

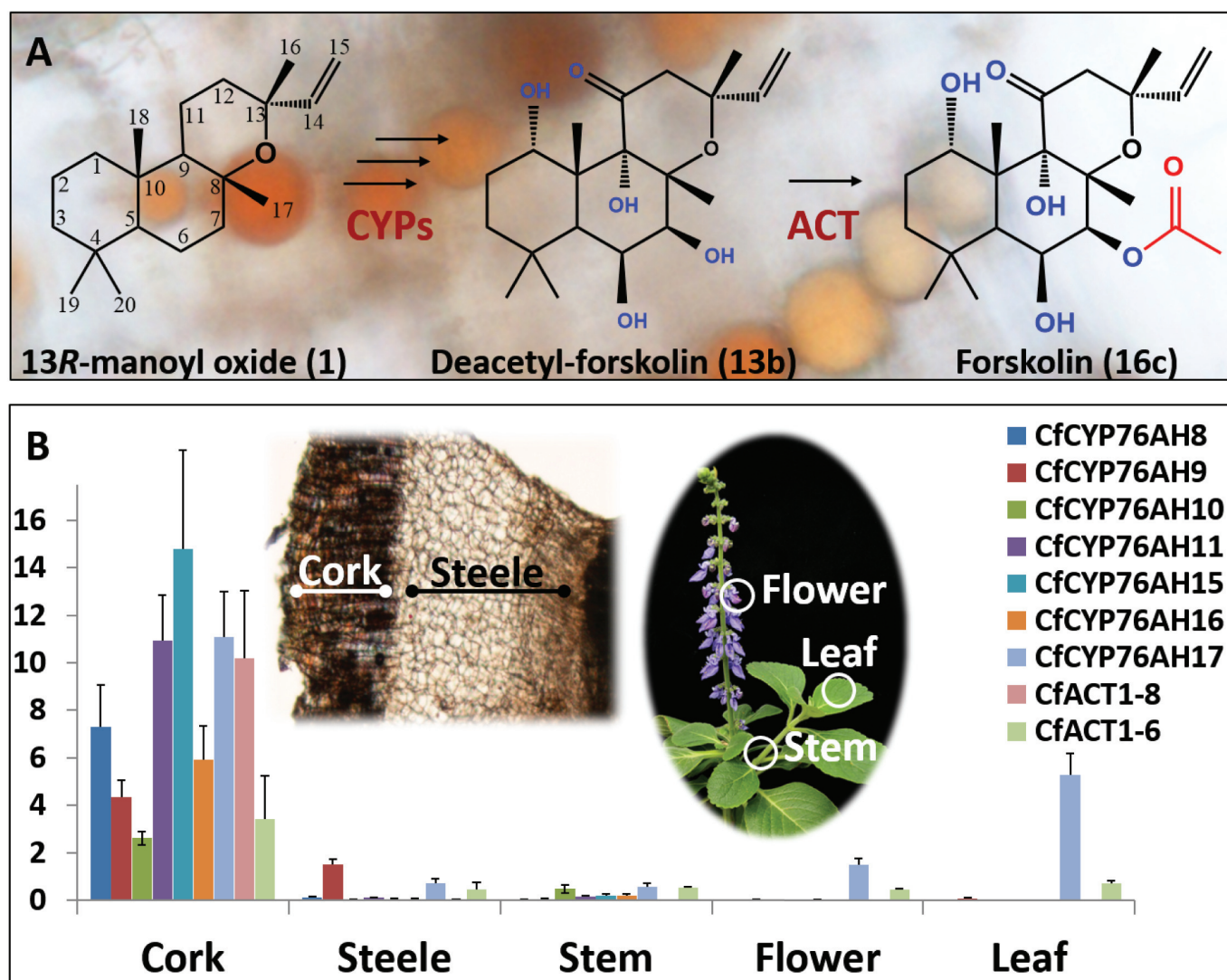


---

## Figures and figure supplements

Total biosynthesis of the cyclic AMP booster forskolin from *Coleus forskohlii*

**Irini Pateraki et al**



**Figure 1.** Biosynthesis of forskolin in the root cork cells of *C. forskohlii*. (A) Scheme showing the structures of the diterpene precursor 13R-manoyl oxide, deacetyl-forskolin and forskolin on a background of root cork cells with forskolin containing oil bodies. (B) Transcript profiles of biosynthetic candidate genes in selected tissues of *C. forskohlii* as shown on the illustrations.

DOI: [10.7554/eLife.23001.003](https://doi.org/10.7554/eLife.23001.003)

The following source data is available for figure 1:

**Source data 1.** cDNAs identified in the *C. forskohlii* root cork transcriptome and cloned during this work, with the GeneBank accession numbers.

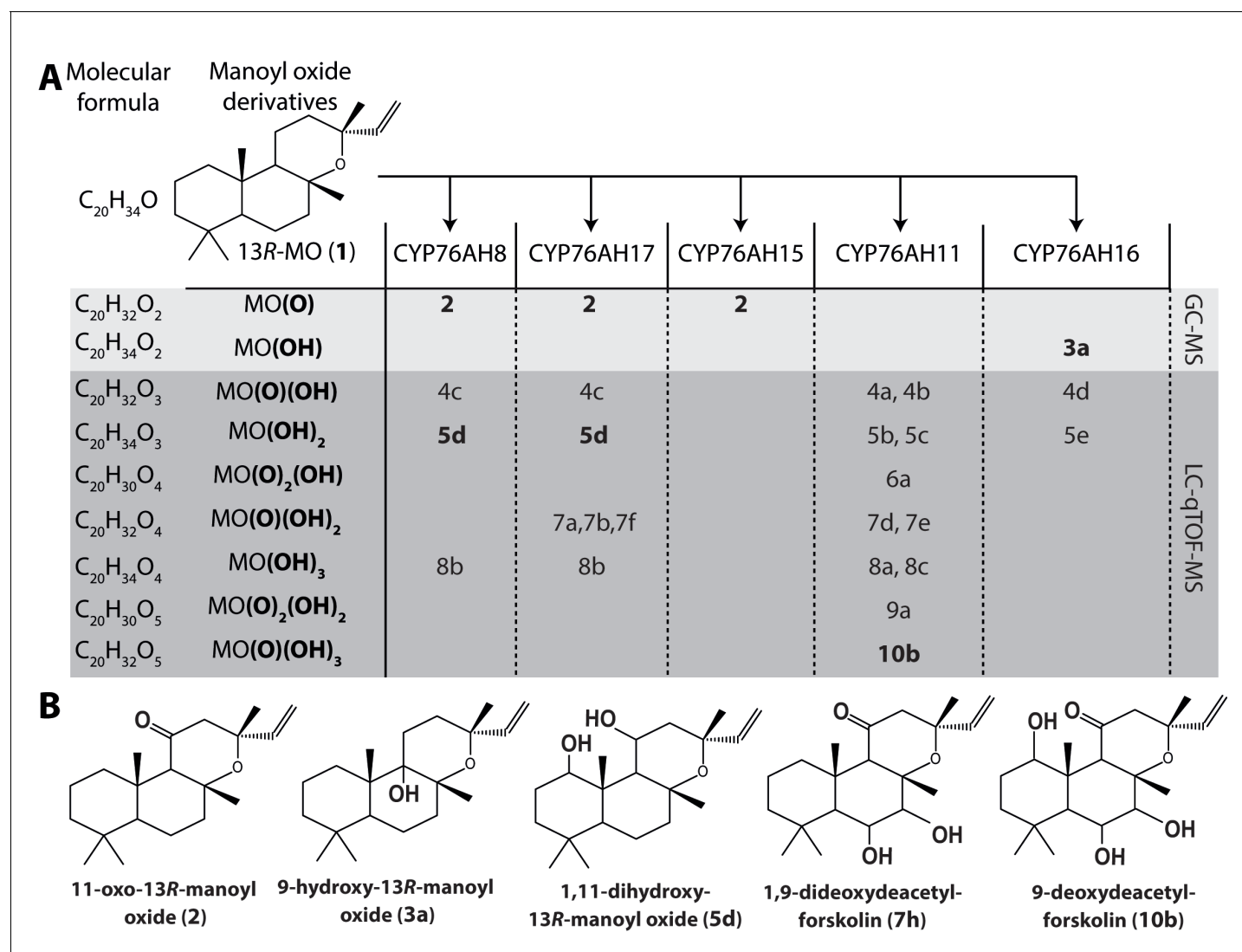
DOI: [10.7554/eLife.23001.004](https://doi.org/10.7554/eLife.23001.004)

**Source data 2.** Table of FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values of the first 20 most abundant cDNAs identified in the root cork transcriptome library.

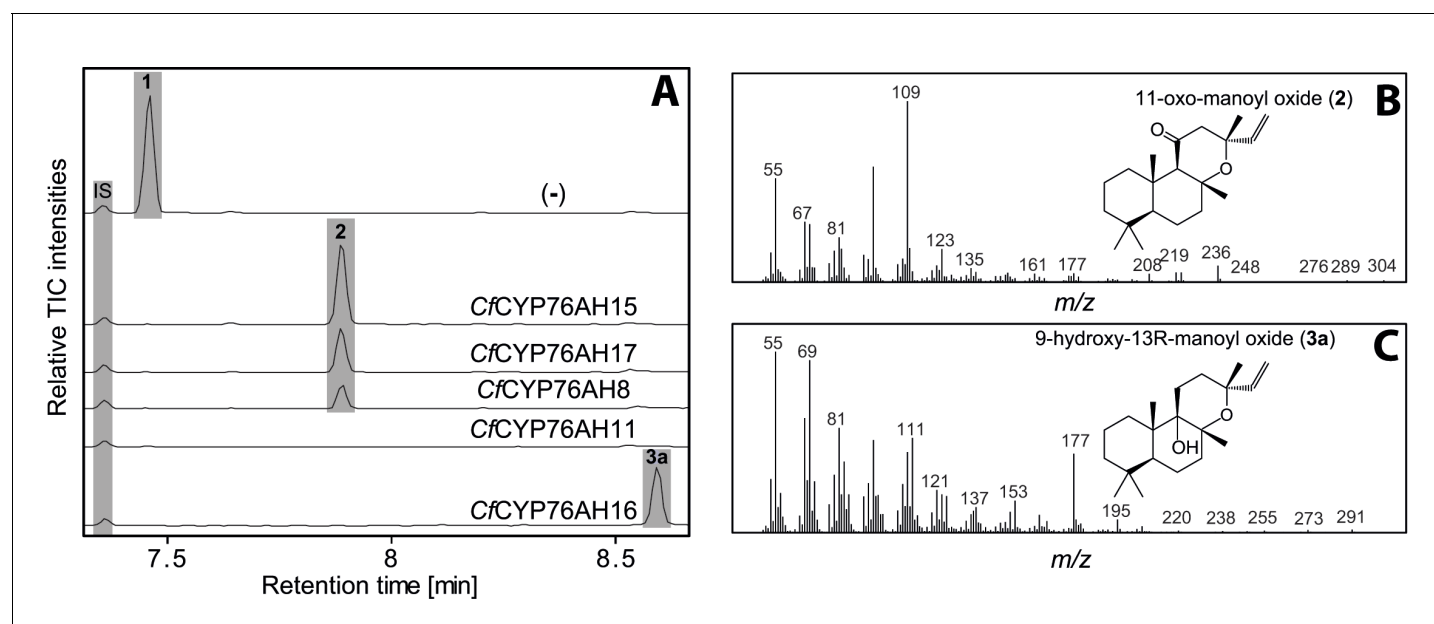
DOI: [10.7554/eLife.23001.005](https://doi.org/10.7554/eLife.23001.005)

**Source data 3.** Table of primers used in this study.

DOI: [10.7554/eLife.23001.006](https://doi.org/10.7554/eLife.23001.006)

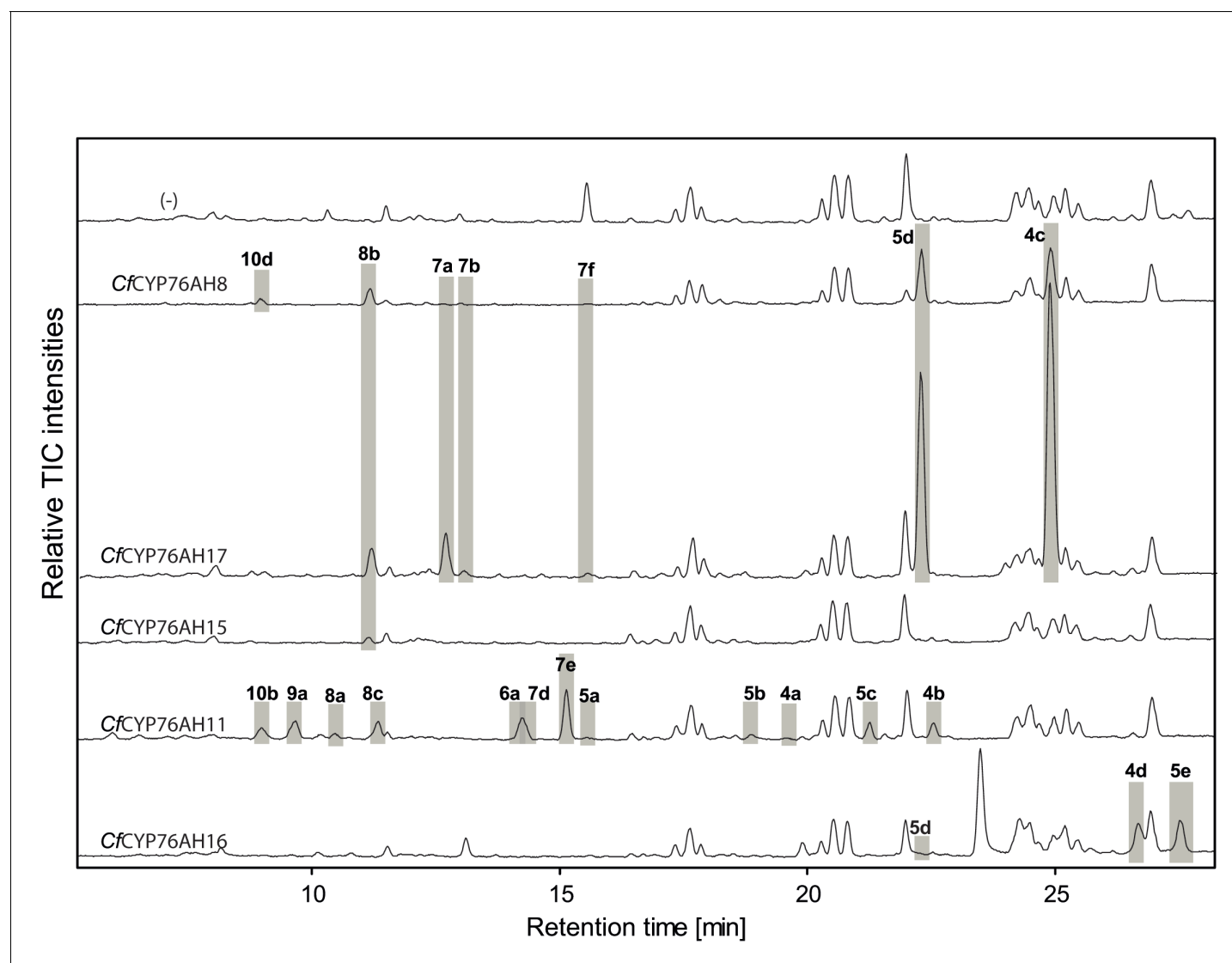


**Figure 2.** 13R-manoyl oxide oxide-derived hydroxylated products formed following transient expression of CfCYP76AHs in *N. benthamiana* leaves. (A) Molecular formulas of the hydroxylated products obtained using different CfCYP76AHs. The number of hydroxylations of each compound was deduced from its accurate molecular mass (<5 ppm, [Supplementary file 1](#)) as determined by LC-qTOF-MS or NMR. Each different compound is marked by a number. (B) Chemical structures of the compounds marked with numbers in bold in A as determined by NMR ([Tables 1 and 2](#)). MO: manoyl oxide  
DOI: 10.7554/eLife.23001.007



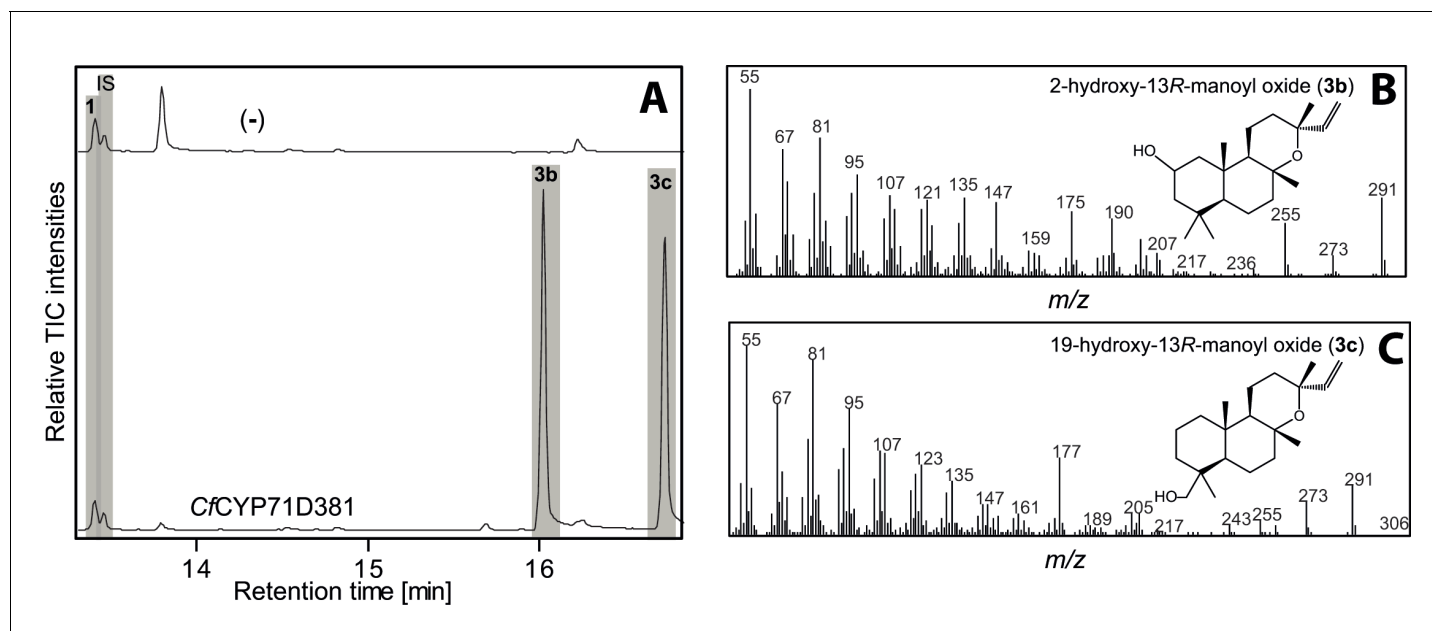
**Figure 3.** GC-MS analysis of 13R-manoyl oxide (1) derived diterpenoids obtained by transient expression of CYP76AHs from *C. forskohlii* in *N. benthamiana* leaves. **(A)** GC-MS total ion chromatograms (TIC) of extracts from *N. benthamiana* transiently expressing CfCXS, CfGGPPS, CfTPS2 and CfTPS3 (13R-manoyl oxide biosynthesis) genes in combination with water (-), CfCYP76AH15, CfCYP76AH17, CfCYP76AH8, CfCYP76AH11 or CfCYP76AH16. 1-eicosene was used as internal standard (IS). 13R-manoyl oxide (1) was identified only in (-), indicating that it is further metabolized by the CfCYP76AH15, CfCYP76AH17, CfCYP76AH8, CfCYP76AH11 and CfCYP76AH16 enzymes. **(B)**  $m/z$  spectrum of 11-oxo-13R-manoyl oxide (2). **(C)**  $m/z$  spectrum of 9-hydroxy-13R-manoyl oxide (3a). The structure of both compounds was verified by NMR analysis (Tables 1 and 2). The compounds have been identified previously in *C. forskohlii* as putative intermediates in the *in planta* biosynthesis of forskolin (Asada et al., 2012). For each combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.008](https://doi.org/10.7554/eLife.23001.008)



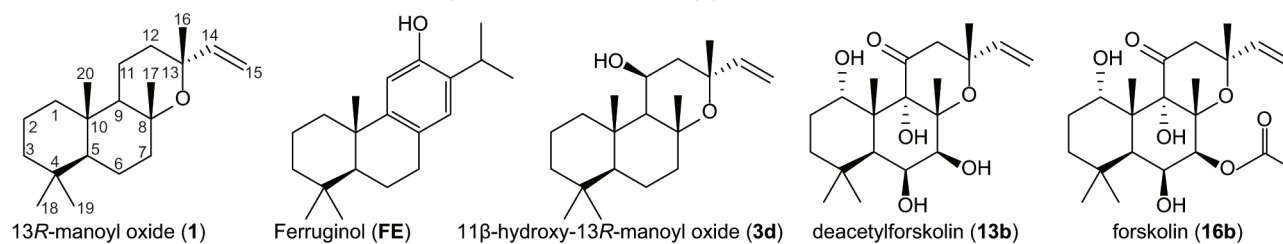
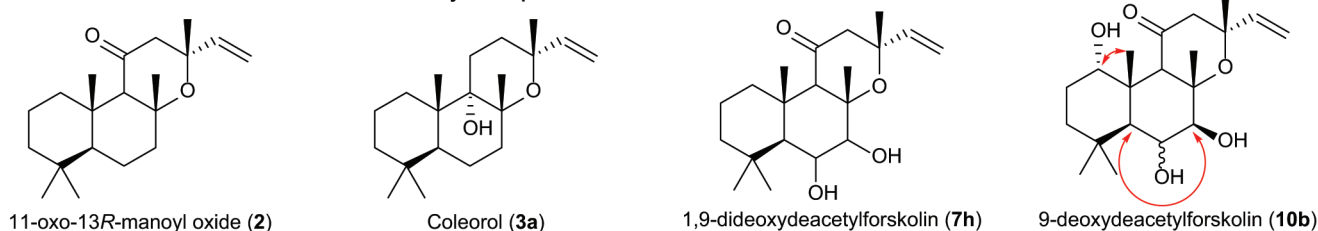
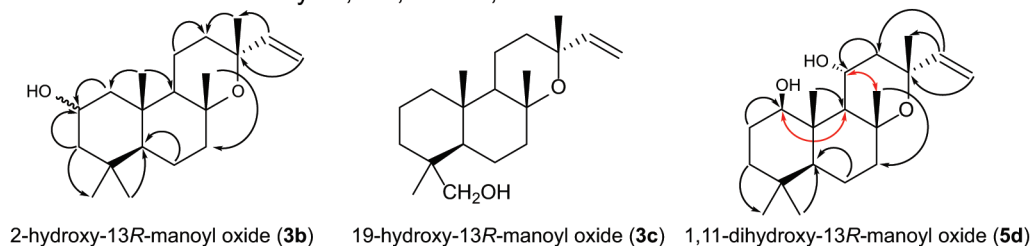
**Figure 3—figure supplement 1.** LC-qTOF-MS analysis of 13R-manoyl oxide-derived diterpenoids obtained by transient expression of *C. forskohlii* CYP76AH encoding genes in *N. benthamiana* leaves. Total ion chromatograms (TIC) of extracts expressing the 13R-manoyl oxide biosynthesis genes (*CfCXS*, *CfGGPPS*, *CfTPS2*, *CfTPS3*) in combination with water (-), *CfCYP76AH8*, *CfCYP76AH17*, *CfCYP76AH15*, *CfCYP76AH11* or *CfCYP76AH16* are shown. 13R-manoyl oxide-derived oxygenated compounds formed (marked with grey bars) and their identity including their molecular formulas was confirmed by their accurate mass (5 ppm tolerance, **Supplementary file 1**). The identity of 1,11-dihydroxy-13R-manoyl oxide (**5d**) and 9-deoxydeactylforskolin (**10b**) was confirmed by NMR analysis (**Figure 4** and **Tables 1** and **2**). No 13R-manoyl oxide-derived diterpenoids were detected in the water control (-). For each combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.009](https://doi.org/10.7554/eLife.23001.009)



**Figure 3—figure supplement 2.** GC-MS analysis of 13R-manoyl oxide-derived diterpenoids following transient expression in *N. benthamiana* leaves of the *C. forskohlii* gene encoding CfCYP71D281 together with genes encoding the required enzymes for biosynthesis of 13R-manoyl oxide (CfCXs, CfGGPPS, CfTPS2, CfTPS3). (A) GC-MS total ion chromatograms (TIC) of extracts from *N. benthamiana* transiently expressing 13R-manoyl oxide biosynthesis genes in combination with water (-) or CfCYP71D381, respectively. 1-Eicosene was used as internal standard (IS) and 13R-manoyl oxide (1) was identified in both (-) and the CfCYP71D381 samples. Compounds 3b and 3c were identified in extracts from *N. benthamiana* leaves expressing CfCYP71D381 together with the genes in 13R-manoyl oxide biosynthesis. CfCYP71D381 efficiently converted compound 1 to a mixture of two mono-hydroxylated 13R-manoyl oxide derivatives (3b and 3c). Structural elucidation by NMR (Figure 4 and Tables 1 and 2) showed hydroxylation of 1 at positions C-2 (3b) and C-19 (3c). These hydroxylation positions do not coincide with those found in forskolin and to our knowledge have not been observed in other diterpenoids known from *C. forskohlii*. (B)  $m/z$  spectrum of 2-hydroxy-13R-manoyl oxide (3b). (C)  $m/z$  spectrum of 19-hydroxy-13R-manoyl oxide (3c). For each combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.010](https://doi.org/10.7554/eLife.23001.010)

**A** Molecule structures confirmed by authentic standard(s)**B** Molecule structures confirmed by comparison of  $C^{13}$  chemical shift with reference values**C** Molecule structures obtained by  $H^1$ ,  $C^{13}$ , HMBC, and NOESY measurements

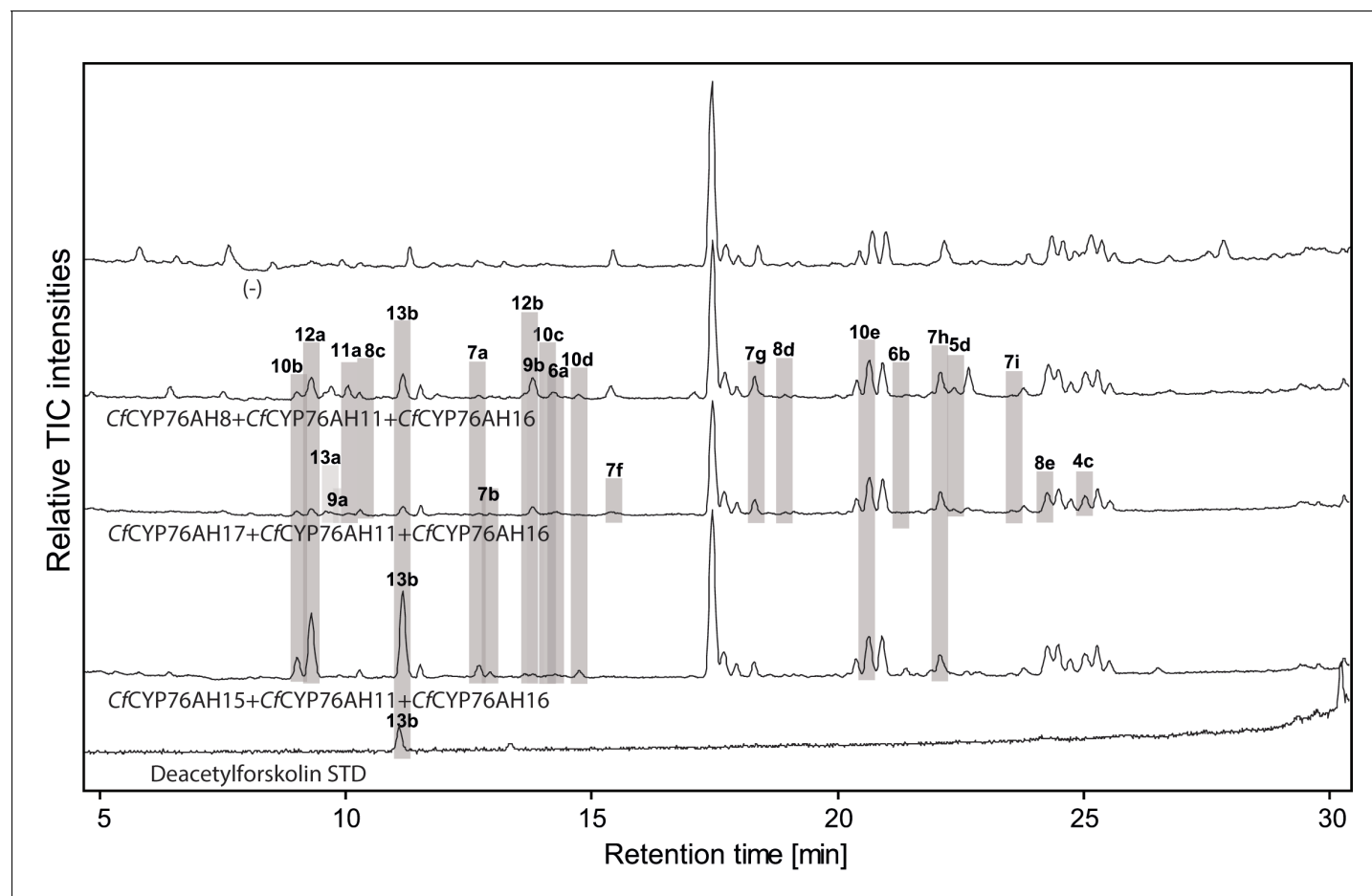
**Figure 4.** Structures of key compounds presented in this work. (A) Compounds confirmed using authentic standards. (B) Compounds which structure was confirmed/identified by comparison of  $^{13}C$  NMR data with existing literature. (C) Compounds which structure was confirmed/identified by HMBC and NOE correlations for assigning position of OH-groups (marked in red), whereas couplings identified in the previously uncharacterized compounds 3b, 3c and 5d are marked in black. All other molecular structures were confirmed by  $^{13}C$  chemical shifts in comparisons to reference values (Table 1, Figure 4—source data 1).

DOI: [10.7554/eLife.23001.011](https://doi.org/10.7554/eLife.23001.011)

The following source data is available for figure 4:

**Source data 1.** NMR spectra's of selected 13R-manoyl oxide derived molecules.

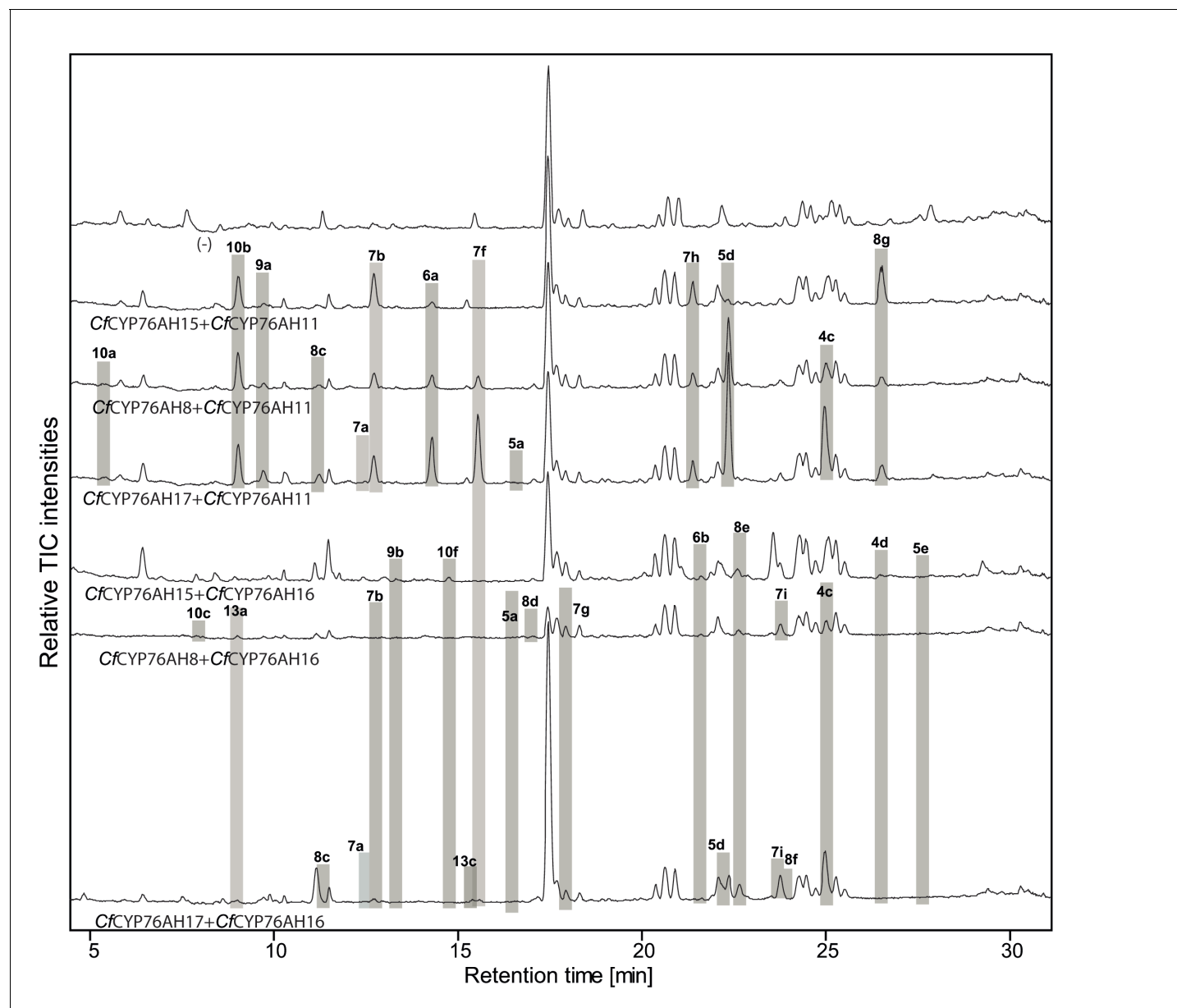
DOI: [10.7554/eLife.23001.012](https://doi.org/10.7554/eLife.23001.012)



**Figure 5.** LC-qTOF-MS analysis of 13R-manoyl oxide-derived diterpenoids obtained by transient expression of combinations of *C. forskohlii* CYP encoding genes, together with genes encoding the required enzymes for biosynthesis of 13R-manoyl oxide in *N. benthamiana* leaves. Total ion chromatograms (TIC) of extracts expressing the 13R-manoyl oxide biosynthesis genes (*CfCXS*, *CfGGPPS*, *CfTPS2*, *CfTPS3*), in combination with (from the top) water (-), *CfCYP76AH8* + *CfCYP76AH11* + *CfCYP76AH16*, *CfCYP76AH17* + *CfCYP76AH11* + *CfCYP76AH16*, and *CfCYP76AH15* + *CfCYP76AH11* + *CfCYP76AH16* are shown. Hydroxylated 13R-manoyl oxide-derived diterpenoids (marked with grey bars) and their identity including their molecular formulas were confirmed by accurate mass (5 ppm tolerance, **Supplementary file 1**). Compounds present in trace amounts are not marked. The identity of 1,11-dihydroxy-13R-manoyl oxide (**5d**), 9-deoxydeacetylforskolin (**10b**) and 1,9-dideoxydeacetylforskolin (**7h**) was confirmed by NMR analysis (**Figure 4** and **Tables 1** and **2**), whereas the identity of deacetylforskolin (**13b**) was confirmed by comparison to an authentic chemically synthesized standard. No 13R-manoyl oxide-derived diterpenoids were identified in the water control (-). For each combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

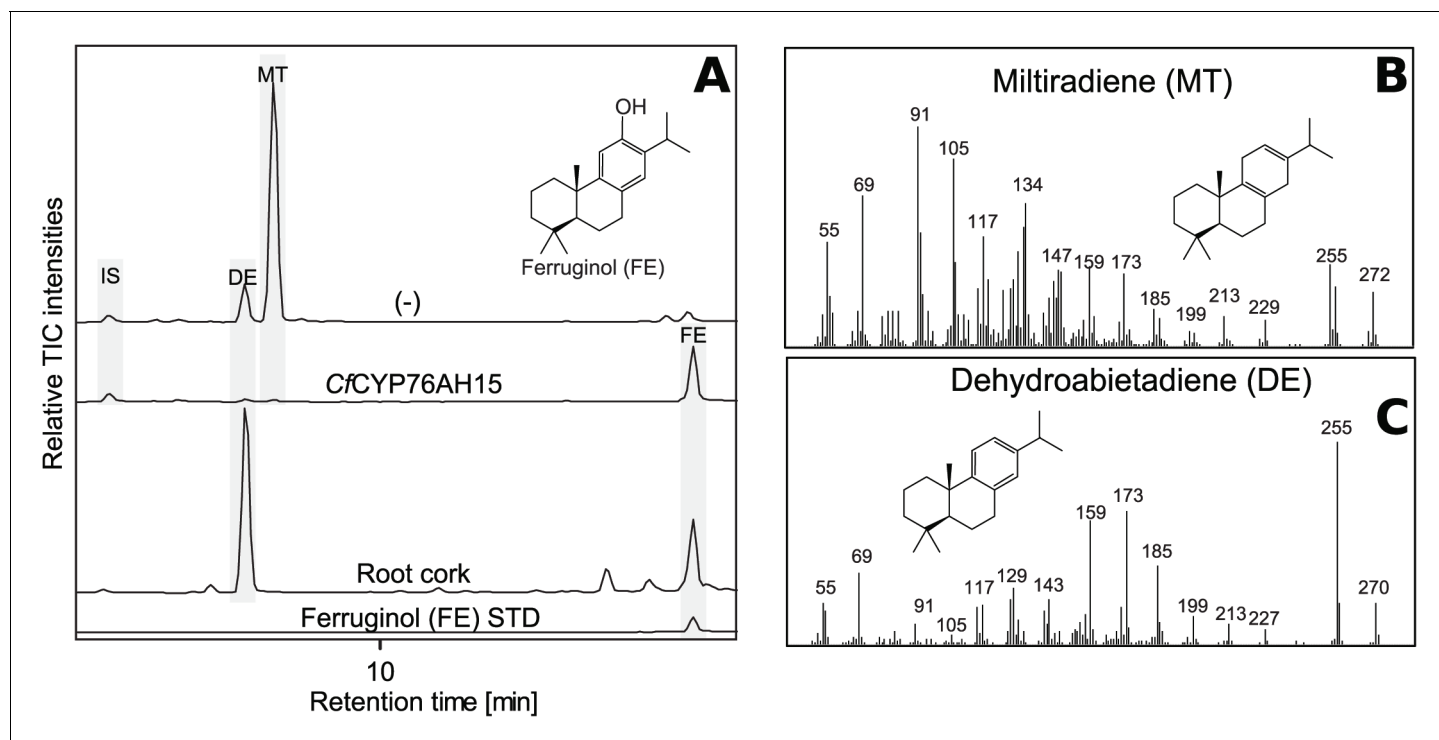
DOI: [10.7554/eLife.23001.015](https://doi.org/10.7554/eLife.23001.015)





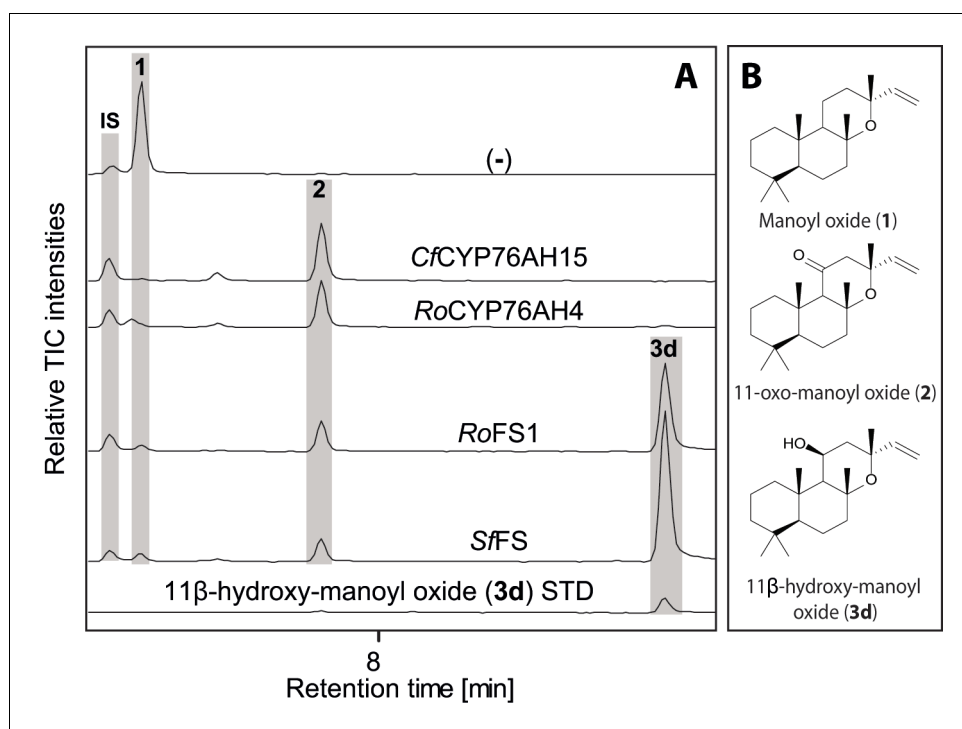
**Figure 5—figure supplement 1.** LC-qTOF-MS analysis of 13R-manoyl oxide-derived diterpenoids obtained by transient expression of combinations of *C. forskohlii* CYP76AH encoding genes in *N. benthamiana* leaves. Total ion chromatograms (TIC) of extracts expressing the 13R-manoyl oxide biosynthesis genes (*CfCXs*, *CfGGPPS*, *CfTPS2*, *CfTPS3*) in combination with (from the top) water (-), *CfCYP76AH15* + *CfCYP76AH11*, *CfCYP76AH8* + *CfCYP76AH11*, *CfCYP76AH17* + *CfCYP76AH11*, *CfCYP76AH15* + *CfCYP76AH16*, *CfCYP76AH8* + *CfCYP76AH16* and *CfCYP76AH17* + *CfCYP76AH16* are shown. Oxygenated 13R-manoyl oxide-derived diterpenoids (marked with grey bars) and their identity including their molecular formulas were confirmed by their accurate mass (5 ppm tolerance, **Supplementary file 1**). Compounds present in trace amounts are not marked. The identity of 1,11-dihydroxy-13R-manoyl oxide (**5d**), 9-deoxydeacetylforskolin (**10b**), 1,9-dideoxydeacetylforskolin (**7h**) was confirmed by NMR analysis (**Figure 4** and **Tables 1** and **2**). No 13R-manoyl oxide-derived diterpenoids were detected in the water control (-). For each combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.016](https://doi.org/10.7554/eLife.23001.016)



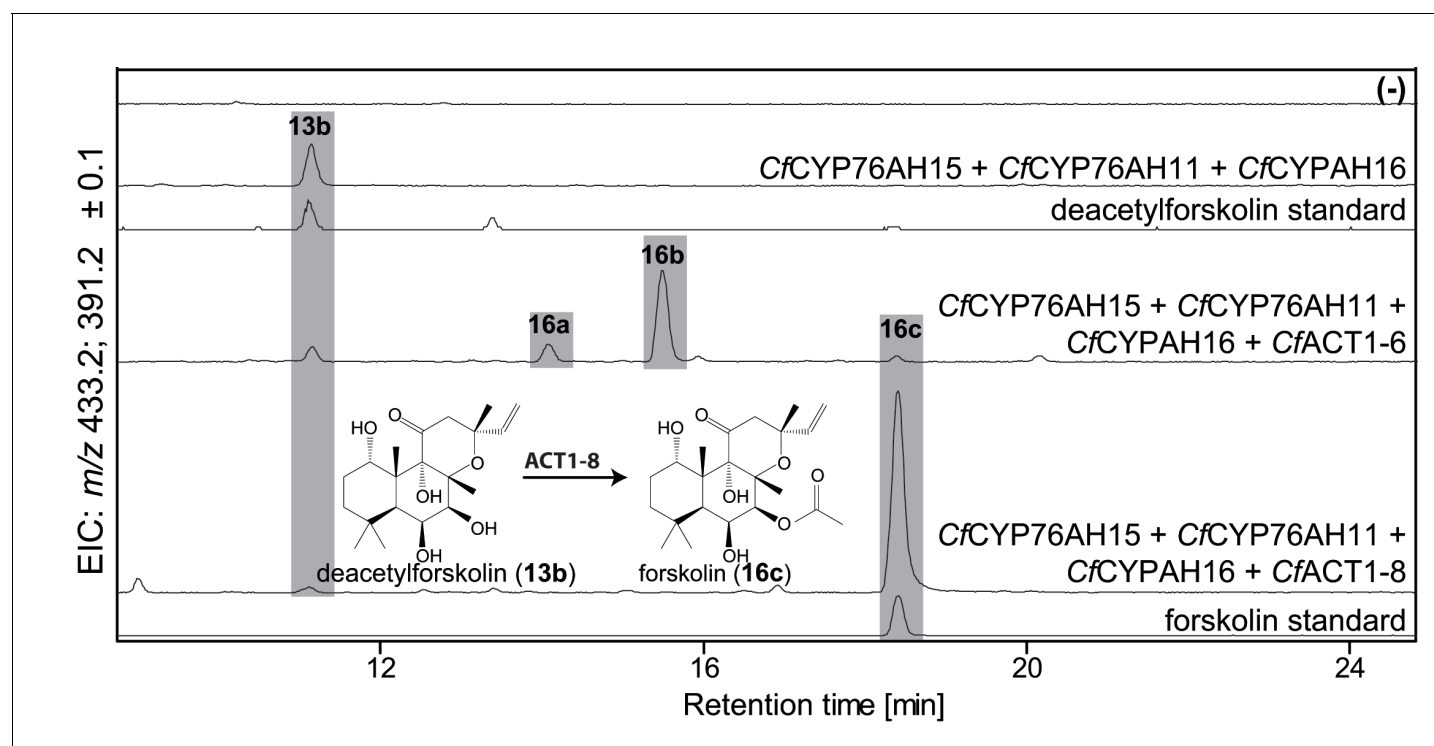
**Figure 6.** GC-MS analysis of miltiradiene-derived diterpenoids obtained by transient expression of *CfCYP76AH15* in *N. benthamiana* leaves. (A) Total ion chromatograms (TIC) of extracts transiently expressing *CfCXS*, *CfGGPPS*, *CfTPS1* and *CfTPS3* (miltiradiene biosynthesis genes) in combination with water (-) or *CfCYP76AH15*. Dehydroabietadiene (DE) and miltiradiene (MT) were observed in the (-) extract, whereas ferruginol was observed in extracts from tissue expressing the miltiradiene biosynthesis genes together with *CfCYP76AH15*. In root cork extracts, ferruginol was detected together with dehydroabietadiene. Presence of ferruginol was confirmed by comparison to an authentic standard (Ignea et al., 2016a), while presence of miltiradiene (B) and dehydroabietadiene (C) were confirmed by comparison of *m/z* spectra with previously characterized compounds (Andersen-Ranberg et al., 2016). For every combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: 10.7554/eLife.23001.017



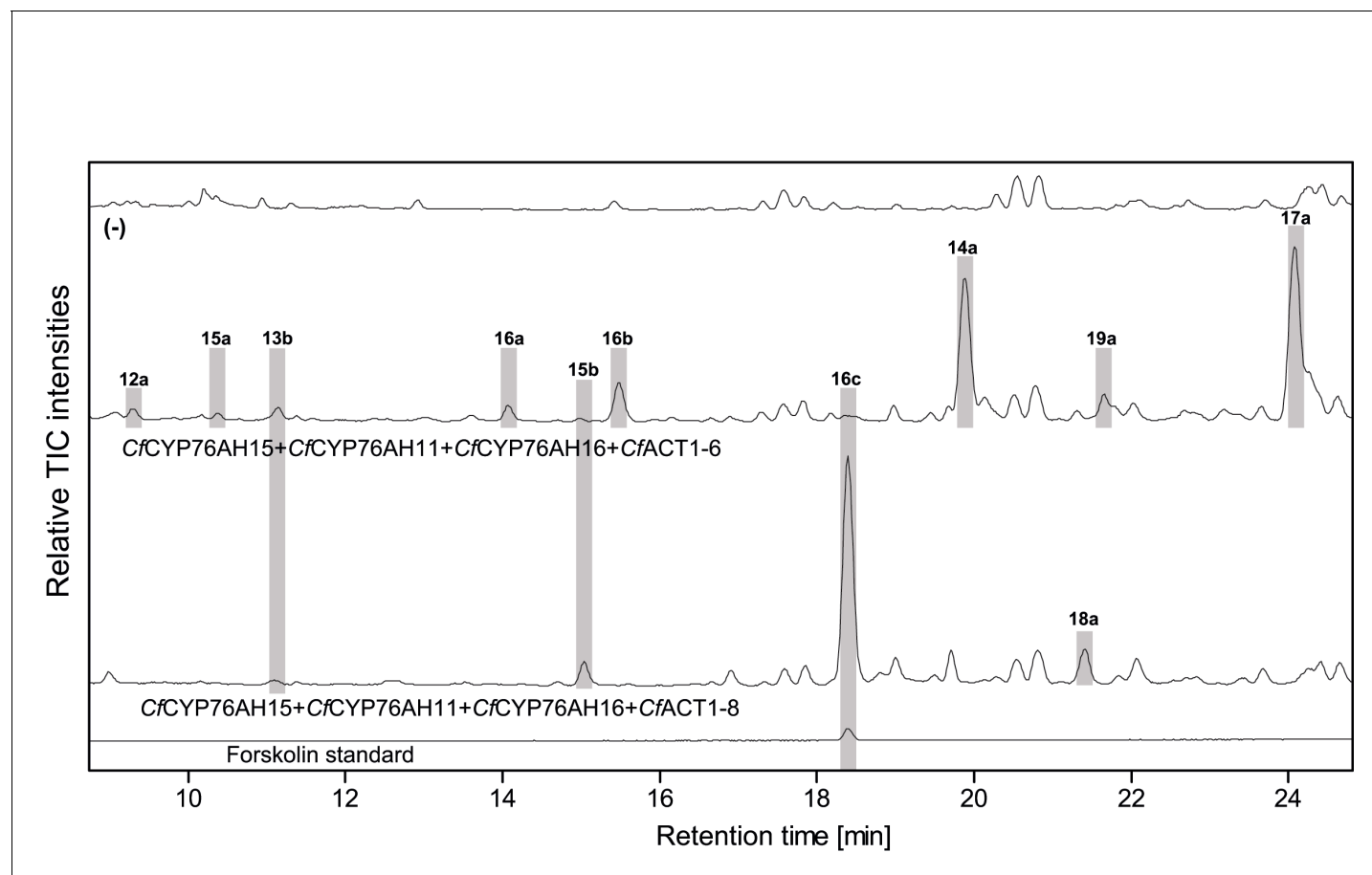
**Figure 7.** GC-MS analysis of 13R-manoyl oxide-derived diterpenoids obtained by transient expression of CYP76AHs in *N. benthamiana* leaves. **(A)** Total ion chromatograms (TIC) of extracts transiently expressing CfCXS, CfGGPPS, CfTPS2 and CfTPS3 (13R-manoyl oxide biosynthesis genes) in combination with water (-), CfCYP76AH15, RoCYP76AH4, RoFS1 and SpFS are shown. 13R-manoyl oxide was observed in the (-) extracts, while 11-oxo-13R-manoyl oxide (2) was observed in the CfCYP76AH15, RoCYP76AH4, RoFS1 and SfFS extracts. 11-Hydroxy-13R-manoyl oxide (3d) is observed only in extracts expressing the RoFS1 and SfFS1 genes. Presence of 11-hydroxy-13R-manoyl oxide was verified by comparison to an authentic standard (Ignea et al., 2016b) while identification of 11-oxo-13R-manoyl oxide was confirmed by comparison of *m/z* spectra with a previously characterized compound (2). The results show RoCYP76AH4 has an activity similar to CfCYP76AH15, able to convert efficiently and specifically 13R-manoyl oxide to 2. RoFS1, as well as SfFS, can also convert 13R-manoyl oxide to 2 but they catalyze the synthesis of an additional product, 11-hydroxy-13R-manoyl oxide (3d). For every combination, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: 10.7554/eLife.23001.018



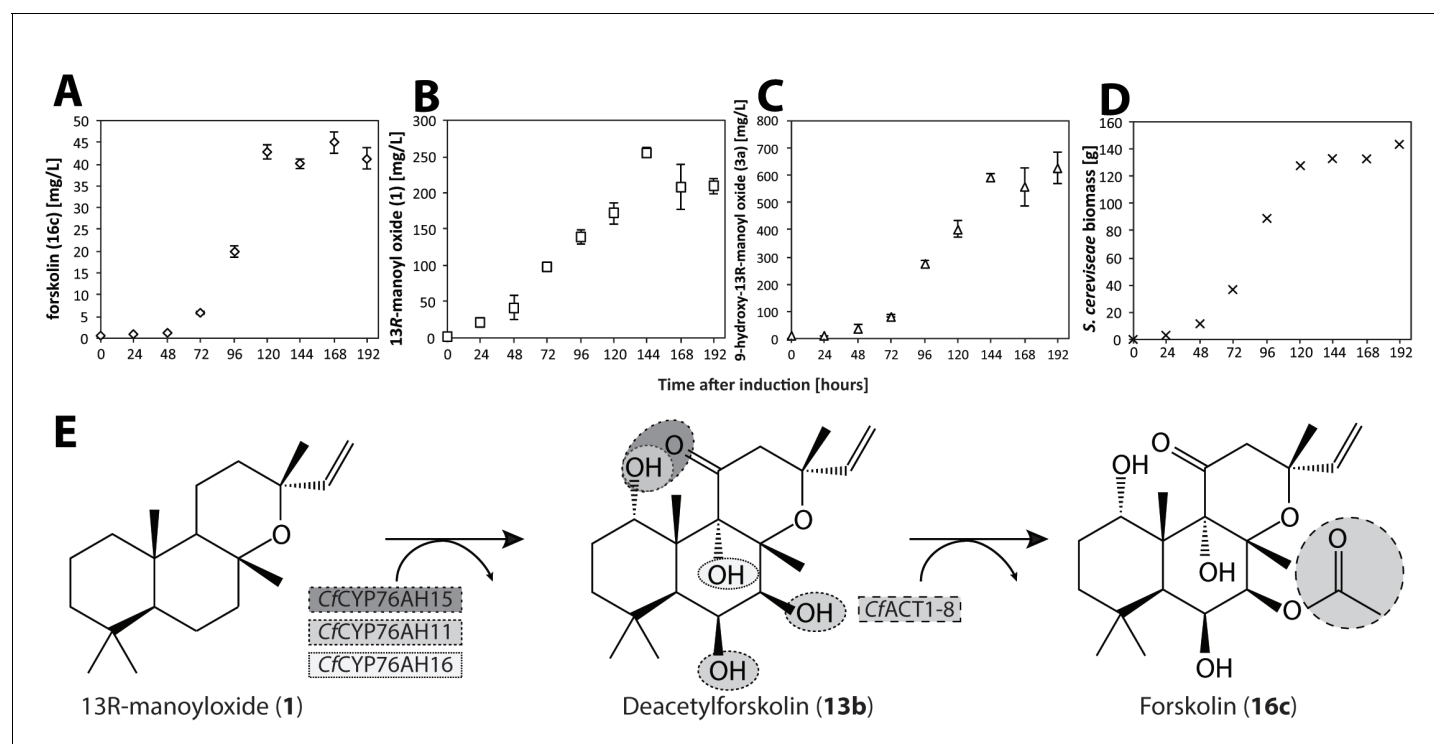
**Figure 8.** De novo biosynthesis of forskolin by transient expression of *C. forskohlii* genes in *N. benthamiana* as monitored by LC-MS-based extracted ion chromatograms (EIC). To monitor both deacetylforskolin (13b) and forskolin (16c), the EIC were selected as the sum of  $m/z$  433.2  $\pm$  0.1 and  $m/z$  391.2  $\pm$  0.1. Chromatograms represent LC-MS analysis of extracts from leaves expressing the 13R-manoyl oxide biosynthesis genes (*CfDXS*, *CfGGPPS*, *CfTP2* and *CfTPS3*) in combination with (from the top): water (-); *CfCYP76AH15*, *CfCYP76AH11* and *CfCYP76AH16*; *CfCYP76AH15*, *CfCYP76AH11*, *CfCYP76AH16* and *CfACT1-6*; *CfCYP76AH15*, *CfCYP76AH11*, *CfCYP76AH16* and *CfACT1-8*, shown together with authentic standards (deacetylforskolin and forskolin). Forskolin (16c) was identified together with two other acetylated compounds (e.g. 16a, 16b) with the same molecular mass in leaves expressing *CfACT1-6* together with forskolin-specific CYPs (**Supplementary file 1**). When *CfACT1-8* was expressed instead of *CfACT1-6*, a predominant accumulation of forskolin was observed, with a drastic reduction of non-specific acetylated products. For all combinations, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.019](https://doi.org/10.7554/eLife.23001.019)



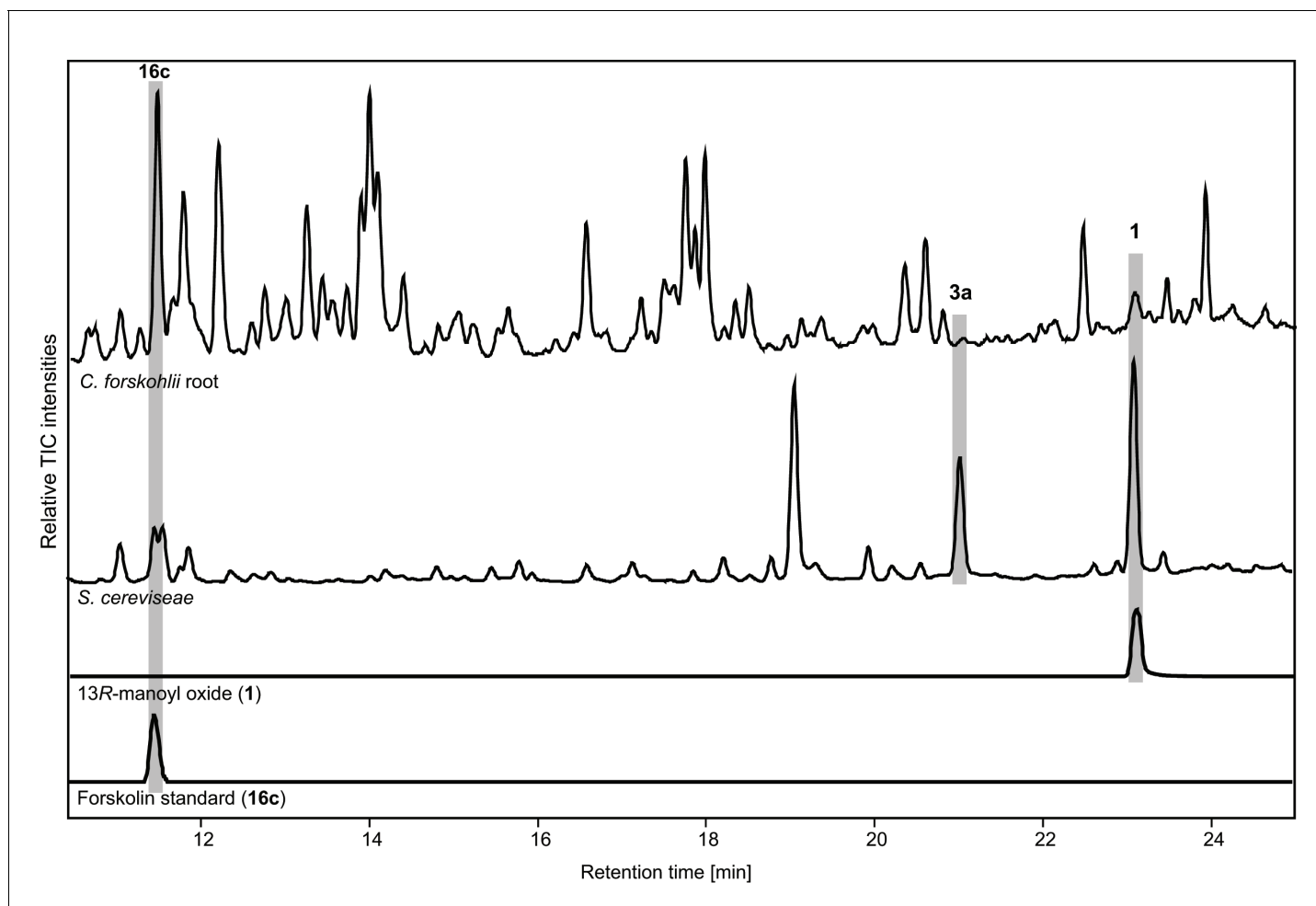
**Figure 9.** LC-qTOF-MS analysis of 13*R*-manoyl oxide-derived diterpenoids obtained by transient expression of combinations of *C. forskohlii* CYP and ACT encoding genes in *N. benthamiana* leaves. Total ion chromatograms (TIC) from extracts expressing the 13*R*-manoyl oxide biosynthesis genes (*CfCXS*, *CfGGPPS*, *CfTPS2*, *CfTPS3*) in combination with (from the top) water (-), *CfCYP76AH15* + *CfCYP76AH11* + *CfCYP76AH16* + *CfACT1-6*, and *CfCYP76AH15* + *CfCYP76AH11* + *CfCYP76AH16* + *CfACT1-8* are shown. Hydroxylated and acetylated 13*R*-manoyl oxide-derived diterpenoids (marked with grey bars) and their identity, including their molecular formulas, was confirmed by their accurate mass (5 ppm tolerance, **Supplementary file 1**). Compounds present in trace amounts are not marked. The identity of deacetylforskolin (**13b**) and forskolin (**16c**) was confirmed by comparison to authentic standards. No 13*R*-manoyl oxide-derived diterpenoids were detected in the water control (-). For all combinations, extracts from leaves of three different *N. benthamiana* plants have been analyzed and representative chromatograms are shown.

DOI: [10.7554/eLife.23001.020](https://doi.org/10.7554/eLife.23001.020)



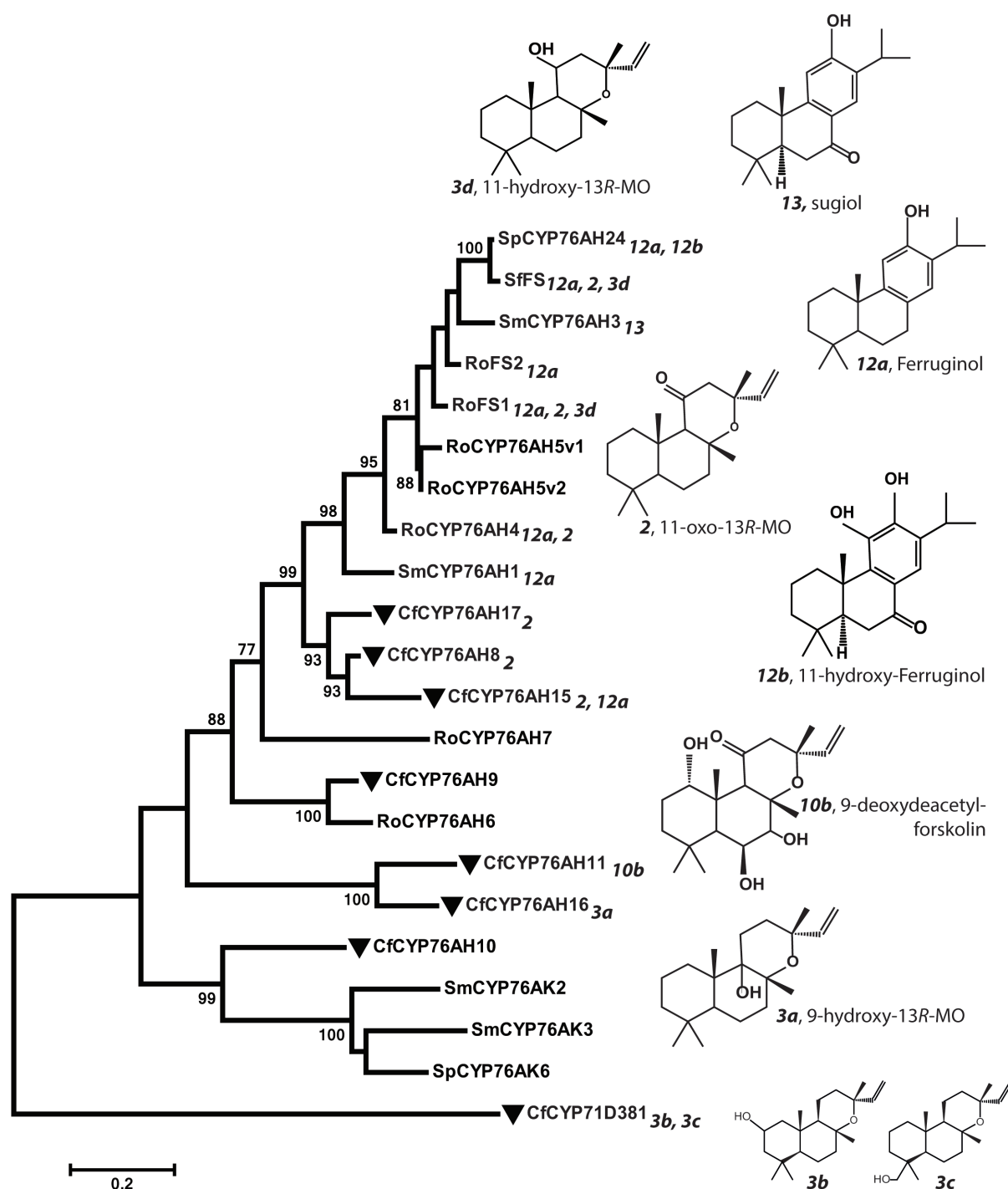
**Figure 10.** Forskolin production in *S. cerevisiae* following stable genomic integration of codon-optimized *C. forskohlii* genes. (A) Forskolin (16c) accumulation in a fermenter batch using the EVST21543 strain (expressing *CfCYP76AH15*, *CfCYP76AH11*, *CfCYP76AH16* and *CfACT1-8* encoding genes in the EFSC4498 *S. cerevisiae* strain, optimized for the production of 13R-manoyl oxide [Andersen-Ranberg et al., 2016]). (B) 13R-manoyl oxide (1) accumulation in EVST21543 strain. (C) 9-Hydroxy-13R-manoyl oxide (3a) accumulation in EVST21543 strain. (D) EVST21543 strain biomass monitored during the fermentation process. (E) The biosynthetic pathway used for the production of forskolin in yeast. The fermentation event occurred once, and a triplicate of samples were analysed from each time course.

DOI: [10.7554/eLife.23001.021](https://doi.org/10.7554/eLife.23001.021)



**Figure 10—figure supplement 1.** Comparison of metabolite profiles between fermenter grown yeast culture of the EVST21543 strain and *C. forskohlii* root extract analyzed by LC-MS. Forskolin (**16c**) and 13*R*-manoyl oxide (**1**) were identified based on co-elution with standards and 9-hydroxy-13*R*-manoyl oxide (**3a**) was identified based on the presence of the  $[M+Na]^+$  ion, 329.2457 ( $C_{20}H_{34}O_2Na^+$ ,  $\Delta$  0.7 ppm).

DOI: [10.7554/eLife.23001.022](https://doi.org/10.7554/eLife.23001.022)



**Figure 11.** Phylogeny of known full-length CYP76AHs. The enzymes used are listed below with their accession numbers or source of publication: CfCYP76AH15, KT382358; CfCYP76AH17, KT382360; CfCYP76AH8, KT382348; CfCYP76AH11, KT382349; CfCYP76AH16, KT382359; CfCYP76AH9, KT382347; CfCYP76AH10, KT382346; CfCYP71D381, KT382342; RoFS1, AJQ30187 (Božić et al., 2015); SmCYP76AH3, KR140168 (Guo et al., 2016); RoFS2, AJQ30188 (Božić et al., 2015); SfFS, AJQ30186 (Božić et al., 2015); RoCYP76AH4, (Zi and Peters, 2013); RoCYP76AH5v1, (Zi and Peters, 2013); RoCYP76AH5v2, (Zi and Peters, 2013); RoCYP76AH6, (Zi and Peters, 2013); RoCYP76AH7, (Zi and Peters, 2013); SmCYP76AH1, AGN04215

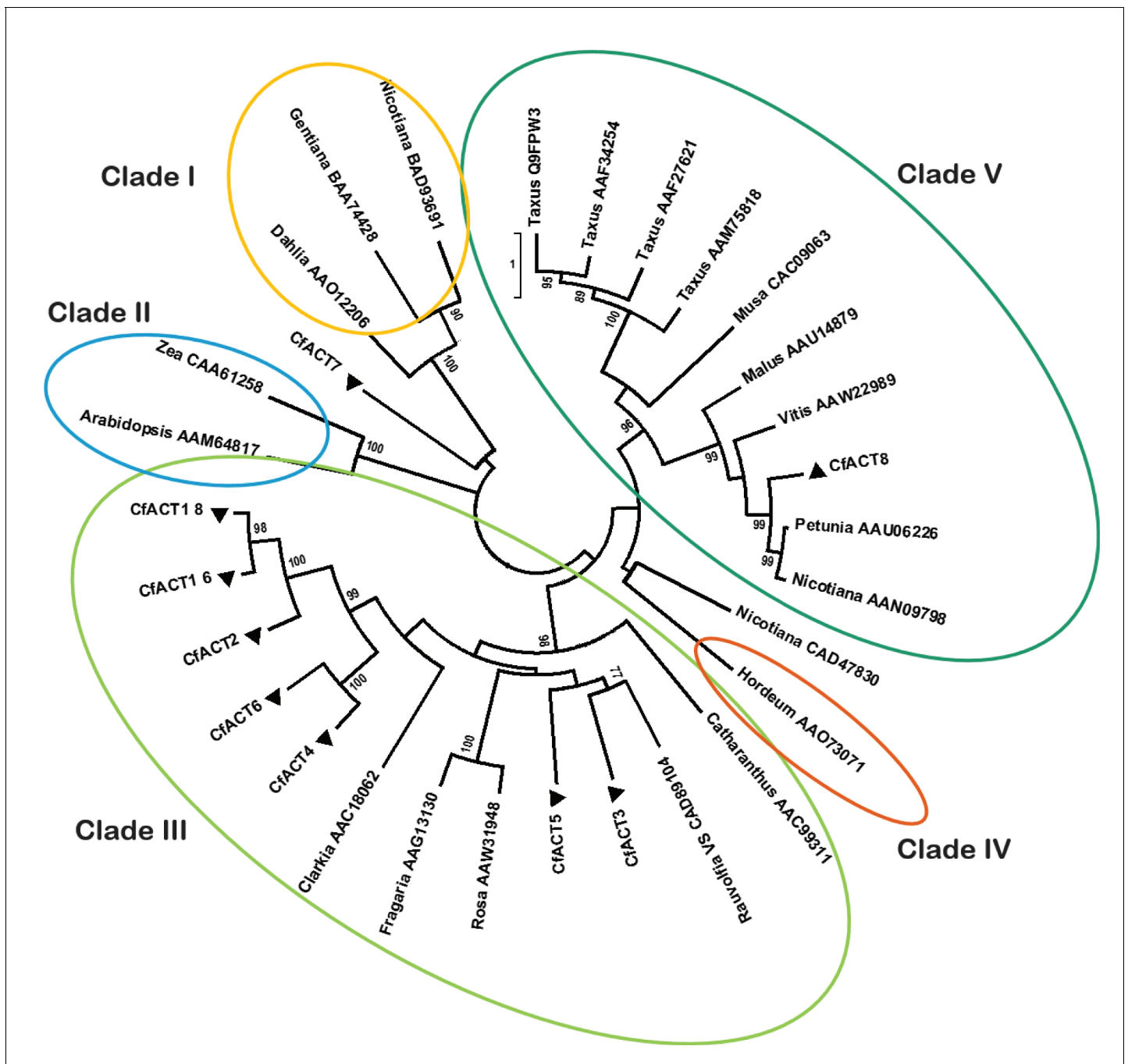
Figure 11 continued on next page



Figure 11 continued

(Guo et al., 2013); SpCYP76AH24, ALM25796 (Igneal et al., 2016a). *Coleus forskohlii* enzymes are indicated by a solid black triangle. CfCYP71D381 was chosen as a root because it can accept 13R-manoyl oxide as a substrate, but does not catalyze the synthesis of forskolin-related products. The number subscripts indicated at each enzyme refer to their respective enzymatic products, the structures of which are given on the right. Only the main products of each enzymes are mentioned. MO stands for manoyl oxide.

DOI: [10.7554/eLife.23001.023](https://doi.org/10.7554/eLife.23001.023)



**Figure 12.** Phylogenetic tree of CfACT encoding candidate genes together with BAHD family acyltransferase representatives from all clades according to *D'Auria (2006)*. Accession numbers of the non-*Coleus forskohlii* selected protein sequences are shown next to the tree taxon names, while *C. forskohlii* peptide accession numbers are provided in *Figure 1—source data 1*. The analysis only includes functionally characterized members. *Coleus forskohlii* enzymes are indicated by a solid black triangle. The majority of the selected CfACTs belong to Clade III, which includes mainly members which accept a diverse range of hydroxylated substrates and use acetyl-CoA as the main acyl donor (*D'Auria, 2006*). Interestingly, the ACTs known to be involved in Taxol biosynthesis belong to Clade V.

DOI: [10.7554/eLife.23001.024](https://doi.org/10.7554/eLife.23001.024)