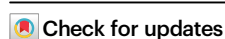


Divergent multifunctional P450s-empowered biosynthesis of bioactive tripterifordin and cryptic atiserenoids in *Aconitum* implies convergent evolution

Received: 9 December 2024

Accepted: 9 June 2025

Published online: 01 July 2025



A list of authors and their affiliations appears at the end of the paper

Diterpenoids exhibit remarkable structural diversity and bioactivities, shaped primarily by the tandem actions of skeleton-forming terpene synthases (TPSs) and cytochrome P450 monooxygenases. The *ent*-kaurene and *ent*-atiserene are labdane-derived diterpene scaffolds for the biosynthesis of diverse bioactive diterpenoids and diterpene alkaloids, including the clinically used analgesic 3-acetylaconitine and anti-arrhythmic guan-fu base A in *Aconitum* spp., yet what and how P450s drive their structural and functional diversification remain largely unexplored. Here, via mining the transcriptomes of *Aconitum carmichaelii* and *Aconitum coreanum* followed by functional validation, we discover nine TPSs capable of furnishing the *ent*-kaurene, *ent*-atiserene or 16 α -hydroxy-*ent*-kaurene scaffold and 14 divergent P450s, of which eight are multifunctional, catalyzing oxidation at seven different sites of the scaffolds. Protein analysis and mutagenesis experiments reveal key residues tuning the P450 activity and product profiles, shedding light on how they diverge functionally. The discovered TPSs and P450s enable combinatorial biosynthesis of tripterifordin, guan-fu diterpenoid A, and 14 novel atiserenoids, with some exhibiting allelopathic activity. Intriguingly, *Tripterygium wilfordii*-derived tripterifordin was detected in *A. coreanum*, revealing convergent biosynthesis of tripterifordin. Our findings highlight P450 plasticity in selective diterpene oxidation and allow for accessing tripterifordin and derivatives, paving ways for elucidating downstream diterpenoid pathways via synthetic biology.

With over 18,000 structures and 126 distinct scaffolds, diterpenes (C₂₀) constitute one large subclass of the terpene family¹. Plants are a major source of bioactive diterpenes, with notable examples ranging from anticancer agents - paclitaxol², ingenol mebutate³, triptolide⁴ and oridonin⁵ and tripterifordin⁶ with anti-HIV potential to sweetener

stevioside⁷, plant hormone gibberellins⁸ and phytoalexins momilactones^{9,10} (Supplementary Fig. 1). Many of the bioactive diterpenes contain a bicyclic core and are collectively referred to as labdane-related diterpenes (LRDs) (Supplementary Fig. 1), which encompass approximately 7000 compounds and represent a

 e-mail: huangac@sustech.edu.cn

substantial proportion of natural diterpenes discovered to date¹¹. Diterpenes are biosynthesized from (*E,E,E*)-geranylgeranyl diphosphate (GGPP) derived from the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway in plastids in plants^{12–14}. The labdane bicyclic core is

constructed via protonation of the linear GGPP by class II diterpene synthases (diTPSs)¹⁵, yielding bicyclic diterpene diphosphates such as copalyl diphosphate (CPP), *ent*-CPP, *syn*-CPP, or 8 α -hydroxy-CPP (Fig. 1). Subsequent class I diTPS-catalyzed ionization of the cyclized

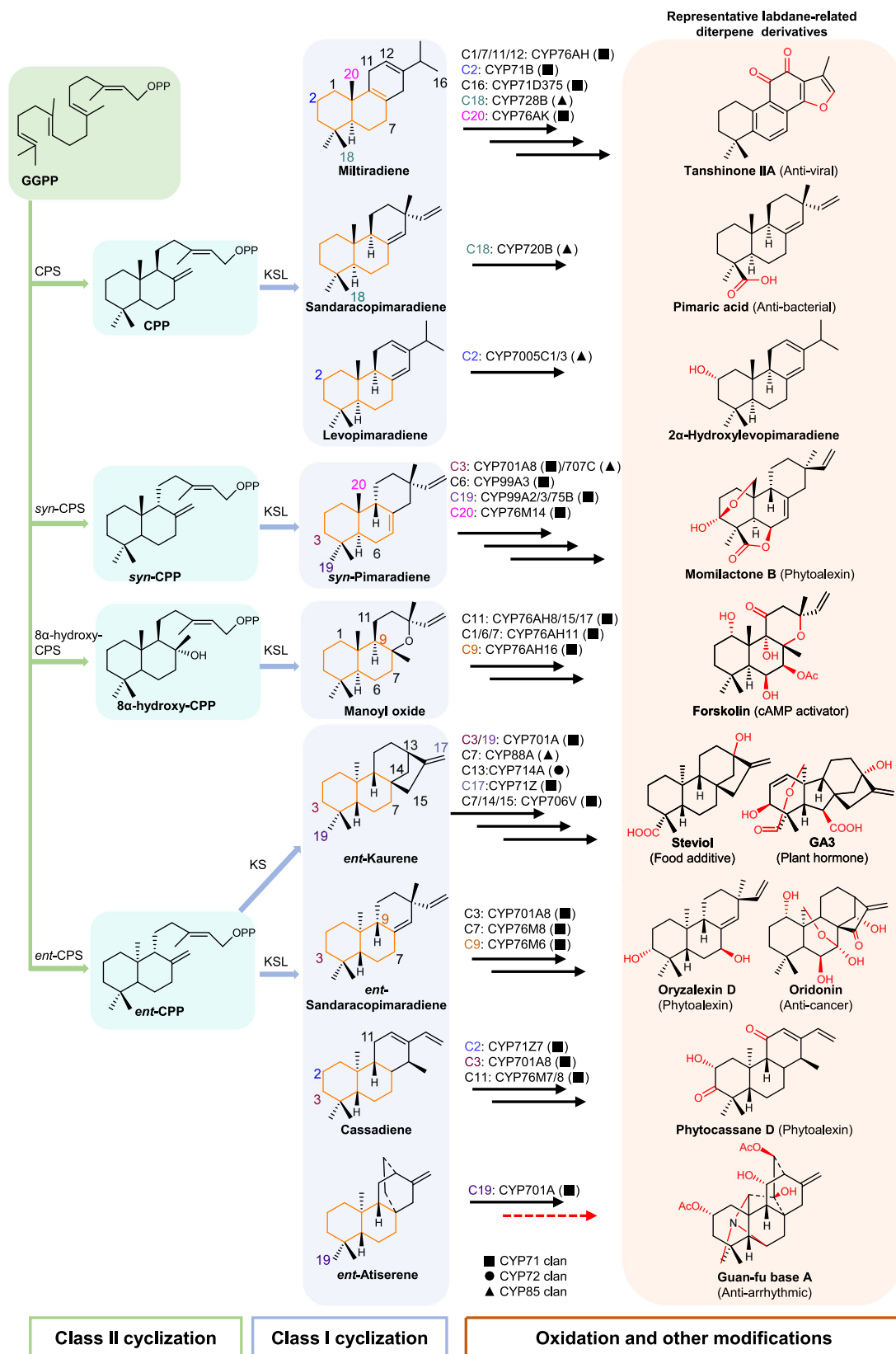


Fig. 1 | Representative labdane-related diterpene scaffolds and CYPs involved in selective oxidations identified in plants. Only one CYP701A has been found to act on the C19 position of *ent*-atiserene so far. GGPP, geranylgeranyl diphosphate. CPP copalyl diphosphate.

labdane diphosphates followed by carbocationic cyclization generates miscellaneous LRD skeletons (Supplementary Fig. 2) for further modifications¹⁶. Due to the rigidity of the bicyclic core, carbocation cyclization of the labdane diterpene diphosphates can generate relatively limited skeletal diversity.

The structural diversification of LRDs is to a large extent driven by downstream modifications of the skeletons, in particular oxidation by cytochrome P450 monooxygenases (CYPs)¹⁷. Previous studies have identified ~123 plant P450s (summarized in Supplementary Data 1) from 22 CYP subfamilies capable of acting on the different sites of LRD skeletons (Supplementary Fig. 3), with the majority of the CYP subfamilies (including CYP76^{18,19}, CYP75B²⁰, CYP70I²¹, CYP706V²¹, CYP81A²² and CYP82D²³) falling into the CYP71 clan, five (CYP88A, CYP720B, CYP728B, CYP7005C and CYP707C^{20,24}) in the CYP85 and only two (CYP714 and CYP72D)²⁵ in the CYP72 clan (Fig. 1). Although a large number of P450s have been identified to be able to act on different LRD skeletons, P450s acting on the *ent*-atiserene scaffold are rarely reported in plants, with only one promiscuous P450 in the CYP701A subfamily from plants²⁶ found to be able to act on the C19 position so far. A few other oxidases from bacteria are also capable of acting on the *ent*-atiserene scaffold in platencin biosynthesis^{27,28}, with those acting on the C7 position non-P450s²⁷ while those on the C19 position likely P450s²⁸. This suggests that the biosynthesis of *ent*-atiserene diterpenes remains scarcely investigated, albeit they constitute a large branch of LRDs with considerable bioactivities²⁹.

Aconitum carmichaelii and *Aconitum coreanum* are two medicinal plants in the Ranunculaceae family that have been widely practiced in Chinese traditional medicine, with *A. carmichaelii* known as “Fuzi” for healing shock resulting from heart-related diseases³⁰ and *A. coreanum* as “Guan Bai fu” for treating stroke, pain, rheumatic arthritis, and arrhythmia³¹. They are rich in bioactive C19 and C20 diterpene alkaloids derived from *ent*-kaurene and *ent*-atiserene skeletons (Supplementary Fig. 4)^{30–32}. Some of the diterpene alkaloids have been developed into drugs approved by the Chinese Food and Drug Administration (CFDA) for clinical applications, e.g., 3-acetylacetonine for analgesia³² and Guan-fu base A (also named acehytisine) for antiarrhythmia^{31,33}. The diterpene alkaloids present in these two *Aconitum* spp. are highly oxidized at various sites of the skeletons with nitrogen incorporated³², indicative of the remarkable latent oxidation capacity encoded by their genomes that is yet to be unlocked.

Recent studies have identified some diTPSs from *A. carmichaelii* and other *Aconitum* spp. capable of generating LRD scaffolds including *ent*-kaurene, *ent*-atiserene, *ent*-13-*epi*-sandaracopimaradiene, and Z-abienol^{34,35}. However, the enzymes responsible for downstream modifications of these diterpene skeletons are not yet discovered. Whether and how P450s drive the structural diversification of diterpenes in *Aconitum* spp. remain unclear. Therefore, mining the *Aconitum* spp. genome for functional P450s acting on LRD skeletons will shed light on the catalytic basis underlying diterpene structural diversification and diterpene alkaloid biosynthesis in *Aconitum* spp., and expand the enzymatic tools for the synthesis of complex bioactive diterpenoids.

Here, via RNA-seq and de novo assembly of the transcriptomes of *A. carmichaelii* and *A. coreanum* followed by systematic transcriptome mining for TPSs and coexpressed P450s, we identified nine diTPSs capable of generating *ent*-kaurene, *ent*-atiserene, and 16 α -hydroxy-*ent*-kaurene, together with 14 functional P450s in the CYP71, 85, and 72 clans that could catalyze oxidation at seven different sites of the diterpene scaffolds. In addition, eight of the 14 P450s are multifunctional, capable of acting on multiple sites of the *ent*-atiserene and *ent*-kaurene scaffolds. To date, only a small number of multifunctional P450s acting on LRD skeletons have been identified before (Supplementary Data 1), the number and functional diversity of the multifunctional P450s discovered here from *A. carmichaelii* and *A. coreanum* are unexpected and highlight the remarkable plasticity of

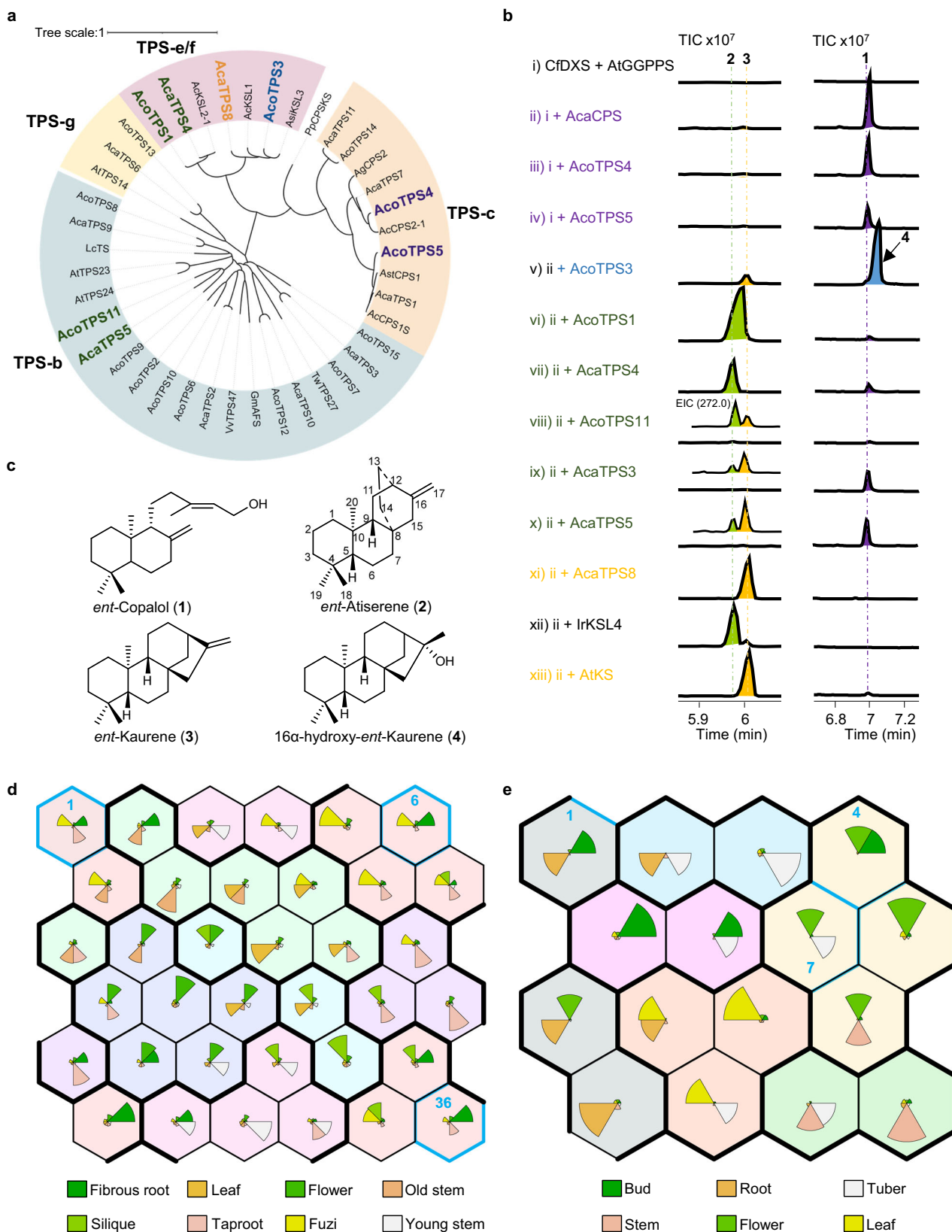
P450s in diterpene biosynthesis. The functional TPSs and P450s from *Aconitum* spp. further open avenues to access bioactive tripterifordin⁶, *ent*-kaurene, and *ent*-atiserene diterpenoids with novel structures through combinatorial biosynthesis. Our findings thus lay the foundation for uncovering the diterpene alkaloid biosynthetic pathways and showcase the substantial latent biosynthetic capacity for cryptic diterpene biosynthesis in *Aconitum* spp.

Results

Discovery of *Aconitum* terpene synthases (TPSs) catalyzing the formation of *ent*-atiserene, *ent*-kaurene, and 16 α -hydroxy-*ent*-kaurene

To mine TPSs and P450s from *Aconitum* spp. genomes, we performed RNA sequencing (RNA-seq) of the different tissues of *A. carmichaelii* (leaf, young stem, old stem, fibrous root, tap root, Fuzi, flower, and silique) and *A. coreanum* (leaf, stem, root, tuber, bud, and flower) and generated the transcriptomics data via de novo assembly of the RNA-seq data (Supplementary Fig. 5 and Supplementary Data 2). Using Hidden Markov-based hmmsearch with TPS C-terminus protein domain (Pfam:PF03936.16), we identified a total of 26 full-length TPSs, 11 AcaTPSs from the transcriptomes of *A. carmichaelii* and 15 AcoTPSs from those of *A. coreanum* (Supplementary Data 3). Phylogenetic analysis of the TPSs with functionally characterized mono-, sesqui-, di-, and sester-TPSs separated the TPSs into four major clades (Fig. 2a, Supplementary Fig. 6, and Supplementary Data 4), corresponding to plant TPS-b, c, e/f, and g subfamilies³⁶. DiTPS activities have been observed from the TPS-c, e/f, and g subfamilies, but not often in TPS-b one³⁷, indicative of diTPS activities for those TPS candidates in TPS-c, e/f, and g subfamilies. We successfully cloned nine AcaTPSs and 14 AcoTPSs from the cDNA libraries of *A. carmichaelii* and *A. coreanum* tissues, respectively, and constructed the pEAQ-HT expression vectors³⁸ with the cloned fragments for functional analysis using transient expression in *Nicotiana benthamiana*³⁹.

To facilitate validation of the functions of candidate TPSs in diterpene biosynthesis, we also cloned genes encoding the MEP pathway upstream rate-limiting enzyme 1-deoxy-D-xylose-5-phosphate synthase from *Coleus forskohlii* (CfDXS)⁴⁰, the geranylgeranyl diphosphate synthase from *Arabidopsis thaliana* (AtGGPPS)⁴¹, and the previously characterized *ent*-copalyl diphosphate synthase (AcaCPS) from *A. carmichaelii*⁴² into pEAQ-HT vectors to boost the metabolic flux towards diterpene biosynthesis. By coexpressing the candidate AcaTPSs or AcoTPSs with CfDXS + AtGGPPS or CfDXS + AtGGPPS + AcaCPS in *N. benthamiana* leaves followed by metabolite analysis of the leaf extracts, we identified two *ent*-CPP synthases (CPSS), i.e., AcoTPS4 and AcoTPS5, as indicated by the detection of a new peak corresponding to copalol (**1**) by gas chromatography-mass spectrometry (GC-MS) (Fig. 2b), likely resulted from the hydrolysis of the enzymatic product *ent*-CPP by endogenous phosphatases in *N. benthamiana*. Further coexpressing CfDXS, AtGGPPS, and AcaCPS with other candidate TPSs identified seven functional class I diTPSs capable of ionizing *ent*-CPP to yield new peaks **2–4** (Supplementary Figs. 7 and 8), amongst which five TPSs (AcoTPS1/11, and AcaTPS3/4/5) could yield **2**, AcaTPS8 generated **3**, and AcoTPS3 gave **4**. Compounds **2–3** were identified as *ent*-atiserene (**2**) and *ent*-kaurene (**3**) based on NIST library search (Supplementary Fig. 8) and comparing their retention times and mass spectra with those detected by coexpressing CfDXS, AtGGPPS, AcaCPS with characterized *ent*-atiserene synthase from *Isodon rubescens* (IrKSL4)⁴³ and *ent*-kaurene synthase (KS) from *A. thaliana* (AtKS)⁴⁴. Compound **4** was tentatively identified as 16 α -hydroxy-*ent*-kaurene based on NIST library search, and its structure was later confirmed by NMR-based structure elucidation following isolation. Amongst the five TPSs with atiserene synthase (AS) activities, AcoTPS1 and AcaTPS4, also named AcoAS1 and AcaAS1, respectively, yielded significantly higher levels of **2** than the other three that



exhibited AS activity. The protein sequences of AcaAS1 and AcaTPS8 (namely AcaKS) and their activities in synthesizing **2** and **3** were consistent with those of the previously reported atiserene and kaurene synthases (i.e., AckSL2-1 and AckSL3-1) from *A. carmichaelii*³⁴. DiTPSs with activity in synthesizing 16 α -hydroxy-ent-kaurene were previously

found in different plant species^{45,46} and bacteria⁴⁷ but not from *Aconitum* spp. The discovery of AcoTPS3 capable of generating **4** suggests that **4** may be an alternative precursor scaffold for downstream oxidation and modifications en route to the biosynthesis of diterpenoids or diterpene alkaloids in *A. coreanum*.

Fig. 2 | Identification and coexpression analysis of diterpene synthases from *A. coreanum* and *A. carmichaelii*. **a** Phylogenetic tree of 26 candidate TPSs identified from *Aconitum* spp. transcriptomes and other characterized TPSs of different subclasses from different plant species (details given in Supplementary Data 2 and 3). The different clades of TPS subfamilies are shaded in different colors. Functional TPS are highlighted in different colors: purple, copalyl diphosphate synthase (CPS) activity; green, atiserene synthase (AS) activity; orange, kaurene synthase (KS) activity; blue, 16 α -hydroxy-*ent*-kaurene synthase activity. Note that AcoTPS11, AcaTPS3, AcaTPS5 showing both AS and KS activity in **b** are colored in green. Details of the trees are zoomed in Supplementary Fig. 1. **b** Comparative GC-MS total ion chromatograms (TICs) of extracts of *N. benthamiana* leaves expressing upstream enzymes CfDXS + AtGGPPS and candidate diTPSs. Product peaks are filled with different colors: *ent*-copalol (purple), 16 α -hydroxy-*ent*-kaurene (blue), *ent*-atiserene (green) and *ent*-kaurene (yellow). *AcaCPS*, *ent*-copalyl diphosphate

synthase from *A. carmichaelii*; *IrKSL4*, *ent*-atiserene synthase from *Isodon rubescens*; *CfDXS*, 1-deoxy-D-xylulose-5-phosphate synthase from *C. forskohlii*; *AtGGPPS*, geranylgeranyl pyrophosphate synthase from *A. thaliana*. Displayed are representative traces of multiple ($n > 3$) replicates with similar results. The floating traces in viii, ix and x are zoomed in view of the designated time range of the extracted ion (272) chromatograms corresponding to the TICs. **c** Chemical structures of *ent*-copalol, *ent*-atiserene, *ent*-kaurene and *ent*-kauran-16 α -ol. **d** SOM code plot of the expression profiles of genes from the eight *A. carmichaelii* tissues (**d**) and six *A. coreanum* ones (**e**). Each hexagon node represents transcripts of *A. carmichaelii* (**d**) and *A. coreanum* (**e**) with similar expression profiles, with the average expression profile plotted within each node. The nodes containing the functional AcaTPS and AcoTPS genes are circled in blue with the corresponding node numbers shown. Clusters containing nodes with the same color shading on the clustering map indicate they share relatively similar expression profiles.

Multifunctional P450s from *Aconitum* spp. enable biosynthesis of cryptic atiserenoids

Having identified the functional TPSs that generate LRD scaffolds and show tissue-specific expression patterns (Supplementary Fig. 9), we next performed coexpression analysis on the RNA-seq of *Aconitum* spp. tissues using self-organising map (SOM)⁴⁸ and Pearson coefficient analysis, with the goal of identifying functional P450s that can act on the LRD scaffolds generated by the AcaTPSs and AcoTPSs. SOM analysis of the RNA-seq transcripts resulted in 6 clusters and 36 nodes for *A. carmichaelii* and 6 clusters and 16 nodes for *A. coreanum* (Fig. 2d,e). The functional AcaKS and AcaASs are mainly located in three separate nodes (1 and 6) within the same cluster 1, suggesting that their expression patterns are similar in *A. carmichaelii*, with high expression observed in fibrous roots (Fig. 2d and Supplementary Data 5). In comparison, the P450s coexpressed with functional AcoTPSs are located separately into three nodes (1, 4, 7) within two clusters. AcoTPS11 (AcoAS2) and AcoTPS3 that generate **4** are in node 4, whilst AcoTPS1 (AcoAS1) and AcoTPS5 (AcoCPS2) in node 1 with transcripts showing higher expression in roots (Fig. 2e and Supplementary Data 5).

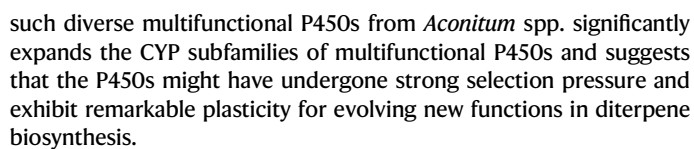
We next carried out Pearson coefficient analysis to identify the P450s that show high correlations with functional TPSs in their corresponding nodes from the SOM analysis. The coexpressed P450s were then ranked based on their Pearson coefficient values, and the highly ranked ones (with a cutoff of 0.9) were cloned for functional validation using transient co-expression in *N. benthamiana*^{49,50}. We cloned a total of 58 coexpressed genes annotated as P450s from six SOM nodes and screened their activities on *ent*-kaurene and *ent*-atiserene by coexpressing them individually with CfDXS+AtGGPPS +AcaCPS+AtKS/IrKSL1 (Supplementary Data 6 and 7). A total of 12 functional P450s were identified to be able to consume **2** or **3** and generate new products as evident from the presence of new peaks detected from liquid chromatography-mass spectrometry (LC-MS) and GC-MS analyses showing characteristic diterpene-related mass fragments and spectra (Fig. 3a and Supplementary Figs. 10–17). Amongst the 12 functional P450s, AcoCYP82C1 and AcaCYP82C1 could only oxidize **3** rather than **2**, whereas the other 10 could act on both scaffolds. Interestingly, all the functional P450s acting on **2** or **3** could generate multiple products (Fig. 3a, b and Supplementary Figs. 10 and 12), suggesting that they might be multifunctional. Further coexpressing other candidate P450s combinatorially with the 12 functional ones identified two additional CYP716B and CYP93A subfamilies P450s (AcoCYP716B1 and AcaCYP93A1) that could act with AcoCYP725A1 to convert **2** to mostly **A13**, or with AcoCYP93A1 to form **A15**, together with a few other tandem catalyses enabled by two functional P450s (Fig. 3a, b and Supplementary Fig. 11).

The 14 functional P450s belong to seven diverse CYP subfamilies spanning across the CYP71, 85 and 72 clans, including CYP93A, CYP71A, CYP701A, CYP725A, CYP82C, CYP716B and CYP72A (Fig. 3c and Supplementary Fig. 18). Most functional P450s from *A. carmichaelii* and *A. coreanum* appear to be homologous pairs except AcoCYP725A1 and

AcoCYP716B1, which are derived from *A. coreanum* only and clustered with the CYP720B subfamily enzymes that are known to oxidize the C18 methyl group of tricyclic LRD scaffolds such as pimarane⁵¹ in the CYP85 clan of the phylogenetic tree (Fig. 3c and Supplementary Fig. 18). AcoCYP72A1 and AcaCYP72A1 belong to the CYP72 clan, whilst other functional P450s clustered together with characterized P450s capable of oxidizing different LRD scaffolds at varied sites in three major subclades of the CYP71 clan (Fig. 3c and Supplementary Fig. 18), suggesting that the functional P450s identified here are indeed highly diverged.

To understand the exact functions of these diverged P450s in driving the structural diversification of **2**, we scaled up the transient coexpression of different functional P450s individually or combinatorially with genes for synthesizing **2** or **3** in *N. benthamiana* via large scale agro-infiltration, followed by extraction and purification³⁹. We successfully isolated 18 oxidized products of **2** and **3** for NMR-based structural elucidation. After comprehensive spectroscopic analyses of the isolated compounds, the structures of atiserenoids **A4–18** and kaurenoids **K9–11** were elucidated as depicted in Fig. 3d (Supplementary Figs. 35–78 and Supplementary Data 8 and 9). The structures of **K1–3** and **A1–3** were tentatively identified as the different forms of C19 oxidized products and **K6** as *ent*-kauran-15-ol by comparing their GC-MS mass spectra with those reported in literature⁵². Of all the compounds isolated, 13 atiserenoids (**A4–6**, **A8–9** and **A11–18**) and one kaurenoid (**K9**) are new compounds reported here for the first time with their structures established based on NMR elucidation with key 2D correlations shown in Supplementary Figs. 35–43, 45–50, 52–75 and 76–78. The NMR data of the known compounds (**A7**, **A10**, **K10–11**) were also consistent with those of the previously reported^{53–56}. The structures of the oxidized products of **2** and **3** indicate that the 14 functional P450s from *Aconitum* spp. discovered here could enable 12 regio- and stereo-selective oxidative modifications at a total of seven different sites (C3, C15, C16, C17, C18, C19 and C20) of **2** and **3**. This accounts for ~39% of all accessible oxidation forms at the different carbon positions of **2** based on the structures of the atiserenoids that have been isolated so far^{29,57,58} (Supplementary Fig. 19).

More importantly, six functional P450s in the CYP93A, CYP71A, CYP725A, and CYP72A subfamilies are multifunctional, capable of oxidizing **2** at more than one site. Notably, AcoCYP93A1 and AcoCYP725A1 could oxidize **2** at three and four distinct sites to generate **A8–10** and **A7**, **13**, **15**, respectively, when coexpressed with genes for synthesizing **2** in *N. benthamiana*. Multifunctional P450s that oxidize LRD scaffolds at different sites are relatively rare, with the majority reported in the CYP76 family, including the CYP76AH subfamily involved in the oxidation of miltiradiene at C7, C11, C12 positions in tanshinone biosynthesis, and the CYP76M ones from rice that oxidize LRD skeletons at C6, C7, C9, C11 and C20 positions in the biosynthesis of plant antitoxins¹⁹. Additionally, a few other subfamilies, such as CYP71Z, CYP71Z7 for the biosynthesis of phytocassanes⁵⁹ are also multifunctional (Fig. 1 and Supplementary Data 1). The discovery of



We next performed protein sequence and structure analyses, followed by site-directed mutagenesis experiments, with a view to probing how the multifunctional P450s may have diverged to give rise to the varied

Fig. 3 | Divergent multifunctional P450s from *Aconitum* spp. catalyzing LRD oxidation. Heatmaps showing the product profiles of functional P450s and their combinations acting on *ent*-atiserene (**a**) or *ent*-kaurene (**b**). The intensity was based on the relative abundances of peak areas of extracted ions for different compounds detected from LC-MS analysis (Supplementary Figs. 10–12): *m/z* 271.2412 for **A2**, **A4**, **A7–A10**; *m/z* 287.2369 for **A1**, **A11** – **A15**; *m/z* 301.2160 for **A6**, **A16**, and **A17**; *m/z* 303.2316 for **A3** and **A5**; *m/z* 317.2475 for **A18**. The oxidation site was indicated on the right. **c** The phylogenetic tree of functional P450s from *A. coreanum* (purple) and *A. carmichaelii* (orange) and other previously characterized plant P450s with their substrate scaffolds indicated on the right) reveals significant functional divergence. Bootstrap values of the tree were generated based on 1000

sequence resamplings. The corresponding P450 subfamily numbers were labeled at the node of the different clades. **d** Schematic depiction of the structures of oxidation products generated from combinatorial biosynthesis of the functional CYPs and TPSs identified in this study. Color shadings indicate products generated by the designated CYPs or CYP combinations. The dotted circle indicates that **K6** can be generated by CYP82C as the dominant product. Blue or yellow box on the right indicates the **A15** or **A13** can be generated by the combinations of CYP725A + CYP93A or CYP725A + CYP716B at higher yield, respectively. Compounds with novel structures are labeled in red and their NMR spectra were provided in Supplementary Figs. 35–86. *, structure tentatively identified based on EI-MS spectrum and relative retention time.

functions. Based on the structurally characterized *ent*-atisane diterpenoids documented²⁹, over 60% of the reported structures have oxidation at C3 of the *ent*-atiserene scaffold, suggesting that the C3 position might be a relatively early oxidation site during the biosynthesis of *ent*-atisane diterpenoids in nature. The five multifunctional P450s in the CYP71A, CYP72A, and CYP93A subfamilies identified from *Aconitum* spp. could all oxidize **2** at the C3 position but yielded different product profiles, which allows us to probe how their sequence and structural variation might lead to functional divergence. Given that CYP93A1 yields the most diverse product profile with oxidation occurring at three different sites (C3, C6, and C7) and potentially represents an evolutionary mid-point that can be more easily altered functionally towards yielding a specific product, we chose CYP93A1 for the initial mutagenesis study.

We first generated the predicted structure of AcoCYP93A1 using AlphaFold2 and docked the substrate **2** into the active site using Autodock tool to map the key amino acid sites interacting with **2**. A total of 16 key amino acid sites spanning across the six substrate recognition sites (SRSs) together with the heme binding site within 5 Å of **2** were identified (Fig. 4a and Supplementary Figs. 20 and 21). Sequence alignment of the five multifunctional P450s (i.e., AcoCYP71A1, AcaCYP71A1, AcoCYP72A1, AcaCYP72A1 and AcoCYP93A1) with characterized C3 oxidases from rice (OsCYP71Z6/7) and *Calophyllum plumiforme* (CpCYP96A1)²⁰ further reveal 13 key sequence variation sites that could potentially contribute to their functional divergence (Supplementary Fig. 22). We next selected seven key variation sites that could best interact with the ligand for further site-directed mutagenesis in AcoCYP93A1, i.e., W145, I241, V334, A335, I404, I515, and V516 (Supplementary Fig. 23).

The seven key amino acids are positioned around the ligand **2**, interacting closely with the different carbon positions of the skeleton, such that the substrate is in appropriate conformation for the carbon positions in close vicinity to the heme group to be oxidized (Fig. 4a,b and Supplementary Fig. 23). Changing the amino acid residues in these seven key sites may result in fine-tuning of the relative conformation of ligand **2** in the active site, and sequentially the relative positions of the carbon skeletons to the heme group. We next mutated the residues at these seven sites individually to different amino acid forms in light of their sizes, polarities, and electric charges (Fig. 4 and Supplementary Fig. 24).

We obtained a total of 26 single-site mutants of AcoCYP93A1 and evaluated their functions in oxidizing **2** by coexpressing the pEAQ-HT constructs carrying the AcoCYP93A1 mutants with gene combinations for the biosynthesis of **2** in *N. benthamiana* leaves, followed by extraction and metabolite analysis using GC-MS. Intriguingly, our results indicated that the seven key amino acid sites indeed affected the catalytic activity and product profile of AcoCYP93A1 to a great extent (Fig. 4c,d and Supplementary Fig. 25). Notably, mutants AcoCYP93A1^{I241S} and AcoCYP93A1^{I404L} exhibited enhanced catalytic activity and specificity, with yields of C3 hydroxylated product **A10** increased by around 2.4 and 3.6 folds, respectively (Supplementary Fig. 25a) and its relative proportion in the enzymatic products (**A10/A8–10**) increased significantly (from 68% to 94% and 92%) compared with the wildtype AcoCYP93A1 (Fig. 4c,d and Supplementary

Fig. 25b). Other mutants at I404 also showed altered product profiles with significant decrease of **A10** (Fig. 4c,d and Supplementary Fig. 25a). In particular, mutating I404 to smaller amino acids such as alanine (AcoCYP93A1^{I404A}) or glycine (AcoCYP93A1^{I404G}) led to major product profile change, yielding substantially less **A10** with some other new products, including **M1–5** with mass fragment *m/z* = 288 as detected by GC-MS (Fig. 4c,d and Supplementary Figs. 24a and 26). We successfully isolated sufficient amount of **M1** for NMR analysis following large-scale transient expression of AcoCYP93A1^{I404A} with **2** biosynthesis genes in *N. benthamiana* leaves and established the structure of **M1** as *ent*-atiserene-14-ol based on key HMBC correlation between the hydroxylated C14 hydrogen and C7 and C13 (Fig. 4c, Supplementary Figs. 81–83 and Supplementary Data 9). Our results suggest that I404 is critical in controlling the conformations of **2** for C3 oxidation, thus constituting a notable key site for tuning the P450 catalytic activity and outcomes for functional divergence. Mutations in other key amino acid sites likely also affect the positioning of substrate **2** in the active site, and subsequently the carbon positions interacting with the heme group for hydrogen abstraction, leading to further product profile change and functional divergence (Fig. 4 and Supplementary Fig. 27).

Biosynthesis and identification of *Aconitum* diterpenoids including tripterifordin and guan-fu diterpenoid A implicate convergent evolution

The functional TPSs and P450s from *Aconitum* spp. have opened avenues to access diverse diterpenoids through combinatorial biosynthesis. With the *ent*-kaurene and *ent*-atiserene synthases and the 14 functional P450s in hand, we have obtained and characterized a total of 25 diterpenoids, amongst which compounds **A16** and **A17** could be detected from *A. carmichaelii* roots with **A16** also detected from *A. coreanum* tuber, stems and leaves (Fig. 3 and Supplementary Fig. 28). Besides **2** and **3** which can serve as scaffolds for oxidative modifications, compound **4** with a C16 hydroxy moiety readily installed may also be an alternative scaffold for further decoration. Diterpenoids such as guan fu diterpenoid A⁶⁰ and tripterifordin that exhibit cell culture-based anti-HIV activity and were previously isolated from *Tripterium wilfordii*⁶ both carry the 16 α hydroxy moiety, with additional oxidation at C20 or C19 and C20, hence are oxidative products derived from **4**.

Next, we examined if it's possible to synthesize these compounds using AcoTPS3 which generates **4** together with other C19 and C20 oxidases. We first coexpressed the genes for **4** synthesis with C20 oxidase AcaCYP71A1 in *N. benthamiana*. We could detect the formation of two new peaks with mass fragments at *m/z* = 306 (**T1**) and *m/z* = 304 (**T2**), corresponding to the installation of a carbonyl and hydroxyl group, respectively. Given that AcaCYP71A1 has been validated to be a C20 oxidase, it is highly likely that **T2** is a C20 aldehyde and **T1** is a C20 alcohol as depicted in Fig. 5, thus **T1** was tentatively identified as guan-fu diterpenoid A and its structure later confirmed via NMR-based structural elucidation following isolation and comparison with previously reported NMR data⁶⁰ (Supplementary Fig. 84 and Supplementary Data 9). Similarly, coexpressing the **4** biosynthesis genes with the C19 oxidase-coding genes AcoCYP701A2 or

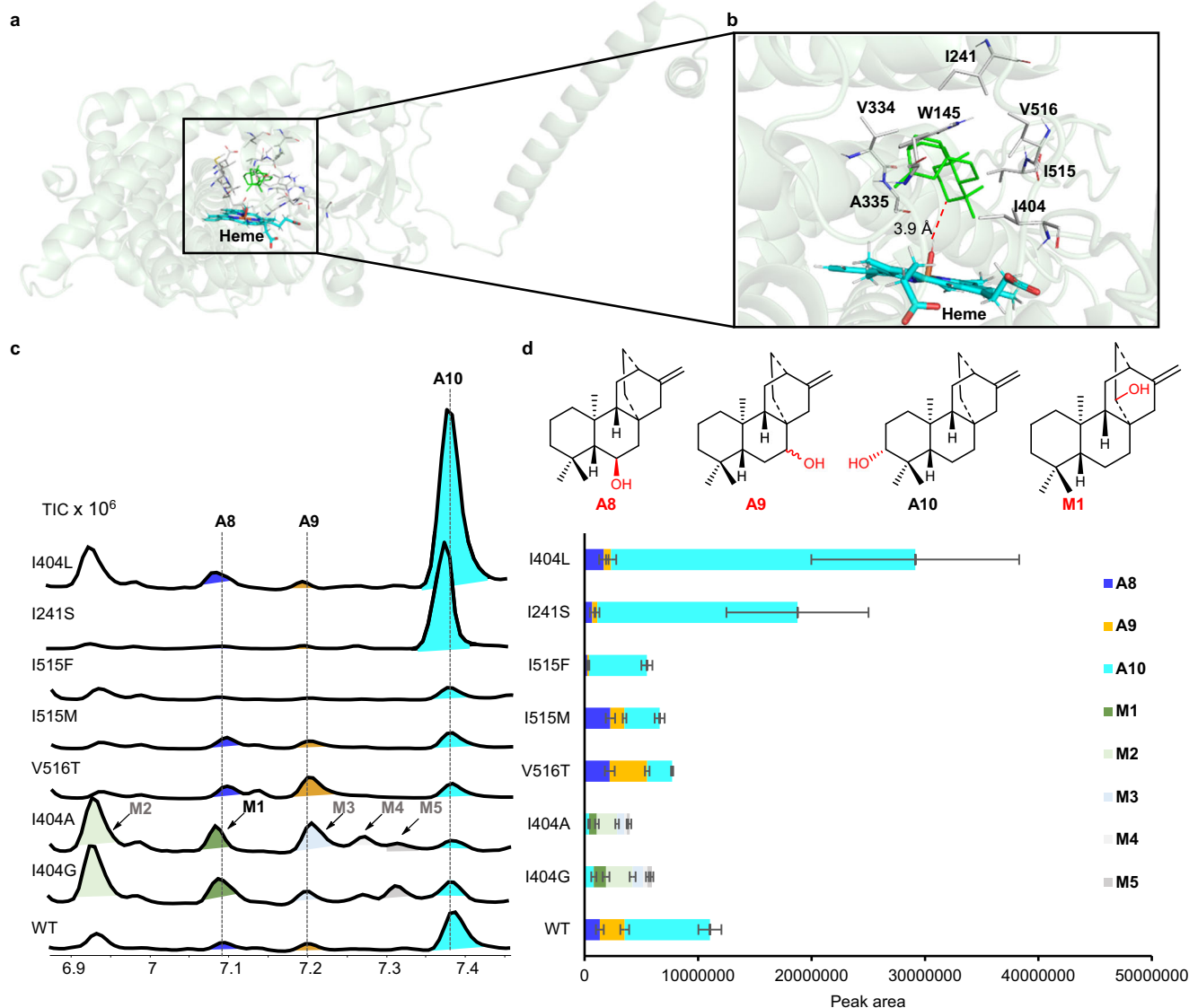


Fig. 4 | Identification of key amino acid sites contributing to *Aconitum* P450 functional divergence. **a** Predicted AcoCYP93A1 structure with *ent*-atiserene docked in the active site. **b** Zoom-in view of interactions among *ent*-atiserene, key residues and heme in the catalytic pocket of AcoCYP93A1. **c** Comparative GC-MS TICs of the oxidation products of *ent*-atiserene by AcoCYP93A1 and its mutants when coexpressed with CfDXS+AtGGPPS+AcaCPS. Displayed are representative

traces of multiple ($n > 3$) replicates with similar results. **d** Relative yields and compositions of the different oxidation products of the AcoCYP93A1 and its mutants detected (**c**) based on their TIC abundance. Data are shown as the mean $\pm$ s.d., $n = 3$ biological replicates. Compounds with novel structures are labeled in red.

AcaCYP701A2-1 could also generate new peaks (**T3** and **T4**) with mass fragments $m/z = 306$ (**T3**) and $m/z = 304$ (**T4**), corresponding to the C19 alcohol (**T3**) and aldehyde (**T4**) as shown in Fig. 5. Further coexpressing genes for **4** biosynthesis (*CfDXS*, *AtGGPPS*, *AcaCPS*, *AcoTPS3*), C20 (*AcaCYP71A1*) and C19 (*AcoCYP701A2* or *AcaCYP701A2-1*) oxidation in *N. benthamiana* yielded new peaks (**T5–9**). **T7** was identified as tripterifordin by comparing its GC-MS and LC-MS retention times and mass spectra with those of the commercially purchased authentic standard (Supplementary Fig. 29). **T8** and **T9** were determined as 16 α -hydroxy-*ent*-kaurane-19,20-dial (**T8**) and 16 α -hydroxy-*ent*-kaurane-20,19-hemiacetal (**T9**), respectively, based on NMR elucidation and comparison to the previously reported data^{61,62} (Supplementary Figs. 85 and 86 and Supplementary Data 9). **T5** and **T6** were tentatively identified as the other 16 α -hydroxy-*ent*-kaurane-19,20-hemiacetal (**T5**) and C19 aldehyde and C20 alcohol (**T6**), respectively, as inferred from their mass fragments ($m/z = 320$) and relative retention times (Fig. 5 and Supplementary Figs. 29 and 30). It is worth noting that **T1**, **T2**, and

T7 could be detected from the tissues of these two *Aconitum* plants (Supplementary Fig. 31). Intriguingly, tripterifordin (**T7**) isolated from *T. wilfordii* in the Celastraceae could also be detected from the tuber of *A. coreanum* [–1.2 mg/g, dry weight (dw)] in the Ranunculaceae (Supplementary Fig. 31), suggesting that there is selection pressure driving the convergent evolution of tripterifordin biosynthesis in *T. wilfordii* and *A. coreanum*.

Aconitum diterpenoids can be allelopathic

Plant growth is mediated by conserved diterpene hormone gibberellins (GAs), which are also LRDs produced universally in plants⁶³. Many LRDs^{9,64,65}, e.g., momilactones from rice^{66,67}, have been found to serve as allelopathic compounds capable of modulating plant growth, potentially via interfering with GA signaling. The substantial structure similarity between *ent*-atiserene and the GA precursor *ent*-kaurene led us to wonder if they may have a role in regulating plant growth. We therefore tested a few *ent*-atiserenoids (**A10**, **A11**, **A13**, **A15**) that were

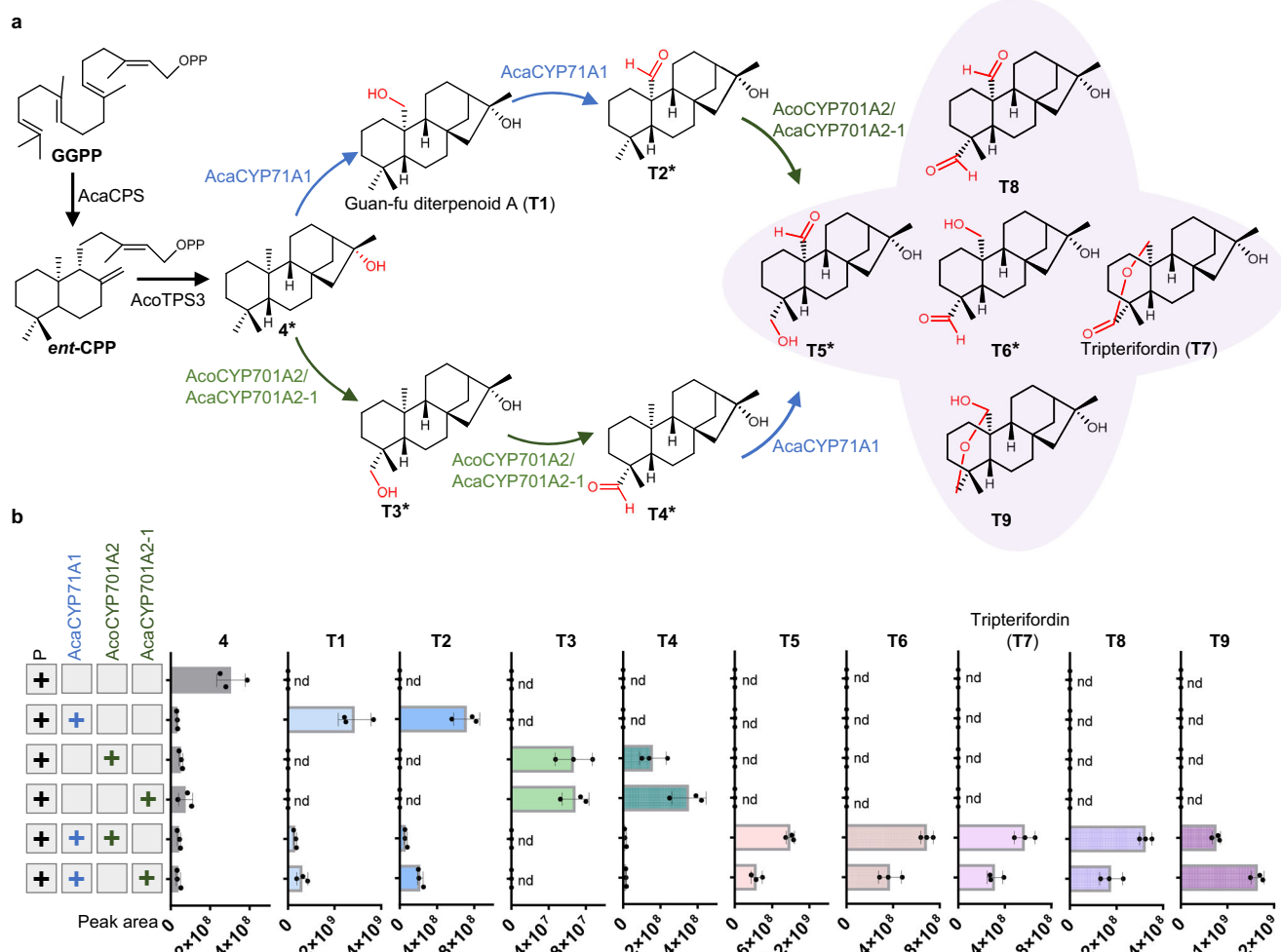


Fig. 5 | Biosynthesis of tripterifordin and guan-fu diterpenoid A. **a** Biosynthesis of tripterifordin with the multifunctional P450s via two possible routes. **b** Relative abundance of enzymatic products detected from extracts of *N. benthamiana* leaves expressing different enzyme combinations. P, CfdXS + AtGGPPS + AcaCPS + AcoTPS3. The abundances are peak areas of extracted ions for the different

compounds: **4** (273.2577), **T1/3** (289.2526), **T2/4** (287.2369), **T5/6/9** (303.2319), **T7/8** (319.2267). Data are shown as the mean $\pm$ s.d., $n = 3$ biological replicates; n.d., not detected. *, structure tentatively identified based on EI-MS spectrum and relative retention time.

isolated in good quantity from *N. benthamiana* leaves together with tripterifordin (**T7**) on the root growth regulatory activity of *A. thaliana* (Fig. 6a,b). The herbicide paclobutrazol (PAC) was used as a positive control, and growth hormone gibberellin GA3 as a negative control. We found that the five compounds we tested could all inhibit *Arabidopsis* root growth to different extents compared with the mock treatment, albeit the inhibition is not as potent as herbicide PAC (Fig. 6c, d and Supplementary Fig. 32). Since *Aconitum* spp. harbor the latent capacity to produce structurally diverse LRDs and LRD-derived alkaloids like mesaconitine (~7.3 mg/g, dw) and hyaconitine (~12.8 mg/g, dw) produced at *A. carmichaelii* roots could be secreted into root exudates to exert allelopathic effects (Supplementary Fig. 33), it is possible that these plants may outcompete neighboring plants via producing allelopathic LRDs.

Discussion

Cytochrome P450 monooxygenases are the largest family of enzymes participating in oxidative modifications of different classes of natural products and shape plant specialized metabolism for plant adaptation⁶⁸, as exemplified by their roles in driving the structural diversification of diterpenes¹⁷. Plants of different lineages have evolved lineage-specific P450 subfamilies to decorate the specialized metabolites present. The diterpene synthases that generate LRD scaffolds

1-4 from our systematic screening suggest that diterpenoids and diterpene alkaloids present in *A. carmichaelii* and *A. coreanum* are likely derived from **2-4**. The functional P450s characterized in our work are highly divergent with some subfamilies appearing to be lineage-specific for LRDs oxidation in *Aconitum* spp., thus expanding the CYP subfamilies in diterpene biosynthesis. For instance, the CYP93A subfamily was previously reported to be involved in isoflavone biosynthesis⁶⁹. The discovery of the functional AcoCYP93A1 and AcaCYP93A1 was the first time that CYP93A subfamily P450 members were shown to be capable of oxidizing LRD scaffolds. The CYP82C subfamily enzymes AcoCYP82C1 and AcaCYP82C1 are similar. Previously, a CYP82C P450 was found to catalyze oxidation of coumarin 8-methoxypsoralen⁷⁰, although other CYP82 subfamilies, such as CYP82G1 from *A. thaliana* could act on linear terpenoids⁷¹. The CYP93A and CYP82C subfamily P450s identified here are in the same CYP71 clan as the CYP701A subfamily kaurene oxidases (KOs) that catalyze multistep oxidation of *ent*-kaurene in gibberellin biosynthesis (Fig. 3b and Supplementary Figs. 10a,b and 12a, b), suggesting that these two subfamilies share close evolutionary relationship with the KOs. In contrast to the CYP71 clan, which encompasses diverse CYP subfamilies in diterpene biosynthesis, the CYP85 clan only has five CYP subfamilies found to be involved in LRDs biosynthesis so far, including the CYP725A subfamily known to oxidize the taxane scaffold¹⁷. Our

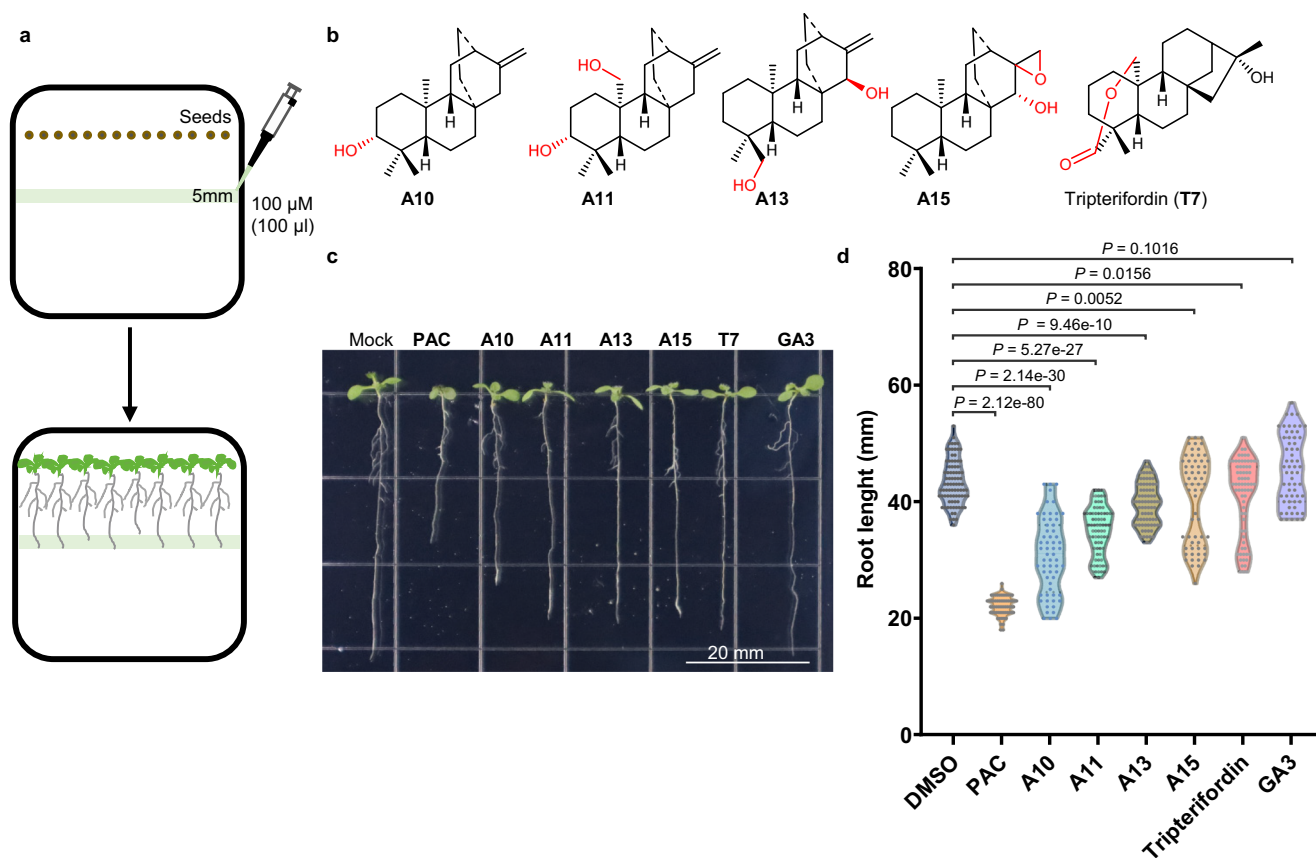


Fig. 6 | Cryptic atiserenoids and tripterifordin exhibit allelopathic activity.

a Schematic diagram depicting the root inhibition assay. Plants grown on each plate were subjected to only one compound treatment. **b** Compounds used for the assay. **c** Phenotypes of representative 10-day-old *A. thaliana* seedlings grown on MS plates treated with the designated compounds at the indicated concentrations. Note that the representative seedlings were taken from the plates of different treatments in

Supplementary Fig. 32 and placed in one plate for photographing and visualization. **d** Violin plots showing the overall root lengths of *A. thaliana* seedlings treated with different compounds. Each violin plot contains at least 63 points corresponding to the root length of individual seedlings. Statistics were performed using one-way ANOVA analysis to compare significant differences between control (DMSO) and the different treatments. PAC paclobutrazol. GA3 gibberellic acid 3.

work showed that CYP725A and CYP716 subfamily enzymes could both act on **2** and **3** (Supplementary Figs. 11a and 13a), indicative of more diverse substrates from these two subfamily enzymes.

The divergent CYP subfamilies are quite promiscuous as they could act on **2–4**, likely owing to the relatively loose active site pockets. The oxidation sites of substrates rely a lot on how the key amino acids collectively position the substrates towards the heme group. Mutation of key amino acids could result in a change of the carbon position pointing to the heme group and altered product profiles, leading to neofunctionalization and further diversifying the diterpenoid structures. Intriguingly, the promiscuity of multifunctional P450s identified here seemed to be able to be rectified to a great extent when expressed with other functional P450s, as evident from the formation of one dominant enzymatic product when coexpressing AcoCYP72A1 + AcaCYP71A1 or AcoCYP725A1 + AcaCYP93A1/AcaCYP716B1/AcoCYP72A1 (Fig. 3a and Supplementary Fig. 11a). Although the reason why coexpressing two multifunctional P450s could increase the product specificity is not yet clear, this appears to be an effective strategy to retain both substrate promiscuity and product specificity of multifunctional P450s at the same time without compromising one another.

The promiscuity of P450s also bespeaks significant plasticity for functional reshaping. The C20 oxidase of *ent*-kaurene (AcaCYP71A1/AcoCYP71A1) discovered in this work belongs to the CYP71A subfamily (Supplementary Fig. 18), whereas that from *T. wilfordii* (TwCYP71BE271)⁷² belongs to the CYP71BE subfamily. AcaCYP71A1 and TwCYP71BE271 could both catalyze C20 oxidation of **4** for the biosynthesis of

tripterifordin, yet these two subfamilies appear to be independently evolved from an ancestral null-C20 oxidase of **4**, as the two species were evolutionarily very distant, with *T. wilfordii* emerging ~13 MYA later (Supplementary Fig. 34) and tripterifordin was only found in *Tripterium* of the Celastraceae and *Ligularia* of the Senecioneae⁷³ before. This implies that the biosynthesis of tripterifordin in *A. coreanum* and *T. wilfordii* is likely a convergent evolution case⁷⁴, although we couldn't rule out the possibility that the C20 oxidases in the two species evolved from a functional ancestral C20 oxidase of **4** in the CYP71 family.

Furthermore, the promiscuity of individual P450s in terms of substrate and oxidation sites provides the basis for further activity tuning towards designer oxidation of specific carbon positions of LRD scaffolds, as manifested by the synthesis of a series of kaurenoids and atiserenoids via combinatorial biosynthesis and mutagenesis experiments. Considering *ent*-kaurene and *ent*-atiserene are precursors for numerous bioactive diterpenoids and diterpene alkaloids^{11,29,32}, the discovery of the TPSs and multifunctional P450s in this work will expand current stock in the enzymatic toolkit for regio- and stereo-selective oxidation of LRD scaffolds. This allows for the synthesis of LRDs as demonstrated by the novel kaurenoids and atiserenoids synthesized in our work. Collectively, the newly synthesized LRDs here can be used for building in-house natural product libraries for drug screening. The functional enzymes discovered here will pave the way for elucidation of downstream biosynthetic pathways of diterpene alkaloids in *Aconitum* spp. and serve as biosynthetic parts for the green and sustainable biosynthesis of bioactive diterpenoids using synthetic biology in the future.

Methods

Plant materials and RNA-seq library preparation and sequencing

N. benthamiana plants were cultivated in a growth chamber at 25 °C with a 16-hour light and 8-hour dark photoperiod. *A. carmichaelii* was sourced from Jiangyou, Sichuan, in June 2020, and *A. coreanum* was collected from Fushun, Liaoning, in May 2020 and cultivated in a growth room. Different plant tissues, including tuber, roots, stems, leaves, flowers, and buds from *A. coreanum* and taproots, lateral roots/fuzi, fibrous roots, old stems, young stems, flowers, and siliques from *A. carmichaelii* were harvested for RNA extraction and metabolite analysis. Total RNA was extracted from the different tissues using an RTM RNA Library Prep Kit for Illumina platform (NEB, USA) NA extraction kit (Huayueyang Biotech, Beijing, China) according to the manufacturer's instructions. The integrity and concentration of total RNA were assessed using 1.0% agarose gel electrophoresis and a Fragment Analyzer or Agilent 2100 Bioanalyzer (Agilent, USA). High-quality total RNA was used for cDNA library preparation with NEB-Next® Ultra following the manufacturer's instructions. The cDNA libraries of all samples were verified by Qubit quantification (Thermo Fisher Scientific, USA), agarose gel electrophoresis, and Agilent High Sensitivity DNA analysis (Agilent, USA), and then subjected to RNA sequencing on the Illumina HiSeq 2500 platform (Illumina, USA) with 150 bp paired-end reads.

De novo assembly of the transcriptomes of *Aconitum* spp

A total of 24 *A. carmichaelii* and 18 *A. coreanum* RNA-seq datasets were generated from plant tissues, each with three biological replicates. The raw data was filtered with Fastp (v0.20.0)⁷⁵ with: (i) Q30 quality filtering; (ii) TruSeq adapter trimming; and (iii) species-specific trimming (*A. carmichaelii*: 15 bp/10 bp 5'/3'-trim; *A. coreanum*: 15 bp/5 bp 5'/3'-trim). This retained ~95% of reads (125–130 bp average length, 4–6 Gb total bases/sample, Supplementary Data 2). For de novo transcriptome assembly, we combined quality-filtered reads from all samples per species and performed de novo assembly using Trinity (v2.8.5)⁷⁶. CD-HIT (v4.8.1)⁷⁷ was employed to remove redundant sequences, resulting in 831,643 assembled transcripts for *A. carmichaelii* and 548,987 for *A. coreanum* as reference transcriptome (contig Ex90N50 values reaches 1,844 bp for *A. carmichaelii* and 1,777 bp for *A. coreanum*). The clean data of each sample were mapped to the reference transcriptomes by Bowtie2⁷⁸ and the expression level of genes was calculated by RSEM⁷⁹ via the Trinity's integrated workflow. The candidate coding regions in assembled transcript sequences were identified using Transdecoder (v5.5.0)⁸⁰, with the corresponding protein sequences identified based on open reading frames in the assembled transcripts. Transcriptome completeness was evaluated using BUSCO (v5.8.0)⁸¹ with the viridiplantae_odb10 dataset. Both assemblies showed exceptional completeness (*A. carmichaelii*: 97.4%; *A. coreanum*: 99.5% complete BUSCOs), with minimal fragmented (*A. carmichaelii*: 2.4%; *A. coreanum*: 0.5%) (Supplementary Data 2). The functional annotation of the de novo assembled transcriptomes was carried out by Trinotate v3.1.1⁸², which integrated homology searches (BLAST+/SwissProt), protein domain identification (HMMER/PFAM), and functional classification through multiple databases (eggNOG/GO/KEGG).

Identification of *Aconitum* terpene synthases and their co-expressed P450s

To identify the *Aconitum* Terpene synthases (TPS), we mined the *Aconitum* reference transcriptomes for the TPS gene family using HMMsearch with hidden Markov Models profiles including Terpene_synth_C (PF03936). A total of 11 AcaTPS and 15 AcoTPS assembled transcripts with represented full-length sequences were identified. To identify candidate P450s involved in oxidizing the diterpene skeletons, we performed gene co-expression network analysis⁸³ with self-organizing maps (SOMs) using the Kohonen package⁸⁴. Within the co-expressed modules (SOM nodes), we further calculated the Pearson correlation coefficients

between functional TPSs and other genes. Those genes annotated as P450s (Pfam: PF00067) and showing a high Pearson correlation coefficient with suitable protein sequence lengths (>300 amino acids) were selected as candidate P450 genes for further functional validation.

Phylogenetic analysis

Amino acids sequences of TPSs (Supplementary Data 3) and CYPs (Supplementary Data 6 and 7) from transcriptomics analysis in this work and those of the previously characterized ones were downloaded from NCBI and aligned by using clusterW. The neighbor-joining trees were constructed using MEGA11⁸⁵ using the poisson model with 1000 bootstrap resampling and visualized using chipLOT software⁸⁶.

Cloning and heterologous expression of *Aconitum* TPSs and P450s in *N. benthamiana*

The TransScript® One-Step gDNA Removal and cDNA Synthesis Super-Mix kit from Transgen (Beijing, China) was utilized for cDNA synthesis. Subsequently, candidate full-length TPS and P450 genes were amplified from the cDNA libraries of *A. carmichaelii* and *A. coreanum* using Phanta Flash Super-Fidelity DNA Polymerase (Vazyme, Nanjing, China) and gene-specific primers (Supplementary Data 10 and 11) via polymerase chain reaction (PCR). PCR products were purified and cloned into the pEAQ-HT vector with the ClonExpress Ultra One Step Cloning Kit (Vazyme, Nanjing, China). The correctly sequenced plasmids were transformed into *Agrobacterium* (LBA4404) via the heat shock method. Transformed cells were plated on selection agar media (LB agar with 50 µg/mL kanamycin, 50 µg/mL streptomycin, and 25 µg/mL rifampicin) and cultured at 28 °C for 72 h. Single colonies harboring the target vectors were inoculated into 5 mL of LB medium containing rifampicin (25 µg/mL) and kanamycin (50 µg/mL) and cultured overnight at 28 °C with shaking at 200 rpm. The bacterial cultures were centrifuged at 1574 g for 10 minutes, and pellets resuspended in 5 mL infiltration buffer (10 mM MES buffer, pH 5.6, 10 mM MgCl₂, 150 µM acetosyringone). The suspensions, adjusted to an OD₆₀₀ of 0.2, were incubated for 2 hours at room temperature before being infiltrated into *N. benthamiana* leaves using a 1 mL syringe. Five days post-infiltration, leaves were harvested, lyophilized, and analyzed using LC-MS and GC-MS. DNA sequence of IrKSL4 (KX580633) was synthesized by Sangon Biotech Incorporation. Genes of AtKS (NP_178064) and AtKO (AF047719) were cloned from the cDNA libraries of *A. thaliana*.

Metabolite extraction and mass spectrometry-based analysis

Dried leaf materials (15 mg) were ground into powder with 3 mm steel beads, and then extracted with 1 mL of ethyl acetate (EA) with sonication for 40 minutes. After centrifugation at 14,167 g for 10 minutes, 600 µL of the EA extract was transferred to a 2 mL sample vial for gas chromatography mass spectrometry (GC-MS) analysis. The GC-MS analysis was performed on an Agilent 8890-5977B system (Agilent Technologies, Santa Clara, CA, USA) equipped with a Zebron ZB-SHT Inferno capillary column (30 m × 0.25 mm × 0.10 µm). Helium was the carrier gas. The flow rate was set at 1.5 mL/min. The sample injection volume was 1 µL in splitless mode. Electron impact (EI) energy was 70 eV. The GC temperature program started at 60 °C for 2 minutes, ramped to 240 °C at 30 °C/min and held for 2 minutes, then ramped to 340 °C at 40 °C/min and held for 2 minutes. The solvent delay was set at 2.8 minutes, and full-scan MS data were acquired in the range of 50 to 800 Dalton. Ion source and quadrupole temperatures were set at 230 °C and 150 °C, respectively. Electron impact (EI) energy was 70 eV. GC-MS data were analyzed using Agilent MassHunter Qualitative Analysis 10.0 software, and the comparative GC-MS chromatograms shown in the Figures were representative ones from at least three biological replicates with similar results.

For ultra-high-performance liquid chromatography high-resolution mass spectrometry (UHPLC-HRMS) analysis, 100 µL of the EA extract was dried under nitrogen, reconstituted in 100 µL methanol, and centrifuged

at 14,167 g for 5 minutes. The supernatant (70 μ L) was analyzed with a Kinetex EVO C18 column (100 mm $\times$ 2.1 mm i.d., 1.7 μ m, 100 Å, Phenomenex) in a Vanquish UPLC system coupled with a Q-Exactive orbitrap mass spectrometer (Thermo Fisher Scientific). The mobile phases used are 0.1% formic acid in water (A) and acetonitrile (B). The flow rate was 0.3 mL/min. The injection volume was 5 μ L. Two gradient elution programs (Method 1 and 2) were used. Method 1 is as follows: 0–2 min at 10% B, increased to 56% B at 6 min, then to 90% B at 11.5 min, 100% B at 12 min, and held until 13 min, decreased to 10% B at 13.5 min, and equilibrated at 10% B for 1 min. Method 2 was used for detecting **T7** from plant extracts and its gradient as follow: 0–2 min at 10% B, increased to 40% B at 5 min, to 45% B at 9 min, to 46% at 13 minutes, and to 100% at 15 minutes and maintained for 2 minutes. At 17.5 minutes, B was reduced to 10% and held for 1 minute. The flow rate decreased from 0.3 mL/min to 0.2 mL/min from 2 min to 5 min and held until 13 minutes, and then increased to 0.3 mL/min at 15 minutes, and for the remainder of the time, the flow rate was kept at 0.3 mL/min. The authentic standard of tripterifordin was purchased from Chengdu MUST Bio-technology Co., Ltd. The MS capillary and heated electrospray ionization (HESI) source temperatures were set to 320 °C and 370 °C, respectively, with a source energy of 4.0 kV. Data were acquired in positive mode within a mass range of 250–700 m/z. LC-MS data were analyzed using Thermo Scientific XCalibur software. The comparative LC-MS extracted ion chromatograms (EICs) shown in the Figures were representative ones from at least three biological replicates with similar results.

For quantitative analysis, freeze-dried *A. carmichaelii* and *A. coreanum* tissues (10 mg) were homogenized using 3-mm steel beads, extracted with methanol (1 mL) with sonication for 1 h, then centrifuged at 14,167 g for 10 min. Aliquot of the supernatant (700 μ L) was analyzed by UHPLC-HRMS (Method 1) and target compounds quantified using the calibration curves of **T7**, mesaconitine, and hypaconitine generated (all showed $R^2 \geq 0.99$) in Supplementary Figs. 31 and 33.

Mesaconitine and hypaconitine were also quantified from the root exudates of *A. carmichaelii*. Three-week-old *A. carmichaelii* seedlings germinated from tuber roots were individually transferred into 50 mL sterile culture tubes containing 20 mL of B5 liquid medium with roots fully submerged. B5 medium without seedlings was used as a blank control. All tubes were incubated in a greenhouse (22 °C, 16/8 h light/dark cycle) for seven days, with fresh B5 medium supplemented daily. After the incubation for seven days, 10 mL medium from each tube was collected, extracted with an equal volume of ethyl acetate (1:1, v/v) at room temperature overnight. The organic extracts were evaporated to dryness using rotovap followed by a gentle nitrogen stream. The resulting residue was reconstituted in methanol (1 mL) for subsequent quantitative LC-MS analysis.

Large-scale extraction, isolation and purification of oxidation products from *N. benthamiana* leaves for NMR analysis

Large-scale agroinfiltration of gene sets into *N. benthamiana* were performed using the previously described method³⁹. After 5 days of infiltration, plant leaves were harvested and lyophilized. The dried leaf materials were ground into powder using a mortar and pestle, followed by extraction with ethyl acetate with sonication three times. Organic extracts were filtered and combined extracts dried and purified with repeated silica and C18 column chromatography, as well as semi-preparative liquid chromatography using the UltiMate 3000 system (Thermo Scientific, USA) to yield the pure compounds (**A4–A18**, **K9–11**, **M1**, **T1**, and **T8–T9**) for NMR analysis. More details can be found in the Supplementary Information.

NMR analysis

¹H and ¹³C along with 2D NMR experiments were performed with a Bruker Avance NEO 600 MHz spectrometer (Bruker Biospin, Karlsruhe, Germany). Samples were dissolved in CDCl₃ and NMR data acquired at 298 K. Chemical shifts were referenced to residual solvent

signals. Detailed spectra and structural assignments of *ent*-atiserene and *ent*-kaurene oxidation products can be found in Supplementary Figs. 35–86 and Supplementary Data 8 and 9.

Molecular docking and site-directed mutagenesis

The protein structure of AcoCYP93A1 was modeled using AlphaFold2 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). Docking was performed with AutoDock Tool⁸⁷ using default parameters, we selected the ligand model orientation with hydroxylation sites near the heme group for further investigation. Visualization of the modeling and docking was enabled with PyMol (Open-Source; PyMOL Molecular Graphics System; Schrödinger). Mutagenesis of selected residues in AcoCYP93A1 were constructed via overlap PCR using primers listed in Supplementary Data12.

Plant growth inhibition assay

Arabidopsis thaliana ecotype Col-0 seeds were sterilized in 1.5 mL Eppendorf tubes using a 1 mL solution of 80% (v/v) bleach and 0.025% (v/v) Tween-20 for 10 minutes. The bleach solution was decanted, and the seeds were rinsed eight times with 1 mL of sterile water. Seeds were then stratified in the dark at 4 °C for 2 days, then placed on 1/2 $\times$ Murashige and Skoog (MS) vitamin medium (PhytoTechnology Laboratories) plates (containing 1.1% agar and 1% sucrose with pH at 5.7). Assayed compounds were dissolved in DMSO to the required final concentration 100 μ M. The solutions of tested compounds (100 μ L) were applied evenly onto the sterile filter paper strips (5 mm wide and 10 cm long) placed at the center of the plates. An equal volume of DMSO was used as a control. The MS plates were placed vertically in a growth chamber and cultivated at 22 °C with 50% humidity under a photoperiod of 16 hours light/8 hours dark. After 10 days of cultivation, root lengths (root tip to hypocotyl) of *A. thaliana* were recorded. The experiments were conducted in three biological replicates for each compound tested.

Statistics and reproducibility

Data are presented as mean $\pm$ SD. The significance analysis was performed by one-way ANOVA analysis conducted using GraphPad Prism 9. The sample sizes (n) chosen were appropriate for the statistical analyses, and p values and sample sizes (n) are indicated and described in individual figures, Supplementary Figs. and the relevant legends. No data were excluded. The design and data collection of experiments were randomized. The results of all key experiments were reproducibly confirmed.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RNA-seq data from this study have been deposited in the NCBI Sequence Read Archive under the BioProject accessions PRJNA1187837 (*Aconitum carmichaelii*, <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1187837/>) and PRJNA1187874 (*Aconitum coreanum*, <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1187874/>). NMR, mass spectra, and other data that support the findings of this study are provided in the Supplementary Information/Source Data and Supplementary Data files. Source data are provided with this paper.

References

- Roncero, A. M., Tobal, I. E., Moro, R. F., Díez, D. & Marcos, I. S. Halimane diterpenoids: sources, structures, nomenclature and biological activities. *Nat. Prod. Rep.* **35**, 955–991 (2018).
- Jiang, B. et al. Characterization and heterologous reconstitution of *Taxus* biosynthetic enzymes leading to baccatin III. *Science* **0**, eadj3484 (2024).

3. Luo, D. et al. Oxidation and cyclization of casbene in the biosynthesis of *Euphorbia* factors from mature seeds of *Euphorbia lathyris* L. *Proc. Natl. Acad. Sci. USA* **113**, E5082–E5089 (2016).
4. Gao, J. et al. Triptolide: pharmacological spectrum, biosynthesis, chemical synthesis and derivatives. *Theranostics* **11**, 7199–7221 (2021).
5. Liu, X., Xu, J., Zhou, J. & Shen, Q. Oridonin and its derivatives for cancer treatment and overcoming therapeutic resistance. *Genes Dis.* **8**, 448–462 (2021).
6. Chen, K. et al. Anti-AIDS Agents, 4. Tripterifordin, a novel anti-HIV principle from *Tripterygium wilfordii*: Isolation and structural elucidation. *J. Nat. Prod.* **55**, 88–92 (1992).
7. Wang, J., Li, S., Xiong, Z. & Wang, Y. Pathway mining-based integration of critical enzyme parts for de novo biosynthesis of steviolglycosides sweetener in *Escherichia coli*. *Cell Res.* **26**, 258–261 (2016).
8. Hedden, P. The current status of research on gibberellin biosynthesis. *Plant Cell Physiol.* **61**, 1832–1849 (2020).
9. De La Peña, R. & Sattely, E. S. Rerouting plant terpene biosynthesis enables momilactone pathway elucidation. *Nat. Chem. Biol.* **17**, 205–212 (2021).
10. Yu, J., Tu, X. & Huang, A. C. Functions and biosynthesis of plant signaling metabolites mediating plant–microbe interactions. *Nat. Prod. Rep.* **39**, 1393–1422 (2022).
11. Peters, R. J. Two rings in them all: The labdane-related diterpenoids. *Nat. Prod. Rep.* **27**, 1521–1530 (2010).
12. Zerbe, P. et al. Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant Physiol.* **162**, 1073–1091 (2013).
13. Liu, X., Zhang, P., Zhao, Q. & Huang, A. C. Making small molecules in plants: A chassis for synthetic biology-based production of plant natural products. *J. Integr. Plant Biol.* **65**, 417–443 (2023).
14. Huang, A. C. & Osbourn, A. Plant terpenes that mediate below-ground interactions: prospects for bioengineering terpenoids for plant protection. *Pest Manag. Sci.* **75**, 2368–2377 (2019).
15. Köksal, M., Hu, H., Coates, R. M., Peters, R. J. & Christianson, D. W. Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. *Nat. Chem. Biol.* **7**, 431–433 (2011).
16. Zhou, K. et al. Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. *J. Biol. Chem.* **287**, 6840–6850 (2012).
17. Bathe, U. & Tissier, A. Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry* **161**, 149–162 (2019).
18. Heskes, A. M. et al. Biosynthesis of bioactive diterpenoids in the medicinal plant *Vitex agnus-castus*. *Plant J.* **93**, 943–958 (2018).
19. Wang, Q. et al. Characterization of CYP76M5-8 indicates metabolic plasticity within a plant biosynthetic gene cluster. *J. Biol. Chem.* **287**, 6159–6168 (2012).
20. Mao, L. et al. Genomic evidence for convergent evolution of gene clusters for momilactone biosynthesis in land plants. *Proc. Natl. Acad. Sci. USA* **117**, 12472–12480 (2020).
21. Sun, Y. et al. A chromosome-level genome assembly reveals that tandem-duplicated CYP706V oxidase genes control oridonin biosynthesis in the shoot apex of *Isodon rubescens*. *Mol. Plant* **16**, 517–532 (2023).
22. Wang Js, et al. A cytochrome P450 CYP81AM1 from *Tripterygium wilfordii* catalyses the C-15 hydroxylation of dehydroabiatic acid. *Planta* **254**, 95 (2021).
23. Zhang, Y. et al. Tandemly duplicated CYP82Ds catalyze 14-hydroxylation in triptolide biosynthesis and precursor production in *Saccharomyces cerevisiae*. *Nat. Commun.* **14**, 875 (2023).
24. Forman, V. et al. A gene cluster in *Ginkgo biloba* encodes unique multifunctional cytochrome P450s that initiate ginkgolide biosynthesis. *Nat. Commun.* **13**, 5143 (2022).
25. Gao, J. et al. CYP72D19 from *Tripterygium wilfordii* catalyzes C-2 hydroxylation of abietane-type diterpenoids. *Plant Cell Rep.* **42**, 1733–1744 (2023).
26. Morrone, D., Chen, X., Coates Robert, M. & Peters Reuben, J. Characterization of the kaurene oxidase CYP701A3, a multi-functional cytochrome P450 from gibberellin biosynthesis. *Biochem J.* **431**, 337–347 (2010).
27. Dong, L.-B. et al. Cryptic and stereospecific hydroxylation, oxidation, and reduction in platensimycin and platencin biosynthesis. *J. Am. Chem. Soc.* **141**, 4043–4050 (2019).
28. Smanski, M. J. et al. Expression of the platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. *J. Nat. Prod.* **75**, 2158–2167 (2012).
29. Drummond, G. J., Grant, P. S. & Brimble, M. A. ent-Atisane diterpenoids: isolation, structure and bioactivity. *Nat. Prod. Rep.* **38**, 330–345 (2021).
30. Zhou, G. et al. A review on phytochemistry and pharmacological activities of the processed lateral root of *Aconitum carmichaelii* Debeaux. *J. Ethnopharmacol.* **160**, 173–193 (2015).
31. Yin, T.-P., Yan, Y.-F. & He, J.-M. *Aconitum coreanum* Rapaics: Botany, traditional uses, phytochemistry, pharmacology, and toxicology. *Open Chem.* **20**, 1263–1282 (2022).
32. Shen, Y., Liang, W.-J., Shi, Y.-N., Kennelly, E. J. & Zhao, D.-K. Structural diversity, bioactivities, and biosynthesis of natural diterpenoid alkaloids. *Nat. Prod. Rep.* **37**, 763–796 (2020).
33. Sun, J. et al. Guanfu base A, an antiarrhythmic alkaloid of *Aconitum coreanum*, is a CYP2D6 inhibitor of human, monkey, and dog isoforms. *Drug Metab. Dispos.* **43**, 713–724 (2015).
34. Mao, L. et al. Functional identification of the terpene synthase family involved in diterpenoid alkaloids biosynthesis in *Aconitum carmichaelii*. *Acta Pharm. Sin. B* **11**, 3310–3321 (2021).
35. Tian, M. et al. Functional diversity of diterpene synthases in *Aconitum* plants. *Plant Physiol. Biochem* **202**, s (2023).
36. Chen, F., Tholl, D., Bohlmann, J. & Pichersky, E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* **66**, 212–229 (2011).
37. Wang, Z., Nelson, D. R., Zhang, J., Wan, X. & Peters, R. J. Plant (di)terpenoid evolution: from pigments to hormones and beyond. *Nat. Prod. Rep.* **40**, 452–469 (2023).
38. Sainsbury, F., Thuenemann, E. C. & Lomonosoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).
39. Huang, A. C. et al. Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci.* **114**, E6005–E6014 (2017).
40. Gnanasekaran, T. et al. Heterologous expression of the isopimaric acid pathway in *Nicotiana benthamiana* and the effect of N-terminal modifications of the involved cytochrome P450 enzyme. *J. Biol. Eng.* **9**, 24 (2015).
41. Ruiz-Sola, M. Á et al. *Arabidopsis* geranylgeranyl diphosphate synthase 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *N. Phytol.* **209**, 252–264 (2016).
42. Gao, J. et al. Transcriptome analysis of *Aconitum carmichaelii* identifies genes involved in terpenoid, alkaloid and phenylpropanoid biosynthesis. *Int J. Agric Biol.* **22**, 710–720 (2019).
43. Jin, B. et al. Functional diversification of kaurene synthase-like genes in *Isodon rubescens*. *Plant Physiol.* **174**, 943–955 (2017).
44. Fleet, C. M. et al. Overexpression of AtCPS and AtKS in *Arabidopsis* confers increased ent-kaurene production but no increase in bioactive gibberellins. *Plant Physiol.* **132**, 830–839 (2003).
45. Su, P. et al. Probing the single key amino acid responsible for the novel catalytic function of ent-kaurene oxidase supported by NADPH-cytochrome P450 reductases in *Tripterygium wilfordii*. *Front. Plant Sci.* **8**, (2017).
46. Irmisch, S. et al. One amino acid makes the difference: the formation of ent-kaurene and 16 α -hydroxy-ent-kaurane by diterpene synthases in poplar. *BMC Plant Biol.* **15**, 262 (2015).

47. Rudolf, J. D., Dong, L.-B., Manoogian, K. & Shen, B. Biosynthetic origin of the ether ring in platensimycin. *J. Am. Chem. Soc.* **138**, 16711–16721 (2016).
48. Nikkilä, J. et al. Analysis and visualization of gene expression data using Self-Organizing Maps. *Neural Netw.* **15**, 953–966 (2002).
49. Huang, A. C. et al. A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* **364**, eaau6389 (2019).
50. Zhou, Q. et al. Insight into neofunctionalization of 2,3-oxidosqualene cyclases in B,C-ring-opened triterpene biosynthesis in quinoa. *N. Phytol.* **241**, 764–778 (2024).
51. Geisler, K., Jensen, N. B., Yuen, M. M. S., Madilao, L. & Bohlmann, J. Modularity of conifer diterpene resin acid biosynthesis: P450 enzymes of different CYP720B clades use alternative substrates and converge on the same products. *Plant Physiol.* **171**, 152–164 (2016).
52. Lin, G.-M. & Voigt, C. A. Design of a redox-proficient *Escherichia coli* for screening terpenoids and modifying cytochrome P450s. *Nat. Catal.* **6**, 1016–1029 (2023).
53. Bell, R. A., Ireland, R. E. & Mander, L. N. Experiments directed toward the total synthesis of terpenes. IX. The total synthesis of (±)-hibaene and the oxygenation of some tetracyclic terpenes. *J. Org. Chem.* **31**, 2536–2542 (1966).
54. Song, L., Zhu, G., Liu, Y., Liu, B. & Qin, S. Total synthesis of atisane-type diterpenoids: application of diels–alder cycloadditions of podocarpane-type unmasked ortho-benzoquinones. *J. Am. Chem. Soc.* **137**, 13706–13714 (2015).
55. Baker, K. M. et al. Diterpenes. Part X. Some transformations of phyllocladene and isophyllocladene. *J. Chem. Soc. Perkin Trans. 1* 190–194 (1972).
56. Fraga, B. M., de Alfonso, I., Gonzalez-Vallejo, V. & Guillermo, R. Microbial transformation of two 15 α -hydroxy-ent-kaur-16-ene diterpenes by *Mucor plumbeus*. *Tetrahedron* **66**, 227–234 (2010).
57. Wang, Y. et al. Diterpenoids from the whole plant of *Euphorbia wallichii* and their protective effects on H₂O₂-induced BV-2 microglial cells injury. *Bioorg. Chem.* **128**, 106067 (2022).
58. Liang, Y. et al. Bioactive diterpenoids from the stems of *Euphorbia antiquorum*. *J. Nat. Prod.* **82**, 1634–1644 (2019).
59. Kitaoka, N., Wu, Y., Xu, M. & Peters, R. J. Optimization of recombinant expression enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis. *Appl. Microbiol. Biotechnol.* **99**, 7549–7558 (2015).
60. Yang, C.-H., Wang, X.-C., Tang, Q.-F., Liu, W.-Y. & Liu, J.-H. A new diterpenoid and a new diterpenoid alkaloid from *Aconitum coreanum*. *Helv. Chim. Acta* **91**, 759–765 (2008).
61. Shaari, K., Gray, A.I., Hockless, D.C.R., Waterman, P.G. & White, A.H. Structural studies of Kaurane Diterpenes From *Ryparosa acuminata* (Flacourtiaceae). *Aust. J. Chem.* **46**, 739–742 (1993).
62. Duan, H. et al. Immunosuppressive terpenoids from extracts of *Tripterygium wilfordii*. *Tetrahedron* **57**, 8413–8424 (2001).
63. Hedden, P. & Thomas Stephen, G. Gibberellin biosynthesis and its regulation. *Biochem. J.* **444**, 11–25 (2012).
64. Li, H.-Y. et al. Labdane-type diterpenoids from *Leonurus japonicus* and their plant-growth regulatory activity. *J. Nat. Prod.* **82**, 2568–2579 (2019).
65. Liu, Y. et al. Labdane diterpenoids from the heartwood of *Leucosceptrum canum* that impact on root growth and seed germination of *Arabidopsis thaliana*. *Phytochemistry* **206**, 113531 (2023).
66. Xu, M. et al. Genetic evidence for natural product-mediated plant–plant allelopathy in rice (*Oryza sativa*). *N. Phytol.* **193**, 570–575 (2012).
67. Kato, T. et al. Momilactones, growth inhibitors from rice, *oryza sativa* L. *Tetrahedron Lett.* **14**, 3861–3864 (1973).
68. Nelson, D. & Werck-Reichhart, D. A P450-centric view of plant evolution. *Plant J.* **66**, 194–211 (2011).
69. Xia, Y. et al. Two CYP93A enzymes play a dual role in isoflavonoid biosynthesis in *Glycine max* L. *Plant Physiol. Biochem.* **203**, 108073 (2023).
70. Kruse, T. et al. In planta biocatalysis screen of P450s identifies 8-methoxypsoralen as a substrate for the CYP82C subfamily, yielding original chemical structures. *Chem. Biol.* **15**, 149–156 (2008).
71. Lee, S. et al. Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **107**, 21205–21210 (2010).
72. Wang, J. et al. Discovery of key cytochrome P450 monooxygenase (C20ox) enables the complete synthesis of tripterifordin and neotripterifordin. *ACS Catal.* **15**, 2690–2702 (2025).
73. Gobu, F.-R. et al. Isolation, structure elucidation, and immunosuppressive activity of diterpenoids from *Ligularia fischeri*. *J. Nat. Prod.* **80**, 2263–2268 (2017).
74. Pichersky, E. & Lewinsohn, E. Convergent evolution in plant specialized metabolism. *Annu. Rev. Plant Biol.* **62**, 549–566 (2011).
75. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
76. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
77. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
78. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
79. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma.* **12**, 323 (2011).
80. Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
81. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral Genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
82. Bryant, D. M. et al. A Tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* **18**, 762–776 (2017).
83. Orme, A. et al. A noncanonical vacuolar sugar transferase required for biosynthesis of antimicrobial defense compounds in oat. *Proc. Natl. Acad. Sci.* 201914652 (2019).
84. Wehrens, R. & Buydens, L. M. C. Self- and Super-organizing Maps in R: The Kohonen Package. *J. Stat. Softw.* **21**, 1–19 (2007).
85. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
86. Xie, J. et al. Tree Visualization By One Table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res.* **51**, W587–w592 (2023).
87. Morris, G. M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

Acknowledgements

We are grateful to all the members of the A.C. Huang Lab for discussion and suggestions. We thank Dr Ziwei Dai (Southern University of Science and Technology) for helping with SOM analysis. We acknowledge the technical support from the computing cluster and mass spectrometry platform of SUSTech-PKU Institute of Plant and Food Science. This

study was supported by the Guangdong Basic and Applied Basic Research Foundation (2023A1515012550), the open project of BGI-Shenzhen (BGIRSZ20210014) and in part by the Shenzhen Science and Technology Program (grant no. 20231120191353002 to A.C.H.; KCXFZ20211020174802004 to A.C.H.; ZDSYS20230626091659010 to A.C.H.; JCYJ202408133000227 to Q.Z.), the National Science Foundation of China (grant no. 32370298 to A.C.H.), the open project of State Key Laboratory of Quality Research in Chinese Medicine funded by Macau Science and Technology Development Fund (Macau University of Science and Technology, 006/2023/SKL to G.Y.Z. and A.C.H.) and the Macao Science and Technology Development Fund (0019/2022/AGJ to G.Y.Z.).

Author contributions

A.C.H. conceived, designed and supervised this study. F.L. designed and conducted the majority of the experiments; Q.Z. performed bioinformatics analyses including de novo transcriptome assembly and mining; F.C. performed NMR-based structural elucidation; X.L. extracted *A. coreanum* RNA, cloned and validated a few functional genes; T.C. acquired sequencing data; G.Y.Z. acquired NMR data. F.L., Q.Z., F.C. and A.C.H. analyzed the results and wrote the manuscript.

Competing interests

A.C.H., F.L., Q.Z. and T.C. have filed a patent (202510153988.3) on the discovery and application of compound **A11**. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-61188-0>.

Correspondence and requests for materials should be addressed to Ancheng C. Huang.

Peer review information *Nature Communications* thanks Reuben Peters, Olof Emanuelsson, and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, modified publication 2026

Fei Luo^{1,5}, Qian Zhou^{1,5}, Feilong Chen¹, Xinyu Liu¹, Tsan-Yu Chiu^{2,3}, Guo-Yuan Zhu⁴ & Ancheng C. Huang¹✉

¹Shenzhen Key Laboratory of Plant Genetic Engineering and Molecular Design, SUSTech-PKU Institute of Plant and Food Science, Department of Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen, Guangdong, China. ²BGI Research, Shenzhen, China. ³HIM-BGI Omics Center, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China. ⁴State Key Laboratory of Quality Research in Chinese Medicine, Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Macau, People's Republic of China.

⁵These authors contributed equally: Fei Luo, Qian Zhou. ✉e-mail: huangac@sustech.edu.cn