatic tissue (Figure 7), where
the distribution pattern correlated with concentric circles of secretory ducts visualized from
histochemical staining (Figure 6). Both epikunzeaol and epidihydrocostunolide were not detected in
the images, indicating that concentrations were below the limit of detection using this methodology.
The root sections were also screened for seven related guanolides from \( T. \) garganica (Table 3).
Structures of the related guanolides and the MALDI imaging results from these are shown in Figure
S4. The seven structurally similar sesquiterpenoids, thapsigargin, nortrilobolide, trilobolide,
thapsivillosin I and thapsivillosin L all showed distributions that coincided with thapsigargin,
including a similar pattern of a more intense \([\text{M+K}]^+\) adduct.

**Localization of mRNA for \( TgTPS2 \) and \( TgCYP76AE2 \)**

Following the specific localization of thapsigargin in the roots and likely in secretory ducts it was
investigated if the production of the compound takes place here as well. Cellular localization of
transcripts encoding \( TgTPS2 \) and \( TgCYP76AE2 \), involved in the biosynthesis, was investigated by
in tube *in situ* PCR. \( TgTPS2 \) and \( TgCYP76AE2 \) transcripts were found to show the same spatial
expression pattern in the investigated tissue and were detected solely in the epithelial cells lining the
secretory ducts (Figure 6A and 6B). The presence of the transcripts were visualized by the color reaction of the alkaline phosphates bound to the antibody specific to the DIG group incorporated during the amplification of the PCR product from the specific produced cDNA using the reverse specific primers recognizing $TgTPS2$ and $TgCYP76AE2$ respectively. In *T. garganica* roots this reaction only took place in the epithelial cells of the secretory ducts when specific reverse primers where included. When no specific reverse primers were used no color reaction were detected. This reaction clearly shows that the ducts follow the circular growth pattern of the parenchymatic tissue. The transcripts show a spatial distribution similar to the m/z corresponding to thapsigargin as is seen in Figure 7.
In this study, \textit{TgCYP76AE2} was found to catalyze the oxidations that lead to the formation of a lactone ring hereby converting epikunzeaol into epidihydrocostunolide (1) (Figure 1). The structure of 2 and 3 confirm the proposed structure of 1 and also finally confirm the structure of epikunzeaol produced by \textit{TgTPS2} as seen in Figure 1. The structure of kunzeaol previously published as the product of \textit{TgTPS2} did not benefit from the Cope rearrangement study included here and the stereochemistry was therefore not final (Pickel et al., 2012). In general, Cope rearrangements are stereoselective and only one product would be expected (Setzer, 2008; Adio, 2009). However, here upon injection of 1 into a GC-MS two elemanolides, 2 and 3, were formed in almost equal amounts (Figure S2). This type of rearrangement has been studied for sesquiterpenoids belonging to the germacrenes and germacranolides. Germacrenes are characteristic by their backbone structure, which is a 10-membered open ring as seen in epikunzeaol (Figure 1). Conversion of germacranolides to elemanolides by Cope rearrangement has previously been used to structurally elucidate the heat labile germacranolides (Fischer and Mabry, 1967; Takeda, 1974; Raucher et al., 1986; de Kraker et al., 2001; Barrero et al., 2002; Adio, 2009). At high temperatures or low pH germacrenes via Cope rearrangement establish equilibrium with elemenes (Takeda, 1974; de Kraker et al., 2002; Pickel et al., 2012). Elemenes are characteristic by the lack of a bond between C2 and C3 (Figure 4). The unexpected formation of two products after Cope-rearrangement of 1 might be explained by the \textit{cis}-fused lactone and instability of the 10-membered decadiene ring. Most reported studies on Cope rearrangements have been performed on \textit{trans}-fused germacranolides, where the preferred chair-chair-conformation of the intermediate is easily accessible (Setzer, 2008; Adio, 2009). One example, however, with Cope rearrangement of a \textit{cis}-fused germacranolide is described. In this example, also only one product is formed in high yield (Appendino and Gariboldi, 1983).

The biochemical lactone ring reaction observed here is also known from Asteraceae where the biosynthesis of (+)-costunolide follow the same type of reaction. In order to establish the lactone ring in (+)-costunolide, two P450s are needed. First three consecutive hydroxylations are performed on C-12 by one P450 (GAO), generating germacra-1(10),4,11(13)-trien-12-ol followed by germacra-1(10),4,11(13)-trien-12-al and resulting in germacra-1(10),4,11(13)-trien-12-oic acid (Nguyen et al., 2010; Cankar et al., 2011; Liu et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014; Liu et al., 2014) (Figure 7). A second P450 (COS) hydroxylates germacra-1(10),4,11(13)-trien-12-oic acid at C-6, resulting in spontaneous lactone formation (Ikezawa et al., 2011; Liu et al.,}
The lactone formation in *T. garganica* is simpler than the formation reported in the Asteraceae species, since epikunzeaol already has a hydroxyl group at C-6 and therefore only requires one P450. Thus, the triple oxidation on C-12 to the carboxylic acid leads to a spontaneous lactone ring formation. Consequently, it is not possible to detect the intermediates that are expected to be an alcohol, aldehyde and acid. The suggested mechanism of epidihydrocostunolide biosynthesis in Figure 7 is based on the previously described mechanism for costunolide in Asteraceae (Nguyen et al., 2010; Cankar et al., 2011; Ikezawa et al., 2011; Liu et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014).

Neither of the two enzymes CYP71BL1 (GAO) and CYP71BL2 (COS) which, catalyze 8β- and 6α-hydroxylation of germacrene A acid, respectively, shows high identity to TgCYP76AE2 with 32.2 % for GAO and 36.6 % for COS (Ikezawa et al., 2011). This is despite Apiaceae and Asteraceae are closely related families, exemplifying limitations with prediction of CYP functionality based on phylogenetic relationship. Sesquiterpenoid lactones are a broad class of compounds and other mechanisms for formation of a lactone ring have been reported. The biosynthesis of the lactone ring in the sesquiterpenoid artemisinin for instance is different to what is found for (+)-costunolide and epidihydrocostunolide. Here an aldehyde dehydrogenase and an aldehyde reductase are involved in addition to a P450. These jointly participate in the formation of an acid that is then non-enzymatically converted into the lactone ring, as opposed to the (+)-costunolide and dihydrocostunolide biosynthesis where ring formation is achieved solely by P450 catalysis (Teoh et al., 2006; Liu et al., 2011).

The P450s, GAO and COS, from Asteraceae that catalyze the lactone ring formation in (+)-costunolide are from the CYP71 family. While the CYP71 family of the CYP71 clan has had much focus, the CYP76 family is now emerging as another major participant in the biosynthesis of specialized metabolites especially terpenoids. CYP76’s have been found to participate in both mono- sesqui- and diterpenoid biosynthesis (Collu et al., 2001; Guo et al., 2013; Weitzel and Simonsen, 2015). In sesquiterpenoid biosynthesis, CYP76F has been described in *Santalum album* as part of the santalol and bergamotol biosynthesis (Diaz-Chavez et al., 2013; Celedon et al., 2016). The CYP71 clan is becoming more of a continuum and it is expected that future sequencing will add to this and make it even harder to distinguish between the families and subfamilies of this clan. Investigation of P450s involved in sesquiterpenoid metabolism are expanding from the CYP71 family to the CYP71 clan, as seen in Figure 2, and in the future probably also beyond that.
Stereochemistry and non-conjugation
A characteristic difference between the guaianolides from Asteraceae and Apiaceae, like thapsigargin, is the α-disposal of the C-6-O bond in Asteraceae whereas it is β-disposed in Apiaceae (Simonsen et al., 2013). The present study support that the β-C-6-O bond characteristic for thapsigargin is introduced already at the very first step of sesquiterpenoid biosynthesis with the formation of epikunzeaol. TgCYP76AE2 converts epikunzeaol to epidihydrocostunolide, which differs from dihydrocostunolide by the β-disposed C-6-O bond. Since no epimerization of C-6 is likely under the oxidative transformation of C-12 into a carboxylic acid the previous suggested structure of epikunzeaol as the 6α-hydroxygermacrene must be considered unlikely (Pickel et al., 2012). The suggested structure was only based on assignment of major signals in the spectrum, which might explain the erroneous assignment of stereochemistry.

The expression of genes, involved in sesquiterpene biosynthesis, in N. benthamiana has previously been reported to result in glycosylation or other conjugations of the produced sesquiterpenoids. For artemisinic acid in the artemisinin pathway, it was seen that this was conjugated to a diglucoside (van Herpen et al., 2010). When the genes in the costunolide biosynthesis was expressed in N. benthamiana costunolide was conjugated to glutathione or cysteine (Liu et al., 2011). Analysis of products of TgCYP76AE2 did not reveal any conjugation and were therefore detectable by GC-MS. In contrast to costunolide, epidihydrocostunolide does not possess a C-11-C-13 double bond preventing a conjugation to a thiol group.

Downstream pathway towards thapsigargin
Costunolide has been suggested as a precursor of many of the studied sesquiterpene lactones in the Asteraceae (de Kraker et al., 2002). The finding of a P450 from T. garganica, which synthesizes epidihydrocostunolide, indicates that epidihydrocostunolide could have a similar role in Thapsia and that the unusual stereochemistry at C-6 is introduced at a very early stage. Further modification of epidihydrocostunolide to produce thapsigargin would require a mechanism to produce a 5- and 7-ring closure of the 10-membered ring of epidihydrocostunolide. The exact reaction and enzyme responsible are yet unknown. A putative route could be the oxidation mediated ring closure similar to what has been described for the biosynthesis of lathyranes in Euphorbiaceae species (Luo et al., 2016). Furthermore, several hydroxylations and various acetylations of the backbone are required. A number of these reactions are likely to be performed by P450s.

Specialized tissue for storage of sesquiterpenoids
Terpenoids are often located in specialized storage structures in plants, which include oil glands, secretory ducts, laticifers, trichomes and vacuoles (Fahn, 1988; Chadwick et al., 2013). Cells adjacent to or harboring these compartments have been shown to be involved in the biosynthesis of terpenoids that are then transported into the storage compartment (Olsson et al., 2009; Lange, 2015). *T. garganica* is closely related to the genus *Daucus* and has a taproot like carrots (Weitzel et al., 2014). In carrots, it was shown that lipophilic compounds including specialized metabolites were primarily localized in extracellular, long schizogenous hydrophobic oil ducts that were located in the periderm/pericyclic parenchyma tissue (Esau, 1940; Schuphan and Boek, 1960; Deutschmann, 1969; Garrod and Lewis, 1980). The secretory ducts in *Daucus* are highly organized and connected throughout the phloem of roots and exhibit a concentric ring pattern in horizontal sections (Deutschmann, 1969; Bowes and Mauseth, 2008). Likewise, a positive correlation between the number of ducts and amount of terpenoids has been demonstrated (Garrod and Lewis, 1980; Senalik and Simon, 1986).

NADI was chosen as a possible stain for terpenoids. The use of NADI has been known since the 1880’s and has since been used to detect oxidized compounds. David and Carde, 1964 were the first to report the use of NADI for staining of terpenoids, unfortunately the exact mechanism nor the specificity were described. Throughout the years, the exact mechanism has been much debated and the specificity for terpenoids is still uncertain. NADI does however, function as a useful method in the current setting to visualize secretory ducts whether the stain is specific towards oxidized and/or lipophilic compounds, including terpenoids.

The finding that transcripts from *TgTPS2* and *TgCYP76AE2* are exclusively found in epithelial cells surrounding secretory ducts in the middle part of the root supports their involvement in the biosynthesis of thapsigargin. Intermediates from the biosynthesis of thapsigargin were not identified in extracts from *T. garganica* in this study nor in previous studies. This could signify a specific and efficient or possibly even a channeled biosynthetic pathway indicating metabolon formation of the enzymes (Møller, 2010). The presence of thapsigargin in highly specific locations also supports the function of ducts as storage compartments in the roots, and possibly also ducts in the fruits of *T. garganica*, which might explain the high content in these two organs. Similar observations were made in stem cross-sections from the conifer *Picea sitchensis* upon methyl jasmonate treatment (Zulak and Bohlmann, 2010). An antibody against the diterpene synthase, levopimaradiene/abietadiene synthase allowed detection of a fluorescence signal in the epithelial cells of cortical and traumatic resin ducts 2 days after methyl jasmonate treatment. In addition, this
indicated the importance of these epithelial cells in the biosynthesis of specialized metabolites. In a following study in *Picea sitchensis* the importance of the epithelial cells were further implemented (Abbott et al., 2010; Hamberger et al., 2011). Here epithelial cells were isolated with laser microdissection, studied by RT-qPCR and found to be enriched in a variety of CYP720s including PsCYP720B4 involved in biosynthesis of isopimaric acid and abietic acid. This approach is of high relevance to the further elucidation of the thapsigargin biosynthetic pathway.

Focus in the present study has been directed towards localizing specific tissue in *T. garganica*, which could biosynthesize and store sesquiterpenoids such as thapsigargin. As it was previously shown that the amount of secretory ducts correlate with the amount of terpenoids and especially lipophilic terpenoids, cells lining these are an obvious target for biosynthesis studies (Senalik and Simon, 1986). The finding that thapsigargin is stored in secretory ducts in the roots and the enzymes involved in the biosynthesis are present in the surrounding cells opens up new possibilities. The data presented here suggest that future studies, including identification of enzymes involved in specialized metabolism and specialized transporters in the cells lining the secretory ducts would benefit from the transcriptomics of these.

**Conclusion**

The roots of the Mediterranean plant *Thapsia garganica* (dicot, Apiaceae), also known as Deadly carrot, have been shown to have secretory ducts that likely contain the highly toxic compound thapsigargin and similar sesquiterpenoid lactones. Through transient expression of *TgTPS2* (epikunzeaol synthase) and *TgCYP76AE2* in *Nicotiana benthamiana* it was shown that *TgTPS2* (epikunzeaol synthase) and *TgCYP76AE2* converts epikunzeaol into epidihydrocostunolide, compounds which are possible intermediates in thapsigargin biosynthesis. Transcripts from *TgTPS2* (epikunzeaol synthase) and *TgCYP76AE2* were found in the epithelial cells lining the secretory ducts and nowhere else in the root. This emphasizes the involvement of these specific cells in the biosynthesis of thapsigargin and other sesquiterpene lactones in *T. garganica*.

The findings presented here enable the full elucidation of the biosynthesis of thapsigargin biosynthesis in the cells lining the secretory ducts. The findings also show the power of *in situ* PCR combined with MALDI-TOF imaging.
Materials and Methods

Plant material

Thapsia garganica L. roots used for cDNA synthesis to obtain genes of interest were collected in June 2010, 25 km SSW of Bari, Italy (GPS 40.89625, 16.706139).

For NADI staining, in tube in situ PCR and MALDI-MSI 2 year old T. garganica plants were used. These had been grown in Ibiza, and were purchased from ThapsIbiza S. L.

Identification and cloning of genes

The ORF of P450s belonging to the CYP71 clade was found in the transcriptome of the Thapsia garganica root (SRX096991) (Pickel et al., 2012) based on BLAST searches using a set of CYP71 clan members as previously described (Dueholm et al., 2015). The P450s are available on NCBI; TgCYP71AH8 (KX826939), TgCYP71AS14 (KX845553), TgCYP71AT12 (KX826940), TgCYP71AU89 (KX845548), TgCYP71AU90 (KX845552), TgCYP71BK1 (KX826941), TgCYP71BK6 (KX845546), TgCYP71D183 (KX845554), TgCYP71D311 (KX845555), TgCYP71D319_ortholog (KX845550), TgCYP76AE1 (KX826942), TgCYP76AE2 (KX826943), TgCYP76AE8 (KX845545), TgCYP76AF7 (KX845549), TgCYP76B79 (KX845547) and TgCYP706C30_ortholog (KX845551). TgCYP71AJ5 (KP191555) and TgCYP71AJ14 (KP191558) were also tested previously (Dueholm et al., 2015). The discovery of TgTPS2 (JQ290345) has been described in (Pickel et al., 2012). Full-length gene sequences were obtained from a cDNA library of T. garganica root material.

Forward and reverse primers for all genes were designed with USER-overhangs, to enable cloning into a pEAQ USER compatible version of the pEAQ-HT vector (Table S1,) (Luo et al., 2016). pEAQ-HT harbors the viral suppressor p19 and was kindly provided by George Lomonosonoff (John Innes Research Centre, Norwich, UK) (Peyret and Lomonossoff, 2013). USER cloning was performed as depicted previously (Nour-Eldin et al., 2006). A truncated version of Arabidopsis thaliana HMGR (GenBank J04537), described previously (Cankar et al., 2015), was kindly provided Katarina Cankar (Wageningen University, The Netherlands).

TcCYP71AV2 (KC441527.1), CiCYP71AV8 (HQ166835.1), CcCYP71AV9 (KF752448.1), LsCYP71BL2 (HQ439599.1), CiCYP71BL3 (JF816041.1), TcCYP71BL4 (KC441528.1) and CcCYP71BL5 (KF752451.1) were furthermore blasted into the T. garganica root transcriptome. Setting the expectation value to 1E-100 no hits were available.


Phylogeny

In total 53 full-length sequences of functional characterized cytochromes P450 related to sesquiterpenoid biosynthesis and cytochromes P450 from *T. garganica* were used to build the phylogenetic tree. The full list is in the supplementary file as part of the alignment, and the NCBI numbers for the *Thapsia* genes are given above.

All obtained full length sequences were aligned using default options in MUSCLE and ClustalW (Edgar, 2004), as implemented in the software Geneious 10.0.5 (www.geneious.com) followed by manual modification. Phylogenetic analyses were conducted using maximum likelihood. Default options for PhyML based on the substation model LG, in Geneious 10.0.5 was chosen (Guindon et al., 2010). All maximum likelihood trees (ML) were obtained using 1,000 replicates of random taxon addition sequence. All characters were included in the analyses. Clade support was assessed using non-parametric bootstrap re-sampling. Bootstrap analysis (Felsenstein, 1985) was carried out using 1,000 replicates. We defined bootstrap percentages (BS) < 50 % to be unsupported, between 50 % and 74 % as weak support, between 75 % and 89 % BS as supported, and scores of greater than 90 % BS as strong support. Alignments supporting the tree are given in supplementary file 1 (for Figure 2). The *Thapsia* gene cinnamate-4-hydroxylase (*Tg*C4H) was used to root the tree.

Expression of *TgTPS2* and P450s in *Nicotiana benthamiana*

*N. benthamiana* plants were grown from seeds at 24°C/19°C (day/night) for five weeks before transformation. The transformation of *Agrobacterium tumefaciens* and infiltration of *N. benthamiana* with *A. tumefaciens* followed the protocol described by (Bach et al., 2014). In short, 10 ml LB containing kanamycin, rifampicin and carbenicillin was inoculated with several agrobacterium colonies containing the plasmid of interest. Cultures were grown ON at 28°C and 200 rpm. Cell pellets were washed twice with water before final resuspension in water followed by a dilution to OD$_{600}$ 0.5. Resuspended *A. tumefaciens* carrying plasmids containing *At*HMGR, *TgTPS2* or *TgCYPs* were mixed 1:1:1 and infiltrated into leaves of at least three *N. benthamiana* plants by use of a syringe. Plants were placed at 24°C/19°C (day/night) and harvested five days after infiltration. As controls, plants were infiltrated with *A. tumefaciens* carrying plasmids with no additional genes, *At*HMGR, and *At*HMGR plus *TgTPS2*.

The ~100 plants needed for purification of products 1, 2 and 3 were infiltrated by use of vacuum. Three *A. tumefaciens* cultures containing *At*HMGR, *TgTPS2* or *TgCYP76AE2* were grown ON at 28°C and 200 rpm in 500 ml LB (containing kanamycin, rifampicin and carbenicillin) from 20 ml
starter cultures. Cell pellets were washed twice with water before final resuspension in water followed by a dilution to OD$_{600}$ 0.5. Resuspended A. tumefaciens carrying plasmids containing AtHMGR, TgTPS2 or TgCYP76AE2 were mixed 1:1:1. Plants were submerged in a 1 L suspension of A. tumefaciens and infiltrated by use of vacuum at 50-100 mbar for 1 min (Andersen-Ranberg et al., 2016).

**GC-MS detection of sesquiterpenoids**

Two leaf discs (Ø=3cm) from N. benthamiana, expressing genes from T. garganica were extracted with 1.2 ml hexane for GC-MS analysis to provide one sample, a minimum of three biological replicates were examined. Hexane extracts were analyzed on a Shimadzu GCMS-QP2010 using an Agilent HP-5MS UI, 20 m, 0.18 mm diameter × 0.18 μm film thickness column kept at a pressure of 66.7 kPa giving a column flow of 1 mL/min(Drew et al., 2012). The injection port temperature was set to 160°C or 250°C to find the optimal temperature (Andersen et al., 2015). The oven temperature was set to 60°C for 3 min, and then increased to 160°C with a rate of 7°C/min. The temperature was further increased to 300°C at a rate of 50°C/min, held for 5 min, finally increased to 320°C at 50°C/min and maintained for 3 min. The carrier gas was H$_2$ and the ionization electron energy was 70 eV. The ion source temperature was 230°C with an interface temperature 280°C. The total run time was 11.67 min. All data were analyzed using the Shimadzu software Lab Solutions, GCMS Solutions version 2.50 SU3, with the 2008 libraries provided by NIST and Wiley.

**Analytical LC-MS for detection of novel sesquiterpenoids**

Two leaf discs with a diameter of 3 cm were ground in liquid nitrogen to provide one samples, a minimum of three biological replicates was examined. Samples were extracted with 1200 μl 80 % methanol and sonicated for 30 min. Before LC-MS analysis samples were filtered through a 0.45 μm filter. Analytical LC-MS was carried out using an Agilent 1100 Series LC (Agilent Technologies, Germany) coupled to a Bruker HCT-Ultra ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). A Gemini-NX column (Phenomenex; 3 μm, C18, 110A, 2 × 150 mm) maintained at 35°C was used for separation. The mobile phases were; A, water with 0.1 % (v/v) formic acid; B, acetonitrile with 0.1 % (v/v) formic acid. The gradient program was: 0 to 1 min, isocratic 12 % B; 1 to 33 min, linear gradient 12 to 80 % B; 33 to 35 min, linear gradient 80 to 99 % B; 35 to 38 min, isocratic 99 % B; 38-47 min, isocratic 12 % B. The flow rate was 0.2 ml min$^{-1}$. The mass spectrometer was run in positive mode and the mass range m/z 100-1000 was acquired.
Preparative GC-MS - purification of Cope rearranged sesquiterpenoids (2 and 3)

For isolation of compounds 2 and 3, a large-scale hexane extraction was made from *N. benthamiana* leaves expressing *At*HMGR, *Tg*TPS2 and *Tg*CYP76AE2. The leaves from approximately 130 five-week-old plants were used. The sample was subjected to an initial separation on a silica column and eluted with hexane:ethyl acetate at 13% ethyl acetate. The final purification was done on an Agilent 7890B GC installed with an Agilent 5977A inert MSD, GERSTEL Preparative Fraction Collector (PFC) AT 6890/7890 and a GERSTEL CIS 4C Bundle injection port. For separation by GC, a RESTEK Rtx-5 column (30 m × 0.53 mm ID × 1 µm df) with H2 as the carrier gas was used. At the end of this column was a split piece with a split of 1:100 to the MS and the PFC, respectively. A sufficient amount of sesquiterpene product for NMR analysis (0.5-1 mg) was obtained by 100 repeated injections of 5 µL of extract. The injection port was put in solvent vent mode with a carrier gas flow of 100 mL/min until 0.17 min. This was combined with an injection speed of 1.5 mL/min. The purge flow was set to 3 mL/min from 0.17 min to 2.17 min. The injection temperature was held at 40°C for 0.1 min, followed by ramping at 12°C/sec until 320°C, which was held for 2 min. The column flow was set to 7.5 mL, which was held constant throughout the GC program. The GC program was set to hold at 60°C for 1 min, ramp 20°C/min to 320°C, which was held for 3 min. Temperature of the transfer line from GC to PFC and the PFC itself was set to 250°C. The PFC was set to collect the peak of product 2 and 3 by their retention time identified by the MS. The MS was set in scan mode from *m/z* 35 to *m/z* 500, with a threshold of 150. Solvent cut-off was set to 4 min, and the temperature of the MS source and the MS quadrupole was set to 300°C and 150°C, respectively. Traps were kept at -20°C and rinsed with chloroform-d (Euriso-top, 99.8 atom % D).

Purification of epidihydrocostunolide (1)

For isolation of compound 1, a large-scale hexane extraction was made from *N. benthamiana* leaves expressing truncated *At*HMGR, *Tg*TPS2 and *Tg*CYP76AE2. The leaves from approximately 100 five-week-old plants were used. The crude hexane extract was subjected to an initial preparative separation on a Biotage Isolera autoflasher using a 10 g, 50 µm diol column and eluted stepwise with hexane:ethyl acetate (80:20) with 2% increments of ethyl acetate. Final isolation of 1 was achieved by semi-preparative HPLC on a 250 × 10 mm, 7 µm Nucleosil PEI column (Macherey-Nagel) eluted isocratic with hexane on a Waters 600 HPLC equipped with a Waters 996 PDA detector.
**1H and 13C NMR spectroscopic analysis**

NMR-spectra were acquired using a 600 MHz Bruker Avance III HD NMR spectrometer (1H operating frequency 600.13 MHz) equipped with a Bruker SampleJet sample changer and a cryogenically cooled gradient inverse triple-resonance 1.7-mm TCI probe-head (Bruker Biospin, Rheinstetten, Germany) optimized for 13C and 1H. Samples were analyzed at 300 K. Proton spectra, at 600.03 MHz, were obtained using 30°-pulses, a spectral width of 12 kHz, collecting 16 scans with a length of 65536 data points with a relaxation delay of 1.0 sec. Carbon spectra were acquired at 150.88 MHz with 30°-pulses, a spectral width of 36 kHz, collecting 256 scans with a length of 65536 data points and with a relaxation delay of 2.0 sec. The 13C nuclei were 1H-decoupled using the Waltz-16 composite pulse-decoupling scheme. FID's were exponentially multiplied with a line-broadening factor of 1.0 Hz before Fourier transformation. The 2D experiments were recorded using Bruker standard parameter settings. The isolated product 1 was dissolved in MeCN-d₃ (99.8 atom % D), while 2 and 3 were dissolved in CDCl₃ (99.8 atom % D) prior to NMR analysis.

**Histochemical analysis staining**

Taproots (~ 2 cm in diameter) were dug up just prior to sectioning. Cross-sections of roots were cut in 60 µm thick sections with a vibratome (Microm, HM 650 V). Histochemical analysis was performed with NADI reagent (1% naphthol + 1% dimethyl-p-phenylenediamine + 0.05M phosphate buffer, pH 7.2) reported to stain terpenoids blue (David and Carde, 1964). Sections were immersed in phosphate buffer 0.05 M (pH 7.2) buffer or NADI solution immediately after cutting and left for minimum 30 min before analysis. Samples were mounted on glass slides and images of the sections were obtained under a Leica DMR HC microscope through x20, x40 dry objectives and an x100 oil immersion objective.

**MALDI-MSI**

Prior to MALDI analysis, thapsigargin had been verified to be present in the roots by HPLC as previously described (Christensen et al., 1984; Pickel et al., 2012).

**Tissue preparation** Root tissue was snap-frozen in a cooling bath (2-propanol:dry-ice) and tissue blocks from the middle part of the root were mounted to the chuck using Tissue-Tek® O.C.T. compound. Tissues were sectioned on a cryosectioner (Leica 3050S cryostat). Sections were cut to 60 µm thick, ensuring no O.C.T. came into contact with the sectioned tissue. Sections were transferred onto pre-chilled Menzel-Gläser Superfrost Plus 25 mm × 75 mm × 1.0 mm glass slides and gently adhered to pre-mounted double-sided adhesive carbon tape (Agar Scientific, UK).
frozen slide with sections was transferred into a chilled 50 mL falcon tube then freeze dried for 24h using a freeze dryer (ScanVac CoolSafe), set to −95°C and an operating pressure of 1 mBar. Samples were stored in a vacuum desiccator prior to matrix deposition.

**Matrix Deposition** Matrix, 2,5-dihydroxybenzoic acid (DHB) was sublimed onto tissue sections using a custom built sublimation apparatus at temperatures of 130-140°C, at vacuum pressures less than 0.1 mBar for a period of 15 minutes generating a matrix coverage of 0.3 ± 0.1 mg/cm². An ice slurry was used to cool the sample cold finger. Alternatively, DHB was deposited onto tissues by spray deposition using a HTX Imaging TM Sprayer (HTX Technologies). DHB in acetone/water (50 mg/mL, 95:5) was sprayed at 30°C using 4 passes, a solvent flow rate of 150 µL/min, a nozzle velocity of 1200 mm/min, with alternate passes at 90° offset, 2 mm track spacing and 1 mm offset for repeat passes, nitrogen sheath gas pressure was set to 10 psi. After deposition, samples were then stored in a desiccator prior to analysis.

**Optical Image** Optical images of tissue sections were acquired using an EPSON Photosmart 4400 flatbed scanner using EPSON Scan Version 3.04A with a setting of 4800 d.p.i.

**Mass Spectrometer** For spatial mass spectrometric analysis, a SolariX XR 7 Tesla Hybrid ESI / MALDI FT-ICR-MS (Bruker Daltonik) was used. The instrument was operated in the Positive ion mode using optimized instrumental settings across the mass range 100-2000 m/z, with the instrument set to Broadband mode with a Time Domain for Acquisition of 2M providing an estimated resolving power of approximately 260000 at 400 m/z using a total of 1 ICR cell fills. The instrument was calibrated to less than 1 ppm tolerance against elemental Red Phosphorous clusters using a Quadratic calibration curve across the mass range 216-1951 m/z. The laser was set between 38-50 % power using the minimum spot size with Smart Walk enabled using a width of 40 µm, grid increment of 10 and offset of 1 using a random walk pattern resulting in ablation spots of approximately 40-50 µm in diameter. A total of 500-750 shots were fired per spectra at a frequency of 2 KHz within a 50 × 50 µm array.

**Data Analysis** Acquired mass spectrometry data was analysed using Compass FlexImaging 4.1 (Build 116, Bruker Daltonik GmbH). Images were either normalized to Root Median Square (RMS) or Total Ion Chromatogram (TIC) and brightness optimization employed to enhance visualization of the distribution of selected compounds. Individual spectra were analyzed and recalibrated using Bruker Compass DataAnalysis 4.3 (Build 110.102.1532) to internal lock masses of known DHB clusters; C_{14}H_{9}O_{6} = 273.039364 and C_{21}H_{13}O_{9} = 409.055408 m/z. Peak lists were generated using S/N threshold = 4 and 0.15 % base peak height threshold.
**In tube in situ RT-PCR on root tissue sections**

The *in situ* PCR was performed as previously described (Athman et al., 2014), with some modifications. Root tissue was left for 4 hours in ice cold FAA. 75μm thick root cross-sections were cut with a vibratome and transferred into tubes containing, RNAsin (Promega, N2515) as the RNase inhibitor. A DNase treatment with DNase 1 and 10x Turbo DNase (Qiagen 79254) was extended to overnight at 37°C. Sections were incubated for 15 min in 0.5M EDTA at 70°C, rinsed with ice cold nuclease free water then incubated on ice for 15 min in 2mg Pepsin in 0.1M HCl prior to the reverse transcription step to remove cross-linking. Reverse transcription was carried out using the Sensicript Reverse Transcription Kit (Qiagen 205213). The first cycling conditions were 5 min at 65°C followed by 1 min at 4°C. RNAsin and the reverse transcriptase enzyme were then added and the following cycling parameters used: 60 min at 42°C, 95°C for 5min, 1 min at 4°C for 1 min and the tubes were then placed directly on ice. The *in situ* PCR followed using TaKaRa Ex Taq (Clontech, RR001A), 5mM dNTPs and PCR DIG labelling mix (Sigma-Aldrich, 11585550910) in the PCR solution.

The negative controls followed the preparation of the test samples, but with the reverse transcriptase enzyme omitted. A positive control was carried out with the ribosomal 18S transcript. The primers used for the cDNA synthesis step [reverse (R) only] and PCR [forward (F) and reverse] are listed in Table S2; the products amplified by these primers were sequenced to confirm their specificity. The cycling parameters were: initial denaturation at 95°C for 2 min, 40 cycles of 95°C for 10 sec, X°C for 30 sec, 72°C for 1 min (X = 58°C for *Tg*CYP76AE2, 56°C for *Tg*TPS2 and 18S). The final extension was at 72°C for 5 min and then held at 4°C. Sections were incubated on ice for 30 min in a blocking solution (1% bovine serum albumin in 1X phosphate buffer) to prevent background staining. Anti-DIG-AP Fab fragments were then added followed by a 60 min incubation at room temperature. BM purple AP substrate (Sigma-Aldrich, 00000011442074001) was used for colorimetric staining. Sections were stained for 1.5 hours and mounted in Antifading solution (Citifluor, AF3-25). Slides were viewed with a Leica DMR HC microscope.

**Acknowledgements**

George Lomonosonoff (John Innes Research Centre, Norwich, UK) provided the pEAQ-HT plasmid. Katarina Cankar provided the truncated HMGR (Wageningen University, Wageningen, The Netherlands). MALDI MSI was conducted at Metabolomics Australia (School of BioSciences, The University of Melbourne, Australia), a NCRIS initiative under Bioplatforms Australia Pty Ltd.
(Figure 8)
Table 1: $^{13}$C NMR results for 1, 2 and 3. Samples were analyzed at 300 K. The isolated product 1 was dissolved in MeCN-d$_3$ (99.8 atom % D), while 2 and 3 were dissolved in CDCl$_3$ (99.8 atom % D).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>#C</td>
<td>δC (ppm)</td>
<td>δC (ppm)</td>
<td>δC (ppm)</td>
</tr>
<tr>
<td>1</td>
<td>128.7</td>
<td>148.7</td>
<td>147.2</td>
</tr>
</tbody>
</table>

Figure 8 A: Proposed biosynthesis of 1. TgCYP76AE2 is suggested to catalyze three hydroxylations to obtain epidihydrocostunolide. The intermediates are based on knowledge from the costunolide pathway. B: The costunolide pathway is as follows. Step one is GAS (germacrene A synthase) followed by three consecutive hydroxylations by GAO (germacrene A oxidase, CYP71AV2-9, Nguyen et al., 2010; Cankar et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014). Finally, a one hydroxylation of C6 by COS (costunolide synthase, CYP71BL2-5, Ikekawa et al., 2011; Liu et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014) to yield costunolide.
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>26.4</td>
<td>110.9</td>
<td>111.2</td>
</tr>
<tr>
<td>3</td>
<td>40.6</td>
<td>116.4</td>
<td>114.4</td>
</tr>
<tr>
<td>4</td>
<td>137.7</td>
<td>143.5</td>
<td>144.6</td>
</tr>
<tr>
<td>5</td>
<td>127.1</td>
<td>54.8</td>
<td>51.9</td>
</tr>
<tr>
<td>6</td>
<td>79.3</td>
<td>81.2</td>
<td>81.2</td>
</tr>
<tr>
<td>7</td>
<td>46.2</td>
<td>40.6</td>
<td>37.2</td>
</tr>
<tr>
<td>8</td>
<td>23.7</td>
<td>18.6</td>
<td>18.9</td>
</tr>
<tr>
<td>9</td>
<td>37.7</td>
<td>37.6</td>
<td>32.1</td>
</tr>
<tr>
<td>10</td>
<td>139.4</td>
<td>39.0</td>
<td>37.7</td>
</tr>
<tr>
<td>11</td>
<td>41.5</td>
<td>41.3</td>
<td>40.6</td>
</tr>
<tr>
<td>12</td>
<td>180.6</td>
<td>179.3</td>
<td>178.9</td>
</tr>
<tr>
<td>13</td>
<td>10.9</td>
<td>9.2</td>
<td>10.2</td>
</tr>
<tr>
<td>14</td>
<td>22.2</td>
<td>18.6</td>
<td>25.5</td>
</tr>
<tr>
<td>15</td>
<td>17.4</td>
<td>26.0</td>
<td>26.1</td>
</tr>
</tbody>
</table>
Table 2. $^1$H NMR results for 1, 2 and 3. Samples were analyzed at 300 K. The isolated product 1 was dissolved in MeCN-d$_3$ (99.8 atom % D), while 2 and 3 were dissolved in CDCl$_3$ (99.8 atom % D).

<table>
<thead>
<tr>
<th></th>
<th>1 ppm $\delta_H$ (J in Hz)</th>
<th>2 ppm $\delta_H$ (J in Hz)</th>
<th>3 ppm $\delta_H$ (J in Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.00 m, overlaid</td>
<td>5.80 dd, J=17.5, 10.8 Hz, 1 H</td>
<td>5.93 dd, J=17.6, 11.0 Hz, 1 H</td>
</tr>
<tr>
<td>2a</td>
<td>2.10 m, overlaid H</td>
<td>4.97 d, J=10.5 Hz, 1 H</td>
<td>4.98 d, J=17.2 Hz, 1 H</td>
</tr>
<tr>
<td>2b</td>
<td>2.27 m, overlaid H</td>
<td>4.98 d, J=18.0 Hz, 1 H</td>
<td>4.99 d, J=11.7 Hz, 1 H</td>
</tr>
<tr>
<td>3a</td>
<td>2.04 m, overlaid H</td>
<td>5.02 s, 1 H</td>
<td>4.85 s, 1 H</td>
</tr>
<tr>
<td>3b</td>
<td>2.23 m, overlaid H</td>
<td>5.07 s, 1 H</td>
<td>4.99 s, 1 H</td>
</tr>
<tr>
<td>5</td>
<td>4.98 d, J=10.1, overlaid H</td>
<td>2.13 d, J=2.6 Hz, 1 H</td>
<td>2.56 d, J=4.0 Hz, 1 H</td>
</tr>
<tr>
<td>6</td>
<td>5.14 dd, J=9.8, 4.6, 1H</td>
<td>4.45 t, J=3.3 Hz, 1 H</td>
<td>4.50 t, J=4.8 Hz, 1 H</td>
</tr>
<tr>
<td>7</td>
<td>2.49 m, overlaid H</td>
<td>2.29 - m, 1 H</td>
<td>2.57 - m, 1 H</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.36</td>
<td>2.65</td>
</tr>
<tr>
<td>8</td>
<td>1.59 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8a</td>
<td></td>
<td>1.37 qd, J=13.9, 2.90 Hz, 1 H</td>
<td>1.42 - m, 1 H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.52</td>
</tr>
<tr>
<td>8b</td>
<td></td>
<td>1.64 dddd, J=13.5, 6.3, 3.3 Hz, 1 H</td>
<td>1.61 - m overlaid, 1 H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.67</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td>1.56 - m overlaid, 2 H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.61</td>
</tr>
<tr>
<td>9a</td>
<td>1.70 m, overlaid H</td>
<td>1.55 m, overlaid H</td>
<td></td>
</tr>
<tr>
<td>9b</td>
<td>2.45 d, J=13.0, 1H</td>
<td>1.53 m, overlaid H</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>2.94 m, overlaid H</td>
<td>2.76 quin, J=7.0 Hz, 1 H</td>
<td>2.79 quin, J=7.3 Hz, 1 H</td>
</tr>
<tr>
<td>13</td>
<td>1.13 s, 3H</td>
<td>1.21 d, J=7.3 Hz, 3 H</td>
<td>1.22 d, J=7.3 Hz</td>
</tr>
<tr>
<td>14</td>
<td>1.53 s, 3H</td>
<td>1.20 s, 3 H</td>
<td>1.03 s, 3 H</td>
</tr>
<tr>
<td>15</td>
<td>1.67 s, 3H</td>
<td>1.87 s, 3 H</td>
<td>1.85 s, 3 H</td>
</tr>
</tbody>
</table>
Table 3. Calculated and observed masses (incl. error expressed in ppm) of Na and K adducts of thapsigargin, epikunzeaol, epidihydrocostunolide and a selection of known guaianolides from *T. garganica* in the MALDI experiment.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Epikunzeaol</td>
<td>C₁₅H₂₆O₂</td>
<td>245.18759</td>
<td>-</td>
<td>-</td>
<td>261.16152</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Epidihydrocostunolide</td>
<td>C₁₅H₂₂O₂</td>
<td>257.15120</td>
<td>-</td>
<td>-</td>
<td>273.12514</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nortrilobolide</td>
<td>C₂₆H₃₆O₁₀</td>
<td>531.22007</td>
<td>531.22110</td>
<td>-1.94</td>
<td>547.19401</td>
<td>547.19380</td>
<td>0.38</td>
</tr>
<tr>
<td>Trilobolide</td>
<td>C₂₇H₃₈O₁₀</td>
<td>545.23572</td>
<td>545.23370</td>
<td>3.70</td>
<td>561.20966</td>
<td>561.20730</td>
<td>-0.01</td>
</tr>
<tr>
<td>Thapsivillosin L</td>
<td>C₃₀H₄₂O₁₂</td>
<td>617.25685</td>
<td>-</td>
<td>-</td>
<td>633.23079</td>
<td>633.23240</td>
<td>-2.55</td>
</tr>
<tr>
<td>Thapsivillosin I</td>
<td>C₃₁H₄₂O₁₂</td>
<td>629.25685</td>
<td>629.25830</td>
<td>-2.31</td>
<td>645.23079</td>
<td>645.2327</td>
<td>-2.97</td>
</tr>
<tr>
<td>Thapsivillosin J</td>
<td>C₃₁H₄₄O₁₂</td>
<td>631.27250</td>
<td>-</td>
<td>-</td>
<td>647.24644</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thapsigarginic</td>
<td>C₃₂H₄₆O₁₂</td>
<td>645.28815</td>
<td>645.2864</td>
<td>2.71</td>
<td>661.26209</td>
<td>661.26040</td>
<td>-0.01</td>
</tr>
<tr>
<td>Thapsigargin</td>
<td>C₃₄H₅₀O₁₂</td>
<td>673.31945</td>
<td>673.3186</td>
<td>1.26</td>
<td>689.29339</td>
<td>689.2924</td>
<td>1.43</td>
</tr>
<tr>
<td>Thapsivillosin C</td>
<td>C₃₅H₅₂O₁₂</td>
<td>687.34588</td>
<td>-</td>
<td>-</td>
<td>703.30904</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

SUPPLEMENTAL DATA

Supplemental Table 1. Primers for cloning into the USER version of the vector pEAQ.

Supplemental Table 2. In tube *in situ* primers.

Supplemental Figure 1. $^{13}$C and $^1$H NMR spectra of product 1, 2 and 3.

Supplemental Figure 2. Cope-rearrangement of 1 upon GC-MS analysis.

Supplemental Figure 3. MALDI-MSI analysis of *T. garganica* taproot section and structures of metabolites.

Supplemental Figure 3-2. MALDI-MSI analysis of *T. garganica* taproot section.

Supplemental Figure 4. Alignment for the figure 2 tree.
Figure legends

**Figure 1.** The structure of epikunzeaol and epidihydrocostunolide (1). Both metabolites are suggested intermediates in the thapsigargin pathway in *T. garganica*.

**Figure 2:** The tree shows the phylogeny of P450s from *Thapsia garganica* (within CYP71 clade) and cytochromes related to sesquiterpenoid biosynthesis. The notes with *Thapsia* genes are marked in blue. The only two groups in the tree that can be assigned to a biochemical function are the germacrene A oxidase group CYP71AVx and the costunolide synthase group CYP71BLx. However, both groups are from the related species within the Asteraceae family thus cannot be used in general for functional identification. Only bootstrap values higher than 50 are shown in the tree.

**Figure 3.** GC-MS analysis of hexane extracts from *N. benthamiana* expressing in A *AtnHMGR* and *TgTPS2* with a GC-MS injection port temperature of 160°C (red) and of 250°C (black), in B *AtnHMGR* alone (black), *AtnHMGR* and *TgTPS2* (red), *AtnHMGR*, *TgTPS2* and *TgCYP76AE2* (blue) with a GC-MS injection port temperature of 250°C. * Denotes epikunzeaol thermal rearrangement products. C: Mass spectra the elemanolides 2 and 3.

**Figure 4:** A: Extracted Ion Chromatograms of epikunzeaol (1) [epikunzeaol disaccharide+Na]+ m/z 569.74 (red) and epidihydrocostunolide [2M+Na]+ m/z 491.32 (blue) of the LC-MS analysis of methanol extract of *AtnHMGR* alone, *AtnHMGR* plus *TgTPS2*, and *AtnHMGR* plus *TgTPS2* co-expressed with *TgCYP76AE2*. B: The MS/MS spectrum of the epikunzeaol peak at 18 minutes detected as the epikunzeaol-disaccharide. The MS spectrum of epidihydrocostunolide at 27 minutes shows a m/z of 491.3258 corresponding to a sodium adduct and a dimer of m/z 235.1741. C: The structure of epikunzeaol disaccharide including the fragmentation pattern.

**Figure 5:** Thermal rearrangement of 1 into the two compounds 2 and 3. The Cope rearrangement is well known for transforming germacradi-1(10),4-dienolides into elemanolides.

**Figure 6 A:** *In situ* PCR of *TgCYP76AE2* and *TgST2* in *Thapsia garganica* L. root cross-sections. All samples are stained with BM purple; blue/purple colour indicates presence of DIG.
labelled cDNA; brown indicates the absence of the amplified cDNA target. 1-2: negative controls (TgCYP76AE2) where reverse transcription (RT) was omitted. 3-4: detection of TgCYP76AE2 mainly in epithelial cells surrounding secretory ducts within vascular cambium. 5-6: negative controls (TgTPS2) where RT was omitted. 7-8: detection of TgTPS2 mainly in epithelial cells surrounding secretory ducts within vascular cambium as with TgCYP76AE2. 9-10: positive control; 18S ribosomal RNA to show staining in all cell types. B: NADI staining of Thapsia garganica L. root cross-sections 11-14: blue indicates presence of oxygenated or lipophilic compounds (e.g. terpenoids) concentrated in epithelial cells & secretory ducts. Scale bars represent 100 μm. p = periderm, ph = phloem, x = xylem, ep = epithelial cells, sd = secretory ducts, vc = vascular cambium, s = starch grains.

Figure 7. MALDI-MSI analysis of T. garganica taproot section. A: optical image of taproot section with sublimed DHB matrix, B: distribution of thapsigargin Na adduct, [M+Na]^+ m/z 673.3186 (calc. 673.31945, 1.26 ppm error), C: distribution of thapsigargin K adduct, [M+K]^+ m/z 689.2924 (calc. 689.29339, 1.43 ppm error). Images normalized to RMS and scaled to 0-60 % of maximum signal intensity using FlexImaging 4.1 to enhance visualization. Results demonstrate thapsigargin to be localized to concentric circles similar to the pattern of secretory ducts.

Figure 8. A: Proposed biosynthesis of 1. TgCYP76AE2 is suggested to catalyze three hydroxylations to obtain epidihydrocostunolide. The intermediates are based on knowledge from the costunolide pathway shown in B. B: The costunolide pathway. Step one is GAS (germacrene A synthase (Bouwmeester et al., 2002)) followed by three consecutive hydroxylation's by GAO (germacrene A oxidase, CYP71AV2-9, (Nguyen et al., 2010; Cankar et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014)). Finally, a hydroxylation of C6 by COS (costunolide synthase, CYP71BL2-5, (Ikezawa et al., 2011; Liu et al., 2011; Ramirez et al., 2013; Eljounaidi et al., 2014)) to yield costunolide.


Sanz JF, Castellano G, Marco JA (1990) Sesquiterpene lactones from Artemisia herba-alba. Phytochemistry 29: 541-545


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


Tu Y (2011) The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nature Medicine 17: 1217-1220


