

Coexpression analysis identifies two oxidoreductases involved in the biosynthesis of the monoterpene acid moiety of natural pyrethrin insecticides in *Tanacetum cinerariifolium*¹

Haiyang Xu^a, Gaurav D. Moghe^{b,2}, Krystle Wiegert-Rininger^c, Anthony L. Schillmiller^d, Cornelius S. Barry^c, Robert L. Last^{b,e}, Eran Pichersky^{a3}

^aDepartment of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA.

^bDepartment of Biochemistry, Michigan State University, East Lansing MI, USA

^cDepartment of Horticulture, Michigan State University, East Lansing, MI, USA.

^dMass Spectrometry and Metabolomics Core Facility, Michigan State University, East Lansing, MI, USA

^eDepartment of Plant Biology, Michigan State University, East Lansing MI, USA

ORCID IDs: 00000000269540000 (H.X.); 0000-0002-8761-064X (GM); 0000-0001-7069-7120 (A.L.S.); 0000-0003-4685-0273 (C.S.B.); 0000-0001-6974-9587 (R.L.L.) 0000-0002-4343-1535 (E.P.)

¹This work was supported by the National Science Foundation collaborative research grants 1565355 to EP and 1565232 to RLL.

²Present address: Section of Plant Biology, School of Integrative Plant Sciences, Cornell University, Ithaca, NY USA

³Address correspondence to: lelx@umich.edu

One Sentence Summary: A set of dehydrogenases are involved in the synthesis of trans-chrysanthemoid acid, the terpene moiety of the natural insecticide pyrethrins.

Running title: biosynthesis of *trans*-chrysanthemic acid

Key words: pyrethrins, *trans*-chrysanthemic acid, *Tanacetum cinerariifolium*, alcohol dehydrogenase, aldehyde dehydrogenase, oxidoreductase, biosynthetic pathway.

Contact information for all authors:

Haiyang Xu: hyxu@umich.edu

Gaurav Moghe: gdm67@cornell.edu

Krystle Wiegert-Rininger: wiegertk@msu.edu

Anthony Schilmiller: schilmil@msu.edu

Cornelius S. Barry: barrycs@msu.edu

Robert Last: lastr@msu.edu

Eran Pichersky: lelx@umich.edu

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: Eran Pichersky (lelx@umich.edu).

H.X, R.L.L and E.P. designed the experiments; H.X., G.M., A.L.S and K.W.R. conducted the experiments; H.X., C.S.B, R.L.L. and E.P wrote the article; all authors edited the article.

ABSTRACT

Flowers of *Tanacetum cinerariifolium* produce a set of compounds known collectively as pyrethrins, which are commercially important pesticides that are strongly toxic to flying insects but not to most vertebrates. A pyrethrin molecule is an ester consisting of either *trans*-chrysanthemic acid or its modified form, pyrethric acid, and one of three alcohols – jasmolone, pyrethrolone, and cinerolone – that appear to be derived from jasmonic acid. Chrysanthemyl diphosphate synthase (CDS), the first enzyme involved in the synthesis of *trans*-chrysanthemic acid, was previously characterized and its gene isolated. *TcCDS* produces free *trans*-chrysanthemol, in addition to *trans*-chrysanthemyl diphosphate, but the enzymes responsible for the conversion of *trans*-chrysanthemol to the corresponding aldehyde and then to the acid have not been reported. We used an RNAseq-based approach and co-expression correlation analysis to identify several candidate genes encoding putative *trans*-chrysanthemol and *trans*-chrysanthemal dehydrogenases. We functionally characterized the proteins encoded by these genes using a combination of *in vitro* biochemical assays and heterologous expression *in planta* to demonstrate that *TcADH2* encodes an enzyme that oxidizes *trans*-chrysanthemol to *trans*-chrysanthemal, while *TcALDH1* encodes an enzyme that oxidizes *trans*-chrysanthemal into *trans*-chrysanthemic acid. Transient co-expression of *TcADH2* and *TcALDH1* together with *TcCDS* in *Nicotiana benthamiana* leaves results in the production of *trans*-chrysanthemic acid as well as several other side-products. The majority (58%) of *trans*-chrysanthemic acid was glycosylated or otherwise modified. Overall, these data identify key steps in the biosynthesis of pyrethrins and demonstrate the feasibility of metabolic engineering to produce components of these defense compounds in a heterologous host.

INTRODUCTION

A small group of plants in the Asteraceae family, including *Tanacetum cinerariifolium* (formerly *Chrysanthemum cinerariifolium*), make insecticides known collectively as pyrethrins, with the highest levels of production observed in flowers (Casida, 1973; Casida and Quistad, 1995; Crombie, 1995). Due to their effective toxicity against a wide range of insect species, low toxicity to warm-blooded animals and propensity for degradation by sunlight and oxidation, pyrethrins have been used for pest control since medieval times (Casida and Quistad, 1995; Katsuda, 1999). Commercial production of natural pyrethrins involves drying and grinding the flowers and dissolving the powder in an organic solvent, which can be used directly for spraying. However, the concentration of pyrethrins in the powder is low at about 2% (Casida, 1973). Synthetic pyrethrins, called pyrethroids, which are chemically similar but not identical to natural pyrethrins are cheaper and used at much higher quantities worldwide. However, these synthetic products are generally less biodegradable and photolabile, persisting in the environment longer than natural pyrethrins and thus give rise to the emergence of resistance among insects. In addition, some pyrethroids are toxic to mammals and fish (Katsuda, 2012).

Natural pyrethrins comprise a group of six compounds. The type I pyrethrins – pyrethrin I, jasmolin I, and cinerin I – are esters of the monoterpene *trans*-chrysanthemic acid with one of the three respective fatty acid-derived alcohols pyrethrolone, jasmolone, and cinerolone. The type II pyrethrins – pyrethrin II, jasmolin II, and cinerin II – are esters of pyrethric acid with one of these respective three alcohols. Pyrethric acid is identical to *trans*-chrysanthemic acid, except that its C8 position has a methylated carboxyl group (Figure 1A).

The synthesis of both the acid and alcohol moieties of pyrethrins is not fully understood. With respect to the acid moiety, Rivera et al (2001) demonstrated that *T. cinerariifolium* flowers contain the enzyme *trans*-chrysanthemyl diphosphate synthase (CDS, EC 2.5.1.67), which condenses two dimethyl allyl diphosphate (DMAPP) molecules via what is known as an irregular C1'-2-3 linkage to form mostly *trans*-chrysanthemyl diphosphate (CDP) (Figure 1B) as well as a small amount of lavandulyl diphosphate. The enzymes responsible for the subsequent hypothetical steps are unknown. The conversion of CDP to *trans*-chrysanthemol could be catalyzed by a member of the terpene synthase (TPS) family similar to those TPS catalyzing the conversion of geranyl diphosphate (GPP) to the alcohol monoterpenes geraniol and linalool in

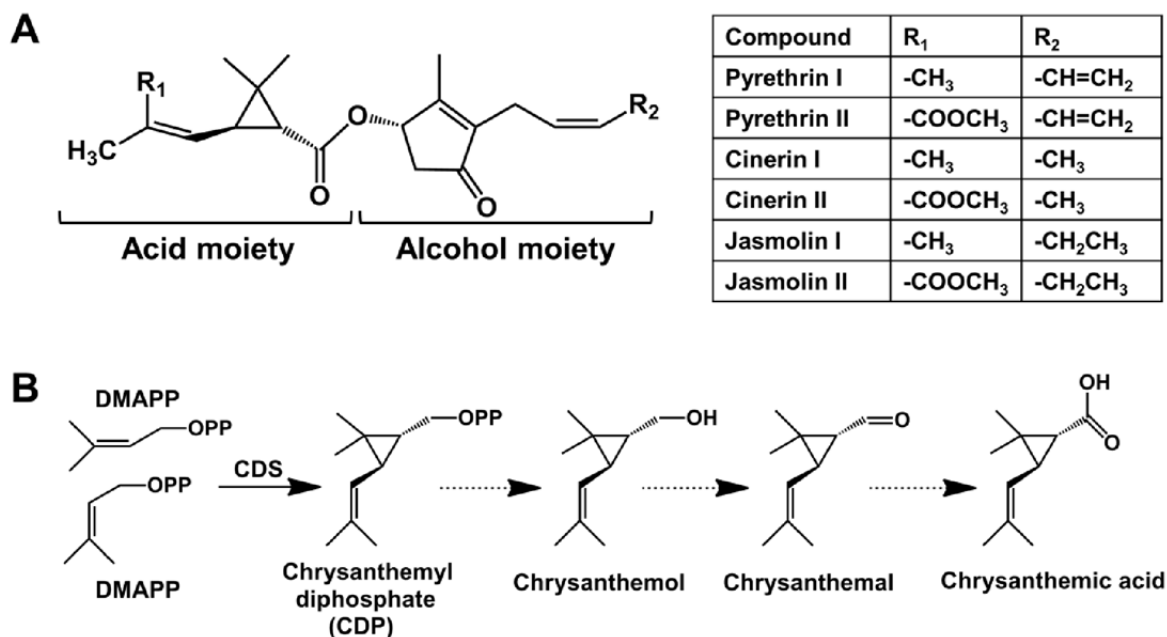


Figure 1. A. Structures of pyrethrins. B. Proposed pathway for the biosynthesis of *trans*-chrysanthemic acid.

many species (Chen et al., 2011), or CDP could be hydrolyzed by phosphatases to give *trans*-chrysanthemol, similar to the conversion of GPP to geraniol in rose flowers (Magnard et al., 2015). However no such enzymatic activities were reported thus far in *T. cinerariifolium*. A recent study by Yang et al (2014) reported that prolonged incubation of *E. coli*-produced recombinant CDS protein with low concentrations of DMAPP led to detection of *trans*-chrysanthemol in addition to CDP, suggesting that CDS can produce *trans*-chrysanthemol *in vivo*. However, the enzymes responsible for the subsequent oxidation reactions of *trans*-chrysanthemol to *trans*-chrysanthemal and then to *trans*-chrysanthemic acid (Figure 1B) had not yet been reported.

To facilitate characterization of pyrethrin biosynthesis, a transcriptome assembly of *Tanacetum cinerariifolium* was generated from RNAseq analysis of leaf and flower tissues harvested at different stages of development. Candidate genes for *trans*-chrysanthemic acid biosynthesis were identified based on co-expression analysis with two previously functionally identified genes in pyrethrin biosynthesis, *TcCDS* and *TcGLIP*, the latter being the gene encoding a GDSL-family lipase that combines the acid and alcohol moieties into an ester (Kikuta

et al., 2012). This analysis led to the identification of genes *TcADH2* and *TcALDH1*, encoding two flower-expressed oxidoreductases that respectively catalyze the two sequential oxidation steps from *trans*-chrysanthemol to *trans*-chrysanthemic acid. The discovery of these enzymes facilitated the reconstruction of *trans*-chrysanthemic acid biosynthesis from the precursor DMAPP, both *in vitro* and *in vivo* by transient expression of multiple genes in *Nicotiana benthamiana*.

RESULTS

Identification of pyrethrins and their terpenoid precursors in *Tanacetum cinerariifolium* flowers

The flower heads of *Tanacetum cinerariifolium* consist of a collection of ray florets on the inside and disc florets on the outside, with both types set on a receptacle (Ramirez et al., 2013). The pyrethrin and terpenoid precursor content in *Tanacetum cinerariifolium* leaves and flowers of different developmental stages (Figure 2A) were determined by analysis of methyl *tert*-butyl ether (MTBE)-extracted macerated tissues by gas chromatography – mass spectrometry (GC-MS). As previously observed (Kikuta et al., 2012; Ramirez et al., 2012), pyrethrin content in flowers increased as they matured (Figure 2B), and leaves contained negligible amounts of pyrethrins compared to the amounts observed in flowers. Also as previously described (Casida, 1973), pyrethrin I was the most abundant pyrethrin (Figure 2C). In addition to pyrethrins, we also observed *trans*-chrysanthemol, *trans*-chrysanthemal, and *trans*-chrysanthemic acid in floral extracts, and quantitation of *trans*-chrysanthemic acid indicated that its concentration also increased as the flower matured (Figure 2D,E).

Identification of candidate genes involved in *trans*-chrysanthemic acid biosynthesis

To identify the genes encoding the enzymes responsible for the conversion of *trans*-chrysanthemol to *trans*-chrysanthemal and *trans*-chrysanthemal to *trans*-chrysanthemic acid, transcriptome assemblies were constructed from RNASeq libraries constructed from eight different *T. cinerariifolium* tissue samples: leaves, flowers at stage 1, flowers at stage 2, flowers at stage 3, ray florets at stage 4, disk florets at stage 4, ray florets at stage 5 and disk florets at stage 5 (Figure 3A). Interrogating our database set (<http://sativa.mcdb.lsa.umich.edu/blast/>) for oxidoreductases with plant alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) sequences (see Materials and Methods) identified 12 transcripts encoding putative alcohol dehydrogenases (named as *TcADH1* ~ *TcADH12*) and three transcripts for putative aldehyde dehydrogenases (named as *TcALDH1* ~ *TcALDH3*). Four of the putative alcohol dehydrogenase genes, *TcADH1*, *TcADH2*, *TcADH4* and *TcADH6*, and one putative aldehyde dehydrogenase

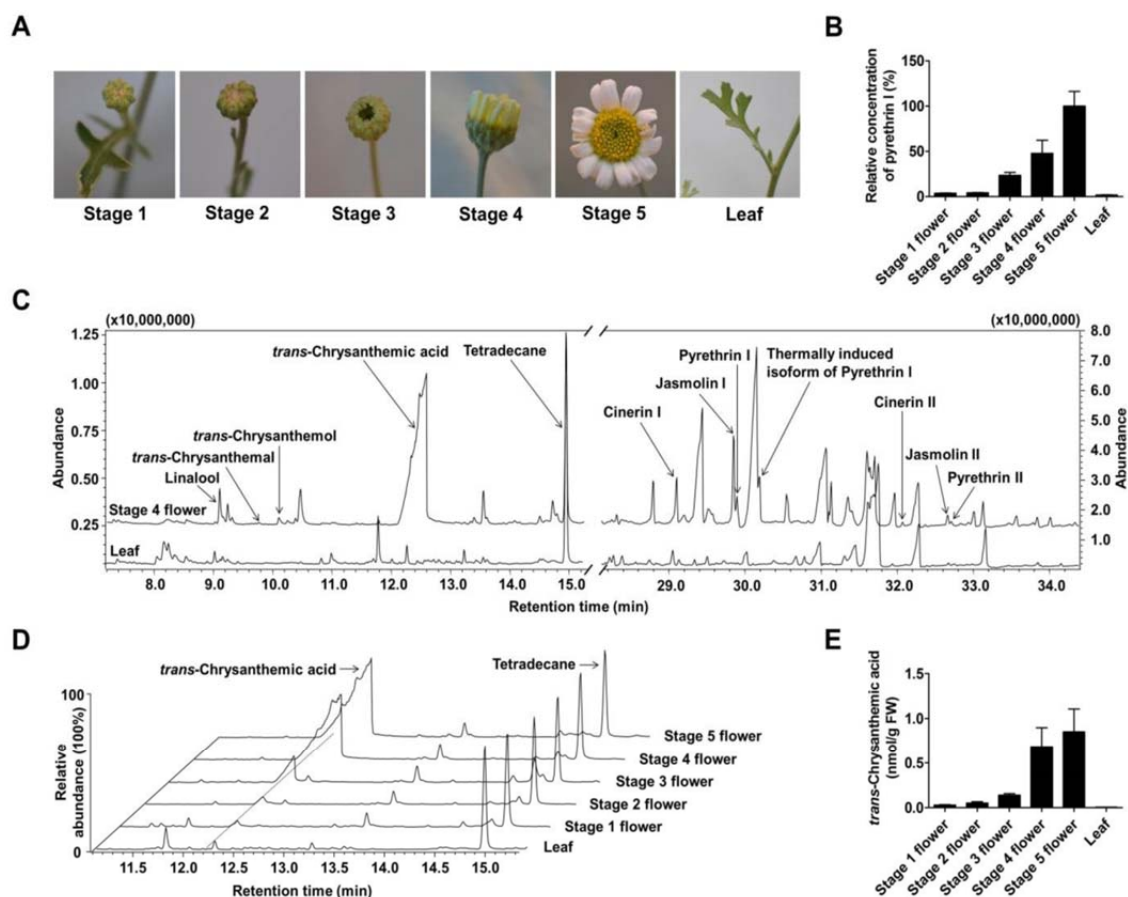


Figure 2. GC-MS analysis of pyrethrins and terpenoids from *Tanacetum cinerariifolium* leaves and flowers at different stages of development.

A. Flowers of different stages of development and a leaf of *T. cinerariifolium*.

B. Changes in relative concentrations of pyrethrin I during floral development. Pyrethrin I is the most abundant pyrethrin in the flower, and changes in concentrations of other pyrethrins follow the same pattern as those of pyrethrin I.

C. GC-MS chromatogram of total ion mode of MTBE extracts from leaves and flowers harvested at stage 4. In each flower/leaf comparison, samples are shown with the same relative y-axis scale, but the 7.2-15.2 min section is shown at a smaller scale to magnify the peaks. Peaks identified as terpenoids and internal standard (tetradecane) are labeled.

D. GC-MS chromatogram (total ion mode) of MTBE extracts from leaves and flowers of different stages of development, showing the *trans*-chrysanthemic acid levels in each sample.

E. Concentrations of *trans*-chrysanthemic acid in the leaf and in different stages of flowers. Quantification was achieved by normalization of the peaks in (D) to the tetradecane internal standard and comparison to a standard curve of authentic *trans*-chrysanthemic acid ($n = 3$; means $\pm$ SD).

160 gene, *TcALDH1*, showed the highest degree of co-expression with the known genes of pyrethrin
 161 biosynthesis *TcCDS* or *TcGLIP* (Figure 3B; Supplemental Table S1). However, the much lower

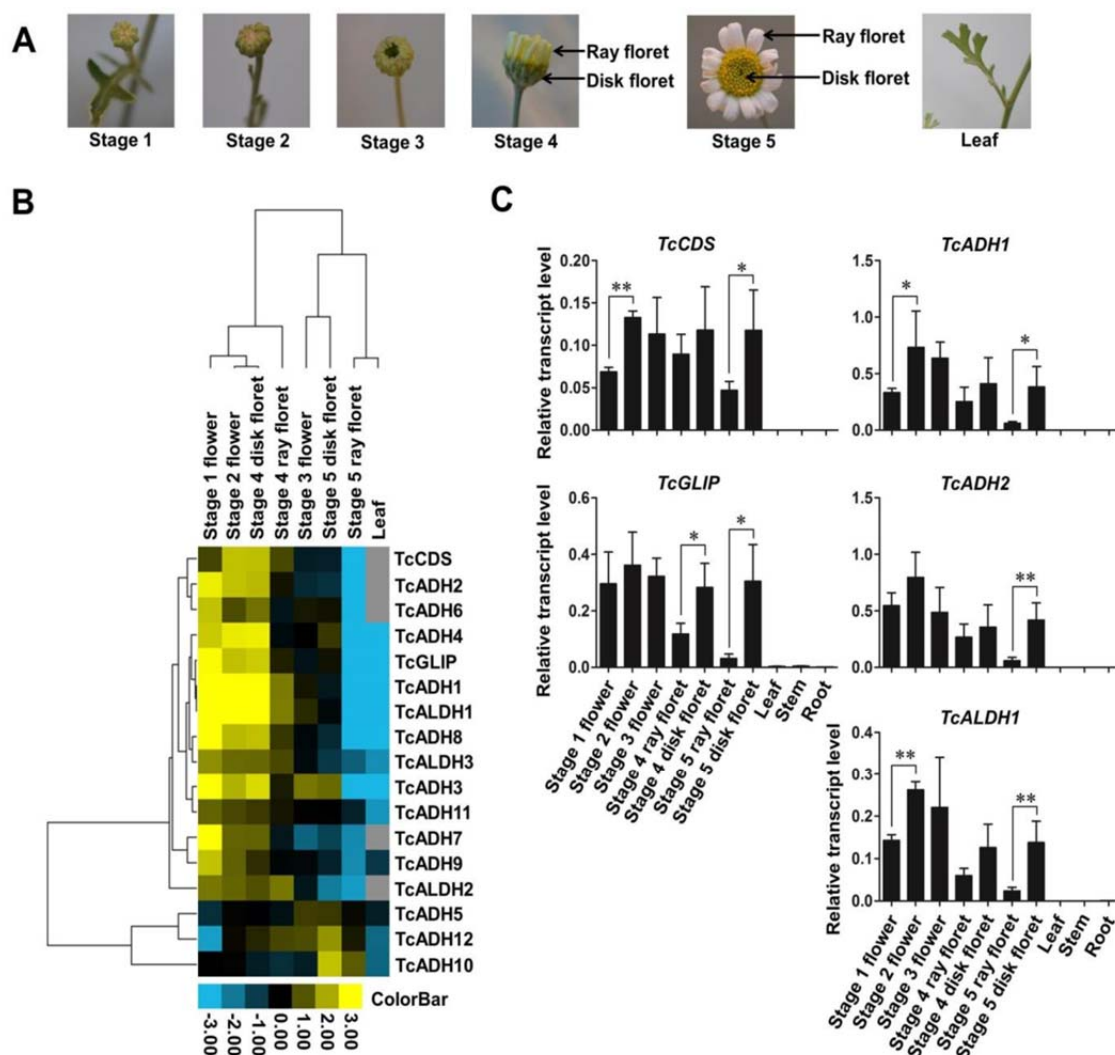


Figure 3. Identification of candidate *ADH* and *ALDH* genes for *trans*-chrysanthemic acid biosynthesis.

A. Images of *T. cinerariifolium* flowers of different stages and of leaves from which RNA samples were obtained for RNAseq analysis.

B. Average-linkage hierarchical clustering of relative transcript abundance of putative *ADH*s and *ALDH*s with *TcCDS* and *TcGLIP* based on number of reads of each transcript in each RNAseq library. Tree and heat map generated by the Cluster 3.0 software (see Materials and Methods).

C. Verification of levels of expression of *TcCDS*, *TcGLIP*, *TcADH1*, *TcADH2*, and *TcALDH1* by qRT-PCR. Transcript levels are expressed relative to that of *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) in each sample ($n = 4$; means $\pm$ SD). ** $p < 0.01$, * $p < 0.05$. The differences between leaf, stem and root datapoints and any flower datapoints are all significant at $p < 0.001$.

162 relative transcript abundance of *TcADH4* and *TcADH6* compared with that of *TcCDS* and

TcGLIP in flowers of *T. cinerariifolium* argued against their involvement in *trans*-chrysanthemic acid biosynthesis.

The expression patterns of *TcADH1*, *TcADH2*, and *TcALDH1* deduced from the RNAseq read frequencies were confirmed by qRT-PCR together with the transcript levels of *TcCDS* and *TcGLIP* (Figure 3C). The transcripts of all five genes were significantly more abundant in floral tissues compared with leaves, roots or stems. Generally, transcripts increased initially from floral stage 1 to floral stage 2 and then remained at similar levels or decreased somewhat, particularly in ray florets, a pattern consistent with the previously reported localization of pyrethrin accumulation (Kikuta et al., 2012; Ramirez et al., 2012).

Phylogenetic analysis that included *T. cinerariifolium* ADHs and other ADHs with assigned functions revealed that *TcADH1* and *TcADH2* are most closely related to terpene modifying ADHs, including ADH1 from *Artemisia annua*, annotated in NCBI as artemisinic alcohol dehydrogenase (accession AEI16475) and 8-hydroxygeraniol oxidoreductase from *Catharanthus roseus* (Miettinen et al., 2014) (Figure 4A). Similarly, *TcALDH1* is most closely related (Figure 4B) to ALDH1 from *Artemisia annua*, an enzyme that converts artemisinic aldehyde and dihydroartemisinic aldehyde to artemisinic acid and dihydroartemisinic acid, respectively (Teoh et al., 2009). No prediction of organelle targeting was obtained for the inferred *TcADH1*, *TcADH2*, and *TcALDH1* protein sequences using the TargetP 1.1 Server (<http://www.cbs.dtu.dk/services/TargetP/>) and WoLF PSORT (<https://wolfpsort.hgc.jp/>) programs (Nakai and Horton, 1999; Emanuelsson et al., 2007).

***TcADH2* is a *trans*-chrysanthemol dehydrogenase**

Based on results of the co-expression and phylogenetic analyses that identified *TcADH1* and *TcADH2* as candidates for the conversion of *trans*-chrysanthemol to *trans*-chrysanthemal, we tested their activities *in vitro*. *TcADH1* and *TcADH2* proteins with fused N-terminal HIS₆ tag were expressed in *Escherichia coli* and purified. To obtain the *trans*-chrysanthemol substrate – which is not commercially available – for the enzymatic assays, we used recombinant *TcCDS* to convert DMAPP to CDP, which was subsequently hydrolyzed by the addition of commercial alkaline phosphatase to give *trans*-chrysanthemol (Figure 5A). As previously reported (Rivera et al., 2001; Yang et al., 2014), the *CDS*-catalyzed reaction also produces small amounts of

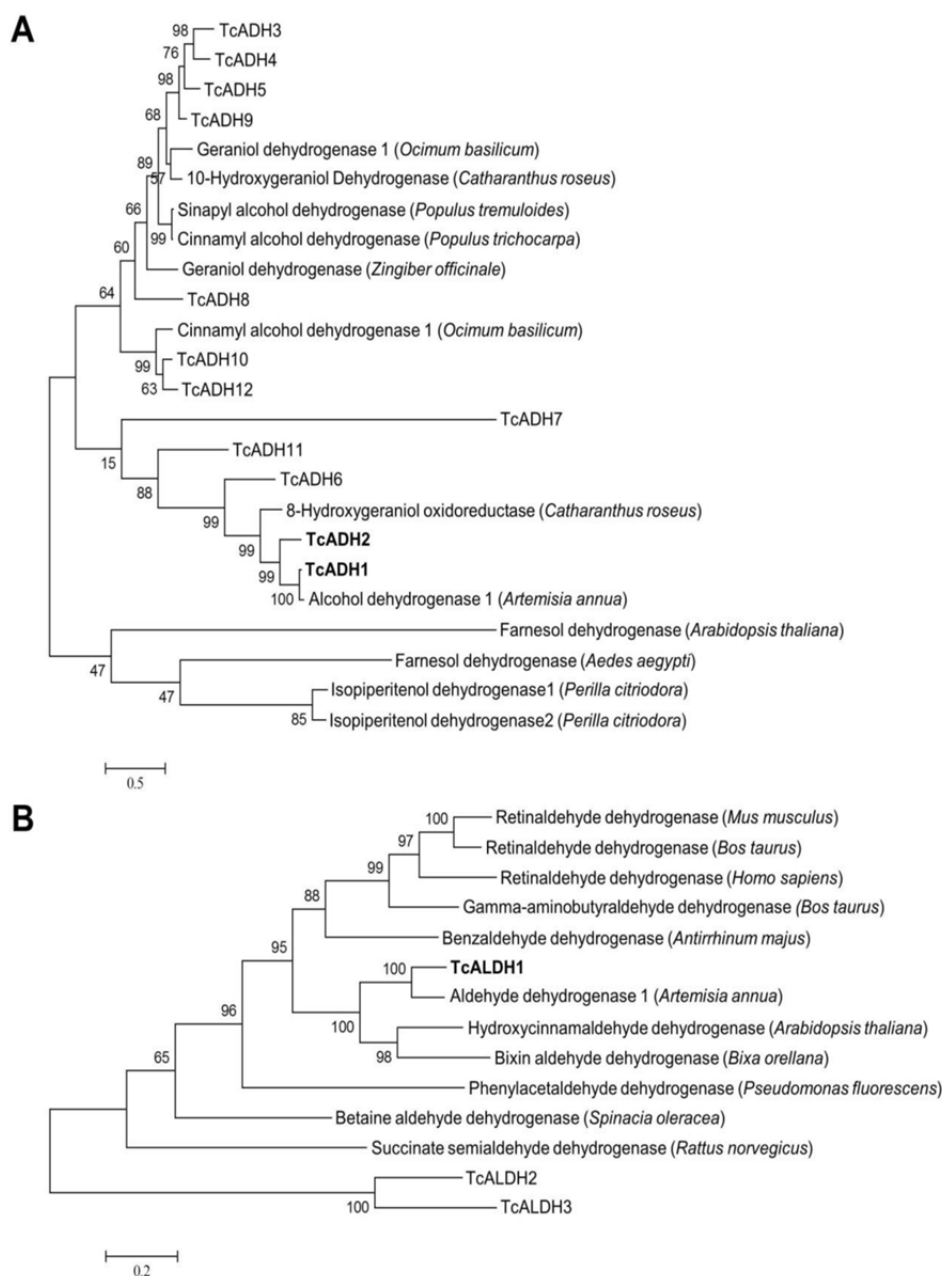


Figure 4. Phylogenetic analysis of candidate *T. cinerariifolium* dehydrogenases for *trans*-chrysanthemic acid biosynthesis based on protein sequences.

A. Phylogenetic tree for TcADH1 and TcADH2.

B. Phylogenetic tree for TcALDH1.

The protein sequences from other species are of functionally characterized enzymes whose sequences were identified by BLAST search to be most closely related to the *T. cinerariifolium* sequences. Phylogenetic analyses were conducted in MEGA7 (Kumar et al., 2016) with the following parameters; multiple sequence alignment with ClustalW, phylogenetic construction with the Maximum Likelihood method and bootstrap tests of 1000 replicates.

194 lavandulyl diphosphate, causing the *trans*-chrysanthemol preparation to contain a small amount

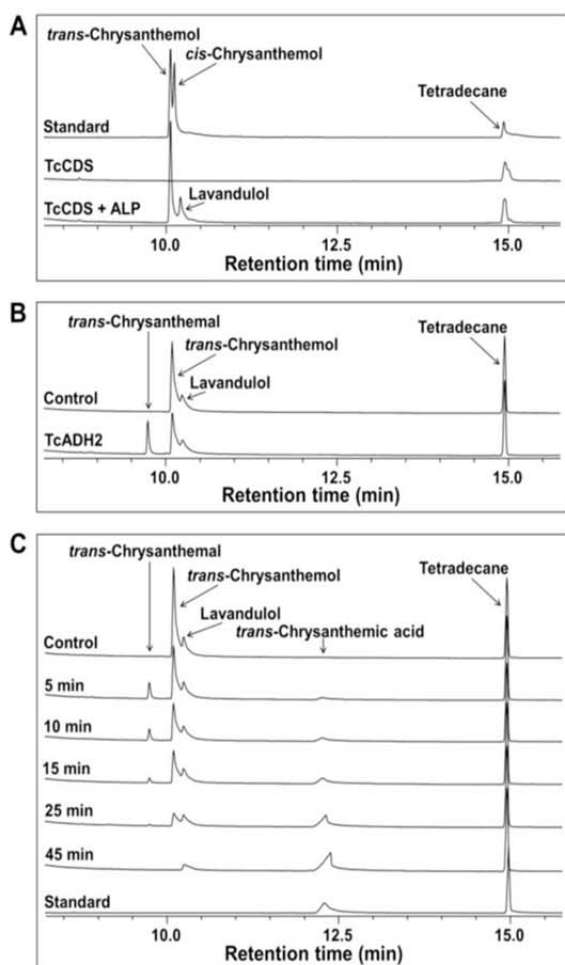


Figure 5. Gas chromatography analyses of products obtained in *in vitro* biochemical assays of TcADH2 and TcALDH1. For all assays analyzed, reactions products were extracted with 100 μ L MTBE and run on Rxi-5Sil column. Tetradecane was used as internal standard.

A. Synthesis of *trans*-chrysanthemol substrate. Top chromatograph trace: 1 mM of a commercially available standard of *trans*- and *cis*-chrysanthemol mixture. Middle trace: Reaction products obtained by incubating 30 μ g recombinant TcCDS with 2.5 mM DMAPP in 50 μ L reaction for 24 hours at 30°C. Bottom chromatograph trace: Reaction products obtained by incubating the products of the TcCDS-catalyzed condensation of DMAPP with 5 units of alkaline phosphatase (ALP) for 1 hour at 37°C.

B. *In vitro* production of *trans*-chrysanthemol. Reaction products obtained by incubating 0.64 mM *trans*-chrysanthemol and 1.5 mM NAD^+ with 5 μ L eluted protein from empty vector (top chromatograph trace) or 1.25 μ g purified TcADH2 (bottom chromatograph trace) in 60 μ L reaction volume for 5 min.

C. Production of *trans*-chrysanthemic acid from *trans*-chrysanthemol in a coupled assay containing 0.64 mM *trans*-chrysanthemol and 1.5 mM NAD^+ with 1.25 μ g purified TcADH2 and 6.00 μ g purified TcALDH1 in 60 μ L reaction volume for 5, 10, 15, 25, and 45 min. A control reaction was performed using 5 μ L eluted protein from empty vector. Bottom chromatograph trace: 0.3 mM of commercial *trans*-chrysanthemic acid.

195 of lavandulol (Figure 5A). Under our assay conditions, we did not detect *trans*-chrysanthemol
12

production by CDS for up to 24 h (Figure 5A).

Recombinant TcADH1 and TcADH2 were initially tested with a variety of alcohol substrates at 0.3 mM concentration and 1mM NAD⁺ or NADP⁺ (Table 1). With both TcADH1 and TcADH2, NAD⁺ was a more effective co-factor; with NADP⁺, product yield with each substrate was <5% compared with NAD⁺. TcADH1, however, had no activity with *trans*-chrysanthemol, instead showing its highest level of activity with nerol and reduced activity with geraniol (48%, Table 1), both monoterpene compounds. TcADH2 had the highest activity with *trans*-chrysanthemol, but it also had some activity (≤70% compared with its activity with *trans*-chrysanthemol) with several other alcohols (Table 1). GC-MS analysis of reaction products showed that TcADH2 converted *trans*-chrysanthemol to *trans*-chrysanthemal, but did not oxidize lavandulol (Figure 5B). Kinetic analysis revealed that TcADH2 had a K_m value of $236.0 \pm 5.8 \mu\text{M}$ and a turnover rate of $0.75 \pm 0.0032 \text{ s}^{-1}$ for *trans*-chrysanthemol while the K_m value for NAD⁺ was $192.6 \pm 8.7 \mu\text{M}$.

TcALDH1 converts *trans*-chrysanthemal to *trans*-chrysanthemic acid

Based on results of the co-expression and phylogenetic analyses that identified TcALDH1 as a candidate for the conversion of *trans*-chrysanthemal to *trans*-chrysanthemic acid, we tested its activity *in vitro* with *trans*-chrysanthemal and several other substrates. *trans*-chrysanthemal was produced in a coupled enzymatic method employing TcCDS, alkaline phosphatase and TcADH2 (Figure 5A and see Materials and Methods). N-terminal tagged TcALDH1 was purified from *E. coli* showed highest activity with *trans*-chrysanthemal substrate but also displayed substantial activity with several additional aliphatic and aromatic substrates (Table 1). The enzyme had a K_m value of $4.6 \pm 1.8 \mu\text{M}$ for *trans*-chrysanthemal when NAD⁺ was provided, and a K_m value of $4.4 \pm 2.2 \mu\text{M}$ for *trans*-chrysanthemal in the presence of NADP⁺ (Table 2). Both NAD⁺ and NADP⁺ served as cofactors for the enzyme, with a K_m of $20.4 \mu\text{M}$ for NAD⁺ and $68.6 \mu\text{M}$ for NADP⁺ (Table 2). Incubation of *trans*-chrysanthemol with both TcADH2 and TcALDH1 led to production of *trans*-chrysanthemic acid (Figure 5C).

Transient co-expression of TcADH2 and TcALDH1 with CDS in *Nicotiana benthamiana* results in *trans*-chrysanthemic acid production

To test the activities of TcADH2 and TcALDH1 *in planta*, *N. benthamiana* leaves were infiltrated with *Agrobacterium tumefaciens* strains harboring plasmids containing *TcCDS*, *TcADH2* and *TcALDH1*. For controls, we infiltrated *N. benthamiana* leaves with *TcCDS* alone, *TcCDS* and *TcADH2*, or *EGFP* (Enhanced Green Fluorescent Protein, see Materials and Methods) as control. In these experiments, the complete open reading frame of *TcCDS* was used, including the transit peptide that was shown to direct the protein to the plastids (Yang et al., 2014). Transformed leaves were harvested 10 days after infiltration and the products were extracted and analyzed by GC-MS. As expected, transient expression of *TcCDS* alone resulted in the production of *trans*-chrysanthemol at 18.3 nmol/g FW and trace amounts of *trans*-chrysanthemic acid (Figure 6). Notably, co-expression of *TcCDS* with *TcADH2* in *N. benthamiana* leaves resulted in production of *trans*-chrysanthemic acid at 328.0 nmol/g FW (Figure 6), a 48-fold increase over expression of *TcCDS* alone. Finally, when all three genes – *TcCDS*, *TcADH2* and *TcALDH1* - were co-expressed in *N. benthamiana* leaves, the concentration of *trans*-chrysanthemic acid was 818.4 nmol/g FW, a 122-fold increase over expression of *TcCDS* alone.

In addition to *trans*-chrysanthemic acid and its precursors, additional volatile derivatives of these metabolites were identified. A peak with a mass fragmentation spectrum (MS) similar to *trans*-chrysanthemol (Supplemental Figure S1C) was found in *N. benthamiana* leaves expressing *TcCDS*; it eluted with RT of 30.16 min, compared to a RT of 10.08 for *trans*-chrysanthemol, and was labeled as “Unknown 1” in Figure 6. Despite the lack of an authentic standard for this compound, which precluded determination of its actual concentration, we could measure changes in its relative concentrations. When *TcCDS*-expressing *N. benthamiana* leaf extract was treated with glycosidase, Unknown 1 was no longer observed (Supplemental Figure S2A). When the same extract was treated with NaOH, there were no changes in Unknown 1 and *trans*-chrysanthemol (Supplemental Figure S2B).

In leaves expressing both *TcCDS* and *TcADH2*, “Unknown 1” concentration was reduced by 30-fold. In addition, a new compound judged to be related to *trans*-chrysanthemic acid by its MS and designated as “Unknown 2” was detected with an RT of 15.49 (Figure 6; Supplemental Figure S1D). In leaves expressing *TcCDS*, *TcADH2* and *TcALDH1*, the levels of “Unknown 2” increased 7-fold over those observed in leaves expressing *TcCDS* and *TcADH2* without

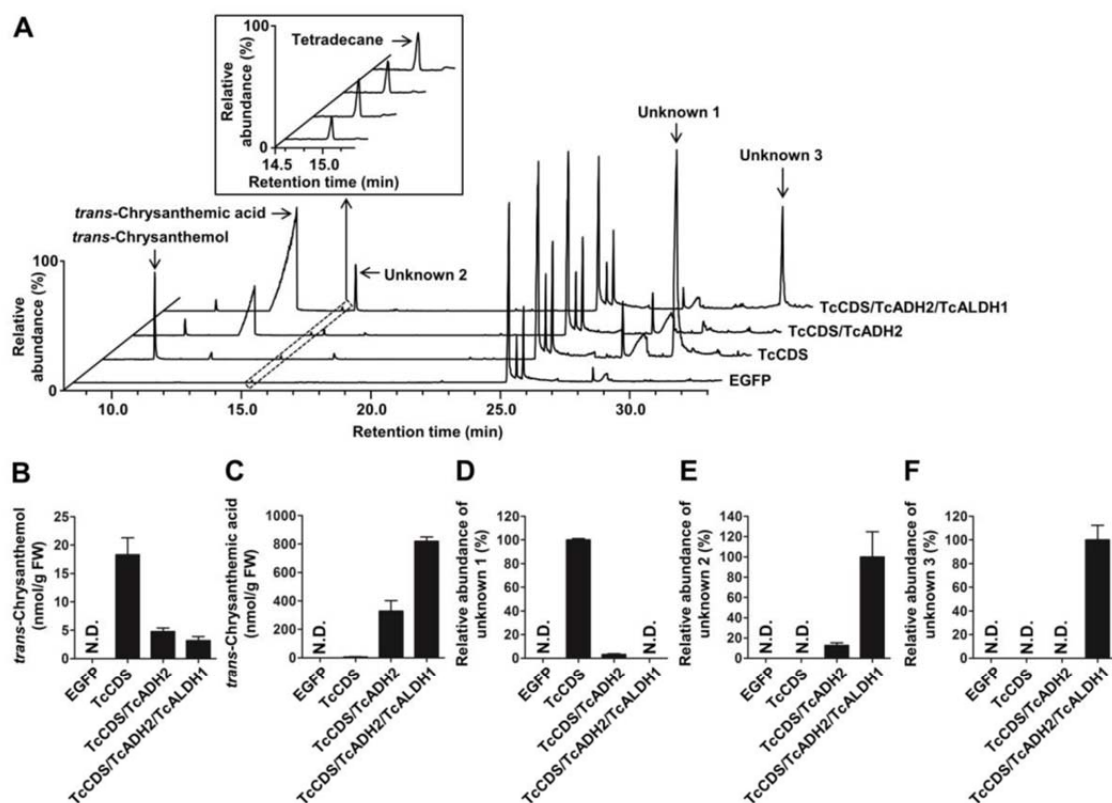


Figure 6. Production of *trans*-chrysanthemol, *trans*-chrysanthemol, and related compounds in *N. benthamiana* leaves transiently expressing TcCDS, TcADH2, and TcALDH1 proteins. A. GC-MS chromatograms of MTBE extracts of *N. benthamiana* leaves expressing EGFP (control), TcCDS, TcCDS and TcADH2, and TcCDS with TcADH2 and TcALDH1. For terpenes, $m/z = 123$ was monitored, and for the internal control tetradecane, $m/z = 198$ was monitored. Peaks related to *trans*-chrysanthemol, *trans*-chrysanthemol and internal standard are labeled. B. Concentrations of free *trans*-chrysanthemol and (C) free *trans*-chrysanthemol in *N. benthamiana* leaves expressing the indicated constructs were determined by comparison with an authentic standard. D-F. The relative levels of Unknown 1 (D), Unknown 2 (E), and Unknown 3 (F) in *N. benthamiana* leaves expressing the indicated constructs. For each compound, the plant material expressing a specific construct that showed the highest levels (average of three biological replicates) was set at 100%. The data in B–F represent mean $\pm$ SD from triplicate biological replicates. N.D. - not detected.

TcALDH1. In addition, the leaves expressing all three genes had a new *trans*-chrysanthemol acid-related compound, designated Unknown 3, with an RT of 31.98 (Figure 6, Supplemental Figure S1E). Unknown 2 and Unknown 3 were eliminated when the extract was treated with either glycosidase or NaOH (Supplemental Figure S3A,B), indicating that they are likely to be esters of *trans*-chrysanthemol. It is notable that the concentration of *trans*-chrysanthemol

was always below detection levels in all *N. benthamiana* infiltrated leaves, whether expressing *TcCDS* by itself, *TcCDS* and *TcADH2*, or *TcCDS*, *TcADH2*, and *TcALDH1*.

To search for non-volatile *trans*-chrysanthemol and *trans*-chrysanthemic acid conjugates, aliquots of the leaf MTBE extracts from these experiments were dried, dissolved in 70% acetonitrile, and analyzed by LC-qToF-MS. The sample from leaves expressing *TcCDS* alone contained a peak with *m/z* 803.3742 (Figure 7B). The mass spectrum-mass spectrum (MS/MS) of this metabolite (Figure 7E) was identical to that of a compound produced in *N. tabacum* expressing *TcCDS* (Yang et al., 2014) that these workers putatively identified as *trans*-chrysanthemylmalonylglucoside. In samples expressing *TcCDS* + *TcADH2*, a peak with *m/z* 831.3328 was detected with an MS/MS spectrum that closely matched the spectrum of another monoterpene acid glucoside, geranyl-6-*O*-malonyl- β -D-glucopyranoside (Yang et al., 2011). Based on this similarity, the *m/z* 831.3328 compound is putatively identified as the dimer ion (2M-H) of a *trans*-chrysanthemic acid malonylglucoside conjugate (Figure 7C,F). When *TcCDS*, *TcADH2*, and *TcALDH1* were co-expressed in *N. benthamiana*, the level of a product with identical elution time and accurate mass increased ~150-fold based on peak area compared to *TcCDS* and *TcADH2* expression without *TcALDH1* (Figure 7C,D). To determine the portion of *trans*-chrysanthemic acid present as malonylated glucoside in plants simultaneously expressing the three enzymes, we used GC-MS to compare the amount of free *trans*-chrysanthemic acid in extract hydrolyzed with 0.4N NaOH at 80 °C (which left no detectable *trans*-chrysanthemic acid malonylglucoside as determined by LC-qToF-MS, see Supplemental Figure S4) with that in non-hydrolyzed extract, and found that 58.0% of the total *trans*-chrysanthemic acid in this extract was esterified, including with malonylated glucoside (Figure 7G). Since the concentration of free *trans*-chrysanthemic acid in these leaves was determined to be 818.4 nmol/g FW (Figure 6C), the total amount of *trans*-chrysanthemic acid produced in these leaves can be calculated to be 1946.6 nmol/g FW.

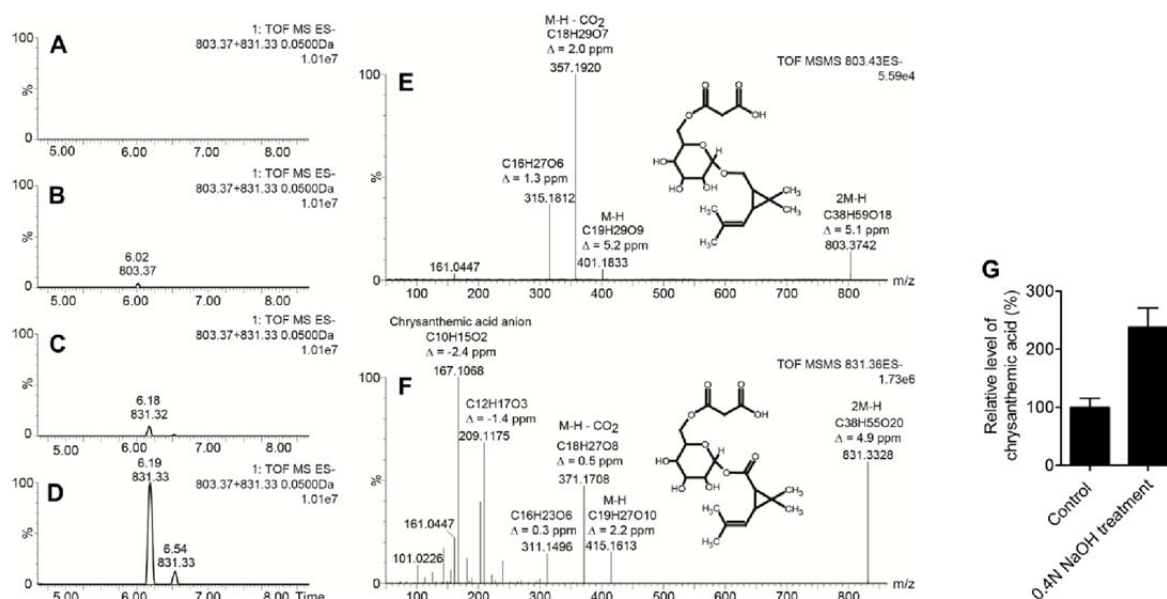


Figure 7. LC-MS analysis of *N. benthamiana* leaves simultaneously expressing the three enzymes TcCDS, TcADH2, and TcALDH1. Extracted ion chromatograms of m/z 803.37 and 831.33 are shown for: A. EGFP single expression control, B. TcCDS, C. TcCDS + TcADH2, D. TcCDS + TcADH2 + TcALDH1. Chromatograms are all scaled the same, as indicated by the ion current (1.01×10^7) in the upper right corner of each chromatogram. The sample for panel D was diluted 10-fold compared to the other three samples due to high concentration of m/z 831.33 in the sample. MS/MS spectra are shown for ions m/z 803.37 (E) and m/z 831.33 (F) along with the proposed compound structures based on exact mass, fragmentation pattern, and similarity to previously published data. G. Relative levels of *trans*-chrysanthemic acid in *N. benthamiana* leaves expressing all TcCDS, TcADH2 and TcALDH1 with or without sodium hydroxide treatment.

DISCUSSION

Co-expression analysis and biochemical assays indicate that TcADH2 and TcALDH1 catalyze reactions in the *trans*-chrysanthemic acid biosynthetic pathway

T. cinerariifolium is an important commercial source of the biodegradable natural pyrethrin insecticides, which are very efficient against flying insects and safe for humans and other vertebrates. It has been shown that pyrethrin biosynthesis begins at early stages in the developing achenes (dry fruits) in the inflorescence and reaches peak accumulation in the mature achene (Ramirez et al., 2012). Our observations on the pattern of accumulation of pyrethrins in the inflorescence (Figure 2B) were consistent with these previous observations. Based on this

information, we proceeded to perform transcriptomic profiling on RNA samples collected from five different stages of developing inflorescences as well as from leaf material. The last two stages of floral development, stages 4 and 5, afforded enough material to do separate analysis on ray florets (flowers on the outside perimeter, which have large petals) and disk florets (flowers on the inside, with small petals).

The transcriptomic profiling data (assembled transcripts and the number of reads from each transcript) were used to perform co-expression cluster analysis (Eisen et al., 1998) to identify candidate ADH and ALDH genes whose expression patterns most resembled that of *TcCDS*, the gene encoding the key enzyme in the pathway for *trans*-chrysanthemic acid (Rivera et al., 2001; Yang et al., 2014), as well as that of *TcGLIP*, encoding the enzyme that forms the final pyrethrin product (Kikuta et al., 2012). Both genes were shown to have peak transcript levels at the earliest developmental stage (stage 2) that could be examined (Ramirez et al., 2012). This analysis identified *TcADH1*, *TcADH2*, and *TcALDH1* as the strongest candidate enzymes for the synthesis of the terpene moiety of pyrethrins. The results of the qRT-PCR analysis of the levels of transcripts of these three genes, as well as those of *TcCDS* and *TcGLIP* (Figure 3C), were generally consistent with the read frequencies obtained in the transcriptomic profiling experiments (Supplemental Table 1). However, the transcript levels we measured for all five gene do not show a steep decline in later stages of floral development as was reported previously for *TcCDs* and *TcGLIP* (Ramirez et al., 2012), although the discrepancy may be due to differences in the delineation of developmental stages and consequent differences in the actual ages of the materials examined.

Based on the identification of *TcADH1*, *TcADH2*, and *TcALDH1* as the most likely candidates to be involved in conversion of *trans*-chrysanthemol to *trans*-chrysanthemic acid, we proceeded to produce recombinant proteins from these three genes and test the catalytic activities of these proteins in *in vitro* assays. *TcADH1* exhibited no activity with *trans*-chrysanthemol, but *TcADH2* was able to catalyze the conversion of *trans*-chrysanthemol to *trans*-chrysanthemal with a K_m value of 236 μM for *trans*-chrysanthemol (Table 2). Similar *in vitro* assays demonstrated that *TcALDH1* catalyzes the conversion of *trans*-chrysanthemal to *trans*-chrysanthemic acid, with a K_m value of 4.4 μM for *trans*-chrysanthemal. Transient co-expression of *TcADH2* and *TcALDH1* together with *TcCDS* in *N. benthamiana* leaves resulted in appreciable amounts of *trans*-chrysanthemic acid produced - 1946.6 nmol/g FW - further

demonstrating the ability of TcADH2 and TcALDH1 to catalyze the sequential oxidation of *trans*-chrysanthemol to *trans*-chrysanthemic acid.

Chrysanthemic acid can be made efficiently by *in planta* heterologous expression of only *TcCDS*, *TcADH2*, and *TcALDH1*

Yang et al. (2014) reported that transgenic tobacco (*N. tabacum*) plants expressing only *TcCDS* produce *trans*-chrysanthemol and *trans*-chrysanthemylmalonylglucoside. Our analysis of *N. benthamiana* leaves transiently expressing *TcCDS* identified the production of these two metabolites as well as a compound by GC-MS that appears to be a modification of *trans*-chrysanthemol (Unknown 1). When extracts of *N. benthamiana* leaves transiently expressing *TcCDS* was treated with glycosidase, both *trans*-chrysanthemylmalonylglucoside and Unknown 1 disappeared, and the amount of *trans*-chrysanthemol increased by 35-fold, indicating that most of the *trans*-chrysanthemol in these leaves is in a glycone form. These results indicate (as also noted by Yang et al., 2014) that *trans*-chrysanthemol is produced directly by TcCDS, and/or that phosphatases present in the host plant can hydrolyze the two phosphate groups from *trans*-chrysanthemyl diphosphate to give *trans*-chrysanthemol. But in addition to *trans*-chrysanthemol and its derivatives, we also detected trace amounts of *trans*-chrysanthemic acid in the *N. benthamiana* leaves transiently expressing *TcCDS* (Figure 6), showing that endogenous enzymes found in plant hosts are capable of further modifying the heterologously produced *trans*-chrysanthemol by additional oxidation reactions to give the corresponding acid. This observation is similar to the observed conversion of the monoterpene alcohol geraniol to geranial and geranic acid in transgenic tomato fruits (Davidovich-Rikanati et al., 2007).

The co-expression of *TcCDS* with *TcADH2*, and particularly the coexpression of these two genes with *TcALDH1*, greatly enhanced the production of *trans*-chrysanthemic acid in *N. benthamiana* leaves (Figure 6). While *trans*-chrysanthemylmalonylglucoside and Unknown 1 (a modified *trans*-chrysanthemol) were no longer detected in plants expressing all three genes, the malonylglucoside ester of *trans*-chrysanthemic and two other esters of this acid, Unknown 2 and Unknown 3, were present at a combined concentration exceeding the levels of free *trans*-chrysanthemic acid by a factor of 1.5. These results indicate that in addition to dehydrogenases that can act on *trans*-chrysanthemic acid precursors, *N. benthamiana* leaves also contain

enzymes that can use *trans*-chrysanthemic acid as a substrate and modify it further. The observation that endogenous enzymes in a plant cell engineered to make compounds not previously present in the cell can react with such new compounds has been made before (Lewinsohn and Gijzen, 2009). For a successful genetic engineering attempt to reconstruct the pyrethrin biosynthetic pathway in a heterologous system, it will be necessary to counteract such non-productive side reactions.

Biosynthesis of pyrethric acid

In three of the six pyrethrins that *T. cinerariifolium* synthesizes (pyrethrin II, cenerin II, and jasmolin II), pyrethric acid rather than *trans*-chrysanthemic acid constitute the acid moiety. Pyrethric acid is identical to *trans*-chrysanthemic acid except that the C8 position (sometimes referred to as C10) has a methylated carboxyl group, and is therefore likely derived from *trans*-chrysanthemic acid by a series of enzymatic reactions involving first an hydroxylation of C8, then two successive oxidations of C8 to give an aldehyde and then a carboxyl group, and finally a carboxylmethylation. The hydroxylation of C8 is likely to be catalyzed by an enzyme of the cytochrome P450 oxidoreductase family. The next two oxidations reactions, however, are equivalent to the reactions catalyzed by TcADH2 and TcALDH1, respectively. Our co-expression analysis identified only *TcADH1*, *TcADH2* and *TcALDH1* as likely candidate genes for involvement in the synthesis of the terpene moiety of pyrethrins. *TcADH6* was also identified as potential candidates, but the low levels of its transcripts in pyrethrin-producing tissue made it less likely to be involved in *trans*-chrysanthemic acid biosynthesis. Therefore, it is possible that the TcADH6 protein is involved in biosynthesis of pyrethric acid by catalyzing the conversion of the C8 alcohol to the C8 aldehyde, as pyrethrins containing the pyrethric acid moiety are much less abundant than those containing *trans*-chrysanthemic acid. It is also possible that TcADH1, which had no activity with *trans*-chrysanthemol, catalyzes this reaction, and perhaps even TcADH2 possesses this activity. It is notable that all these three ADHs are evolutionarily closely related to 8-hydroxygeraniol oxidoreductase from *Catharanthus roseus*, an enzyme that catalyzes the same C8 alcohol oxidation reaction on another monoterpenoid (Miettinen et al., 2014). Finally, it is possible that TcALDH1 catalyzes the conversion of the C8 aldehyde to give a carboxyl group. The lack of suitable substrates presently prevents the testing of these

hypotheses, but the availability of transgenic plants producing some of the precursors in the pathway to pyrethric acid may ameliorate this problem.

MATERIALS AND METHODS

Plant materials and chemicals

The various tissues including flowers, leaves, stem, and roots were collected from *Tanacetum cinerariifolium* grown in a greenhouse with a 16/8 hour day/night photoperiod. Harvested tissues were flash frozen in liquid nitrogen and stored at -80°C until use. All commercial chemicals were purchased from Sigma-Aldrich with the exception of 8-hydroxygeraniol which was obtained from Santa Cruz Biotechnology, and *trans*-chrysanthemol and *trans*-chrysanthemal.

trans-chrysanthemol was produced via a enzymatic method by TcCDS plus alkaline phosphatase (ALP) from DMAPP, and its concentration calculated according on GC-MS based on the standard curve of commercial mixture of *trans* and *cis*-chrysanthemol. *trans*-chrysanthemal was produced by TcADH2 and NAD⁺ from *trans*-chrysanthemol, extracted by MTBE from the reaction volume, dried, and dissolved in water.

GC-MS analysis of pyrethrins and terpenoids from *Tanacetum cinerariifolium* leaves and flowers of different stages of development

Leaf and flower tissues of different stages of development were cut into small pieces, 100 mg of which were transferred into a tube containing 200 µL MTBE with 0.01 ng/µl tetradecane as internal standard. The tube was vortexed for 3 min at maximum speed, incubated at room temperature for 2 hours, and the MTBE phase was then collected. Sequential extractions with MTBE on test samples indicated that in the initial extraction >94% of the chrysanthemic acid partitioned into the MTBE phase. The MTBE extracts were analyzed by GC-MS. The measurement of *trans*-chrysanthemic acid was preformed based on the corresponding standard curve.

RNAseq analysis

Total RNA was extracted from leaf and flower parts at different stages of development using the Total RNA Isolation Kit from Omega. The quantity and quality of extracted RNA for sequencing was determined with Qubit (Thermo Fisher Scientific, USA) and Bioanalyzer (Agilent Technologies, Palo Alto, California). Libraries for all 8 samples (flowers at stage 1, flowers at stage 2, flowers at stage 3, ray florets at stage 4, disk florets at stage 4, ray florets at stage 5, disk florets at stage 5, and leaf) were made using the KAPA stranded RNA-seq library preparation kit and sequenced at the Michigan State University Genomics Core on the Illumina HiSeq 2500 with HiSeq SBS reagents in 2x150 bp format. Initial reads have been deposited in NCBI (bioproject accession numbers PRJNA399494).

To assemble the reads, paired end RNA-seq reads were trimmed by quality scores using Trimmomatic v0.32 (Bolger et al., 2014) with the following parameters (*PE -threads 4 -phred33 ILLUMINACLIP:all_adapters_combined.fasta:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 HEADCROP:10 MINLEN:50*). Prior to generating a full-scale RNA-seq assembly, we tested the impact of changing the kmer value for *de novo* assembly using a set of normalized reads obtained using *insilico_read_normalization.pl* function provided in Trinity v2.2.0 (Haas et al., 2013), with the following options (*--seqType fq --JM 100G --max_cov 50 --pairs_together --SS_lib_type RF --CPU 10 --PARALLEL_STATS --KMER_SIZE 25 --max_pct_stdev 200*). The assemblies were analyzed for completeness by screening the Metazoan and Eukaryote comparison databases of BUSCO v1.22 (Simao et al., 2015) using default parameters. This analysis revealed k=31 to be the most optimal value of k for assembling complete transcripts. This result was also supported by analysis of the length distribution of transcripts using the *abyss_fac* function in AbySS 1.9.0 (Simpson et al., 2009), with k=31 producing the longest transcripts (N50=1810 bp). Thus, we performed an assembly of all RNA-seq reads with Trinity v2.2.0 using the following parameters: (*--seqType fq --KMER_SIZE 31 --max_memory 120G --SS_lib_type RF --CPU 16 --min_kmer_cov 2*). This assembly was used to identify ADH-like and ALDH-like transcripts using BLAST.

The expression levels of all transcripts were estimated using the *align_and_estimate_abundance.pl* function of the Trinity software, using the following parameters (*--transcripts Trinity.fasta --prep_reference --left PYR*_R1_*.fastq --right PYR*_R2_*.fastq --est_method RSEM --aln_method bowtie --SS_lib_type RF --thread_count 10 --max_ins_size 1000 --trinity_mode --seqType fq*), which estimated transcript abundance using the RSEM function (Li and Dewey, 2011). These expression values were used to identify transcripts correlated with CDSases and GLIP transcripts across all tissues.

Identifying ADH and ALDH sequences in the *T. cinerariifolium* transcriptome

To find candidate ADH and ALDH genes in our *T. cinerariifolium* transcriptome database, we queried it on the publicly accessible site (<http://sativa.mcdb.lsa.umich.edu/blast/>) with various plant ADH and ALDH sequences (e.g., cinnamyl alcohol dehydrogenase from *Populus trichocarpa*, accession number ACC63874; geraniol dehydrogenase from *Ocimum basilicum*, Q2KNL6; benzaldehyde dehydrogenase from *Antirrhinum majus*, ACM89738; aldehyde dehydrogenase 1 from *Artemisia annua*, ACR61719.1) using the TBLASTN function. We also screened the entire annotated database for the designation “alcohol dehydrogenase” and “aldehyde dehydrogenase” and added transcripts identified in this way to our list of candidates.

Co-expression analysis

The number of reads from each transcript in the various RNAseq analyses of RNA samples from different stages of development and parts of the plant were used to perform cluster analysis using the freely available Cluster 3.0 software program (<http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm#ctv>). This program is based on an algorithm developed by Eisen et al., (1998), and the analysis results in a tree and heat map that matches each gene with another gene whose expression pattern matches best with that of the first one.

Quantitative RT-PCR analysis of *TcCDS*, *TcGLIP*, *TcADH1*, *TcADH2* and *TcALDH1* transcript Levels

For real-time RT-PCR analysis of transcripts in different tissues, RNA was isolated using Total RNA Isolation Kit from Omega with a DNA digestion step. cDNAs were prepared using High Capacity cDNA Reverse Transcription Kit from Thermo Fisher Scientific following the manufacturer's instructions. Primer design and real-time PCR were performed following the manufacturer's instructions. Assays were performed using four independent biological replicates. The relative amounts of transcripts for different genes were normalized to glyceraldehyde-3-P dehydrogenase (GAPDH) transcript levels using LinRegPCR software (<http://www.hartfaalcentrum.nl/index.php?main=files&fileName=LinRegPCR.zip&description=LinRegPCR:%20qPCR%20data%20analysis&sub=LinRegPCR>). The C_T mean values for GAPDH are shown in Supplemental Figure S5. The statistics assay (unpaired *t*-test, two-tailed option) was performed via software GraphPad Prism (<https://www.graphpad.com/scientific-software/prism/>).

Generation, expression, and purification of recombinant *TcCDS*, *TcADH1*, *TcADH2* and *TcALDH1*

The open reading frames of *TcADH1*, *TcADH2* and *TcALDH1* were obtained by RT-PCR from prepared cDNA of flowers at stage 1 of *T. cinerariifolium*. The full-length cDNAs were introduced into the expression vector pET28a+ or pHIS8, in each case generating a fusion gene that encoded a “tag” of HIS₆ residues at the N-terminus for expression in *E. coli*. To obtain soluble proteins for expression in *Escherichia coli*, a truncated open reading frame of *TcCDS*, missing the first 50 codons, was obtained by RT-PCR from prepared cDNA of flowers at stage 1 of *T. cinerariifolium*. The truncated *TcCDS* ORF was inserted into the pEXP5-CT/TOPO vector following the manufacturer's instructions, generating a fusion gene that encodes a “tag” of HIS₆ residues at the C-terminus for expression in *E. coli*. The purification of recombinant proteins were performed as described previously (Xu et al., 2013). All constructs were transformed into *E. coli* BL21(+) cells. Transformed *E.coli* cells were grown in Luria-Bertani medium containing appropriate antibiotics until optical density of the culture at 600 nm reached 0.6 and then

recombinant gene expression was induced with the addition of IPTG to a final concentration of 0.15 mM and cells grown for overnight at 16°C. The resulting His-tagged fusion proteins were purified using Ni-NTA affinity columns.

Enzymatic assays of recombinant TcCDS

The enzymatic assay for TcCDS broadly followed the protocol described by Rivera et al. (2001) as follows: The reaction was initiated by adding 30 µg of affinity-purified His-tagged enzyme in a final volume of 50 µL of assay buffer (pH 7.5) containing 50 mM Tris-HCl, 2 mM DTT, 5 mM MgCl₂, and 2.5 mM DMAPP. The assay was incubated at 30°C for 24 hours. To analyze the production of CDP in this assay by GC-MS, hydrolysis of CDP to *trans*-chrysanthemol by 5 units of Roche rAPid alkaline phosphatase (Sigma) was preformed following the manufacturer's instructions at 37°C for 2 h. Reaction products were extracted with 100 µL MTBE and the MTBE extract was injected and analyzed by GC-MS.

Enzymatic assays of recombinant TcADH1 and TcADH2

For assaying of substrate specificity of TcADH1 and TcADH2, reactions were initiated by adding 1.25 µg of affinity-purified His-tagged enzyme in a final volume of 50 µL of assay buffer (pH 8.0) containing 50 mM Tris-HCl, 2 mM DTT, 1 mM NAD⁺ and 0.3 mM of selected alcohols. The assays were incubated at 30°C for 10 min, after which reaction products were extracted with 100 µL MTBE. The MTBE extract was analyzed by GC-MS for product and remaining substrate.

To determine the kinetic parameters of TcADH2, a similar protocol was followed. The K_m value for NAD⁺ was determined by using 0.64 mM *trans*-chrysanthemol whereas the K_m value for *trans*-chrysanthemol was determined with 1 mM NAD⁺. K_m and k_{cat} values were calculated from initial rate data by using the hyperbolic regression analysis method in Hyper32 software (version 1.0.0, <http://hyper32.software.informer.com/>).

Enzymatic assays of recombinant TcALDH1

For testing the substrate specificity of TcALDH1, reactions were initiated by adding 2.5 µg of affinity-purified His-tagged enzyme in a final volume of 50 µL of assay buffer (pH 8.5) containing 50 mM Tris-HCl, 2 mM DTT, 1 mM NADP⁺ and 40 µM selected aldehydes. The assays were incubated at 30°C for 10 min after which reaction products were extracted with 100 µL MTBE. The MTBE extract was analyzed by GC-MS. The reaction rate was calculated according to the decrease of substrate based on the corresponding standard curve.

To determine the kinetic parameters of TcALDH1, a similar protocol was followed. The K_m value for NAD⁺ and NADP⁺ was determined by using 40 µM *trans*-chrysanthemal whereas the K_m value for *trans*-chrysanthemal was determined with 1mM NAD⁺ or NADP⁺. The reaction rate was calculated according to production of *trans*-chrysanthemic acid based on the corresponding standard curve. K_m and k_{cat} values were calculated as described above.

Plasmid construction for transient expression in *N. benthamiana* leaves

The complete open reading frame of *EGFP* was amplified from pSAT6-EYFP-N1 vector and used as control gene in this study. The *EGFP*, *TcCDS*, *TcADH2* and *TcALDH1* genes were spliced into pEAQ-HT binary vector between AgeI and XhoI restriction sites using the NEBuilder HiFi DNA Assembly Cloning Kit (<https://www.neb.com/products/e5520-nebuilder-hifi-dna-assembly-cloning-kit>, NEB) according to manufacturer's instructions to express them under the control of 35S promoter.

Transient expression in leaves of *N. benthamiana*

Agrobacterium tumefaciens strain GV3101 infiltration (agro-infiltration) was performed as described previously (Sainsbury et al., 2009). Briefly, *A. tumefaciens* strain GV3101 was grown at 28°C at 200 rpm for 24 hours in LB medium containing kanamycin (50 mg L⁻¹), rifampicin (50 mg L⁻¹) and gentamycin (25 mg L⁻¹). Cells were collected by centrifugation for 15 min at 3000 *g* at 20°C, and then resuspended in 10 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer containing 10 mM MgCl₂ and 100 µM acetosyringone (4'-hydroxy-3',5'-dimethoxyacetophenone) to a final OD₆₀₀ of 0.4, followed by incubation at 20°C for 3 hours. For co-infiltration, equal number of cells from each of the cultures of strains harboring different

binary plasmids were mixed together, collected by centrifugation, and resuspended as above to a final OD₆₀₀ of 0.4 per strain. *N. benthamiana* plants were grown from seeds on soil in a greenhouse with 16/8 hours day/night photoperiod at 25°C. Leaves of 4-week-old *N. benthamiana* plants were infiltrated using a 2 ml syringe without a needle. *N. benthamiana* plants transformed with the binary vector harboring EGFP alone were used as control. The infiltrated leaves were collected 10 days after infiltration.

To analyze the compounds produced in the leaves, the harvested plant materials were flash-frozen and ground into a fine power in liquid nitrogen. Three grams of powder was extracted with 4 ml MTBE. The extracts were briefly vortexed for 3 min at maximum speed and then incubated at room temperature with shaking at 50 rpm for 3 h, followed by centrifugation for 15 min at 8000 *g*. The MTBE layer was transferred to a fresh vial, dehydrated using anhydrous Na₂SO₄ and concentrated by evaporating the solvent to a final volume of about 0.3 ml. Analysis of the samples was performed with an Rxi-5Sil column on a Shimadzu QP-2010 GC-MS system.

GC-MS and LC-MS analyses

Analytes from 1 µl samples were separated by Shimadzu QP-2010 GC-MS system equipped with the Rxi-5Sil column (30 m × 0.25 mm × 0.25 µm film thickness, RESTEK, USA) using helium as the carrier gas at a flow rate of 1.4 ml min⁻¹. The injector was used in split mode at ratio of 1:2 with the inlet temperature set to 240°C. The initial oven temperature of 50°C was increased after holding for 3 min to 110°C at a rate of 10°C min⁻¹, then increased to 150 °C at a rate of 5°C min⁻¹, held for 3 min at 150 °C, increased to 300°C at a rate of 10°C min⁻¹, and finally held for 3 min at 300°C. Compounds were identified by comparison of mass spectra and retention time with those of the authentic standards, when available, or with known retention indices and mass fragmentation from the literature and NIST library.

Each MTBE extract (4 mL) was evaporated by BUCHI Rotavapor and dissolved in 0.5 ml of 70% acetonitrile/30% water for analysis using a Waters Xevo G2-XS Q-Tof mass spectrometer interfaced with a Waters Acquity binary solvent manager and 2777c autosampler. Samples (5 µL each) were injected onto an Acquity BEH C18 UPLC column (2.1 x 100 mm, 1.7 µm particle size; Waters Corp) at 40°C. Initial conditions were 0.3 ml/min of 99% solvent A (water + 0.1% formic acid) and 1% solvent B (acetonitrile). Following injection, solvent B was

increased in a linear gradient over 16 min to 99%, followed by a hold at 99% B for 2 min, then return to 99% A at 18.01 min and equilibrate for 2 min before starting the next sample. Ions were generated by electrospray ionization in negative-ion mode. Capillary voltage was 2.0 kV, sample cone voltage was 40 V and source temperature was 100°C. Desolvation temperature was 350°C and desolvation gas flow was 600 L/hr. Data were acquired using a data-independent MS^e method providing both non-fragmenting and fragmenting conditions for each run, and lock mass correction was performed using a leucine enkephalin standard. MS/MS was performed for selected ions in separate runs.

Hydrolysis of modified *trans*-chrysanthemol and *trans*-chrysanthemic acid

Acylglucosides such as chrysanthemic acid glucoside can typically be hydrolyzed by base, but chrysanthemyl glucosides cannot. Some non-specific glycosidases can hydrolyze both classes of compounds. We therefore used both and glycosidase treatments to hydrolyze volatile and non-volatile derivatives of *trans*-chrysanthemol and *trans*-chrysanthemic acid produced in *N. benthamiana* leaves. Samples were prepared by flash-freezing leaves and grounding them into a fine power in liquid nitrogen. For base hydrolysis, homogenates (0.5 g per sample) were mixed completely with 50 µL 4N NaOH and incubated at 80°C for 20 min, followed by neutralization with 50 µL 4N HCl, and then adding 1 mL MTBE containing 0.01 ng/µl tetradecane as internal standard and vortexing at maximum speed for 4 min. MTBE extracts were transferred to a fresh vial, dehydrated and concentrated to a final volume of about 0.2 mL for GC-MS analysis, or dried and dissolved in 0.3 ml of 70% acetonitrile/30% water for LC-MS analysis. For glycosidase treatment, homogenates were treated with 4 mg of β-glucosidase (EC 3.2.1.21, purchased from Sigma) at 37°C for 1.5 hour, and then extracted and analyzed as described above. GC-MS analysis was used to measure volatile compounds as described above. LC-MS analysis was performed to check for disappearance of the non-volatile *trans*-chrysanthemic acid conjugates also as described above.

Accession Numbers

The sequence data used in this study can be obtained from NCBI with the following GenBank accession numbers: TcADH1 – MF497443; TcADH2 – MF497444; TcALDH1 – MF497445. The bioproject accession numbers for the RNAseq data is PRJNA399494

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Mass spectral analysis of Unknown 1, Unknown 2 and Unknown 3 produced in *N. benthamiana* leaves expressing *TcCDS*, *TcADH2* and *TcALDH1*.

Supplemental Figure S2. GC-MS analysis of hydrolysis assays of tissues of *N. benthamiana* leaves expressing *TcCDS*.

Supplemental Figure S3. GC-MS analysis of hydrolysis assays of tissues of *N. benthamiana* leaves expressing *TcCDS*, *TcADH2* and *TcALDH1*.

Supplemental Figure S4. LC-MS analysis of hydrolysis assays of tissues of *N. benthamiana* leaves expressing *TcCDS*, *TcADH2* and *TcALDH1*.

Supplemental Figure S5. The C_T mean values of *TcGAPDH*.

Supplemental Table S1. Transcript Abundance for Selected Candidate in the RNASeq database of leaves and flowers at different stages from *Tanacetum cinerariifolium*.

ACKNOWLEDGMENTS

We thank Prof. Daniel Jones (Michigan State University) for help with the analysis of non-volatiles compounds in *N. benthamiana* leaves transiently expressing various *T. cinerariifolium* genes.

665
666
667
668
669
670

Figure Legends:

Figure 1. A. Structures of pyrethrins. B. Proposed pathway for the biosynthesis of *trans*-chrysanthemic acid.

Figure 2. GC-MS analysis of pyrethrins and terpenoids from *Tanacetum cinerariifolium* leaves and flowers at different stages of development.

A. Flowers of different stages of development and a leaf of *T. cinerariifolium*.

B. Changes in relative concentrations of pyrethrin I during floral development. Pyrethrin I is the most abundant pyrethrin in the flower, and changes in concentrations of other pyrethrins follow the same pattern as those of pyrethrin I.

C. GC-MS chromatogram of total ion mode of MTBE extracts from leaves and flowers harvested at stage 4. In each flower/leaf comparison, samples are shown with the same relative y-axis scale, but the 7.2-15.2 min section is shown at a smaller scale to magnify the peaks. Peaks identified as terpenoids and internal standard (tetradecane) are labeled.

D. GC-MS chromatogram (total ion mode) of MTBE extracts from leaves and flowers of different stages of development, showing the *trans*-chrysanthemic acid levels in each sample.

E. Concentrations of *trans*-chrysanthemic acid in the leaf and in different stages of flowers.

Quantification was achieved by normalization of the peaks in (D) to the tetradecane internal standard and comparison to a standard curve of authentic *trans*-chrysanthemic acid ($n = 3$; means $\pm$ SD).

Figure 3. Identification of candidate *ADH* and *ALDH* genes for *trans*-chrysanthemic acid biosynthesis.

A. Images of *T. cinerariifolium* flowers of different stages and of leaves from which RNA samples were obtained for RNAseq analysis.

B. Average-linkage hierarchical clustering of relative transcript abundance of putative ADHs and ALDHs with *TcCDS* and *TcGLIP* based on number of reads of each transcript in each RNASeq library. Tree and heat map generated by the Cluster 3.0 software (see Materials and Methods).

C. Verification of levels of expression of *TcCDS*, *TcGLIP*, *TcADH1*, *TcADH2*, and *TcALDH1* by qRT-PCR. Transcript levels are expressed relative to that of *GAPDH* (glyceraldehyde-3-

phosphate dehydrogenase) in each sample ($n = 4$; means $\pm$ SD). ** $p < 0.01$, * $p < 0.05$. The differences between leaf, stem and root datapoints and any flower datapoints are all significant at $p < 0.001$.

Figure 4. Phylogenetic analysis of candidate *T. cinerariifolium* dehydrogenases for *trans*-chrysanthemic acid biosynthesis based on protein sequences.

A. Phylogenetic tree for TcADH1 and TcADH2.

B. Phylogenetic tree for TcALDH1.

The protein sequences from other species are of functionally characterized enzymes whose sequences were identified by BLAST search to be most closely related to the *T. cinerariifolium* sequences. Phylogenetic analyses were conducted in MEGA7 (Kumar et al., 2016) with the following parameters; multiple sequence alignment with ClustalW, phylogenetic construction with the Maximum Likelihood method and bootstrap tests of 1000 replicates.

Figure 5. Gas chromatography analyses of products obtained in *in vitro* biochemical assays of TcADH2 and TcALDH1. For all assays analyzed, reactions products were extracted with 100 μ l MTBE and run on Rxi-5Sil column. Tetradecane was used as internal standard.

A. Synthesis of *trans*-chrysanthemol substrate. Top chromatograph trace: 1 mM of a commercially available standard of *trans*- and *cis*-chrysanthemol mixture. Middle trace: Reaction products obtained by incubating 30 μ g recombinant TcCDS with 2.5 mM DMAPP in 50 μ L reaction for 24 hours at 30°C. Bottom chromatograph trace: Reaction products obtained by incubating the products of the TcCDS-catalyzed condensation of DMAPP with 5 units of alkaline phosphatase (ALP) for 1 hour at 37°C.

B. *In vitro* production of *trans*-chrysanthemol. Reaction products obtained by incubating 0.64 mM *trans*-chrysanthemol and 1.5 mM NAD^+ with 5 μ l eluted protein from empty vector (top chromatograph trace) or 1.25 μ g purified TcADH2 (bottom chromatograph trace) in 60 μ L reaction volume for 5 min.

C. Production of *trans*-chrysanthemic acid from *trans*-chrysanthemol in a coupled assay containing 0.64 mM *trans*-chrysanthemol and 1.5 mM NAD^+ with 1.25 μ g purified TcADH2 and 6.00 μ g purified TcALDH1 in 60 μ L reaction volume for 5, 10, 15, 25, and 45 min. A control

reaction was performed using 5 µl eluted protein from empty vector. Bottom chromatograph trace:
0.3 mM of commercial *trans*-chrysanthemic acid.

Figure 6. Production of *trans*-chrysanthemol, *trans*-chrysanthemic acid, and related compounds in *N. benthamiana* leaves transiently expressing TcCDS, TcADH2, and TcALDH1 proteins.

A. GC-MS chromatograms of MTBE extracts of *N. benthamiana* leaves expressing *EGFP* (control), *TcCDS*, *TcCDS* and *TcADH2*, and *TcCDS* with *TcADH2* and *TcALDH1*. For terpenes, $m/z = 123$ was monitored, and for the internal control tetradecane, $m/z = 198$ was monitored. Peaks related to *trans*-chrysanthemol, *trans*-chrysanthemic acid and internal standard are labeled.

B. Concentrations of free *trans*-chrysanthemol and (C) free *trans*-chrysanthemic acid in *N. benthamiana* leaves expressing the indicated constructs were determined by comparison with an authentic standard.

D-F. The relative levels of Unknown 1 (D), Unknown 2 (E), and Unknown 3 (F) in *N. benthamiana* leaves expressing the indicated constructs. For each compound, the plant material expressing a specific construct that showed the highest levels (average of three biological replicates) was set at 100%.

The data in B–F represent mean $\pm$ SD from triplicate biological replicates. N.D. - not detected.

Figure 7. LC-MS analysis of *N. benthamiana* leaves simultaneously expressing the three enzymes TcCDS, TcADH2, and TcALDH1. Extracted ion chromatograms of m/z 803.37 and 831.33 are shown for: A. *EGFP* single expression control, B. *TcCDS*, C. *TcCDS* + *TcADH2*, D. *TcCDS* + *TcADH2* + *TcALDH1*. Chromatograms are all scaled the same, as indicated by the ion current (1.01×10^7) in the upper right corner of each chromatogram. The sample for panel D was diluted 10-fold compared to the other three samples due to high concentration of m/z 831.33 in the sample. MS/MS spectra are shown for ions m/z 803.37 (E) and m/z 831.33 (F) along with the proposed compound structures based on exact mass, fragmentation pattern, and similarity to previously published data. G. Relative levels of *trans*-chrysanthemic acid in *N. benthamiana* leaves expressing all *TcCDS*, *TcADH2* and *TcALDH1* with or without sodium hydroxide treatment.

Table 1: Relative activities of recombinant TcADH1, TcADH2 and TcALDH1 with selected substrates

Alcohols	TcADH1	TcADH2	Aldehydes	TcALDH1
<i>cis</i> -3-Hexenol	N.D. ^a	N.D.	<i>trans</i> -2-Heptenal	47
Perillyl alcohol	N.D.	35	Octanal	90
Nerol	100 ^b	42	Dodecanal	47
Geraniol	48	38	Citral	20
(<i>S</i>)- β -citronellol	8	22	Perillaldehyde	40
<i>trans,trans</i> -Farnesol	N.D.	N.D.	Farnesal	54
<i>trans</i> -Chrysanthemol	N.D.	100 ^c	<i>trans</i> -Chrysanthemal	100 ^c
8-Hydroxygeraniol	N.D.	70	(<i>S</i>)-(-)-Citronellal	84
Benzyl alcohol	N.D.	N.D.	Benzaldehyde	90
Cinnamyl alcohol	N.D.	N.D.	<i>trans</i> -Cinnamaldehyde	8

The activities of TcADH1 and TcADH2 were measured with 0.3 mM alcohols and 1mM NAD⁺.

The activities of TcALDH1 were measured with 0.04 mM aldehydes and 1mM NADP⁺. Data are expressed as relative mean percentages from triplicate independent assays.

^aN.D., not detectable (< 5% of highest activity)

^b100% relative activity, corresponds to 0.44 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ of citral.

^c100% relative activity, corresponds to 0.50 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ of *trans*-chrysanthemal.

^e100% relative activity, corresponds to 0.11 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ of *trans*-chrysanthemic acid.

Table 2: Kinetic properties of recombinant TcADH2 and TcALDH1

Enzyme	Substrate	$K_m(\mu\text{M})$	$K_{cat}(\text{s}^{-1})$	$K_{cat}/K_m (\text{M}^{-1}\text{s}^{-1})$
TcADH2	<i>trans</i> -chrysanthemol ^a	236.0 ± 5.8	0.75 ± 0.0032	3186.2
	NAD ^{+b}	192.6 ± 8.7	0.64 ± 0.0034	3345.5
TcALDH1	<i>trans</i> -chrysanthemal ^c	4.4 ± 2.2	0.11 ± 0.0049	25122.4
	NADP ^{+d}	20.4 ± 7.1	0.090 ± 0.0013	4391.8
	<i>trans</i> -chrysanthemal ^e	4.6 ± 1.8	0.096 ± 0.0032	20992.3
	NAD ^{+f}	68.6 ± 17.1	0.086 ± 0.0012	1256.07

Data are presented as mean ± SD from triplicate independent assays. All assays were performed on GC-MS.

^a Kinetic parameters were determined with 1 mM NAD⁺.

^b Kinetic parameters were determined with 0.64 mM *trans*-chrysanthemol.

^c Kinetic parameters were determined with 1 mM NADP⁺.

^d Kinetic parameters were determined with 0.040 mM *trans*-chrysanthemal.

^e Kinetic parameters were determined with 1 mM NAD⁺.

^f Kinetic parameters were determined with 0.040 mM *trans*-chrysanthemal.

Parsed Citations

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Casida, J. E (1973) Pyrethrum, the Natural Insecticide, Academic Press, New York

Casida, J.E., and Quistad, G.B (1995) Pyrethrum Flowers: Production, Chemistry, Toxicology, and Uses (New York: Oxford University Press)

Crombie, L (1995) Chemistry of the pyrethrins. In Pyrethrum Flowers (Casida, J.E. and Quistad, G.B. eds). Oxford: Oxford University Press, pp. 123–193

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Davidovich-Rikanati R, Sitrit Y, Tadmor Y, Iijima Y, Bilenko N, Bar E, Carmona B, Fallik E, Dudai N, Simon JE, Pichersky E, Lewinsohn E (2007) Enrichment of tomato flavor by diversion of the early plastidial terpenoid pathway. Nat Biotechnol 25: 899-901

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863-14868

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nature Protocols 2: 953-971

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8: 1494-1512

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Iijima Y, Wang G, Fridman E, Pichersky E (2006) Analysis of the enzymatic formation of citral in the glands of sweet basil. Arch Biochem Biophys 448: 141-149

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Katsuda Y (1999) Development and future prospects for pyrethroid chemistry. Pestic. Sci 55:

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

775–782

Katsuda Y (2012) Progress and future of pyrethroids. Top Curr Chem 314: 1-30

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kikuta Y, Ueda H, Takahashi M, Mitsumori T, Yamada G, Sakamori K, Takeda K, Furutani S, Nakayama K, Katsuda Y, Hatanaka A, Matsuda K (2012) Identification and characterization of a GDSL lipase-like protein that catalyzes the ester-forming reaction for pyrethrin biosynthesis in *Tanacetum cinerariifolium*- a new target for plant protection. Plant Journal 71: 183-193

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33: 1870-1874

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Lewinsohn E, Gijzen M (2009) Phytochemical diversity: The sounds of silent metabolism. Plant Science 176: 161-169

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Magnard JL, Roccia A, Caissard JC, Vergne P, Sun P, Hecquet R, Dubois A, Hibrand-Saint Oyant L, Jullien F, Nicole F, Raymond O, Huguët S, Baltenweck R, Meyer S, Claudel P, Jeauffre J, Rohmer M, Foucher F, Huguéney P, Bendahmane M, Baudino S (2015) Plant volatiles. Biosynthesis of monoterpene scent compounds in roses. Science 349: 81-83

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Pollier J, Woittiez L, van der Krol S, Lugan R, Ilc T, Verpoorte R, Oksman-Caldentey KM, Martinoia E, Bouwmeester H, Goossens A, Memelink J, Werck-Reichhart D (2014) The seco-iridoid pathway from *Catharanthus roseus*. Nat Commun 5: 3606

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Moghe GD, Leong BJ, Hurney S, Jones AD, Last RL (2017) Evolutionary routes to biochemical innovation revealed by integrative analysis of a plant-defense related specialized metabolic pathway. Elife 6

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends in Biochemical Sciences 24: 34-35

Ramirez AM, Saillard N, Yang T, Franssen MC, Bouwmeester HJ, Jongsma MA (2013) Biosynthesis of sesquiterpene lactones in pyrethrum (*Tanacetum cinerariifolium*). PLoS One 8: e65030

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Ramirez AM, Stoop G, Menzel TR, Gols R, Bouwmeester HJ, Dicke M, Jongsma MA (2012) Bidirectional secretions from glandular trichomes of pyrethrum enable immunization of seedlings. Plant Cell 24: 4252-4265

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Rivera SB, Swedlund BD, King GJ, Bell RN, Hussey CE, Jr., Shattuck-Eidens DM, Wrobel WM, Peiser GD, Poulter CD (2001) Chrysanthemyl diphosphate synthase: isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from *Chrysanthemum cinerariaefolium*. Proc Natl Acad Sci U S A 98: 4373-4378

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Sainsbury F, Thuenemann EC, Lomonosoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol J 7: 682-693

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210-3212

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only Title Only Author and Title](#)

Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19: 1117-1123

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Xu H, Zhang F, Liu B, Huhman DV, Sumner LW, Dixon RA, Wang G (2013) Characterization of the formation of branched short-chain fatty acid:CoAs for bitter acid biosynthesis in hop glandular trichomes. Mol Plant 6: 1301-1317

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

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

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yang T, Gao L, Hu H, Stoopen G, Wang C, Jongsma MA (2014) Chrysanthemyl diphosphate synthase operates in planta as a bifunctional enzyme with chrysanthemol synthase activity. J Biol Chem 289: 36325-36335

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)