Biosynthesis and Bioactivity of Anti-Inflammatory Triterpenoids in *Calendula officinalis*

Golubova D[1,2‡], Salmon M[1‡*], Su H[1,3‡], Tansley C[1,3], Kaithakottil GG[1], Linsmith G[1], Schudoma C[1], Swarbreck D[1], O'Connell MA[*2], Patron NJ[*1,3]

[1] Engineering Biology, Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ
[2] School of Chemistry, Pharmacy and Pharmacology, University of East Anglia, Norwich, NR4 7TJ
[3] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA

Supplementary Information

Supplementary Fig. 1: Metabolite analysis of *Calendula officinalis* by GC-MS
Supplementary Fig. 2: Metabolite analysis of *Calendula officinalis* by LC-MS
Supplementary Fig. 3: GC-MS standards and mass spectra for triterpene monols
Supplementary Fig. 4: GC-MS standards and mass spectra for triterpene diols and acids
Supplementary Fig. 5: GC-MS standards and mass spectra for faradiol fatty acid esters
Supplementary Fig. 6: Semi-preparative uHPLC chromatograms of the methanol extract of *Calendula officinalis* ray florets
Supplementary Fig. 7: GC-MS analysis of compounds purified from *Calendula officinalis* ray floret extracts
Supplementary Fig. 8: GC-MS analysis of *Calendula officinalis* ray floret fractions
Supplementary Fig. 9: Structure of 3-*O*-palmitoyl faradiol (faradiol palmitate)
Supplementary Fig. 10: ^{1}H NMR spectra of faradiol palmitate
Supplementary Fig. 11: ^{13}C NMR spectra of faradiol palmitate
Supplementary Fig. 12: ^{1}H-^{1}H COSY NMR (600 MHz, CDCl3, 298 K) of faradiol palmitate
Supplementary Fig. 13: ^{1}H-^{13}C-HSQC-edited NMR (600 MHz, CDCl3, 298 K) of faradiol palmitate
Supplementary Fig. 14: ^{1}H-^{13}C HMBC NMR (600 MHz, CDCl3, 298 K) of faradiol palmitate
Supplementary Fig. 15: Effects of *C. officinalis* extracts and triterpenoids on the viability of human monocytic (THP-1) cells
Supplementary Fig. 16. Concentration-dependent responses to floral extracts and faradiol
Supplementary Fig. 17: Phylogenetic analysis of plant oxidosqualene cyclases
Supplementary Fig. 18: Chromosomal location of *Calendula officinalis OSC* genes
Supplementary Fig. 19: Differential expression of *Calendula officinalis OSC* genes
Supplementary Fig. 20: GC-MS analysis of *N. benthamiana* leaf extracts infiltrated with constructs expressing taraxasterol synthases (TXSSs)
Supplementary Fig. 21: Bayesian phylogenetic tree of mixed amyrin synthases (MAS) and taraxasterol synthases (TXSS)
Supplementary Fig. 22 Genomic locations of *Calendula officinalis*, *Helianthus annuus* (sunflower), *Lactuca sativa* (lettuce) and *Cynara cardunculus* (globe artichoke) *TXSS* genes
Supplementary Fig. 23: GC-MS profile of *Calendula officinalis* tissues
Supplementary Fig. 24: GC-MS profile of *Taraxacum kok-saghyz* (Russian dandelion) tissues
Supplementary Fig. 25: GC-MS analysis of *N. benthamiana* expressing TXSS mutants
Supplementary Fig. 26: Maximum-likelihood tree of plant cytochrome p450s (CYPs)
Supplementary Fig. 27: Differential expression of *Calendula officinalis CYP* genes
Supplementary Fig. 28. Genomic location and synteny of *Calendula officinalis* (pot marigold) genes encoding CYP716A392 and CYP716A393.
Supplementary Fig. 29: GC-MS analysis of *Nicotiana benthamiana* expressing TXSS and CYPs
Supplementary Fig. 30: GC-MS profiling of Asteraceae floral extracts
Supplementary Fig. 31: GC-MS analysis of *Nicotiana benthamiana* expressing CYP mutants
Supplementary Fig. 32. Maximum-likelihood tree of plant acyltransferases.
Supplementary Fig. 33: Differential expression of *Calendula officinalis ACT* genes
Supplementary Fig. 34: GC-MS analysis of *Nicotiana benthamiana* expressing TXSS and ACTs
Supplementary Fig. 35: GC-MS analysis of *Nicotiana benthamiana* expressing TXSS, CYPs and ACTs
Supplementary Fig. 36: Genomic location and synteny of *Calendula officinalis ACT* genes

Supplementary Fig. 37: Peak area analysis showing the accumulation of Ψ-taraxasterol, faradiol and faradiol palmitate through *Calendula officinalis* floral development
Supplementary Table 1: Triterpenes detected in *Calendula officinalis*
Supplementary Table 2: Table of genome assembly statistics
Supplementary Table 3: *Calendula officinalis OSC* transcripts
Supplementary Table 4: Taxa and accession numbers of sequences used for positive selection analysis in Figure 5B.
Supplementary Data 1: NMR data for faradiol palmitate
Supplementary Data 2: Comparison of experimental and literature assignment of NMR data for faradiol palmitate
Supplementary Data 3: Statistical tests used in Figure 2
Supplementary Data 4: Candidate *OSC* genes identified in the *Calendula officinalis* genome
Supplementary Data 5: Coding and translated protein sequences of *Calendula officinalis OSC* genes
Supplementary Data 6: OSCs used in the phylogenetic analyses in Figure 3
Supplementary Data 7: Statistical tests used in Figure 4E
Supplementary Data 8: Statistical tests used in Figure 5E
Supplementary Data 9: Statistical tests used in Figure 6D and Supplementary Figure 31
Supplementary Data 10: Statistical tests used in Figure 7B
Supplementary Data 11: Plasmids used in this study
Supplementary Data 12: Primers used in this study
Supplementary Methods: Methods for genome assembly, annotation and differential gene expression analysis