Plant Physiology®

Plant Physiology, 2024, 196, 1407–1425
https://doi.org/10.1093/plphys/kiae399
Advance access publication 25 July 2024
Research Article

Roles of three cytochrome P450 monooxygenases in triterpene biosynthesis and their potential impact on growth and development

Caiqiong Yang,[1,2,*] Rayko Halitschke,[3] Sarah E. O'Connor,[2] Ian T. Baldwin[1,*]

[1]Department of Molecular Ecology, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, Jena D-07745, Germany
[2]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, Jena D-07745, Germany
[3]Mass Spectrometry and Metabolomics, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, Jena D-07745, Germany
*Author for correspondence: baldwin@ice.mpg.de (I.T.B.), cyang@ice.mpg.de (C.Y.).
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/General-Instructions) is: Caiqiong Yang (cyang@ice.mpg.de).

Abstract

Pentacyclic triterpenoids, recognized for their natural bioactivity, display complex spatiotemporal accumulation patterns within the ecological model plant *Nicotiana attenuata*. Despite their ecological importance, the underlying biosynthetic enzymes and functional attributes of triterpenoid synthesis in *N. attenuata* remain unexplored. Here, we show that 3 cytochrome P450 monooxygenases (NaCYP716A419, NaCYP716C87, and NaCYP716E107) from *N. attenuata* oxidize the pentacyclic triterpene skeleton, as evidenced by heterologous expression in *Nicotiana benthamiana*. NaCYP716A419 catalyzed a consecutive 3-step oxidation reaction at the C28 position of β-amyrin/lupeol/lupanediol, yielding the corresponding alcohol, aldehyde, and carboxylic acid. NaCYP716C87 hydroxylated the C2α position of β-amyrin/lupeol/lupanediol/erythrodiol/oleanolic acid/betulinic acid, while NaCYP716E107 hydroxylated the C6β position of β-amyrin/oleanolic acid. The genes encoding these 3 CYP716 enzymes are highly expressed in flowers and respond to induction by ABA, MeJA, SA, GA_3, and abiotic stress treatments. Using VIGS technology, we revealed that silencing of *NaCYP716A419* affects the growth and reproduction of *N. attenuata*, suggesting the ecological significance of these specialized metabolite biosynthetic steps.

Introduction

Triterpenoids, a class of isoprenoid compounds, are universally present in all eukaryotic organisms. They constitute a highly diverse group of polycyclic molecules with a wide range of biological functions, spanning both primary and secondary metabolism. These compounds are elaborated not only in medicinal plants, like *Glycyrrhiza glabra* and *Panax vietnamensis* (Dou et al. 2001; Tran et al. 2001; Seki et al. 2008, 2011), but are also commonly found in crops such as legumes and oats (Geisler et al. 2013; Hu et al. 2021; Yu et al. 2022). Triterpenoids often serve as integral components of plant defense mechanisms, playing central roles in direct defense responses (González-Coloma et al. 2011; Leveau et al. 2019). Their frequent inclusion in traditional medicinal practices underscores their long-recognized pharmacological significance (Wang 2021; Reed et al. 2023). Furthermore, triterpene saponins are extensively used in various food, cosmetics, and pharmaceutical industrial sectors (Güçlü-Üstündağ and Mazza 2007; Sharma et al. 2023). Therefore, elucidation of the potential pathways for triterpenoid biosynthesis at a molecular level can facilitate the engineering of these pathways into heterologous hosts or enhance their production in native hosts to improve crop resilience.

Triterpene biosynthesis begins with the acetylation of coenzyme A (Co-A) and proceeds through the mevalonic acid pathway, involving various oxidative cyclizations catalyzed by oxidosqualene cyclases (OSCs). These OSCs produce many triterpene structure scaffolds that range from acyclic to polycyclic compounds. Among these, pentacyclic triterpene skeletons are the most common. The currently identified OSCs are involved in the synthesis of pentacyclic triterpenes include β-amyrin and lupeol synthase (Khakimov et al. 2015; Moses et al. 2015a; Jo et al. 2017), and some are multifunctional OSCs reported to synthesize unique triterpene scaffolds, such as lupanediol, taraxasterol, δ-amyrin, germanicol, and others (Ito et al. 2011; Wang et al. 2011; Moses et al. 2015b; Andre et al. 2016; Srisawat et al. 2019). These triterpene scaffolds are subsequently decorated through modifications, including additions of hydroxyl, ketone, aldehyde, or carboxyl groups, and glycosylations.

Cytochrome P450 (CYP) monooxygenases perform several modifications of triterpene scaffolds that can occur at various positions (Ghosh 2017; Romsuk et al. 2022). To date, approximately 50 P450 enzymes have been reported to act on plant pentacyclic triterpene scaffolds, and the majority belongs to the CYP716 family. Additionally, some members of the CYP51, CYP71, CYP72, CYP87, CYP88, and CYP93 families have also been reported to modify pentacyclic triterpenes (Ghosh 2017). The biochemical properties of multiple members of the CYP716 family have been characterized (Carelli et al. 2011; Moses et al. 2015b; Miettinen et al. 2017; Misra et al. 2017; Yasumoto et al. 2017; Sandeep et al. 2019; Romsuk et al. 2022). Most catalyze a series of 3 consecutive oxidation reactions at the C28 position of β-amyrin/lupeol scaffolds, resulting in the sequential introductions of hydroxyl, aldehyde, and carboxyl groups at the C28 position (Carelli et al. 2011; Fukushima et al. 2011; Misra et al. 2017). Furthermore, certain plant CYP716A genes

Received May 10, 2024. Accepted July 6, 2024.

also catalyze oxidation reactions at carbon atoms other than C28 in β-amyrin scaffolds. For instance, in Arabidopsis (*Arabidopsis thaliana*), AtCYP716A2 catalyzes hydroxylation at C22α (Yasumoto et al. 2016), while in *Artemisia annua*, AaCYP716A14v2 catalyzes oxidations at C3 (Moses et al. 2015b). AcCYP716A111 from *Aquilegia coerulea* and PgCYP716A141 from *Platycodon grandiflorus* both catalyze hydroxylations at C16β (Miettinen et al. 2017). Members of the CYP716E subfamily, such as SlCYP716E26 from *Solanum lycopersicum* and CaCYP716E41 from *Centella asiatica*, have been biochemically characterized as C6β-hydroxylases that accept α/β-amyrin and oleanolic/ursolic/maslinic acids as substrates (Miettinen et al. 2017; Yasumoto et al. 2017). CaCYP716C11 from *C. asiatica* and OeCYP716C67 from *Olea europaea* catalyze C2α hydroxylations of oleanolic acid, 6β-hydroxy oleanolic acid, or ursolic acid (Miettinen et al. 2017; Alagna et al. 2023). Furthermore, GuCYP88D6 from licorice catalyzes C11 oxidations (Seki et al. 2008), while GmCYP93E3 catalyzes C24 hydroxylations of β-amyrin (Seki et al. 2008; Moses et al. 2014). In monocotyledonous plants such as oats, AsCYP51H10 is a multifunctional enzyme, catalyzing both hydroxylations and epoxidations of β-amyrin, resulting in the formation of 12,13β-epoxy-3β,16β-dihydroxy-oleanane (Geisler et al. 2013). Although the biosynthetic pathways of triterpenes have been elucidated in some plants, many of the enzymes responsible for triterpene structural modifications remain to be fully elucidated in most plants.

Nicotiana attenuata is a native annual wild tobacco species that grows in the Great Basin Desert, Utah, United States, and is a model for native plant–environment ecological interactions (Kessler and Baldwin 2004; Joo et al. 2021; Yang et al. 2023a; You et al. 2023). While triterpenoid compounds have been reported from *Nicotiana* species (Popova et al. 2018, 2019, 2020), our understanding of their biosynthesis remains rudimentary. Previously, we reported that *N. attenuata* inducibly accumulates pentacyclic triterpenoid compounds, primarily derived from lupane and oleanane scaffolds, in young plant organs or flowers. We identified NaOSC1 as being required for the biosynthesis of the triterpene scaffolds, lupeol, β-amyrin, lupanediol, dammarenediol II, and taraxasterol, while NaOSC2 predominantly synthesizes β-amyrin (Yang et al. 2023b). Here, we identify and characterize *N. attenuata* CYP enzymes from 6 distinct subfamilies: CYP716A (NaCYP716A419, NaCYP716A420), CYP716E (NaCYP716E107, NaCYP716E108), CYP716D (NaCYP716D93, NaCYP716D94), CYP716C (NaCYP716C87), CYP716H (NaCYP716H6), and CYP88B (NaCYP88B4). Through heterologous expression in *N. benthamiana*, we found NaCYP716A419, NaCYP716C87, and NaCYP716E107 to be triterpene modifying enzymes, with NaCYP716A419 catalyzing a consecutive 3-step oxidation reaction at the C28 position of β-amyrin, lupeol, and lupanediol; NaCYP716C87 hydroxylating the C2α position of β-amyrin, erythrodiol, oleanolic acid, lupeol, betulinic acid, and lupanediol; and NaCYP716E107 hydroxylating the C6β position of β-amyrin or oleanolic acid. Additionally, using tobacco rattle virus (TRV)-induced gene silencing (VIGS), we further characterized that *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* are potentially involved in *N. attenuata*'s growth and development.

Results

N. benthamiana-based in vivo screening of candidate P450 enzymes identifies 3 functional enzymes

Based on the CYP hidden Markov model, PF00067, we identified a total of 234 complete P450 cytochrome enzymes in *N. attenuata*, among which 10 were annotated as β-amyrin oxidases (Supplementary Fig. S1). Microarray data sourced from the *Nicotiana attenuata* Data Hub (http://nadh.ice.mpg.de/NaDH/) revealed that *NIATv7_g15098* and *NIATv7_g07976* were predominantly expressed in roots, *NIATv7_g01943* was primarily expressed in various floral organs, and *NIATv7_g06757* and *NIATv7_g33874* were mainly expressed in seeds, with some expression in floral organs. In contrast, *NIATv7_g18201* and *NIATv7_g15096* were expressed in both roots and floral organs, while *NIATv7_g17429*, *NIATv7_g33423*, and *NIATv7_g03976* were primarily expressed in stems and leaves (Fig. 1A). We conducted a phylogenetic analysis of these P450 enzymes in comparison to known P450 oxidases involved in triterpenoid biosynthesis (Supplementary Table S1). The results revealed that NIATv7_g01943 and NIATv7_g18201 clustered on the same branch as MtCYP716A12, known for C28 oxidation activity in *Medicago truncatula* (Fig. 1A). NIATv7_g01943 and NIATv7_g18201 share 87% and 85% amino acid sequence similarities with MtCYP716A12, respectively (Supplementary Fig. S2). NIATv7_g15096 and NIATv7_g15098 clustered with SlCYP716E26, which exhibits C6β oxidation activity (Fig. 1A). Their amino acid sequences were 62% and 63% identical to SlCYP716E26, with similarities of 80% and 81% (Supplementary Fig. S2). NIATv7_g06757 and NIATv7_g33874 clustered with GuCYP88D6 (Fig. 1A), and their amino acid sequences were 40% identical to GuCYP88D6, with similarities of 61% (Supplementary Fig. S2). NIATv7_g07976 clustered on the same branch with CaCYP716C11 (Fig. 1A) with 63% identity and 78% similarity (Supplementary Fig. S2).

The NIATv7_g03976, NIATv7_g33423, and NIATv7_g17429 displayed low sequence similarities of less than 40% when compared to known triterpene oxidases. These 10 candidate CYP450 enzymes were submitted to the P450 Nomenclature Committee for naming, and NIATv7_g01943 and NIATv7_g18201 were designated as NaCYP716A419 and NaCYP716A420, in the CYP716A subfamily. NIATv7_g15096 and NIATv7_g15098 were designated as NaCYP716E107 and NaCYP716E108, belonging to the CYP716E subfamily. NIATv7_g07976 was designated NaCYP716C87, a member of the CYP716C subfamily. NIATv7_g03976 was designated as NaCYP716H6 in the CYP716H subfamily. NIATv7_g33423 and NIATv7_g17429 were placed in the CYP716D subfamily and designated as NaCYP716D93 and NaCYP716D94, respectively. Lastly, NIATv7_g06757 and NIATv7_g33874 were placed in the CYP88C and CYP88B subfamilies and named NaCYP88C16 and NaCYP88B4, respectively (Supplementary Table S2).

Previously, our research identified the roles of *NaOSC1* and *NaOSC2* in the biosynthesis of triterpene scaffolds in *N. attenuata* (Yang et al. 2023b). We conducted Spearman's correlation analysis to examine associations among the expressions of *NaOSC1* and *NaOSC2* in different tissues and CYP450 candidates. Strong positive correlations were found between *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* with *NaOSC2* and between *NaCYP716C87* and *NaCYP88C16* with *NaOSC1* (Fig. 1B). Based on these correlations, we hypothesize that these CYP450 enzymes might function in the downstream pathway of triterpenes following NaOSC1 or NaOSC2, respectively. Except for *NaCYP88C16*, all were successfully cloned into the heterologous expression vector 3Ω1 (Cárdenas et al. 2019; Hong et al. 2022).

To determine the oxidation activity of these P450 enzymes toward simple triterpenes in *N. attenuata*, we conducted combination experiments in which these candidate genes were coexpressed with OSCs. In *N. benthamiana*, heterologous expression of *NaOSC2* results in abundant production of β-amyrin, while heterologous expression of *NaOSC1* yields a diverse array of products, including 11 compounds represented by lupeol and lupanediol. Due to the

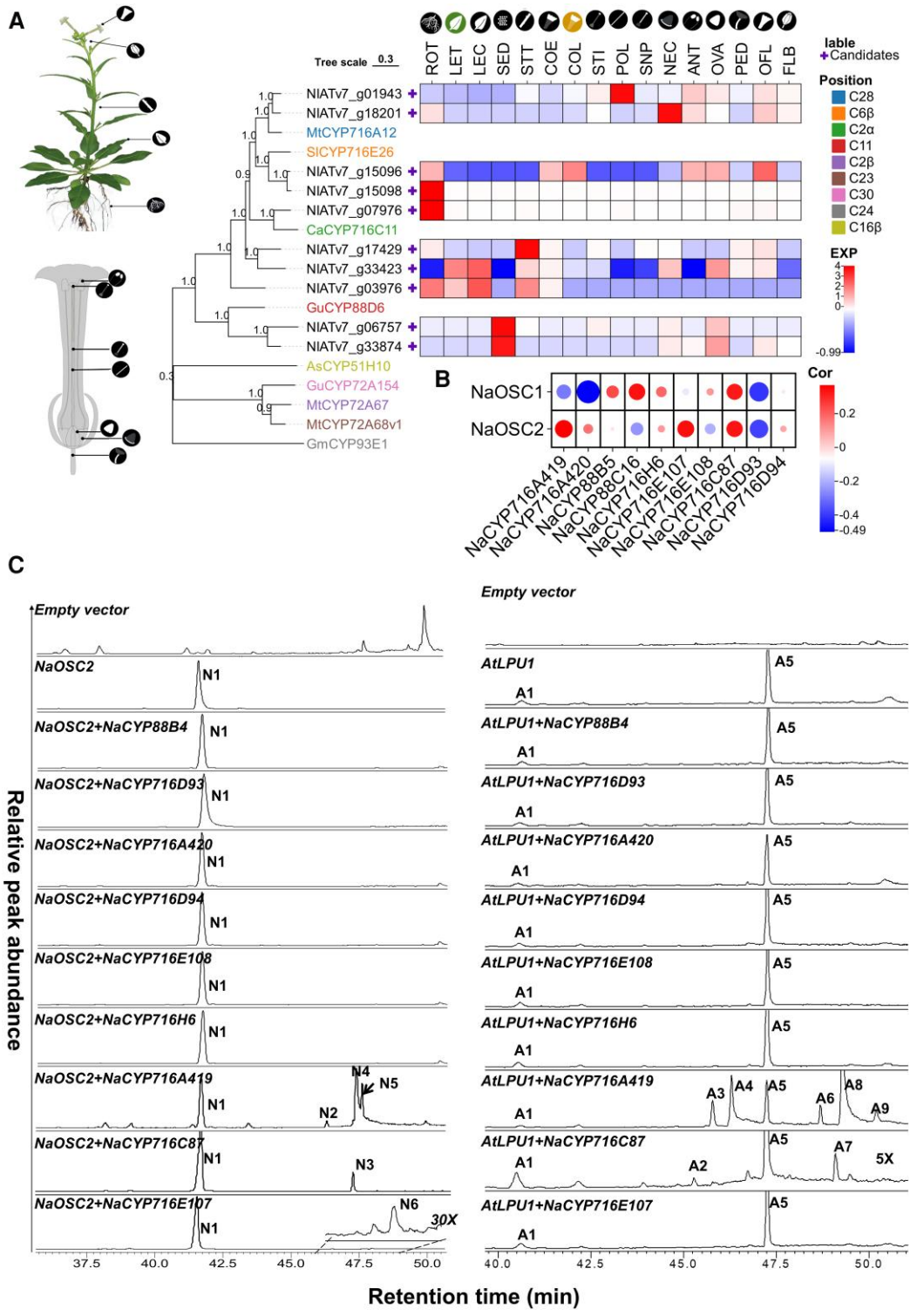

Figure 1. Screening of CYP enzymes from the triterpenoid biosynthesis pathway in *N. attenuata*. **A)** Phylogenetic tree and tissue-specific expression of candidate CYP enzymes. The protein sequences of known triterpenoid biosynthetic enzymes were obtained from the NCBI GeneBank, while the protein sequences and microarray data of candidate enzymes were sourced from the *Nicotiana attenuata* Data Hub (Brockmöller et al. 2017). ROT, root treated with *Manduca sexta* oral secretion; LET, leaf treated with *M. sexta* oral secretion; LEC, leaf control; SED, seed; STT, stem treated with *M. sexta* oral secretion; COE, corolla early; COL, corolla late; STI, stigma; POL, pollen tubes; SNP, style without pollination; ANT, anthers; NEC, nectaries; OVA, ovary; PED, pedicels; OFL, opening flower; FLB, flower buds. In the phylogenetic tree, the color-coded P450 enzymes represent those already reported to have triterpenoid modification functions. Exp, relative expression levels of genes. **B)** Correlation analysis of candidate genes with OSC tissue expression data. The microarray data used for correlation analysis were obtained from the *Nicotiana attenuata* Data Hub. The depth of color represents the correlation coefficient (Cor), while the data points' size indicates the significance level. **C)** GC-MS chromatograms of extracts of *N. benthamiana* leaves coexpressing CYP450 candidate genes and OSC. Selective ion monitoring (SIM) at *m/z*: 189, 203, 131, 216, and 320 for the combinations *AtLPU1* + *CYP* and *m/z* 189, 203, 216, and 320 for the combinations *NaOSC2* + *CYP*. "30x" and "5x" indicate that the chromatograms at this location have been zoomed in by a factor of 30 and 5, respectively. Triterpene compounds obtained by heterologous expression are labeled as peaks N1 to N6 (*NaOSC2* + *CYP*) and A1 to A9 (*AtLPU1* + *CYP*). Each treatment was conducted with at least 3 replications.

diverse and generally low levels of products obtained from the heterologous expression of NaOSC1 in *N. benthamiana* (Yang et al. 2023b), we opted to substitute NaOSC1 with an enzyme from *A. thaliana*, lupeol synthase 1 (AtLPU1), known to produce higher levels of lupeol and lupanediol when heterologously expressed in *N. benthamiana* (Segura et al. 2000).

Leaves expressing *NaOSC2* or *AtLUP1* were used as negative controls. Gas chromatography (GC)-MS analysis revealed distinct chromatographic peaks for NaCYP716C87, NaCYP716A419, and NaCYP716E107 compared to the control group (Fig. 1C; Supplementary Fig. S3). Specifically, the expression of *NaCYP716C87* and *NaCYP716A419* generated novel peaks with high yields when coexpressed with *NaOSC2* (peaks N2 to N5) or *AtLUP1* (peaks A2 to A4 and A6 to A9), whereas *NaCYP716E107* showed only weak activity when coexpressed with *NaOSC2* (peak N6) but not with *AtLUP1* (Fig. 1C; Supplementary Fig. S3).

NaCYP716A419 is a C28 oxidase

Comparisons of retention times and electron impact MS (EI-MS) spectra with authentic commercial standards revealed that the EI-MS and retention times of peaks N2 and N4 matched those of erythrodiol and oleanolic acid (Fig. 2A; Supplementary Table S3). Additionally, the EI-MS and retention times of peaks A3 and A4 matched those of betulin and betulinic acid (Fig. 2B; Supplementary Table S4). Due to the absence of reliable standards, our identification of oleanolic aldehyde (N5) is based on existing EI-MS data and the relative elution order of these triterpenes in GC (Misra et al. 2017) (Supplementary Table S3 and Fig. S4). In *N. benthamiana* leaves coexpressing *NaOSC2* and *NaCYP716A419*, 3 new peaks were observed (Fig. 2A), corresponding to 3 C28 oxidized products of β-amyrin (N1), namely erythrodiol (N2), putative oleanolic aldehyde (N5), and oleanolic acid (N4). Meanwhile, we also compared the triterpenoid profiles when *NaCYP716A419* was coexpressed with either empty vector (EV) or *NaOSC2* in *N. benthamiana* leaves, demonstrating that the β-amyrin substrate present in *N. benthamiana* leaf tissue does not yield detectable levels of erythrodiol (N2), oleanolic aldehyde (N5), and oleanolic acid (N4) (Supplementary Fig. S5A). Expression of the EV in *N. benthamiana* leaves and supplemented with erythrodiol (N2) as a substrate also did not result in detectable levels of oleanolic aldehyde (N5) and oleanolic acid (N4), indicating that oleanolic aldehyde (N5) and oleanolic acid (N4) are not products of endogenous enzymes in *N. benthamiana* leaves (Supplementary Fig. S5B).

In *N. benthamiana* leaves coexpressing *AtLUP1* and *NaCYP716A419*, 5 new peaks were observed (Fig. 2B), representing 2 C28 oxidized forms of lupeol (A1), betulin (A3), and betulinic acid (A4), as well as 3 putative C28 oxidized forms of lupanediol (A5), including putative compounds 28-hydroxy lupanediol (A6: (1R,3aS,5aR,5bR,9S,11aR)-3a-(hydroxymethyl)-1-(2-hydroxypropan-2-yl)-5a,5b,8,8,11a-pentamethylicosahydro-1H-cyclopenta[a]chrysen-9-ol), 28-aldehyde lupanediol (A9: (1R,3aS,5aR,5bR,9S,11aR)-9-hydroxy-1-(2-hydroxypropan-2-yl)-5a,5b,8,8,11a-pentamethylicosahydro-3aH-cyclopenta[a]chrysene-3a-carbaldehyde), and 28-carboxy lupanediol (A8: (1R,3aS,5aR,5bR,9S,11aR)-9-hydroxy-1-(2-hydroxypropan-2-yl)-5a,5b,8,8,11a-pentamethylicosahydro-3aH-cyclopenta[a]chrysene-3a-carboxylic acid). These identifications were based on characteristic ions in the EI-MS spectra and the oxidative function of NaCYP716A419 (Supplementary Table S4). From these data, we infer that NaCYP716A419 is an enzyme capable of catalyzing oxidation reactions at the C28 position of β-amyrin (N1), lupeol (A1), and lupanediol (A5) triterpene scaffolds, resulting in the formation of hydroxyl, aldehyde, and carboxyl groups (Fig. 2, A and B).

NaCYP716C87 is a C2α hydroxylase

The phylogenetic analysis (Fig. 1) revealed that NaCYP716C87 clustered with CaCYP716C11, a CYP monooxygenase from *C. asiatica* reported having C2α oxidation activity of oleanolic acid to produce the product, maslinic acid (Miettinen et al. 2017). The high degree of amino acid sequence identity (63%) and similarity (78%) suggested a similar function. To verify the C2α oxidation activity of NaCYP716C87, we heterologously expressed the *NaCYP716C87* in *N. benthamiana* leaves and fed them with the oleanolic acid substrate. The results revealed peak N7, with retention time and GC-MS spectrum consistent with the maslinic acid standard (Fig. 3A). When *NaOSC2*, *NaCYP716A419*, and *NaCYP716C87* are simultaneously expressed in *N. benthamiana*, a strong maslinic acid (N7) peak was detected as well as a notable decrease in oleanolic acid compared to leaves expressing only *NaOSC2* and *NaCYP716A419* (Fig. 3B; Supplementary Fig. S6). This indicates that NaCYP716C87 also functions as a C2α hydroxylase. Concurrently, coexpression of *NaOSC2*, known to possess β-amyrin synthase activity, with *NaCYP716C87* in *N. benthamiana* leaves resulted in the production of β-amyrin and another new peak, N3 (Figs. 1B and 3C). N3 exhibits *m/z* 218, 203, and 189, which is similar to the fragmentation pattern of β-amyrin (Fig. 3D; Supplementary Fig. S7). Moreover, N3 has *m/z* 277 and 235 that match the steps of AB ring loss ions after retro-Diels–Alder (RDA) cleavage of maslinic acid. Therefore, by comparing their EI-MS spectra and fragmentation pathways with those of β-amyrin and maslinic acid (Supplementary Figs. S7 and S8; Fig. 3D), N3 was inferred to be putative 2α-hydroxyl β-amyrin. Moreover, a new peak, N8, appeared when erythrodiol was fed to *N. benthamiana* leaves expressing *NaCYP716C87* (Fig. 3C). By comparing its EI-MS spectra with those of maslinic acid and erythrodiol (Supplementary Figs. S8 and S9; Fig. 3D), N8 was suggested to be putative 2α-hydroxy erythrodiol.

In *N. benthamiana* leaves expressing *AtLUP1*, 3 distinct peaks, namely N1, A1, and A5, were observed (Fig. 4). Peaks A1 and A5 correspond to the major products of AtLUP1, lupeol, and lupanediol, respectively (Segura et al. 2000). Coexpression of *AtLUP1* and *NaCYP716C87* in *N. benthamiana* resulted in the appearance of 2 new peaks, designated as A2 and A7 (Fig. 4; Supplementary Fig. S3). Comparison of their EI-MS spectra and fragmentation pathways with those of lupeol (A1) (Supplementary Fig. S10), lupanediol (Supplementary Fig. S11), and maslinic acid (Supplementary Fig. S8) revealed characteristic fragments indicative of potential C2α hydroxylation. Additionally, peak A2 exhibited the characteristic fragment 189 of lupeol, while peak A7 displayed the characteristic fragment 131 of lupanediol (Fig. 4; Supplementary Figs. S10 and S11). Therefore, we inferred that peak A2 corresponds to putative 2α-hydroxy lupeol and peak A7 corresponds to putative 2α-hydroxy lupanediol. Upon coexpression of *AtLUP1*, *NaCYP716C87*, and *NaCYP716A419* in *N. benthamiana*, peaks A2 and A7 disappeared, while peaks A10 and A11 emerged (Fig. 4; Supplementary Fig. S12). Based on the EI-MS spectra and inferred fragmentation pathways of oleanolic acid (N4, Supplementary Fig. S13), betulinic acid (A4, Supplementary Fig. S14), and maslinic acid (N7, Supplementary Fig. S8), compounds bearing a C28 carboxyl group exhibit a characteristic fragment at *m/z* 320. Combining the previous findings (Fig. 3), peak A10 is inferred to be putative 2α-hydroxyl betulinic acid (alphitolic acid), whereas peak A11 is suggested to be putative 2α-hydroxy28-carboxy lupanediol ((1R,3aS,5aR,5bR,9R,11aR)-9,10-dihydroxy-1-(2-hydroxypropan-2-yl)-5a,5b,8,8,11a-pentamethylicosahydro-3aH-cyclopenta[a]chrysene-3a-carboxylic acid).

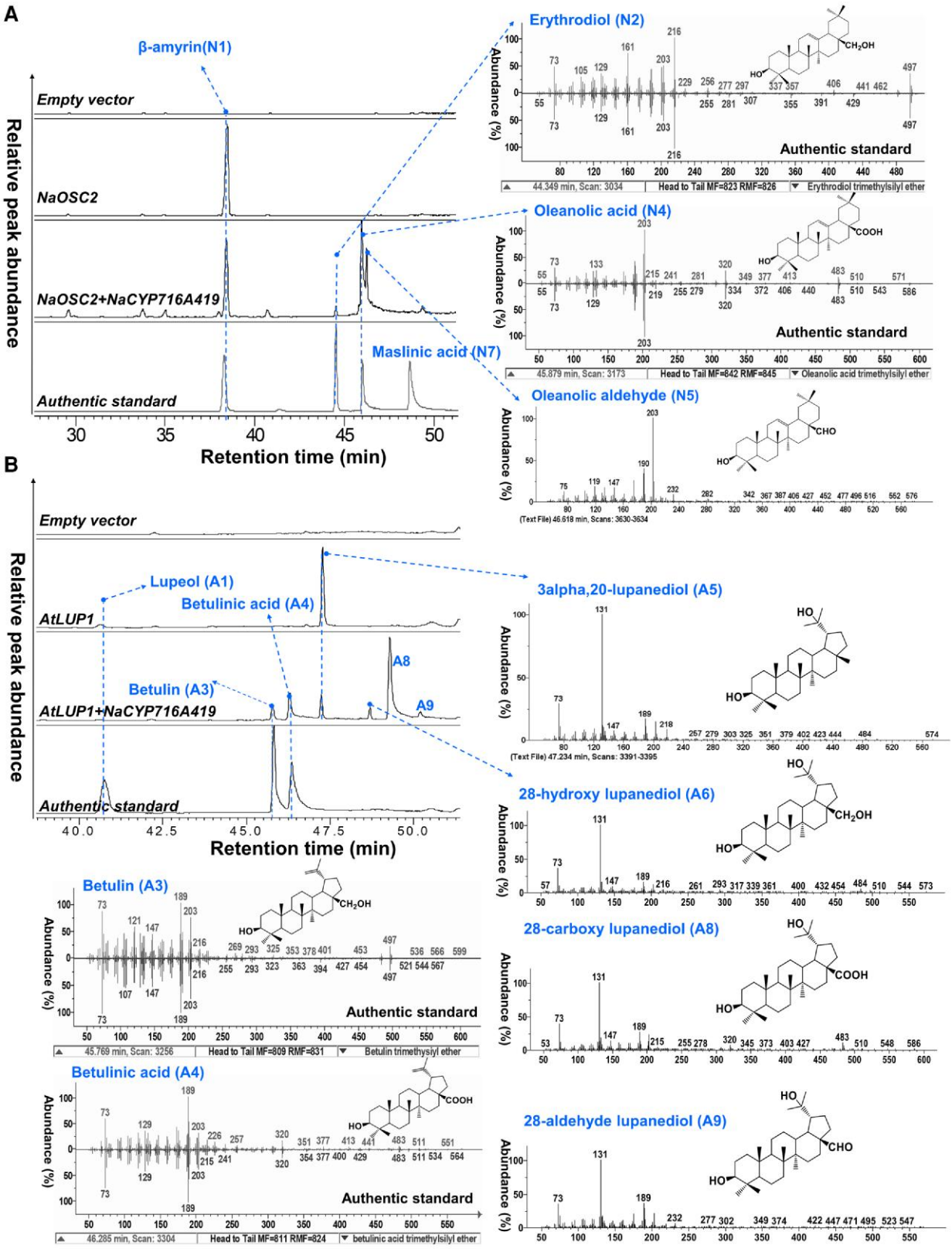

Figure 2. Characterization of NaCYP716A419 as a C28 oxidase. **A)** GC-MS profiles (SIM at *m/z*: 189, 203, 216, and 320) and EI-MS spectra of trimethylsilylated triterpenoids of extracts of *N. benthamiana* leaves coexpressing *NaCYP716A419* with *NaOSC2*. The authentic standard β-amyrin (N1), erythrodiol (N2), oleanolic acid (N4), and maslinic acid (N7) were used for product identification. **B)** GC-SIM-MS profiles (*m/z*: 189, 203, 131, 216, and 320) and EI-MS spectra of trimethylsilylated triterpenoids of extracts of *N. benthamiana* leaves coexpressing *CYP716A419* with *AtLUP1*. Each treatment was conducted with at least 3 replications. The authentic standard lupeol (A1), betulin (A3), and betulinic acid (A4) were used for product identification.

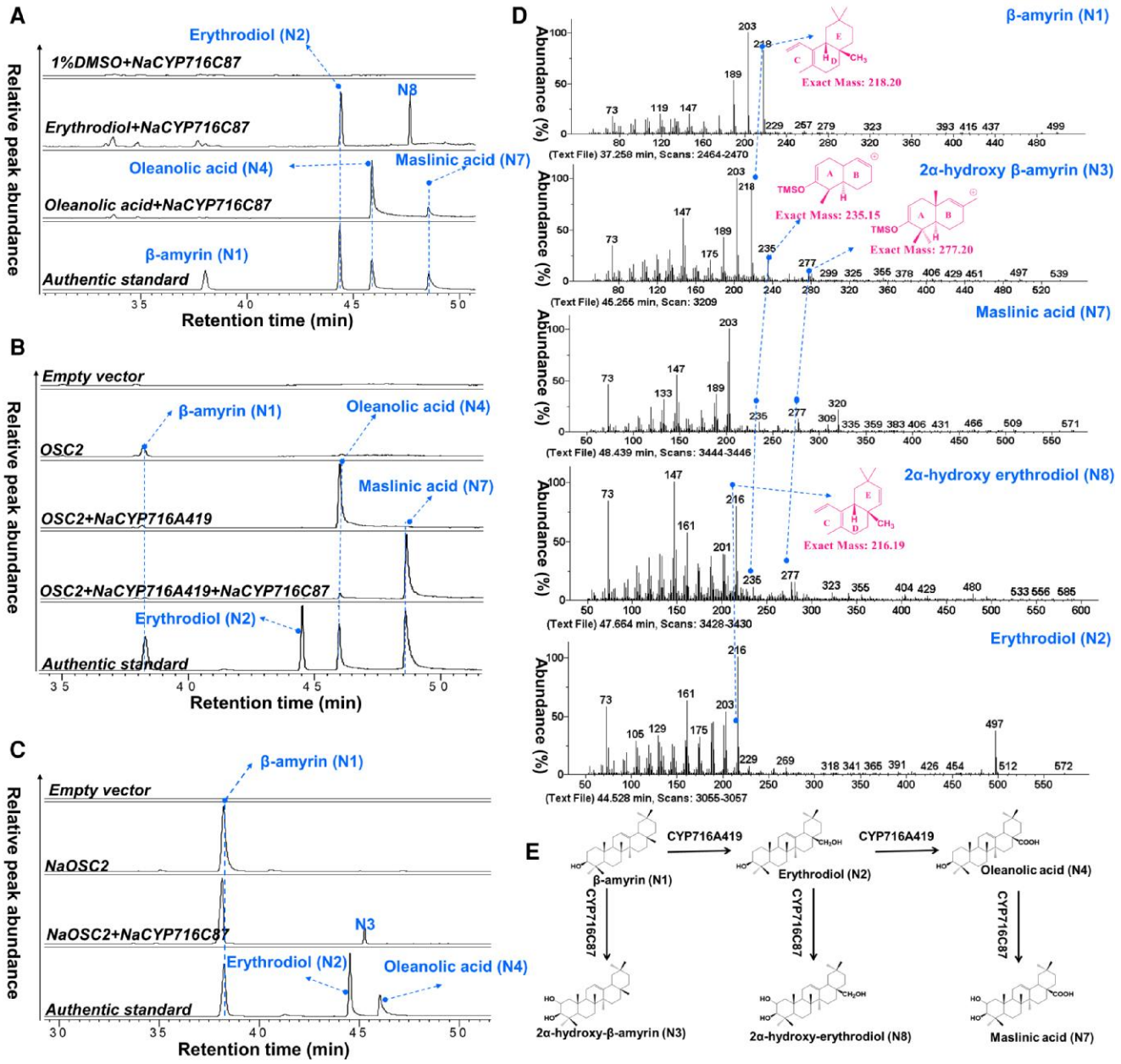

Figure 3. NaCYP716C87 is a C2α oxidase. **A)** GC-SIM-MS (*m*/z: 189, 203, 216, and 320) chromatograms of extracts from *N. benthamiana* leaves expressing only *NaCYP716C87* and supplemented with erythrodiol (N2) and oleanolic acid (N4) substrates (100 μM in 1% DMSO). One percent DMSO was added to *N. benthamiana* leaves expressing *NaCYP716C87* as a control. **B)** GC-SIM-MS (*m*/z: 189, 203, 216, and 320) chromatograms of acidic extracts of *N. benthamiana* leaves coexpressing *NaOSC2*, *NaCYP716A419*, and *NaCYP716C87*. Leaves expressing EVs, *NaOSC2*, and *NaOSC2*+ *NaCYP716A419* were employed as controls. **C)** GC-SIM-MS (*m*/z: 189, 203, 216, and 320) chromatograms of extracts of *N. benthamiana* leaves coexpressing *NaOSC2* and *NaCYP716C87*. Leaves expressing EVs or only *NaOSC2* were employed as controls. **D)** EI-MS spectra of trimethylsilylated 2α-hydroxy β-amyrin (N3), maslinic acid (N7), 2α-hydroxy erythrodiol (N8), β-amyrin (N1), and erythrodiol. **E)** Structures of the substrate, intermediates, and products of the oxidation reaction step from β-amyrin to maslinic acid by NaCYP716A419 and NaCYP716C87. Each treatment was conducted with at least 3 replications. The authentic standard β-amyrin (N1), erythrodiol (N2), oleanolic acid (N4), and maslinic acid (N7) were used for product identification in **A) and B)**. The authentic standard β-amyrin (N1), erythrodiol (N2), and oleanolic acid (N4) were used for product identification in **C)**.

In summary, CYP716C87 functions as a C2α hydroxylase, which accepts β-amyrin (N1), erythrodiol (N2), and oleanolic acid (N4) (Fig. 3E), as well as lupeol (A1), lupanediol (A5), betulinic acid (A4), and putative 28-carboxy lupanediol (A8) as substrates (Fig. 4).

NaCYP716E107 is a C6β hydroxylase

Amino acid sequence analysis revealed a 62% identity and 80% similarity between NaCYP716E107 and SlCYP716E26 (Supplementary Fig. S15), a hydroxylase that catalyzes the β-amyrin C6β hydroxylation reaction and produces daturadiol as a product (Yasumoto et al. 2017).To determine if NaCYP716E107 has C6β hydroxylation functionality given the lack of daturadiol standards, we cloned *SlCYP716E26* from tomato roots. Then, we conducted heterologous expression in *N. benthamiana* by coexpressing the enzyme individually with *NaOSC2* or in combination with *NaOSC2*/*NaCYP716A419* (Fig. 5, A and B). The leaves expressing *NaOSC2* and *SlCYP716E26* produced 2 peaks, β-amyrin (N1) and N6 (Fig. 5A; Supplementary Fig. S16). By comparing the retention times

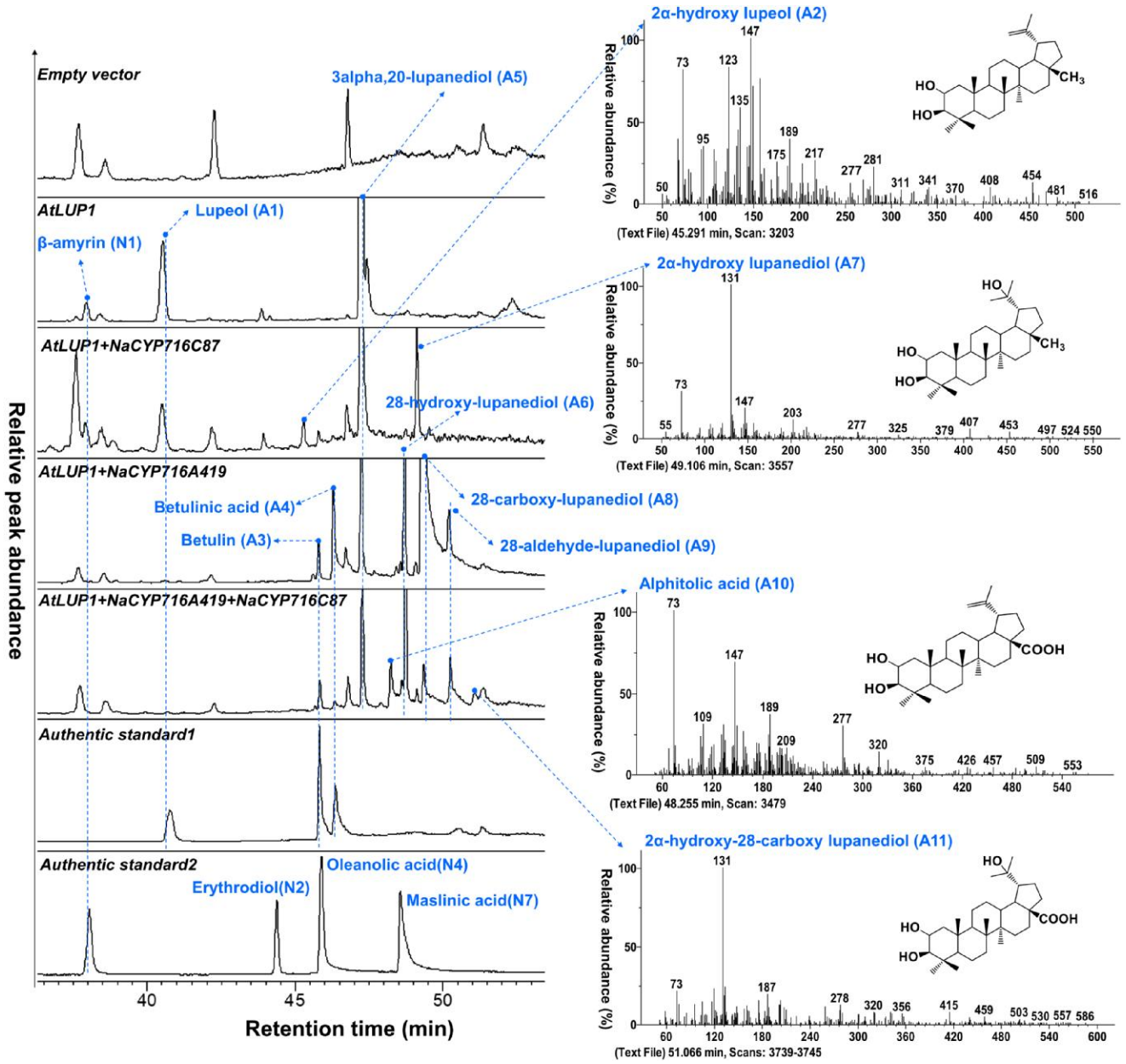

Figure 4. GC-MS chromatograms (SIM: 131, 189, 203, 320, and 216) and EI-MS spectra of TMS derivitization of extracts of *N. benthamiana* leaves coexpressing *AtLUP1* and *NaCYP716C87*. Leaves expressing EVs, *AtLUP1*, and *AtLUP1*+ *NaCYP716A419* were employed as controls. Each treatment was conducted with at least 3 replications. The lupeol (A1), betulin (A3), and betulinic acid (A4) were used in authentic standard 1. The β-amyrin (N1), erythrodiol (N2), oleanolic acid (N4), and maslinic acid (N7) were used in authentic standard 2.

and EI-MS spectra with those reported products for *SlCYP716E26* (Yasumoto et al. 2017), we confirmed that N6 is daturadiol (Supplementary Table S3). Leaves coexpressing *NaOSC2* and *NaCYP716E107* yielded a peak at the same position. Upon GC-MS analysis, the peak matched with daturadiol (N6) obtained from the coexpression of *NaOSC2* and *SlCYP716E26*. However, the yield was reduced by 95% compared to when *SlCYP716E26* was coexpressed with *NaOSC2* (Fig. 5A; Supplementary Fig. S16A). CaCYP716E41 from *C. asiatica* is another enzyme demonstrated to catalyze C6β hydroxylation of triterpenes. It shares high sequence similarity with NaCYP716E107 and SlCYP716E26 (Supplementary Fig. S15). However, unlike NaCYP716E107 and SlCYP716E26, CaCYP716E41 does not catalyze C6β hydroxylation of β-amyrin. Instead, it catalyzes oleanolic acid to yield 2 peaks: 6β-hydroxyl oleanolic acid and putative incompletely derivatized 6β-hydroxyl oleanolic acid. Furthermore, CaCYP716E41 can also catalyze maslinic acid to produce 6β-hydroxy maslinic acid (Miettinen et al. 2017). To investigate whether NaCYP716E107 and SlCYP716E26 can accept other triterpene scaffolds, we supplemented leaves expressing *NaCYP716E107* or *SlCYP716E26* with substrates such as erythrodiol (N2), oleanolic acid (N4), or maslinic acid (N7). No new peaks were detected when erythrodiol (N2) and maslinic acid (N7) were added to the leaves (Fig. 5C). However, upon the addition of oleanolic acid (N4), a new peak, N9, was detected in leaves expressing either *NaCYP716E107* or *SlCYP716E26*, albeit in a low yield (Fig. 5C; Supplementary Fig. S16B). We hypothesized that the efficiency of substrate uptake into plant cells might be too low to provide sufficient substrate for the enzymes. Therefore, under the condition of not providing oleanolic acid and maslinic acid, we coexpressed *NaOSC2*/*NaCYP716A419* or *NaOSC2*/*NaCYP716A419*/*CYP716C87* with *NaCYP716E107* or *SlCYP716E26* (Fig. 5D). Coexpression of *NaOSC2*/

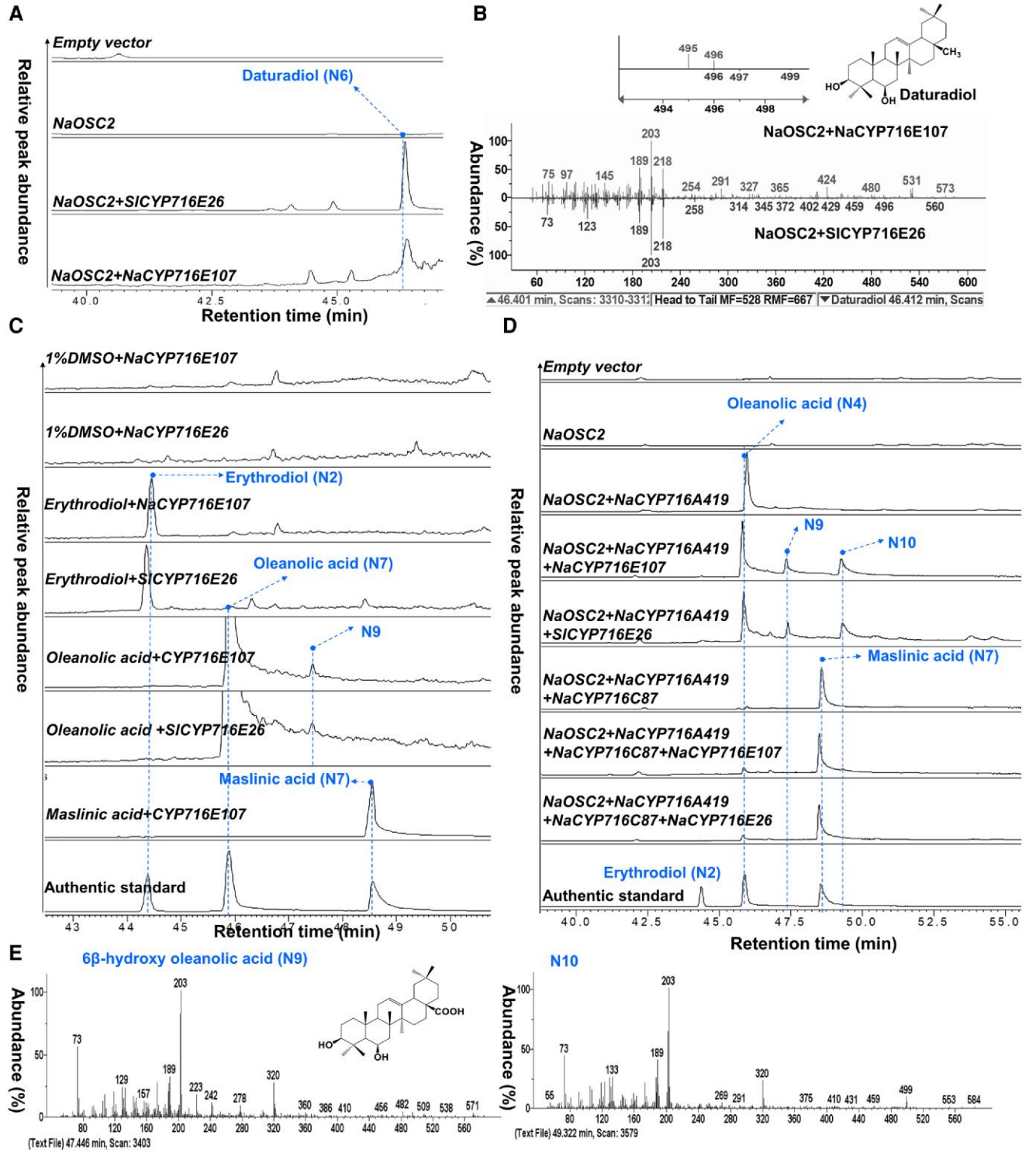

Figure 5. CYP716E107 is a C6β oxidase. **A)** GC-SIM-MS (*m/z*: 189, 203, and 218) profile of extracts of *N. benthamiana* leaves coexpressing *NaOSC2*/*CYP716E107* (*SlCYP716E26*). *SlCYP716E26* is a P450 monooxygenase cloned from *S. lycopersicum*, responsible for the C6β oxidation of β-amyrin (Yasumoto et al. 2017). **B)** GC-MS spectrum of daturadiol. **C)** GC-SIM-MS (*m/z*: 189, 203, and 320) profiles of extracts of *N. benthamiana* leaves fed with different substrates. **D)** GC-SIM-MS (*m/z*: 189, 203, and 320) profiles of extracts of *N. benthamiana* leaves coexpressing *NaCYP716E107* or *SlCYP716E26* with *NaOSC2*/*NaCYP716A419* or *NaOSC2*/*NaCYP716A419*/*NaCYP716C87* (the results only show the profiles of acidic extracts). **E)** EI-MS spectra of trimethylsilylated 6β-hydroxy oleanolic acid (N9) and N10. Each treatment was conducted with at least 3 replications. Erythrodiol (N2), oleanolic acid (N4), and maslinic acid (N7) were used as authentic standards.

NaCYP716A419 with *NaCYP716E107* or *SlCYP716E26* resulted in N9 in greater abundance and another new peak, N10 (Fig. 5D; Supplementary Fig. S6).

Based on the comparison of the EI-MS spectra of N10 with the fragmentation pathway of oleanolic acid (Fig. 5E; Supplementary Fig. S13), as well as the similarity of N9 and N10 to the products

of CaCYP716E41 reported in Miettinen et al. (2017), we speculate that N9 is 6β-hydroxy oleanolic acid, and N10 is incompletely derivatized 6β-hydroxy oleanolic acid. However, even the expression of *NaCYP716E107* or *SlCYP716E26* with the coexpression of *NaOSC2/NaCYP716A419/NaCYP716C87* did not produce detectable quantities of 6β-hydroxy maslinic acid (Fig. 5D; Supplementary Fig. S6). We also attempted coexpression of *NaCYP716E107* with *AtLUP1* or *AtLUP1/NaCYP716A419* but did not observe any new peaks (Fig. 1; Supplementary Figs. S3 and S12).

Expression patterns of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87*

To investigate the potential functions of these CYP716 enzymes in *N. attenuata*, we conducted reverse transcription quantitative PCR (RT-qPCR) analysis to examine their expression and induction patterns. Consistent with the results obtained from the microarray analysis in Fig. 1, the relative transcript levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* were found to be higher in flowers, followed by roots, with the lowest expression observed in leaves. Notably, the expression level of the *NaCYP716C87* gene in roots was comparable to that in flowers (Fig. 6A). Additionally, we investigated the responses of 3 *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* to various phytohormones and abiotic stresses, including abscisic acid (ABA), salicylic acid (SA), gibberellin A3 (GA_3), and methyl jasmonate (MeJA), as well as PEG6000 (mimicking drought stress), NaCl (salt stress), and Na_2CO_3 (mimicking alkaline stress) (Fig. 6, B to E). After 4-h treatments with ABA, SA, GA_3, and MeJA, the relative transcript levels of *NaCYP716A419* in the roots increased by 1.7-, 6.5-, 15.3-, and 4.9-fold, respectively, compared to the control group (Fig. 6B). In contrast, *NaCYP716E107* exhibited a weaker response to these 4 phytohormones. ABA, GA_3, and MeJA induced *NaCYP716E107* most strongly 7 h after treatment, resulting in a 4-fold increase in relative transcript levels for ABA, a 3.5-fold increase for GA_3, and a 3-fold increase for MeJA treatments, compared to the control. No significant induction of *NaCYP716C87* transcripts was observed within the first 9 h of MeJA treatment. However, 4 h after treatments with ABA, SA, and GA_3, the relative transcript levels of *NaCYP716C87* increased by 6.1-, 117-, and 41.7-fold, respectively, compared to the control group (Fig. 6B). In response to PEG6000 treatments, simulating drought stress, *NaCYP716E107*, and *NaCYP716C87* transcript levels increased 3- and 5-fold, respectively, compared to the control group (0% PEG) (Fig. 6C). In contrast, *NaCYP716A419* was largely unresponsive. In response to NaCl treatments, relative transcript levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* all showed reductions when compared to the control group (0 mM NaCl) (Fig. 6D). Specifically, treatments with 50 and 100 mM NaCl resulted in 19% and 37% decreases in *NaCYP716A419*, 25% and 36% decreases in *NaCYP716E107*, and 94% and 97% decreases in *NaCYP716C87* relative transcript levels. Under alkaline stress (Na_2CO_3 treatment), only transcripts of *NaCYP716A419* and *NaCYP716E107*increased significantly. Specifically, at 100 and 200 mM Na_2CO_3 treatments, *NaCYP716A419* increased 10- and 17-fold, respectively, and *NaCYP716E107* increased 41- and 56-fold, respectively (Fig. 6E). *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* appear to be responsive to a variety of environmental stressors in *N. attenuata*. The response of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* to hormones and abiotic stresses exhibits notable differences, suggesting they may possess distinct biological functions. *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* all show a response to GA_3 and ABA or MeJA treatments, indicating their potential involvement in growth-related or defense-related processes.

Silencing *NaCYP716A419* attenuates *N. attenuata* growth and reproductive performance

To further elucidate the biological functions of *NaCYP716A419*, *NaCYP716C87*, and *NaCYP716E107* in *N. attenuata*, we employed a VIGS technique using TRV (Saedler and Baldwin 2004). Each of the CYP450 genes was individually silenced in separate replicate plants. Additionally, plants inoculated with TRV vectors harboring EV constructs were included as negative controls and *NaOSC1*- and *NaOSC2*-harboring VIGS vectors served as positive controls (*NaOSC1/2_VIGS*). The visibly apparent bleaching phenotype of tomato phytoene desaturase (PDS)-silenced plants (*PDS_VIGS*) was used to monitor the onset and spread of gene silencing.

Given that most plant-specialized metabolites function in defense or stress responses, phenotyping of plant growth under intense competition and nutrient limitations provides a useful means of amplifying the growth-related costs associated with metabolite production (Baldwin and Hamilton 2000; Yang et al. 2023a). Hence, a companion–plant approach was used, in which 2 plants matched in initial size were grown in a single pot, with 1 plant silenced in the expression of different triterpene biosynthesis genes (*NaOSC1/2*, *NaCYP716A419*, *NaCYP716E107*, or *NaCYP716C87*) and the other was a control EV-inoculated plant. This setup facilitates the quantification of subtle differences in plant growth and reproductive performance resulting from silencing these triterpene biosynthesis genes. Each combination was performed with 30 replicates. Within 14 d after *Agrobacterium* infiltration, plants that did not exhibit viral infection symptoms were discarded. At 25 d post *Agrobacterium* infiltration, newly grown leaves of *PDS_VIGS* plants became completely bleached, indicating the successful induction of virus-mediated gene silencing (Fig. 7A). qPCR analysis of VIGS plant leaves revealed that, compared to EV plants, *NaOSC1/2-VIGS* plants had relative transcript levels reduced by 56% and 44% for *NaOSC1* and *NaOSC2*, respectively. *NaCYP716A419*-silenced plants (*CYP716A419_VIGS*) exhibited a 45% reduction in *NaCYP716A419* transcript levels, while *NaCYP716E107*-silenced plants (*CYP716E107_VIGS*) showed a 74% reduction in *NaCYP716E107* transcript levels. *NaCYP716C87*-silenced plants (*CYP716C87_VIGS*) displayed a 48% reduction in *NaCYP716C87* transcript levels (Supplementary Fig. S17). After 34 d of *Agrobacterium* inoculation, we observed significant growth inhibitions in *CYP716A419_VIGS* plants (Fig. 7A). This inhibition was evident in reduced stalk height and rosette diameter (Fig. 7B) compared to the neighboring EV plants (Fig. 7A). As the plants entered the flowering stage, these growth differences became more pronounced. In *CYP716A419_VIGS* plants, the leaves became slender, and flower buds were aborted (Fig. 7A). Statistical analysis revealed that rosette diameters of *CYP716A419_VIGS* plants were 24% smaller than those of both neighboring EV plants and the EV–EV plants (Fig. 7B). Furthermore, *CYP716A419_VIGS* plants were 33% thinner in stalk diameter, 29% smaller in stalk height, and matured 59% fewer seed capsules compared to their neighboring EV plants. These reductions were 31%, 29%, and 51%, respectively, of those of the EV–EV control group (Fig. 7B). While significant reductions in rosette diameters and stalk heights were also observed in plants silenced in *NaCYP716E107*, and silencing *NaOSC1/2* resulted in a notable reduction in rosette diameters compared to EV plants, these plants did not show significant differences in capsule numbers compared to the EV controls (Fig. 7B). Due to the poor reproductive performance of *CYP716A419_VIGS* plants in terms of capsule numbers, we evaluated the pollen viability of these plants. Metabolic staining of pollen revealed lower pollen viability in *CYP716A419_VIGS* plants compared to both EV plants and *NaOSC1/2_VIGS* plants (Supplementary Fig. S18).

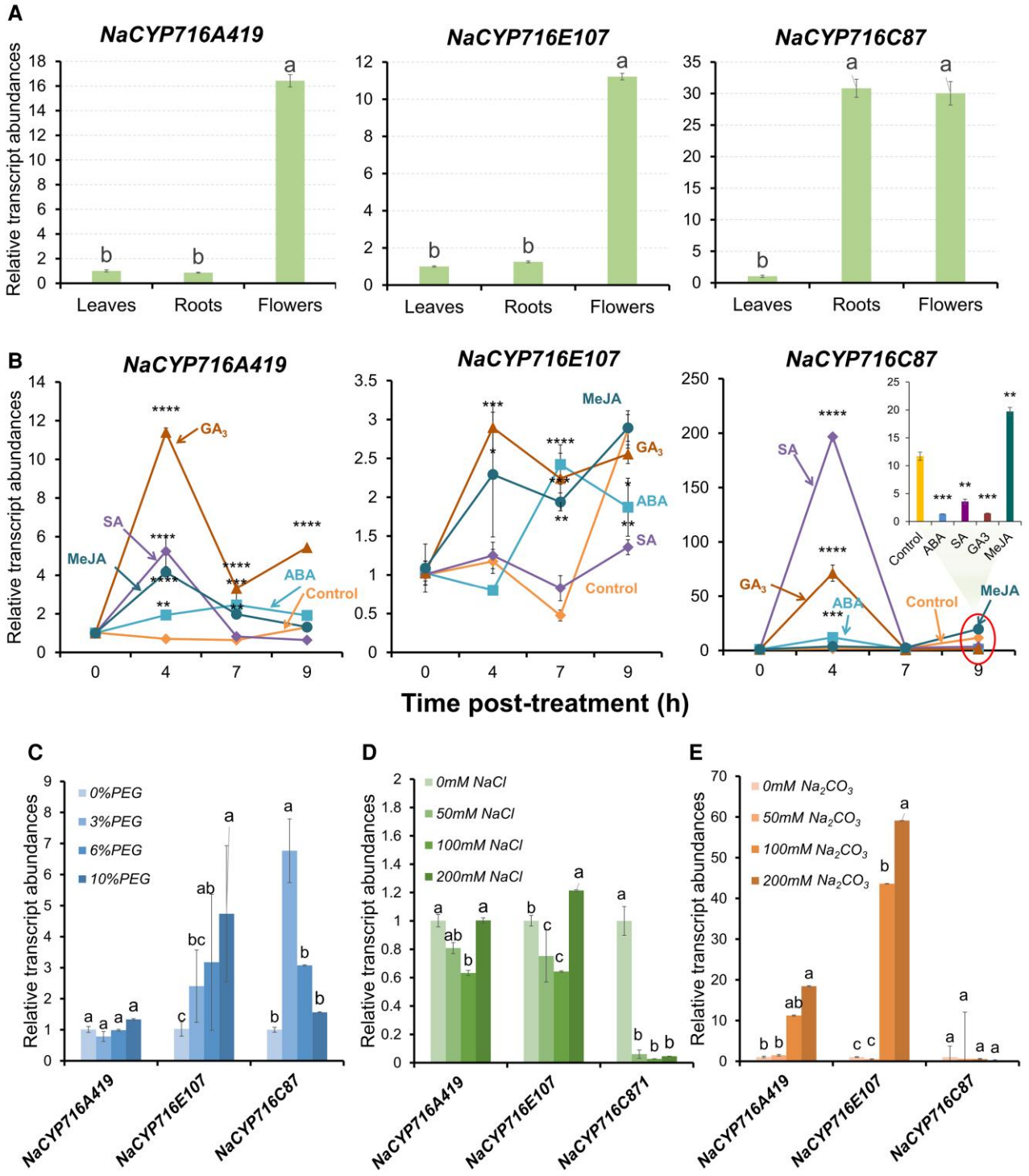

Figure 6. Responses of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* transcripts to hormone and abiotic stress treatments. **A)** Relative transcript levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* in different tissues. **B)** Relative transcript levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* after phytohormone treatment. Inset provides statistical details of the 9-h harvest for *NaCYP716C87*. The circle indicates the relative transcript abundance of *NaCYP716C87* after 9 h of hormone treatment, which is expanded in the bar chart at the top right corner of **B)**. **C)** Relative transcript levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* after PEG6000 treatments. **D)** Relative transcription levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* after NaCl treatment. **E)** Relative transcription levels of *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87* after Na_2CO_3 treatment. Results of ANOVAs with Tukey's test are shown in **A)** and **C to E)** ($n = 3$, mean ± SE; different lowercase letters indicate statistically significant differences between 2 groups of samples at the $P < 0.05$ level). Results of Student's *t* tests are shown in **B)** ($n = 3$, mean ± SE, $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$).

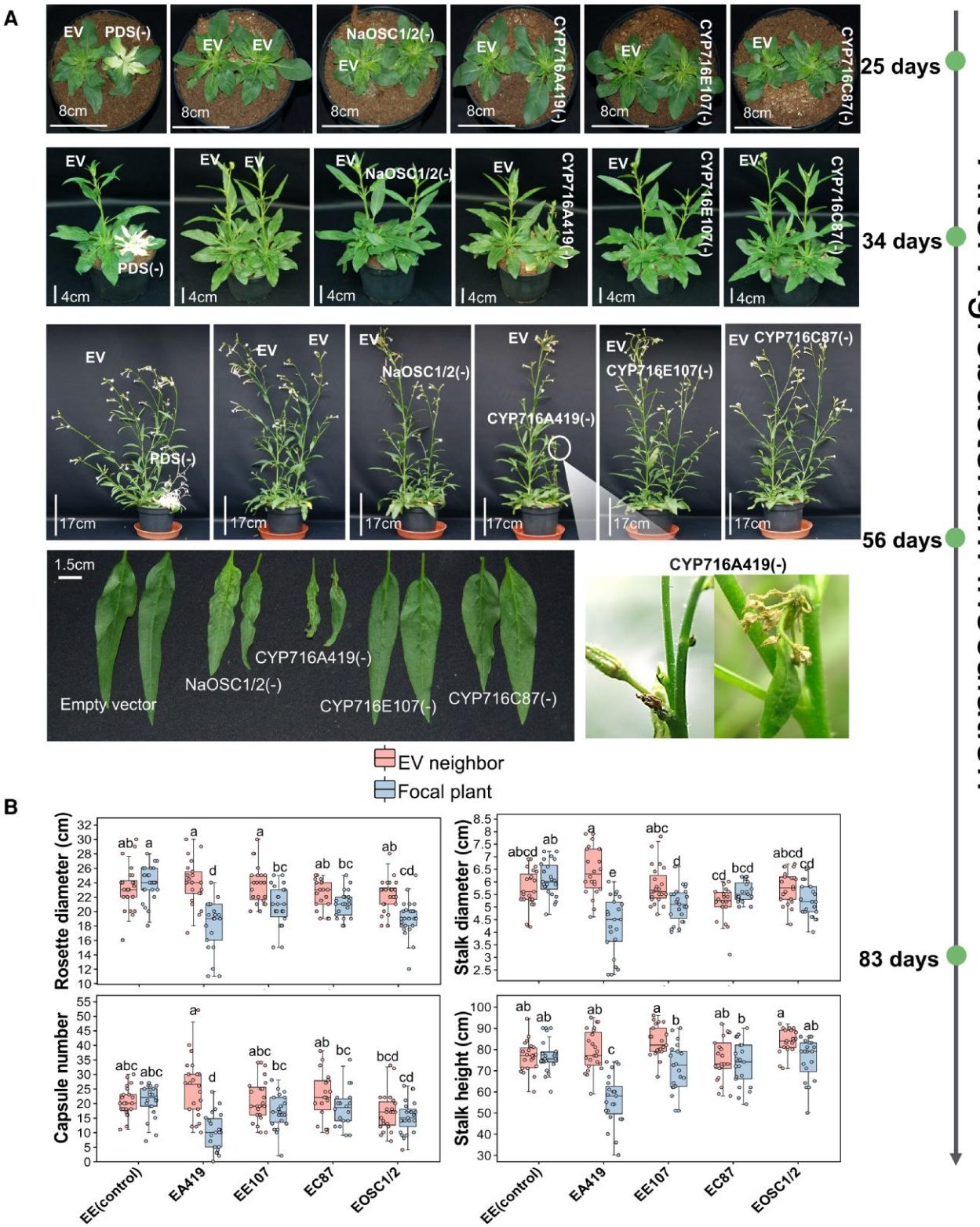

Figure 7. Silencing key triterpenoid biosynthesis genes attenuates the growth and fitness of *N. attenuata*. **A)** The growth phenotypes of VIGS plants after *Agrobacterium* inoculation. Only *CYP716A419_VIGS* plants aborted flowers. **B)** Fitness of VIGS plants (*NaOSC1/2_VIGS*, *CYP716A419_VIGS*, *CYP716E107_VIGS*, or *CYP716C87_VIGS*). Focal plants refer to the plants in which *NaOSC1/2*, *NaCYP716A419*, *NaCYP716E107*, or *NaCYP716C87* was silenced. Neighbor plants refer to initially size-matched plants inoculated with *Agrobacterium* carrying an EV construct. The control group, labeled EE(control), consists of 2 EV plants. EA419 indicates that the neighbor plants are EV plants, and the focal plants are *CYP716A419_VIGS* plants. Similar labeling schemes for the other combinations: EE107: EV and *CYP716E107_VIGS* plant pairs; EC87: EV and *CYP716C87_VIGS* plant pairs; EOSC1/2: EV and *NaOSC1/2_VIGS* plants pairs. Results of 2-tailed ANOVAs with Tukey's tests are shown ($n = 21$ to 24, mean ± SE, different lowercase letters indicate statistically significant differences between 2 groups of samples at the $P < 0.05$ level). The central line within the box represents the median of the data. The upper and lower boundaries of the box denote the upper and lower quartiles of the data. The lines above and below the box, known as whiskers, signify the variability of the data (error bar). The whiskers' length is set at 1.5 times the interquartile range. The points represent specific observations in the data set. The values beyond 1.5 times the interquartile range are considered as outliers.

Triterpenoids in *N. attenuata*

In leaf tissues, only β-amyrin was detected. Compared to EV, the levels of β-amyrin in *CYP716A419_VIGS*, *CYP716AC87_VIGS*, and *CYP716E107_VIGS* plants increased by 37%, 30%, and 10%, respectively. Conversely, in *NaOSC1/2_VIGS* plants, β-amyrin levels decreased by 77% (Fig. 8). In the floral organs of VIGS plants, β-amyrin and the products of NaCYP716A419 (erythrodiol and oleanolic aldehyde) were detected (Fig. 8). Silencing *NaCYP716A419* led to a 26% increase in β-amyrin content in the corolla, 26% in the pollen, and 17% in the thecal coverings (theca) of the anther heads that enclose the pollen grains until dehiscence. These increases were associated with decreases in erythrodiol, by 24%, 23%, and 13%, and oleanolic aldehyde, by 26%, 26%, and 12%, in these respective floral tissues. In the case of

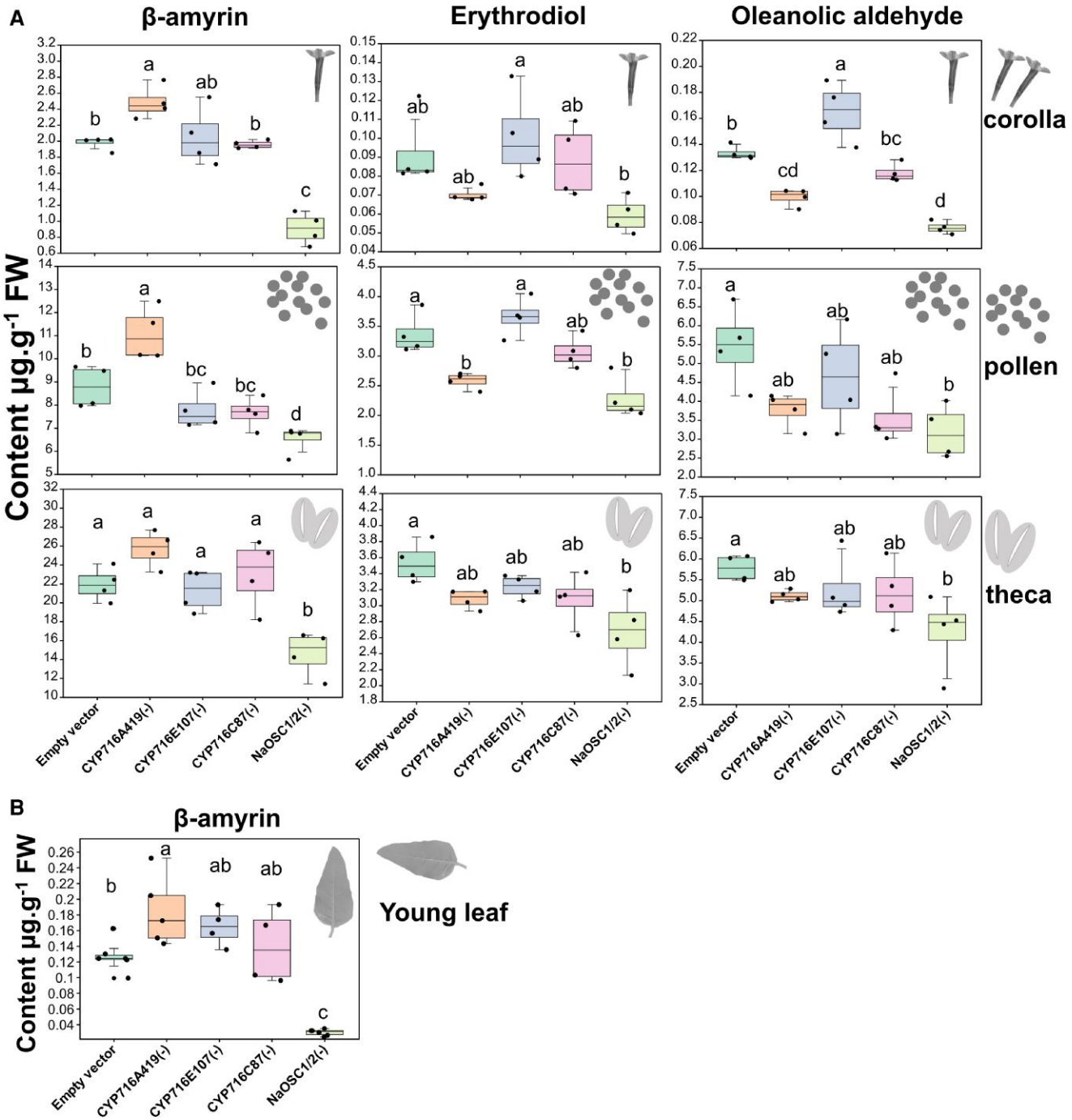

Figure 8. Triterpene contents in CYP450-silenced plants. **A)** Triterpene contents in flower parts. **B)** Triterpene contents in leaves. Samples from 4 individual plants were mixed to form a single replicate; a total of 4 to 6 replicates for each tissue type were used. ANOVAs with Tukey's tests are shown ($n = 4\sim6$, mean ± SE, different lowercase letters indicate statistically significant differences between 2 groups of samples at the $P < 0.05$ level). The central line within the box represents the median of the data. The upper and lower boundaries of the box denote the upper and lower quartiles of the data. The lines above and below the box, known as whiskers, signify the variability of the data (error bar). The whiskers' length is set at 1.5 times the interquartile range. The points represent specific observations in the data set. The values beyond 1.5 times the interquartile range are considered as outliers.

NaCYP716E107 silencing, only oleanolic aldehyde increased by 25% in the corolla, while the other components remained unchanged. In *NaOSC1* and *NaOSC2* cosilenced plants, there were significant reductions in triterpenoid contents when compared to the EV control. Specifically, β-amyrin decreased by 54% in the corolla, 26% in the pollen, and 33% in the theca. Erythrodiol showed a 36% reduction in the corolla, a 32% reduction in the pollen, and a 24% reduction in the theca. Meanwhile, oleanolic aldehyde decreased by 43% in the corolla, 42% in the pollen, and 27% in the theca (Fig. 8). Notably, these compounds did not exhibit significant changes in plants with *NaCYP716C87* silenced (Fig. 8).

We did not detect maslinic acid, 2α-hydroxy β-amyrin (products of NaCYP716C87), or 6β-hydroxy oleanolic acid (products of NaCYP716E107) in the flowers of *N. attenuata*. To explore whether the enzyme products accumulate in *N. attenuata*, and considering that ABA, GA_3, SA, and MeJA can induce the expression of the *NaCYP716C87*, *NaCYP716A419*, and *NaCYP716E107* genes in the root, we attempted to detect the triterpenoids in plant roots 72 h after induction with ABA, GA_3, SA, and MeJA. No products of NaCYP716C87 and NaCYP716E107 were observed; instead, peaks suspected to be derivatives of lupanediol were detected in these hormone-induced plants (Supplementary Fig. S19). However, it is noteworthy that oleanolic acid (N4), maslinic acid (N7), and 6β-hydroxy oleanolic acid (N9) were detected in the roots colonized for 5 wk by arbuscular mycorrhizal fungi (AMF) (Supplementary Fig. S20). These results suggest that NaCYP716A419, NaCYP716C87, and NaCYP716E107 can produce several of their respective products in *N. attenuata*, but their production depends on specific developmental stages and specific inducing factors.

Discussion

In this study, we characterized the functions of several P450 enzymes from *N. attenuata* in *N. benthamiana* (Fig. 1). The work revealed that NaCYP716A419, NaCYP716C87, and NaCYP716E107 modify pentacyclic triterpene skeletons, specifically at the C28, C2α, and C6β positions of β-amyrin, lupeol, or lupanediol, respectively.

The CYP716A subfamily has been extensively studied in other plants and is known to catalyze a sequential 3-step oxidation at the C28 position of β-amyrin, lupeol, and α-amyrin, forming a hydroxyl group, aldehyde, and carboxyl group, respectively (Carelli et al. 2011; Fukushima et al. 2011; Miettinen et al. 2017; Misra et al. 2017; Yasumoto et al. 2017; Alagna et al. 2023). Certain members of the CYP716A subfamily, such as BpCYP716A180 from *Betula platyphylla*, have been suggested to catalyze C28 oxidations on lupanediol (Zhou et al. 2016). *NaCYP716A419*, when coexpressed with *NaOSC2* in *N. benthamiana*, yielded erythrodiol (N2), oleanolic acid (N4), and oleanolic aldehyde (N5). Additionally, when coexpressed with *AtLUP1* in *N. benthamiana*, the enzyme resulted in the production of betulin (A3), betulinic acid (A4), 28-hydroxy lupanediol (A6), and 28-carboxy lupanediol (A8), indicating that NaCYP716A419 also functions as a C28 oxidase.

In contrast to the well-characterized CYP716A subfamily enzymes, the enzymatic function of the CYP716C subfamily remains largely unknown. To date, only CaCYP716C11 from *C. asiatica* (Miettinen et al. 2017), AmCYP716C53 from *Avicennia marina* (Nakamura et al. 2018), OeCYP716C67 from *O. europaea* (Alagna et al. 2023), and CpCYP716C49 from *Crataegus pinnatifida* (Dai et al. 2019), as well as LsCYP716C55 from *Lagerstroemia speciosa* (Sandeep et al. 2019), have been characterized. These enzymes catalyze the hydroxylation of the C2α position of oleanolic acid, ursolic acid, or betulinic acid, thereby forming maslinic acid, corosolic acid, and alphitolic acid (Miettinen et al. 2017; Nakamura et al. 2018; Dai et al. 2019; Alagna et al. 2023). However, there have been limited studies indicating the activity of CYP716C subfamily members on other triterpene alcohol substrates. NaCYP716C87 catalyzes the hydroxylation of the C2α position of oleanolic acid and betulinic acid, indicating that NaCYP716C87 is also a C2α hydroxylase. When NaCYP716C87 is coexpressed with *NaOSC2* or *AtLUP1* in *N. benthamiana* (Fig. 1), it yields 2α-hydroxy β-amyrin (N3), 2α-hydroxy lupeol (A2), and A7 (2α-hydroxy lupanediol). Feeding erythrodiol to the leaves expressing *NaCYP716C87* results in the production of 2α-hydroxy erythrodiol. These results suggested that some CYP716C family members can also catalyze C2α hydroxylation of triterpene alcohols.

Many members of the CYP716E subfamily have not yet been functionally characterized. To date, only 2 members, SlCYP716E26 and CaCYP716E41, have been reported to participate in the biosynthesis of triterpenoids (Miettinen et al. 2017; Yasumoto et al. 2017). In this study, NaCYP716E107 was found to act on the β-amyrin and oleanolic acid skeletons, but not on lupeol or lupanediol (Fig. 5). Previous reports showed that SlCYP716E26 catalyzes hydroxylation at the C6β position of β-amyrin, forming daturadiol (N6), but its activity toward other substrates was not investigated (Yasumoto et al. 2017). Conversely, CaCYP716E41 was reported to not catalyze β-amyrin but to catalyze hydroxylation at the C6β position of oleanolic acid and maslinic acid (Miettinen et al. 2017). In our study, we cloned SlCYP716E26 from tomato and compared it with the products of NaCYP716E107, demonstrating that NaCYP716E107 also functions as a C6β hydroxylase, catalyzing β-amyrin to form daturadiol (N6). Furthermore, both NaCYP716E107 and SlCYP716E26 exhibit the capability to catalyze oleanolic acid, aligning with the functional role of CaCYP716E41, yielding 2 distinct products (putative 6β-hydroxy oleanolic acid and its partially derivatized form). However, these enzymes are incapable of catalyzing maslinic acid substrates, a function unique to CaCYP716E41. Such disparities may be attributed to considerable disparities in their protein sequences, particularly at positions 251 to 260 (Supplementary Fig. S15).

In this study, other enzymes (NaCYP716A420, NaCYP716E108, NaCYP716D93, NaCYP716D94, NaCYP716H6, NaCYP88B4, and NaCYP88C16) did not show activity toward β-amyrin or lupeol (Fig. 1). We attempted coexpression of these inactive enzyme genes with *NaCYP716A419*, *NaCYP716C87*, or *NaCYP716E107* in *N. benthamiana*, but still no products were detected (Supplementary Figs. S21 and S22). Given that some CYP450 enzymes only accept substrates that have been modified with specific substituents, we cannot rule out the possibility that these enzymes may be involved in triterpene biosynthesis. For instance, MtCYP72A61v2 and MtCYP72A68v2 exhibit C-22β and C-23 oxidation activity, respectively, toward 24-hydroxy-β-amyrin and ursolic acid, but they do not catalyze the oxidation of β-amyrin (Fukushima et al. 2013). GuCYP72A154 from *Glycyrrhiza uralensis* is a C30 oxidase capable of catalyzing the 3-step oxidation of 11-oxo-β-amyrin to glycyrrhizin instead of directly oxidizing the β-amyrin scaffold (Seki et al. 2011). Furthermore, some members of the CYP716 family can catalyze the oxidation of tetracyclic triterpenes (Ghosh 2017). For example, ginseng's GuCYP716A47 (renamed GuCYP716U1) catalyzes the conversion of dammarenediol II into protopanaxadiol (Han et al. 2011), and GuCYP716A53v2 (renamed GuCYP716S1v2) catalyzes the transformation of protopanaxadiol into protopanaxatriol (Han et al. 2012). CsCYP88L2 and CsCYP81Q58 from *Cucumis sativus* catalyze 2 consecutive oxidation reactions in the cucurbitacin biosynthesis pathway (Shang et al. 2014). Our previous research has also confirmed the presence of dammarenediol, taraxasterol,

and several other unidentified triterpene scaffolds in *N. attenuata* (Yang et al. 2023b). Hence, we cannot exclude the possibility that NaCYP716A420, NaCYP716E108, NaCYP716D93, NaCYP716D94, NaCYP716H6, NaCYP88B4, and NaCYP88C16 may exhibit catalytic activity toward these triterpenes. *NaCYP716A419*, *NaCYP716C87*, and *NaCYP716E107* exhibit the highest expression in flowers. However, in the flowers of VIGS plants in which these enzymes were individually knocked down, only in *CYP716A419_VIGS* plants could erythrodiol and oleanolic aldehyde be detected in modest decreases: we did not observe the enzyme products of NaCYP716C87 and NaCYP716E107. Given the transcript elicitations of these enzymes by GA_3, ABA, and MeJA treatments in roots (Fig. 6B), we also attempted to detect the triterpenoid products of NaCYP716C87 and NaCYP716E107 after hormone inductions, but without success (Supplementary Fig. S19). Metabolomic analysis suggests the possible existence of glycosylated compounds derived from lupeol or β-amyrin in *N. attenuata* (Yang et al. 2023b). The enzyme products identified in this study are intermediates; hence, the enzyme products of NaCYP716C87 and NaCYP716E107 may have been metabolized into downstream compounds, thwarting the detection of the intermediates. Additionally, detectable levels of oleanolic acid (N4), maslinic acid (N7), 6β-hydroxy oleanolic acid (N9), and N10 were found in roots after AMF colonization. This indicates that these compounds may be detected in *N. attenuata* only under specific growth conditions, and this provides an alternative explanation for the lack of product detection for NaCYP716C87 and NaCYP716E107.

We also observed that *CYP716A419-VIGS* plants exhibited clear reductions in capsule numbers, as well as decreased stem height, diameter, and rosette diameter, along with attenuated pollen viability. However, silencing the initial enzyme of the triterpenoid biosynthetic pathway, *NaOSC1/2_VIGS* plants did not result in similarly severe effects. In a previous study in *N. attenuata*, we observed a similar phenomenon with a different biosynthetic pathway. Silencing the specific geranylgeranyl pyrophosphate (GGPPS) that controls diterpene biosynthetic flux into 17-hydroxygeranyllinalool diterpene glycoside biosynthesis did not result in developmental abnormalities. However, silencing genes involved in later steps in this pathway, such as *UGT74P3*, *UGT74P5*, *CYP736A304*, or *CYP736A305*, led to the accumulation of 17-hydroxygeranyllinalool or geranyllinalool, resulting in severe developmental defects (Heiling et al. 2021; Li et al. 2021). In many plants, intermediates of triterpene biosynthesis accumulate during normal growth and development, and manipulating their biosynthesis can lead to morphological and physiological effects (Guhling et al. 2006; Ohyama et al. 2007). For instance, *A. thaliana* plants that overexpressed thalianol synthase (THAS) are dwarfed but produce longer roots (Field and Osbourn 2008). Mutants of the marneral synthase (MRN1) gene in Arabidopsis display delayed embryogenesis, late flowering, rotund leaves, and abnormal seed morphology (Field et al. 2011). In oats, an increase in β-amyrin in the roots leads to shorter roots with a super-hairy phenotype (Kemen et al. 2014). In addition, disrupting the oxidation, glycosylation, or acylation steps in triterpene synthesis can also lead to developmental phenotypes. For example, the loss of function of mutants in genes *MtCYP716A12* and *UGT73F3* in *M. truncatula* truncates the production of hemolytic saponins and results in a dwarfing phenotype (Naoumkina et al. 2010; Carelli et al. 2011). Some triterpenoid saponins can regulate growth-related hormones, such as chromosaponin I (CSI), a γ-pyronyl-triterpenoid saponin isolated from peas (*Pisum sativum*) and other leguminous plants (Kudou et al. 1992; Tsurumi et al. 1992), which specifically interacts with the auxin influx carrier AUX1, thereby altering responses of Arabidopsis roots to auxin and ethylene (Rahman et al. 2001). Hence, we hypothesize that the observed phenotypic alterations could be attributed to either the toxic effects arising from the accumulation of intermediate products due to the silencing of *NaCYP716A419* or the absence of downstream growth-related triterpenoid compounds.

In summary, heterologous expression in *N. benthamiana* revealed that NaCYP716A419, NaCYP716E107, and NaCYP716C87 catalyze the oxidation of β-amyrin, lupeol, lupanediol, or their downstream compound skeletons at C28, C6β, and C2α positions. We attempted to validate these enzyme functions in *N. attenuata* by silencing *NaCYP716A419*, *NaCYP716E107*, and *NaCYP716C87*. We attempted to detect these enzyme products in hormone- or stress-induced *N. attenuata* roots. We found maslinic acid (N7), 6β-hydroxy oleanolic acid (N9), and N10 in AMF-harboring roots. While their concentrations were low, these products accumulated in *N. attenuata*. Additionally, our study indicated that *CYP716A419_VIGS* plants exhibited significant growth inhibition and reduced capsule numbers under competitive conditions. However, due to the transient and unstable silencing of VIGS, we can only conclude that NaCYP716A419 likely has a notable impact on the growth of *N. attenuata*. The specific functions and regulatory mechanisms of how NaCYP716A419 influences the growth and development of *N. attenuata* will require additional work with stable gene knockouts.

Materials and methods

Chemicals

MeJA (product ID: W341002), ABA (product ID: A4906), SA (product ID: 84210), α-amyrin (product ID: 53017), β-amyrin (product ID: 09236), erythrodiol (product ID: 09258), lupeol (product ID: 18692), betulinic acid (product ID: 91466), oleanolic acid (product ID: 42515), *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide (MSTFA) (product ID: 69479), and *N*,*N*-dimethylformamide (DMF) (product ID: 09258) were purchased from Sigma-Aldrich (St. Louis, Missouri, United States), and GA_3 (product ID: art. no. 7464.2) was purchased from Carl Roth (Karlsruhe, Germany).

Cultivation and elicitation of *N. attenuata* plants

Nicotiana attenuata Torr. Ex Watts. seeds from the 31st-generation inbred line were utilized as the wild-type (WT) genotype in all experiments. Seed germination and plant growth followed the protocols outlined in previous reports (Krügel et al. 2002), with a day/night cycle of 16 h (26 to 28 °C) and 8 h (22 to 24 °C) in a glasshouse at the Max Planck Institute for Chemical Ecology, Jena, Germany.

For phytohormone and abiotic stress inductions, 10-d-old seedlings were transferred to a substrate composed of soil balls and sand in a 1:3 ratio for 3 wk.

Twenty microliters of lanolin containing 100 μM MeJA, ABA, SA, or GA_3 was applied to the basal ends of seedling stems. An equivalent amount of pure lanolin served as a negative control. Roots were collected at 0, 4, 7, and 9 h after treatments for transcript analysis. Roots were collected after 72-h treatments for triterpenoid analysis. For abiotic stress treatments (salt, alkali, and drought stress), 50 mL of different concentrations of NaCl (0, 50, 100, and 200 mM), Na_2CO_3 (0, 50, 100, and 200 mM), and PEG6000 (0%, 3%, 6%, and 10% [w/v]) was added to the sand to create salt, alkali, and drought stress conditions, respectively. After 2 d of stress treatments, root samples were harvested for transcript analysis.

Phylogenetic analysis and sequence alignment

The protein sequences of CYP enzymes were obtained from the NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and the *Nicotiana attenuata* Data Hub (Brockmöller et al. 2017). Details regarding CYP enzymes related to triterpene biosynthesis are provided in Supplementary Table S1. Sequence alignment was conducted using ClustalW, and a phylogenetic tree was constructed through MEGA11 (Tamura et al. 2021) employing the neighbor-joining method with 1,000 bootstrap replicates. The tree's esthetic enhancement was accomplished using tvBOT (Xie et al. 2023).

Cultivation of *N. benthamiana* plants

Nicotiana benthamiana seeds were directly sown in 9 × 9 cm pots filled with soil and placed in a glasshouse at the Max Planck Institute for Chemical Ecology, Jena, Germany. The glasshouse maintained a day/night cycle of 16 h (26 to 28 °C) and 8 h (22 to 24 °C) for 3 wk. Following this, the 3-wk-old plants were transferred to the VIGS chamber, with a 16/8-h day/night cycle (22 °C during the day/night) and 40% relative humidity. The plants were kept in the VIGS chamber for 1 or 2 wk to reach the suitable stage for heterologous expression.

Heterologous expression of candidate enzymes in *N. benthamiana*

Full-length complementary DNAs (cDNAs) of selected CYP and *NaOSC2* enzymes from *N. attenuata*, *AtLUP1* from Arabidopsis (*A. thaliana*), and *SlCYP716E26* from tomato (*S. lycopersicum*) were cloned (digested with BsaI) into a 3Ω1 expression vector (Cárdenas et al. 2019; Hong et al. 2022) using the ClonExpress II One Step Cloning Kit (Vazyme) with the primers listed in Supplementary Table S5, and the vectors were transformed into *Agrobacterium tumefaciens* strain *GV3101*. A heterologous expression followed established protocols (Yang et al. 2023b). *Agrobacterium* strains carrying the gene constructs were cultured in 10 mL of LB medium supplemented with antibiotics (100 μg/mL rifampicin and 250 μg/mL spectinomycin) at 28 °C for 24 h. After centrifugation, the supernatant was discarded, and the cell pellet was resuspended in 5 mL of infiltration buffer (50 mM MES, 2 mM Na_3PO_4, 10 mM $MgCl_2$, and 100 μM acetosyringone). The cell suspension was then diluted to an OD_{600} of 0.6 for single enzymes and 0.4 for enzyme combinations (0.4 for every individual enzyme of the combinations). Infiltration was performed on the abaxial surface of 4- to 5-wk-old *N. benthamiana* leaves using a needle-free syringe. Triterpenoid analysis was conducted 5 d after *Agrobacterium* infiltration, and leaves were collected from 3 individual plants. For substrate feeding, 1 mL of 100 μM substrates (dissolved in 1% DMSO) was infiltrated into the abaxial surface of leaves 3 d after *Agrobacterium* inoculation. Leaf samples were collected for triterpenoid analysis 3 d after substrate injection.

VIGS

VIGS experiments were performed according to a published protocol optimized for VIGS in *N. attenuata* (Saedler and Baldwin 2004) with the *pBINTRA* and *pTV00* VIGS expression system. Briefly, 260 to 300 bp mRNA (CDS or UTR) sequences of enzymes were amplified by PCR using Q5 High-Fidelity DNA Polymerase (New England Biolabs) with primers listed in Supplementary Table S6 and cut by BamHI and SalI (New England Biolabs). The DNA fragments were cloned into a *pTV00* vector using T4 DNA ligase (Promega) and transformed into *A. tumefaciens* strain GV3101.

Two size-matched 20-d-old seedlings were transplanted into individual 2-L pots and placed in a controlled environment chamber with a 16/8-h day/night cycle (22 °C during the day/night) and 40% relative humidity. Four days after transplantation, one of the seedlings from each pair was inoculated with *Agrobacterium* harboring the pBINTRA vector along with a VIGS vector containing gene fragments related to triterpene biosynthesis (*pTV-OSC1/2*, *pTV-CYP716A419*, *pTV-CYP716E107*, or *pTV-CYP716C87*). Thirty plants were inoculated with each vector. The neighboring plants of these individuals were inoculated with *Agrobacterium* containing EVs and served as negative controls within each group. Pots with 2 EV plants were used as negative controls. Plants inoculated with *Agrobacterium* carrying the PDS gene served as positive controls, and the onset of leaf bleaching was used to time leaf harvest. Fourteen days after *Agrobacterium* inoculation, plants without apparent viral phenotypes were removed from the study. Each experimental group retained 21 to 24 replicate plants.

Transcript abundance analysis

The total RNA was extracted from the plant tissues of *N. attenuata* by utilizing the plant RNA purification kit (Macherey-Nagel) following the manufacturer's instructions. cDNA was generated from total RNA using PrimeScript RT Master Mix (Takara Bio Inc., Japan). RT-qPCR was performed on a Stratagene Mx3005P qPCR machine using a Takyon No ROX SYBR 2X MasterMix Blue dTTP (Eurogentec, Seraing, Belgium). The housekeeping gene IF-5α from *N. attenuata* was used as an internal reference. The primers used for RT-qPCR are listed in Supplementary Table S7.

GC-MS analysis of triterpenes

Triterpene extraction followed established procedures outlined in a prior publication (Field and Osbourn 2008) with minor modifications. For triterpene extraction from leaves and corollas of *N. attenuata*, approximately 1 g of fresh plant tissue (corolla excluding the reproductive organs and leaf disks prepared with a 2-cm-diameter hole punch) was subjected to 2 sequential 1-min washes in 10 mL of *n*-hexane containing 200 ng of α-amyrin as internal standard. The resulting extracts were vacuum dried and subsequently saponified in 500 μL of a saponification buffer (20% KOH [w/v] in 50% EtOH [v/v] with 0.5 mg/mL butylated hydroxytoluene) for 2 h at 65 °C. Subsequently, 100 μL of 10 M HCl was added to the aqueous solution to lower the pH to below 2.0, and the mixture was extracted 3 times with 500 μL of hexane. The resulting extracts were vacuum dried and subjected to derivatization using a mixture of MSTFA and DMF before GC-MS analysis.

For the preparation of pollen and thecal samples, flowers with opened corollas with matured anthers were selected, and anther heads were collected into 5-mL Eppendorf tubes (EP tubes). Following the addition of 1 mL of double-distilled water (ddH_2O), samples were vortexed for 30 s followed by a 5-min incubation period. The thecae were then separated into another 5-mL EP tube using a small spatula. After centrifugation at 11,000 × *g* for 15 min at 4 °C to remove the water layer, 2 steel beads (2 mm diameter) were added to the tube. The sample was rapidly frozen in liquid nitrogen and homogenized in a ball mill (Genogrinder 2000; SPEX CertiPrep) for 60 s at a rate of 1,100 strokes per minute.

For triterpene extraction of pollen, thecae, and *N. benthamiana* leaves, frozen and ground fresh samples (30 mg for pollen and thecae, and 100 mg for *N. benthamiana* leaves) with 200 ng α-amyrin internal standard were saponified in 300 μL of the saponification buffer described above at 65 °C for 2 h. Subsequently, 70 μL of 10 M HCl was added to the aqueous solution to lower the pH below 2.0, followed by 3 extractions with 300 μL of hexane. The extracts were then

concentrated and derivatized with MSTFA and DMF before GC-MS analysis. For *N. benthamiana* leaves requiring separate detection of alkaline and acidic extracts, after saponification, 300 μL of hexane was added to the saponification reaction mixture. The extraction procedure was repeated 3 times, and the resulting extracts were combined to obtain the alkaline extract. Following saponification, the pH of the reaction mixture was adjusted to around 2 by adding 10 M HCl. Subsequently, 300 μL of hexane was added to the saponification reaction mixture, and the extraction procedure was repeated 3 times. The resulting extracts were combined to obtain the acidic extract. Alkaline and acidic extracts were then derivatized separately for GC-MS analysis. The GC-MS analysis was conducted using the same instrument, columns, temperature programs, and MS settings as in previous studies (Yang et al. 2023b).

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 23 (IBM Inc., Chicago, Illinois, United States). Statistical differences among groups were determined using ANOVA followed by Tukey's honestly significant difference (HSD) post hoc test. The significance of differences between the 2 sample groups was assessed using the Student's *t* test. A significance level of $P \leq 0.05$ was considered statistically significant for all comparisons. The Kruskal–Wallis test was used when the data did not meet the normality or homogeneity of variance assumptions required for ANOVA. The box plots and correlation analysis were conducted using Chiplot (https://www.chiplot.online/).

Accession numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers NaOSC1 (LOC109226501), NaOSC2 (LOC109226503), CYP716A419 (LOC109234271), CYP716E107 (LOC109209296), CYP716E108 (LOC109209295), CYP716C87 (LOC109239546), CYP716D93 (LOC109205849), CYP716D94 (OIT28874), CYP716H6 (LOC109230529), CYP716A420 (LOC109236645), CYP88B4 (LOC109228169), and CYP88C16 (LOC109244115).

Acknowledgments

We acknowledge Dr. Xincong Jiang for photographing the pollen. We acknowledge Prof. David R. Nelson for his contributions in naming candidate CYP450 enzymes in the manuscript.

Author contributions

C.Y. conceived the study, conducted the experiments, and analyzed the data. I.T.B. and S.E.O. supervised the project. C.Y. and I.T.B. wrote the manuscript. R.H. provided valuable technical support for metabolite analysis. All the authors commented on the manuscript.

Supplementary data

The following materials are available in the online version of this article.

Supplementary Figure S1. p450 enzyme from *N. attenuata* plants.

Supplementary Figure S2. Sequence alignment of *N. attenuata* CYP450 candidates with CYP450 enzymes from the triterpenoid biosynthetic pathway.

Supplementary Figure S3. The levels of triterpenes after coexpression of *AtLPU1/NaOSC2* with different CYP450s.

Supplementary Figure S4. Primary fragmentation pathways of TMS derivatives of oleanolic aldehyde.

Supplementary Figure S5. The triterpenoid metabolite profiles of *N. benthamiana* leaves expressing either EV or *NaCYP716A419*.

Supplementary Figure S6. The levels of triterpenes after coexpression of *NaOSC2* and *NaCYP716A419* with *NaCYP716E107*, *SlCYP716E26*, or *NaCYP716C87*.

Supplementary Figure S7. Primary fragmentation pathways of TMS derivatives of β-amyrin.

Supplementary Figure S8. Primary fragmentation pathways of TMS derivatives of maslinic acid.

Supplementary Figure S9. Primary fragmentation pathways of TMS derivatives of erythrodiol.

Supplementary Figure S10. Primary fragmentation pathways of TMS derivatives of lupeol.

Supplementary Figure S11. Primary fragmentation pathways of TMS derivatives of lupanediol.

Supplementary Figure S12. The levels of triterpenes after coexpression of *AtLPU1* and *NaCYP716A419* with *NaCYP716E107* or *NaCYP716C87*.

Supplementary Figure S13. Primary fragmentation pathways of TMS derivatives of oleanolic acid.

Supplementary Figure S14. Primary fragmentation pathways of TMS derivatives of betulinic acid.

Supplementary Figure S15. Alignment of protein sequences of NaCYP716E107, SlCYP716E26, and CaCYP716E41.

Supplementary Figure S16. The levels of triterpenes after expressed *NaCYP716E107* or *SlCYP716E26*.

Supplementary Figure S17. Silencing efficiency of 5 triterpene biosynthesis genes in VIGS plants.

Supplementary Figure S18. 2,3,5-Triphenyltetrazolium chloride (TTC) staining of pollen from EV, *NaOSC1/2_VIGS*, and *CYP716A419_VIGS* plants.

Supplementary Figure S19. GC-MS chromatograms of extracts of *N. attenuata* roots with phytohormone treatments.

Supplementary Figure S20. Products of NaCYP716A419, NaCYP716C87, and NaCYP716E107 in *N. attenuata* roots (infected with AMF for 5 wk).

Supplementary Figure S21. GC-MS chromatograms of reaction products of *NaOSC2/CYP716A419* coexpressed with other CYP450 candidates, aligned with standards.

Supplementary Figure S22. GC-MS chromatograms of reaction products of *NaOSC2/CYP716A419/CYP716C87* or *NaOSC2/CYP716A419/CYP716E107* coexpressed with other CYP450 candidates in *N. benthamiana*, aligned with authentic standards.

Supplementary Table S1. CYP enzymes used for phylogenetic tree.

Supplementary Table S2. CYP enzymes from *N. attenuata*.

Supplementary Table S3. The level of confidence for β-amyrin derivates identification.

Supplementary Table S4. The level of confidence for lupeol and lupanediol derivates identification.

Supplementary Table S5. Primer sequences used for the design of constructs for transient expression.

Supplementary Table S6. Primer sequences for VIGS.

Supplementary Table S7. Primer sequences for qPCR.

Funding

This work was supported by the China Scholarship Council (no. 201906910083) and the Max Planck Society.
Conflict of interest statement. None declared.

Data availability

All data are incorporated into this article and its online supplementary material.

References

Alagna F, Reed J, Calderini O, Thimmappa R, Cultrera NGM, Cattivelli A, Tagliazucchi D, Mousavi S, Mariotti R, Osbourn A, *et al.* OeBAS and CYP716C67 catalyze the biosynthesis of health-beneficial triterpenoids in olive (*Olea europaea*) fruits. *New Phytol*. 2023:238(5): 2047–2063. https://doi.org/10.1111/nph.18863

Andre CM, Legay S, Deleruelle A, Nieuwenhuizen N, Punter M, Brendolise C, Cooney JM, Lateur M, Hausman J-F, Larondelle Y, *et al.* Multifunctional oxidosqualene cyclases and cytochrome P450 involved in the biosynthesis of apple fruit triterpenic acids. *New Phytol*. 2016:211(4):1279–1294. https://doi.org/10.1111/nph.13996

Baldwin IT, Hamilton W. Jasmonate-induced responses of *Nicotiana sylvestris* results in fitness costs due to impaired competitive ability for nitrogen. *J Chem Ecol*. 2000:26(4):915–952. https://doi.org/10.1023/A:1005408208826

Brockmöller T, Ling Z, Li D, Gaquerel E, Baldwin IT, Xu S. *Nicotiana attenuata* Data Hub (NaDH): an integrative platform for exploring genomic, transcriptomic and metabolomic data in wild tobacco. *BMC Genomics*. 2017:18(1):79. https://doi.org/10.1186/s12864-016-3465-9

Cárdenas PD, Sonawane PD, Heinig U, Jozwiak A, Panda S, Abebie B, Kazachkova Y, Pliner M, Unger T, Wolf D, *et al.* Pathways to defense metabolites and evading fruit bitterness in genus *Solanum* evolved through 2-oxoglutarate-dependent dioxygenases. *Nat Commun*. 2019:10(1):5169. https://doi.org/10.1038/s41467-019-13211-4

Carelli M, Biazzi E, Panara F, Tava A, Scaramelli L, Porceddu A, Graham N, Odoardi M, Piano E, Arcioni S, *et al. Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell*. 2011:23(8):3070–3081. https://doi.org/10.1105/tpc.111.087312

Dai Z, Liu Y, Sun Z, Wang D, Qu G, Ma X, Fan F, Zhang L, Li S, Zhang X. Identification of a novel cytochrome P450 enzyme that catalyzes the C-2α hydroxylation of pentacyclic triterpenoids and its application in yeast cell factories. *Metab Eng*. 2019:51:70–78. https://doi.org/10.1016/j.ymben.2018.10.001

Dou D-Q, Chen Y-J, Liang L-H, Pang F-G, Shimizu N, Takeda T. Six new dammarane-type triterpene saponins from the leaves of *panax ginseng*. *Chem Pharm Bull*. 2001:49(4):442–446. https://doi.org/10.1248/cpb.49.442

Field B, Fiston-Lavier A-S, Kemen A, Geisler K, Quesneville H, Osbourn AE. Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc Natl AcadScie, USA*. 2011:108-(38):16116–16121. https://doi.org/10.1073/pnas.1109273108

Field B, Osbourn AE. Metabolic diversification—independent assembly of operon-like gene clusters in different plants. *Science*. 2008:320(5875):543–547. https://doi.org/10.1126/science.1154990

Fukushima EO, Seki H, Ohyama K, Ono E, Umemoto N, Mizutani M, Saito K, Muranaka T. CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol*. 2011:52(12):2050–2061. https://doi.org/10.1093/pcp/pcr146

Fukushima EO, Seki H, Sawai S, Suzuki M, Ohyama K, Saito K, Muranaka T. Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant Cell Physiol*. 2013:54(5):740–749. https://doi.org/10.1093/pcp/pct015

Geisler K, Hughes RK, Sainsbury F, Lomonossoff GP, Rejzek M, Fairhurst S, Olsen CE, Motawia MS, Melton RE, Hemmings AM, *et al.* Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc Natl Acad Sci USA*. 2013:110(35): E3360–E3367. https://doi.org/10.1073/pnas.1309157110

Ghosh S. Triterpene structural diversification by plant cytochrome P450 enzymes. *Front Plant Sci*. 2017:8:1886. https://doi.org/10.3389/fpls.2017.01886

González-Coloma A, López-Balboa C, Santana O, Reina M, Fraga BM. Triterpene-based plant defenses. *Phytochem Rev*. 2011:10(2): 245–260. https://doi.org/10.1007/s11101-010-9187-8

Güçlü-Üstündağ Ö, Mazza G. Saponins: properties, applications and processing. *Crit Rev Food Sci Nutr*. 2007:47(3):231–258. https://doi.org/10.1080/10408390600698197

Guhling O, Hobl B, Yeats T, Jetter R. Cloning and characterization of a lupeol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of Ricinus communis. *Arch Biochem Biophys*. 2006:448(1-2):60–72. https://doi.org/10.1016/j.abb.2005.12.013

Han J-Y, Hwang H-S, Choi S-W, Kim H-J, Choi Y-E. Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol*. 2012:53(9):1535–1545. https://doi.org/10.1093/pcp/pcs106

Han J-Y, Kim H-J, Kwon Y-S, Choi Y-E. The cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol*. 2011:52(12):2062–2073. https://doi.org/10.1093/pcp/pcr150

Heiling S, Llorca LC, Li J, Gase K, Schmidt A, Schäfer M, Schneider B, Halitschke R, Gaquerel E, Baldwin IT. Specific decorations of 17-hydroxygeranyllinalool diterpene glycosides solve the autotoxicity problem of chemical defense in *Nicotiana attenuata*. *Plant Cell*. 2021:33(5):1748–1770. https://doi.org/10.1093/plcell/koab048

Hong B, Grzech D, Caputi L, Sonawane P, López CER, Kamileen MO, Hernández Lozada NJ, Grabe V, O'Connor SE. Biosynthesis of strychnine. *Nature*. 2022:607(7919):617–622. https://doi.org/10.1038/s41586-022-04950-4

Hu C, Tang Y, Snooks HD, Sang S. Novel steroidal saponins in oat identified by molecular networking analysis and their levels in commercial oat products. *J Agric Food Chem*. 2021:69(25): 7084–7092. https://doi.org/10.1021/acs.jafc.1c02728

Ito R, Mori K, Hashimoto I, Nakano C, Sato T, Hoshino T. Triterpene cyclases from *Oryza sativa* L.: cycloartenol, parkeol and achilleol B synthases. *Org Lett*. 2011:13(10):2678–2681. https://doi.org/10.1021/ol200777d

Jo H-J, Han JY, Hwang H-S, Choi YE. β-Amyrin synthase (EsBAS) and β-amyrin 28-oxidase (CYP716A244) in oleanane-type triterpene saponin biosynthesis in *Eleutherococcus senticosus*. *Phytochemistry*. 2017:135:53–63. https://doi.org/10.1016/j.phytochem.2016.12.011

Joo Y, Kim H, Kang M, Lee G, Choung S, Kaur H, Oh S, Choi JW, Ralph J, Baldwin IT, *et al.* Pith-specific lignification in *Nicotiana attenuata* as a defense against a stem-boring herbivore. *New Phytol*. 2021:232(1):332–344. https://doi.org/10.1111/nph.17583

Kemen AC, Honkanen S, Melton RE, Findlay KC, Mugford ST, Hayashi K, Haralampidis K, Rosser SJ, Osbourn A. Investigation of triterpene synthesis and regulation in oats reveals a role for β-amyrin in determining root epidermal cell patterning. *Proc Natl Acad Sci USA*. 2014:111(23):8679–8684. https://doi.org/10.1073/pnas.1401553111

Kessler A, Baldwin IT. Herbivore-induced plant vaccination. Part I. The orchestration of plant defenses in nature and their fitness consequences in the wild tobacco *Nicotiana attenuata*. *Plant J*.

2004:38(4):639–649. https://doi.org/10.1111/j.1365-313X.2004.02076.x

Khakimov B, Kuzina V, Erthmann PØ, Fukushima EO, Augustin JM, Olsen CE, Scholtalbers J, Volpin H, Andersen SB, Hauser TP, *et al.* Identification and genome organization of saponin pathway genes from a wild crucifer, and their use for transient production of saponins in *Nicotiana benthamiana*. *Plant J*. 2015:84(3):478–490. https://doi.org/10.1111/tpj.13012

Krügel T, Lim M, Gase K, Halitschke R, Baldwin IT. Agrobacterium-mediated transformation of *Nicotiana attenuata*, a model ecological expression system. *Chemoecology*. 2002:12(4):177–183. https://doi.org/10.1007/PL00012666

Kudou S, Tonomura M, Tsukamoto C, Shimoyamada M, Uchida T, Okubo K. Isolation and structural elucidation of the major genuine soybean saponin. *Biosci Biotechnol Biochem*. 1992:56(1):142–143. https://doi.org/10.1271/bbb.56.142

Leveau A, Reed J, Qiao X, Stephenson MJ, Mugford ST, Melton RE, Rant JC, Vickerstaff R, Langdon T, Osbourn A. Towards take-all control: a C-21β oxidase required for acylation of triterpene defence compounds in oat. *New Phytol*. 2019:221(3):1544–1555. https://doi.org/10.1111/nph.15456

Li J, Halitschke R, Li D, Paetz C, Su H, Heiling S, Xu S, Baldwin IT. Controlled hydroxylations of diterpenoids allow for plant chemical defense without autotoxicity. *Science*. 2021:371(6526):255–260. https://doi.org/10.1126/science.abe4713

Miettinen K, Pollier J, Buyst D, Arendt P, Csuk R, Sommerwerk S, Moses T, Mertens J, Sonawane PD, Pauwels L, *et al.* The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. *Nat Commun*. 2017:8(1):14153. https://doi.org/10.1038/ncomms14153

Misra RC, Sharma S, Sandeep GA, Chanotiya CS, Ghosh S. Two CYP716A subfamily cytochrome P450 monooxygenases of sweet basil play similar but nonredundant roles in ursane- and oleanane-type pentacyclic triterpene biosynthesis. *New Phytol*. 2017:214(2):706–720. https://doi.org/10.1111/nph.14412

Moses T, Pollier J, Faizal A, Apers S, Pieters L, Thevelein Johan M, Geelen D, Goossens A. Unraveling the triterpenoid saponin biosynthesis of the African shrub Maesa lanceolata. *Mol Plant*. 2015a:8(1):122–135. https://doi.org/10.1016/j.molp.2014.11.004

Moses T, Pollier J, Shen Q, Soetaert S, Reed J, Erffelinck ML, Van Nieuwerburgh FC, Vanden Bossche R, Osbourn A, Thevelein JM, *et al.* OSC2 and CYP716A14v2 catalyze the biosynthesis of triterpenoids for the cuticle of aerial organs of *Artemisia annua*. *Plant Cell*. 2015b:27(1):286–301. https://doi.org/10.1105/tpc.114.134486

Moses T, Thevelein JM, Goossens A, Pollier J. Comparative analysis of CYP93E proteins for improved microbial synthesis of plant triterpenoids. *Phytochemistry*. 2014:108:47–56. https://doi.org/10.1016/j.phytochem.2014.10.002

Nakamura M, Linh TM, Lien LQ, Suzuki H, Mai NC, Giang VH, Tamura K, Thanh NV, Suzuki H, Misaki R, *et al.* Transcriptome sequencing and identification of cytochrome P450 monooxygenases involved in the biosynthesis of maslinic acid and corosolic acid in *Avicennia marina*. *Plant Biotechnol*. 2018:35(4):341–348. https://doi.org/10.5511/plantbiotechnology.18.0810a

Naoumkina MA, Modolo LV, Huhman DV, Urbanczyk-Wochniak E, Tang Y, Sumner LW, Dixon RA. Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Cell*. 2010:22(3):850–866. https://doi.org/10.1105/tpc.109.073270

Ohyama K, Suzuki M, Masuda K, Yoshida S, Muranaka T. Chemical phenotypes of the hmg1 and hmg2 mutants of arabidopsis demonstrate the in-planta role of HMG-CoA reductase in triterpene biosynthesis. *Chem Pharm Bull*. 2007:55(10):1518–1521. https://doi.org/10.1248/cpb.55.1518

Popova V, Ivanova T, Stoyanova A, Georgiev V, Hristeva T, Nikolova V, Docheva M, Nikolov N, Damianova S. Phytochemicals in leaves and extracts of the variety "Plovdiv 7" of Bulgarian oriental tobacco (*Nicotiana tabacum* L.). *Trends Phytochem Res*. 2018:2:27–36. https://doi.org/20.1001.1.25883623.2018.2.1.4.2

Popova VT, Ivanova TA, Stoyanova AS, Nikolova VV, Docheva MH, Hristeva TH, Damyanova ST, Nikolov NP. Chemical constituents in leaves and aroma products of *Nicotiana rustica* L. tobacco. *Int J Food Studies*. 2020:9(1):146–159. https://doi.org/10.7455/ijfs/9.1.2020.a2

Popova V, Ivanova T, Stoyanova A, Nikolova V, Hristeva T, Docheva M, Nikolov N, Iliev I. Polyphenols and triterpenes in leaves and extracts from three *Nicotiana* species. *J Appl Biol Biotechnol*. 2019:7(5):45–49. https://doi.org/10.7324/JABB.2019.70508

Rahman A, Ahamed A, Amakawa T, Goto N, Tsurumi S. Chromosaponin I specifically interacts with AUX1 protein in regulating the gravitropic response of arabidopsis roots. *Plant Physiol*. 2001:125(2):990–1000. https://doi.org/10.1104/pp.125.2.990

Reed J, Orme A, El-Demerdash A, Owen C, Martin LBB, Misra RC, Kikuchi S, Rejzek M, Martin AC, Harkess A, *et al.* Elucidation of the pathway for biosynthesis of saponin adjuvants from the soapbark tree. *Science*. 2023:379(6638):1252–1264. https://doi.org/10.1126/science.adf3727

Romsuk J, Yasumoto S, Seki H, Fukushima EO, Muranaka T. Identification of key amino acid residues toward improving the catalytic activity and substrate specificity of plant-derived cytochrome P450 monooxygenases CYP716A subfamily enzyme for triterpenoid production in *Saccharomyces cerevisiae*. *Front Bioeng Biotechnol*. 2022:10:955650. https://doi.org/10.3389/fbioe.2022.955650

Saedler R, Baldwin IT. Virus-induced gene silencing of jasmonate-induced direct defences, nicotine and trypsin proteinase-inhibitors in *Nicotiana attenuata*. *J Exp Bot*. 2004:55(395):151–157. https://doi.org/10.1093/jxb/erh004

Sandeep MR, Chanotiya CS, Mukhopadhyay P, Ghosh S. Oxidosqualene cyclase and CYP716 enzymes contribute to triterpene structural diversity in the medicinal tree banaba. *New Phytol*. 2019:222(1):408–424. https://doi.org/10.1111/nph.15606

Segura MJR, Meyer MM, Matsuda SPT. *Arabidopsis thaliana* LUP1 converts oxidosqualene to multiple triterpene alcohols and a triterpene diol. *Org Lett*. 2000:2(15):2257–2259. https://doi.org/10.1021/ol006016b

Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T. Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc Natl Acad Sci USA*. 2008:105(37):14204–14209. https://doi.org/10.1073/pnas.0803876105

Seki H, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, Fukushima EO, Akashi T, Aoki T, Saito K, *et al.* Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell*. 2011:23(11):4112–4123. https://doi.org/10.1105/tpc.110.082685

Shang Y, Ma Y, Zhou Y, Zhang H, Duan L, Chen H, Zeng J, Zhou Q, Wang S, Gu W, *et al.* Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science*. 2014:346(6213):1084–1088. https://doi.org/10.1126/science.1259215

Sharma K, Kaur R, Kumar S, Saini RK, Sharma S, Pawde SV, Kumar V. Saponins: a concise review on food related aspects, applications and health implications. *Food Chem Adv*. 2023:2:100191. https://doi.org/10.1016/j.focha.2023.100191

Srisawat P, Fukushima EO, Yasumoto S, Robertlee J, Suzuki H, Seki H, Muranaka T. Identification of oxidosqualene cyclases from the

medicinal legume tree *Bauhinia forficata*: a step toward discovering preponderant α-amyrin-producing activity. *New Phytol*. 2019:224(1):352–366. https://doi.org/10.1111/nph.16013

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 2021:38(7):3022–3027. https://doi.org/10.1093/molbev/msab120

Tran QL, Adnyana IK, Tezuka Y, Nagaoka T, Tran QK, Kadota S. Triterpene saponins from Vietnamese ginseng (*Panax vietnamensis*) and their hepatocytoprotective activity. *J Nat Prod*. 2001:64(4):456–461. https://doi.org/10.1021/np000393f

Tsurumi S, Takagi T, Hashimoto T. A γ-pyronyl-triterpenoid saponin from *Pisum sativum*. *Phytochemistry*. 1992:31(7):2435–2438. https://doi.org/10.1016/0031-9422(92)83294-9

Wang P. Natural and synthetic saponins as vaccine adjuvants. *Vaccines (Basel)*. 2021:9(3):222. https://doi.org/10.3390/vaccines9030222

Wang Z, Guhling O, Yao R, Li F, Yeats TH, Rose JK, Jetter R. Two oxidosqualene cyclases responsible for biosynthesis of tomato fruit cuticular triterpenoids. *Plant Physiol*. 2011:155(1):540–552. https://doi.org/10.1104/pp.110.162883

Xie J, Chen Y, Cai G, Cai R, Hu Z, Wang H. Tree visualization by one table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res*. 2023:51(W1):W587–W592. https://doi.org/10.1093/nar/gkad359

Yang C, Bai Y, Halitschke R, Gase K, Baldwin G, Baldwin IT. Exploring the metabolic basis of growth/defense trade-offs in complex environments with *Nicotiana attenuata* plants cosilenced in *NaMYC2a/b* expression. *New Phytol*. 2023a:238(1):349–366. https://doi.org/10.1111/nph.18732

Yang C, Halitschke R, O'Connor SE. OXIDOSQUALENE CYCLASE 1 and 2 influence triterpene biosynthesis and defense in *Nicotiana attenuata*. *Plant Physiol*. 2023b:194(4):2580–2599. https://doi.org/10.1093/plphys/kiad643

Yasumoto S, Fukushima EO, Seki H, Muranaka T. Novel triterpene oxidizing activity of *Arabidopsis thaliana* CYP716A subfamily enzymes. *FEBS Lett*. 2016:590(4):533–540. https://doi.org/10.1002/1873-3468.12074

Yasumoto S, Seki H, Shimizu Y, Fukushima EO, Muranaka T. Functional characterization of CYP716 family P450 enzymes in triterpenoid biosynthesis in tomato. *Front Plant Sci*. 2017:8:21. https://doi.org/10.3389/fpls.2017.00021

You Y, Ray R, Halitschke R, Baldwin G, Baldwin IT. Arbuscular mycorrhizal fungi-indicative blumenol-C-glucosides predict lipid accumulations and fitness in plants grown without competitors. *New Phytol*. 2023:238(5):2159–2174. https://doi.org/10.1111/nph.18858

Yu B, Patterson N, Zaharia LI. Saponin biosynthesis in pulses. *Plants*. 2022:11(24):3505. https://doi.org/10.3390/plants11243505

Zhou C, Li J, Li C, Zhang Y. Improvement of betulinic acid biosynthesis in yeast employing multiple strategies. *BMC Biotechnol*. 2016:16(1):59. https://doi.org/10.1186/s12896-016-0290-9