

Review

A Review of Plant-Derived Diterpenoid Biosynthesis: From Structural Scaffold Diversity and Lineage-Associated Distribution to Enzyme Mining and Discovery Strategies

Yalan Zhao ^{1,2,*}, Mengyao Li ³, Shasha Zuo ⁴, Xiulin Han ^{2,*}  and Yupeng Liang ^{5,*}¹ Department of Basic Courses, Zunyi Medical and Pharmaceutical College, Zunyi 563006, China² Key Laboratory of Microbial Diversity in Southwest China, Ministry of Education, Yunnan Institute of Microbiology, School of Life Sciences, Yunnan University, Kunming 650500, China³ Department of Pharmacy, Zunyi Medical and Pharmaceutical College, Zunyi 563006, China; limyao2020@163.com⁴ School of Basic Medical Sciences, Medicine & Technology College of Zunyi Medical University, Zunyi 563000, China; 15120297265@163.com⁵ Academy for Advanced Interdisciplinary Studies, Nankai University, Tianjin 300350, China

* Correspondence: zhaoyalan2016yu@163.com (Y.Z.); xlhan@ynu.edu.cn (X.H.); 9820250113@nankai.edu.cn (Y.L.)

Abstract

Plant diterpenoids are a diverse class of natural products with important ecological roles and wide applications in the pharmaceutical, agricultural, food additive, and chemical industries. Biosynthesis represents a primary strategy for accessing these valuable compounds. However, the identification of downstream tailoring enzymes (hereafter referred to as tailoring enzymes) involved in diterpenoid biosynthetic pathways remains a major bottleneck, particularly in non-model plant species with limited genomic resources. This review summarizes current strategies for discovering plant diterpenoid biosynthetic pathways and recent advances in elucidating their metabolic routes. We further highlight the lineage-biased distribution of diterpene scaffolds across plant taxa. We propose that scaffold enrichment in specific evolutionary lineages, when integrated with enzyme family expansion and functional divergence, may provide a complementary framework for prioritizing candidate tailoring enzymes. Importantly, scaffold enrichment alone cannot establish enzyme function or evolutionary causality; rather, it provides a complementary layer of evidence that can guide future experimental investigation. Future perspectives include predictive substrate–enzyme mapping, computational and generative design of cytochrome P450 enzymes, and the integration of enzyme discovery, structural modeling, and heterologous chassis engineering.

Keywords: plant diterpenoids; diterpene biosynthesis; diterpene scaffold; diterpene synthases; cytochrome P450 monooxygenases



Academic Editor: Jixun Zhan

Received: 16 June 2026

Revised: 19 July 2026

Accepted: 25 July 2026

Published: 30 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Diterpenoids (C₂₀) are a major class of plant specialized metabolites and represent an important component of plant chemical diversity. More than 30,000 natural diterpenoids have been reported to date [1], some of which have important applications in medicine, food additives, and the fragrance industry (Figure 1). Diterpenoids often possess highly complex structures and multiple chiral centers. According to the degree of cyclization of their carbon scaffolds, they can be broadly classified into acyclic, bicyclic, tricyclic, tetracyclic, and

macrocyclic types [2]. Such structural diversity underlies the wide range of biological activities of diterpenoids, but also makes their biosynthetic pathways difficult to elucidate. Paclitaxel provides a representative example: its complete biosynthetic pathway has long been considered a major challenge in natural product biosynthesis, and key advances have only recently been achieved after decades of research [3].

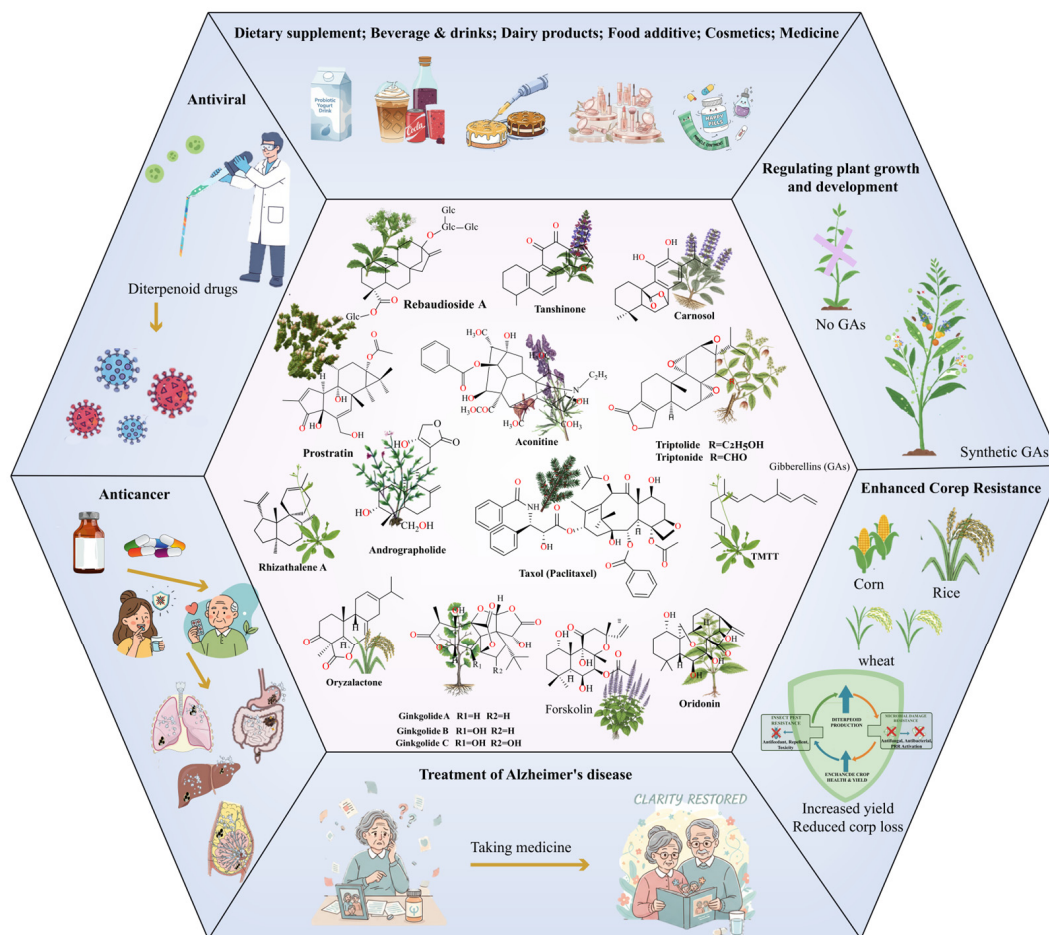


Figure 1. Major plant-derived diterpenoids and their ecological functions and applications.

The chemical diversity of diterpenoids is generated through a modular biosynthetic logic. Diterpene synthases (diTPSs) first cyclize the universal precursor geranylgeranyl diphosphate (GGPP) into a relatively limited set of core carbon scaffolds [4], such as kaurane- and clerodane-type scaffolds. Although more than 60,000 diterpenoids conforming to the C20 rule have been described, they are derived from fewer than 1000 basic scaffold types [1]. The expansion from a limited scaffold repertoire to a large number of specialized metabolites is largely driven by multi-site and sequential modifications of diterpene scaffolds by downstream tailoring enzymes, particularly cytochrome P450 monooxygenases (P450s). Therefore, in several experimentally characterized diterpenoid pathways, expansion and functional divergence of P450 families have been shown to contribute to chemical diversification by generating lineage-specific oxidative modification capacities. However, whether similar evolutionary patterns broadly apply across plant lineages remains unclear [5].

Although the general biosynthetic model of core scaffold formation followed by downstream modification is increasingly clear, complete elucidation of individual diterpenoid pathways remains challenging. To date, systematic characterization of plant diterpenoid pathways remains concentrated on a limited number of representative compounds

and compound classes, including paclitaxel [3], tanshinones [6–10], forskolin [11], andrographolide [12,13], ginkgolides [14], casbene [15], and steviol glycosides [16]. For most diterpenoids, especially with respect to functional annotation of downstream modifying enzymes, two major limitations remain. First, many diterpenoids are produced by non-model plants, for which high-quality genomes, multi-omics datasets, and stable genetic transformation systems are often unavailable. This restricts metabolic gene cluster mining and integrated multi-omics screening. Second, P450s constitute a large enzyme family with broad substrate adaptability and extensive functional redundancy. Tissue-specific expression, induced expression, or co-expression patterns alone are often insufficient to accurately identify the key P450s responsible for pathway-specific modifications from large candidate pools [17,18]. Improving the prioritization of candidate tailoring enzymes has therefore become a central bottleneck in the elucidation of plant diterpenoid biosynthetic pathways. We further highlight the lineage-biased distribution of diterpene scaffolds across plant taxa and propose a metabolite-centered, evolution-guided framework for candidate enzyme prioritization. In contrast to conventional gene-centered workflows, this framework starts from scaffold enrichment in specific plant lineages and integrates this information with gene family expansion and functional divergence to narrow candidate enzyme pools.

This bottleneck is closely associated with the ecological and adaptive evolution of plant specialized metabolites (PSMs). As sessile organisms, plants have evolved chemical defense systems centered on PSMs to respond to diverse biotic and abiotic stresses [19]. Diterpenoid accumulation is often inducible, tissue-specific, and environmentally dependent. For example, the biosynthesis of 17-hydroxygeranyllinalool diterpene glycosides (17-HGL-DTGs) in *Nicotiana attenuata* directly mediates resistance to the tobacco hornworm, *Manduca sexta* [20]. Monocots can specifically produce diterpenoid phytoalexins upon pathogen infection [21], while drought stress can reshape the terpenoid metabolic network in *Pinus elliottii* [22]. These examples indicate that diterpenoid metabolism is not uniformly distributed across the plant kingdom but is shaped by lineage evolution, niche adaptation, and environmental selection pressures. Accordingly, enzyme families involved in diterpenoid post-scaffold modification may have undergone gene duplication, expansion, and functional divergence in certain plant lineages, thereby contributing to lineage-specific metabolic capacities.

Compared with highly specialized terminal diterpenoid derivatives, which often have narrow taxonomic distributions, diterpene core scaffolds are relatively more stable across plants and often show enrichment patterns in particular lineages. Thus, core scaffolds not only provide a basis for structural classification, but also offer useful clues for tracing the evolution of associated biosynthetic genes. In this review, we systematically summarize the structural diversity of plant-derived diterpenoids, the lineage distribution of their core scaffolds, and recent progress in the elucidation of representative diterpenoid biosynthetic pathways. Building on this synthesis, we further propose a candidate enzyme discovery strategy that integrates plant chemotaxonomy with molecular phylogenetics. The enrichment of particular diterpene scaffolds in specific plant lineages can provide an initial prioritization signal. This signal can then be integrated with evidence of lineage-specific expansion and functional divergence in tailoring enzyme families derived from multispecies genomic or transcriptomic datasets. We clarify that this metabolite-centered, evolution-guided framework is intended to complement, rather than replace, established gene-centered approaches such as gene cluster mining, co-expression analysis, and homology-based prediction. By integrating lineage-associated scaffold enrichment with enzyme family expansion and functional divergence, it provides an additional layer of evidence for generating testable hypotheses and prioritizing candidate genes within large tailoring enzyme families, particularly P450s, before biochemical and genetic validation.

2. Current Strategies for Identifying Plant-Derived Diterpenoid Biosynthetic Pathways

Multi-omics datasets, including genomics, transcriptomics, metabolomics, and proteomics, have become an important foundation for elucidating plant specialized metabolic pathways [23,24]. For plant-derived diterpenoids, the central challenge is not only to identify diTPs responsible for scaffold formation, but also to select the key enzymes that truly participate in downstream oxidation, glycosylation, acylation, or rearrangement reactions from large and functionally complex families of modifying enzymes. To address this challenge, current pathway-discovery studies mainly use metabolic gene cluster mining, transcriptome co-expression analysis, homology-based prediction, proteomics-based screening, and spatiotemporal expression analysis (Figure 2).

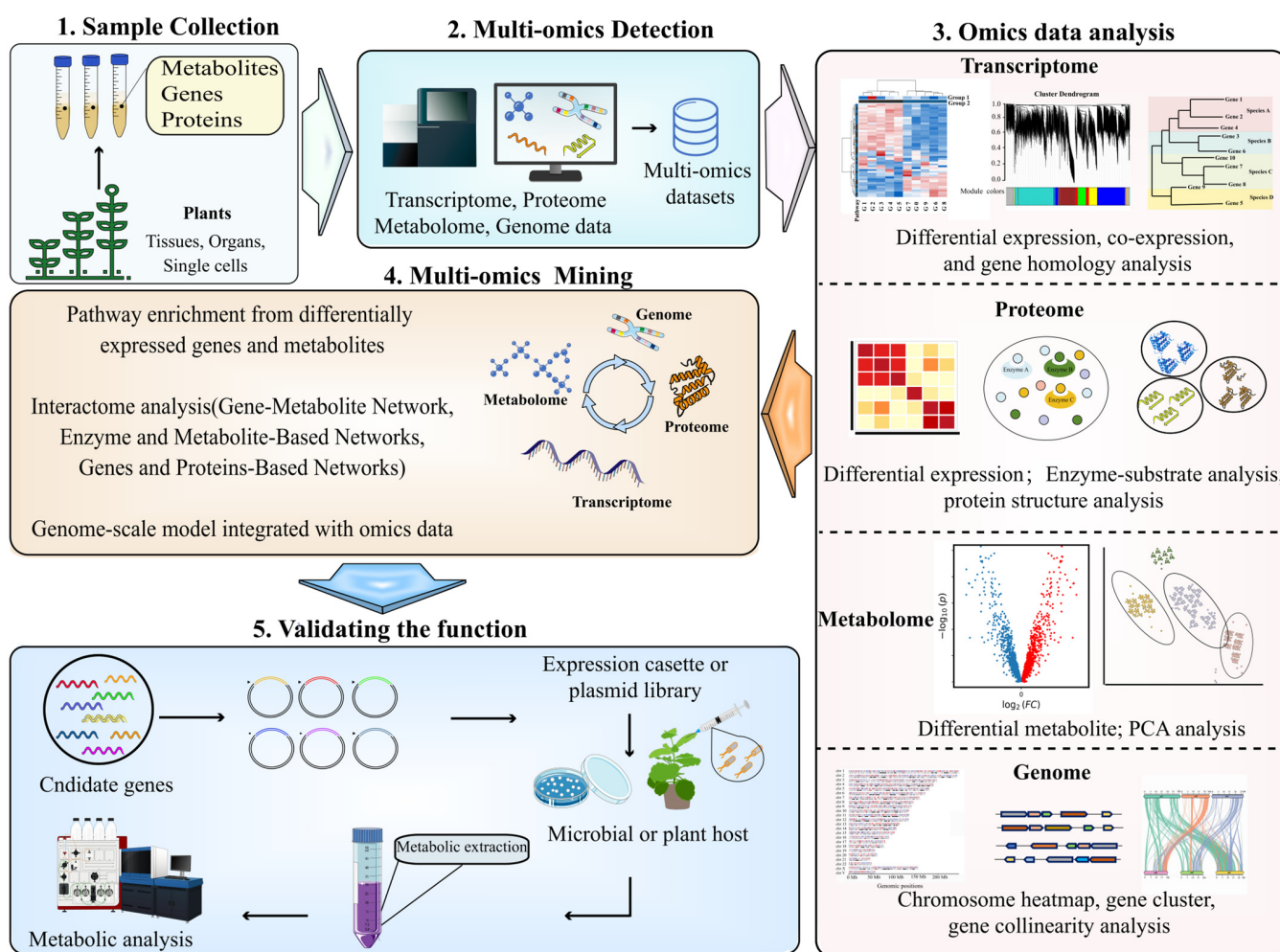


Figure 2. Schematic workflow of multi-omics approaches for the discovery and functional validation of diterpenoid biosynthetic genes.

2.1. Mining Strategies Based on Metabolic Gene Clusters

Plant metabolic gene clusters (MGCs) generally refer to sets of physically neighboring genes in the genome that participate in the same metabolic pathway. Compared with microorganisms, known MGCs in plants appear to be more limited in distribution and more difficult to identify. Nevertheless, gene clustering has been shown to serve as an important evolutionary and regulatory mechanism in some diterpenoid biosynthetic pathways [25]. Genome mining is one of the major approaches for discovering MGCs. For example, Forman et al. (2022) identified a gene cluster containing five P450 genes in the *Ginkgo biloba* genome [14]. This cluster is tightly linked to the scaffold-forming enzyme gene *GbLPS* and

shows root-enriched co-expression. Subsequent functional studies demonstrated that the cluster encodes key oxygenases involved in ginkgolide biosynthesis. In *Salvia miltiorrhiza*, gene family evolutionary analysis enabled the identification of CYP71D373 and CYP71D375, which are located near diterpene synthase genes, and revealed their roles in catalyzing D-ring formation during tanshinone biosynthesis [10].

In addition to linear gene proximity, plants can also coordinate gene regulation through higher-order chromosomal organization. Some metabolic gene clusters may span relatively long linear distances but become spatially associated through three-dimensional chromatin topology, thereby forming coordinated transcriptional states [26]. For example, the gene cluster associated with the *ent*-10-oxodepressin diterpenoid phytoalexin pathway on rice chromosome 7, which includes *OsECBS* and multiple *CYP71Z* genes, exhibits strong co-expression under induced conditions [27]. In recent years, the development of bioinformatic tools such as plantiSMASH 2.0 has improved the efficiency of MGC prediction in plant genomes and has also facilitated the preliminary screening of gene clusters associated with diterpenoid pathways [28].

However, MGCs do not represent a universal pattern of gene organization in plant diterpenoid biosynthesis. Some pathway genes may be dispersed across different chromosomal regions, and genomic proximity does not necessarily indicate functional association. For example, in *Scutellaria barbata*, although some P450 genes are located near diterpene synthase genes, functional validation did not support their involvement in the corresponding diterpenoid pathway [29]. Therefore, candidate gene prediction based solely on genomic location carries inherent risks and generally requires integration with transcriptome co-expression, metabolite accumulation patterns, and functional validation.

2.2. Prediction Strategies Based on Co-Expression Analysis

Co-expression analysis is based on the premise that genes involved in the same metabolic pathway often show similar or coordinated expression patterns. By comparing transcriptomic datasets across different tissues, developmental stages, or induced conditions, researchers can identify candidate genes whose expression is highly correlated with known pathway genes or with the accumulation patterns of target metabolites. Compared with MGC mining, co-expression analysis does not depend on genomic location and is therefore better suited for identifying non-clustered metabolic pathway genes.

Inducible expression datasets provide a typical application scenario for co-expression analysis in diterpenoid pathway elucidation. In tanshinone biosynthesis, for example, transcriptome analysis of Ag⁺-induced *S. miltiorrhiza* hairy roots enabled the identification of candidate genes whose expression patterns were highly correlated with the known key enzyme CYP76AH1. This approach led to the characterization of CYP76AH3 and CYP76AK1, which are responsible for ferruginol C-11 hydroxylation and subsequent oxidation reactions [7,9]. Similarly, in *Isodon lophanthoides* var. *graciliflorus*, screening for candidate genes induced by methyl jasmonate (MeJA) and co-expressed with upstream diTPSs led to the identification of several CYP76AH subfamily members involved in the formation of ferruginol derivatives [30]. These examples indicate that, under conditions involving induction, tissue-specific expression, or marked accumulation of target metabolites, co-expression analysis can effectively narrow the candidate gene pool and improve the efficiency of key enzyme screening.

However, co-expression alone does not directly establish enzyme function. For large gene families such as P450s, UDP-dependent glycosyltransferases (UGTs), and BAHD acyltransferases, many genes may respond synchronously under the same treatment condition. These may include enzymes that truly participate in the target pathway, as well as genes associated with stress responses, developmental regulation, or other metabolic processes.

Therefore, co-expression analysis usually needs to be combined with metabolite profiling, phylogenetic analysis, and in vivo or in vitro functional validation.

2.3. Prediction Strategies Based on Homology

Homology-based prediction based on sequence similarity is a classical approach for mining conserved pathway enzymes. This strategy typically uses functionally characterized enzymes as query sequences and applies algorithms such as BLAST(v2.14.0) and HMMER (v3.3.2) to identify orthologs or paralogs in target species. Candidate genes are then evaluated by integrating expression patterns and functional validation to determine whether they participate in the pathway of interest.

Homology-based prediction is particularly useful for identifying diTPSs. For example, Brückner, Božić, and colleagues used known copalyl diphosphate synthase (CPS) and kaurene synthase-like (KSL) genes as reference sequences and successfully cloned PdiTPS family members involved in the formation of the carnosic acid scaffold from sage and rosemary [31,32]. For downstream modifying enzymes, Zi, Peters, and colleagues used the tanshinone oxidase CYP76AH1 as a template to identify four homologous genes, CYP76AH4–7, from the rosemary transcriptome, and confirmed that CYP76AH4 has similar oxidative activity [33].

Nevertheless, homology-based prediction depends strongly on the number and functional coverage of available reference enzymes. For novel scaffolds, unusual modification reactions, or lineage-specific pathways, key enzymes may originate from highly diverged paralogous branches or may show relatively low sequence similarity to known enzymes. In such cases, reliance on sequence similarity alone may miss genuine candidate enzymes or incorrectly assign homologs with similar activities but unrelated pathway roles as target enzymes. Therefore, homology-based prediction is better suited for preliminary screening of conserved reaction types and should not be used as the sole basis for elucidating new and complex diterpenoid pathways.

2.4. Functional Screening Based on Proteomics

Unlike transcriptomics, which provides indirect evidence, proteomics can more directly reflect enzyme abundance and biochemical activity. In particular, when combined with chemical probe-based approaches, proteomics can enable targeted capture of substrate-binding proteins and thereby provide evidence that is closer to the functional level for candidate enzyme screening.

In the elucidation of the steviol glycoside pathway, researchers first screened 68 candidate UGTs through global proteomic analysis and then designed an ST-Dayne photoaffinity probe based on the structure of steviol. This probe labeled substrate-binding proteins and ultimately helped identify UGT73E1 as the key enzyme responsible for C-4 glucosylation [34]. This activity-based proteomic strategy, known as activity-based protein profiling (ABPP), provides an effective tool for addressing functional redundancy within large gene families.

Proteomics-based screening also has several limitations. First, target protein abundance, membrane localization, and extraction efficiency can all affect detection. Second, for membrane-associated enzymes such as P450s, protein extraction, maintenance of protein stability, and in vitro activity validation remain technically challenging. Third, chemical probe design depends on prior knowledge of substrate or intermediate structures, whereas key intermediates in many unknown diterpenoid pathways have not yet been defined.

Here, we systematically compare five strategies for discovering genes involved in plant specialized metabolic pathways, as summarized in Table 1. Collectively, these five strategies represent different trade-offs between the efficiency of high-throughput screening

and the certainty of functional validation, and their practical application often requires an integrated, multi-strategy approach.

Table 1. Summary of major strategies for discovering plant diterpenoid biosynthetic pathways.

Strategy	Main Evidence/Data	Typical Application	Strengths	Limitations	Refs.
Metabolic gene cluster mining	Genome assembly, gene proximity, collinearity, cluster prediction	Identification of physically clustered pathway genes	Effective when pathway genes are genomically clustered	Many plant pathway genes are dispersed; proximity does not prove function	[25,26]
Co-expression analysis	Transcriptomes across tissues, stages, or treatments	Screening genes co-expressed with known diTPSs or metabolite accumulation	Useful for non-clustered pathways; narrows candidate pools	Stress-responsive or tissue-specific genes may generate false positives	[17,18]
Homology-based prediction	BLAST/HMMER, known characterized enzymes	Identification of conserved diTPSs or related P450s	Simple and effective for conserved reaction types	Poor performance for novel reactions or highly diverged enzymes	[35,36]
Spatiotemporal expression association	Tissue-specific or inducible expression plus metabolite localization	Linking metabolite accumulation sites with candidate genes	Useful when metabolites accumulate in specific tissues or conditions	Expression overlap alone cannot establish biochemical function	[37]
Proteomics/ABPP	Protein abundance, substrate probes, activity-based labeling	Functional-level screening of substrate-binding enzymes	Closer to enzyme activity than transcriptomics	Probe design and membrane-protein detection are technically difficult	[38]

2.5. Integrated Strategies Based on Expression Specificity and Evolutionary Associations

The accumulation of PSMs often shows clear tissue specificity, developmental-stage specificity, and inducible responses. Based on the principle of metabolite–transcript co-localization, when a metabolite accumulates strongly in a particular tissue or under a specific condition, candidate enzyme genes that are also highly expressed in the same tissue or condition are more likely to participate in the biosynthesis of that metabolite.

Pateraki et al. focused on root cork cells of *Coleus forskohlii*, which are enriched in forskolin precursors, and used comparative transcriptomics to identify CYP76AH subfamily members and acyltransferases that were specifically highly expressed in this tissue. These genes enabled reconstruction of the forskolin pathway [39]. Similarly, in rosemary, carnosic acid accumulates mainly in young leaves. The high expression of *RoKSL2–RoCPS2* is consistent with the spatial distribution of carnosic acid, and the young leaf-enriched *RoCYP76AK7/8* enzymes have also been shown to participate in the biosynthesis of this diterpenoid [40].

However, spatiotemporal expression specificity alone is not sufficient to exclude interference from numerous candidate P450s within the same enzyme family. In recent years, some studies have begun to incorporate molecular evolutionary perspectives in addition to expression-based associations. In elucidating the oridonin pathway, Sun et al. not only examined highly expressed genes in shoot tips, the major biosynthetic site, but also found that the CYP706V subfamily involved in complex oxidative modifications had undergone marked lineage-specific expansion in the genus *Isodon* [41]. Notably, the expansion of the CYP706V subfamily corresponds to the abundant accumulation of kaurane-type diterpenoids in this lineage. This case suggests that the enrichment of specific diterpene scaffolds or scaffold-derived compounds in plant lineages may be associated with the expansion and functional divergence of related modifying enzyme families.

Previous studies have generally suggested that the diversification of PSMs is closely linked to gene duplication, functional divergence or neofunctionalization, and lineage-specific expansion [42–45]. Therefore, integrating chemotaxonomic and molecular phylogenetic information with conventional co-expression analysis, homology-based prediction, and tissue-specific expression analysis may improve the reliability of prioritization of candidate tailoring enzymes from an evolutionary perspective. This integrated strategy may be particularly useful for complex diterpenoid pathways that are specific to certain genera or families, as it can provide a more targeted set of candidate genes for subsequent functional validation.

3. Prioritizing Candidate Tailoring Enzymes Through Lineage-Biased Scaffold Distribution and Gene Family Expansion

The strategies described above have advanced the elucidation of plant-derived diterpenoid biosynthetic pathways from different perspectives. The proposed framework differs from existing workflows mainly in its direction of inference. Metabolic gene cluster mining, co-expression analysis, homology-based prediction, proteomics-based screening, and spatiotemporal expression analysis are primarily gene-centered or enzyme-centered approaches: they first identify candidate genes and then evaluate whether these candidates explain the formation of target metabolites. By contrast, the framework proposed here is metabolite-centered: it first examines whether particular diterpene scaffolds or scaffold-derived compounds are enriched in specific plant lineages, and then evaluates whether relevant tailoring enzyme families show lineage-specific expansion, divergence, or expression patterns, providing evolutionary context for interpreting these chemical distributions. MGC mining facilitates the identification of pathway genes that are physically clustered in the genome; co-expression analysis helps screen candidate genes with coordinated expression patterns; homology-based prediction is useful for tracing enzymes involved in conserved reaction types; and proteomics can provide evidence that is closer to the functional level. However, most of these approaches still begin with “genes or enzymes,” that is, they first search for potential pathway enzymes within candidate gene sets and then validate their functions experimentally. For large modifying enzyme families such as P450s, which often show broad substrate adaptability and rapid functional divergence, sequence similarity, expression correlation, or genomic location alone is insufficient to fully explain the lineage-biased patterns of plant diterpenoid chemical diversity. It is therefore useful to further consider metabolite structures and lineage distributions themselves, and to connect the enrichment of specific diterpene scaffolds with the expansion and divergence of candidate modifying enzyme families. Based on this rationale, this review proposes a data-driven framework that uses the distribution of diterpene core scaffolds as a guide and gene family expansion as an important reference for prioritizing candidate tailoring enzymes.

3.1. Data Curation, Scaffold Extraction, Structural Annotation, and Statistical Analysis of Lineage-Associated Diterpenoid Distributions

To investigate the lineage distribution of plant-derived diterpenoids from a structural perspective, we systematically collected and curated structural information, source information, and scaffold characteristics of diterpenoid compounds from publicly available natural product databases and published literature up to June 2025. The initial dataset contained 64,132 diterpenoid-related records. Records were retained only when they met the following criteria: (i) clear plant-derived source information was available; (ii) sufficient structural information was provided to determine the core carbon skeleton; and (iii) the carbon framework met the C₂₀ diterpenoid criterion. Records lacking reliable structural in-

formation, organismal source information, plant origin, or a definable core carbon scaffold, as well as obvious non-diterpenoid entries, were excluded.

Data redundancy was removed through a multi-step curation workflow based on compound names (including common synonyms), molecular formulas, structural identifiers (SMILES and InChIKey where available), carbon-number features, and source records. Ambiguous or borderline cases were manually inspected. Exact duplicate database entries were consolidated, whereas records linking the same compound to different plant sources were retained as distinct compound–source records. After curation, 45,177 plant-derived diterpenoid records were retained, representing 589 canonical C20 scaffolds (Supplementary Table S1).

The carbon skeletons of diterpenoids were extracted using RDKit following a previously reported three-step procedure [46]. First, bonds between carbon atoms and heteroatoms were cleaved to separate the original molecular structures into multiple carbon-containing fragments. Second, all bonds within each fragment were converted into single bonds, with stereochemical information removed. Finally, the largest fragment containing exactly $5n$ carbon atoms ($n = 4$, corresponding to 20 carbons) was selected as the core carbon scaffold of each diterpenoid (Supplementary Table S2). This procedure enables structural comparison based on carbon skeleton topology rather than variations in oxidation state, substitution patterns, or stereochemistry.

To further evaluate structural diversification associated with oxidative modification, the oxidation level of diterpenoid molecules was estimated based on oxygen-containing functional groups and oxidation sites located on the core carbon scaffold. Oxygen-containing sites were classified according to their bonding patterns. Carbon atoms bonded to oxygen through a single bond were classified as hydroxylation sites, carbon atoms double-bonded to oxygen were classified as carbonylation sites, and carboxyl groups containing both C–O and C=O bonds were classified as carboxylation sites. The degree of molecular oxygenation was approximated using the oxygen-to-carbon ratio. All structural calculations were performed using RDKit, and complete oxidation-related results are provided in Supplementary Table S3.

Taxonomic information was standardized based on the original source records of each compound and cross-referenced with accepted plant taxonomy where possible. Species-level information was organized into genus, family, and higher taxonomic categories, enabling comparative analyses of diterpenoid scaffold distributions and structural characteristics among different plant lineages. Dataset integrity was further evaluated through cross-checking against published diterpenoid reviews and manual inspection of randomly selected records.

A reproducible statistical workflow was implemented to quantify lineage-associated scaffold distributions. Undefined and unclassified taxonomic labels were excluded from inferential lineage analyses. For the overall association analysis, lineages represented by at least 50 records and scaffolds represented by at least 20 records were retained. Pearson contingency statistics were calculated; because 86.9–94.8% of expected cell counts were below five, significance was determined by fixed-margin Monte Carlo/permutation sampling (999 replicates for family and genus analyses and 499 replicates for the species-label analysis). Cramér's V was reported as the association effect size.

For each eligible lineage–scaffold pair, a one-sided hypergeometric test, equivalent to a one-sided Fisher exact test under fixed margins, compared the observed count with the count expected from the complete 45,177-record empirical background. p values were adjusted separately at the family, genus, and species-label levels using the Benjamini–Hochberg procedure. Reported enriched pairs required at least five observed records, fold enrichment greater than one, and $q < 0.05$. Extremely small probabilities below double-

precision reporting limits are reported as $q < 1 \times 10^{-300}$. Extremely small q values reflect statistical overrepresentation within the curated record-level dataset and should not be interpreted as correspondingly strong biological evidence.

Because database and literature records are non-independent and are influenced by research focus, species availability, analytical accessibility, database coverage, and publication bias, the curated dataset does not represent a uniform or random survey of plant diterpenoid diversity. Record-level permutation and enrichment analyses were therefore complemented by a binary named-species occurrence analysis, in which each species–scaffold combination was counted once. Spearman correlations between lineage record count and scaffold richness were used to assess potential sampling effects. These analyses quantify overrepresentation relative to the curated empirical background but do not estimate absolute abundance in nature or establish enzyme function or evolutionary causality. The resulting patterns should therefore be treated as hypotheses requiring validation through broader taxonomic sampling, phylogenetically informed comparative analyses, genomics, expression profiling, biochemical characterization, and genetic experiments.

3.2. Lineage Preferences and Oxidative Modification Features of Plant-Derived Diterpene Scaffolds

Based on the curated dataset, source statistics indicate that green plants (Viridiplantae) represent a major reported source of natural diterpenoid diversity. In the original diterpenoid records, green plants accounted for a relatively high proportion of the dataset compared with mammals, fungi, bacteria, and other groups (Figure 3A). Although this proportion may be influenced by research attention in natural product chemistry and the scope of database coverage, it nevertheless suggests that plants represent an important source of diterpenoid chemical diversity. Among the 45,177 plant-derived compound–source records, 589 core skeletons conforming to the C20 rule were identified. Linear, bicyclic, tricyclic, tetracyclic, and macrocyclic scaffolds constituted the major scaffold types (Figure 3B) and showed certain lineage-biased patterns across plant groups. For example, bicyclic scaffolds were relatively abundant in Asteraceae and Taxaceae, tetracyclic kaurane-type scaffolds were prominent in the genus *Isodon* within Lamiaceae, and macrocyclic scaffolds were mainly found in Euphorbiaceae and *Taxus*-related groups. Further comparison of compound numbers and scaffold-type numbers showed that, in some plant groups, a large number of diterpenoid compounds are often derived from a relatively limited number of scaffold types. For example, in Lamiaceae, the ratio between compound number and scaffold-type number was relatively high, with an average of approximately 50 compounds corresponding to one scaffold type (Figure 3C–E). This pattern suggests that the diversification of plant diterpenoids does not depend solely on the continual generation of new carbon scaffolds, but may rely more strongly on the sustained modification and derivatization of existing core scaffolds. Therefore, “diterpene scaffold conservation” does not imply a lack of scaffold variation. Rather, it indicates that, within certain plant lineages, a limited set of core scaffolds can give rise to numerous derivatives through post-scaffold modification reactions.

Quantitative analysis supported a non-random association between taxonomic lineage and common scaffold distribution. At the family level, 40,005 records from 59 eligible families and 98 common scaffolds yielded Cramér’s $V = 0.465$ (permutation $p \leq 0.001$). At the genus level, 32,795 records from 152 genera and 100 scaffolds yielded Cramér’s $V = 0.484$ ($p \leq 0.001$). The species-label analysis included 15,256 records, 151 labels, and 94 scaffolds and yielded Cramér’s $V = 0.477$ ($p \leq 0.002$). Because 86.9–94.8% of expected cells were below five, permutation p values rather than asymptotic chi-square p values were used for inference (Supplementary Table S8).

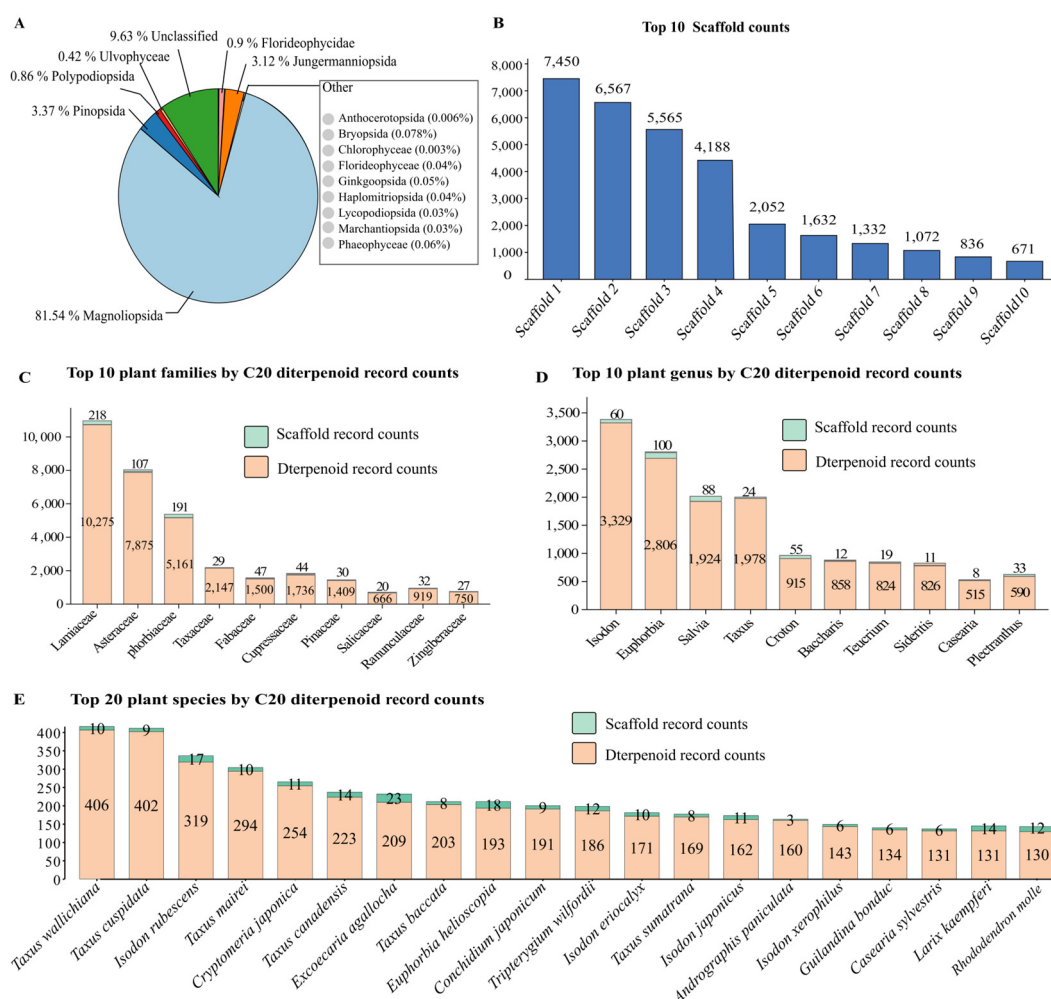


Figure 3. Counts of curated diterpenoid compound–source records and canonical scaffold diversity across taxonomic groups. **(A)** Distribution of curated diterpenoid records among major source groups. **(B)** The ten most frequently reported canonical C20 scaffolds; the corresponding structures are provided in Supplementary Table S2. **(C–E)** Curated diterpenoid compound–source record counts and canonical scaffold counts in the top 10 families, top 10 genera, and top 20 species ranked by diterpenoid record count, respectively; detailed data are provided in Supplementary Table S4. The reported record counts represent minimum estimates based on currently available database and literature records and should not be interpreted as absolute abundances in nature.

Pairwise enrichment tests identified lineage–scaffold combinations contributing to the overall association. In *Isodon*, kaurane-type scaffold 2 occurred 2351 times compared with 483.9 expected under the complete empirical background (4.86-fold enrichment; odds ratio = 21.46; BH $q < 1 \times 10^{-300}$). Other illustrative associations included abietane-type scaffold 4 in *Salvia* (999 observed versus 178.4 expected; 5.60-fold; odds ratio = 13.57; $q < 1 \times 10^{-300}$), scaffold 9 in *Euphorbia* (784 versus 51.9; 15.10-fold; odds ratio = 315.55; $q < 1 \times 10^{-300}$), and scaffold 20 in *Taxus* (342 versus 17.3; 19.83-fold; odds ratio = 173.46; $q < 1 \times 10^{-300}$) (Supplementary Tables S9–S11).

The named-species occurrence analysis provided a conservative sensitivity assessment. Scaffold 2 was recorded in 57 of 60 named *Isodon* species, compared with 11.5 species expected from the empirical species background (4.97-fold enrichment; odds ratio = 88.11; $q = 1.38 \times 10^{-35}$). Scaffold 4 was recorded in 102 of 155 named *Salvia* species (4.57-fold; odds ratio = 14.18; $q = 2.52 \times 10^{-48}$). Representative *Euphorbia* and *Taxus* associations were also retained (Supplementary Table S12).

The statistical results require cautious interpretation. Observed scaffold richness was strongly correlated with record count at the family (Spearman's $\rho = 0.850$, $p = 7.32 \times 10^{-71}$), genus ($\rho = 0.750$, $p = 1.72 \times 10^{-219}$), and species-label levels ($\rho = 0.615$, $p < 1 \times 10^{-300}$). Thus, the reported associations are database- and literature-supported overrepresentation patterns. They provide quantitative candidate clues for enzyme family prioritization but do not establish natural prevalence, enzyme function, or an evolutionary causal relationship.

3.3. An Evidence-Guided Framework for Prioritizing Candidate Tailoring Enzymes by Integrating Scaffold-Derived Chemical Diversity and Gene Family Evolution

Following the descriptive criteria defined in Section 3.2, lineage-biased distributions of diterpene scaffolds and their derivatives can be used to identify potential chemical diversity hotspot lineages. Diterpenoids have long been used as chemotaxonomic markers within Lamiaceae [47]. Structurally diverse kaurane-type diterpenoids are particularly well represented in *Isodon* [48,49]. Although kaurane-type scaffolds are widely distributed throughout the plant kingdom because of their roles in gibberellin biosynthesis, structurally diversified specialized metabolites derived from these scaffolds are particularly well represented in the genus *Isodon*, including *I. rubescens*. Here, “potential chemical diversity hotspot” refers to a lineage showing statistically supported overrepresentation relative to the curated empirical background. It refers to a lineage in which the reported diversity or relative representation of derivatives derived from a particular scaffold is comparatively high. This pattern should not, however, be interpreted as evidence of natural prevalence or evolutionary causality.

Representative diterpenoid pathways provide different levels of evidence linking chemical diversification with tailoring enzyme evolution. In *Isodon rubescens*, one of the species particularly rich in kaurane diterpenoids (Figure 4A), Sun et al. identified lineage-specific expansion and tandem duplication of the CYP706V subfamily. Functional studies further demonstrated that these expanded CYP706V members catalyze diverse site-specific oxidations of *ent*-kaurene and its derivatives, representing an example in which gene family evolution and kaurane diterpenoid diversification are supported within the same biosynthetic context [41]. In *Salvia*, abietane-type diterpenoids are relatively well represented in several species in the curated dataset (Figure 4B). Functional divergence within the CYP76AK subfamily, including the retention, modification, and loss of ancestral catalytic activities, has been proposed to contribute to lineage-associated variation in abietane-type diterpenoid profiles [50]. Together with the *Isodon* example, the *Salvia* system supports the biological plausibility of integrating scaffold distributions with gene family evolution for candidate enzyme prioritization. These examples should nevertheless be regarded as evidence-guided case studies rather than universal evolutionary rules.

Collectively, the *Isodon* and *Salvia* studies validate different components of the proposed framework: the *Isodon* example connects lineage-associated gene family evolution with experimentally characterized catalytic functions; the *Salvia* example demonstrates how functional retention, modification, and loss within an enzyme subfamily can contribute to clade-associated metabolic variation. Together, these findings support the biological plausibility of integrating metabolite distributions with enzyme family evolution, but they do not demonstrate that all diterpenoid-rich lineages follow a common expansion–neofunctionalization–chemical enrichment process.

On the basis of these precedents and the descriptive chemical distribution patterns identified in the curated dataset, we propose the following testable hypothesis: in some plant lineages characterized by comparatively high scaffold-derived chemical diversity, downstream tailoring-enzyme families may have undergone lineage-specific expansion, retention, or functional divergence, thereby contributing to their capacity to generate diverse derivatives of particular diterpene scaffolds. This hypothesis does not imply

that scaffold enrichment necessarily causes enzyme family expansion, that expansion invariably results in neofunctionalization, or that gene family size alone predicts metabolic diversity. Based on this hypothesis, a metabolite-centered workflow for candidate enzyme prioritization can be proposed (Figure 5).

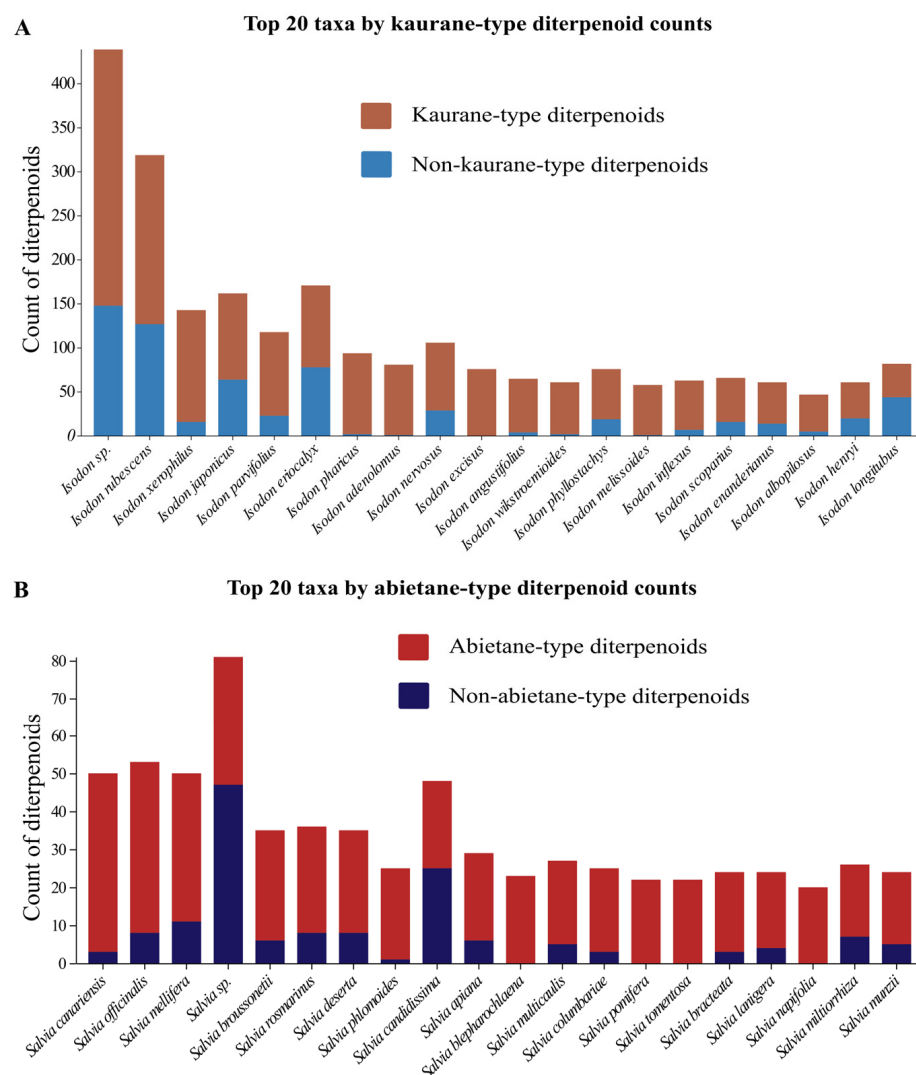


Figure 4. Curated diterpenoid compound–source record counts for the top 20 *Isodon* taxa and the top 20 *Salvia* taxa. **(A)** Kaurane-type and non-kaurane-type diterpenoid record counts for the 20 *Isodon* taxa ranked by kaurane-type diterpenoid record count; detailed data are provided in Supplementary Tables S5 and S6. **(B)** Abietane-type and non-abietane-type diterpenoid record counts for the 20 *Salvia* taxa ranked by abietane-type diterpenoid record count; detailed data are provided in Supplementary Table S7. The stacked bars represent curated compound–source record counts based on currently available database and literature records and should not be interpreted as absolute metabolite abundances in planta.

The central feature of this strategy is therefore not the direct prediction of enzyme function from metabolite distribution. Rather, lineage-biased scaffold-derived diversity provides an additional layer of evidence that can be integrated with comparative gene family evolution and conventional gene-centered analyses to narrow the candidate pool. Enzyme function must ultimately be established through biochemical and/or genetic validation, whereas broader evolutionary relationships require comparative genomic and phylogenetic evidence, appropriately controlled quantitative analyses, and consideration of unequal sampling and research effort.

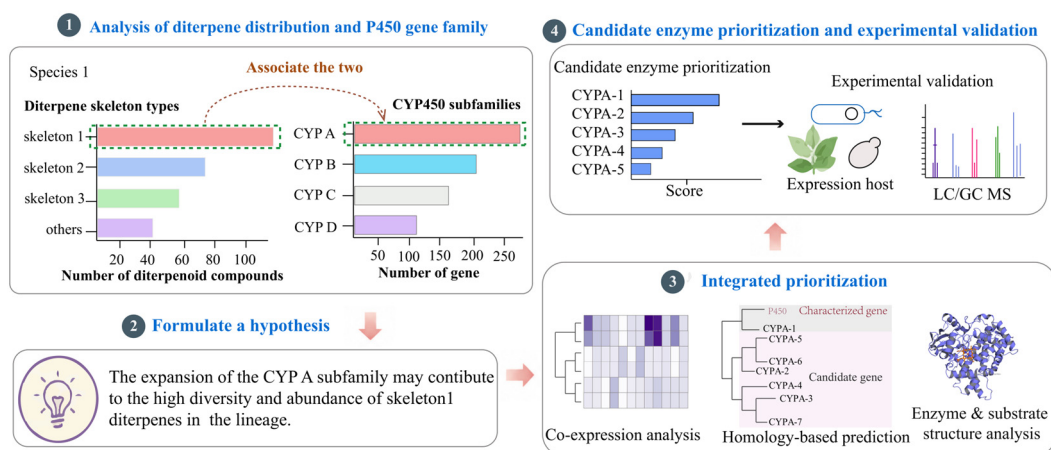


Figure 5. Evidence-guided workflow for prioritizing candidate tailoring enzymes in diterpenoid biosynthesis by integrating lineage-biased scaffold distributions with gene family evolution.

4. Recent Advances in Elucidating Plant-Derived Diterpenoid Biosynthetic Pathways

This section summarizes recent progress in the elucidation of biosynthetic pathways for acyclic, bicyclic, tricyclic, tetracyclic, taxane-type, and macrocyclic diterpenoids based on the ring-system characteristics of their core scaffolds [51–53]. Although the elucidation of plant diterpenoid biosynthesis has long been centered on a limited number of well-studied medicinal pathways, such as tanshinones, forskolin, oridonin, ginkgolides, and paclitaxel [3,9,14,39,41,54], recent studies have increasingly expanded pathway discovery to phylogenetically diverse and less-studied plant lineages, including furanoditerpenoids from *Panicum virgatum* [55], marrubiin from *Marrubium vulgare* [56], clerodane diterpenoids from *Scutellaria* [29], salvinatorin A from *Salvia divinorum* [57], macrocyclic diterpenoids from Euphorbiaceae [58], and diterpenoid alkaloids from *Aconitum* [59]. These examples demonstrate that diterpenoid pathway discovery is no longer limited to a few canonical medicinal compounds. Instead, comparative genomics, transcriptomics, metabolomics, co-expression analysis, and phylogenetic analysis are increasingly revealing pathway enzymes from non-model or less-studied lineages. Rather than listing compounds individually, this section focuses on the common biosynthetic logic of different scaffold classes, the major types of modifying enzymes involved, and the implications of these pathways for the proposed candidate enzyme prioritization framework based on scaffold distribution and gene family expansion.

4.1. Biosynthesis of Acyclic Diterpenoids

Acyclic diterpenoids retain linear carbon-chain structures, and their chemical diversity is mainly generated through terminal oxidative modifications. Multiple plant-derived diTPSs have been shown to convert GGPP into geranylinalool (GL) (Figure 6).

Representative enzymes include GrTPS5 from *Grindelia robusta* [60], NaGLS from *N. attenuate* [61], AtGES from *Arabidopsis thaliana* [62], ZmTPS2 from *Zea mays* [63], and TwGES from *Tripterygium wilfordii* [64]. Downstream modifications in acyclic diterpenoid pathways are typically coordinated by P450s and glycosyltransferases, which together regulate product activity and toxicity. In tobacco, for example, GL undergoes hydroxylation by CYP736A family members, followed by glycosylation catalyzed by the UGT74P family, ultimately producing the defensive metabolites 17-HGL-DTGs [20]. In *Croton stellatopilosus*, *Z. mays*, and *Gossypium hirsutum*, members of the CYP97C [65], CYP92C [63], and CYP82G [66] families have been identified as participating in the biosynthesis of plaunotol and the volatile signaling molecule 4,8,12-trimethyltrideca-1,3,7,11-tetraene (TMTT), a C16-homoterpene. Overall,

modifying enzymes involved in acyclic diterpenoid pathways are frequently associated with plant defense and volatile signal production, providing useful clues for screening candidate P450s and UGTs from tissue-specific or inducible expression datasets.

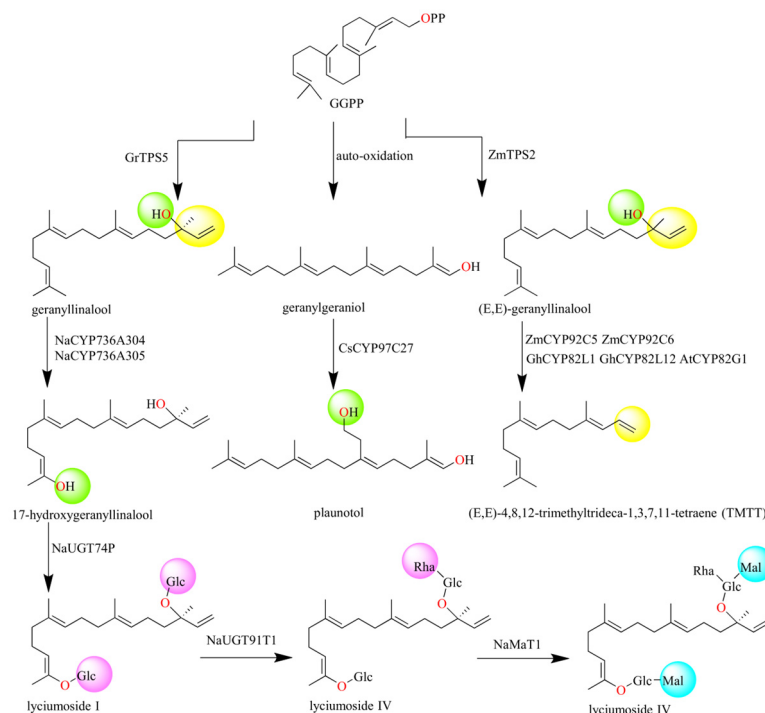


Figure 6. Characterized diTPSs and P450s in the biosynthesis of acyclic diterpenoids. In the structures, Glc, Rha, and Mal denote glucosyl, rhamnosyl, and malonyl groups, respectively. Comprehensive data are available in Supplementary Tables S13–S16.

4.2. Biosynthesis of Bicyclic Diterpenoids

The increased stereochemical complexity of bicyclic diterpenoids provides additional sites for downstream modification. Forskolin and andrographolide are among the best-characterized examples of this scaffold class (Figure 7A).

Forskolin accumulates specifically in root cork cells of *C. forskohlii*. Its biosynthesis begins with scaffold formation by CfTPS2 and CfTPS3 and is followed by coordinated modifications catalyzed by CYP76AH subfamily enzymes and acyltransferases, ultimately yielding the final product [11]. In contrast, andrographolide biosynthesis is initiated by ApCPS1/2-catalyzed conversion of GGPP into *ent*-copalyl diphosphate, after which downstream oxidative networks and UDP-dependent glycosyltransferases contribute to structural maturation (Figure 7B) [12–14]. In addition, several P450 enzymes identified in *Vitex agnus-castus*, *Panicum virgatum*, and *Marrubium vulgare* are available in Supplementary Table S14.

4.3. Biosynthesis of Tricyclic Diterpenoids

Tricyclic diterpenoids mainly include abietane- and pimarane-type scaffolds. Representative products include tanshinones, carnosic acid, triptolide, triptonide, ginkgolides, and rice diterpenoid phytoalexins. These scaffolds are generally constructed by CPS/KSL-type diterpene synthases and then undergo complex modifications, including sequential oxidation, aromatization, lactonization, ring rearrangement, or C–C bond cleavage. They therefore provide important systems for studying functional divergence among P450 enzymes. Among abietane-type diterpenoids, the biosynthetic pathways of tanshinones, carnosic acid, triptolide, and triptonide represent well-characterized examples (Figure 8).

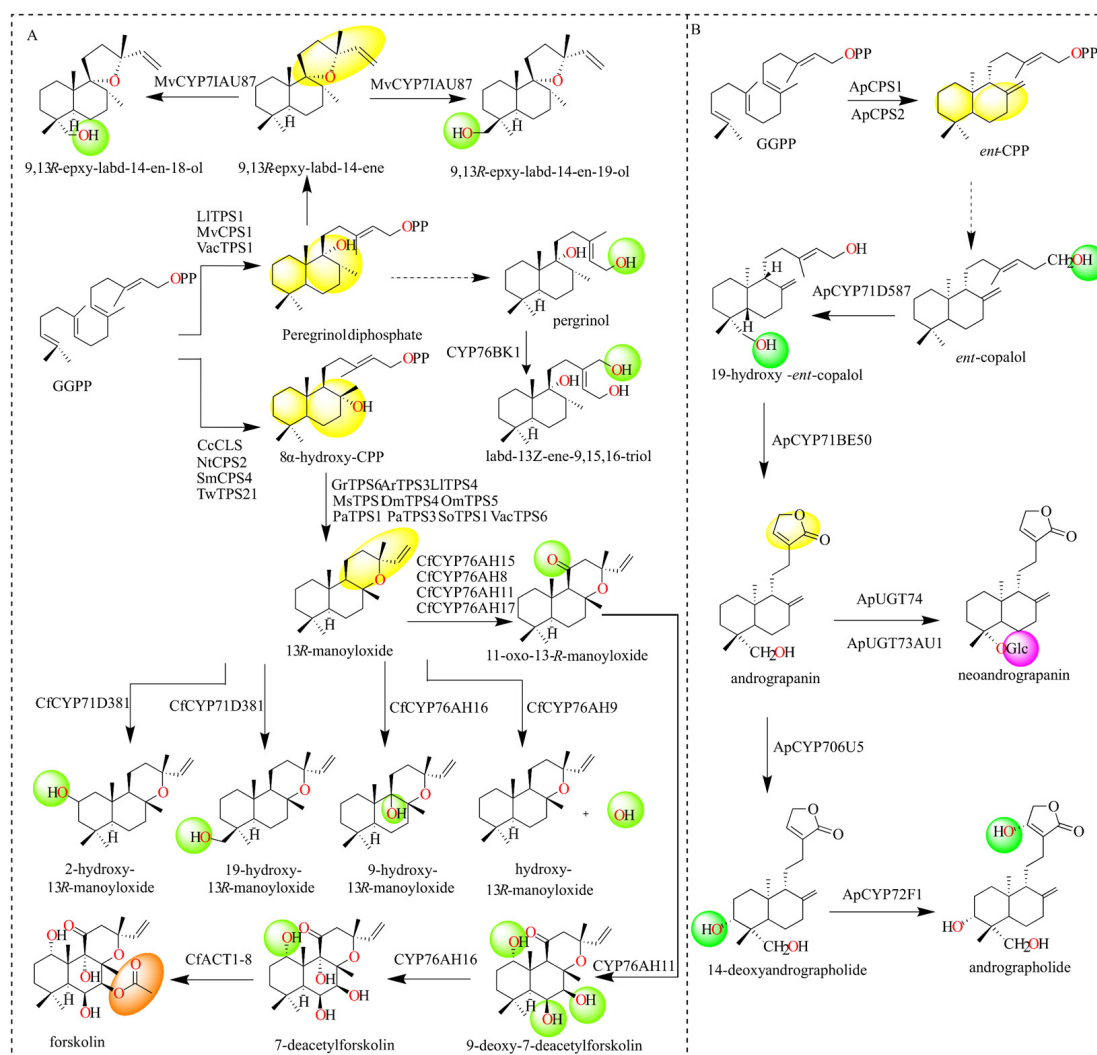


Figure 7. (A) Characterized diTPSs and P450s in the biosynthesis of bicyclic diterpenoids. **(B)** The biosynthetic pathway for andrographolide and the enzymes whose functions have been characterized in the pathway. The “Glc” represents the glucose group. Comprehensive data are available in Supplementary Tables S13–S16.

In *S. miltiorrhiza*, SmCPS1 and SmKSL1 construct the tanshinone diene scaffold, after which enzymes from the CYP76AH [6–8], CYP76AK [9], and CYP71D [10] subfamilies form a complex oxidative network that mediates a series of reactions from aromatization to D-ring lactonization. In the carnosic acid pathway, members of the CYP76AH and CYP76AK subfamilies show clear substrate promiscuity and can convert ferruginol into different products, such as pisiferic acid. Although triptolide and triptonide are also abietane-related diterpenoids, their downstream modification routes have diverged and involve enzymes such as CYP728B70 [67], CYP71BE [68], and CYP82D [69]. Progress in the study of pimarane-type compounds has mainly focused on ginkgolides and phytoalexins (Figure 9).

Forman et al. elucidated key early steps in the biosynthesis of ginkgolides [14]. This pathway is initiated by GbLPS and proceeds through a complete route cooperatively catalyzed by five co-expressed P450 enzymes, including CYP7005C, CYP867E/K, and CYP720B. The pathway involves unusual C–C bond cleavage and sequential hydroxylation, ultimately producing ginkgolides C and D. The formation of rice diterpenoid phytoalexins depends on sequential oxidation catalyzed by lineage-specific P450 systems, including CYP76M [70,71], CYP99A [72], and CYP71Z [73] enzymes. Recent functional characteriza-

tion of KSLX-OL and the CYP71Z family has further refined the metabolic map of defensive diterpenoids in rice [74,75]. Overall, these metabolic networks are also consistent with a biosynthetic logic in which specific P450 subfamilies modify particular scaffolds to expand diterpenoid diversity.

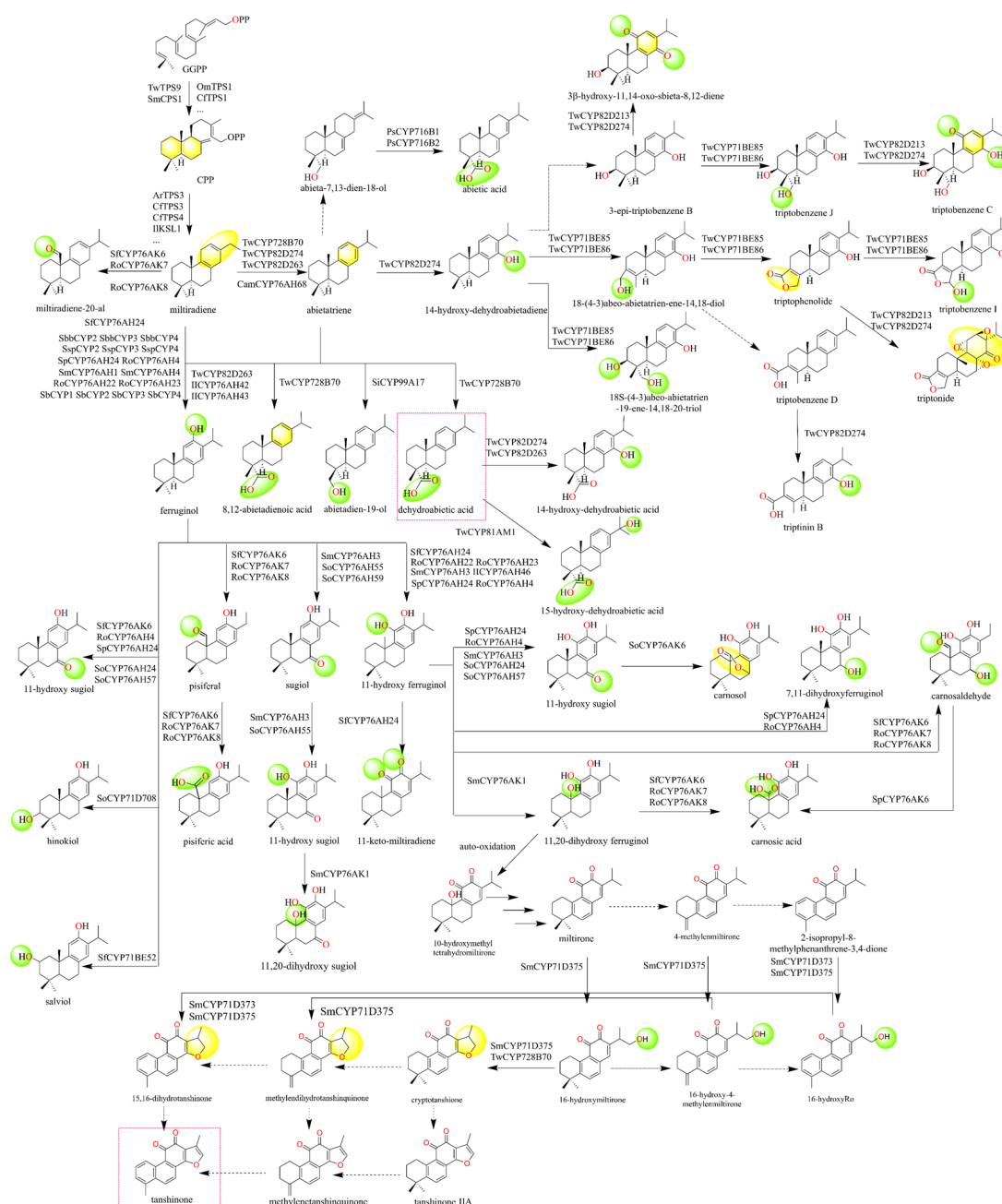


Figure 8. Biosynthesis of abietane-type tricyclic diterpenoids. The red box highlights the target end products of tanshinone and carnosic acid biosynthesis. Solid arrows indicate biochemically characterized steps; dashed arrows indicate uncharacterized steps; and multiple arrows indicate multistep reactions. Comprehensive data are available in Supplementary Tables S13–S16.

4.4. Biosynthesis of Tetracyclic Diterpenoids

Representative tetracyclic diterpene scaffolds include the *ent*-kaurane- and *ent*-atiserane-type scaffolds (Figure 10).

The *ent*-kaurane scaffold serves as a precursor in gibberellin biosynthesis and reflects an evolutionary branch point from primary metabolism toward specialized metabolism [76]. In *Stevia rebaudiana*, intermediates from the gibberellin pathway are recruited and modified

through hydroxylation by KA13H and a series of glycosylation reactions catalyzed by glycosyltransferases such as SrUGT85C2 and SrUGT74G1, leading to the formation of high-intensity sweet steviol glycosides [16].

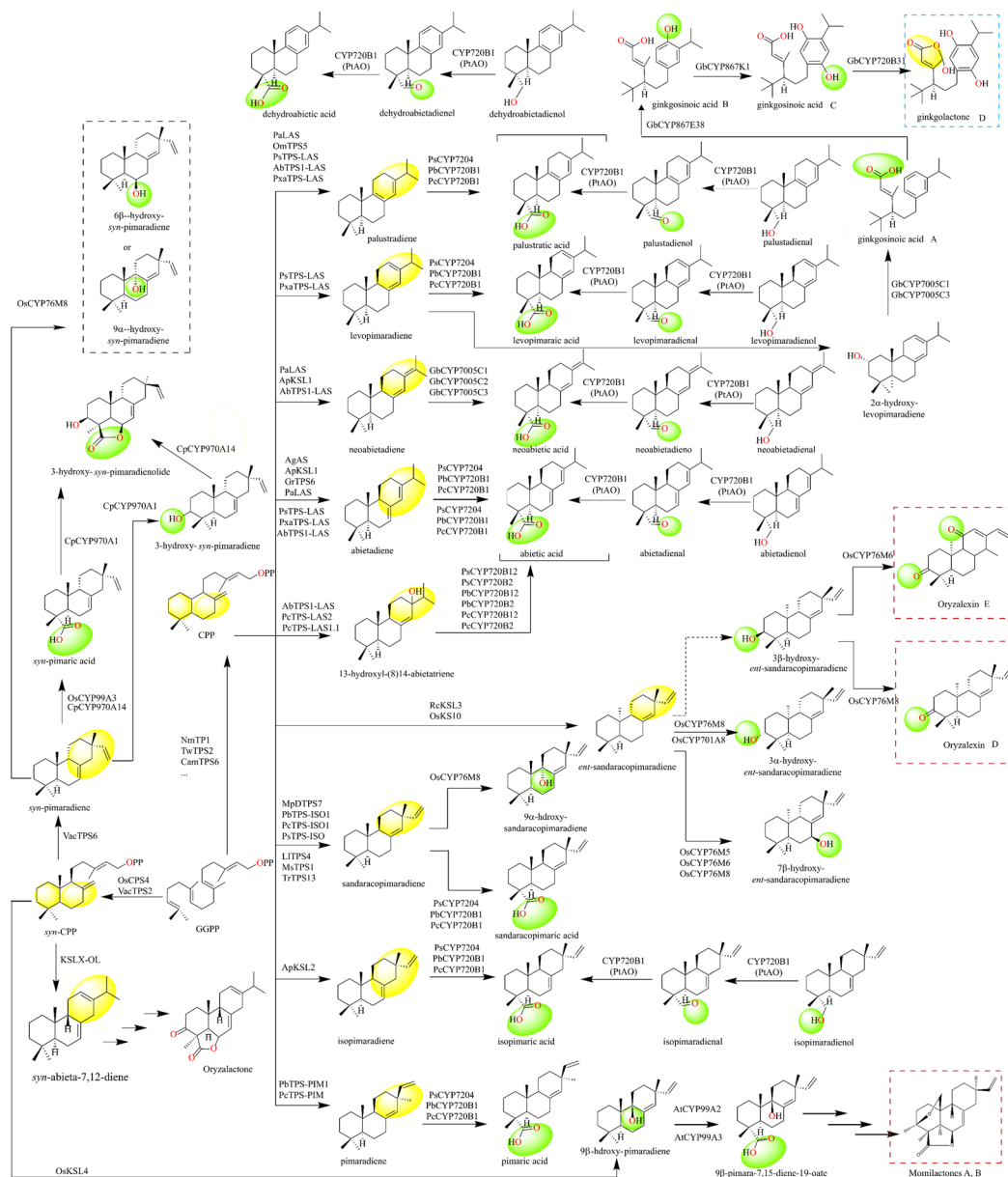


Figure 9. Summary of the biosynthesis of characterized pimarane-type tricyclic diterpenoids. Red boxes indicate end products accessible through the currently reconstructed diterpenoid phytoalexins pathway. Blue boxes indicate end products accessible through the currently reconstructed ginkgolide pathway. Comprehensive data are available in Supplementary Tables S13–S16.

An alternative glycosylation route requiring two additional steps has also been identified in *Rubus suavisissimus* and *Angelica keiskei* [77]. In addition, in *I. rubescens*, *ent*-kaurene is subjected to site-specific oxidation by the lineage-expanded CYP706V subfamily to produce oridonin, which has strong anti-inflammatory activity [41]. This pathway shows a clear evolutionary divergence from the gibberellin biosynthetic route, in which post-scaffold modification is mainly mediated by CYP701. In *I. rubescens*, the kaurane scaffold is generated through the coordinated activity of IrCPS4/5 and IrKSL5, whereas the subsequent key oxidation steps are primarily mediated by the CYP706V P450 subfamily that has specifically expanded in this species [41,54].

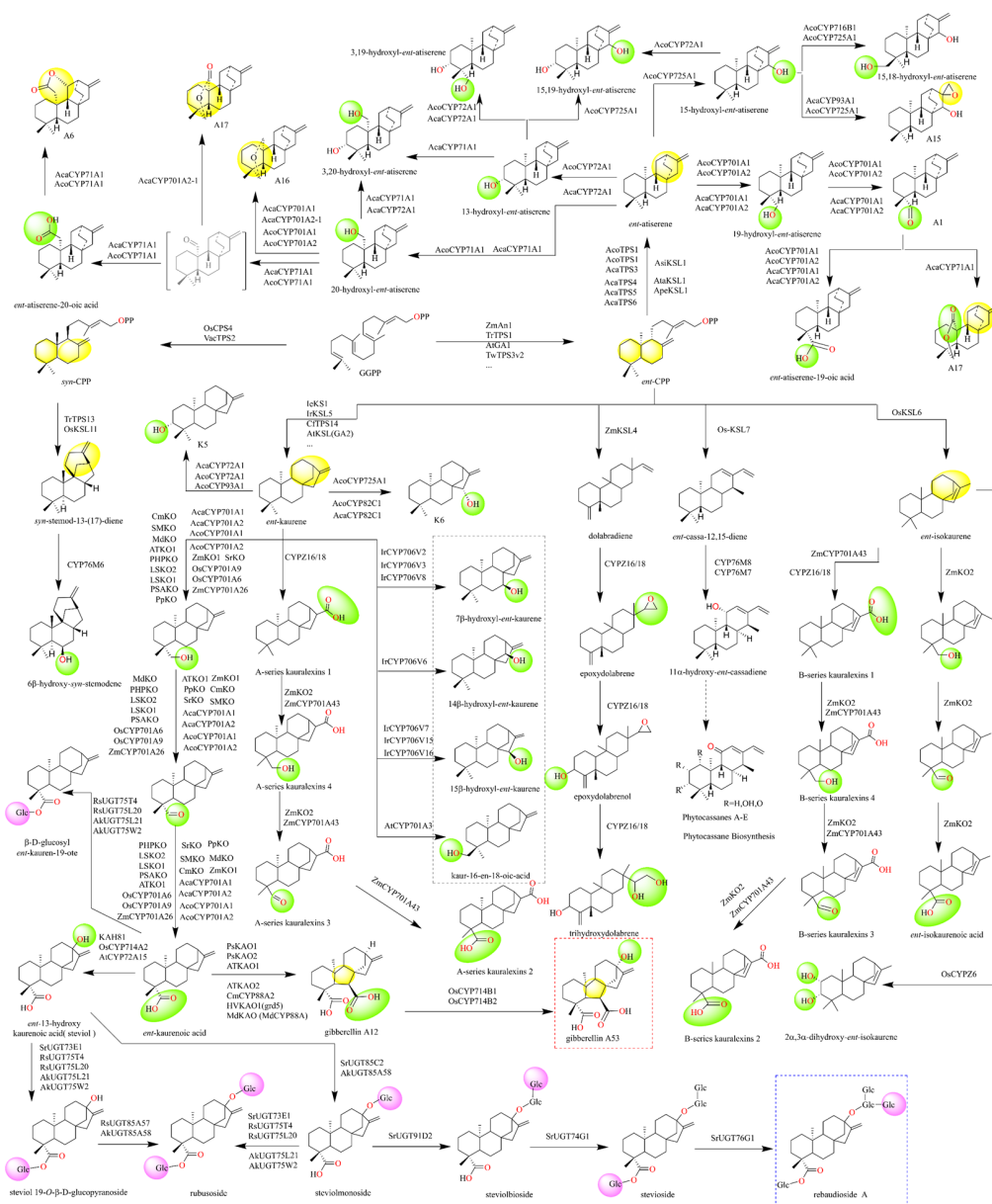


Figure 10. Summary of biosynthesis of characterized tetracyclic diterpenoids. The red box indicates end products accessible through the currently reconstructed gibberellin pathway. Blue boxes indicate end products accessible through the currently reconstructed stevioside pathway. Gray boxes indicate currently synthesizable intermediate products in the biosynthesis pathway of oridonin. Comprehensive data are available in Supplementary Tables S13–S16.

For *ent*-atiserene-type diterpenoids, Luo et al. (2025) made substantial progress in elucidating diterpenoid biosynthesis in plants of the genus *Aconitum* [78]. They identified nine diTPSs and 14 P450 enzymes from *Aconitum carmichaelii* and *Aconitum coreanum*. Using a combinatorial biosynthesis strategy, they successfully constructed tripterifordin, guan-fu diterpenoid A, and 14 new atiserenoid diterpenes in a yeast chassis. This study demonstrates the combinatorial potential of P450 enzymes in multi-site scaffold oxidation.

4.5. Biosynthesis of Taxane-Type Diterpenoids

Taxane-type diterpenoids are best exemplified by paclitaxel, whose characteristic 6/8/6 tricyclic scaffold and highly oxygenated structure have made its biosynthesis a long-standing challenge in natural-product research [79]. The pathway begins with taxadiene synthase (TXS), which converts geranylgeranyl diphosphate into taxa-4(5),11(12)-diene [80].

Following decades of investigation, several recent landmark studies have resolved key steps that were previously unclear (Figure 11). In particular, CYP725A4 was identified as the enzyme responsible for the formation of the characteristic oxetane ring, leading to a revision of earlier models concerning the timing of oxetane formation during paclitaxel biosynthesis [79]. Subsequently, Zhang et al. discovered the previously unidentified C1 hydroxylase (TOT) and C9 oxidase through heterologous pathway reconstruction, which reduced the core enzyme set for baccatin III biosynthesis to nine [3,81]. Elucidation of the taxane pathway illustrates that complete reconstruction of complex diterpenoid pathways often depends on the coordinated integration of multiple P450s, acyltransferases, and heterologous expression systems. Because taxane diterpenoids are strongly enriched within the genus *Taxus*, these pathways also support the strategy of tracing candidate modifying enzyme families from lineages showing scaffold-specific enrichment.

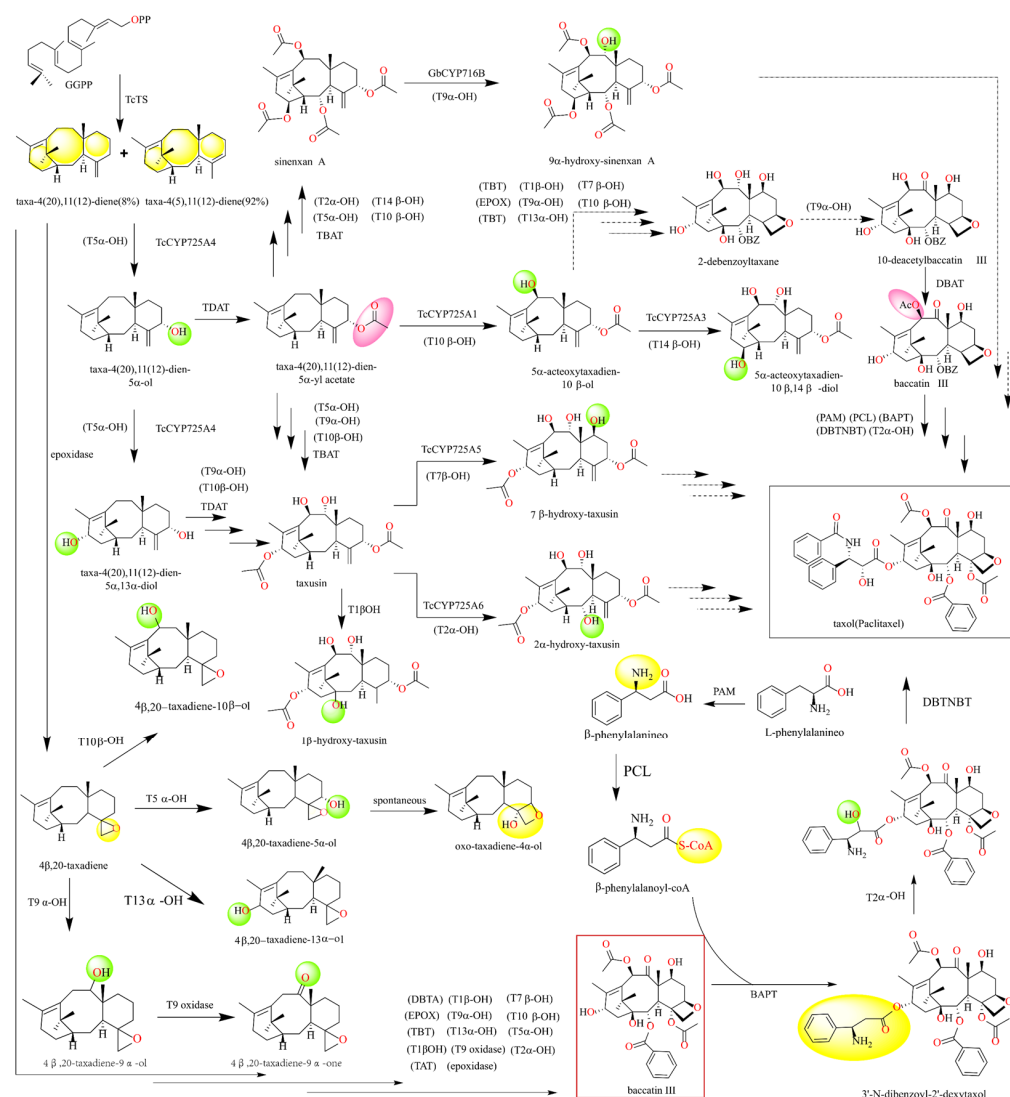


Figure 11. Biosynthesis of paclitaxel. The red box shows baccatin III, an important precursor in the biosynthesis of paclitaxel. The black box shows paclitaxel, the end product. TXS: taxadiene synthase; T5αOH: taxane 5α-hydroxylase; TAT: taxadiene-5α-ol-O-acetyltransferase; T10βOH: taxane 10β-hydroxylase; T13αOH: taxane 13α-hydroxylase; T2αOH: taxane 2α-hydroxylase; T9αOH: taxane 9α-hydroxylase; T7βOH: taxane 7β-hydroxylase; T1βOH: taxane 1β-hydroxylase; TBT: taxane-2α-O-benzoyltransferase; DBAT: 10-deacetyl baccatin III-10-O-acetyltransferase; BAPT: baccatin III-3-amino,13-phenylpropanoyltransferase; DBTNBT: 3'-N-debenzoyl-2'-deoxytaxol-N-benzoyltransferase; PAM: phenylalanine aminomutase; PCL: β-phenylalanine coenzyme A ligase. Comprehensive data are available in Supplementary Tables S13–S16.

5. Conclusions and Future Perspectives

This review integrates the structural diversity, lineage-associated distribution, scaffold types, and oxidative modification features of plant-derived diterpenoids to address a central challenge in diterpenoid biosynthesis: identifying the enzymes responsible for scaffold formation and downstream modification from large candidate enzyme families. From 64,132 diterpenoid-related structures and source records, 45,177 curated plant-derived diterpenoid records were retained, representing 589 canonical C₂₀ scaffolds. Their distribution was highly uneven across plant lineages, with heavily studied taxa contributing a substantial proportion of the available records.

Quantitative analyses demonstrated significant associations between taxonomic lineage and common scaffold distribution at the family, genus, and species-label levels, with Cramér's V values of 0.465, 0.484, and 0.477, respectively. Representative enriched lineage-scaffold relationships included kaurane-type scaffold 2 in *Isodon*, abietane-type scaffold 4 in *Salvia*, scaffold 9 in *Euphorbia*, and scaffold 20 in *Taxus*. These patterns were retained in the more conservative named-species occurrence analysis, supporting lineage-associated organization of currently documented diterpenoid diversity. However, scaffold richness was also strongly correlated with record number, indicating that the observed patterns reflect both biological specialization and variation in sampling intensity, database coverage, and research effort. They should therefore be interpreted as database- and literature-supported overrepresentation patterns rather than estimates of natural abundance or evidence of evolutionary causality.

Within these limitations, integrating scaffold distribution, lineage enrichment, oxidative modification features, and enzyme family evolution provides an evidence-guided framework for prioritizing candidate enzymes, including diTPSs, P450s, acyltransferases, and glycosyltransferases. Broadly enriched scaffolds may direct attention toward conserved or expanded enzyme clades, whereas lineage-restricted oxidative derivatives may highlight recently diversified tailoring enzymes. This framework is intended to complement, rather than replace, gene-cluster mining, co-expression analysis, homology-based prediction, structural modeling, and experimental validation. Accordingly, scaffold enrichment and enzyme family expansion should be treated as a prioritization hypothesis until supported by genomic, transcriptomic, biochemical, and genetic evidence.

Future progress will depend on translating lineage-level hypotheses into experimentally verified enzyme functions and expanding catalytic diversity beyond that sampled by natural evolution. Three directions are particularly important: improving substrate-enzyme mapping, applying generative approaches to enzyme design, and integrating computational discovery, enzyme engineering, and experimental validation into an iterative closed-loop workflow.

5.1. From Finding Similar Sequences to Predicting Substrate-Enzyme Compatibility

For downstream modifying enzymes such as P450s, the key issue is not merely whether a candidate gene belongs to a known enzyme family, but whether its active site can accommodate a specific diterpenoid scaffold, stabilize a particular reaction conformation, and catalyze selective reactions at specific carbon atoms. Computational resources and heme-guided studies—including PCPD [87], PlantP450Dock [88], and the identification of *ent*-kaurene C-14-hydroxylating P450s [89], have demonstrated that heme incorporation, substrate docking, target carbon-heme distances, MD stability, and MM/PBSA decomposition can significantly enhance the interpretability of candidate P450 screening. In addition, general enzyme-substrate models such as ESP [90], EZSpecificity [91], and EnzymeCAGE [92] suggest that future functional prediction of diterpenoid P450s could further integrate protein language models, substrate 3D conformations, active site geometry, and

evolutionary information. This approach advances the predictive goal from “whether an enzyme may act on a class of diterpenoids” to “whether an enzyme can catalyze oxidation at specific carbon atoms of a specific diterpenoid scaffold.”

To achieve more effective substrate–enzyme mapping, it is also necessary to develop diterpenoid representation systems that connect chemical structures with biosynthetic logic. Traditional SMILES, molecular fingerprints, or general structural similarity metrics are suitable for database searches, but they struggle to encode diterpenoid skeleton rearrangements, conserved carbon atom correspondences, oxidation site positions, and possible reaction sequences. Previous studies on triterpenoids have attempted to establish a unified representation of core skeletons and modification sites by standardizing atom and bond position encoding, providing a computational reference framework for terpenoid natural products [93]. Extending this concept to diterpenoids could establish a standardized diterpenoid scaffold coordinate system to encode conserved carbon atoms, ring junctions, migrating bonds, oxidation sites, and glycosylation or acylation positions. If this system is further integrated with known diterpene synthase and modifying enzyme reaction rules, it may allow inference of potential biosynthetic pathways from metabolite structures and provide more mechanistically interpretable inputs for candidate enzyme ranking.

5.2. From Predicting Enzymes to Designing or Engineering Enzymes

Substrate–enzyme mapping can reveal structural determinants of regioselectivity and substrate compatibility, providing useful constraints for P450 engineering and active-site design [94]. Existing cases of diterpenoid P450s provide preliminary support for this design logic. For example, heme-guided screening identified bacterial P450s capable of C-14 hydroxylation of *ent*-kaurane diterpenoids by evaluating the distance between the target carbon and the heme iron [89]. Molecular-dynamics simulations further showed that productive catalysis requires a stable, catalytically favorable substrate orientation rather than a plausible static docking pose alone [89]. Similarly, mutations in CYP76AH1 altered substrate binding and expanded its oxidation profile from mainly C12 oxidation to C11 and C12 hydroxylation and C7 oxidation [95,96]. These findings highlight the importance of considering substrate orientation, pocket plasticity, and persistent near-attack conformations in P450 design.

Existing design cases in non-diterpenoid or general heme proteins also support this approach. Studies of flavonoid P450 pocket design, artificial heme enzymes, and NovoChrome collectively demonstrate that the distal pocket of heme proteins can be remodeled via computational design, ancestral sequence information, key pocket residue constraints, or experimental evolution, thereby altering substrate positioning and reaction selectivity [97–99]. For diterpenoid pathway-related P450s, this perspective, together with enzyme–substrate prediction rules, suggests that mutations should not be guided solely by conserved P450 motifs or overall sequence similarity. Instead, design should focus on the target diterpenoid skeleton, target carbon atoms for oxidation, heme–substrate distances, oxygen transfer angles, pocket hydrophobicity, and electron transfer compatibility.

The design logic for P450s has the potential to address several long-standing challenges in diterpenoid pathway reconstruction, including insufficient regioselectivity, uncontrolled sequential oxidation, difficulty in accumulating unstable intermediates, poor expression of plant P450s in microbial hosts, and incompatibility with heterologous CPRs or redox partners. Multiple cases of diterpenoid P450s have shown that these challenges do not occur in isolation. CYP720B1/PtAO can sequentially oxidize diterpene alcohol and diterpene aldehyde intermediates, demonstrating the multi-substrate and multi-step oxidative capacity of P450s in resin acid biosynthesis [100]. CYP88A can also catalyze consecutive oxidation steps from *ent*-kaurenoic acid to GA12 in gibberellin biosynthesis [101]. Such

sequential oxidative capacity is advantageous for pathway compression, but in heterologous reconstruction it may also lead to excessive consumption of intermediates or make product distributions difficult to control. Taxadiene 5 α -hydroxylase/CYP725A4 further illustrates that diterpenoid P450s, even when responsible for a defined pathway step, may generate multiple by-products, and the target hydroxylated product is not necessarily the major product [102]. Therefore, future design of diterpenoid P450s should aim not only to enhance catalytic activity, but also to restrict undesired binding poses, off-target oxidation sites, and overoxidation routes.

In particular, when a specific plant lineage is known to be enriched in a given diterpenoid skeleton and certain P450 subfamilies are expanded within that lineage, potential reaction templates can first be identified from natural candidate enzymes. Structural modeling, substrate docking, molecular dynamics, and generative pocket design can then be used to propose engineering strategies. In this way, lineage signals, natural enzyme discovery, and generative design can form a continuous framework. Plant evolution provides the candidate space, substrate–enzyme models provide a basis for prioritization, and structural design further expands or optimizes the available enzymatic functions.

In addition, chassis optimization [103–105] and the introduction of non-natural catalytic modules [106–108] will become important complements to diterpenoid pathway engineering. Plant P450s often face limitations in heterologous hosts, including membrane localization, electron transfer, and NADPH supply. Therefore, yeast platforms optimized for redox balance, plant transient expression systems, photosynthetic chassis, and cell-free systems may each play important roles in different contexts. Studies on artificial metalloenzymes and non-natural cofactors further suggest that the binding affinity between protein scaffolds and cofactors can be optimized, enabling abiotic catalysis in complex cytoplasmic environments and allowing the regulation of selectivity between competing reactions [106–108]. Although these non-natural catalytic systems do not directly mimic plant diterpenoid biosynthesis, they provide potential tools for late-stage diversification of diterpenoid skeletons and for transformations that are difficult to achieve using natural enzymes alone.

In the long term, Further work is needed to move plant diterpenoid biosynthesis research from “discovering individual missing enzymes” toward “predicting, designing, and reconstructing entire reaction networks.” Nevertheless, with the refinement of diterpenoid structure databases, the accumulation of enzyme functional data, and advances in protein structure prediction and generative models, research on plant-derived diterpenoid biosynthesis is expected to gradually shift from empirical discovery toward mechanism-driven prediction and design. The evidence-guided framework integrating scaffold enrichment with gene family evolution provides a starting point for such a closed-loop system. It first addresses where diterpenoid chemical innovation occurs during plant evolution and which enzyme families may have participated in this process. Subsequently, substrate–enzyme mapping and structural modeling can further determine which candidate enzymes are likely to catalyze specific reactions on particular diterpenoid skeletons. Finally, generative design and engineered chassis enable researchers to test, optimize, and even reshape these catalytic capabilities.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules31152653/s1>. References [109–219] are included in the Supplementary Materials.

Author Contributions: Conceptualization, Y.Z. and Y.L.; validation, X.H.; investigation, M.L.; data collection, S.Z. and Y.Z.; writing—original draft preparation, Y.Z. and M.L.; writing—review and editing, Y.Z.; visualization, Y.Z.; supervision, M.L. and X.H.; project administration, Y.L.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Doctoral Research Start-up Fund of Zunyi Medical and Pharmaceutical College, grant number zyyz-bsqi (2026) 09.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated and analyzed in this study are provided in the Supplementary Materials.

Acknowledgments: The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

diTPSs	Diterpene synthases
GGPP	Geranylgeranyl diphosphate
P450s	Cytochrome P450 monooxygenases
PSMs	Plant specialized metabolisms
17-HGL-DTGs	17-Hydroxygeranyllinalool diterpene glycosides
MGCs	Metabolic gene clusters
MeJA	Methyl jasmonate
UGTs	UDP-dependent glycosyltransferases
BAHD	BAHD acyltransferases
CPS	Copalyl diphosphate synthase
KSL	Kaurene synthase-like
ABPP	Activity-based protein profiling
GL	Geranyllinalool
Glc	Glucosyl
Rha	Rhamnosyl
Mal	Malonyl
TMTT	4,8,12-Trimethyltrideca-1,3,7,11-tetraene
<i>ent</i> -CPP	<i>ent</i> -Copalyl diphosphate
ADHs	Alcohol dehydrogenases
TXS	Taxadiene synthase
T5 α OH	Taxane 5 α -hydroxylase
TAT	Taxadiene-5 α -ol-O-acetyl transferase
T10 β OH	Taxane 10 β -hydroxylase
T13 α OH	Taxane 13 α -hydroxylase
T2 α OH	Taxane 2 α -hydroxylase
T9 α OH	Taxane 9 α -hydroxylase
T7 β OH	Taxane 7 β -hydroxylase
T1 β OH	Taxane 1 β -hydroxylase
TBT	Taxane-2 α -O-benzoyltransferase
DBAT	10-Deacetyl baccatin III-10-O-acetyltransferase
BAPT	Baccatin III-3-amino,13-phenylpropanoyltransferase
T2' α OH	Taxane 2' α -hydroxylase
DBTNBT	3'-N-Debenzoyl-2'-deoxytaxol-N-benzoyltransferase
PAM	Phenylalanine aminomutase
PCL	β -Phenylalanine coenzyme A ligase
MD	Molecular Dynamics
MM/PBSA	Molecular Mechanics/Poisson–Boltzmann Surface Area

References

- Mathieu, D.T.; Schlecht, N.; van Aalst, M.; Shebek, K.M.; Busta, L.; Babineau, N.; Ebenhöf, O.; Hamberger, B. Rule-based deconstruction and reconstruction of diterpene libraries: Categorizing patterns & unravelling the structural landscape. *bioRxiv* **2024**. [\[CrossRef\]](#)
- Hanson, J.R.; Nichols, T.; Mukhrish, Y.; Bagley, M.C. Diterpenoids of terrestrial origin. *Nat. Prod. Rep.* **2019**, *36*, 1499–1512. [\[CrossRef\]](#) [\[PubMed\]](#)
- Jiang, B.; Gao, L.; Wang, H.; Sun, Y.; Zhang, X.; Ke, H.; Liu, S.; Ma, P.; Liao, Q.; Wang, Y.; et al. Characterization and heterologous reconstitution of *Taxus* biosynthetic enzymes leading to baccatin III. *Science* **2024**, *383*, 622–629. [\[CrossRef\]](#) [\[PubMed\]](#)
- Christianson, D.W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **2017**, *117*, 11570–11648. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mokhosoev, I.M.; Astakhov, D.V.; Terentiev, A.A.; Moldogazieva, N.T. Cytochrome P450 monooxygenase systems: Diversity and plasticity for adaptive stress response. *Prog. Biophys. Mol. Biol.* **2024**, *193*, 19–34. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gao, W.; Hillwig, M.L.; Huang, L.; Cui, G.H.; Wang, X.Y.; Kong, J.Q.; Yang, B.; Peters, R.J. A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. *Org. Lett.* **2009**, *11*, 5170–5173. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guo, J.; Zhou, Y.J.; Hillwig, M.L.; Shen, Y.; Yang, L.; Wang, Y.J.; Zhang, X.A.; Liu, W.J.; Peters, R.J.; Chen, X.Y.; et al. CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12108–12113. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ma, Y.; Ma, X.H.; Meng, F.Y.; Zhan, Z.L.; Guo, J.; Huang, L.Q. RNA interference targeting CYP76AH1 in hairy roots of *Salvia miltiorrhiza* reveals its key role in the biosynthetic pathway of tanshinones. *Biochem. Bioph. Res. Commun.* **2016**, *477*, 155–160. [\[CrossRef\]](#)
- Guo, J.; Ma, X.H.; Cai, Y.; Ma, Y.; Zhan, Z.L.; Zhou, Y.J.; Liu, W.J.; Guan, M.X.; Yang, J.; Cui, G.H.; et al. Cytochrome P450 promiscuity leads to a bifurcating biosynthetic pathway for tanshinones. *New Phytol.* **2016**, *210*, 525–534. [\[PubMed\]](#)
- Ma, Y.; Cui, G.H.; Chen, T.; Ma, X.H.; Wang, R.S.; Jin, B.L.; Yang, J.; Kang, L.P.; Tang, J.F.; Lai, C.J.S.; et al. Expansion within the CYP71D subfamily drives the heterocyclization of tanshinones synthesis in *Salvia miltiorrhiza*. *Nat. Commun.* **2021**, *12*, 685. [\[CrossRef\]](#) [\[PubMed\]](#)
- Salehi, B.; Staniak, M.; Czopek, K.; Stepień, A.; Dua, K.; Wadhwa, R.; Chellappan, D.K.; Sytar, O.; Brestic, M.; Bhat, N.G.; et al. The therapeutic potential of the labdane diterpenoid forskolin. *Appl. Sci.* **2019**, *9*, 4089. [\[CrossRef\]](#)
- Li, Y.; Lin, H.X.; Wang, J.; Yang, J.; Lai, C.J.S.; Wang, X.; Ma, B.W.; Tang, J.F.; Li, Y.; Li, X.L.; et al. Glucosyltransferase capable of catalyzing the last step in neoandrographolide biosynthesis. *Org. Lett.* **2018**, *20*, 5999–6002. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sun, W.; Leng, L.; Yin, Q.G.; Xu, M.M.; Huang, M.K.; Xu, Z.C.; Zhang, Y.J.; Yao, H.; Wang, C.X.; Xiong, C.; et al. The genome of the medicinal plant *Andrographis paniculata* provides insight into the biosynthesis of the bioactive diterpenoid neoandrographolide. *Plant J.* **2019**, *97*, 841–857. [\[CrossRef\]](#) [\[PubMed\]](#)
- Forman, V.; Luo, D.; Geu-Flores, F.; Lemcke, R.; Nelson, D.R.; Kampranis, S.C.; Staerk, D.; Moller, B.L.; Pateraki, I. A gene cluster in *Ginkgo biloba* encodes unique multifunctional cytochrome P450s that initiate ginkgolide biosynthesis. *Nat. Commun.* **2022**, *13*, 5143. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kirby, J.; Nishimoto, M.; Park, J.G.; Withers, S.T.; Nowroozi, F.; Behrendt, D.; Rutledge, E.J.G.; Fortman, J.L.; Johnson, H.E.; Anderson, J.V.; et al. Cloning of casbene and neocembrene synthases from Euphorbiaceae plants and expression in *Saccharomyces cerevisiae*. *Phytochemistry* **2010**, *71*, 1466–1473. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, J.F.; Li, S.Y.; Xiong, Z.Q.; Wang, Y. Pathway mining-based integration of critical enzyme parts for de novo biosynthesis of steviolglycosides sweetener in *Escherichia coli*. *Cell Res.* **2016**, *26*, 258–261. [\[PubMed\]](#)
- Usadel, B.; Obayashi, T.; Mutwil, M.; Giorgi, F.M.; Bassel, G.W.; Tanimoto, M.; Chow, A.; Steinhauser, D.; Persson, S.; Provart, N.J. Co-expression tools for plant biology: Opportunities for hypothesis generation and caveats. *Plant Cell Environ.* **2009**, *32*, 1633–1651. [\[CrossRef\]](#) [\[PubMed\]](#)
- Saito, K.; Hirai, M.Y.; Yonekura-Sakakibara, K. Decoding genes with coexpression networks and metabolomics—‘majority report by precogs’. *Trends Plant Sci.* **2008**, *13*, 36–43. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, Y.Z.; Deng, T.; Sun, L.; Landis, J.B.; Moore, M.J.; Wang, H.C.; Wang, Y.H.; Hao, X.J.; Chen, J.J.; Li, S.H.; et al. Phylogenetic patterns suggest frequent multiple origins of secondary metabolites across the seed-plant ‘tree of life’. *Natl. Sci. Rev.* **2021**, *8*, nwaa105. [\[PubMed\]](#)
- Li, J.C.; Halitschke, R.; Li, D.P.; Paetz, C.; Su, H.C.; Heiling, S.; Xu, S.Q.; Baldwin, I.T. Controlled hydroxylations of diterpenoids allow for plant chemical defense without autotoxicity. *Science* **2021**, *371*, 255–260. [\[CrossRef\]](#) [\[PubMed\]](#)
- Murphy, K.M.; Zerbe, P. Specialized diterpenoid metabolism in monocot crops: Biosynthesis and chemical diversity. *Phytochemistry* **2020**, *172*, 112289. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, Y.; Diao, S.; Ding, X.; Sun, J.; Luan, Q.; Jiang, J. Transcriptional regulation modulates terpenoid biosynthesis of *Pinus elliottii* under drought stress. *Ind. Crops Prod.* **2023**, *202*, 116975. [\[CrossRef\]](#)

23. Jacobowitz, J.R.; Weng, J.K. Exploring uncharted territories of plant specialized metabolism in the postgenomic Era. *Annu. Rev. Plant Biol.* **2020**, *71*, 631–658. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Singh, K.S.; van der Hooft, J.J.J.; van Wees, S.C.M.; Medema, M.H. Integrative omics approaches for biosynthetic pathway discovery in plants. *Nat. Prod. Rep.* **2022**, *39*, 1876–1896. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Osbourn, A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. *Trends Genet.* **2010**, *26*, 449–457. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Nützmann, H.-W.; Doerr, D.; Ramírez-Colmenero, A.; Sotelo-Fonseca, J.E.; Wegel, E.; Di Stefano, M.; Wingett, S.W.; Fraser, P.; Hurst, L.; Fernandez-Valverde, S.L.; et al. Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 13800–13809. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Liang, J.; Shen, Q.Q.; Wang, L.P.; Liu, J.; Fu, J.Y.; Zhao, L.; Xu, M.M.; Peters, R.J.; Wang, Q. Rice contains a biosynthetic gene cluster associated with production of the casbane-type diterpenoid phytoalexin *ent-10-oxodepressin*. *New Phytol.* **2021**, *231*, 85–93. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Del Pup, E.; Owen, C.; Luo, Z.; Augustijn, H.E.; Draisma, A.; Polturak, G.; Kautsar, S.A.; Osbourn, A.; van der Hooft, J.J.J.; Medema, M.H. plantSMASH 2.0: Improvements to detection, annotation, and prioritization of plant biosynthetic gene clusters. *J. Mol. Biol.* **2026**, *438*, 169798. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Li, H.X.; Wu, S.; Lin, R.X.; Xiao, Y.R.; Morotti, A.L.M.; Wang, Y.; Galilee, M.; Qin, H.W.; Huang, T.; Zhao, Y.; et al. The genomes of medicinal skullcaps reveal the polyphyletic origins of clerodane diterpene biosynthesis in the family Lamiaceae. *Mol. Plant* **2023**, *16*, 549–570. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Du, Z.Y.; Peng, Z.Q.; Yang, H.; Wu, H.S.; Sun, J.; Huang, L.L. Identification and functional characterization of three cytochrome P450 genes for the abietane diterpenoid biosynthesis in *Isodon lophanthoides*. *Planta* **2023**, *257*, 90. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Božić, D.; Papaefthimiou, D.; Brückner, K.; de Vos, R.C.H.; Tsoleridis, C.A.; Katsarou, D.; Papanikolaou, A.; Pateraki, I.; Chatzopoulou, F.M.; Dimitriadou, E.; et al. Towards elucidating carnosic acid biosynthesis in Lamiaceae: Functional characterization of the three first steps of the pathway in *Salvia fruticosa* and *Rosmarinus officinalis*. *PLoS ONE* **2015**, *10*, e0124106. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Brückner, K.; Božić, D.; Manzano, D.; Papaefthimiou, D.; Pateraki, I.; Scheler, U.; Ferrer, A.; de Vos, R.C.H.; Kanellis, A.K.; Tissier, A. Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (*Rosmarinus officinalis*) glandular trichomes. *Phytochemistry* **2014**, *101*, 52–64. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Zi, J.C.; Peters, R.J. Characterization of CYP76AH4 clarifies phenolic diterpenoid biosynthesis in the Lamiaceae. *Org. Biomol. Chem.* **2013**, *11*, 7650–7652. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Li, W.; Zhou, Y.; You, W.; Yang, M.; Ma, Y.; Wang, M.; Wang, Y.; Yuan, S.; Xiao, Y. Development of photoaffinity probe for the discovery of steviol glycosides biosynthesis pathway in *Stevia rebaudiana* and rapid substrate screening. *ACS Chem. Biol.* **2018**, *13*, 1944–1949. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Tian, W.; Skolnick, J. How Well is enzyme function conserved as a function of pairwise sequence identity? *J. Mol. Biol.* **2003**, *333*, 863–882. [\[CrossRef\]](#) [\[PubMed\]](#)
36. de Crecy-Lagard, V.; Amorin de Hegedus, R.; Arighi, C.; Babor, J.; Bateman, A.; Blaby, I.; Blaby-Haas, C.; Bridge, A.J.; Burley, S.K.; Cleveland, S.; et al. A roadmap for the functional annotation of protein families: A community perspective. *Database* **2022**, *2022*, baac062. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Tissier, A. Glandular trichomes: What comes after expressed sequence tags? *Plant J.* **2012**, *70*, 51–68. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Cravatt, B.F.; Wright, A.T.; Kozarich, J.W. Activity-based protein profiling: From enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383–414. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Pateraki, I.; Andersen-Ranberg, J.; Jensen, N.B.; Wubshet, S.G.; Heskes, A.M.; Forman, V.; Hallström, B.; Hamberger, B.; Motawia, M.S.; Olsen, C.E.; et al. Total biosynthesis of the cyclic AMP booster forskolin from *Coleus forskohlii*. *eLife* **2017**, *6*, e23001. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Han, D.L.; Li, W.L.; Hou, Z.W.; Lin, C.F.; Xie, Y.; Zhou, X.F.; Gao, Y.; Huang, J.W.; Lai, J.B.; Wang, L.; et al. The chromosome-scale assembly of the *Salvia rosmarinus* genome provides insight into carnosic acid biosynthesis. *Plant J.* **2023**, *113*, 819–832. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Sun, Y.W.; Shao, J.; Liu, H.L.; Wang, H.; Wang, G.Y.; Li, J.H.; Mao, Y.P.; Chen, Z.; Ma, K.; Xu, L.; 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* **2023**, *16*, 517–532. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Weng, J.K. The evolutionary paths towards complexity: A metabolic perspective. *New Phytol.* **2014**, *201*, 1141–1149. [\[PubMed\]](#)
43. Moghe, G.D.; Last, R.L. Something Old, Something New: Conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiol.* **2015**, *169*, 1512–1523. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Torrens-Spence, M.P.; Fallon, T.R.; Weng, J.K. A Workflow for studying specialized metabolism in nonmodel eukaryotic organisms. *Method. Enzymol.* **2016**, *576*, 69–97. [\[CrossRef\]](#)
45. Weng, J.K.; Philippe, R.N.; Noel, J.P. The rise of chemodiversity in plants. *Science* **2012**, *336*, 1667–1670. [\[CrossRef\]](#) [\[PubMed\]](#)

46. Zeng, T.; Chen, Y.X.X.; Jian, Y.X.; Zhang, F.; Wu, R.B. Chemotaxonomic investigation of plant terpenoids with an established database (TeroMOL). *New Phytol.* **2022**, *235*, 662–673. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Barbosa Silva Cavalcanti, A.; Costa Barros, R.P.; Costa, V.C.d.O.; Sobral da Silva, M.; Fechine Tavares, J.; Scotti, L.; Scotti, M.T. Computer-aided chemotaxonomy and bioprospecting study of diterpenes of the lamiaceae family. *Molecules* **2019**, *24*, 3908. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Tundis, R.; Peruzzi, L.; Menichini, F. Phytochemical and biological studies of *Stachys* species in relation to chemotaxonomy: A review. *Phytochemistry* **2014**, *102*, 7–39. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Piozzi, F.; Bruno, M.; Rosselli, S.; Maggio, A. The diterpenoids from the genus *Hyptis* (Lamiaceae). *Heterocycles* **2009**, *78*, 1413–1426. [\[CrossRef\]](#)
50. Hu, J.; Qiu, S.; Wang, F.; Li, Q.; Xiang, C.L.; Di, P.; Wu, Z.; Jiang, R.; Li, J.; Zeng, Z.; et al. Functional divergence of CYP76AKs shapes the chemodiversity of abietane-type diterpenoids in genus *Salvia*. *Nat. Commun.* **2023**, *14*, 4696. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Zerbe, P.; Bohlmann, J. Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.* **2015**, *33*, 419–428. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Mafu, S.; Zerbe, P. Plant diterpenoid metabolism for manufacturing the biopharmaceuticals of tomorrow: Prospects and challenges. *Phytochem. Rev.* **2018**, *17*, 113–130.
53. Peters, R.J. Two rings in them all: The labdane-related diterpenoids. *Nat. Prod. Rep.* **2010**, *27*, 1521–1530. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Jin, B.L.; Cui, G.H.; Guo, J.; Tang, J.F.; Duan, L.X.; Lin, H.X.; Shen, Y.; Chen, T.; Zhang, H.B.; Huang, L.Q. Functional diversification of kaurene synthase-like genes in *Isodon rubescens*. *Plant Physiol.* **2017**, *174*, 943–955. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Muchlinski, A.; Jia, M.R.; Tiedge, K.; Fell, J.S.; Pelot, K.A.; Chew, L.; Davisson, D.; Chen, Y.X.; Siegel, J.; Lovell, J.T.; et al. Cytochrome P450-catalyzed biosynthesis of furanoditerpenoids in the bioenergy crop switchgrass (*Panicum virgatum* L.). *Plant J.* **2021**, *108*, 1053–1068. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Zerbe, P.; Chiang, A.; Dullat, H.; O’Neil-Johnson, M.; Starks, C.; Hamberger, B.; Bohlmann, J. Diterpene synthases of the biosynthetic system of medicinally active diterpenoids in *Marrubium vulgare*. *Plant J.* **2014**, *79*, 914–927. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Li, H.; Sun, Y.; Xu, W.; Sun, B.; Wu, S.; Li, C.; Zhao, J.; Martin, C.; Tatsis, E.C. Comparative genome analysis provides a foundation for defining salvinorin a biosynthesis in *Salvia divinorum*. *Nat. Commun.* **2026**, *17*, 3414. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Czechowski, T.; Forestier, E.; Swamidatta, S.H.; Gilday, A.D.; Cording, A.; Larson, T.R.; Harvey, D.; Li, Y.; He, Z.; King, A.J.; et al. Gene discovery and virus-induced gene silencing reveal branched pathways to major classes of bioactive diterpenoids in *Euphorbia peplus*. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2203890119. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Mao, L.; Jin, B.; Chen, L.; Tian, M.; Ma, R.; Yin, B.; Zhang, H.; Guo, J.; Tang, J.; Chen, T.; et al. Functional identification of the terpene synthase family involved in diterpenoid alkaloids biosynthesis in *Aconitum carmichaelii*. *Acta Pharm. Sin. B* **2021**, *11*, 3310–3321. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Zerbe, P.; Chiang, A.; Yuen, M.; Hamberger, B.R.; Hamberger, B.; Draper, J.A.; Britton, R.; Bohlmann, J. Bifunctional *cis*-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *J. Biol. Chem.* **2012**, *287*, 12121–12131. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Falara, V.; Alba, J.M.; Kant, M.R.; Schuurink, R.C.; Pichersky, E. Geranylinalool synthases in solanaceae and other angiosperms constitute an ancient branch of diterpene synthases involved in the synthesis of defensive compounds. *Plant Physiol.* **2014**, *166*, 428–441. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Morrone, D.; Chen, X.M.; Coates, R.M.; Peters, R.J. Characterization of the kaurene oxidase CYP701A3, a multifunctional cytochrome P450 from gibberellin biosynthesis. *Biochem. J.* **2010**, *431*, 337–344. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Richter, A.; Schaff, C.; Zhang, Z.W.; Lipka, A.E.; Tian, F.; Köllner, T.G.; Schnee, C.; Preiss, S.; Irmisch, S.; Jander, G.; et al. Characterization of biosynthetic pathways for the production of the volatile homoterpenes DMNT and TMTT in *Zea mays*. *Plant Cell* **2016**, *28*, 2651–2665. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Su, P.; Hu, T.; Liu, Y.; Tong, Y.; Guan, H.; Zhang, Y.; Zhou, J.; Huang, L.; Gao, W. Functional characterization of NES and GES responsible for the biosynthesis of (E)-nerolidol and (E,E)-geranylinalool in *Tripterygium wilfordii*. *Sci. Rep.* **2017**, *7*, 40851. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Sintupachee, S.; Ngamrojanayanich, N.; Sitthithaworn, W.; De-Eknamkul, W. Molecular cloning, bacterial expression and functional characterisation of cytochrome P450 monooxygenase, CYP97C27, and NADPH-cytochrome P450 reductase, CPR I, from *Croton stellatopilosus* Ohba. *Plant Sci.* **2014**, *229*, 131–141. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Lee, S.; Badieyan, S.; Bevan, D.R.; Herde, M.; Gatz, C.; Tholl, D. Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21205–21210. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Tu, L.C.; Su, P.; Zhang, Z.R.; Gao, L.H.; Wang, J.D.; Hu, T.Y.; Zhou, J.W.; Zhang, Y.F.; Zhao, Y.J.; Liu, Y.; et al. Genome of *Tripterygium wilfordii* and identification of cytochrome P450 involved in triptolide biosynthesis. *Nat. Commun.* **2020**, *11*, 971. [\[CrossRef\]](#) [\[PubMed\]](#)

68. Hansen, N.L.; Kjaerulff, L.; Heck, Q.K.; Forman, V.; Staerk, D.; Moller, B.L.; Andersen-Ranberg, J. cytochrome P450s catalyze the methyl shift and epoxidations in the biosynthesis of triptonide. *Nat. Commun.* **2022**, *13*, 5011. [[CrossRef](#)] [[PubMed](#)]
69. Zhang, Y.; Gao, J.; Ma, L.; Tu, L.; Hu, T.; Wu, X.; Su, P.; Zhao, Y.; Liu, Y.; Li, D.; et al. Tandemly duplicated CYP82Ds catalyze 14-hydroxylation in triptolide biosynthesis and precursor production in *Saccharomyces cerevisiae*. *Nat. Commun.* **2023**, *14*, 875. [[CrossRef](#)] [[PubMed](#)]
70. Wu, Y.; Wang, Q.; Hillwig, M.L.; Peters, R.J. Picking sides: Distinct roles for CYP76M6 and CYP76M8 in rice oryzalexin biosynthesis. *Biochem. J.* **2013**, *454*, 209–216. [[CrossRef](#)] [[PubMed](#)]
71. Swaminathan, S.; Morrone, D.; Wang, Q.; Fulton, D.B.; Peters, R.J. CYP76M7 Is an *ent*-Cassadiene C11 α -Hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell* **2009**, *21*, 3315–3325. [[CrossRef](#)] [[PubMed](#)]
72. Wang, Q.A.; Hillwig, M.L.; Peters, R.J. CYP99A3: Functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice. *Plant J.* **2011**, *65*, 87–95. [[PubMed](#)]
73. Kitaoka, N.; Wu, Y.S.; Xu, M.M.; Peters, R.J. Optimization of recombinant expression enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 7549–7558. [[CrossRef](#)] [[PubMed](#)]
74. Kariya, K.; Mori, H.; Ueno, M.; Yoshikawa, T.; Teraishi, M.; Yabuta, Y.; Ueno, K.; Ishihara, A. Identification and evolution of a diterpenoid phytoalexin oryzalactone biosynthetic gene in the genus *Oryza*. *Plant J.* **2024**, *118*, 358–372. [[CrossRef](#)] [[PubMed](#)]
75. Kitaoka, N.; Zhang, J.; Oyagbenro, R.K.; Brown, B.; Wu, Y.; Yang, B.; Li, Z.; Peters, R.J. Interdependent evolution of biosynthetic gene clusters for momilactone production in rice. *Plant Cell* **2021**, *33*, 290–305. [[PubMed](#)]
76. Hedden, P.; Thomas, S.G. Gibberellin biosynthesis and its regulation. *Biochem. J.* **2012**, *444*, 11–25. [[CrossRef](#)] [[PubMed](#)]
77. Sun, Y.W.; Chen, Z.; Li, J.X.; Li, J.H.; Lv, H.J.; Yang, J.Y.; Li, W.W.; Xie, D.G.; Xiong, Z.Q.; Zhang, P.; et al. Diterpenoid UDP-Glycosyltransferases from Chinese Sweet Tea and *Ashitaba* complete the biosynthesis of rubusoside. *Mol. Plant* **2018**, *11*, 1308–1311. [[CrossRef](#)] [[PubMed](#)]
78. Luo, F.; Zhou, Q.; Chen, F.; Liu, X.; Chiu, T.Y.; Zhu, G.Y.; Huang, A.C. Divergent multifunctional P450s-empowered biosynthesis of bioactive tripterifordin and cryptic atiserenoids in *Aconitum* implies convergent evolution. *Nat. Commun.* **2025**, *16*, 5857. [[CrossRef](#)] [[PubMed](#)]
79. Zhao, Y.; Liang, F.Y.; Xie, Y.M.; Duan, Y.T.; Andeadelli, A.; Pateraki, I.; Makris, A.M.; Pomorski, T.G.; Staerk, D.; Kampranis, S.C. Oxetane ring formation in taxol biosynthesis is catalyzed by a bifunctional Cytochrome P450 enzyme. *J. Am. Chem. Soc.* **2024**, *146*, 801–810. [[PubMed](#)]
80. Koksai, M.; Jin, Y.H.; Coates, R.M.; Croteau, R.; Christianson, D.W. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* **2011**, *469*, 116–120. [[PubMed](#)]
81. Zhang, Y.; Wiese, L.; Fang, H.; Alseekh, S.; Perez de Souza, L.; Scossa, F.; Molloy, J.J.; Christmann, M.; Fernie, A.R. Synthetic biology identifies the minimal gene set required for paclitaxel biosynthesis in a plant chassis. *Mol. Plant* **2023**, *16*, 1951–1961. [[CrossRef](#)] [[PubMed](#)]
82. Wang, H.B.; Wang, X.Y.; Liu, L.P.; Qin, G.W.; Kang, T.G. Tiglane diterpenoids from the Euphorbiaceae and Thymelaeaceae families. *Chem. Rev.* **2015**, *115*, 2975–3011. [[CrossRef](#)] [[PubMed](#)]
83. Vasas, A.; Hohmann, J. Diterpenes: Isolation, structure, biological activity, and synthesis (2008–2012). *Chem. Rev.* **2014**, *114*, 8579–8612. [[CrossRef](#)] [[PubMed](#)]
84. Luo, D.; Callari, R.; Hamberger, B.; Wubshet, S.G.; Nielsen, M.T.; Andersen-Ranberg, J.; Hallström, B.M.; Cozzi, F.; Heider, H.; Moller, B.L.; 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* **2016**, *113*, E5082–E5089. [[CrossRef](#)] [[PubMed](#)]
85. Schotte, C.; Florean, M.; Czechowski, T.; Gilday, A.; Alam, R.M.; Ploss, K.; Wurlitzer, J.; Li, Y.; Sonawane, P.; Graham, I.A.; et al. Identification of BAHD-acyltransferase enzymes involved in ingenane diterpenoid biosynthesis. *New Phytol.* **2025**, *247*, 2591–2600. [[CrossRef](#)] [[PubMed](#)]
86. Zhao, W.; Wang, F.; Li, P.; Li, L.; Ernst, L.; Huang, L.; Tian, M.; Lv, W.; Xu, S.; Liu, F.; et al. Two O-acyltransferases from the diterpene biosynthetic gene cluster of *Euphorbia lathyris* contribute to the structural diversity of medicinal macrocyclic diterpenoid esters biosynthesis. *Plant J.* **2025**, *121*, e70003. [[CrossRef](#)] [[PubMed](#)]
87. Wang, H.; Wang, Q.; Liu, Y.Q.; Liao, X.P.; Chu, H.Y.; Chang, H.; Cao, Y.; Li, Z.G.; Zhang, T.C.; Cheng, J.; et al. PCPD: Plant cytochrome P450 database and web-based tools for structural construction and ligand docking. *Syn. Syst. Biotechnol.* **2021**, *6*, 102–109. [[CrossRef](#)]
88. Feng, L.; Niu, C.; Qing, X.; Zhang, C.; Li, C. PlantP450Dock: An automated molecular docking pipeline of plant cytochrome P450s. *bioRxiv* **2026**. [[CrossRef](#)]
89. Lin, X.; Xiao, Z.; Xu, X.; Zhang, X.; Pan, X.; Zhu, C.; He, C.; Feng, L.; Li, F.R.; Xu, H.M.; et al. Discovery and engineering of bacterial P450s for C-14 hydroxylation in *ent*-kaurene diterpenoids. *Nat. Commun.* **2026**, *17*, 3850. [[CrossRef](#)] [[PubMed](#)]
90. Kroll, A.; Ranjan, S.; Engqvist, M.K.M.; Lercher, M.J. A general model to predict small molecule substrates of enzymes based on machine and deep learning. *Nat. Commun.* **2023**, *14*, 2787. [[CrossRef](#)] [[PubMed](#)]

91. Cui, H.; Su, Y.; Dean, T.J.; Yu, T.; Zhang, Z.; Peng, J.; Shukla, D.; Zhao, H. Enzyme specificity prediction using cross-attention graph neural networks. *Nature* **2025**, *647*, 639–647. [[CrossRef](#)] [[PubMed](#)]
92. Liu, Y.; Hua, C.; Xu, M.; Zeng, T.; Rao, J.; Zhang, Z.; Wu, R.; Weng, J.-K.; Coley, C.W.; Zheng, S. A geometric foundation model for enzyme retrieval with evolutionary insights. *Nat. Catal.* **2026**, *9*, 148–160. [[CrossRef](#)]
93. Babineau, N.; Dien Nguyen, L.T.; Mathieu, D.; McCue, C.; Schlecht, N.; Abrahamson, T.; Hamberger, B.; Busta, L. A molecular representation system with a common reference frame for analyzing triterpenoid structural diversity. *Plant Commun.* **2025**, *6*, 101320. [[CrossRef](#)] [[PubMed](#)]
94. Bathe, U.; Tissier, A. Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry* **2019**, *161*, 149–162. [[CrossRef](#)] [[PubMed](#)]
95. Qiu, Y.; Diao, H.; Zheng, Y.; Wu, R. Multiscale simulations on the catalytic plasticity of CYP76AH1. *Front. Chem.* **2021**, *9*, 689731. [[CrossRef](#)] [[PubMed](#)]
96. Mao, Y.; Ma, Y.; Chen, T.; Ma, X.; Xu, Y.; Bu, J.; Li, Q.; Jin, B.; Wang, Y.; Li, Y.; et al. Functional integration of two CYP450 genes involved in biosynthesis of tanshinones for improved diterpenoid production by synthetic biology. *ACS Synth. Biol.* **2020**, *9*, 1763–1770. [[CrossRef](#)] [[PubMed](#)]
97. Wang, Q.; Liu, X.; Zhang, H.; Chu, H.; Shi, C.; Zhang, L.; Bai, J.; Liu, P.; Li, J.; Zhu, X.; et al. Cytochrome P450 enzyme design by constraining the catalytic pocket in a diffusion model. *Research* **2024**, *7*, 0413. [[CrossRef](#)] [[PubMed](#)]
98. Kalvet, I.; Ortmayer, M.; Zhao, J.; Crawshaw, R.; Ennist, N.M.; Levy, C.; Roy, A.; Green, A.P.; Baker, D. Design of heme enzymes with a tunable substrate binding pocket adjacent to an open metal coordination site. *J. Am. Chem. Soc.* **2023**, *145*, 14307–14315. [[CrossRef](#)] [[PubMed](#)]
99. Hou, K.; Huang, W.; Qi, M.; Tugwell, T.H.; Alturaifi, T.M.; Chen, Y.; Zhang, X.; Lu, L.; Mann, S.I.; Liu, P.; et al. *De novo* design of porphyrin-containing proteins as efficient and stereoselective catalysts. *Science* **2025**, *388*, 665–670. [[CrossRef](#)] [[PubMed](#)]
100. Ro, D.K.; Bohlmann, J. Diterpene resin acid biosynthesis in loblolly pine (*Pinus taeda*): Functional characterization of abietadiene/levopimaradiene synthase (PtTPS-LAS) cDNA and subcellular targeting of PtTPS-LAS and abietadienol/abietadienal oxidase (PtAO, CYP720B1). *Phytochemistry* **2006**, *67*, 1572–1578. [[CrossRef](#)] [[PubMed](#)]
101. Helliwell, C.A.; Chandler, P.M.; Poole, A.; Dennis, E.S.; Peacock, W.J. The CYP88A cytochrome P450, *ent*-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2065–2070. [[CrossRef](#)] [[PubMed](#)]
102. Jennewein, S.; Long, R.M.; Williams, R.M.; Croteau, R. Cytochrome P450 taxadiene 5 α -hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. *Chem. Biol.* **2004**, *11*, 379–387. [[CrossRef](#)] [[PubMed](#)]
103. Xiao, H.; Zhang, Y.; Wang, M. Discovery and engineering of Cytochrome P450s for terpenoid biosynthesis. *Trends Biotechnol.* **2019**, *37*, 618–631. [[CrossRef](#)] [[PubMed](#)]
104. Reed, J.; Osbourn, A. Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant Cell Rep.* **2018**, *37*, 1431–1441. [[CrossRef](#)] [[PubMed](#)]
105. Hong, B.; Qiu, L.; Lv, R.; Yu, Z. The sustainable production of terpenoids in cyanobacterial chassis. *Microorganisms* **2025**, *13*, 1342. [[CrossRef](#)] [[PubMed](#)]
106. Zou, Z.; Kalvet, I.; Lozhkin, B.; Morris, E.; Zhang, K.; Chen, D.; Ernst, M.L.; Zhang, X.; Baker, D.; Ward, T.R. *De novo* design and evolution of an artificial metathase for cytoplasmic olefin metathesis. *Nat. Catal.* **2025**, *8*, 1208–1219. [[CrossRef](#)] [[PubMed](#)]
107. Radley, E.J.; Andrews, A.C.; Kalvet, I.; Deng, Y.; Bell, E.L.; Levy, C.W.; Ortmayer, M.; Heyes, D.J.; Megarity, C.F.; Nunez-Franco, R.; et al. A *de novo* CO₂ reductase featuring a cysteine-ligated cobalt porphyrin cofactor. *J. Am. Chem. Soc.* **2026**, *148*, 30194–30201. [[CrossRef](#)] [[PubMed](#)]
108. Huang, J.; Liu, Z.; Bloomer, B.J.; Clark, D.S.; Mukhopadhyay, A.; Keasling, J.D.; Hartwig, J.F. Unnatural biosynthesis by an engineered microorganism with heterologously expressed natural enzymes and an artificial metalloenzyme. *Nat. Chem.* **2021**, *13*, 1186–1191. [[CrossRef](#)] [[PubMed](#)]
109. Zhou, K.; Peters, R.J. Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants. *Phytochemistry* **2009**, *70*, 366–369. [[CrossRef](#)] [[PubMed](#)]
110. Tian, M.; Jin, B.; Chen, L.; Ma, R.; Ma, Q.; Li, X.; Chen, T.; Guo, J.; Ge, H.; Zhao, X.; et al. Functional diversity of diterpene synthases in *Aconitum* plants. *Plant Physiol. Biochem.* **2023**, *202*, 107968. [[CrossRef](#)] [[PubMed](#)]
111. Johnson, S.R.; Bhat, W.W.; Bibik, J.; Turmo, A.; Hamberger, B.R.; Hamberger, B. A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *J. Biol. Chem.* **2019**, *294*, 1349–1362. [[CrossRef](#)] [[PubMed](#)]
112. Misra, R.C.; Garg, A.; Roy, S.; Chanotiya, C.S.; Vasudev, P.G.; Ghosh, S. Involvement of an *ent*-copalyl diphosphate synthase in tissue-specific accumulation of specialized diterpenes in *Andrographis paniculata*. *Plant Sci.* **2015**, *240*, 50–64. [[CrossRef](#)] [[PubMed](#)]
113. Sun, T.P.; Kamiya, Y. The Arabidopsis GA1 locus encodes the cyclase *ent*-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **1994**, *6*, 1509–1518. [[CrossRef](#)] [[PubMed](#)]
114. Yamaguchi, S.; Sun, T.; Kawaide, H.; Kamiya, Y. The GA2 locus of *Arabidopsis thaliana* encodes *ent*-kaurene synthase of gibberellin biosynthesis. *Plant Physiol.* **1998**, *116*, 1271–1278. [[CrossRef](#)] [[PubMed](#)]

115. Jia, Q.; Brown, R.; Köllner, T.G.; Fu, J.; Chen, X.; Wong, G.K.; Gershenzon, J.; Peters, R.J.; Chen, F. Origin and early evolution of the plant terpene synthase family. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2100361119. [[CrossRef](#)] [[PubMed](#)]
116. Shimada, T.; Minato, S.; Hasegawa, Y.; Miyamoto, K.; Minato, Y.; Shenton, M.R.; Okada, K.; Kawaide, H.; Toyomasu, T. Characterization of diterpene synthase genes in *Brachypodium distachyon*, a monocotyledonous model plant, provides evolutionary insight into their multiple homologs in cereals. *Biosci. Biotechnol. Biochem.* **2023**, *88*, 8–15. [[CrossRef](#)] [[PubMed](#)]
117. Bryson, A.E.; Lanier, E.R.; Lau, K.H.; Hamilton, J.P.; Vaillancourt, B.; Mathieu, D.; Yocca, A.E.; Miller, G.P.; Edger, P.P.; Buell, C.R.; et al. Uncovering a multiradiene biosynthetic gene cluster in the Lamiaceae reveals a dynamic evolutionary trajectory. *Nat. Commun.* **2023**, *14*, 343. [[CrossRef](#)] [[PubMed](#)]
118. Li, C.; Wang, S.; Yin, X.; Guo, A.; Xie, K.; Chen, D.; Sui, S.; Han, Y.; Liu, J.; Chen, R.; et al. Functional characterization and cyclization mechanism of a diterpene synthase catalyzing the skeleton formation of Cephalotane-Type diterpenoids. *Angew. Chem. Int. Ed.* **2023**, *62*, e202306020. [[CrossRef](#)]
119. Wu, T.J.; Lin, C.C.; Ma, L.T.; Yang, C.K.; Ho, C.L.; Wang, S.Y.; Chu, F.H. Functional identification of specialized diterpene synthases from *Chamaecyparis obtusa* and *C. obtusa* var. *formosana* to illustrate the putative evolution of diterpene synthases in Cupressaceae. *Plant Sci.* **2024**, *344*, 112080. [[CrossRef](#)] [[PubMed](#)]
120. Ma, L.T.; Wang, C.H.; Hon, C.Y.; Lee, Y.R.; Chu, F.H. Discovery and characterization of diterpene synthases in *Chamaecyparis formosensis* Matsum. which participated in an unprecedented diterpenoid biosynthesis route in conife. *Plant Sci.* **2021**, *304*, 110790. [[CrossRef](#)] [[PubMed](#)]
121. Falara, V.; Pichersky, E.; Kanellis, A.K. A copal-8-ol diphosphate synthase from the angiosperm *Cistus creticus* subsp. *creticus* is a putative key enzyme for the formation of pharmacologically active, oxygen-containing labdane-type diterpene. *Plant Physiol.* **2010**, *154*, 301–310. [[CrossRef](#)] [[PubMed](#)]
122. Papanikolaou, A.S.; Papaefthimiou, D.; Matekalo, D.; Karakousi, C.V.; Makris, A.M.; Kanellis, A.K. Chemical and transcriptomic analyses of leaf trichomes from *Cistus creticus* subsp. *creticus* reveal the biosynthetic pathways of certain labdane-type diterpenoids and their acetylated forms. *J. Exp. Bot.* **2024**, *75*, 3431–3451. [[CrossRef](#)] [[PubMed](#)]
123. Ivamoto-Suzuki, S.T.; Celedón, J.M.; Yuen, M.M.S.; Kitzberger, C.S.G.; Silva, D.D.; Bohlmann, J.; Protasio, P.L.F. functional characterization of *ent*-Copalyl diphosphate synthase and kaurene synthase genes from *Coffea arabica* L. *J. Agric. Food Chem.* **2023**, *71*, 15863–15873.
124. Smith, M.W.; Yamaguchi, S.; Ait-Ali, T.; Kamiya, Y. The first step of gibberellin biosynthesis in pumpkin is catalyzed by at least two copalyl diphosphate synthases encoded by differentially regulated genes. *Plant Physiol.* **1998**, *118*, 1411–1419. [[CrossRef](#)] [[PubMed](#)]
125. Yamaguchi, S.; Saito, T.; Abe, H.; Yamane, H.; Murofushi, N.; Kamiya, Y. Molecular cloning and characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from pumpkin (*Cucurbita maxima*). *Plant J.* **1996**, *10*, 203–213. [[CrossRef](#)] [[PubMed](#)]
126. Zerbe, P.; Hamberger, B.; Yuen, M.M.; Chiang, A.; Sandhu, H.K.; Madilao, L.L.; Nguyen, A.; Hamberger, B.; Bach, S.S.; Bohlmann, J. Gene discovery of modular diterpene metabolism in nonmodel systems. *Plant Physiol.* **2013**, *162*, 1073–1091. [[CrossRef](#)] [[PubMed](#)]
127. Andersen-Ranberg, J.; Kongstad, K.T.; Nielsen, M.T.; Jensen, N.B.; Wurlitzer, J.; Christensen, U.B.; Staerk, D.; Hansen, J.E.; Mortensen, U.H. Expanding the landscape of diterpene structural diversity through stereochemically controlled combinatorial biosynthesis. *Angew. Chem. Int. Ed.* **2016**, *55*, 2142–2146. [[CrossRef](#)]
128. Schepmann, H.G.; Pang, J.; Matsuda, S.P. Cloning and characterization of *Ginkgo biloba* levopimaradiene synthase which catalyzes the first committed step in ginkgolide biosynthesis. *Arch. Biochem. Biophys.* **2001**, *392*, 263–269. [[CrossRef](#)] [[PubMed](#)]
129. Du, G.; Gong, H.Y.; Feng, K.N.; Chen, Q.Q.; Yang, Y.L.; Fu, X.L.; Lu, S.; Zeng, Y. Diterpene synthases facilitating production of the kaurane skeleton of eriocalyxin B in the medicinal plant *Isodon eriocalyx*. *Phytochemistry* **2019**, *158*, 96–102. [[CrossRef](#)] [[PubMed](#)]
130. Li, J.L.; Chen, Q.Q.; Jin, Q.P.; Gao, J.; Zhao, P.J.; Lu, S.; Zeng, Y. *IcPS2* is potentially involved in the biosynthesis of pharmacologically active *Isodon* diterpenoids rather than gibberellin. *Phytochemistry* **2012**, *76*, 32–39. [[CrossRef](#)] [[PubMed](#)]
131. Pelot, K.A.; Hagelthorn, L.M.; Addison, J.B.; Zerbe, P. Biosynthesis of the oxygenated diterpene nezukol in the medicinal plant *Isodon rubescens* is catalyzed by a pair of diterpene synthases. *PLoS ONE* **2017**, *12*, e0176507. [[CrossRef](#)] [[PubMed](#)]
132. Sawada, Y.; Katsumata, T.; Kitamura, J.; Kawaide, H.; Nakajima, M.; Asami, T.; Nakaminami, K.; Kurahashi, T.; Mitsunashi, W.; Inoue, Y.; et al. Germination of photoblastic lettuce seeds is regulated via the control of endogenous physiologically active gibberellin content, rather than of gibberellin responsiveness. *J. Exp. Bot.* **2008**, *59*, 3383–3393. [[CrossRef](#)] [[PubMed](#)]
133. Yang, R.; Du, Z.; Qiu, T.; Sun, J.; Shen, Y.; Huang, L. Discovery and functional characterization of a diverse diterpene synthase family in the medicinal herb *Isodon lophanthoides* Var. *gerardiana*. *Plant Cell Physiol.* **2021**, *62*, 1423–1435. [[CrossRef](#)] [[PubMed](#)]
134. Sallaud, C.; Giacalone, C.; Töpfer, R.; Goepfert, S.; Bakaher, N.; Rösti, S.; Tissier, A. Characterization of two genes for the biosynthesis of the labdane diterpene *Z*-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. *Plant J.* **2012**, *72*, 1–17. [[CrossRef](#)] [[PubMed](#)]

135. Zhou, S.; Zhan, C.; Zhu, J.; Yang, C.; Zhao, Q.; Sun, Y.; Zhou, J.; Shen, S.; Luo, J. Molecular and biochemical evolution of casbene-type diterpene and sesquiterpene biosynthesis in rice. *J. Integr. Plant Biol.* **2025**, *67*, 1105–1118. [\[CrossRef\]](#) [\[PubMed\]](#)
136. Otomo, K.; Kenmoku, H.; Oikawa, H.; König, W.A.; Toshima, H.; Mitsunashi, W.; Yamane, H.; Sassa, T.; Toyomasu, T. Biological functions of *ent*- and *syn*-copalyl diphosphate synthases in rice: Key enzymes for the branch point of gibberellin and phytoalexin biosynthesis. *Plant J.* **2004**, *39*, 886–893. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Prisic, S.; Xu, M.; Wilderman, P.R.; Peters, R.J. Rice contains two disparate *ent*-copalyl diphosphate synthases with distinct metabolic functions. *Plant Physiol.* **2004**, *136*, 4228–4236. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Xu, M.; Hillwig, M.L.; Prisic, S.; Coates, R.M.; Peters, R.J. Functional identification of rice *syn*-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. *Plant J.* **2004**, *39*, 309–318. [\[CrossRef\]](#) [\[PubMed\]](#)
139. Cho, E.M.; Okada, A.; Kenmoku, H.; Otomo, K.; Toyomasu, T.; Mitsunashi, W.; Sassa, T.; Yajima, A.; Yabuta, G.; Mori, K.; et al. Molecular cloning and characterization of a cDNA encoding *ent*-cassa-12,15-diene synthase, a putative diterpenoid phytoalexin biosynthetic enzyme, from suspension-cultured rice cells treated with a chitin elicitor. *Plant J.* **2004**, *37*, 1–8. [\[PubMed\]](#)
140. Xu, M.; Wilderman, P.R.; Morrone, D.; Xu, J.; Roy, A.; Margis-Pinheiro, M.; Upadhyaya, N.M.; Coates, R.M.; Peters, R.J. Functional characterization of the rice kaurene synthase-like gene family. *Phytochemistry* **2007**, *68*, 312–326. [\[CrossRef\]](#) [\[PubMed\]](#)
141. Margis-Pinheiro, M.; Zhou, X.R.; Zhu, Q.H.; Dennis, E.S.; Upadhyaya, N.M. Isolation and characterization of a Ds-tagged rice (*Oryza sativa* L.) GA-responsive dwarf mutant defective in an early step of the gibberellin biosynthesis pathway. *Plant Cell Rep.* **2005**, *23*, 819–833. [\[CrossRef\]](#) [\[PubMed\]](#)
142. Otomo, K.; Kanno, Y.; Motegi, A.; Kenmoku, H.; Yamane, H.; Mitsunashi, W.; Oikawa, H.; Toshima, H.; Itoh, H.; Matsuoka, M.; et al. Diterpene cyclases responsible for the biosynthesis of phytoalexins, momilactones A, B, and oryzalexins A-F in rice. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 2001–2006. [\[CrossRef\]](#) [\[PubMed\]](#)
143. Morrone, D.; Jin, Y.; Xu, M.; Choi, S.Y.; Coates, R.M.; Peters, R.J. An unexpected diterpene cyclase from rice: Functional identification of a stemodene synthase. *Arch. Biochem. Biophys.* **2006**, *448*, 133–140. [\[CrossRef\]](#) [\[PubMed\]](#)
144. Pelot, K.A.; Chen, R.; Hagelthorn, D.M.; Young, C.A.; Addison, J.B.; Muchlinski, A.; Tholl, D.; Zerbe, P. Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* **2018**, *178*, 54–71. [\[CrossRef\]](#) [\[PubMed\]](#)
145. Wyatt, G.; Zerbe, P.; Tiedge, K. Characterization of switchgrass (*Panicum virgatum* L.) PvKSL1 as a levopimaradiene/abietadiene-type diterpene synthase. *Plant Biol.* **2024**, *27*, 698–709. [\[CrossRef\]](#) [\[PubMed\]](#)
146. Hayashi, K.; Kawaide, H.; Notomi, M.; Sakigi, Y.; Matsuo, A.; Nozaki, H. Identification and functional analysis of bifunctional *ent*-kaurene synthase from the moss *Physcomitrella patens*. *FEBS Lett.* **2006**, *580*, 6175–6181. [\[CrossRef\]](#) [\[PubMed\]](#)
147. Martin, D.M.; Fäldt, J.; Bohlmann, J. Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol.* **2004**, *135*, 1908–1927. [\[CrossRef\]](#) [\[PubMed\]](#)
148. Keeling, C.I.; Dullat, H.K.; Yuen, M.; Ralph, S.G.; Jancsik, S.; Bohlmann, J. Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. *Plant Physiol.* **2010**, *152*, 1197–1208. [\[PubMed\]](#)
149. Keeling, C.I.; Weisshaar, S.; Ralph, S.G.; Jancsik, S.; Hamberger, B.; Dullat, H.K.; Bohlmann, J. Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *Plant Biol.* **2011**, *11*, 43. [\[CrossRef\]](#)
150. Hall, D.E.; Zerbe, P.; Jancsik, S.; Quesada, A.L.; Dullat, H.; Madilao, L.L.; Yuen, M.; Bohlmann, J. Evolution of conifer diterpene synthases: Diterpene resin acid biosynthesis in lodgepole pine and jack pine involves monofunctional and bifunctional diterpene synthases. *Plant Physiol.* **2013**, *161*, 600–616. [\[PubMed\]](#)
151. Pateraki, I.; Andersen-Ranberg, J.; Hamberger, B.; Heskes, A.M.; Martens, H.J.; Zerbe, P.; Bach, S.S.; Møller, B.L.; Bohlmann, J.; Hamberger, B. Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. *Plant Physiol.* **2014**, *164*, 1222–1236. [\[CrossRef\]](#) [\[PubMed\]](#)
152. Chen, Y.; Lin, Y.; Qiu, Y.; Li, W.; Shen, Y.; Huang, L. Identification and functional characterization of the diterpene synthase family in *Pogostemon cablin* (Blanco) Ben. *Plant Physiol. Biochem.* **2024**, *216*, 109190. [\[CrossRef\]](#) [\[PubMed\]](#)
153. Johnson, S.R.; Bhat, W.W.; Sadre, R.; Miller, G.P.; Garcia, A.S.; Hamberger, B. Promiscuous terpene synthases from *Prunella vulgaris* highlight the importance of substrate and compartment switching in terpene synthase evolution. *New Phytol.* **2019**, *223*, 323–335. [\[CrossRef\]](#) [\[PubMed\]](#)
154. Keeling, C.I.; Madilao, L.L.; Zerbe, P.; Dullat, H.K.; Bohlmann, J. The primary diterpene synthase products of *Picea abies* levopimaradiene/abietadiene synthase (PaLAS) are epimers of a thermally unstable diterpen. *J. Biol. Chem.* **2011**, *286*, 21145–21153. [\[CrossRef\]](#) [\[PubMed\]](#)
155. Mafu, S.; Karunanithi, P.S.; Palazzo, T.A.; Harrod, B.L.; Rodriguez, S.M.; Mollhoff, I.N.; O'Brien, T.E.; Tong, S.; Fiehn, O.; Tantillo, D.J.; et al. Biosynthesis of the microtubule-destabilizing diterpene pseudolaric acid B from golden larch involves an unusual diterpene synthase. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 974–979. [\[CrossRef\]](#) [\[PubMed\]](#)

156. Jackson, A.J.; Hershey, D.M.; Chesnut, T.; Xu, M.; Peters, R.J. Biochemical characterization of the castor bean *ent*-kaurene synthase(-like) family supports quantum chemical view of diterpene cyclization. *Phytochemistry* **2014**, *103*, 13–21. [\[CrossRef\]](#) [\[PubMed\]](#)
157. Pelot, K.A.; Mitchell, R.; Kwon, M.; Hagelthorn, L.M.; Wardman, J.F.; Chiang, A.; Bohlmann, J.; Ro, D.K.; Zerbe, P. Biosynthesis of the psychotropic plant diterpene salvinorin A: Discovery and characterization of the *Salvia divinorum* clerodienyl diphosphate synthase. *Plant J.* **2017**, *89*, 885–897. [\[CrossRef\]](#) [\[PubMed\]](#)
158. Chen, X.; Berim, A.; Dayan, F.E.; Gang, D.R. A (-)-kolavenyl diphosphate synthase catalyzes the first step of salvinorin A biosynthesis in *Salvia divinorum*. *J. Exp. Bot.* **2017**, *68*, 1109–1122. [\[CrossRef\]](#) [\[PubMed\]](#)
159. Cui, G.; Duan, L.; Jin, B.; Qian, J.; Xue, Z.; Shen, G.; Snyder, J.H.; Song, J.; Chen, S.; Huang, L.; et al. Functional divergence of diterpene syntheses in the medicinal plant *Salvia miltiorrhiza*. *Plant Physiol.* **2015**, *169*, 1607–1618. [\[CrossRef\]](#) [\[PubMed\]](#)
160. Ignea, C.; Ioannou, E.; Georgantea, P.; Loupassaki, S.; Trikkas, F.A.; Kanellis, A.K.; Makris, A.M.; Roussis, V.; Kampranis, S.C. Reconstructing the chemical diversity of labdane-type diterpene biosynthesis in yeast. *Metab. Eng.* **2015**, *28*, 91–103. [\[CrossRef\]](#) [\[PubMed\]](#)
161. Caniard, A.; Zerbe, P.; Legrand, S.; Cohade, A.; Valot, N.; Magnard, J.L.; Bohlmann, J.; Legendre, L. Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. *Plant Biol.* **2012**, *12*, 119. [\[CrossRef\]](#)
162. Takeshi, N.; Jung-Bum, L.; Toshimitsu, H. cDNA cloning, functional expression and characterization of *ent*-copalyl diphosphate synthase from *Scoparia dulcis* L. *Plant Sci.* **2005**, *169*, 760–767. [\[CrossRef\]](#)
163. Lee, J.B.; Ohmura, T.; Yamamura, Y. Functional characterization of three diterpene synthases responsible for tetracyclic diterpene biosynthesis in *Scoparia dulcis*. *Plants* **2022**, *12*, 69. [\[CrossRef\]](#) [\[PubMed\]](#)
164. Qiu, T.; Li, Y.; Wu, H.; Yang, H.; Peng, Z.; Du, Z.; Wu, Q.; Wang, H.; Shen, Y.; Huang, L. Tandem duplication and sub-functionalization of clerodane diterpene synthase originate the blooming of clerodane diterpenoids in *Scutellaria barbata*. *Plant J.* **2023**, *116*, 375–388. [\[CrossRef\]](#) [\[PubMed\]](#)
165. Yu, J.; Shiraiishi, T.; Taizoumbe, K.A.; Karasuno, Y.; Yoshida, A.; Nishiyama, M.; Dickschat, J.S.; Kuzuyama, T. Mechanistic characterization of diterpene synthase pairs for tricyclic diterpenes from Cyanobacteria. *J. Am. Chem. Soc.* **2025**, *147*, 11896–11905. [\[CrossRef\]](#) [\[PubMed\]](#)
166. Sugai, Y.; Ueno, Y.; Hayashi, K.; Oogami, S.; Toyomasu, T.; Matsumoto, S.; Natsume, M.; Nozaki, H.; Kawaide, H. Enzymatic (13)Cl labeling and multidimensional NMR analysis of mitratriene synthesized by bifunctional diterpene cyclase in *Selaginella moellendorffii*. *J. Biol. Chem.* **2011**, *286*, 42840–42847. [\[CrossRef\]](#) [\[PubMed\]](#)
167. Falara, V.; Akhtar, T.A.; Nguyen, T.T.; Spyropoulou, E.A.; Bleeker, P.M.; Schauvinhold, I.; Matsuba, Y.; Bonini, M.E.; Schillmiller, A.L.; Last, R.L.; et al. The tomato terpene synthase gene family. *Plant Physiol.* **2011**, *157*, 770–789. [\[CrossRef\]](#) [\[PubMed\]](#)
168. Tan, H.H.; Xia, M.; Su, P.; Huang, L.Q. Functional characterization of *ent*-kaurane-type diterpenoid synthases from *Stellera chamaejasme*. *China J. Chin. Mater. Med.* **2024**, *49*, 2410–2421.
169. Richman, A.S.; Gijzen, M.; Starratt, A.N.; Yang, Z.; Brandle, J.E. Diterpene synthesis in *Stevia rebaudiana*: Recruitment and up-regulation of key enzymes from the gibberellin biosynthetic pathway. *Plant J.* **1999**, *19*, 411–421. [\[CrossRef\]](#) [\[PubMed\]](#)
170. Ma, L.T.; Lee, Y.R.; Tsao, N.W.; Wang, S.Y.; Zerbe, P.; Chu, F.H. Biochemical characterization of diterpene synthases of Taiwanian cryptomerioides expands the known functional space of specialized diterpene metabolism in gymnosperms. *Plant J.* **2019**, *100*, 1254–1272. [\[CrossRef\]](#) [\[PubMed\]](#)
171. Wildung, M.R.; Croteau, R. A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. *J. Biol. Chem.* **1996**, *271*, 9201–9204. [\[CrossRef\]](#) [\[PubMed\]](#)
172. Inabuy, F.S.; Fischedick, J.T.; Lange, I.; Hartmann, M.; Srividya, N.; Parrish, A.N.; Xu, M.; Peters, R.J.; Lange, B.M. Biosynthesis of diterpenoids in *Tripterygium* adventitious root cultures. *Plant Physiol.* **2017**, *175*, 92–103. [\[CrossRef\]](#) [\[PubMed\]](#)
173. Su, P.; Guan, H.; Zhao, Y.; Tong, Y.; Xu, M.; Zhang, Y.; Hu, T.; Yang, J.; Cheng, Q.; Gao, L.; et al. Identification and functional characterization of diterpene synthases for triptolide biosynthesis from *Tripterygium wilfordii*. *Plant J.* **2018**, *93*, 50–65. [\[PubMed\]](#)
174. Hansen, N.L.; Heskes, A.M.; Hamberger, B.; Olsen, C.E.; Hallström, B.M.; Andersen-Ranberg, J.; Hamberger, B. The terpene synthase gene family in *Tripterygium wilfordii* harbors a labdane-type diterpene synthase among the monoterpene synthase TPS-b subfamily. *Plant J.* **2017**, *89*, 429–441. [\[CrossRef\]](#) [\[PubMed\]](#)
175. Hillwig, M.L.; Xu, M.; Toyomasu, T.; Tiernan, M.S.; Wei, G.; Cui, G.; Huang, L.; Peters, R.J. Domain loss has independently occurred multiple times in plant terpene synthase evolution. *Plant J.* **2011**, *68*, 1051–1060. [\[CrossRef\]](#) [\[PubMed\]](#)
176. Heskes, A.M.; Sundram, T.C.M.; Boughton, B.A.; Jensen, N.B.; Hansen, N.L.; Crocoll, C.; Cozzi, F.; Rasmussen, S.; Hamberger, B.R.; Hamberger, B.; et al. Biosynthesis of bioactive diterpenoids in the medicinal plant *Vitex agnus-castus*. *Plant J.* **2018**, *93*, 943–958. [\[CrossRef\]](#) [\[PubMed\]](#)
177. Harris, L.J.; Sapanaro, A.; Johnston, A.; Pristic, S.; Xu, M.; Allard, S.; Kathiresan, A.; Ouellet, T.; Peters, R.J. The maize *An2* gene is induced by *Fusarium* attack and encodes an *ent*-copalyl diphosphate synthase. *Plant Mol. Biol.* **2005**, *59*, 881–894. [\[CrossRef\]](#) [\[PubMed\]](#)

178. Ignea, C.; Athanasakoglou, A.; Ioannou, E.; Georgantea, P.; Trika, F.A.; Loupassaki, S.; Roussis, V.; Makris, A.M.; Kampranis, S.C. Carnosic acid biosynthesis elucidated by a synthetic biology platform. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 3681–3686. [[CrossRef](#)] [[PubMed](#)]
179. Winkler, R.G.; Helentjaris, T. The Maize Dwarf3 Gene Encodes a Cytochrome P450-Mediated early step in gibberellin biosynthesis. *Plant Cell* **1995**, *7*, 1307. [[CrossRef](#)] [[PubMed](#)]
180. Zhang, Y.; Bian, S.; Liu, X.; Fang, N.; Wang, C.; Liu, Y.; Du, Y.; Timko, M.P.; Zhang, Z.; Zhang, H. Synthesis of cembratriene-ol and cembratriene-diol in yeast via the MVA pathway. *Microb. Cell Fact.* **2021**, *20*, 29. [[CrossRef](#)] [[PubMed](#)]
181. Helliwell, C.A.; Olive, M.R.; Forster, L.G.R.; Peacock, W.J.; Dennis, E.S. Isolation of an ent-kaurene oxidase cDNA from *Cucurbita maxima*. *Aust. J. Plant Physiol.* **2000**, *27*, 1141–1149. [[CrossRef](#)]
182. Schoendorf, A.; Rithner, C.D.; Williams, R.M.; Croteau, R.B. Molecular cloning of a cytochrome P450 taxane 10 β -hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2065–2070. [[CrossRef](#)]
183. Davidson, S.E.; Elliott, R.C.; Helliwell, C.A.; Poole, A.T.; Reid, J.B. The Pea gene NA encodes ent-Kaurenoic acid oxidase. *Plant Physiol.* **2003**, *131*, 335–344. [[CrossRef](#)] [[PubMed](#)]
184. Jennewein, S.; Rithner, C.D.; Williams, R.M.; Croteau, R. Taxoid metabolism: Taxoid 14 β -hydroxylase is a cytochrome P450-dependent monooxygenase. *Arch. Biochem. Biophys.* **2003**, *413*, 262–270. [[CrossRef](#)] [[PubMed](#)]
185. Davidson, S.E.; Smith, J.J.; Helliwell, C.A.; Poole, A.T.; Reid, J.B. The Pea gene LH encodes ent-Kaurene oxidase. *Plant Physiol.* **2004**, *134*, 1123–1134. [[CrossRef](#)] [[PubMed](#)]
186. Humphrey, T.V.; Richman, A.S.; Menassa, R.; Brandle, J.E. Spatial organisation of four enzymes from *Stevia rebaudiana* that are involved in steviol glycoside synthesis. *Plant Mol. Biol.* **2006**, *61*, 47–62. [[CrossRef](#)] [[PubMed](#)]
187. Chau, M.; Croteau, R. Taxol Biosynthesis: Molecular cloning and characterization of a cytochrome P450 taxoid 7 β -Hydroxylase. *Chem. Biol.* **2004**, *427*, 48–57. [[CrossRef](#)]
188. Zhou, F.; Teng, X.; Wang, P.; Zhang, Y.; Miao, Y. Isolation, identification, expression and subcellular localization of PPARG gene in buffalo mammary gland. *Gene* **2020**, *759*, 144981. [[CrossRef](#)] [[PubMed](#)]
189. Chau, M. Molecular cloning and characterization of a cytochrome P450 taxoid 2 α -hydroxylase involved in Taxol biosynthesis. *Arch. Biochem. Biophys.* **2004**, *427*, 48–57. [[CrossRef](#)] [[PubMed](#)]
190. Jennewein, S.; Rithner, C.D.; Williams, R.M.; Croteau, R.B. Taxol biosynthesis: Taxane 13 α -hydroxylase is a cytochrome P450-dependent monooxygenase. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13595–13600. [[CrossRef](#)] [[PubMed](#)]
191. Wang, Q.; Hillwig, M.L.; Wu, Y.; Peters, R.J. CYP701A8: A rice ent-Kaurene oxidase paralog diverted to more specialized diterpenoid metabolism. *Plant Physiol.* **2012**, *158*, 1418–1425. [[CrossRef](#)] [[PubMed](#)]
192. Ro, D.; Arimura, G.; Lau, S.Y.W.; Piers, E.; Bohlmann, J. Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8060–8065. [[CrossRef](#)] [[PubMed](#)]
193. Sakamoto, T.; Miura, K.; Itoh, H.; Tatsumi, T.; Ueguchi-Tanaka, M.; Ishiyama, K.; Kobayashi, M.; Agrawal, G.K.; Takeda, S.; Abe, K.; et al. An overview of gibberellin metabolism enzyme genes and their related mutants in Rice. *Plant Physiol.* **2004**, *134*, 1642–1653. [[CrossRef](#)] [[PubMed](#)]
194. Geisler, K.; Jensen, N.B.; Yuen, M.M.; 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.* **2016**, *171*, 152–164. [[CrossRef](#)] [[PubMed](#)]
195. Hamberger, B.; Ohnishi, T.; Hamberger, B.; Séguin, A.; Bohlmann, J. Evolution of diterpene metabolism: Sitka Spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. *Plant Physiol.* **2011**, *157*, 1677–1695. [[CrossRef](#)] [[PubMed](#)]
196. Guo, X.; Li, X.; Duan, X.; Shen, Y.; Xing, Y.; Cao, Q.; Qin, L. Characterization of sck1, a novel castanea mollissima mutant with the extreme short catkins and decreased gibberellin. *PLoS ONE* **2012**, *7*, e43181. [[CrossRef](#)] [[PubMed](#)]
197. Zhang, N.; Han, Z.; Sun, G.; Hoffman, A.; Wilson, I.W.; Yang, Y.; Gao, Q.; Wu, J.; Xie, D.; Dai, J.; et al. Molecular cloning and characterization of a cytochrome P450 taxoid 9 α -hydroxylase in *Ginkgo biloba* cells. *Biochem. Biophys. Res. Commun.* **2014**, *443*, 938–943. [[CrossRef](#)] [[PubMed](#)]
198. King, A.J.; Brown, G.D.; Gilday, A.D.; Larson, T.R.; Graham, I.A. production of bioactive diterpenoids in the Euphorbiaceae depends on evolutionarily conserved gene clusters. *Plant Cell* **2014**, *26*, 3286–3298. [[CrossRef](#)] [[PubMed](#)]
199. Su, P.; Tong, Y.; Cheng, Q.; Hu, Y.; Zhang, M.; Yang, J.; Teng, Z.; Gao, W.; Huang, L. Functional characterization of ent-copalyl diphosphate synthase, kaurene synthase and kaurene oxidase in the *Salvia miltiorrhiza* gibberellin biosynthetic pathway. *Sci. Rep.* **2016**, *6*, 23057. [[CrossRef](#)] [[PubMed](#)]
200. Scheler, U.; Brandt, W.; Porzel, A.; Rothe, K.; Manzano, D.; Božić, D.; Papaefthimiou, D.; Balcke, G.U.; Henning, A.; Lohse, S.; et al. Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. *Nat. Commun.* **2016**, *7*, 12942. [[CrossRef](#)] [[PubMed](#)]

201. Ko, K.; Lin, F.; Katsumata, T.; Sugai, Y.; Miyazaki, S.; Kawaide, H.; Okada, K.; Nojiri, H.; Yamane, H. Functional identification of a Rice *ent*-Kaurene oxidase, OsKO2, using the *Pichia pastoris* expression system. *Biosci. Biotechnol. Biochem.* **2008**, *72*, 3285–3288. [CrossRef] [PubMed]
202. King, A.J.; Brown, G.D.; Gilday, A.D.; Forestier, E.; Larson, T.R.; Graham, I.A. A Cytochrome P450-mediated intramolecular carbon–carbon ring closure in the biosynthesis of multidrug-resistance-reversing lathyrane diterpenoids. *Chem. Bio. Chem.* **2016**, *17*, 1593–1597. [CrossRef] [PubMed]
203. Ding, Y.; Murphy, K.M.; Poretsky, E.; Mafu, S.; Yang, B.; Char, S.N.; Christensen, S.A.; Saldivar, E.; Wu, M.; Wang, Q.; et al. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nat. Plants* **2019**, *5*, 1043–1056. [CrossRef] [PubMed]
204. He, J.; Chen, Q.; Xin, P.; Yuan, J.; Ma, Y.; Wang, X.; Xu, M.; Chu, J.; Peters, R.J.; Wang, G. CYP72A enzymes catalyse 13-hydrolyzation of gibberellins. *Nat. Plants* **2019**, *5*, 1057–1065. [CrossRef] [PubMed]
205. Su, P.; Guan, H.; Zhang, Y.; Wang, X.; Gao, L.; Zhao, Y.; Hu, T.; Zhou, J.; Ma, B.; Tu, L.; 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.* **2017**, *8*, 1756. [CrossRef] [PubMed]
206. Miyazaki, S.; Katsumata, T.; Natsume, M.; Kawaide, H. The CYP701B1 of *Physcomitrella patens* is an *ent*-kaurene oxidase that resists inhibition by uniconazole-P. *FEBS Lett.* **2011**, *585*, 1879–1883. [CrossRef] [PubMed]
207. Mao, L.; Kawaide, H.; Higuchi, T.; Chen, M.; Miyamoto, K.; Hirata, Y.; Kimura, H.; Miyazaki, S.; Teruya, M.; Fujiwara, K.; et al. Genomic evidence for convergent evolution of gene clusters for momilactone biosynthesis in land plants. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 12472–12480. [CrossRef] [PubMed]
208. Fambrini, M.; Mariotti, L.; Parlanti, S.; Picciarelli, P.; Salvini, M.; Ceccarelli, N.; Pugliesi, C. The extreme dwarf phenotype of the GA-sensitive mutant of sunflower, dwarf2, is generated by a deletion in the *ent*-kaurenoic acid oxidase1 (*HaKAO1*) gene sequenc. *Plant Mol. Biol.* **2011**, *75*, 431–450. [CrossRef] [PubMed]
209. Nomura, T.; Magome, H.; Hanada, A.; Takeda-Kamiya, N.; Mander, L.N.; Kamiya, Y.; Yamaguchi, S. Functional analysis of *Arabidopsis* CYP714A1 and CYP714A2 reveals that they are distinct gibberellin modification enzymes. *Plant Cell Physiol.* **2013**, *54*, 1837–1851. [CrossRef] [PubMed]
210. Magome, H.; Nomura, T.; Hanada, A.; Takeda-Kamiya, N.; Ohnishi, T.; Shinma, Y.; Katsumata, T.; Kawaide, H.; Kamiya, Y.; Yamaguchi, S. CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1947–1952. [CrossRef] [PubMed]
211. Karunanithi, P.S.; Dhanota, P.; Addison, J.B.; Tong, S.; Fiehn, O.; Zerbe, P. Functional characterization of the cytochrome P450 monooxygenase CYP71AU87 indicates a role in marrubiin biosynthesis in the medicinal plant *Marrubium vulgare*. *BMC. Plant Biol.* **2019**, *19*, 114. [CrossRef] [PubMed]
212. Karunanithi, P.S.; Berrios, D.I.; Wang, S.; Davis, J.; Shen, T.; Fiehn, O.; Maloof, J.N.; Zerbe, P. The foxtail millet (*Setaria italica*) terpene synthase gene family. *Plant J.* **2020**, *103*, 781–800. [CrossRef] [PubMed]
213. Kwon, M.; Utomo, J.C.; Park, K.; Pascoe, C.A.; Chiorean, S.; Ngo, I.; Pelot, K.A.; Pan, C.; Kim, S.; Zerbe, P.; et al. Cytochrome P450-Catalyzed Biosynthesis of a Dihydrofuran Neoclerodane in Magic Mint (*Salvia divinorum*). *ACS Catal.* **2022**, *12*, 777–782. [CrossRef]
214. Miller, G.P.; Bhat, W.W.; Lanier, E.R.; Johnson, S.R.; Mathieu, D.T.; Hamberger, B. The biosynthesis of the anti-microbial diterpenoid leubethanol in *Leucophyllum frutescens* proceeds via an all-cis prenyl intermediate. *Plant J.* **2020**, *104*, 693–705. [CrossRef] [PubMed]
215. Li, C.; Yang, L.; Liu, Y.; Xu, Z.; Gao, J.; Huang, Y.; Xu, J.; Fan, H.; Kong, Y.; Wei, Y.; et al. The sage genome provides insight into the evolutionary dynamics of diterpene biosynthesis gene cluster in. *Plants* **2022**, *40*, 111236. [CrossRef]
216. Matsuba, Y.; Zi, J.; Jones, A.D.; Peters, R.J.; Pichersky, E. Biosynthesis of the diterpenoid lycosantalanol via neryl neryl diphosphate in *Solanum lycopersicum*. *PLoS ONE* **2015**, *10*, e0119302. [CrossRef] [PubMed]
217. Wang, Q.; Hillwig, M.L.; Okada, K.; Yamazaki, K.; Wu, Y.; Swaminathan, S.; Yamane, H.; Peters, R.J. Characterization of CYP76M5–8 indicates metabolic plasticity within a plant biosynthetic gene cluster. *J. Biol. Chem.* **2012**, *287*, 6159–6168. [CrossRef] [PubMed]
218. Zhu, Y.; Nomura, T.; Xu, Y.; Zhang, Y.; Peng, Y.; Mao, B.; Hanada, A.; Zhou, H.; Wang, R.; Li, P.; et al. ELONGATED UPPERMOST INTERNODE encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* **2006**, *18*, 442–456. [CrossRef] [PubMed]
219. Wang, J.; Su, P.; Gao, L.; Zhang, Y.; Wang, J.; Tu, L.; Zhao, Y.; Lu, Y.; Yin, Y.; Huang, L.; et al. A cytochrome P450 CYP81AM1 from *Tripterygium wilfordii* catalyses the C-15 hydroxylation of dehydroabietic acid. *Planta* **2021**, *254*, 95. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.