Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Glucosylation of chrysanthemic acid drives the final acylation step in pyrethrin biosynthesis

Sotirios Kampranis

soka@plen.ku.dk

University of Copenhagen https://orcid.org/0000-0001-6208-1684

Chaojie Guo

University of Copenhagen https://orcid.org/0000-0003-3181-2717

Neda Popovic

University of Copenhagen

Mohammed Motawie

University of Copenhagen https://orcid.org/0000-0001-5582-9463

Fernando Geu-Flores

University of Copenhagen https://orcid.org/0000-0002-5735-9810

Irini Pateraki

University of Copenhagen https://orcid.org/0000-0002-7526-2334

Article

Keywords:

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11528003/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Glucosylation of chrysanthemic acid drives the final acylation step in pyrethrin biosynthesis

Chaojie Guo[1], Neda Popović[1*], Mohammed S. Motawia[1], Fernando Geu-Flores[1], Irini Pateraki[1], Sotirios C. Kampranis[1*]

[1]Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

*Correspondence: soka@plen.ku.dk; neda@plen.ku.dk

Abstract

Pyrethrins are biodegradable plant-derived insecticides used in organic agriculture, but commercial production relies on extraction from pyrethrum flowers, constraining supply. Biotechnological production of pyrethrins is hindered by incomplete understanding of their biosynthesis. The final pathway steps were proposed to involve CoA-dependent activation of monoterpenoid acids followed by acylation of rethrolone alcohols. Here, systematic functional evaluation of 73 candidate CoA ligases did not support the proposed CoA-dependent mechanism. Instead, we identify a glucose ester-dependent route involving a *Tanacetum* UDP-glucosyltransferase, *Tco*UGT1, which converts chrysanthemic acid to chrysanthemoyl-β-D-glucose. Biochemical assays demonstrate that the GDSL enzyme *Tc*GLIP uses chrysanthemoyl-β-D-glucose to acylate pyrethrolone, producing pyrethrin I. Heterologous reconstruction in *Saccharomyces cerevisiae* enables jasmolin I formation from supplied precursors, while structural modelling supports recognition of the glucose moiety by *Tc*GLIP. These findings elucidate the final two steps of pyrethrin biosynthesis, expand the known acyl-donor repertoire of GDSL enzymes, and provide a foundation for microbial pyrethrin production.

Introduction

Pyrethrins are potent plant-derived insecticides that have long been used in agriculture. They are produced by the flowers of *Tanacetum cinerariifolium* (Dalmatian pyrethrum) and *Tanacetum coccineum* (Persian pyrethrum), where they contribute to defence against herbivores[1,2].

Their insecticidal properties inspired the development of synthetic pyrethroids with improved stability[2], which are now widely used to control insect pests in agriculture, on livestock and companion animals, and in households[3,4]. However, the environmental persistence of synthetic pyrethroids can lead to their accumulation in sediments and adverse effects on non-target organisms[3,4]. By contrast, pyrethrins degrade rapidly in the environment and are among the relatively few insecticides permitted for use in organic agriculture[1,5,6]. With growing concerns over the impacts of pesticide use on human health and the environment[7-10], natural pyrethrins represent an important alternative for insect pest control[1,11-13].

Despite their advantages, the commercial supply of natural pyrethrins remains constrained by their dependence on extraction from dried pyrethrum flower heads, primarily cultivated in Tasmania and East Africa[14]. Pyrethrum cultivation is susceptible to climatic variability and plant diseases, resulting in fluctuating yields and supply limitations. Establishing pyrethrin biosynthesis in a heterologous host could provide an alternative, potentially scalable production platform, reducing dependence on agricultural extraction[14,15]. However, achieving biotechnological production requires the complete elucidation and functional reconstruction of the pyrethrin biosynthetic pathway[13,15,16]

Pyrethrins are a group of six esters formed by the condensation of a monoterpenoid acid, either chrysanthemic acid or pyrethric acid, with a rethrolone-type alcohol[5]. Their biosynthesis proceeds through two branches that generate the acid and alcohol moieties (**Fig. 1**). The monoterpenoid branch begins in plastids, where chrysanthemyl diphosphate synthase (CDS) catalyzes an atypical head-to-middle condensation of two molecules of dimethylallyl diphosphate (DMAPP) to form chrysanthemyl diphosphate (CDP). CDP is subsequently dephosphorylated to chrysanthemol (**1**) by

Nudix hydrolases[5]. From this common intermediate, alcohol dehydrogenase 2 (ADH2) and aldehyde dehydrogenase 1 (ALDH1) sequentially oxidise chrysanthemol (**1**) to chrysanthemal (**2**) and chrysanthemic acid (**3**), respectively. A second route converts chrysanthemol to pyrethric acid through successive oxidation reactions catalysed by chrysanthemol 10-hydroxylase (CHH), ADH2 and ALDH1, followed by *O*-methylation mediated by 10-carboxychrysanthemic acid 10-methyltransferase (CCMT)[5,17]. In parallel, the rethrolone branch proceeds from α-linolenic acid to cis-jasmone (**4**), which is hydroxylated by jasmone hydroxylase (*Tc*JMH) to jasmolone (**5**) and subsequently oxidised by pyrethrolone synthase (*Tc*PYS) to pyrethrolone (**6**)[18,19].

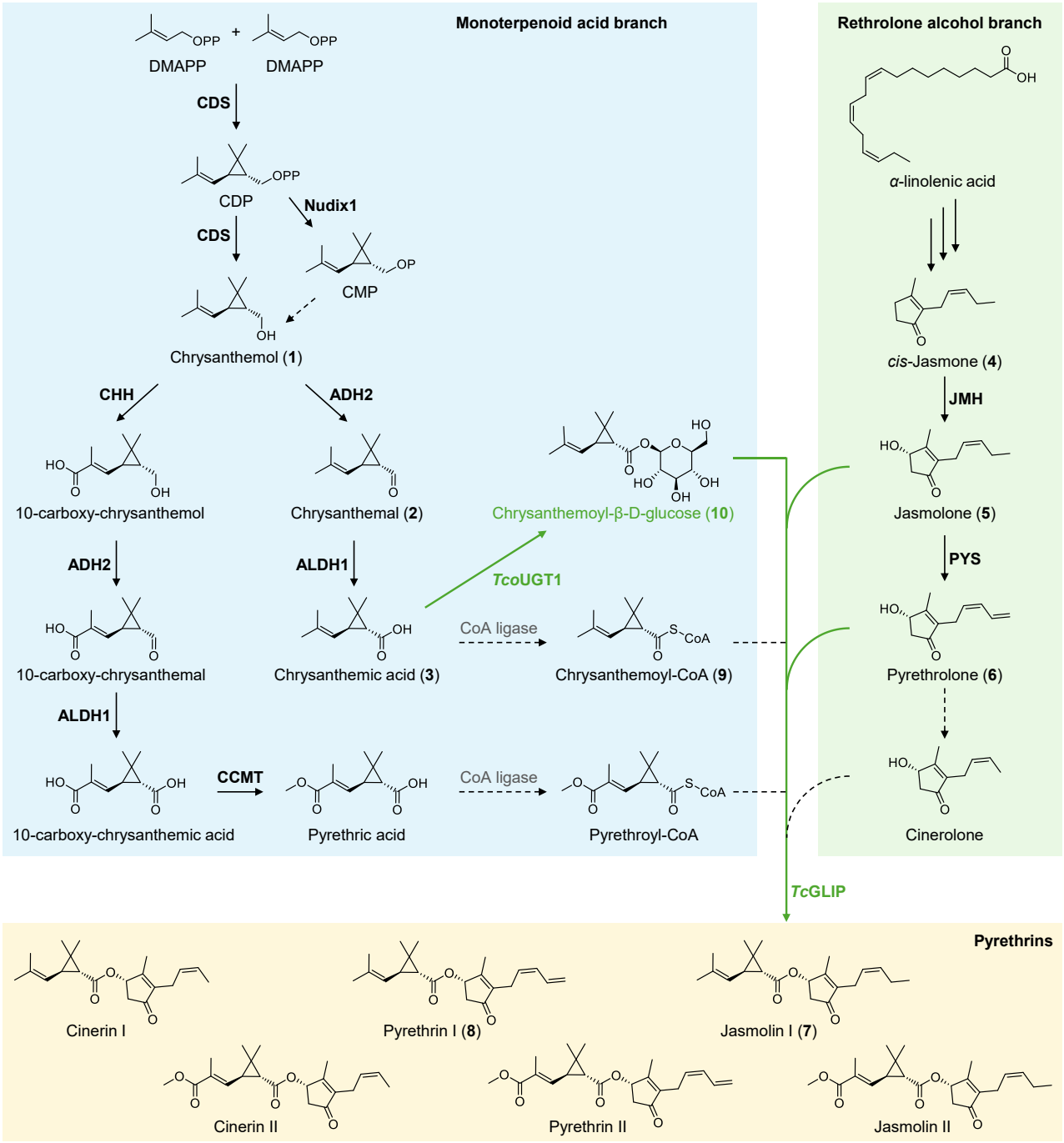


Fig. 1. Established and proposed steps in pyrethrin biosynthesis. The pathway comprises two branches that generate the monoterpenoid acid and rethrolone alcohol moieties, which are joined through *Tc*GLIP-catalysed esterification. In the glucose ester-dependent route identified in this study, the *T. coccineum* UDP-glucosyltransferase *Tco*UGT1 converts chrysanthemic acid (**3**) to chrysanthemoyl-β-D-glucose (**10**), which serves as an acyl donor for *Tc*GLIP-mediated pyrethrin formation. Solid arrows indicate established reactions, whereas dashed arrows represent proposed or incompletely characterized steps. Multiple consecutive arrows indicate successive reactions. The route identified in this study is highlighted in green. Abbreviations: DMAPP, dimethylallyl diphosphate; CDS, chrysanthemyl diphosphate synthase; CDP, chrysanthemyl diphosphate; Nudix, Nudix hydrolase; ADH2, alcohol dehydrogenase 2; ALDH1, aldehyde dehydrogenase 1; CHH,

chrysanthemol 10-hydroxylase; CCMT, 10-carboxychrysanthemic acid 10-methyltransferase; JMH, jasmone hydroxylase; PYS, pyrethrolone synthase; GLIP, GDSL lipase-like protein.

The two branches converge in the final steps of the pathway. The monoterpenoid acids (chrysanthemic acid or pyrethric acid) are first activated and subsequently esterified with different rethrolone alcohols to yield six pyrethrins, including jasmolin I (**7**), jasmolin II, pyrethrin I (**8**), pyrethrin II, cinerin I, and cinerin II[20] (**Fig. 1**). The activation step has been proposed to involve the formation of acyl-CoA thioesters, and the *T. cinerariifolium* acyl-activating enzyme *Tc*AAE1 has been implicated in this reaction[21]. The subsequent esterification is thought to be catalysed by the GDSL lipase-like protein *Tc*GLIP[20]. (**Fig. 1**). However, the proposed activities of *Tc*AAE1 and *Tc*GLIP have only been assigned using *in vitro* experiments[20,21], and have not been validated in a plant system.

Here, we investigate the unresolved final steps of pyrethrin biosynthesis. We find that *Tc*AAE1 does not enhance pyrethrin formation in the plant system *Nicotiana benthamiana* and proceed to discover an alternative CoA ligase enzyme that catalyzes the postulated chrysanthemic acid (**3**) activation. Despite screening 72 additional candidate CoA ligases from *T. cinerariifolium* and *T. coccineum*, we find no evidence supporting the proposed CoA-dependent activation of chrysanthemic acid in plant, yeast, or *in vitro* assays. Instead, we identify a *Tanacetum* UDP-glucosyltransferase, *Tco*UGT1, that converts chrysanthemic acid to 1-O-chrysanthemoyl-β-D-glucopyranose (chrysanthemoyl-β-D-glucose). We demonstrate that this glucose ester serves as an acyl donor for the GDSL enzyme *Tc*GLIP, enabling pyrethrin formation. We validate this route through biochemical assays and heterologous pathway reconstruction in *Saccharomyces cerevisiae*. To our knowledge, this is the first demonstration of a GDSL acyltransferase using a glucose ester as an acyl donor, a substrate class previously associated primarily with SCPL acyltransferases. These findings elucidate the two final steps of pyrethrin biosynthesis and lay a foundation for their biotechnological production.

RESULTS

*Tc*AAE1 does not enhance jasmolin I formation

To evaluate the proposed role of *Tc*AAE1 in the CoA-dependent activation of chrysanthemic acid (**3**)[21], we established a functional assay based on jasmolin I (**7**) production using transient expression in *N. benthamiana* (**Fig. 2a**). Tobacco leaves were co-infiltrated with AGL1 agrobacteria expressing *TcJMH*, *TcGLIP*, and *TcAAE1*, and supplemented with chrysanthemic acid (**3**) and *cis*-jasmone (**4**). Following GC-EI-MS analysis of leaf extracts, we were able to detect production of jasmolin I (**7**), validated by comparison with an authentic standard (**Fig. 2b**). However, we observed that jasmolin I (**7**) was also present in extracts from the control samples, where *TcAAE1* was not expressed (**Fig. 2b**). Furthermore, the levels of jasmolin I (**7**) synthesized in the presence of *Tc*AAE1, *Tc*JMH, and *Tc*GLIP, were similar to the levels produced when only *Tc*JMH and *Tc*GLIP were present (**Fig. 2b**; **Supplementary Fig. 1**). Notably, we observed that jasmolin I (**7**) production was not detectable in the absence of *Tc*GLIP, suggesting that endogenous *N. benthamiana* metabolism can provide an acyl donor for *Tc*GLIP. The absence of an observable contribution from *Tc*AAE1 prompted us to investigate additional candidate enzymes involved in chrysanthemic acid (**3**) activation.

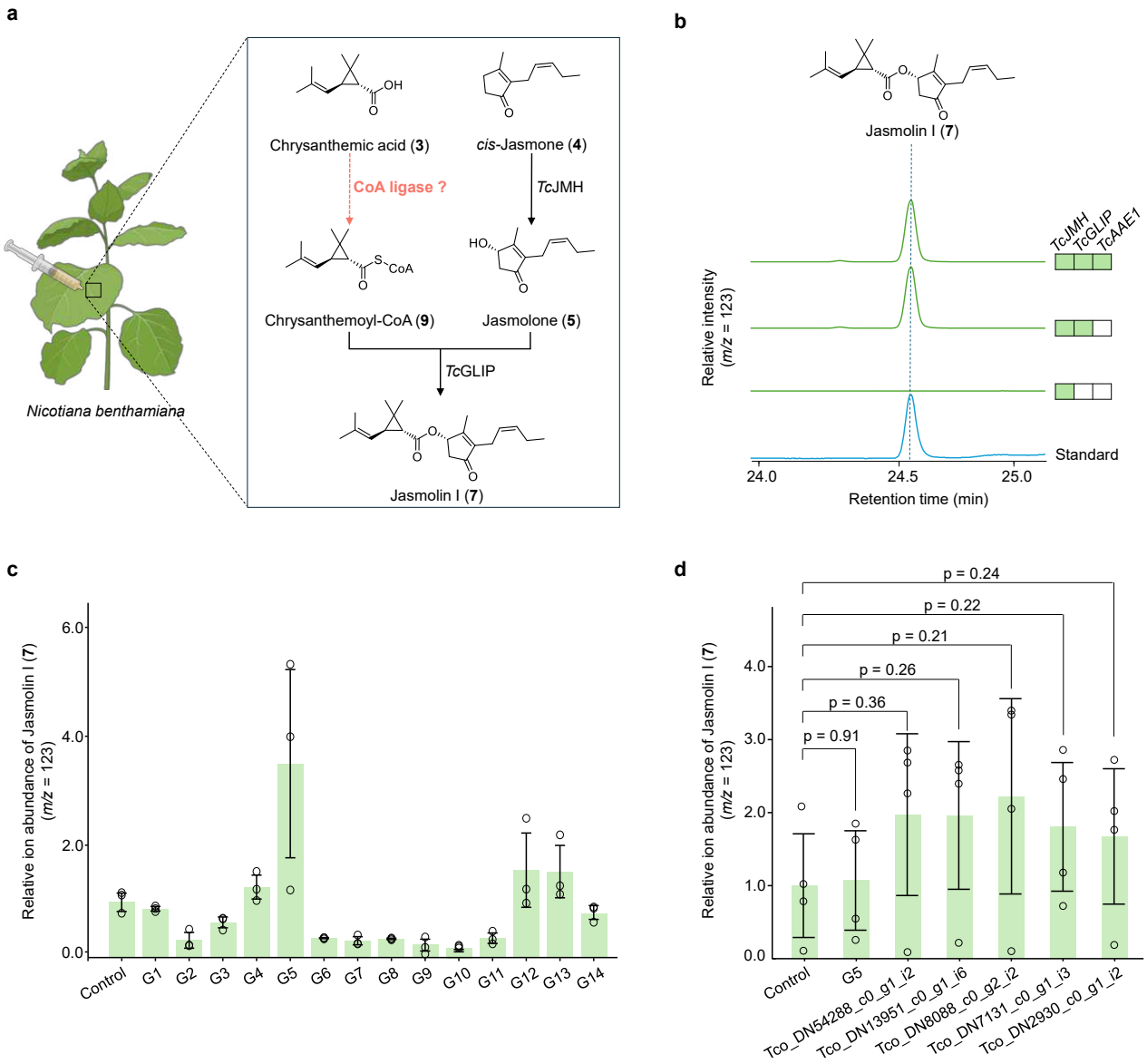


Fig. 2. Functional evaluation of candidate CoA ligases in *N. benthamiana*. **a**. Schematic representation of the functional assay based on transient reconstruction of the proposed jasmolin I (**7**) biosynthetic pathway in *N. benthamiana*. The red dashed arrow indicates the proposed CoA-dependent activation step being evaluated. **b**. GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from tobacco leaves transiently expressing the indicated combinations of *TcJMH*, *TcGLIP* and *TcAAE1*. Leaves were supplemented with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**). The black dashed line indicates the retention time of jasmolin I (**7**). Open squares indicate the absence of the corresponding enzyme. **c**. Relative abundance of jasmolin I (**7**) in tobacco leaves co-expressing *TcJMH*, *TcGLIP* and pools of candidate CoA ligases (G1-G14), determined by GC-EI-MS analysis. Leaves were supplied with 100 mg L^{-1} of chrysanthemic acid (**3**) and 300 mg L^{-1} of *cis*-Jasmone (**4**). Control samples express *TcJMH* and *TcGLIP* without an additional CoA ligase. Jasmolin I (**7**) abundance was normalized to the control, which was set to 1. Data are presented as the mean ± SD of three (n = 3) independent biological replicates. **d**. Relative abundance of jasmolin I (**7**) in leaves co-expressing *TcJMH*, *TcGLIP* and individual CoA ligase candidates from group G5, determined by

GC-EI-MS analysis. Leaves were supplied with 100 mg L^{-1} of chrysanthemic acid (**3**) and 300 mg L^{-1} of *cis*-Jasmone (**4**). Control samples express *TcJMH* and *TcGLIP* without an additional CoA ligase. Jasmolin I (**7**) abundance was normalized to the control, which was set to 1. Data are presented as the mean ± SD of four (n = 4) independent biological replicates. One-way analysis of variance (ANOVA) indicated no statistically significant differences among the groups in panel d.

Systematic identification of candidate CoA ligases for chrysanthemic acid activation

To identify additional CoA ligases potentially involved in pyrethrin biosynthesis, we assembled publicly available RNA-seq datasets from the pyrethrin-producing species *T. cinerariifolium*[22] and *T. coccineum*[23] (**Supplementary Table 1**). A systematic search identified 72 non-redundant candidate CoA-ligases and related acyl-activating enzymes, in addition to *Tc*AAE1. Phylogenetic analysis of these candidates, together with *Tc*AAE1 and three characterized acyl-activating enzymes from *Arabidopsis thaliana* (*At*4CL1, *At*4CL3, and *At*AAE20)[24] resolved eleven major clades (**Supplementary Fig. 2**). Clade 1 included *Tc*AAE1 and *At*AAE20, while clade 2 contained candidates closely related to *At*4CL1 and *At*4CL3. The remaining candidates were distributed among nine additional clades.

To prioritize candidates for functional characterization, we analysed their expression correlations with known pyrethrin biosynthetic genes. Five candidates showed strong positive correlations with *Tc*GLIP (Pearson's $r > 0.8$), nine showed moderate positive correlations ($0.5 < r < 0.8$), and 26 showed weaker positive correlations ($0 < r < 0.5$), whereas 32 candidates were negatively correlated (**Supplementary Fig. 3**). These correlations were used to prioritize candidates for functional characterization. Given the high sequence similarity between the two *Tanacetum* species and the greater availability of *T. coccineum* material, the corresponding coding sequences were amplified from *T. coccineum* flower-head cDNA for subsequent experimental analysis.

To functionally characterize the 72 candidates, we employed the transient expression assay previously established for *Tc*AAE1 in *N. benthamiana*. Candidate enzymes were co-expressed with *TcJMH* and

TcGLIP and tobacco leaves were supplemented with chrysanthemic acid (**3**) and *cis*-jasmone (**4**). Although jasmolin I (**7**) was produced at low levels in the absence of an introduced CoA ligase, we reasoned that expression of a functional *Tanacetum* enzyme would enhance its accumulation above this background. The 72 candidates were divided into 14 groups (G1-G14), each containing five or six enzymes, and prioritized according to their expression correlations with *Tc*GLIP (**Supplementary Fig. 3**). Candidates showing the strongest positive correlations were assigned to the first groups, followed by those with weaker correlations.

Following transient co-expression and GC-MS analysis, jasmolin I (**7**) abundance varied substantially among biological replicates and candidate groups (**Fig. 2c**). Group G5 showed the largest numerical increase in jasmolin I (**7**) accumulation and was selected for further investigation. To determine whether pooled expression had masked the activity of an individual candidate, the enzymes in G5 were co-expressed separately with *TcJMH* and *TcGLIP*. However, none significantly increased jasmolin I (**7**) accumulation above the control (**Fig. 2d**).

To further investigate the candidate CoA ligases in a complementary heterologous system, we established a functional assay in *S. cerevisiae.* We first reconstructed the rethrolone branch to provide the alcohol substrates required for pyrethrin formation. Using strain EGY48[25] as the base, we genomically integrated two codon-optimized copies each of *TcJMH* and *TcPYS*, together with one copy of the *T. cinerariifolium* cytochrome P450 reductase gene (*TcCPR*), generating strain PY1. Following supplementation with 300 mg L^{-1} *cis*-jasmone (**4**), GC-MS analysis revealed the production of jasmolone (**5**), formed through *Tc*JMH-catalysed hydroxylation of *cis*-jasmone, and pyrethrolone (**6**), produced by subsequent *Tc*PYS-mediated oxidation of jasmolone (**Supplementary Fig. 4**). Pyrethrolone (**6**) production was confirmed by comparison with an authentic standard.

We next introduced an episomal plasmid encoding a *Tc*GLIP variant lacking its native 27-amino-acid N-terminal signal peptide and fused to maltose-binding protein (MBP-*Δ*27-*Tc*GLIP)[26,27], into strain PY1. Each of the 73 candidate CoA ligase genes, including *TcAAE1*, was subsequently expressed

individually from a second episomal plasmid, generating yeast strains co-expressing *MBP-Δ27-TcGLIP* and a single candidate CoA ligase. Following supplementation with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**), GC-MS analysis detected no pyrethrin production in any of the resulting strains (**Supplementary Figs. 5 and 6**). As jasmolone (**5**) and pyrethrolone (**6**) could be detected in these samples, confirming that the precursors were available, the lack of pyrethrin production indicated either insufficient activation of chrysanthemic acid by the candidate CoA ligases or poor activity of MBP-*Δ*27-*Tc*GLIP in this system.

To assess the ability of the candidate enzymes to activate chrysanthemic acid independently of *Tc*GLIP, we performed *in vitro* assays using crude protein extracts from EGY48 yeast cells individually expressing *TcAAE1* or one of the 72 additional candidate CoA ligases from episomal plasmids, aiming to detect synthesis of chrysanthemoyl-CoA (**9**). The extracts were incubated with chrysanthemic acid (**3**), CoA, ATP, and $MgCl_2$, and the reaction products were analysed by HPLC-qTOF-MS in negative-ion mode. No signal corresponding to chrysanthemoyl-CoA (**9**) was detected at the expected *m/z* of 916.2124 ± 0.01 ($[M-H]^-$; $C_{31}H_{49}N_7O_{17}P_3S^-$; **Supplementary Fig. 7**). As a positive control, crude extracts from yeast expressing the benzoate-CoA ligase from *Magnetospirillum sp*. *TS-6* (*Ms*BCL)[28] produced detectable benzoyl-CoA from benzoic acid and CoA under comparable assay conditions, confirming that the assay could detect CoA thioester formation (**Supplementary Fig. 8**). These results suggest that none of the tested CoA ligase candidates produced detectable amounts of compound chrysanthemoyl CoA (**9**) under the conditions employed.

Overall, functional evaluation of *Tc*AAE1 and 72 additional candidate CoA ligases using transient expression in *N. benthamiana*, heterologous expression in *S. cerevisiae*, and *in vitro* assays provided no evidence supporting the proposed CoA-dependent activation of chrysanthemic acid.

Identification of an alternative route for pyrethrin biosynthesis

These results prompted us to investigate alternative mechanisms for generating the acyl donor required for pyrethrin biosynthesis. In an increasing number of pathways of plant specialized metabolism, UDP-glucosyltransferases (UGTs) generate glucose esters of carboxylic acids, which subsequently serve as acyl donors for serine carboxypeptidase-like (SCPL) acyltransferases[29]. Therefore, we hypothesized that pyrethrin biosynthesis might involve a glucose ester of chrysanthemic acid, rather than chrysanthemoyl-CoA (**9**), as the acyl donor for the final esterification reaction.

To test this hypothesis, we investigated whether *Tanacetum* encodes a UGT capable of converting chrysanthemic acid (**3**) to chrysanthemoyl-β-D-glucose (**10**) (**Fig. 3a**). We identified six candidate UGT genes through sequence homology searches and amplified their coding sequences from *T. coccineum* flower-head cDNA (**Supplementary Fig. 9**). The corresponding genes were cloned into pET vectors and expressed in *Escherichia coli*. Crude protein extracts were incubated with chrysanthemic acid (**3**) and UDP-glucose, and the reaction products were extracted with methyl *tert*-butyl ether[30] and analysed by HPLC-qTOF-MS. One candidate, hereafter designated *Tco*UGT1, produced a metabolite that co-eluted with an authentic standard of chrysanthemoyl-D-glucose (**10**; **Fig. 3b**). The enzymatic product and authentic standard exhibited matching mass spectra, characterized by a prominent formate adduct ($[M+HCOO]^-$) at *m/z* 375.1661 and a deprotonated molecular ion ($[M-H]^-$) at *m/z* 329.1606 (**Supplementary Fig. 10**). Thus, based on the established stereochemical inversion mechanism of plant UGTs, the enzymatic product was assigned as chrysanthemoyl-β-D-glucose (**10**). No corresponding product was detected in reactions containing extracts from cells carrying the empty vector (**Fig. 3b**).

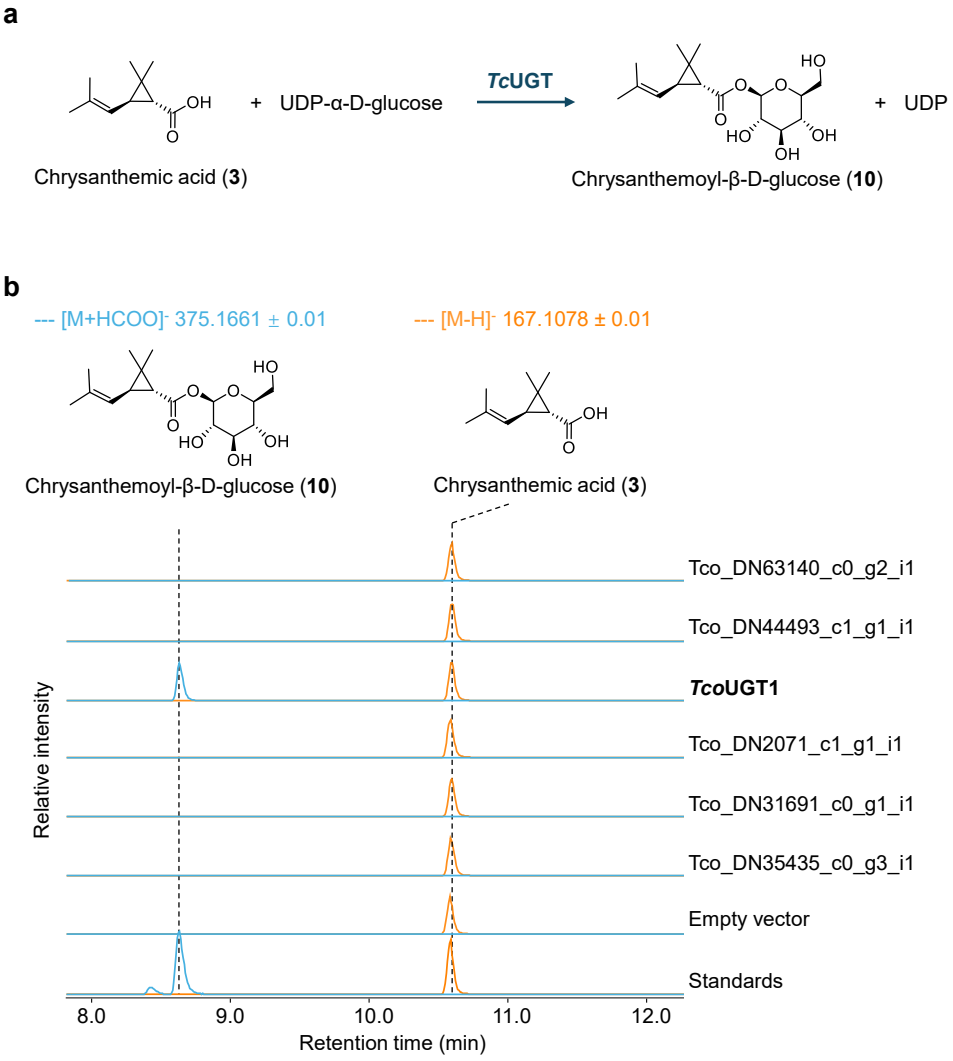


Fig. 3. Identification of *Tco*UGT1 as a chrysanthemic acid glucosyltransferase. **a**. Schematic representation of the *Tco*UGT1-catalysed conversion of chrysanthemic acid (**3**) to chrysanthemoyl-β-D-glucose (**10**) using UDP-α-D-glucose as the glucosyl donor. **b**. UPLC-qTOF-MS extracted ion chromatograms (EICs; negative-ion mode) of *tert*-butyl ether (MTBE) extracts from *in vitro* enzymatic assays of six UGT candidates, showing chrysanthemic acid (**3**; $[M-H]^-$, $m/z = 167.1078 \pm 0.01$) and chrysanthemoyl-β-D-glucose (**10**; $[M+HCOO]^-$, $m/z = 375.1661 \pm 0.01$). *Tco*UGT1 produced a metabolite co-eluting with the synthetic reference standard, whereas no corresponding product was detected in the empty-vector control.

Identification of the esterase partner of *Tco*UGT1

The identification of *Tco*UGT1 as a chrysanthemic acid glucosyltransferase suggested that chrysanthemoyl-β-D-glucose (**10**) could serve as an acyl donor in pyrethrin biosynthesis. As SCPL acyltransferases utilize glucose esters as acyl donors in several pathways of plant specialized metabolism, we investigated whether an SCPL enzyme catalyses the final esterification reaction.

Analysis of the *Tanacetum* transcriptomes identified nine candidate *SCPL* genes co-expressing with known pyrethrin biosynthetic genes (**Supplementary Fig. 11**). Their corresponding coding sequences were amplified from *T. coccineum* flower-head cDNA.

To evaluate their activity, we transiently co-expressed the nine candidates in three groups of three in *N. benthamiana*, together with *TcoUGT1* and *TcJMH* (**Fig. 4a**). Following supplementation with chrysanthemic acid (**3**) and *cis*-jasmone (**4**), GC-EI-MS analysis did not reveal detectable levels of jasmolin I (**7**) formation in leaves expressing any of the SCPL groups (**Fig. 4b**). By contrast, co-expression of *TcGLIP* with *TcoUGT1* and *TcJMH* under the same conditions resulted in detectable jasmolin I (**7**) production (**Fig. 4b**). These findings did not support a role for the tested SCPL candidates in pyrethrin formation.

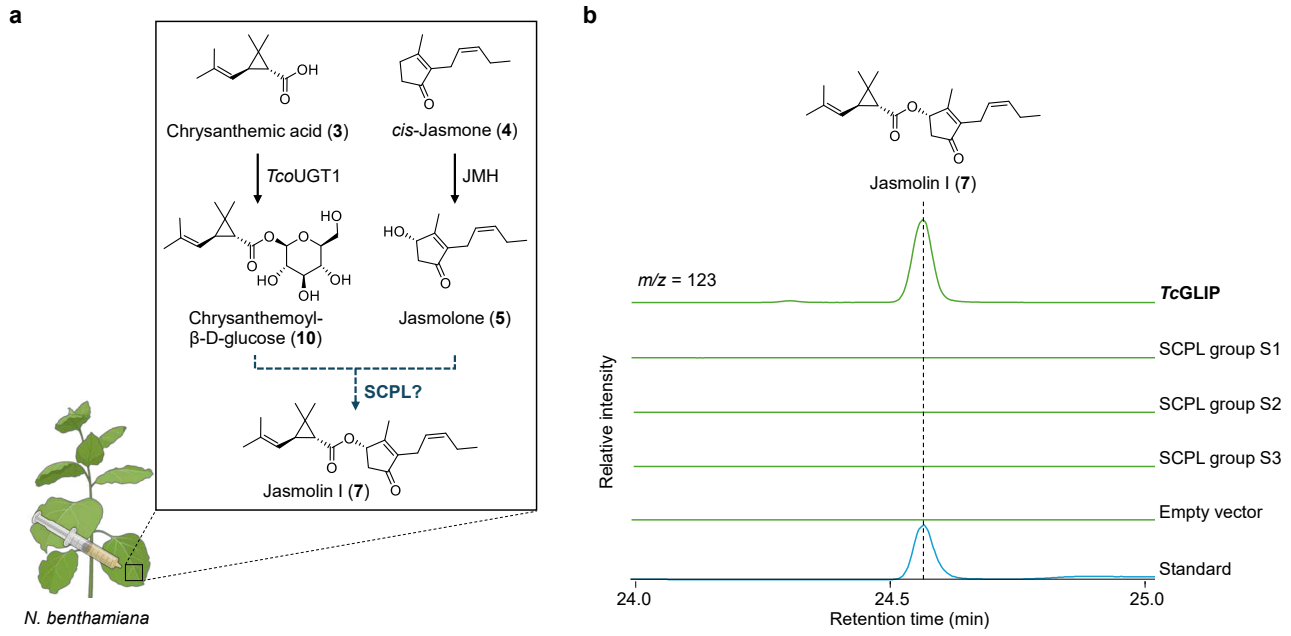


Fig. 4. Evaluation of candidate SCPL acyltransferases in *N. benthamiana*. **a**. Schematic representation of the proposed UGT–SCPL route to jasmolin I (**7**) formation, in which chrysanthemoyl-β-D-glucose (**10**) serves as an acyl donor for SCPL-mediated esterification with jasmolone (**5**). The teal dashed arrow indicates the proposed SCPL-catalysed reaction. **b**. GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from tobacco leaves transiently expressing *TcJMH* and *TcoUGT1*, together with either *TcGLIP* or one of three pools of candidate SCPL acyltransferases (S1-S3), each containing three candidates. Leaves expressing *TcJMH* and *TcoUGT1* together with an empty vector served as the negative control. All leaves were supplemented with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**). Jasmolin I (**7**) was

detected in samples expressing *TcGLIP*, but not in those expressing the SCPL candidate pools or the empty-vector control. An authentic jasmolin I standard is shown for comparison.

*Tc*GLIP forms pyrethrin esters from a glucose ester donor

These findings raised the possibility that *Tc*GLIP catalyses acyl transfer from chrysanthemoyl-β-D-glucose (**10**), rather than chrysanthemoyl-CoA (**9**), to rethrolone alcohols. To investigate this hypothesis, we established an *in vitro* assay combining *Tco*UGT1 and *Tc*GLIP (**Fig. 5a**). Codon-optimized *TcoUGT1* and *MBP-Δ27-TcGLIP* were expressed in *E. coli* BL21. Crude protein extracts containing *Tco*UGT1 were first incubated with chrysanthemic acid (**3**) and UDP-glucose to generate chrysanthemoyl-β-D-glucose (**10**). Purified MBP-*Δ*27-*Tc*GLIP and pyrethrolone (**6**) were subsequently added, and the reaction products were analysed by GC-MS. Pyrethrin I (**8**) was detected in reactions containing both enzymes but was absent from control reactions lacking *Tc*oUGT1 (**Fig. 5b, Supplementary Fig. 12**).

To establish whether the glucose ester could serve directly as an acyl donor, we incubated purified MBP-*Δ*27-*Tc*GLIP with chrysanthemoyl-β-D-glucose (**10**) and pyrethrolone (**6**). GCMS analysis confirmed the synthesis of pyrethrin I (**8**). However, no product was detected when chrysanthemoyl-β-D-glucose was replaced with chrysanthemoyl-CoA (**9**) under the same conditions (**Fig. 5b**). Together, these results demonstrate that *Tc*GLIP catalyses acyl transfer from chrysanthemoyl-β-D-glucose to rethrolone alcohols, establishing a glucose ester-dependent route in pyrethrin biosynthesis. To our knowledge, this represents the first demonstration of a GDSL acyltransferase using a glucose ester as an acyl donor, expanding the known substrate repertoire of this enzyme family.

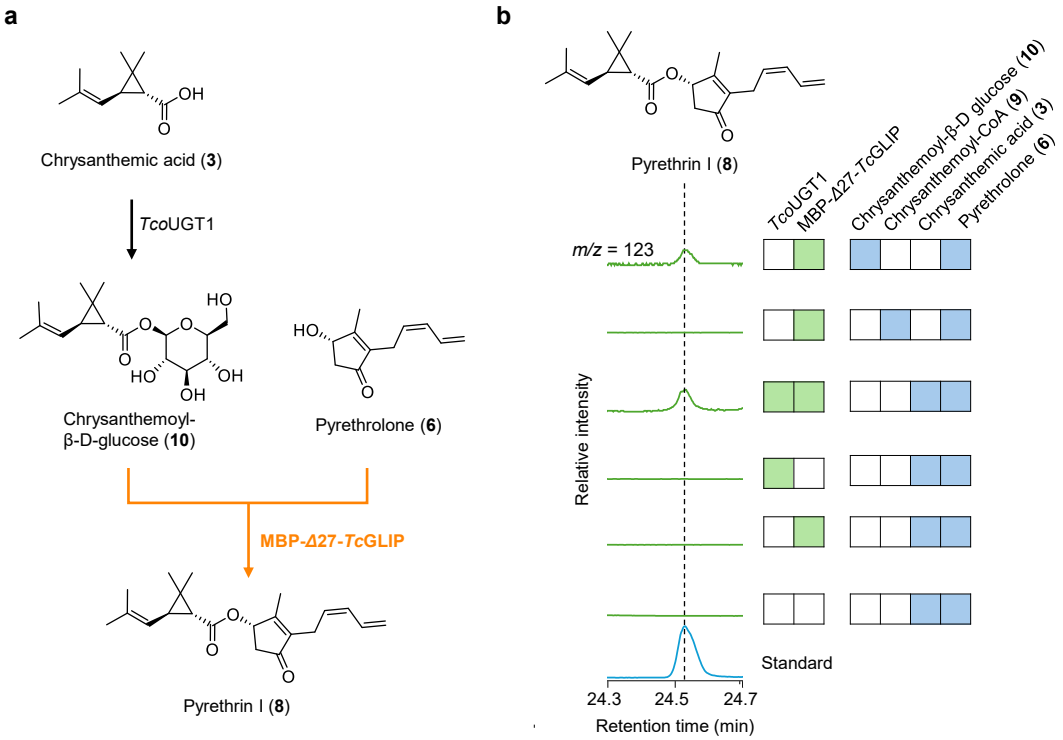


Fig. 5. *Tco*UGT1 and *Tc*GLIP catalyse pyrethrin I synthesis *in vitro*. **a**. Schematic representation of the glucose ester-dependent pathway. *Tco*UGT1 converts chrysanthemic acid (**3**) to chrysanthemoyl-β-D-glucose (**10**), which serves as an acyl donor for *Tc*GLIP-mediated esterification with pyrethrolone (**6**) to produce pyrethrin I (**8**). **b**. GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from *in vitro* reactions containing the indicated combinations of enzymes and substrates. Filled green and blue squares indicate the presence of enzymes and substrates, respectively, whereas open squares indicate their absence. Pyrethrin I (**8**) formation was detected in reactions containing *Tc*GLIP and chrysanthemoyl-β-D-glucose (**10**), but not when the glucose ester was replaced with chrysanthemoyl-CoA (**9**) under the conditions tested. An authentic pyrethrin I (**8**) standard is shown for comparison.

*Tco*UGT1 and *Tc*GLIP catalyze pyrethrin biosynthesis in yeast

To further validate the identified biosynthetic route, we reconstructed pyrethrin biosynthesis in an orthogonal microbial host, *S. cerevisiae* (**Fig. 6a**). To this end, the *TcoUGT1* gene was introduced into the genome of strain PY1, generating strain PY2. Following supplementation with 100 mg L^{-1} chrysanthemic acid (**3**), UPLC-qTOF-MS analysis revealed accumulation of chrysanthemoyl-β-D-glucose (**10**) in PY2 (**Fig. 6b**; **Supplementary Fig. 13**), whereas no corresponding product was detected in the parental strain, PY1, lacking *Tco*UGT1 (**Fig. 6b**). These results demonstrate that

*Tco*UGT1 enables chrysanthemoyl-β-D-glucose (**10**) formation in yeast using the endogenous UDP-α-D-glucose pool.

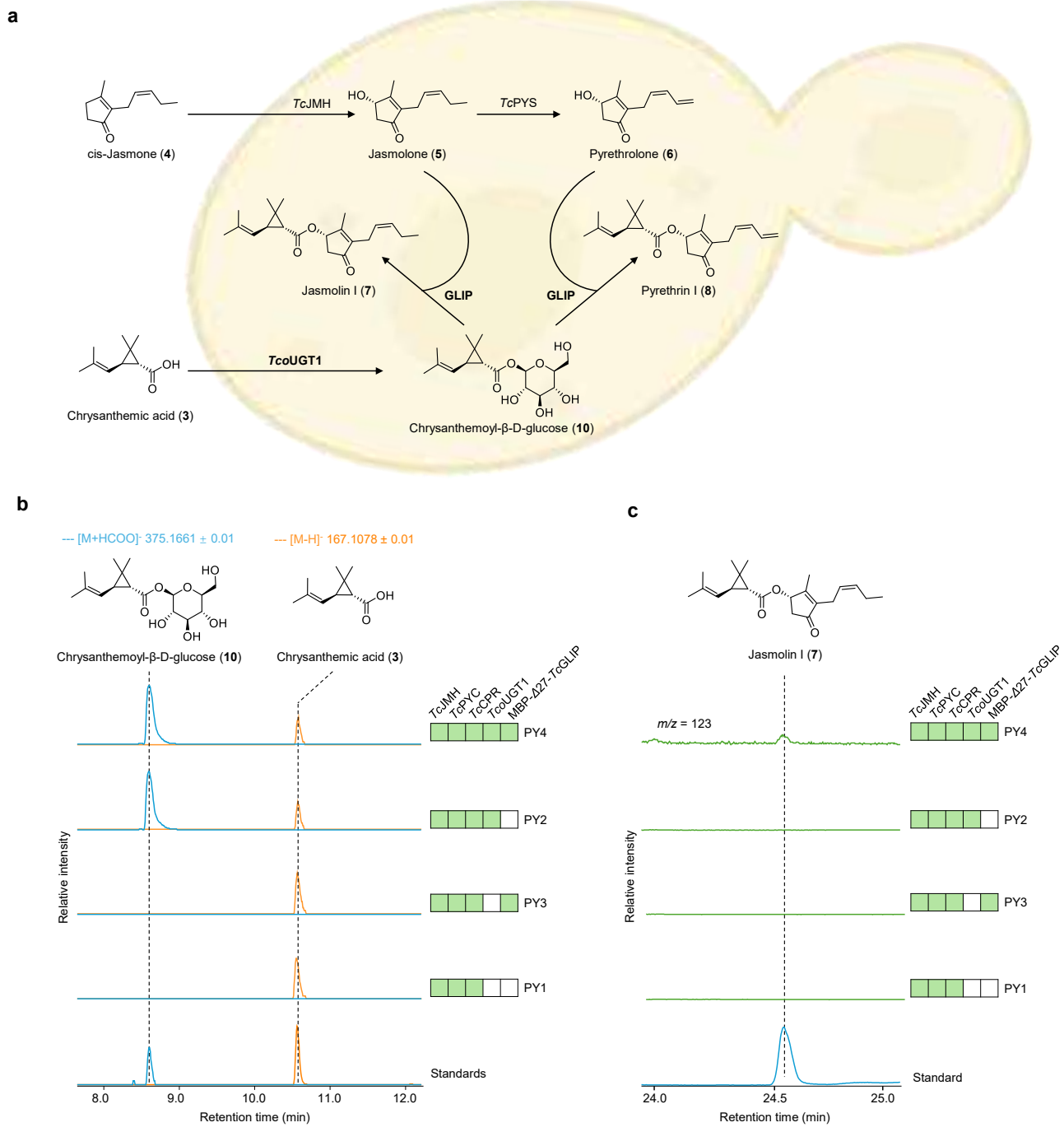


Fig. 6. Heterologous reconstruction of jasmolin I biosynthesis in yeast. a. Schematic representation of the reconstructed pathway in *S. cerevisiae*. Genomically integrated *TcJMH*, *TcPYS*, *TcoUGT1,* and an expression cassette encoding MBP-*Δ*27-*Tc*GLIP enable jasmolin I (**7**) formation from supplied chrysanthemic acid (**3**) and *cis*-jasmone (**4**), via chrysanthemoyl-β-D-glucose (**10**). **b.** The parental strain PY1, carries genomically integrated *TcJMH*, *TcPYS*, and *TcCPR.* Strains PY2, PY3 and PY4 were derived from PY1 by integration of *TcoUGT1,* or *MBP-Δ27-TcGLIP,* or both, respectively. UPLC-qTOF-MS extracted ion chromatograms (EICs;

negative-ion mode) of methanolic extracts from yeast cultures supplemented with 100 mg L^{-1} chrysanthemic acid (**3**), showing chrysanthemic acid ($[M-H]^-$, *m/z* = 167.1078 ± 0.01) and chrysanthemoyl-β-D-glucose ($[M+HCOO]^-$, *m/z* = 375.1661 ± 0.01). Chrysanthemoyl-β-D-glucose (**10**) was detected in strains PY2 and PY4. Authentic chrysanthemic acid (**3**) and chrysanthemoyl-β-D-glucose (**10**) standards are shown for comparison. **c.** GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from yeast strains PY1-PY4 supplemented with 100 mg L^{-1} of chrysanthemic acid (**3**) and 300 mg L^{-1} of *cis*-jasmone (**4**), showing jasmolin I (**7**) formation in strain PY4. An authentic jasmolin I (**7**) standard is shown for comparison.

We subsequently integrated an expression cassette encoding *MBP-Δ27-TcGLIP* into the *XII-5* locus of strain PY1, either alone (strain PY3) or together with *TcoUGT1* (strain PY4). Following supplementation with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1}*cis*-jasmone (**4**), GC-MS analysis demonstrated the production of jasmolin I (**7**) in strain PY4, confirmed by comparison of its retention time and mass spectrum with those of an authentic standard (**Fig. 6c**; **Supplementary Fig. 14**). By contrast, no jasmolin I (**7**) was detected in strains PY2 or PY3, expressing *TcoUGT1* or *MBP-Δ27-TcGLIP* individually, respectively (**Fig. 6c**). Together with the *in vitro* biochemical evidence, these findings establish that *Tco*UGT1 and *Tc*GLIP enable pyrethrin formation through sequential glucose ester formation and acyl transfer, supporting the proposed glucose ester-dependent route to pyrethrin biosynthesis (**Fig. 6a**).

Structural basis of glucose ester recognition by *Tc*GLIP

*Tc*GLIP is the first example of an enzyme of the GDSL family utilizing glucose-activated substrates. To investigate the structural basis for glucose ester recognition by *Tc*GLIP, molecular docking simulations using OpenFold3[31-33] followed by restrained energy minimization were conducted to evaluate the binding mode of chrysanthemic acid-β-D-glucose (**10**) within the catalytic pocket. The resulting structural model predicts a well-defined substrate binding mode, with the glucose moiety being stabilized by interactions with the side chains of Asn168, Gln171, Tyr297, and the backbone carbonyl of Asn55 (**Fig. 7**).

The refined model positioned the chrysanthemoyl carbonyl within the catalytic pocket, with a distance of 2.92 Å between the carbonyl carbon and the catalytic nucleophile Ser40. In this position, the carbonyl oxygen of the substrate is tightly anchored via a network of hydrogen bonds. Specifically, strong coordination is observed with the sidechain of the catalytic nucleophile, Ser40, and the side-chain amide nitrogen of the highly conserved Asn168 residue (**Fig. 7**). This dual coordination mirrors the characteristic electrophilic activation mechanism typical of GDSL family enzymes, effectively polarizing the carbonyl bond of chrysanthemic acid-β-D-glucose (**10**) to elevate the electrophilicity of the adjacent carbonyl carbon. The side-chain hydroxyl oxygen of Ser40 is well-poised for nucleophilic attack, upon initiation by the conserved Ser40-His321-Asp318 charge-relay triad. This structural analysis confirms that the active site of *Tc*GLIP has the correct architecture of residues to bind the glucose moiety of chrysanthemic acid-β-D-glucose (**10**) and position its ester carbonyl group in the proximity of its catalytic machinery.

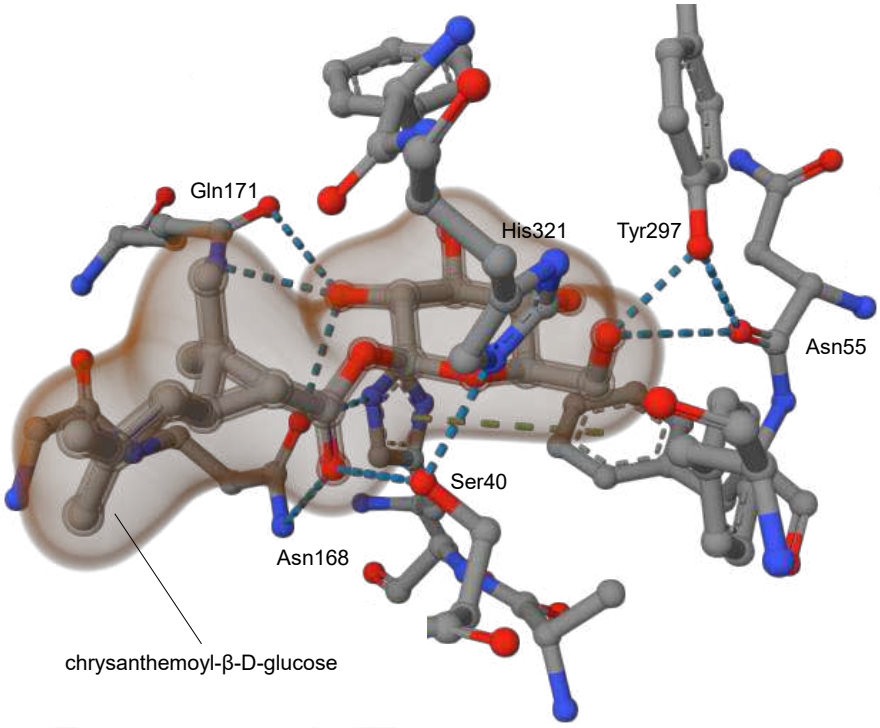


Fig. 7. Predicted interactions of chrysanthemoyl-β-D-glucose (10) with *Tc*GLIP. Structural model of *Tc*GLIP in complex with chrysanthemoyl-β-D-glucose (**10**), generated using OpenFold3 and subjected to restrained energy minimization with OpenMM. The glucose moiety is predicted to be involved in hydrogen-bonding interactions with the side chains of Asn168, Gln171, Tyr297, and the backbone carbonyl of Asn55. The substrate is positioned within the catalytic pocket in proximity to the Ser40-His321-Asp318 catalytic triad,

with a distance of 2.92 Å between the catalytic Ser40 Oγ and the substrate carbonyl carbon. Residue numbering corresponds to full-length *Tc*GLIP. Graphic created with Neurosnap Viewer (www.neurosnap.ai).

DISCUSSION

Here, we identify a glucose ester-dependent route for the final steps of pyrethrin biosynthesis. We show that the *Tanacetum* UDP-glucosyltransferase *Tco*UGT1 converts chrysanthemic acid to chrysanthemoyl-β-D-glucose, which serves as an acyl donor for *Tc*GLIP in the formation of pyrethrin esters. This route was validated through biochemical assays and heterologous reconstruction in *S. cerevisiae*, enabling jasmolin I production from supplied precursors. These findings provide an alternative to the previously proposed CoA-dependent activation of chrysanthemic acid and reveal an unexpected acyl-donor specificity within the GDSL enzyme family.

The proposal that pyrethrin biosynthesis involves a CoA-activated monoterpenoid acid originated from the identification of *Tc*GLIP as a GDSL acyltransferase capable of transferring the chrysanthemoyl group from synthetic chrysanthemoyl-CoA to pyrethrolone in *in vitro* assays[20]. Subsequent studies demonstrated that *Tc*GLIP also accepts truncated thioesters, including chrysanthemoyl-*N*-acetylcysteamine and chrysanthemoyl-pantetheine, indicating that the complete CoA moiety is dispensable for catalysis[34]. However, all donors tested to date have been thioesters, and the form in which chrysanthemic acid is presented to *Tc*GLIP has remained unresolved.

Following the proposal that chrysanthemoyl-CoA is the acyl donor of *Tc*GLIP, the *T. cinerariifolium* acyl-activating enzyme *Tc*AAE1 was proposed to catalyse its formation from chrysanthemic acid[21]. Although this activity was initially reported based on *in vitro* assays, subsequent unpublished experiments by H. Xu and E. Pichersky, cited in a review[5], failed to demonstrate *Tc*AAE1-mediated CoA activation of chrysanthemic acid, either *in vitro* or following transient expression in *N. benthamiana*. Consistent with these observations, we found that *TcAAE1* expression did not enhance jasmolin I formation in *N. benthamiana*, prompting us to investigate alternative CoA ligases potentially involved in chrysanthemic acid activation.

Transcriptome analysis identified 72 additional candidate CoA ligases from *T. cinerariifolium* and *T. coccineum*. Functional evaluation of these candidates in *N. benthamiana*, *S. cerevisiae* and *in vitro* assays provided no evidence supporting chrysanthemoyl-CoA formation or CoA-dependent pyrethrin biosynthesis. Furthermore, jasmolin I was produced in *N. benthamiana* even in the absence of an introduced *Tanacetum* CoA ligase, suggesting that endogenous tobacco metabolism could provide an alternative acyl donor for *Tc*GLIP. Although these negative results do not exclude CoA-dependent activation under other conditions, the consistent absence of supporting evidence across the three experimental approaches prompted us to reconsider whether chrysanthemic acid undergoes CoA-dependent activation during pyrethrin biosynthesis.

An alternative mechanism for activating carboxylic acids in plant specialized metabolism involves the formation of 1-*O*-acyl glucose esters by UDP-glucosyltransferases. These compounds serve as acyl donors for SCPL acyltransferases in several biosynthetic pathways[29]. Notably, previous studies have demonstrated extensive glycosylation of chrysanthemic acid in heterologous plant hosts. During partial reconstruction of the pyrethrin pathway in tomato, approximately 62% of the chrysanthemic acid produced accumulated in glycosylated forms[35], while reconstruction of the pyrethric acid pathway in *N. benthamiana* gave rise to glycosylated forms of several intermediates, including chrysanthemic acid[17]. Although these conjugates have generally been regarded as products of host detoxification rather than pathway intermediates, their formation demonstrates that chrysanthemic acid is recognized by plant glucosyltransferases and raises the possibility that its glycosylation may serve a biosynthetic function.

Here, we identify a *Tanacetum* UDP-glucosyltransferase, *Tco*UGT1, that converts chrysanthemic acid to chrysanthemoyl-β-D-glucose, providing a potential acyl donor for pyrethrin biosynthesis. Although glucose esters are established substrates for SCPL acyltransferases, none of the nine *Tanacetum* SCPL candidates tested supported pyrethrin formation. Instead, we demonstrate that *Tc*GLIP catalyses the transfer of the chrysanthemoyl group from chrysanthemoyl-β-D-glucose to

pyrethrolone, yielding pyrethrin I. Notably, under the same assay conditions, *Tc*GLIP did not produce detectable pyrethrin I when chrysanthemoyl-β-D-glucose was replaced with chrysanthemoyl-CoA. These findings contrast with earlier reports of *Tc*GLIP activity with synthetic CoA thioesters[20,34], and support the glucose ester as an alternative acyl donor for *Tc*GLIP. This biosynthetic route was further validated by heterologous reconstruction in *S. cerevisiae*, where co-expression of *Tco*UGT1 and *Tc*GLIP enabled jasmolin I formation from supplied precursors. Together, these results establish a glucose ester-dependent route to pyrethrin formation, revealing a biosynthetic function for chrysanthemic acid glucosylation.

We further support these findings by molecular docking of chrysanthemoyl-β-D-glucose on *Tc*GLIP. The model shows that chrysanthemoyl-β-D-glucose is accommodated in the *Tc*GLIP catalytic pocket with the glucose moiety contacting Asn168, Gln191 and Tyr297 together with the backbone of Asn55. Furthermore, the chrysanthemoyl carbonyl sits close to the catalytic Ser40 and the predicted oxyanion hole in a position suitable for catalysis. Although this computational model remains to be experimentally validated, it supports the expected role of the glucose moiety in determining substrate recognition.

Our findings expand the known range of acyl donors utilized by GDSL acyltransferases. Although these enzymes were initially characterized as esterases and lipases, several members have been shown to catalyse acyl transfer using chemically diverse donors[29]. For example, the tomato enzyme *Sl*CGT transfers a caffeoyl group from chlorogenic acid, an O-linked ester, to glucarate and galactarate[36], and the sweet potato enzyme *Ib*ICS uses chlorogenic acid as donor in the formation of 3,5-dicaffeoylquinic acid[37]. These examples establish that GDSL acyltransferases can utilize esterified acyl donors rather than CoA thioesters. However, to our knowledge, the use of a 1-*O*-acyl-glucose ester, a donor class primarily associated with SCPL acyltransferases, has not previously been demonstrated for a GDSL enzyme. Our findings therefore extend the catalytic repertoire of GDSL

acyltransferases and suggest that glucose ester-dependent acyl transfer may have a broader role in plant specialized metabolism.

Taken together, our results define a route to pyrethrin formation in which chrysanthemic acid is activated by glucosylation rather than CoA conjugation, and the resulting glucose ester serves as an acyl donor for the final esterification step. These findings elucidate the two final steps in pyrethrin biosynthesis and provide a foundation for producing these insecticides in a microbial host.

METHODS

Transcriptome assembly and annotation

Previously published RNA-seq datasets for *T. cinerariifolium*[22] and *T. coccineum*[23]*cinerariifolium*[22] and *T. coccineum*[23] were downloaded from the China National Genomics Data Center (BioProject PRJCA033633) and NCBI (BioProject PRJNA730666), respectively. The *T. cinerariifolium* dataset comprised fifteen libraries from stems and flower heads at developmental stages 1-4, whereas the *T. coccineum* dataset comprised nine libraries from leaves and flower heads at stages 1 and 4, with three biological replicates per sample type in both datasets. The datasets were reassembled and annotated for candidate gene identification and co-expression analysis using the Galaxy Europe platform[38].

Read quality was assessed using FastQC (Version 0.12.1). Adapter sequences and low-quality bases were removed using Trimmomatic (Version 0.39), applying a Phred quality threshold of 20 and excluding reads shorter than 20 bp. Processed reads were reassessed using FastQC, and quality reports were summarized using MultiQC (Version 1.35). The processed reads were assembled *de novo* using Trinity (Version 2.15.1). Transcript abundance was estimated using the Galaxy tool "Align reads and estimate abundance" (Version 2.15.1), and expression matrices were generated using "Build expression matrix" (Version 2.15.1). Low-expression transcripts were removed using "Filter low expression transcripts" (Version 2.15.1), with a threshold of 1 transcript per million (TPM). The *T.*

cinerariifolium transcriptome comprised 184,925 transcripts representing 95,435 unigenes, whereas the *T. coccineum* transcriptome comprised 172,012 transcripts representing 105,377 unigenes (**Supplementary Table 1**). Assembly quality was assessed using expression-weighted N50 (ExN50) analysis using "Compute contig Ex90N50 statistic and Ex90 transcript count" (Version 2.15.1) and BUSCO (Version 5.8.0). Maximum ExN50 values were 1,701 bp at Ex85 for the *T. cinerariifolium* transcriptome and 1,604 bp at Ex84 for the *T. coccineum* (**Supplementary Fig. 15**). Complete BUSCOs accounted for 99.2% and 96.9% of the respective assemblies (**Supplementary Table 2**).

Protein-coding sequences were predicted using default settings of TransDecoder (Version 5.5.0), yielding 89,632 and 88,163 predicted protein sequences for *T. cinerariifolium* and *T. coccineum*, respectively (**Supplementary Table 2**). Predicted proteins were annotated against the NCBI Conserved Domain Database (version 22 Jan 2018) and Pfam-A (version 22 Jan 2018) using NCBI BLAST + rpsblast (Version 2.16.0) in Galaxy Europe. The annotated assemblies and expression matrices were subsequently used for candidate identification and co-expression analysis.

Candidate identification and co-expression analysis

CoA ligase and related acyl-activating enzyme candidates were identified by searching for annotated transcriptomes using the terms "CoA ligase," "AMP-binding enzyme," "acyl-CoA synthetase," and "acyl-activating enzyme". This search initially yielded more than 400 candidate sequences. Duplicate amino acid sequences and incomplete sequences were excluded, resulting in a set of 73 non-redundant candidates selected for functional characterization. Candidate identifiers and sequences are provided in **Supplementary Data 1**. Co-expression between candidate genes and pyrethrin biosynthetic pathway genes was assessed by calculating Pearson correlation coefficients using TMM-normalized transcript counts in R version 4.4.1. Candidates were classified as strongly positively correlated ($r > 0.8$), moderately positively correlated ($0.5 < r \leq 0.8$) or negatively correlated ($r < 0$).

Candidate genes for UGTs and SCPLs were selected based on their Pearson co-expression coefficients with pyrethrin biosynthetic pathway genes and corresponding annotations.

Phylogenetic analysis

Protein sequences of the identified CoA-ligase candidates and selected reference proteins were aligned using ClustalW implemented in MEGA 12[39]. Poorly aligned N- and C-terminal regions containing extensive gaps were manually removed from the alignment; the conserved internal regions were retained. Phylogenetic relationships were inferred using the Maximum Likelihood method based on the JTT model with gamma-distributed rate variation among sites (JTT + G). The reliability of the inferred tree was assessed using Bootstrap method with 1,000 replicates. Gaps and missing data were treated using the partial deletion option with an 80% site coverage cutoff. Branch lengths represent the estimated number of amino acid substitutions per site. The tree was visualized using iTOL[40].

Chemicals and reference compounds

Chrysanthemol (AB79902) was purchased from A2B Chem LLC. *cis*-Jasmone (W319600), (+)-*trans*-chrysanthemic acid (18509), pyrethrin reference extract (CRM42339), adenosine 5′-triphosphate disodium salt hydrate (A26209), UDP-α-D-glucose disodium salt (670120), and coenzyme A trilithium salt (C3019) were purchased from Sigma-Aldrich. (*S*)-(+)-Pyrethrolone (CSCS00016293816) was purchased from ChemSpace LLC, Ukraine. Pyrethrin I (TRC-P991245) was purchased from LCG Standards GmbH, Germany. Q5U Hot Start High-Fidelity DNA Polymerase (M0515S), DpnI (R0176S), AsiSI (R0630S), Nb.BsmI (R0706S) and USER Enzyme (M5505L) were purchased from New England Biolabs. SuperScript IV Reverse Transcriptase (18090050) was purchased from ThermoFisher.

Chemical synthesis of (1*R*,3*R*)-Chrysanthemoyl CoA

(*1R,3R*)-Chrysanthemoyl-CoA was synthesized using the synthesis described by Kikuta and colleagues[20] without modifications.

Chemical synthesis of 1-O-chrysanthemoyl-β-D-glucopyranose (chrysanthemoyl-β-D-glucose)

The strategy for synthesis of chrysanthemic acid glucose ester is outlined in **Supplementary Fig. 16** and involves the manipulation of D-glucose into its phenylthiol-glucoside according to the reported procedure[41]. Phenylthiol-glucoside was chloroacetylated with chloroacetic acid anhydride using *N*,*N*-dimethyl-4-aminopyridine as catalyst[42,43] to produce phenyl 2,3,4,6-tetra-*O*-chloroacetyl-1-thio-β-D-glucopyranoside from which the thiophenyl group was easily removed according to a reported procedure[44] to afford the 2,3,4,6-tetra-*O*-chloroacetyl-D-glucopyranose. The 2,3,4,6-tetra-*O*-chloroacetyl-D-glucopyranose was then used for the esterification of (+)-*trans*-chrysanthemic acid under the conditions of the Steglich esterification (the use of dicyclohexylcarbodiimide (DCC) as a coupling reagent and 4-dimethylaminopyridine (DMAP) as a catalyst) to obtain the chrysanthemic acid protected-glucose ester (D-2,3,4,6-tetra-O-chloroacetylglucopyranosyloxy (+)-*trans*-chrysanthemate {2,3,4,6-tetra-O-chloroacetyl-D-glucopyranose 1-[2,2-dimethyl-3-(2-methyl-1-propen-1-yl) cyclopropanecarboxylate]}). Finally, specific removal of the chloroacetyl-group from the glucose moiety afforded the target molecule 1-O-chrysanthemoyl-β-D-glucopyranose.

Media

Yeast glucose medium: 20 g L^{-1} glucose, 6.7 g L^{-1} yeast nitrogen base without amino acids (US Biological, Y2025), and 1.3 g L^{-1} dropout supplement lacking uracil, tryptophan, histidine, and leucine (Sigma-Aldrich, Y2001). Yeast Galactose/raffinose (Gal-Raf) medium: 40 g L^{-1} galactose, 20 g L^{-1} raffinose, 1.3 g L^{-1} yeast synthetic drop-out medium supplements (without histidine, leucine, tryptophan and uracil), 6.7 g L^{-1} yeast nitrogen base and 0.2 M 2-morpholinoethanesulfonic acid (MES). The pH was adjusted to 6.5 using 4 M KOH. Yeast extract peptone galactose/raffinose (YPGal-Raf) medium: 10 g L^{-1} yeast extract, 20 g L^{-1} Bacto peptone, 40 g L^{-1} galactose, 20 g L^{-1} raffinose and 0.2 M MES. The pH was adjusted to 6.5 using 4 M KOH. Yeast extract peptone (YEP) medium: 10 g L^{-1} yeast extract, 10 g L^{-1} Bacto peptone, and 5 g L^{-1} NaCl. Terrific Broth (TB) medium containing 24 g L^{-1} yeast extract, 20 g L^{-1} tryptone, 4 mL L^{-1} glycerol, 9.4 g L^{-1} K_2HPO_4, and 2.2 g

492 L^{-1} KH_2PO_4. Lysogeny Broth (LB) medium: 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 5 g L^{-1} $NaCl_2$.
493 The pH was adjusted to 7.2 using 4 M NaOH.

494 **Plant material and RNA extraction**

495 *Tanacetum coccineum* plants were purchased from PLANTORAMA A/S (Copenhagen, Denmark).
496 Immediately after purchase, flower heads were classified into developmental stages 1-5: stage 1,
497 closed buds; stage 2, vertical ray florets; stage 3, horizontal ray florets with the first row of disk florets
498 open; stage 4, three rows of disk florets open; and stage 5, all disk florets open[21]. Flower heads were
499 collected for RNA isolation, snap-frozen in liquid nitrogen, and stored at -80 °C until use. At stages
500 3, 4 and 5, disk and ray florets were collected separately. Total RNA was isolated from *T. coccineum*
501 flower heads stage 3 using the Spectrum Plant Total RNA Kit (Sigma-Aldrich, STRN250-1KT).
502 Complementary DNA was synthesized using reverse transcriptase, according to manufacturer's
503 instructions.

504 **Gene cloning and plasmid construction**

505 All genes used in this study are listed in **Supplementary Data 1.** CoA ligase, UGT, and SCPL
506 candidate genes were amplified from *T. coccineum* cDNA using Q5U DNA polymerase. Primers
507 were designed for USER cloning and are listed in **Supplementary Data 2**. Plasmids were assembled
508 through USER cloning[45], utilizing the enzymes AsiSI and Nb.BsmI. All plasmids used in this study
509 are listed in **Supplementary Data 3**. Constructs were propagated in *E. coli* TOP10, and plasmid DNA
510 was isolated using the E.Z.N.A. Plasmid DNA Mini Kit II (Omega Bio-tek, D6945). For episomal
511 expression in yeast, genes were cloned into pUUS, pHUS, pWUS, or pLUS vectors[25] under the
512 control of GAL1 or CCW12 promoter. For transient expression in *N. benthamiana*, genes were cloned
513 into vector pLIFE33n[46] under the control of the M24 promoter[47]. For recombinant protein production
514 in *E. coli* BL21 (DE3), genes were cloned into pET expression vector under the control of the T7
515 promoter. MBP-*Δ*27-*Tc*GLIP was generated by removing the 27-amino-acid signal peptide and
516 fusing MBP to the N-terminus[20,46] under the control of the M24 promoter[47]. For recombinant protein

production in *E. coli* BL21 (DE3), genes were cloned into pET expression vector under the control of the T7 promoter. MBP-*Δ*27-*Tc*GLIP was generated by removing the 27-amino-acid signal peptide and fusing MBP to the N-terminus[20] through the linker GGGSGGGS.

Yeast integration vectors and transformation procedures were based on previously described methods[46]. All assembled plasmid constructs were verified by Sanger sequencing (GENEWIZ, Germany).

Yeast cultivation

Yeast strains carrying expression plasmids were precultured for 16-18 h at 30 °C in 5 mL glucose medium. Required supplements were added according to the selection markers. For induction, precultures were inoculated at an initial OD_{600} of 0.2 into yeast Gal-Raf medium and incubated at 30 °C and 150 rpm for three days.

Transient gene expression in *N. benthamiana*

The *A. tumefaciens* strain AGL1 was used for all agroinfiltration experiments; cultures were grown overnight at 28 °C in 2 mL YEP medium supplemented with 50 mg L^{-1} of carbenicillin, 50 mg L^{-1} of kanamycin and 25 mg L^{-1} of rifampicin. Cultures were harvested by centrifugation at 3,500 × g for 5 min, washed twice with sterile water, and resuspended in water to an OD_{600} of 1.0. Suspensions were infiltrated into the 4-6 weeks-old *N. benthamiana* leaves[48]. Plants were maintained at 25 °C in a greenhouse under a 16 h light/8 h dark photoperiod.

For expressing CoA ligases candidates in tobacco, the 72 CoA ligases were divided into 14 groups (G1-G14), each containing five or six candidates. *Agrobacterium* cultures carrying the candidates within each group were pooled and co-infiltrated into *N. benthamiana* leaves with cultures carrying *TcGLIP* and *TcJMH*. Candidates from G5 were subsequently tested individually, each together with

TcGLIP and *TcJMH*. Control treatments contained *TcGLIP* and *TcJMH*. The nine candidate SCPL acyltransferases were tested using the same approach, divided into three groups of three candidates each. Each pooled group was co-expressed with *TcJMH* and *TcoUGT1*. Control treatments contained *TcJMH* and *TcoUGT1*, together with empty vector as control. The silencing suppressor p19 was included in all agroinfiltration experiments. The final OD_{600} of each individual culture was adjusted to 1.0, and equal total infiltration volumes were used per plant. Chrysanthemic acid (100 mg L^{-1}) and *cis*-jasmone (300 mg L^{-1}) were supplied at the same respective concentrations across treatments by vacuum feeding. For vacuum feeding, leaves were harvested four days after *Agrobacterium* infiltration and each leaf was placed in a separate Petri dish containing 15 mL substrate solution prepared in 0.25% (v/v) methanol. Leaves were subjected to vacuum infiltration (at 0.85 kPa pressure) for 10 min and subsequently incubated at room temperature for 24 h with approximately 5 mL solution remaining in each dish.

Recombinant protein production and crude-extract preparation

Yeast strain EGY48[25] was used for CoA ligases recombinant protein production. Precultures were grown as described above and transferred into 25 mL MES-buffered induction medium (pH 6.5) with initial OD_{600} of 0.2. Cultures were incubated at 30 °C, 150 rpm for two days. Cells were harvested by centrifugation at 3,500 × *g* for 5 min, washed twice with cold sterile water, and resuspended in 1 mL 50 mM Tris-HCl, pH 7.5. Suspensions were transferred into 2 mL tubes containing 0.25 mL 0.5 mm glass beads. Cells were disrupted using ten cycles of 30 s vortexing followed by 1 min cooling on ice. Lysates were centrifuged at 15,000 × *g* for 10 min at 4 °C, and supernatants were collected and stored in -80 °C freezer.

For recombinant UGT production, *E. coli* SoluBL21™ (Genlantis) cells carrying the corresponding expression constructs were grown overnight at 37 °C in 2 mL LB medium supplemented with 100 μg mL^{-1} carbenicillin. Overnight cultures were used to inoculate 50 mL fresh LB medium and grown at 37 °C with shaking at 220 rpm until an OD_{600} of 0.4-0.8 was reached. Protein expression was induced

with 200 μM IPTG, followed by incubation at 18 °C for 24 h with shaking at 200 rpm. Cells were harvested by centrifugation at 4,000 × g for 5 min, washed twice with cold sterile water, and resuspended in 1 mL of 50 mM Tris-HCl, pH 7.5. Cells were disrupted by five 15 s sonication cycles, separated by 45 s cooling intervals on ice. Lysates were clarified by centrifugation at 14,000× g for 20 min at 4 °C, and the supernatants were collected as crude protein extracts and stored at -80°C until further use.

Protein expression and purification of *Tc*GLIP

For recombinant protein production, *Tc*GLIP lacking the N-terminal 27 amino acids of the predicted signal peptide was expressed as an N-terminal His-tagged maltose-binding protein (MBP) fusion (MBP-*Δ*27-*Tc*GLIP) in *E. coli* SoluBL21™ (Genlantis) cells. Protein expression was induced by 0.1 mM IPTG at 18 °C for 15 h. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 5% glycerol, and 10 mM imidazole), and disrupted by sonication. Following clarification by centrifugation, the recombinant protein was purified using Ni-NTA resin (Qiagen). The resin was washed with buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, 5% glycerol, and 30 mM imidazole, and bound proteins were eluted with the same buffer containing 100 mM imidazole. Protein purity and apparent molecular weight were assessed by SDS-PAGE and Western blotting. The resulting MBP-*Δ*27-*Tc*GLIP preparation was stored at -80 °C until use for *in vitro* enzymatic assays.

***In vitro* enzymatic assays**

In vitro validation of CoA ligase candidates were performed in a total volume of 200 μL containing 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 1 mM CoA, 1 mM chrysanthemic acid, 1 mM pyrethrolone, 20 μL crude extract from each CoA ligase candidate, and 50 μL MBP-*Δ*27-*Tc*GLIP crude extract. Reactions were incubated at 25 °C for 5 h. For CoA ligase assays without GLIP, the MBP-*Δ*27-*Tc*GLIP extract and pyrethrolone were omitted, and the volume of each CoA ligase crude extract was increased to 30 μL.

UGT assays were performed in a total volume of 200 µL containing 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 2 mM UDP-glucose, 2 mM chrysanthemic acid, and 150 µL crude protein extract. Reactions were incubated at 30 °C for 3 h. Controls contained an equal volume of crude extract from cells carrying the empty vector. Following incubation, 100 µL of methanol was added to 100 µL reaction-aliquot, to stop the reaction, and analysed by UPLC-qTOF-MS. For sequential UGT-GLIP assays, the final reaction volume was 200 µL, containing 2 mM pyrethrolone, 50 mM Tris-HCl (pH 7.5), 100 ul of UGT reaction mixture, and 40 µL purified MBP-*Δ*27-*Tc*GLIP. Reactions were incubated for a further 8 h at 25 °C. Control received an equal volume of water instead of MBP-*Δ*27-*Tc*GLIP. Where indicated, synthetic chrysanthemic acid glucoside was used at 2 mM in place of the UGT reaction product. Products were extracted with two volumes of ethyl acetate and injected directly for GC-EI-MS analysis.

Metabolite extraction

For *N. benthamiana*, one leaf disk (18 mm in diameter) was collected, frozen in liquid nitrogen, and ground to a fine powder. Samples were extracted with 300 µL ethyl acetate for GC-EI-MS analysis or 300 µL methanol for UPLC-qTOF-MS analysis. The extracts were sonicated in a water bath for 30 min and subsequently centrifuged at 15,000 × g for 2 min. The resulting supernatants were collected for analysis.

For GC-EI-MS analysis of yeast cultures, 10 mL culture was extracted with 5 mL ethyl acetate by vortexing for 2 min. Samples were centrifuged at 3,500 × g for 5 min, and the upper organic phase was collected. For UPLC-qTOF-MS analysis of yeast cultures, 500 µL of culture was transferred to a 2-mL microcentrifuge tube containing 0.25 mL of 0.5 mm glass beads. 1 mL MTBE was added, and the samples were vortexed for 2 min. Following centrifugation at 15,000 × g for 2 min, the upper phase was collected and dried using a vacuum concentrator for 10 min at 25 °C and 1,000 rpm. The dried extracts were subsequently reconstituted in 300 µL of 80% methanol.

For *in vitro* enzymatic assays, samples were extracted with two volumes of ethyl acetate and vortexed for 1 min. The samples were then centrifuged at 15,000 × g for 2 min, and the upper organic phase was collected for GC-EI-MS analysis. For UPLC-qTOF-MS analysis, samples were extracted with two volumes of MTBE using the same extraction procedure described above. All methanol-dissolved extracts were filtered through a 0.22-μm polyvinylidene difluoride filter before UPLC-qTOF-MS analysis.

GC-EI-MS and UPLC-qTOF-MS analysis

Ethyl acetate extracts (1 μL) were analysed by gas chromatography and electron ionization mass spectrometry (GC-EI-MS) using a Shimadzu Nexis GC-2030 gas chromatograph coupled to a GCMS-QP2020 NX mass spectrometer (Shimadzu, Kyoto, Japan). Injections were made in splitless mode at 180 °C. Separation was performed on an Agilent HP-5ms column (30 m × 0.25 mm internal diameter, 0.25 μm film thickness; Agilent Technologies, Santa Clara, CA, USA). Helium was used as the carrier gas at a linear velocity of 30 cm s^{-1}. The oven temperature was initially set to 50 °C for 1 min, increased to 300 °C at 10 °C min^{-1}, and held at 300 °C for 3.5 min (total run time 29.5 min). Data was processed using Shimadzu GCMS Solution version 4.20.

For UPLC-qTOF-MS analysis, 2 μL of methanol-dissolved extract was injected to the Dionex UltiMate® 3000 Quaternary Rapid Separation UHPLC focused system (Thermo Fisher Scientific, Germering, Germany). A Bruker Daltonics Compact qTOF mass spectrometer containing an electrospray ionization (ESI) interface (Bruker Daltonics, Bremen, Germany) was used as MS detector. Mass spectra were acquired in negative mode for tobacco, yeast and *in vitro* reaction assay samples for detecting non-volatile compounds. Other ESI-MS settings were as follows: Capillary voltage, 4000V; End plate offset, -500V; Drying temperature, 220 °C; Nebulizer pressure, 2.0 bar; Drying gas flow, 8 L min^{-1}. Sodium formate solution was injected at the start of each chromatographic run as mass calibrant. Chromatographic separation was performed on a Phenomenex Kinetex XB-C18 column (100 mm × 2.1 mm i.d., 1.7 μm particle size, 100 Å pore size; Phenomenex Inc.,

Torrance, CA, USA) at a flow rate of 0.3 mL min^{-1} and a column temperature of 40 °C. The mobile phases were Milli-Q water (solvent A) and LC-MS grade acetonitrile (HiPerSolv CHROMANORM®; solvent B), both containing 0.05% (v/v) formic acid. The gradient was increased linearly from 0 min, 2% B; 20 min, 100% B; 26 min, 2% B. Data were processed in Compass DataAnalysis 4.20. The HyStar 3.2 software (Bruker Daltonics) was used for data acquisition.

Protein-ligand complex modelling and structural refinement

The structure of *Tc*GLIP in complex with chrysanthemoyl-β-D-glucose was predicted using OpenFold3 through the Neurosnap platform[31-33]. The two highest-ranked models were independently subjected to restrained energy minimization using OpenMM, with the AMBER14 protein force field and GAFF 2.2.20 parameters for the ligand. The complexes were solvated in explicit TIP3P water with 1.0 nm padding and an ionic strength of 0.15 M. During minimization (up to 1,500 iterations), protein heavy atoms more than 8 Å from the ligand and backbone atoms within the binding pocket were positionally restrained using force constants of 1,500 and 700 kJ mol^{-1} nm^{-2}, respectively, while ligand heavy atoms were weakly restrained (40 kJ mol^{-1} nm^{-2}). The resulting structures were evaluated for protein-ligand interactions, hydrogen-bonding geometry and the orientation of the substrate carbonyl relative to the catalytic Ser40 residue.

Statistical analysis

Experiments were performed using three or four independent biological replicates. Data are presented as the mean ± SD. Statistical analyses were performed in Microsoft Excel (Microsoft 365 MSO) using ANOVA one-way analysis. Replicate numbers and tests used for individual experiments are specified in the figure legends.

Software

Chemical structures were drawn using ChemDraw Professional 25.5.0. Bar charts and heatmaps were generated in R version 4.4.1 using ggplot2 (version 4.0.3) and pheatmap (version 1.0.12) packages,

664 respectively, while line charts were generated using Microsoft Excel (Microsoft 365 MSO).
665 Schematic illustrations were prepared using Microsoft PowerPoint (Microsoft 365 MSO).

666

667 **Data Availability**

668 The data that support the findings of this study is available within the main text and its Supplementary
669 Information file and Supplementary Data files. Data is also available from the corresponding authors
670 upon request.

671

672 **Acknowledgments**

673 We thank Dr. Feiyan Liang (University of Copenhagen) for critically reading the manuscript, Dr.
674 Yao-Tao Duan and Dr. Zhen Fan (University of Copenhagen) for assistance with chemical structures
675 and glucose ester LCMS analysis, and Jack Olsen and Dr. Mariela Ramirez (University of
676 Copenhagen) for their assistance in running analytical instruments. This work was financially
677 supported by the Novo Nordisk Foundation grant NNF23OC0083498 (to S.C.K.), a China
678 Scholarship Council PhD scholarship (Grant no. 202206790009 to C.G.), and the European Union's
679 Horizon Europe programme through a Marie Skłodowska-Curie Postdoctoral Fellowship (BioPesTY,
680 grant agreement no. 101153913 to N.P.). Views and opinions expressed are, however, those of the
681 author(s) only and do not necessarily reflect those of the European Union or the European Research
682 Executive Agency (REA). Neither the European Union nor the granting authority can be held
683 responsible for them.

684

685 **Author Contributions Statement**

S.C.K. conceived the study. C.G., N.P., I.P., F.G.-F., and S.C.K. designed research. C.G., N.P., and M.S.M performed experimental work. C.G., N.P., M.S.M., I.P., and S.C.K. wrote the manuscript. All authors read and commented on the final version of the manuscript.

Competing Interests Statement

The authors declare no competing commercial or non-commercial interests.

REFERENCES

1 Matsuda, K. Understanding pyrethrin biosynthesis: toward and beyond natural pesticide overproduction. *Biochem.Soc. Trans.* **52**, 1927-1937 (2024).

2 Matsuo, N. Discovery and development of pyrethroid insecticides. *Proc. Jpn. Acad. B: Phys. Biol. Sci.* **95**, 378-400 (2019).

3 Li, H., Cheng, F., Wei, Y., Lydy, M. J. & You, J. Global occurrence of pyrethroid insecticides in sediment and the associated toxicological effects on benthic invertebrates: An overview. *J. Hazard. Mater*.**324**, 258-271 (2017).

4 Aznar-Alemany, Ò. & Eljarrat, E. Bioavailability and bioaccumulation of pyrethroid insecticides in wildlife and humans. In *Pyrethroid Insecticides* (ed. Eljarrat, E.) *The Handbook of Environmental Chemistry*, vol. 92 (Springer, Cham, 2020).

5 Lybrand, D. B., Xu, H., Last, R. L. & Pichersky, E. How Plants Synthesize Pyrethrins: Safe and Biodegradable Insecticides. *Trends Plant Sci.* **25**, 1240-1251 (2020).

6 Atkinson, B. L., Blackman, A. J. & Faber, H. The Degradation of the Natural Pyrethrins in Crop Storage. *J. Agric. Food Chem.* **52**, 280-287 (2004).

7 Mew, E. J. *et al.* The global burden of fatal self-poisoning with pesticides 2006-15: Systematic review. *J. Affect. Disord.* **219**, 93-104 (2017).

8 Porras-Rivera, G., Górski, K. & Colin, N. Behavioral biomarkers in fishes: A non-lethal approach to assess the effects of chemical pollution on freshwater ecosystems. *Environ. Res.* **260**, 119607 (2024).

9 Musah, B. I. Impact of pesticides on honeybees (*Apis mellifera*) physiology, foraging and colony: An overview. *Environ. Toxicol. Pharmacol* **125**, 105066 (2026).

10 Niño-de-Guzman-Tito, M., Cuno, K. D. A. & Vilca, F. Z. Emerging contaminants in aquatic ecosystems: A critical review of the physiological and bioindicator responses of macroinvertebrates in water quality index assessment. *Sci. Total Environ.* **1049**, 182113 (2026).

11 Hezakiel, H. E., Thampi, M., Rebello, S. & Sheikhmoideen, J. M. Biopesticides: a Green Approach Towards Agricultural Pests. *Appl. Biochem. Biotechnol.* **196**, 5533-5562 (2024).

12 Khatib, S., Sobeh, M., Faraloni, C. & Bouissane, L. *Tanacetum* species: Bridging empirical knowledge, phytochemistry, nutritional value, health benefits and clinical evidence. *Front. Pharmacol.* **14**, 1169629 (2023).

13 Ríos, F. J. A. & Simonsen, H. T. Beyond *Tanacetum cinerariifolium*: Advances in Pyrethrins Biosynthesis and Prospects for Engineered Cell Factories. *Mol. Biotechnol.* **68**, 2659-2672 (2026).

14 Jeran, N. *et al.* Pyrethrin from Dalmatian pyrethrum (*Tanacetum cinerariifolium* (Trevir.) Sch. Bip.): biosynthesis, biological activity, methods of extraction and determination. *Phytochem. Rev.* **20**, 875-905 (2021).

15 Yu, A., Mao, J. & Xu, N. Metabolic Engineering of Yeasts: A Key Cell Factory Platform for Advanced Biomanufacturing. *J. Fungi* **11**, 863 (2025).

16 Chen, R., Yang, S., Zhang, L. & Zhou, Y. J. Advanced Strategies for Production of Natural Products in Yeast. *iScience* **23**, 100879 (2020).

17 Xu, H. *et al.* Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in *Nicotiana benthamiana*. *New Phytol.* **223**, 751-765 (2019).

18 Li, W., Lybrand, D. B., Zhou, F., Last, R. L. & Pichersky, E. Pyrethrin Biosynthesis: The Cytochrome P450 Oxidoreductase CYP82Q3 Converts Jasmolone To Pyrethrolone. *Plant Physiol.* **181**, 934-944 (2019).

19 Li, W., Zhou, F. & Pichersky, E. Jasmone Hydroxylase, a Key Enzyme in the Synthesis of the Alcohol Moiety of Pyrethrin Insecticides. *Plant Physiol.* **177**, 1498-1509 (2018).

20 Kikuta, Y. *et al.* 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 J.* **71**, 183-193 (2012).

21 Yang, T. *Biosynthesis of monoterpene alcohols, derivatives and conjugates in plants : roles in resistance to western flower thrips*, Wageningen University, (2013).

22 Zeng, T., Li, J., Wang, C. & Li, J. Comparative Transcriptomic Analysis of Pyrethrin and EβF Biosynthesis in *Tanacetum cinerariifolium* Stems and Flowers. *Horticulturae* **11**, 201 (2025).

23 Zeng, T. *et al.* Transcriptional Responses and GCMS Analysis for the Biosynthesis of Pyrethrins and Volatile Terpenes in *Tanacetum coccineum*. *Int. J. Mol. Sci.* **22**, 13005 (2021).

24 Srividya, N. *et al.* Biochemical characterization of acyl activating enzymes for side chain moieties of Taxol and its analogs. *J. Biol. Chem.* **295**, 4963-4973 (2020).

25 Dusseaux, S., Wajn, W. T., Liu, Y., Ignea, C. & Kampranis, S. C. Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl. Acad. Sci. USA* **117**, 31789-31799 (2020).

26 Kikuta, Y. *et al.* Requirement of catalytic-triad and related amino acids for the acyltransferase activity of *Tanacetum cinerariifolium* GDSL lipase/esterase TcGLIP for ester-bond formation in pyrethrin biosynthesis. *Biosci. Biotechnol. Biochem.* **77**, 1822-1825 (2013).

27 Sun, P., Tropea, J. E. & Waugh, D. S. Enhancing the solubility of recombinant proteins in *Escherichia coli* by using hexahistidine-tagged maltose-binding protein as a fusion partner. *Methods Mol. Biol.* **705**, 259-274 (2011).

28 Kawaguchi, K., Shinoda, Y., Yurimoto, H., Sakai, Y. & Kato, N. Purification and characterization of benzoate-CoA ligase from *Magnetospirillum sp.* strain TS-6 capable of aerobic and anaerobic degradation of aromatic compounds. *FEMS Microbiol. Lett.* **257**, 208-213 (2006).

29 Mallavergne, A., Hilbert, J. L. & Gagneul, D. Completing the acylation landscape in plants: GDSL enzymes join the BAHD and SCPL acyltransferase families. *J. Exp. Bot.* **77**, 4249-4268 (2026).

30 Xu, H. *et al.* Coexpression Analysis Identifies Two Oxidoreductases Involved in the Biosynthesis of the Monoterpene Acid Moiety of Natural Pyrethrin Insecticides in *Tanacetum cinerariifolium*. *Plant. Physiol.* **176**, 524-537 (2018).

31 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).

32 Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679-682 (2022).

33 Neurosnap Inc. *Neurosnap: An online platform for computational biology and chemistry*, <https://neurosnap.ai> (2022).

34 Sugisaka, Y., Aoyama, S., Kumagai, K., Ihara, M. & Matsuda, K. TcGLIP GDSL Lipase Substrate Specificity Co-determines the Pyrethrin Composition in *Tanacetum cinerariifolium*. *J. Agric. Food. Chem.* **70**, 8645-8652 (2022).

35 Xu, H. *et al.* Production of trans-chrysanthemic acid, the monoterpene acid moiety of natural pyrethrin insecticides, in tomato fruit. *Metab. Eng.* **47**, 271-278 (2018).

36 Teutschbein, J. *et al.* Identification and localization of a lipase-like acyltransferase in phenylpropanoid metabolism of tomato (*Solanum lycopersicum*). *J. Biol. Chem.* **285**, 38374-38381 (2010).

37 Miguel, S. *et al.* A GDSL lipase-like from Ipomoea batatas catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*. *Commun. Biol.* **3**, 673 (2020).

38 Goecks, J., Nekrutenko, A., Taylor, J. & The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).

39 Kumar, S. *et al.* MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing. *Mol. Biol. Evol.* **41**, msae263 (2024).

40 Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **52**, W78-W82 (2024).

41 Motawia, M. S., Olsen, C. E., Enevoldsen, K., Marcussen, J. & Møller, B. L. Chemical synthesis of 6'-alpha-maltosyl-maltotriose, a branched oligosaccharide representing the branch point of starch. *Carbohydr. Res.* **277**, 109-123 (1995).

42 Møller, B. L., Olsen, C. E., Zhao, Y. & Motawie, M. S. A Convenient Route to Large-Scale Chemical Synthesis of p-Hydroxyphenylacetaldehyde Oxime and Its p-β-d-Glucopyranoside: Key Intermediates and Products in Plant Specialized Metabolism. *ACS Omega* **9**, 44312-44320 (2024).

43 Chaudhary, S. K. & Hernandez, O. 4-dimethylaminopyridine: an efficient and selective catalyst for the silylation of alcohols. *Tetrahedron Lett.* **20**, 99-102 (1979).

44 Motawia, M. S., Marcussen, J. & Møller, B. L. A General Method Based on the Use of N-Bromosuccinimide for Removal of the Thiophenyl Group at the Anomeric Position to Generate A Reducing Sugar with the Original Protecting Groups Still Present. *J. Carbohydr. Chem.* **14**, 1279-1294 (1995).

45 Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T. & Halkier, B. A. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* **35**, e55 (2007).

46 Forman, V. *et al.* A gene cluster in *Ginkgo biloba* encodes unique multifunctional cytochrome P450s that initiate ginkgolide biosynthesis. *Nat Commun* **13**, 5143 (2022).

47 Dey, N. & Maiti, I. B. Structure and promoter/leader deletion analysis of mirabilis mosaic virus (MMV) full-length transcript promoter in transgenic plants. *Plant Mol. Biol.* **40**, 771-782 (1999).

48 Liang, F. *et al.* Elucidation of the final steps in Taxol biosynthesis and its biotechnological production. *Nat. Synth.* **4**, 1212-1222 (2025).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData3Plasmids.xlsx
- SupplementaryData2Primers.xlsx
- SupplementaryData1Genessequences.xlsx
- EditorialpolicychecklistSK.pdf
- ReportingSummarySK.pdf
- SupplementaryInformation.pdf
- rsflattened.pdf