SUPPLEMENTARY INFORMATION

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

Guo et al.

TABLE OF CONTENTS

Supplementary Tables 1-3

Supplementary Figures 1-16

SUPPLEMENTARY TABLES

Supplementary Table 1. Sequencing quality of *T. cinerariifolium* and *T. coccineum* transcriptomes

Average Parameter	*T. cinerariifolium* Before trimming	*T. cinerariifolium* After trimming	*T. coccineum* Before trimming	*T. coccineum* After trimming
Reads	22,868,813	22,398,907	23,053,114	22,713,027
Bases (Gbp)	3.39	3.24	3.39	3.26
GC content (%)	42.60	42.40	42.11	42.11
Q30 (%)	97.71	99.70	98.02	99.68
Adapter removal	No	Yes	No	Yes

Supplementary Table 2. Transcriptome assembly and quality assessment statistics for *T. cinerariifolium* and *T. coccineum*.

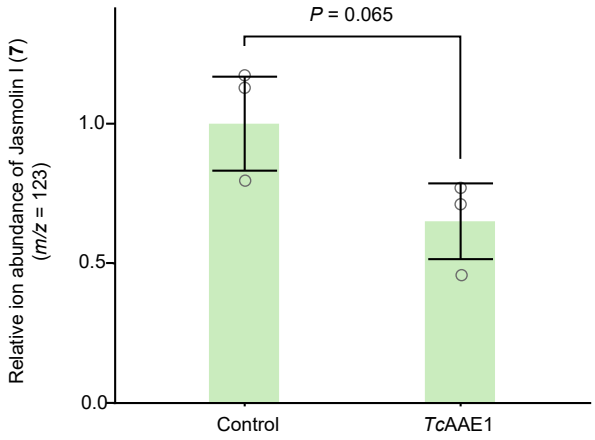
Parameter	*T. cinerariifolium* Uncleaned	*T. cinerariifolium* Cleaned	*T. coccineum* Uncleaned	*T. coccineum* Cleaned
Assembled transcripts	189,541	184,925	191,347	172,012
Unigenes	117,586	95,435	105,768	105,377
ExN50	-	Ex85N50 = 1,701 bp	-	Ex84N50 = 1,604 bp
Complete BUSCOs (%)	-	99.2	-	96.9
Fragmented BUSCOs (%)	-	0.8	-	2.4
Missing BUSCOs (%)	-	0.0	-	0.8
Predicted protein sequences	111,146	89,632	99,321	88,163

-, not determined.

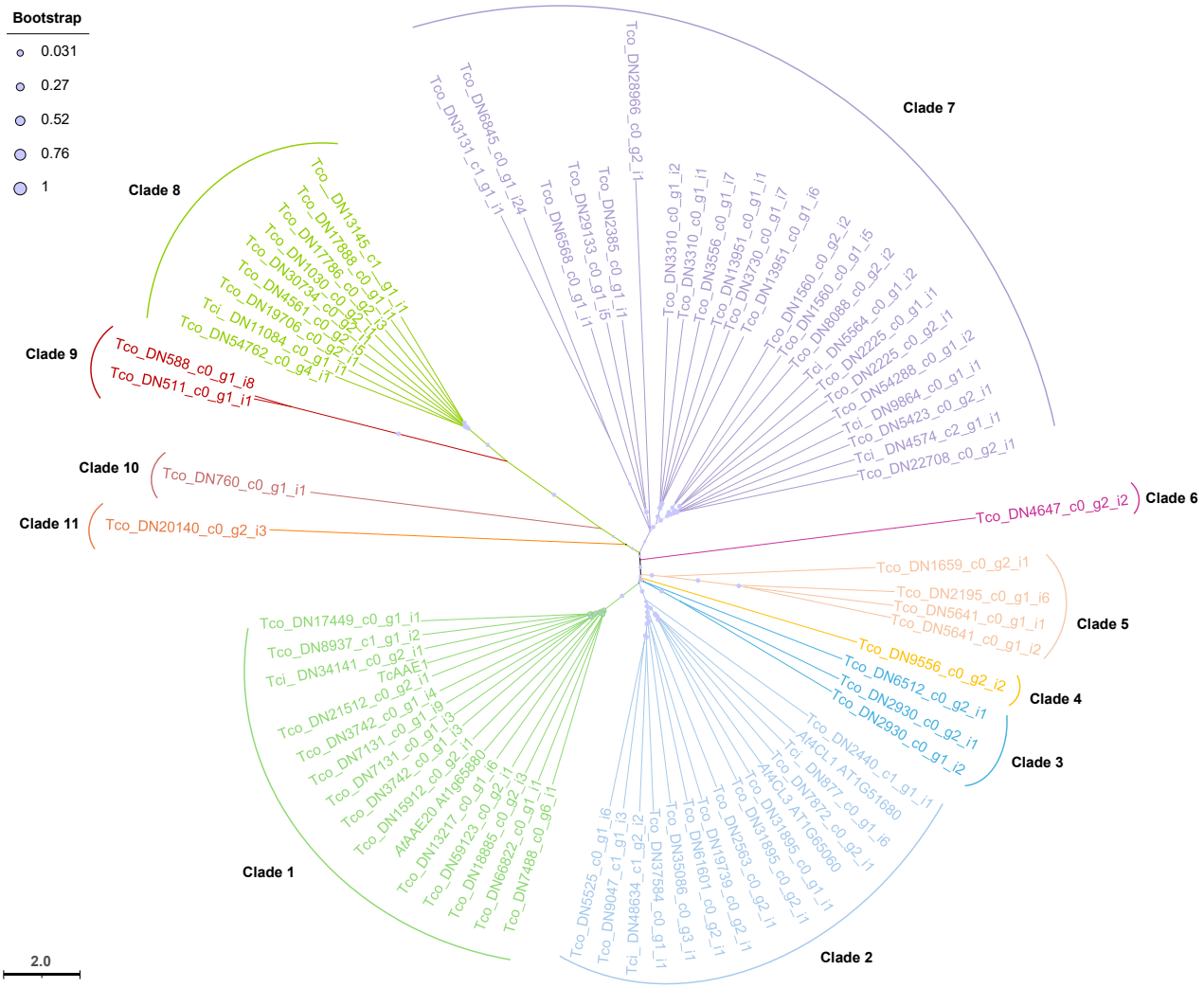
Supplementary Table 3. Strains used in this study

Strains	**Genotype**
EGY48	W303, *MATalpha, ura3, his3, trp1, 6xLexAop::LEU2*
PY1	EGY48, *X4*::P_{GAL1}-*TcJMH*-T_{CYC1}-P_{GAL10}-*TcPYS*-T_{PGL1}-P_{GAL1}-*TcCPR*-T_{FBA1}-P_{GAL1}-*TcJMH*-T_{TDH2}-P_{GAL10}-*TcPYS*- T_{ADH1}
PY2	EGY48, *X4*::P_{GAL1}-*TcJMH*-T_{CYC1}-P_{GAL10}-*TcPYS*-T_{PGL1}-P_{GAL1}-*TcCPR*-T_{FBA1}-P_{GAL1}-*TcJMH*-T_{TDH2}-P_{GAL10}-*TcPYS*- T_{ADH1}, *XII-5*::P_{GAL1}-*TcoUGT1*-T_{CYC1}
PY3	EGY48, *X4*::P_{GAL1}-*TcJMH*-T_{CYC1}-P_{GAL10}-*TcPYS*-T_{PGL1}-P_{GAL1}-*TcCPR*-T_{FBA1}-P_{GAL1}-*TcJMH*-T_{TDH2}-P_{GAL10}-*TcPYS*- T_{ADH1}, *XII-5*:: P_{GAL1}-*MBP-Δ27-TcGLIP* -T_{CYC1}
PY4	EGY48, *X4*::P_{GAL1}-*TcJMH*-T_{CYC1}-P_{GAL10}-*TcPYS*-T_{PGL1}-P_{GAL1}-*TcCPR*-T_{FBA1}-P_{GAL1}-*TcJMH*-T_{TDH2}-P_{GAL10}-*TcPYS*- T_{ADH1}, *XII-5*::P_{GAL1}-*TcoUGT1*-T_{CYC1}-P_{GAL10}-*MBP-Δ27-TcGLIP*-T_{PGL1}

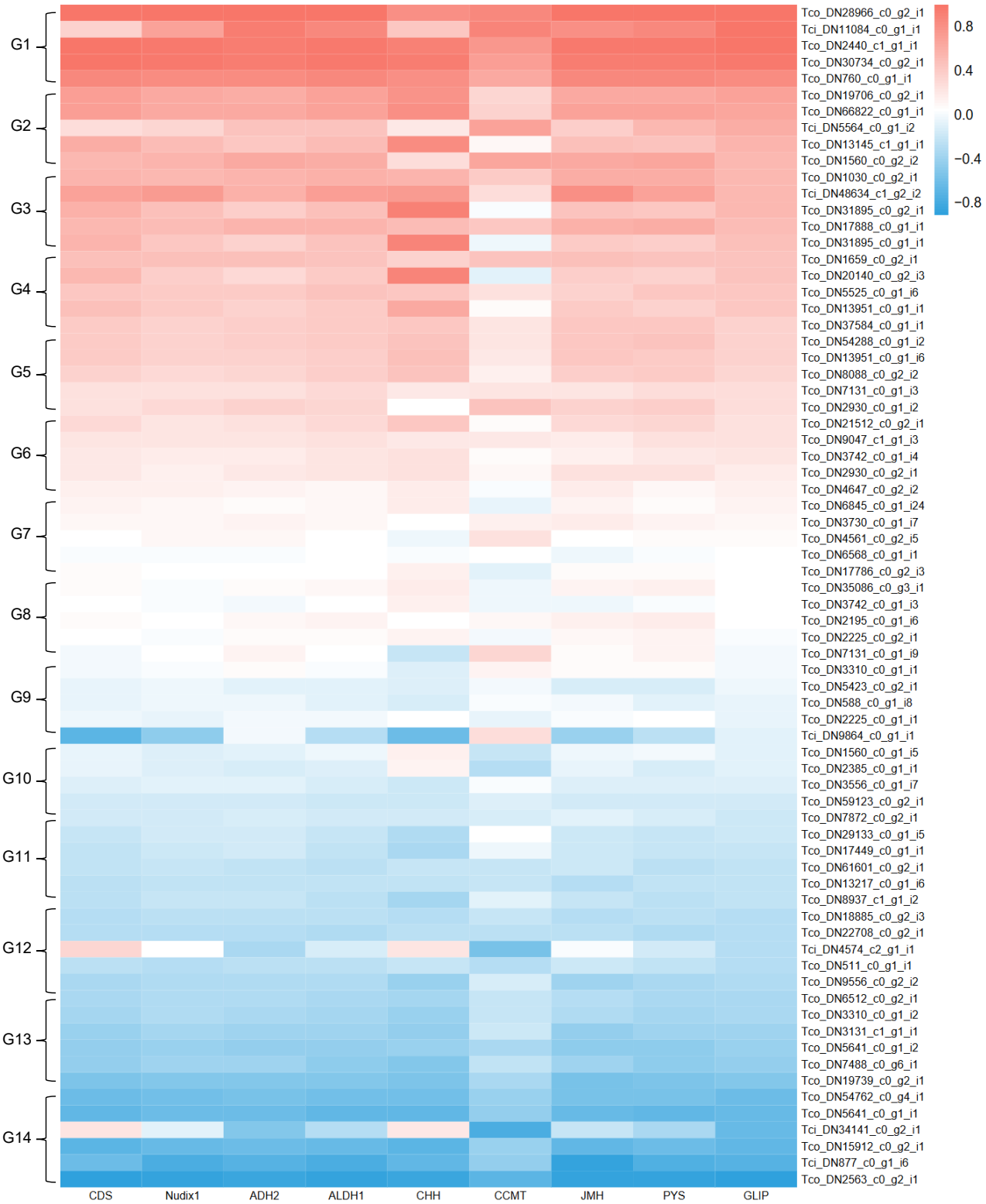
SUPPLEMENTARY FIGURES



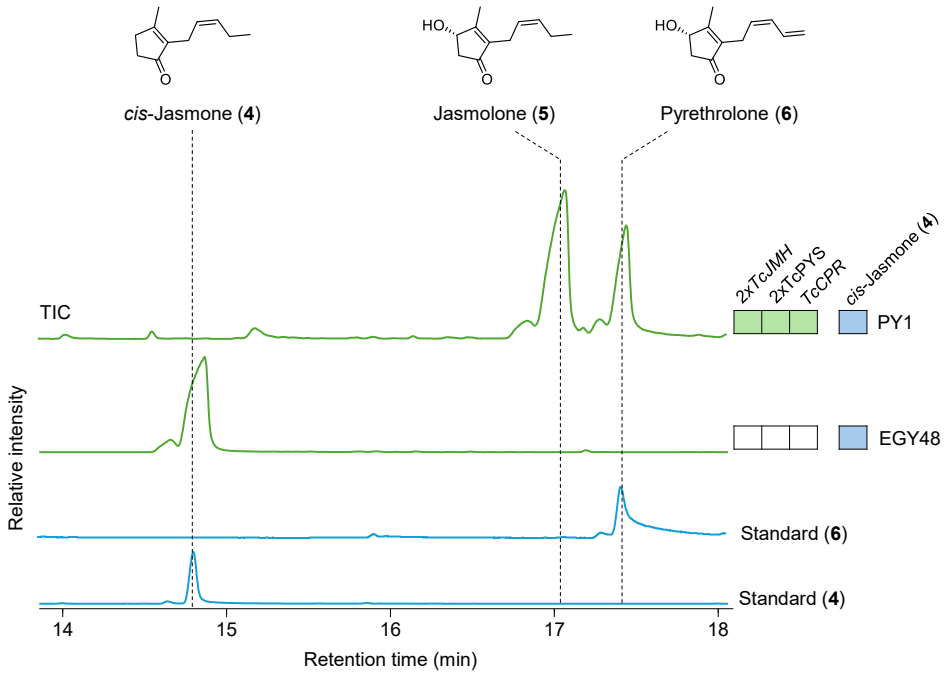
Supplementary Fig. 1. **Effect of transient *TcAAE1* expression on jasmolin I production in *N. benthamiana*.** Relative abundance of jasmolin I (**7**) in tobacco leaves transiently co-expressing *TcJMH* and *TcGLIP*, with or without *TcAAE1*. Leaves were supplemented with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**) and extracts were analyzed by GC-EI-MS. Jasmolin I abundance was normalized to the control expressing *TcJMH* and *TcGLIP* without *TcAAE1* (control = 1). Data are presented as mean ± SD of three independent biological replicates ($n = 3$). No statistically significant difference was observed between the two groups (one-way ANOVA, $P = 0.065$).



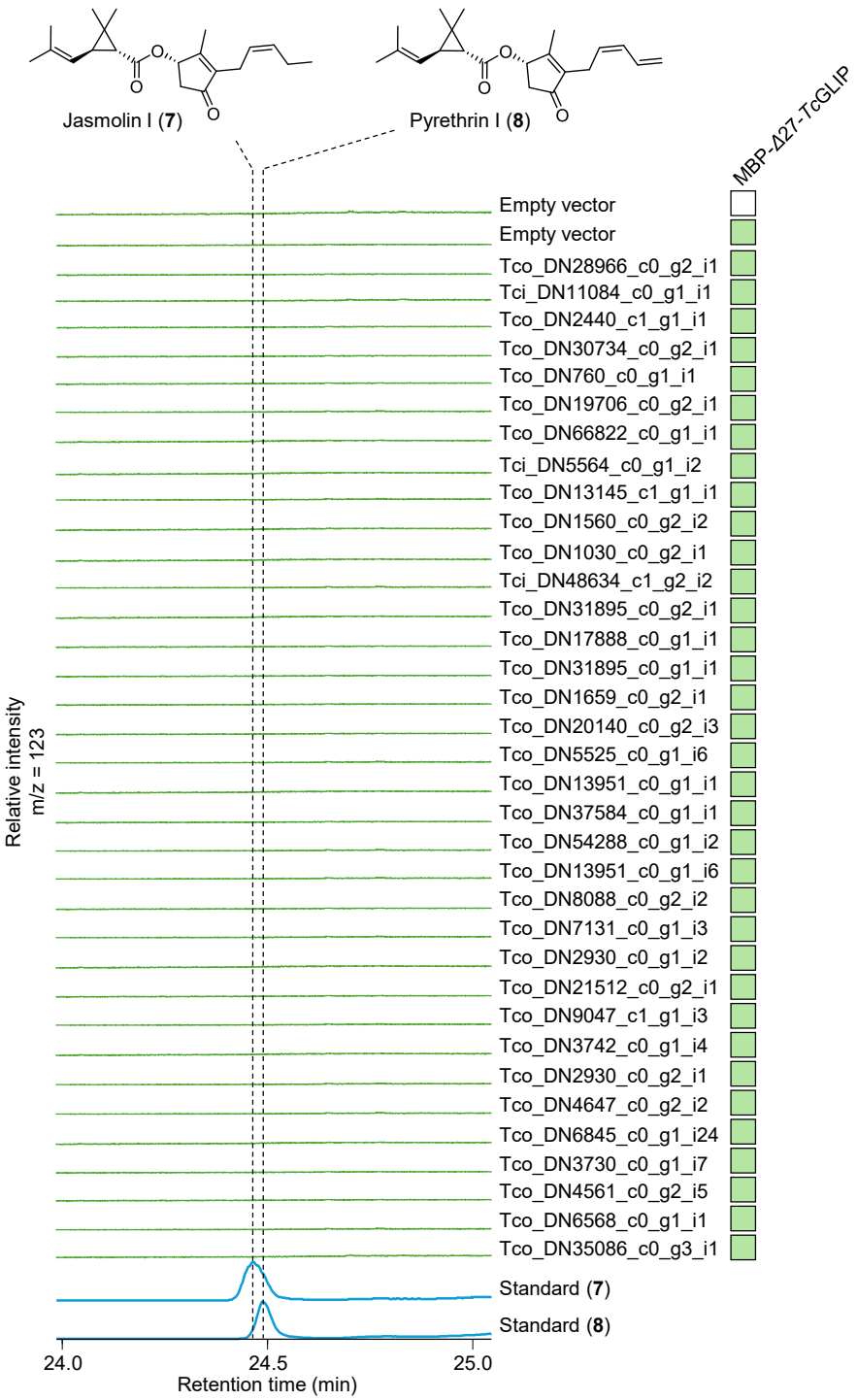
Supplementary Fig. 2. Maximum-likelihood phylogenetic analysis of candidate CoA ligases. Amino acid sequences of 72 candidate CoA ligases identified from *T. cinerariifolium* and *T. coccineum*, together with *Tc*AAE1 and three *Arabidopsis thaliana* reference proteins (*At*4CL1, *At*4CL3 and *At*AAE20), were aligned using ClustalW in MEGA 12. A maximum-likelihood phylogenetic tree was constructed in MEGA 12, and branch support was assessed using 1,000 bootstrap replicates. Bootstrap support values are indicated at the corresponding nodes. The scale bar represents 2.0 amino acid substitutions per site.



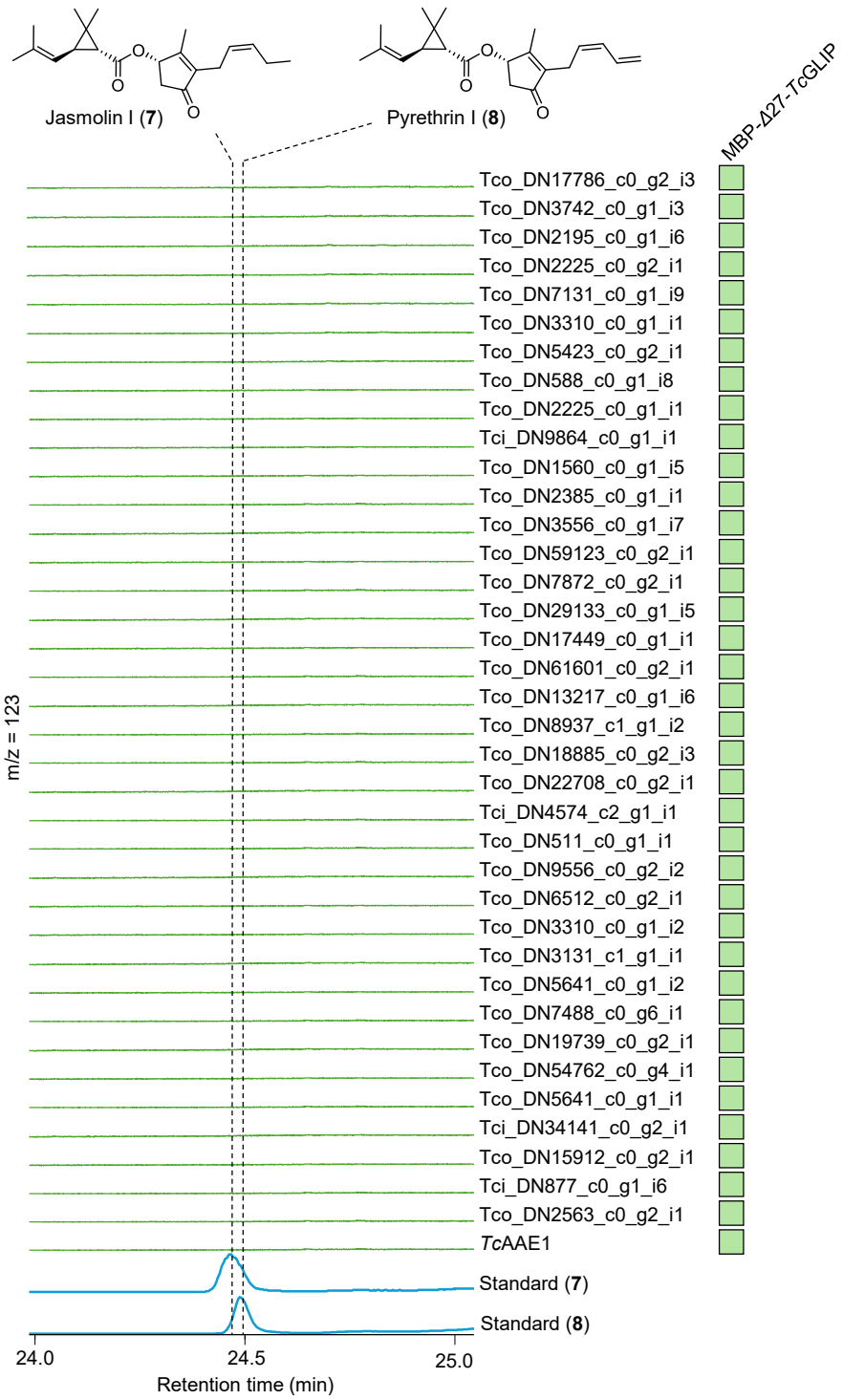
Supplementary Fig. 3. Co-expression analysis of candidate CoA ligases and characterized pyrethrin biosynthetic genes. Heatmap showing Pearson correlation coefficients (*r*) between the expression profiles of 72 candidate CoA ligases and previously characterized pyrethrin biosynthetic genes. Red indicates positive correlations ($r > 0$), blue indicates negative correlations ($r < 0$), and white indicates correlations close to zero. The 72 candidates were organized into 14 groups (G1-G14) according to their correlation with *TcGLIP*, with each group containing five or six candidates. Group assignments are indicated on the left of the heatmap.



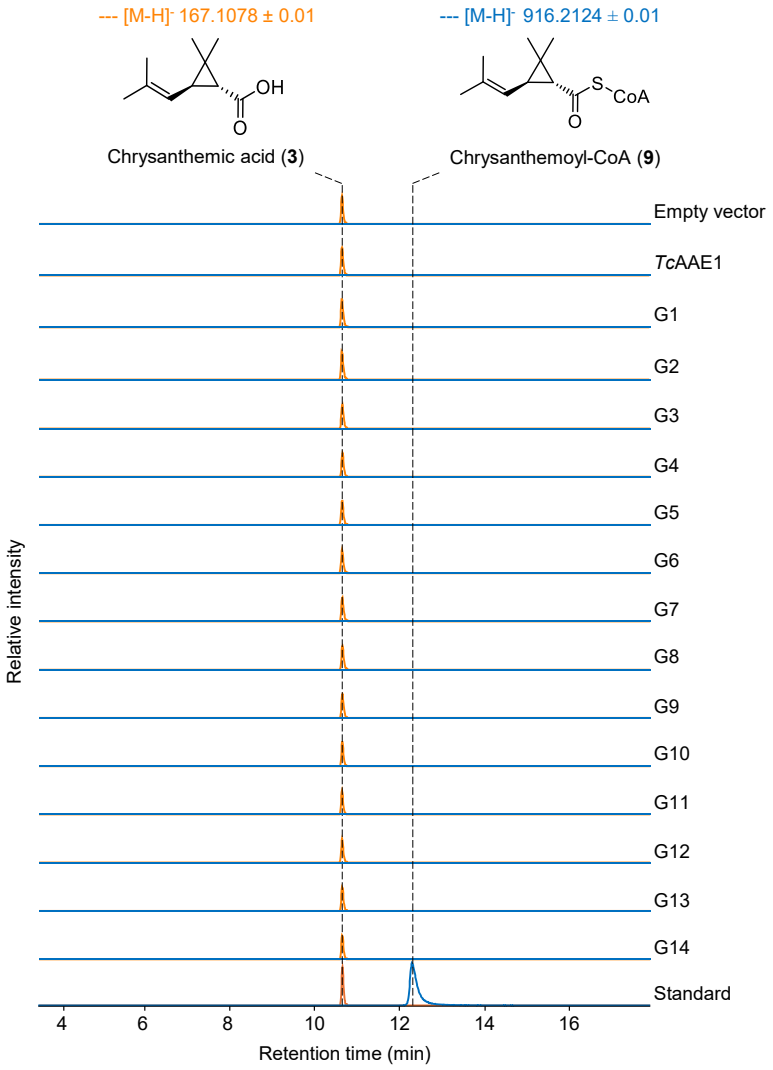
Supplementary Fig. 4. Heterologous reconstruction of rethrolone alcohol biosynthesis in *S. cerevisiae*. GC-EI-MS total ion chromatograms (TICs) of ethyl acetate extracts from engineered yeast strains expressing the indicated combinations of biosynthetic enzymes. Strain PY1 was derived from EGY48 by genomic integration of two copies each of *TcJMH* and *TcPYS* and one copy of *TcCPR*. Yeast cultures were supplemented with 300 mg L^{-1} *cis*-jasmone (**4**). Black dashed lines indicate the retention times of the corresponding compounds. Open squares indicate the absence of the corresponding enzyme. Chromatograms of authentic *cis*-jasmone (**4**) and pyrethrolone (**6**) standards are shown for comparison.



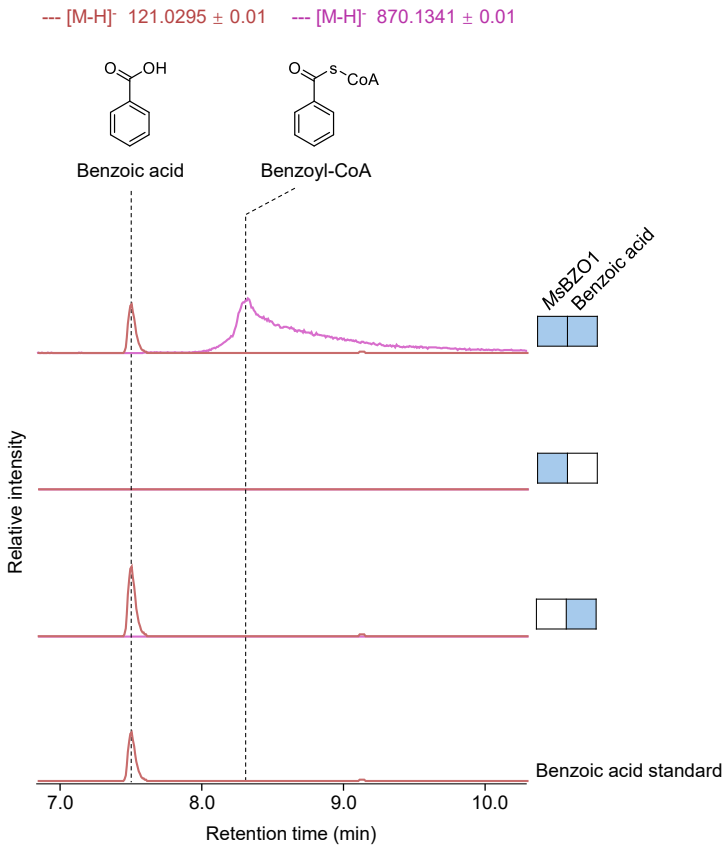
Supplementary Fig. 5. Functional evaluation of candidate CoA ligases in yeast expressing *MBP–Δ27–TcGLIP* (candidates 1-35). GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from yeast strain PY1 episomally expressing *MBP-Δ27-TcGLIP* from the pHUS vector together with individual CoA ligase candidates from the pLUS vector. The first 35 of 72 candidate CoA ligases are shown here and the remaining 37 are presented in Supplementary Fig. 6. Yeast cultures were supplemented with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**). Chromatograms of authentic jasmolin I (**7**) and pyrethrin I (**8**) standards are shown for comparison. Black dashed lines indicate the retention times of the corresponding standards. Open squares indicate the absence of MBP-*Δ*27-*Tc*GLIP.



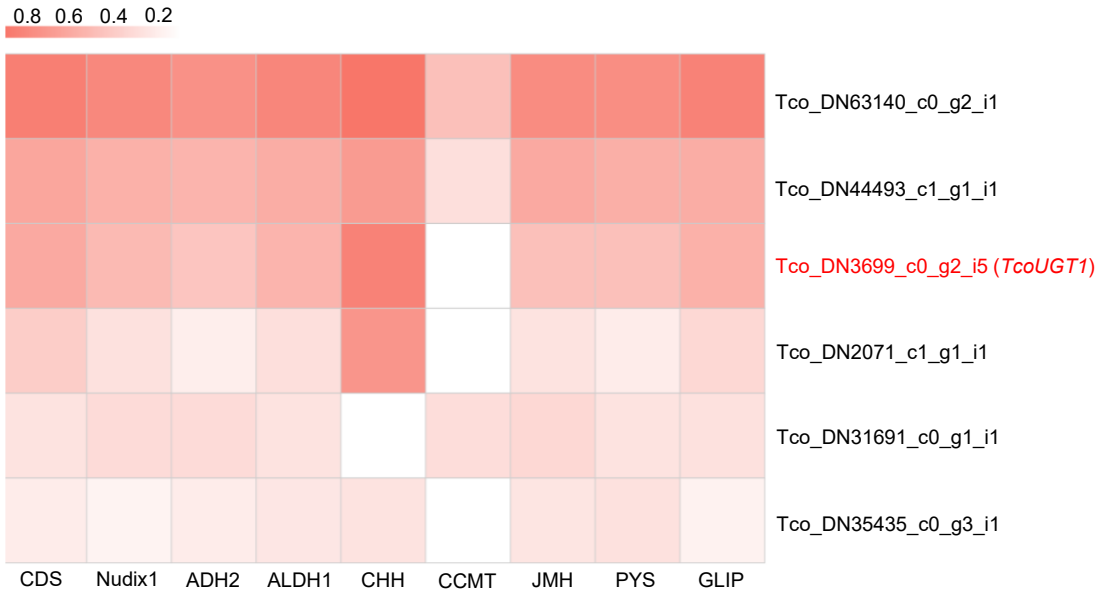
Supplementary Fig. 6. Functional evaluation of candidate CoA ligases in yeast expressing *MBP–Δ27–TcGLIP* (candidates 36-72). GC-EI-MS extracted ion chromatograms (*m/z* = 123) of ethyl acetate extracts from yeast strain PY1 episomally expressing *MBP-Δ27-TcGLIP* from the pHUS vector together with individual CoA ligase candidates from the pLUS vector. The remaining 37 of 72 candidate CoA ligases are shown here, while the first 35 are presented in Supplementary Fig. 5. Yeast cultures were supplemented with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**). Chromatograms of authentic jasmolin I (**7**) and pyrethrin I (**8**) standards are shown for comparison. Black dashed lines indicate the retention times of the corresponding standards.



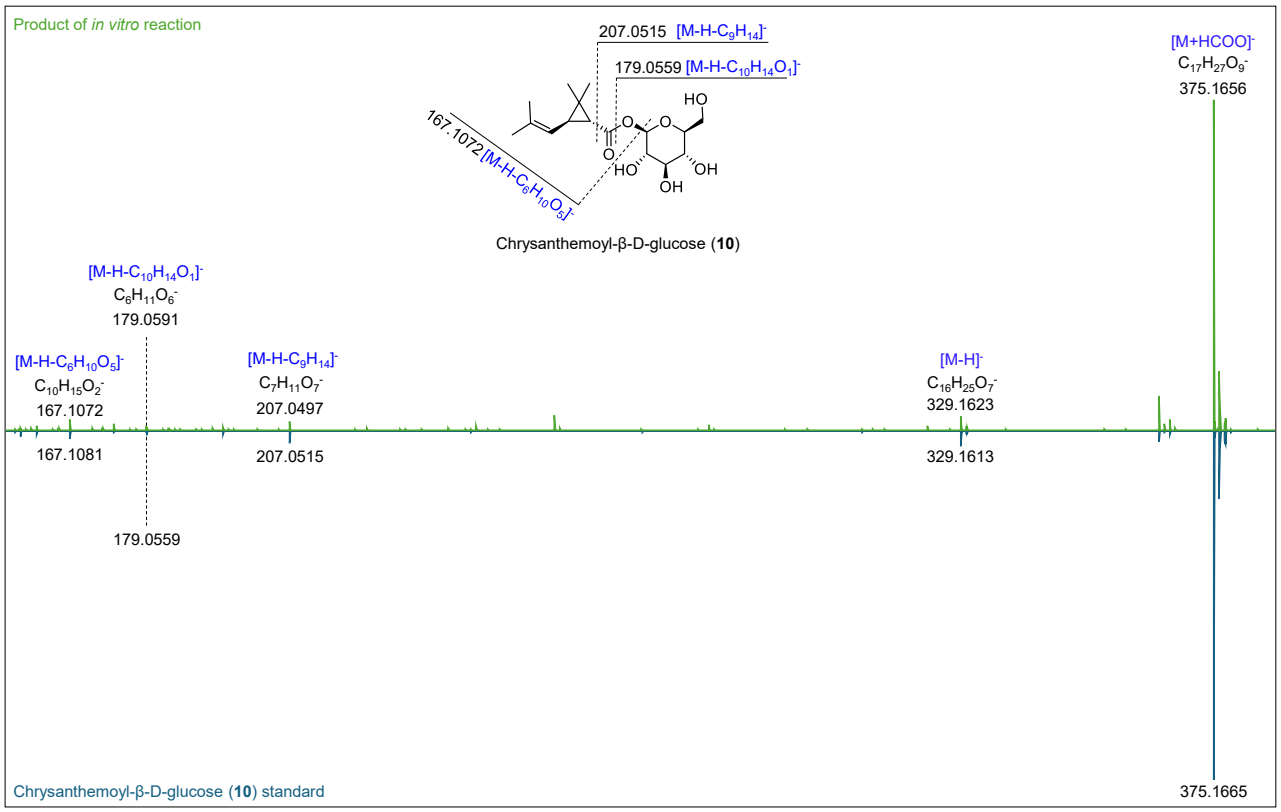
Supplementary Fig. 7. *In vitro* evaluation of candidate CoA ligases for chrysanthemoyl-CoA formation. UPLC-qTOF-MS extracted ion chromatograms (EICs; negative ESI mode) of methanol extracts from *in vitro* reactions, monitoring chrysanthemic acid (**3**) ([M-H]$^-$, *m/z* = 167.1078 ± 0.01) and chrysanthemoyl-CoA (**9**) ([M-H]$^-$, *m/z* = 916.2124 ± 0.01). Reactions containing chrysanthemic acid (**3**) were adjusted to pH 7.5 and incubated at 25 °C for 5 h. Different colours indicate the extracted ion chromatograms corresponding to each compound. Black dashed lines indicate the retention times of the corresponding compounds. Chromatograms of authentic chrysanthemic acid (**3**) and chrysanthemoyl-CoA (**9**) standards are shown for comparison.



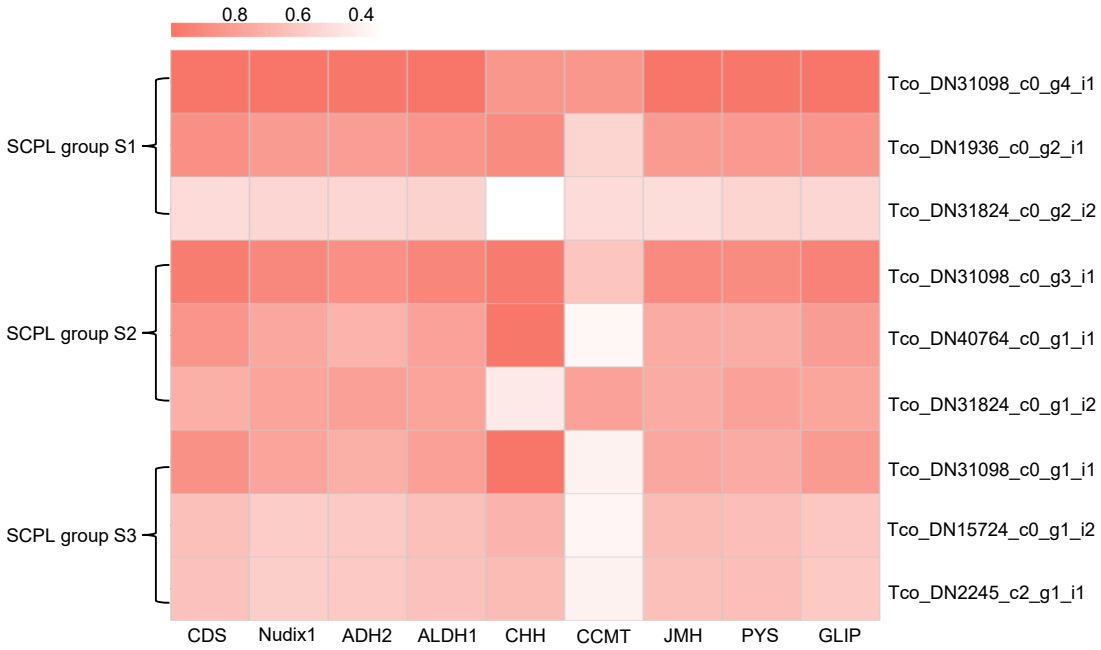
Supplementary Fig. 8. Positive-control assay demonstrating *Ms*BCL-mediated benzoyl-CoA formation. UPLC-qTOF-MS extracted ion chromatograms (EICs; negative ion mode) of methanol extracts from enzymatic assay reactions, monitoring benzoic acid ($[M-H]^-$, m/z = 121.0295 ± 0.01) and benzoyl CoA ($[M-H]^-$, m/z = 870.1341 ± 0.01). Reactions containing crude protein extracts from cells expressing *MsBCL* and benzoic acid were incubated at pH 7.5 and 25 °C for 5 h. Different colours indicate the extracted ion chromatograms corresponding to each compound. Black dashed lines indicate the retention times of the corresponding compounds. Open squares indicate the absence of the corresponding enzyme or substrate.



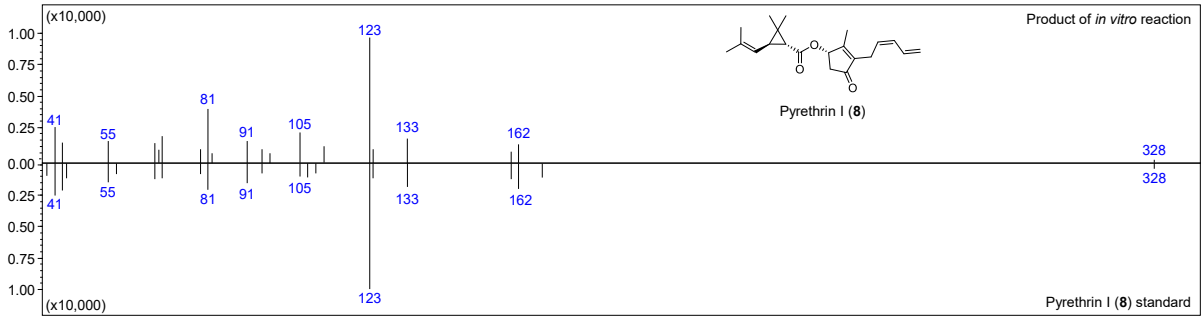
Supplementary Fig. 9. Co-expression analysis of candidate UDP-glucosyltransferases and characterized pyrethrin biosynthetic genes. Heatmap showing Pearson correlation coefficients (*r*) between the expression profiles of six candidate UGTs and previously characterized pyrethrin biosynthetic genes. Red indicates positive correlations ($r > 0$), whereas white represents near zero correlation ($r = 0$).



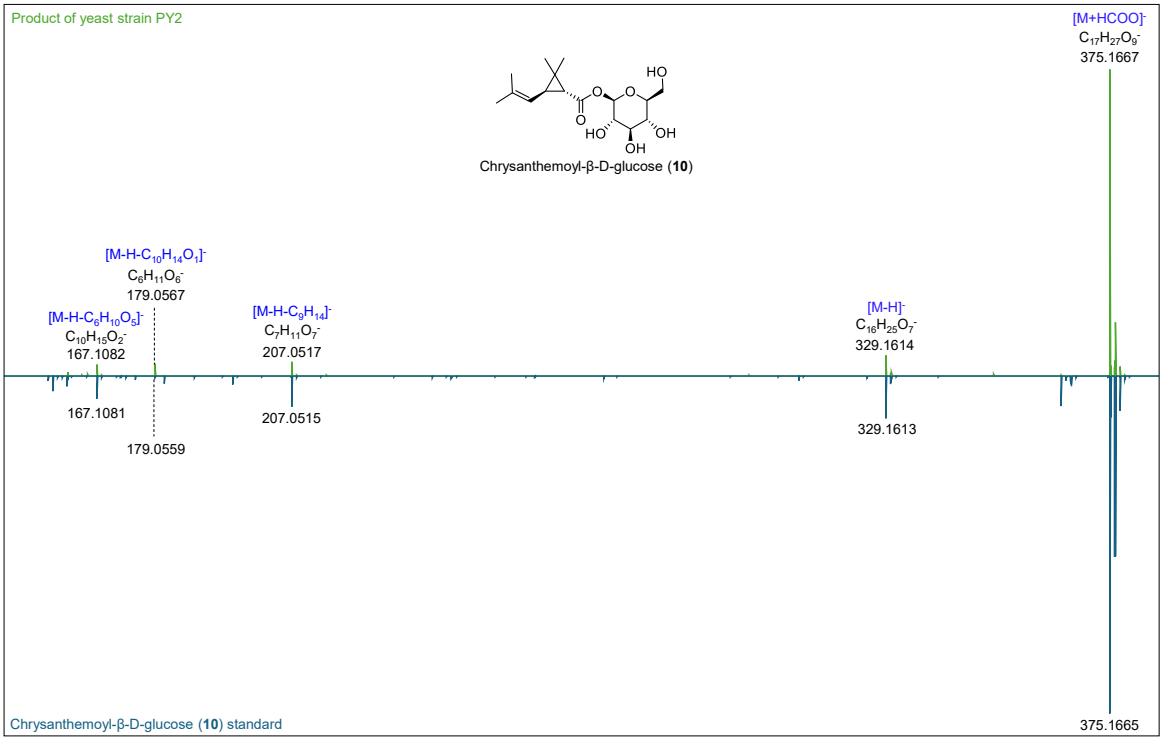
Supplementary Fig. 10. High-resolution mass spectra (HRMS) of chrysanthemoyl-β-D-glucose produced by *Tco*UGT1. HRMS obtained by UPLC-qToF-MS analysis in negative-ion mode of the product obtained from ethyl acetate extracts of *in vitro* reactions containing *Tco*UGT1, compared with the HRMS of a synthetic chrysanthemoyl-β-D-glucose (**10**) reference standard. The enzymatic product exhibited a predominant ion at *m/z* 375.1656, corresponding to the formate adduct ([M+HCOO]$^-$), together with an ion at *m/z* 329.1623 ([M-H]$^-$), and additional ions at *m/z* 207.0497, 179.0591 and 167.1072.



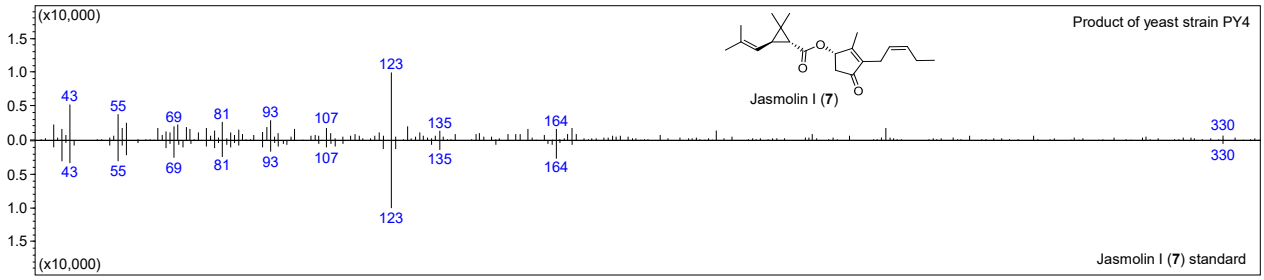
Supplementary Fig. 11. Co-expression analysis of candidate *SCPL* acyltransferases and characterized pyrethrin biosynthetic genes. Heatmap showing Pearson correlation coefficients (*r*) between the expression profiles of nine candidate SCPL acyltransferases and previously characterized pyrethrin biosynthetic genes. The nine candidates were organized into three groups (S1-S3) according to their correlation with *ALDH1*, with each group containing three candidates. Red indicates positive correlations ($r > 0$), whereas white represents near zero correlation ($r = 0$)



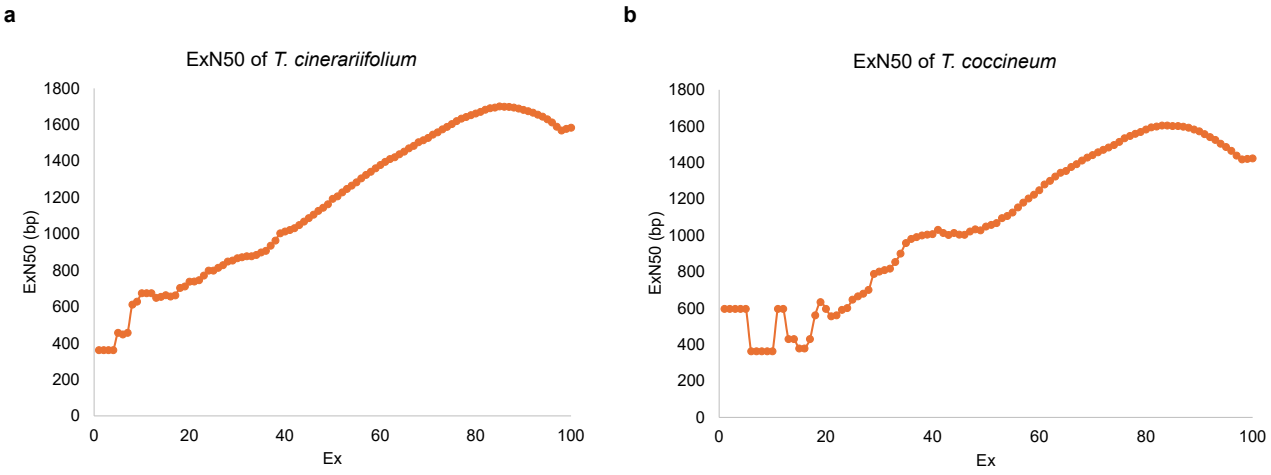
Supplementary Fig. 12. Comparison of electron ionization mass spectra of enzymatically produced pyrethrin I and an authentic standard. GC-EI-MS spectra of pyrethrin I (**8**) detected in ethyl acetate extracts of *in vitro* reactions containing crude protein extracts expressing *TcoUGT1* and *MBP-Δ27-TcGLIP*, compared with the mass spectrum of an authentic pyrethrin I (**8**) standard.



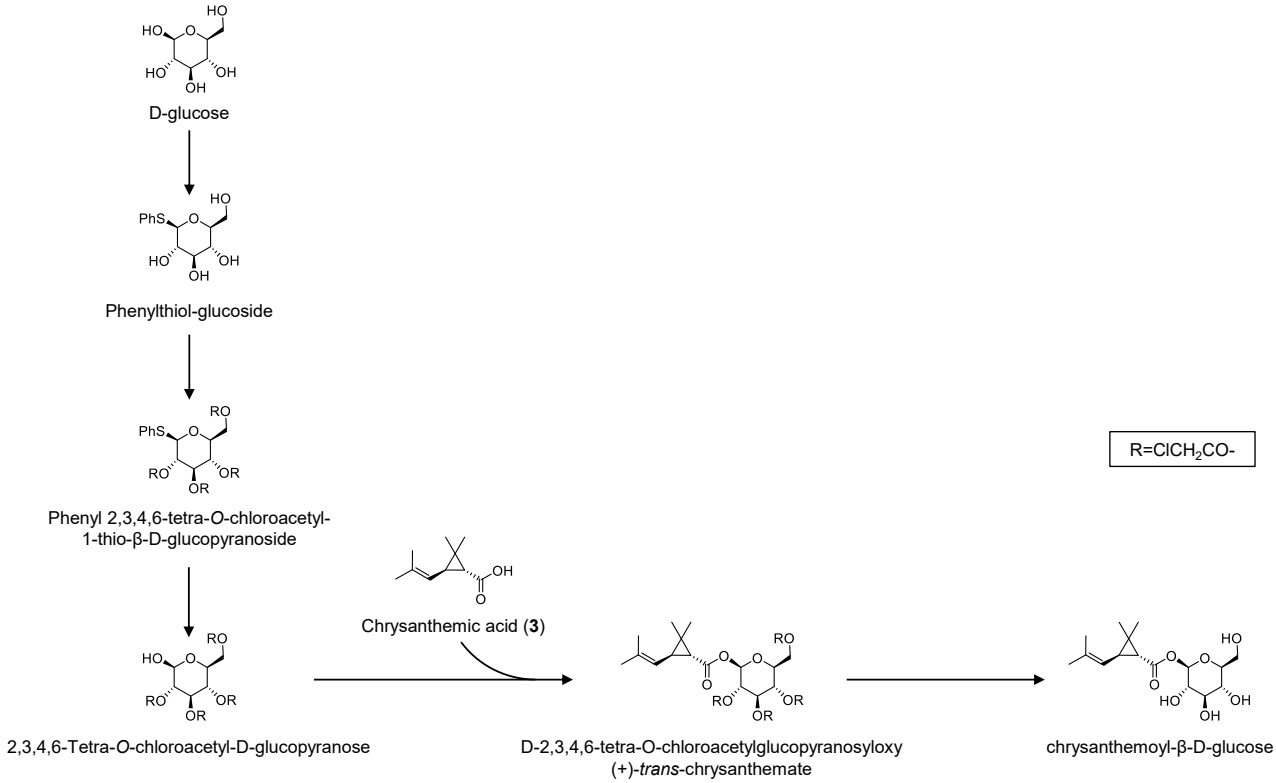
Supplementary Fig. 13. HRMS analysis of chrysanthemoyl-glucose produced in yeast strain PY2. High-resolution mass spectra (HRMS; negative-ion mode) of the compound detected in ethyl acetate extracts of yeast strain PY2 expressing *TcoUGT1*, compared with a synthetic chrysanthemoyl-glucose (**10**) reference standard. The compound detected in PY2 exhibited a predominant ion at *m/z* 375.1667, corresponding to the formate adduct $[M+HCOO]^-$, together with fragment ions *m/z* at 329.1614 ($[M-H]^-$), 207.0517, 179.0567 and 167.1082.



Supplementary Fig. 14. Mass spectra comparison of jasmolin I detected from yeast PY4 with standard. Mass spectra comparison of jasmolin I (**7**) detected from GC-EI-MS extracted ion chromatograms of ethyl acetate extracted yeast strain PY4 with authentic standard. Yeast PY4 was fed with 100 mg L^{-1} chrysanthemic acid (**3**) and 300 mg L^{-1} *cis*-jasmone (**4**).



Supplementary Fig. 15. ExN50 analysis of *T. cinerariifolium* and *T. coccineum* transcriptomes. a. ExN50 analysis of *T. cinerariifolium* transcriptome. Maximum ExN50 value was 1,701 bp at Ex85 for the *T. cinerariifolium* transcriptome. **b**. ExN50 analysis of *T. coccineum* transcriptome and maximum ExN50 value was 1,604 bp at Ex84 for the *T. coccineum* transcriptome.



Supplementary Fig. 16. Chemical synthesis of chrysanthemoyl-β-D-glucose.