Genomic and structural basis for evolution of tropane alkaloid biosynthesis

Article *in* Proceedings of the National Academy of Sciences · April 2023

DOI: 10.1073/pnas.2302448120

9 authors, including:

John D'Auria
Leibniz Institute of Plant Genetics and Crop Plant Research
67 PUBLICATIONS **3,834** CITATIONS
SEE PROFILE

Sheng-Xiong Huang
The Scripps Research Institute
180 PUBLICATIONS **5,482** CITATIONS
SEE PROFILE

PNAS RESEARCH ARTICLE | PLANT BIOLOGY

Genomic and structural basis for evolution of tropane alkaloid biosynthesis

Yong-Jiang Wang[a,1], Tian Tain[a,b,1], Jia-Yi Yu[a,b,1], Jie Li[a,b], Bingyan Xu[a,b], Jianghua Chen[c], John C. D'Auria[d], Jian-Ping Huang[a,2], and Sheng-Xiong Huang[a,2]

Edited by Ian Baldwin, Max-Planck-Institut fur chemische Okologie Abteilung fur Molekulare Okologie, Jena, Germany; received February 12, 2023; accepted March 23, 2023

The tropane alkaloids (TAs) cocaine and hyoscyamine have been used medicinally for thousands of years. To understand the evolutionary origins and trajectories of serial biosynthetic enzymes of TAs and especially the characteristic tropane skeletons, we generated the chromosome-level genome assemblies of cocaine-producing *Erythroxylum novogranatense* (Erythroxylaceae, rosids clade) and hyoscyamine-producing *Anisodus acutangulus* (Solanaceae, asterids clade). Comparative genomic and phylogenetic analysis suggested that the lack of spermidine synthase/*N*-methyltransferase (*En*SPMT1) in ancestral asterids species contributed to the divergence of polyamine (spermidine or putrescine) methylation in cocaine and hyoscyamine biosynthesis. Molecular docking analysis and key site mutation experiments suggested that ecgonone synthases CYP81AN15 and CYP82M3 adopt different active-site architectures to biosynthesize the same product ecgonone from the same substrate in Erythroxylaceae and Solanaceae. Further synteny analysis showed different evolutionary origins and trajectories of CYP81AN15 and CYP82M3, particularly the emergence of *CYP81AN15* through the neofunctionalization of ancient tandem duplication genes. The combination of structural biology and comparative genomic analysis revealed that ecgonone methyltransferase, which is responsible for the biosynthesis of characteristic 2-substituted carboxymethyl group in cocaine, evolved from the tandem copies of salicylic acid methyltransferase by the mutations of critical E216 and S153 residues. Overall, we provided strong evidence for the independent origins of serial TA biosynthetic enzymes on the genomic and structural level, underlying the chemotypic convergence of TAs in phylogenetically distant species.

tropane alkaloids | evolution | biosynthesis | genome | structural biology

Significance

Plants produce structurally diverse and lineage-specific natural products through a coordination of serial enzymes to catalyze sequential reactions. Cocaine and hyoscyamine with similar tropane skeletons are distributed among different families. The independent recruitments of serial biosynthetic enzymes responsible for cocaine and hyoscyamine biosynthesis were revealed on the genomic and structural levels. This rare phenomenon reminds us that the discovery of biosynthetic enzymes based exclusively on enzyme homology or key conserved sites requires caution when investigating the biosynthesis of natural products with similar chemical structure in phylogenetically distant species.

Plants obtain structurally rich and functionally diverse natural products during their evolution (1). It is common to find natural products with similar chemical structures existing in phylogenetically close species. Beyond that, some similar natural products, such as tropane alkaloids (TAs) (2), spiroketal steroids (3), and iridoids (4), show scattered distributions in phylogenetically distant plants. At present, our understanding of the scattered distributions of natural products is dominated by elucidating the evolution of individual biosynthetic enzyme (3–6). The biosynthetic pathway of natural products involves a complex coordination of a set of enzymes which catalyze sequential reactions (7). However, little is known about how a set of enzymes involved in the biosynthesis of natural products exhibiting a scattered distribution pattern evolved in plants. Gene duplication, followed by neofunctionalization, is a major contributor to the diversity of natural products in plants (8). Neofunctionalization can be obtained through a small number of key mutations, which can be identified via protein structural and functional analyses (9, 10). Besides obtaining the neofunctionalized gene, other mutated genetic material generated by gene duplication events can be either lost or remained (11). These genetic traces are subsequently deposited within the plant genome and can be used to trace the ancestral state and locus of a neofunctional gene, in turn, shedding light on the evolutionary process (11). Therefore, the combination of structural analysis and genomic analysis can provide strong evidence to further explore the evolutionary origins and trajectories of biosynthetic enzymes.

Tropane alkaloids can be divided into two major groups: tropinone and ecgonone derivatives (Fig. 1) (12). The tropinone group (i.e., hyoscyamine, scopolamine, calystegine, etc.) mainly exists in Solanaceae, Convolvulaceae, and Rhizophoraceae families (12, 13). The ecgonone group, represented by cocaine, mainly exists in Erythroxylaceae family (12, 13). Among the TAs, hyoscyamine and scopolamine display anticholinergic activity and are clinically used to treat motion sickness, obstetrical analgesia, organophosphate poisoning, and Parkinson's symptoms (14). Cocaine (Goprelto and Numbrino) has been approved by the food and drug administration as a topical anesthetic. After a

Author contributions: J.-P.H. and S.-X.H. designed research; Y.-J.W., T.T., and J.-Y.Y. performed research; Y.-J.W., T.T., J.L., B.X., J.C., J.C.D., and S.-X.H. analyzed data; and Y.-J.W., J.C.D., J.-P.H., and S.-X.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]Y.-J.W., T.T., and J.-Y.Y. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: huangjp1020@163.com or sxhuang@mail.kib.ac.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2302448120/-/DCSupplemental.

Published April 17, 2023.

century of investigation, the complete biosynthetic pathway for hyoscyamine and scopolamine has been fully described, and their de novo production in yeast was achieved (15, 16). Recently, we reported a near-complete biosynthetic route of cocaine (17, 18). Cocaine and hyoscyamine share a similar biosynthetic pathway (Fig. 1), including sharing the common precursors putrescine, *N*-methylputrescine, *N*-methyl-Δ^1-pyrrolinium, 3-oxo-glutaric acid (OGA), 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid (MPOA), and ecgonone (16–18). In contrast to the one-step catalysis of putrescine *N*-methyltransferase (PMT) to form *N*-methylputrescine in solanaceous species, an alternative biosynthetic route for *N*-methylputrescine formation was revealed in Erythroxylaceae (Fig. 1) (18). In addition, the biosynthetic flux to tropinone or ecgonone group is controlled via *En*MT4-catalyzed methylation of ecgonone, thereby avoiding the spontaneous decarboxylation of unstable ecgonone to form tropinone (Fig. 1) (17). The subsequent reduction and esterification of methylecgonone or tropinone drive the formation of cocaine or littorine (the precursor of hyoscyamine) (Fig. 1) (6, 19–21). Methylecgonone reductase (MecgoR) from the aldo-keto reductase (AKR) family (6) and tropinone reductase I (TRI) from the short-chain dehydrogenase/reductase (SDR) family (20), as well as cocaine synthase (CS) from the BAHD acyltransferase (benzylalcohol *O*-acetyl transferase, anthocyanin *O*-hydroxycinnamoyl transferase, *N*-hydroxycinnamoyl anthranilate benzoyl transferase, and deacetylvindoline 4-*O*-acetyltransferase) family (19) and littorine synthase (LS) from the SCPL-AT (serine carboxypeptidase-like acyltransferase) family (21), suggest the independent recruitment of these enzymes in Erythroxylaceae and Solanaceae plants. Previous protein structural and phylogenetic analyses suggested that polyketide synthases in cocaine and hyoscyamine biosynthesis were independently recruited (5). These similar biosynthetic intermediates and different enzyme strategies made us curious about the evolutionary origins and trajectories of serial enzymes involved in the biosynthesis of TAs, especially the characteristic tropane skeletons (i.e., tropinone and ecgonone skeletons in different plants), on both genomic and structural levels.

In this study, protein structural and functional analysis showed that ecgonone synthases (CYP81AN15 and CYP82M3) in Erythroxylaceae and Solanaceae adopted different active-site architectures to biosynthesize the same product, and ecgonone methyltransferase (*En*MT4) most likely evolved from salicylic acid methyltransferase (SAMT) in *Erythroxylum novogranatense*. Two high-quality genome assemblies of cocaine-producing *E. novogranatense* (Erythroxylaceae) and hyoscyamine-producing *Aacutangulus acutangulus* (Solanaceae) were generated. Further comparative genomic and phylogenetic analyses were used to explore the genomic events that led to the independent emergence of a set of TA biosynthetic enzymes in Erythroxylaceae and Solanaceae. Our work revealed layers for the occurrence of natural products with scattered distributions deep into the structural and genomic levels, thus broadening our understanding of the remarkable chemotypic convergence/divergence of natural products in phylogenetically distant plants.

Results

Genome Sequencing, Assembly, and Evolution. To understand the genomic contexts that underpin the emergence of TA biosynthesis in plants, the chromosome-level genomes of hyoscyamine-producing *A. acutangulus* (1.26 Gb) and cocaine-producing *E. novogranatense* (1.30 Gb) were sequenced and assembled by combining Illumina short-read sequencing, nanopore long-read sequencing, and high-throughput chromosome conformation capture (Hi-C) technologies (*SI Appendix*, Fig. S1 and Table S1). The assembly quality of these two genomes was high based on benchmarking universal single-copy orthologs (BUSCO) analysis (93.40 to 94.49% complete BUSCO genes), core eukaryotic gene-mapping approach analysis (94.76 to 96.77% core eukaryotic genes), and short-read mapping analysis (98.88 to 99.83% map rates) (*SI Appendix*, Tables S2–S4).

To better understand the evolutionary relationships between *E. novogranatense* and *A. acutangulus,* we constructed a phylogenetic tree based on 108 single-copy orthologous sequences identified from 11 plants from different families (Fig. 2). This analysis revealed that the possible splitting time between *A. acutangulus* and *E. novogranatense* was approximately 129.4 Mya, corresponding to the time of divergence between the asterids and rosids clades of core eudicots (22). The within-species Ks (synonymous substitutions per synonymous site) distribution of paralogous genes in *A. acutangulus* showed two major peaks at approximately 0.12 and 0.53, respectively (*SI Appendix*, Fig. S2). The peak at 0.53 was consistent with the whole-genome triplication (WGT) event shared by solanaceous species (Fig. 2) (23, 24), while the peak at 0.12 suggested the occurrence of a more recent, species-specific whole-genome duplication (WGD) in *A. acutangulus* following divergence from its last common ancestor with tomato (*Solanum lycopersicum*, Solanaceae), which was further confirmed by intragenomic (*SI Appendix*, Fig. S1) and intergenomic collinearity analysis (*SI Appendix*, Fig. S3). The distribution of Ks values for

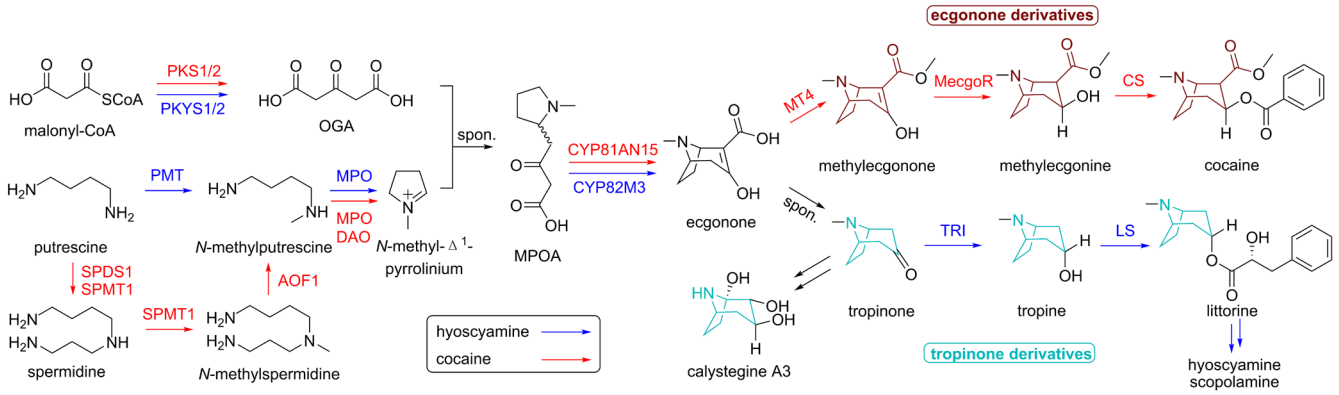

Fig. 1. The similar biosynthetic routes of cocaine and hyoscyamine. PKS: polyketide synthase; PYKS: pyrrolidine ketide synthase; SPDS1: spermidine synthase; PMT: putrescine *N*-methyltransferase; SPMT1: spermidine synthase/*N*-methyltransferase; AOF1: flavin-dependent amine oxidase; MPO: *N*-methylputrescine oxidase; DAO: diamine oxidase; CYP81AN15: ecgonone synthase; CYP82M3: ecgonone synthase; MT4: ecgonone methyltransferase; MecgoR: methylecgonone reductase; TRI: tropinone reductase I; CS: cocaine synthase; LS: littorine synthase.

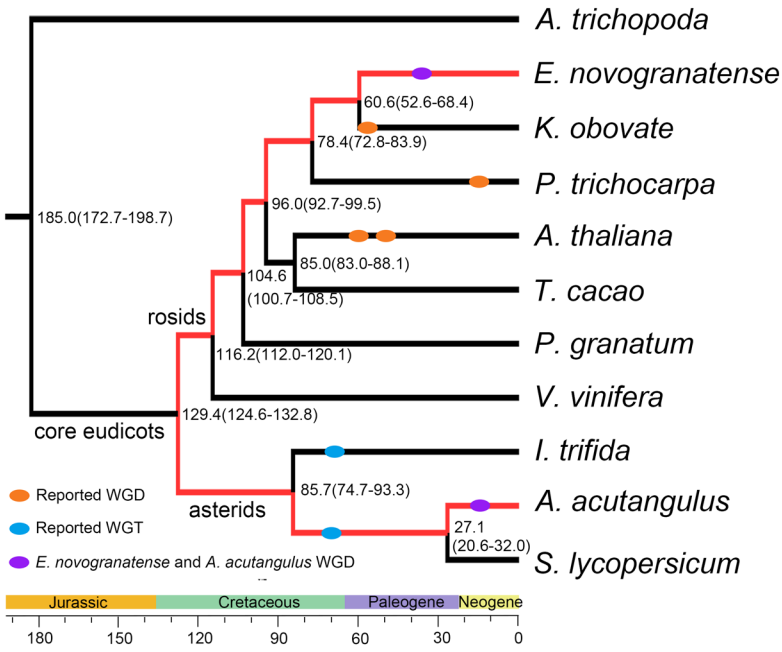

Fig. 2. Inferred phylogenetic tree based on 108 single-copy orthologs of 11 plants. The approximate times of *A. acutangulus* and *E. novogranatense* WGD and other reported WGT/WGD are superimposed on the tree. *Amborella trichopoda* (Amborellaceae); *Kandelia obovate* (Rhizophoraceae); *Populus trichocarpa* (Salicaceae); *Arabidopsis thaliana* (Brassicaceae); *Theobroma cacao* (Malvaceae); *Punica granatum* (Lythraceae); *Vitis vinifera* (Vitaceae); *Ipomoea trifida* (Convolvulaceae); *Solanum lycopersicum* (Solanaceae).

paralogous genes in *E. novogranatense* revealed a WGD event that occurred after its divergence with *Kandelia obovate* (Rhizophoraceae) (*SI Appendix*, Fig. S4). Intergenomic synteny analysis between *E. novogranatense* and grape (*Vitis vinifera*, Vitaceae) revealed a 2:1 syntenic relationship, adding more support to a WGD event in *E. novogranatense* (Fig. 2 and *SI Appendix*, Fig. S5). Subsequently, based on the estimated divergence time in the phylogenetic tree between *A. acutangulus* and tomato (27.1 Mya), and between *E. novogranatense* and *K. obovate* (60.6 Mya), the species-specific WGD events that occurred in *A. acutangulus* and *E. novogranatense* were determined to be around 7.4 to 22.2 Mya and 30.5 to 40.3 Mya, respectively (Fig. 2).

The Origins of TA Pathway Divergence. Methylation of putrescine or spermidine is the first specific step in TA biosynthesis (Fig. 1). In Solanaceae, it was suggested that PMT was derived from spermidine synthase (SPDS) through tandem duplication and neofunctionalization (24). In erythroxylaceous species, the corresponding methylation of spermidine is catalyzed by spermidine synthase/*N*-methyltransferase (*En*SPMT1) (Fig. 1) (18). In addition, bifunctional *En*SPMT1 shows similar catalytic activity in spermidine formation to that of *En*SPDS1 (18). Subsequently, phylogenetic analysis of genome-wide samples of SPMTs and SPDSs from core eudicot species (rosids and asterids clades) showed a long distance between *En*SPMT1 and *En*SPDS1 (*SI Appendix*, Fig. S6). Intriguingly, all core eudicot species analyzed here have SPDSs that phylogenetically group with *En*SPDS1, while the *En*SPMT1 branch consists only of protein sequences from the rosids species (*SI Appendix*, Fig. S6). To elucidate the different distributions of SPMTs and SPDSs, we performed comparative genomic analysis between *E. novogranatense* and grape (*V. vinifera*), which have not undergone recent genome duplication after the shared hexaploidy (or γ) event (25). As a result, *EnSPMT1* and *EnSPDS1* showed synteny with *VvSPMT1* and *VvSPDS1* in grape, respectively (*SI Appendix*, Fig. S7). In grape, the chromosomal regions surrounding *VvSPMT1* (pseudochromosome 17) and *VvSPDS1* (pseudochromosome 01) were generated by the hexaploidy event (25). Intragenomic comparative analysis showed that *VvSPMT1* was the syntenic gene of *VvSPDS1* (*SI Appendix*, Fig. S7). Combined with phylogenetic analysis and intergenomic comparative analysis (*SI Appendix*, Figs. S6 and S7), we propose a straightforward evolutionary history of *En*SPMT1 (Fig. 3*A*): The SPMT branch is derived from the SPDS branch via the hexaploidy event; after the division of rosids and asterids ancestral species, the *SPMT* gene was preserved in the rosids and lost in the asterids.

A flavin-dependent amine oxidase (AOF1) from erythroxylaceous species catalyzes the removal of aminopropyl group in *N*-methylspermidine to form *N*-methylputrescine (Fig. 1) (18). Phylogenetic analysis of genome-wide samples of polyamine oxidases from core eudicot species showed that the AOF1-located branch consists only of protein sequences from the rosids (*SI Appendix*, Fig. S8). Therefore, *AOF* ancestral gene might be lost in the asterids, similar to that of observed for the *SPMT* ancestral gene. The following oxidation of *N*-methylputrescine in Erythroxylaceae and Solanaceae is catalyzed by *N*-methylputrescine oxidases (MPOs) and diamine oxidases (DAOs) (Fig. 1) (18, 26), which are located on synteny blocks well conserved in eudicot plant genomes (*SI Appendix*, Figs. S9 and S10). Although *MPOs* and *DAOs* originated from the same ancestral genes, the lack of initial *SPMT* and *AOF* genes in asterids led to the inability of solanaceous plants to biosynthesize *N*-methylputrescine using a similar biosynthetic pathway in Erythroxylaceae (Fig. 1). The independent recruitment of PMT enzymes responsible for *N*-methylputrescine formation in solanaceous plants resulted in the divergence of polyamine methylation, which is the first specific step of TA biosynthesis in both rosids (Erythroxylaceae) and asterids (Solanaceae) (Figs. 1 and 3*A*).

Independent Evolutionary Origins of *En*PKS1/2 and PYKS1/2. The intermediate OGA, crucial for the formation of the second ring in TAs, is biosynthesized by PYKS and PKS1/2 in Solanaceae and Erythroxylaceae, respectively (Fig. 1) (5, 27, 28). Independent recruitments of *Aa*PYKS1 and *En*PKS1/2 were revealed by structural analysis (5). To further understand the evolutionary origins of

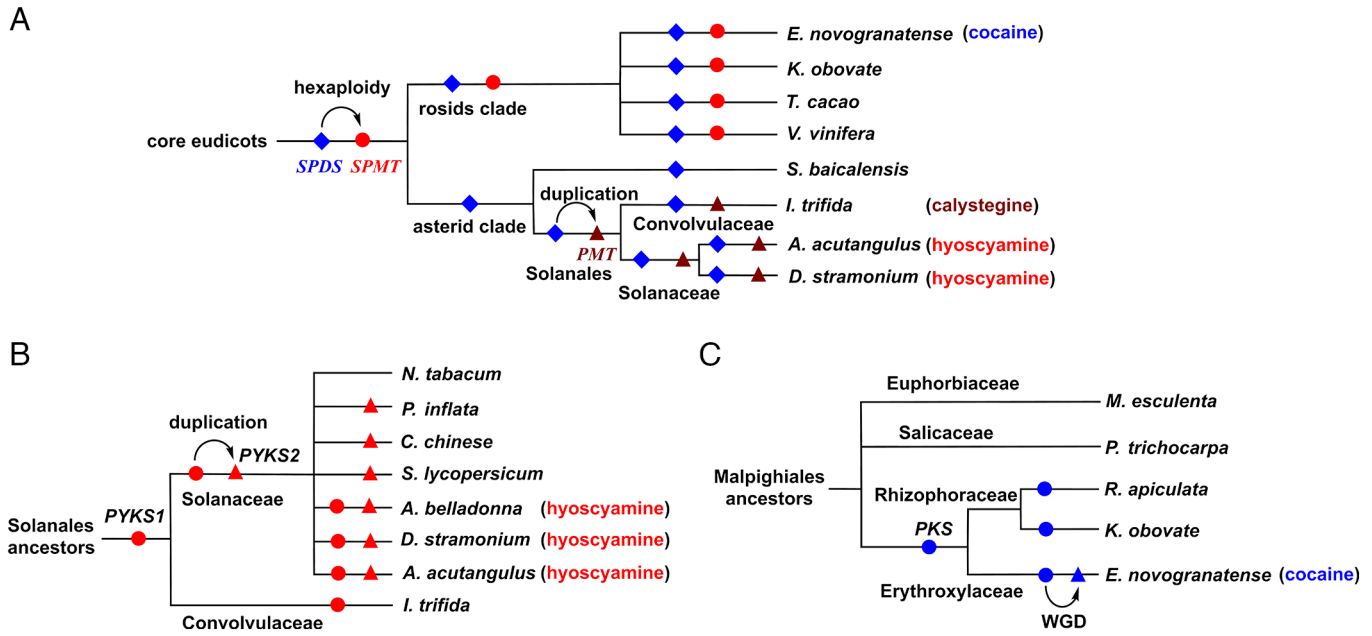

Fig. 3. The evolutionary history of *SPMT*, *PMT*, *PYKS*, and *PKS* genes. (*A*) The evolutionary history of *SPMT* and *PMT* genes in the core eudicots. The red circles, blue diamonds, and brown triangles represent the existence of *SPMT*, *SPDS*, and *PMT* clade genes in the selected species, respectively. (*B*) The evolutionary history of *PYKS* genes in the Solanaceae. The red circles and triangles represent the existence of *PYKS1* and *PYKS2* clade genes in the selected species, respectively. (*C*) The evolutionary history of *PKS* genes in Malpighiales. The blue circles or triangles represent the existence of *PKS* gene in the selected species.

*Aa*PYKS1 and *En*PKS1/2 in their respective genomes, we performed comparative genomic analysis. Unexpectedly, the canonical, syntenic ortholog of *AaPYKS1* was only observed in *Ipomoea trifida* (*ItPYKS*), a species within Convolvulaceae (*SI Appendix*, Fig. S11*A*). The corresponding syntenic *PYKS* (*DsPYKS2*) in the genome of hyoscyamine-producing *Datura stramonium* apparently lost its 5′ exon (29) (*SI Appendix*, Fig. S11*A*). These results indicated ambiguities as to the origin and evolutionary history of the Solanales lineage *PYKS* gene. Subsequently, we traced the origin of functional *DsPYKS1* in *D. stramonium* (27). Intriguingly, the synteny block containing *DsPYKS1* was well conserved in solanaceous plants, except *Nicotiana* species (*SI Appendix*, Fig. S11*B*). The orthologous gene *AaPYKS2* of *DsPYKS1* in *A. acutangulus* showed lateral root-specific expression and similar in vitro catalytic activity in OGA formation to that of *Aa*PYKS1 (*SI Appendix*, Fig. S12), indicating that *A. acutangulus* carried two potentially functionally redundant *PYKS* copies. Combined with intergenomic comparative analysis and synonymous substitution rate analysis between *AaPYKS1* and its paralogous gene *AaPYKS2*, the evolutionary history of *PYKS* was suggested as follows (Fig. 3*B*): Ancestral *PYKS1* emerged in the most recent common ancestor of Solanales and was inherited by both Solanaceae and Convolvulaceae during the divergence of these two families; subsequently, *PYKS1* was subject to a duplication event, resulting in the generation of *PYKS2* (Ks = 0.75), that occurred around the time of the shared Solanaceae WGT. Consequently, both *PYKS1* and *PYKS2* clades were fixed in hyoscyamine-producing *A. acutangulus* and *Atropa belladonna*, while the *PYKS1* clade was pseudogenized (e.g., *DsPYKS2*) or deleted in other solanaceous plant genomes examined here.

The two *PKSs* (*EnPKS*1 and *EnPKS*2) (5) involved in cocaine biosynthesis were harbored within two synteny blocks likely formed by the species-specific WGD in *E. novogranatense* (*SI Appendix*, Fig. S11*C*). Comparative genomics analysis revealed that, although the genomic regions near *EnPKSs* were well conserved within the Malpighiales, syntenic orthologs of *EnPKSs* were only found in Rhizophoraceae (*SI Appendix*, Fig. S11*C*), suggesting that ancestral PKS capable of OGA production emerged in the last common ancestor of Erythroxylaceae and Rhizophoraceae (Fig. 3*C*). The independent acquisitions of OGA-producing PKSs in their respective ancestral species and the increase of functional *PKSs* gene dosages generated by the WGD/WGT events promoted the flux of *N*-methyl-Δ^1-pyrrolinium into pyrrolizidine (e.g., hygrine or cuscohygrine) or TAs (Fig. 1).

Contrasting Active-Site Architectures of Ecgonone Synthases in Cocaine and Hyoscyamine Biosynthesis. The intermediate MPOA, produced by nonenzymatic Mannich condensation of *N*-methyl-Δ^1-pyrrolinium and OGA, is oxidized to form tropane ring by CYP81AN15 and CYP82M3 in cocaine and hyoscyamine biosynthesis, respectively (Fig. 1) (17, 28). Despite CYP81AN15 and CYP82M3 belonging to different CYP450 subfamilies, the overall conserved 3D structural fold in diversified P450 superfamily members made us wonder whether the same active-site architectures were recruited to be responsible for the oxidation of MPOA (Fig. 4*A*). Subsequently, we predicted their protein structures and performed molecular docking with MPOA in the predicted catalytic pockets of CYP81AN15 and CYP82M3 to explore their potential catalytic mechanisms. In CYP81AN15, T115 and Q116 were predicted to fix the carboxy group of MPOA via salt bridge or hydrogen bond, and D372 could form a salt bridge with the *N* atom in the pyrrole ring of MPOA (Fig. 4*B*). To validate the roles of these predicted residues, site-directed mutants of CYP81AN15 were generated. The catalytic activities were almost abolished by the mutation of T115A or Q116A, and similar results were observed when the amide group in Q116 residue was replaced with a carboxyl group by the mutation of Q116E (Fig. 4*C* and *SI Appendix*, Fig. S13), suggesting that T115 and Q116 cooperate to govern the direction of carboxy group of MPOA via salt bridge or hydrogen bond. Large-to-small (D372A) and polar-to-nonpolar (D372V) substitutions of D372 resulted in 93.9% and 97.8% decrease in CYP81AN15 activity, respectively. In contrast, functionally similar substitution (D372E) showed 57.4% catalytic activity of wild CYP81AN15 (Fig. 4*C* and *SI Appendix*, Fig. S13), indicating that the carboxyl group in the D372 residue

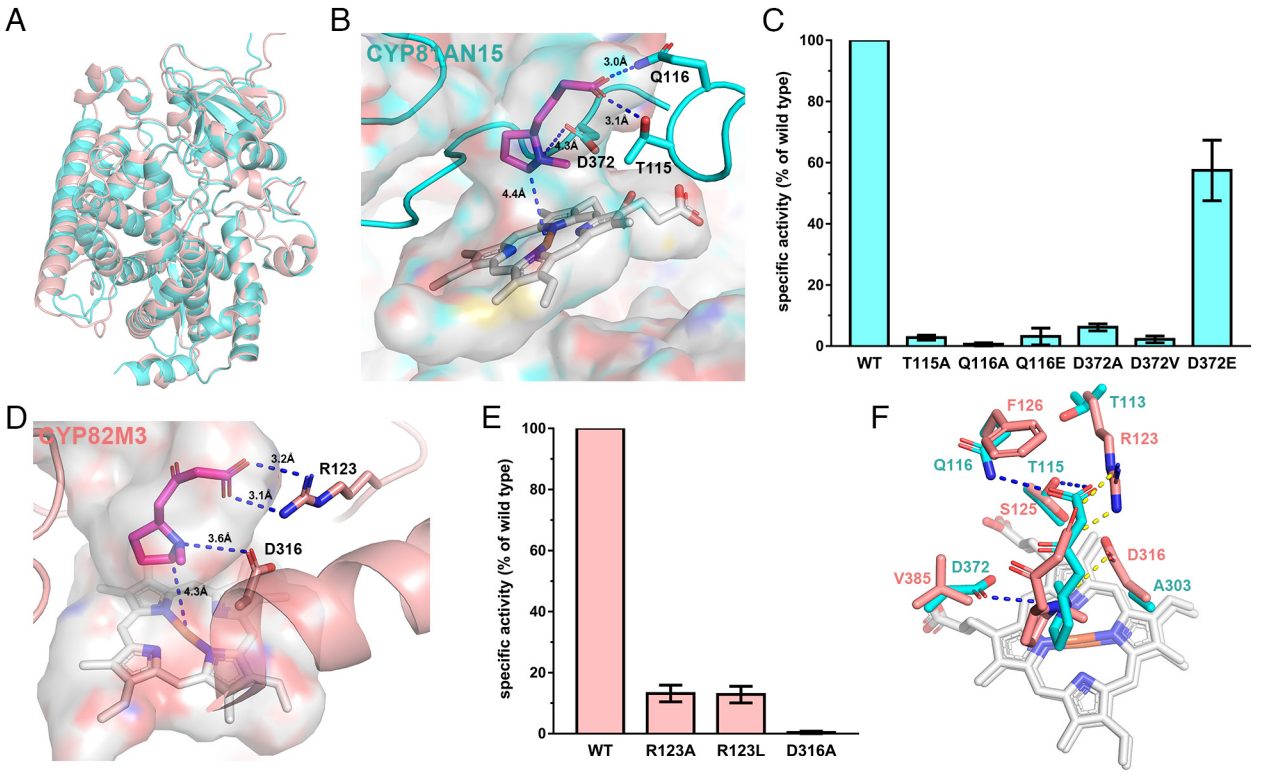

Fig. 4. Identification of crucial amino acid residues for ecgonone synthase activity of CYP81AN15 and CYP82M3. (*A*) Superposition of the predicted *En*CYP81AN15 (cyan) and *Ab*CYP82M3 (pink) structures. (*B*) Molecular docking of the MPOA intermediate with predicted *En*CYP81AN15. (*C*) Product of tropinone catalyzed by wild-type *En*CYP81AN15 or its mutants in *N. benthamiana*. Three biological duplications were performed for each assay. (*D*) Molecular docking of the MPOA intermediate with predicted *Ab*CYP82M3. (*E*) Product of tropinone catalyzed by wild-type *Ab*CYP82M3 or its mutants in vitro. Three biological duplications were performed for each assay. (*F*) Comparison of crucial-site architectures between *En*CYP81AN15 (cyan) and *Ab*CYP82M3 (pink). *N* and *O* atoms are colored blue and red, respectively. *C* atoms of MPOA in *En*CYP81AN15 and *Ab*CYP82M3 are colored cyan and pink, respectively.

plays a critical role in the position control of the pyrrole ring through salt bridge interaction with the *N* atom of the pyrrole ring. Accordingly, the interactions of MPOA with T115, Q116, and D372 in CYP81AN15 allow MPOA to be properly positioned and oxidized to form ecgonone in cocaine biosynthesis (Fig. 4*B*).

In CYP82M3, R123 and D316 were predicted to form salt bridges with the carboxylate and *N* atom of MPOA, respectively, in order to control the position of MPOA (Fig. 4*D*). As a result, its catalytic activity was completely abolished by the mutation of D316A (Fig. 4*E* and *SI Appendix*, Fig. S14). Large-to-small (R123A) and polar-to-nonpolar (R123L) substitutions of R123 resulted in 87.7% and 88.5% decrease in catalytic activity, respectively (Fig. 4*E* and *SI Appendix*, Fig. S14). These results suggested a vital role of the salt bridge interactions between D316, R123, and MPOA (Fig. 4*D*). In particular, the superposition of CYP81AN15 and CYP82M3 protein structures showed that the aspartic acid residues (i.e., D372 in CYP81AN15 and D316 in CYP82M3), interacting with the *N* atom of MPOA to control the position of pyrrole ring, are recruited from two opposite directions (Fig. 4*F*). The different active-site architectures in CYP81AN15 and CYP82M3 suggested independent recruitment of ecgonone synthases in cocaine and hyoscyamine biosynthesis.

Independent Evolution of Ecgonone Synthases CYP81AN15 and CYP82M3. To trace the evolutionary history of ecgonone synthase, we analyzed the chromosomal region surrounding *En*CYP81AN15. Seven homologous proteins adjacent to *En*CYP81AN15 were identified, while the conserved T115, Q116, and D372 residues only coexist in *En*CYP81AN15, indicating that functional *En*CYP81AN15 might be the evolutionary product of tandem replication (*SI Appendix*, Fig. S15). The Ks ratios of these tandem *EnCYP81AN15* paralogs were mainly between 1.1 and 1.7, which suggested that an ancient tandem duplication event occurred (*SI Appendix*, Table S5). In addition, comparative genomic analysis showed that *En*CYP81AN15 localized to a synteny block well conserved in eudicot plant genomes, and multiple adjacent *En*CYP81AN15 homologous genes were identified in this collinearity region of *En*CYP81AN15 in *Manihot esculenta*, *A. thaliana,* and grape (*SI Appendix*, Fig. S15 and Tables S6). Therefore, this tandem replication event likely took place before the divergence of eudicot plants and provided generous CYP81 family homologous genes for the evolution of *EnCYP81AN15*. Subsequently, the key sites of proteins encoded by *EnCYP81AN15* syntenic genes in eudicot plants were analyzed to trace the origin of functional *En*CYP81AN15. Aside from *En*CYP81AN15 in *E. novogranatense*, the key T115, Q116, and D372 residues only coexist in protein sequences from members in the Rhizophoraceae, suggesting that ancestral ecgonone synthase emerged in the last common ancestor between Erythroxylaceae and Rhizophoraceae. Thus, members of these two families' plants share a consistent strategy to biosynthesize the tropane skeleton.

*Aa*Contig2.1388, which is encoded by the syntenic gene of *EnCYP81AN15* in *A. acutangulus* (*SI Appendix*, Fig. S15), did not contain the key T115, Q116, and D372 sites, while CYP82M3 was recruited to be responsible for the biosynthesis of the tropane skeleton in Solanaceae. *Aa*CYP82M3 was only observed on well synteny blocks with other solanaceous species (*SI Appendix*, Fig. S16), suggesting independent recruitment of ecgonone synthase in Solanaceae.

The *E. novogranatense*-Specific Acquisition of Ecgonone Methyltransferase *En*MT4 in Erythroxylaceae. The existence of a 2-substituted carboxymethyl group in cocaine is attributed to

the participation of ecgonone methyltransferase *En*MT4 in *E. novogranatense* (Fig. 1) (17). To trace the origin of *En*MT4, we performed intergenomic comparative genomics analysis of the chromosomal region surrounding *EnMT4*. As a result, there were two significant chromosomal expansion regions (regions A and B) in *E. novogranatense* relative to *K. obovate* and *P. trichocarpa* (Fig. 5*A*). In these two chromosomal expansion regions, four *EnMT4* homologous genes (*EnMT1*, *EnMT3,* and *EnMT14* in region A and *EnMT2* in region B) were found. *EnMT1*, *EnMT3,* and *EnMT14* in region A were syntenic genes of *PtSAMT* (30) in *P. trichocarpa* (Fig. 5*A*). Subsequently, *En*MT1 and *En*MT3 were identified to catalyze the methylation of salicylic acid in vitro and were designated as SAMTs in *E. novogranatense* (*SI Appendix*, Fig. S17). In addition, phylogenetic analysis of SABATH proteins from Malpighiales species showed that *En*MT1–4, *En*MT14, and *Pt*SAMT were located on the same branch (*SI Appendix*, Fig. S18), suggesting that *EnMT4* in expansionary region B might have evolved from *SAMTs* in expansionary region A (Fig. 5*A*).

The analysis of conserved sites in ecgonone methyltransferase *En*MT4 provides insight regarding the evolution of *En*MT4 from SAMTs. We performed molecular docking with ecgonone in the predicted protein structure of *En*MT4. Q25 and W157 were supposed to orient the carboxylate of ecgonone (Fig. 5*B*), which is consistent with the function of conserved Gln and Trp in other SABATH family methyltransferases, such as SAMT (31), IAMT (32), FAMT (33), and LAMT (34). Q25 could form a hydrogen bond with the *N* atom of tropane ring (Fig. 5*B*). As a result, mutations of Q25A or W157A in *En*MT4 lost the activity for methylation of the carboxyl group in substrate (Fig. 5*C* and *SI Appendix*, Fig. S19*A*). For the substrate specificity of ecgonone, the hydroxyl group of tropane ring might be positioned by hydrogen bonds with E216 and S153 (Fig. 5*B*), while the corresponding amino acid residues were hydrophobic Leu and aromatic Tyr for substrate recognition of salicylic acid in SAMT (31). Consequently, the catalytic activities of *En*MT4 for methylation of ecgonone were abolished by the polar-to-nonpolar (E216L) substitution and small-to-large (S153Y) substitutions or significantly decreased by large-to-small (E216A) and polar-to-nonpolar (S153A) substitutions (Fig. 5*C* and *SI Appendix*, Fig. S19*B*). These results indicated that E216 and S153 perform critical roles in *En*MT4 catalysis of methylecgonone formation. In order to clarify that ecgonone methyltransferase *En*MT4 evolved from SAMT *En*MT1 and *En*MT3, we mutated *En*MT1 by adding the crucial amino acid residues E216 and S153 in order to position ecgonone correctly. As expected, mutation of L216E in *En*MT1 obtained the ability to methylate ecgonone (Fig. 5*C* and *SI Appendix*, Fig. S19*C*). In addition, a combination of the L216E and Y153S mutations in *En*MT1 led to a four-fold increase of methylation activity compared with the single mutation L216E in *En*MT1 (Fig. 5*C* and *SI Appendix*, Fig. S19*C*). Thus, these two amino acid substitutions in *En*MT1 enabled the methylation of ecgonone, providing strong evidence for the origin of ecgonone methyltransferase *En*MT4 from SAMTs.

To explore the driving force of *En*MT4 evolution from SAMTs, we analyzed the repeat sequences of these two chromosomal

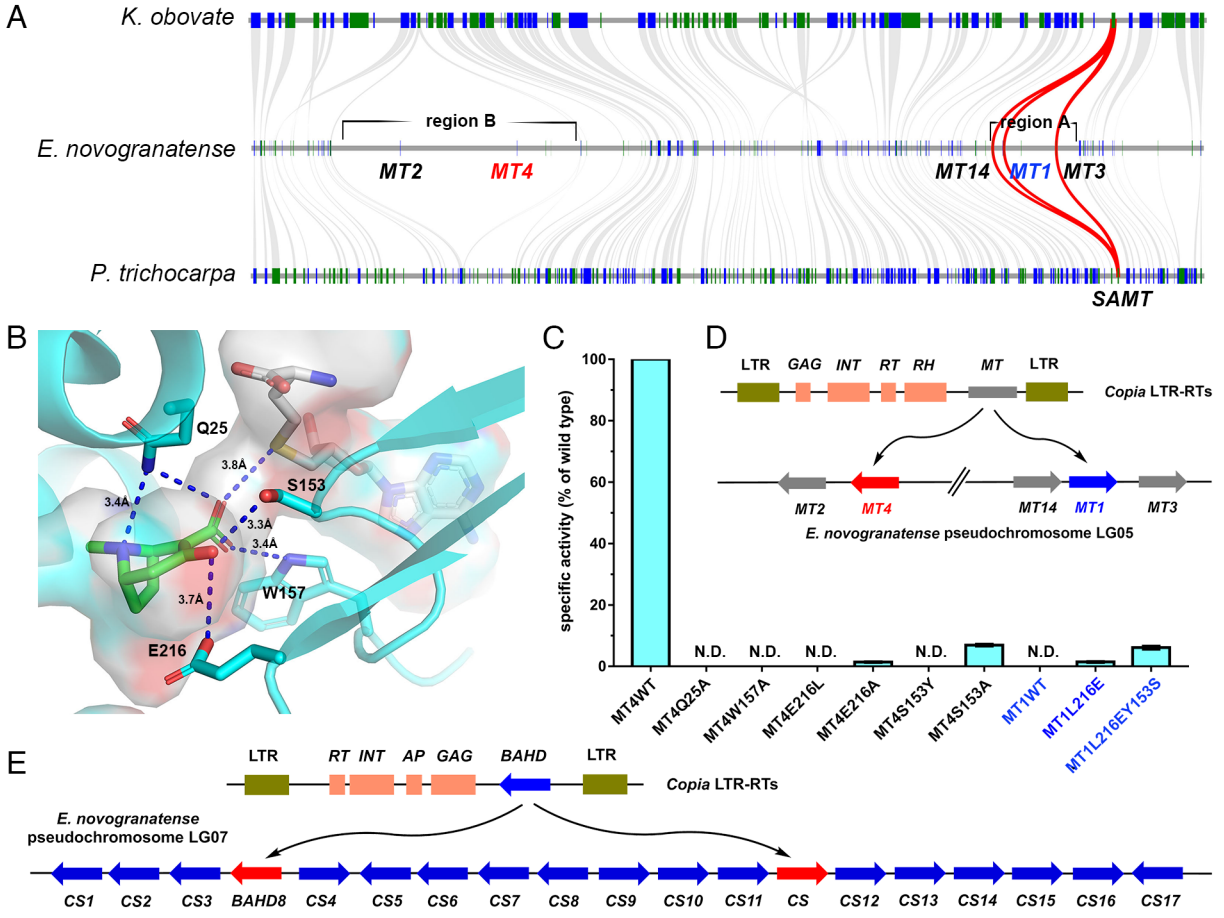

Fig. 5. The evolutions of *En*MT4 and *En*CS. (*A*) Synteny analysis of *EnMT1* and *EnMT4* in *E. novogranatense*, *K. obovate,* and *P. trichocarpa*. (*B*) Molecular docking of the ecgonone intermediate with predicted *En*MT4. *N* and *O* atoms are colored blue and red, respectively. *C* atoms in ecgonone and *En*MT4 residues are colored green and cyan, respectively. (*C*) Product of methylecgonone catalyzed by wild-type *En*MT4 and *En*MT1 or their mutants in vitro. Three biological duplications were performed for each assay. (*D*) The emergence of *EnMT4* in pseudochromosome LG05 mediated by *copia* LTR-RTs (long terminal repeat retrotransposons). (*E*) The emergence of *EnCS* and *EnBAHD8* in pseudochromosome LG07 mediated by *copia* LTR-RTs.

expansion regions (regions A and B) in pseudochromosome LG05, containing *EnMT1*–*EnMT4* and *EnMT14*. These five methyltransferase genes and their surrounding regions were annotated as *copia* LTR-RT-mediated repeat sequence Repeat_10570 (*SI Appendix*, Table S7), which is located on expansion region A. Density analysis of Repeat_10570 in pseudochromosome LG05 revealed that Repeat_10570 was enriched in these two chromosomal expansion regions (regions A and B) (Fig. 5*A* and *SI Appendix*, Fig. S20), suggesting that expansion regions A and B might be generated by *copia* LTR-RTs. Subsequently, a truncated SABATH-MT gene was found in Repeat_10570 (*SI Appendix*, Table S7), indicating that *copia* LTR-RT-mediated genome expansion probably provided abundant templates of inherent *SAMTs* for the evolution of *EnMT4* (Fig. 5*D*). In addition, the Ks ratios among *EnMT4* and its homologous genes were mainly around 0.15 (*SI Appendix*, Table S8), suggesting that this LTR-RT-mediated duplication event happened around 11.8 Mya, which was later than the divergence (60.6 Mya) of Erythroxylaceae and Rhizophoraceae (Fig. 2). Thus, these *E. novogranatense*-specific duplication events probably led to the evolution of *En*MT4 from SAMTs, partially explaining the structural differences in TAs observed between members of the Erythroxylaceae and Rhizophoraceae (*SI Appendix*, Fig. S21).

Independent Evolution of Postmodified Reductases and Acyltransferases. TRI (SDR family) and MecgoR (AKR family) are the enzymes responsible for the reduction of keto moiety on the tropane ring in hyoscyamine and cocaine biosynthetic pathways, respectively (Fig. 1) (6, 20). Synteny comparisons suggested that *TRI* was likely a Solanaceae-specific duplication (*SI Appendix*, Fig. S16). However, *MecgoR* is located on a synteny block well conserved in eudicot plant genomes (*SI Appendix*, Fig. S22). *Sl*MecgoR (*Sl*AKR4B) in tomato has been found to catalyze the reduction of glyceraldehyde, methylglyoxal, or glyoxal (35). Nevertheless, the expression of both *SlMecgoR* and *AaMecgoR* was largely undetectable in roots (23), suggesting a minimal contribution by MecgoR to tropinone reduction in Solanaceae.

Littorine synthase, a member of the SCPL-AT family, catalyzes the esterification of tropane ring (Fig. 1) (21), a reaction critical for shunting the TA precursor skeleton into hyoscyamine and scopolamine. In agreement with this function, we only found syntenic orthologs of *AaLS* in the genome of hyoscyamine-producing *D. stramonium* (29) among all of the available solanaceous genomes (*SI Appendix*, Fig. S23). In cocaine pathway, BAHD acyltransferase CS performs a similar catalytic function to that of LS (Fig. 1) (19, 21). At the genome-wide level, no genes with high predicted sequence similarity to *EnCS* were detected in close relatives of *E. novogranatense* examined here. By comparative genomic analysis of the chromosomal region surrounding *EnCS* (*SI Appendix*, Fig. S24), 18 species-specific tandem BAHD acetyltransferase genes adjacent to *EnCS* were identified (Fig. 5*E* and *SI Appendix*, Table S9). Among these, the *EnBAHD8* gene product was reportedly capable of catalyzing the same reaction as *En*CS (19). Further intragenomic sequence homology investigation implied a probable mechanism by which the *CS* gene emerged in Erythroxylaceae (Fig. 5*E*): An LTR-RT-mediated genome expansion event resulted in tandem BAHD genes on pseudochromosome LG07, among which *EnBAHD8* and *EnCS* evolved for cocaine biosynthesis (*SI Appendix*, Table S10) (19).

Conclusion

The absence of high-quality genome sequences of medicinal TA-producing plants has hindered the understanding of TA biosynthesis and evolution. In this study, we generated high-quality chromosome-level genome assemblies of TA-producing *E. novogranatense* and *A. acutangulus*, thus providing a reliable means of comparing the genetic evolutionary trajectories of TA biosynthetic pathways in these species that diverged as early as 129.4 Mya (Fig. 2). This early differentiation point provided sufficient time for the independent evolution of cocaine and hyoscyamine biosynthesis through gene multiplication, gene loss, tandem duplication, segmental duplication, and key site mutation.

Methylation of putrescine or spermidine catalyzed by SPMT or PMT is the first specific step in cocaine and hyoscyamine biosynthesis, respectively (Fig. 1). Comparative genomic analysis suggested that the shared hexaploidy (or γ) event in the core eudicots prompted the occurrence of SPMT from SPDS (Fig. 3*A*). Despite the lack of SPMT in the asterids, the PMT responsible for the methylation of putrescine probably evolved from tandem copies of SPDS, again in the ancestors of solanaceous species (24). The following specific polyketide synthase (PYKS) and ecgonone synthase (CYP82M3) in Solanaceae were recruited to biosynthesize the tropane skeleton of hyoscyamine (27, 28). Although these PYKS and CYP82M3 are not shared with Erythroxylaceae and Rhizophoraceae, the independent evolution of polyketide synthase (PKS1/2) and ecgonone synthase (CYP81AN15) with different active-site architectures in their common ancestor promoted the emergence of tropane alkaloids in closely related Erythroxylaceae and Rhizophoraceae plants (5). After their species divergence, ecgonone methyltransferase (*En*MT4) evolved via Erythroxylaceae-specific segmental duplication and subsequent mutations of key E216 and S153 residues, thus stabilizing the carboxyl group of ecgonone and directing the biosynthetic flux to ecgonone derivatives (i.e., cocaine and cinnamoylcocaine). Aside from the early tandem duplication for CYP81AN15 evolution, the recent LTR-RT-mediated duplications probably provided numerous templates for the evolution of ecgonone methyltransferase (*En*MT4) and cocaine synthase (CS), which drove the lineage-specific emergence of cocaine in Erythroxylaceae. Overall, the analysis of evolutionary origins and trajectories of tropane skeleton biosynthesis on the genomic and structural levels deepen our understanding of the occurrence of lineage-specific tropane alkaloids in Solanaceae, Erythroxylaceae, and Rhizophoraceae.

In summary, a set of enzymes involved in the biosynthesis of scattered TAs potentially emerged via 1) the evolution of SPMT from SPDS through hexaploidy event in core eudicot species and the emergence of PMT from SPDS once again in the Solanales, 2) parallel recruitment of polyketide synthases and ecgonone synthases, 3) Erythroxylaceae-specific evolution of ecgonone methyltransferase which directs the biosynthetic flux to ecgonone derivatives, and 4) the lineage-specific recruitment of reductases and acyltransferases from distinct families for the postmodification of the tropane skeleton in cocaine and hyoscyamine. This rare example of chemotypic convergence in phylogenetically distant species highlights the remarkable structural and functional flexibility of secondary metabolite enzymes through parallel/convergent evolution, which will accelerate the discovery of more enzymes with diverse catalytic mechanisms.

Materials and Methods

Plant Material and Genome Sequencing. *E. novogranatense* was cultivated in Xishuangbanna Tropical Botanical Garden. *A. acutangulus* and *N. benthamiana* plants were grown at the Kunming Institute of Botany. Young leaves of *E. novogranatense* and *A. acutangulus* were used for genomic DNA extraction following the modified cetyltrimethylammonium bromide (CTAB) protocol (36). The genome sequencing, assembly, and annotation process were described in *SI Appendix, Materials and Methods*.

The Construction of Phylogenetic Trees and Comparative Genomic Analysis of Functional Genes. Orthologous gene families were clustered from proteins of 11 plants from different families by OrthoFinder (37). Protein sequences of 108 single-copy orthologous families were aligned by MAFFT (38) and were then corrected by Gblocks (39). The resulting sequences were used to construct the maximum likelihood tree by IQ-TREE (40) using the best-fit model JTT+F+R4 according to BIC. The divergence time was estimated by MCMCtree program of the PAML package (41). The WGD events were analyzed by MCScanX (42). Ks values of orthologous or paralogous gene pairs were calculated by the Simple Ka/Ks calculator program of TBtools package (43). To time the *E. novogranatense* WGD event, the molecular clock rate (*r*) was calculated to be 6.35E-9 substitutions per synonymous site per year, according to the mean Ks value (0.77) of *E. novogranatense-K. obovate* and their divergent time (60.6 Mya) (Fig. 2) (44). Then, we estimated *E. novogranatense* WGD event happened approximately 30.5 to 40.3 Mya (Ks = 0.45 ± 0.062, T = Ks/2*r*). In the same way, the recent WGD event in *A. acutangulus* happened approximately 7.4 to 22.2 Mya (Ks = 0.12 ± 0.060), according to the mean Ks value (0.22) of *A. acutangulus–S. lycopersicum* and their divergent time (27.1 Mya) (Fig. 2). Intergenomic and intragenomic synteny analyses of genomic regions near functionally characterized genes were performed by MCScanX (42).

In Vitro Characterization of *Aa*PYKS2 with Malonyl-CoA. The standard assay was carried out in a 100-μL mixture which contained potassium phosphate buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8.0), 0.5 mM malonyl-CoA, and 300 ng μL^{-1} *Aa*PYKS2 protein. After incubation at 30 °C for 1 h, reactions were quenched by adding 10 μL 20 % HCl and were then analyzed by HPLC with a YMC-Triart C_{18} column (I.D. 4.6 mm × 250 mm) using water with 0.1% formic acid as solvent A and methanol with 0.1% formic acid as solvent B. The injections were eluted with 5% B for 10 min, with a flow rate of 1.0 mL/min. The products were detected at 242 nm.

In Vitro Characterization of *En*MT1–4 with Salicylic Acid. The assays were carried out in a 200-μL mixture containing potassium phosphate buffer (50 mM K_2HPO_4/KH_2PO_4, pH 7.1), 1 mM salicylic acid, 1 mM SAM, and 100 ng μL^{-1} enzymes. Assay which lacked enzyme served as a negative control (blank). After incubation at 30 °C for 1 h, reactions were quenched by adding 200 μL acetonitrile and were then analyzed by HPLC with a YMC-Triart C_{18} column (I.D. 4.6 mm × 250 mm) using water with 0.1% formic acid as solvent A and acetonitrile as solvent B. The injections were eluted with a 28-min gradient mobile phase program (0.0 to 20.0 min, 10% B to 100% B; 20.0 to 24.0 min, 100% B; 24.1 to 28.0 min, 10% B), with a flow rate of 1.0 mL/min. The products were detected at 305 nm.

Homology Modeling and Molecular Docking. The putative protein structures of *En*MT4, *En*CYP81AN15, and *Ab*CYP82M3 were predicted by AlphaFold2. Molecular dockings of *En*MT4 with (+)-ecgonone, as well as *En*CYP81AN15 and *Ab*CYP82M3 with *S*-MPOA, were performed using AutoDock Vina 1.1.2. The ligands (+)-ecgonone and *S*-MPOA were energy minimized with Molecular Mechanics (MM2) using ChemBio3D Ultra 14.0. The key parameters of algorithm and grid number were set as default. The complex structures with favorable orientation and low binding energy were selected for subsequent analysis. The interactions between ligands and proteins were visualized by PyMOL 2.4 (https://pymol.org/).

In Vitro Characterization of Wild and Mutated EnMT4 with Ecgonone. The assays were carried out in a 100-μL mixture containing potassium phosphate buffer (50 mM K_2HPO_4/KH_2PO_4, pH 8.0), 500 μM ecgonone, 1 mM SAM, and 5 ng μL^{-1} enzymes. After incubation at 35 °C for 10 min, the reactions were quenched by adding 100 μL acetonitrile and were then analyzed by LC-MS with a YMC-Triart C_{18} column (I.D. 4.6 mm × 250 mm) using water with 0.1% formic acid as solvent A and acetonitrile as solvent B. The injections were eluted with 10% B for 6 min, with a flow rate of 1.0 mL/min. The MS data were collected using electron spray ionization in a positive ion mode (mass range: 50 to 400 m/z). The assays were carried out in triplicate, and the specific activities of mutated *En*MT4 were calculated based on the peak area ratios of methylecgonone between mutated and wild *En*MT4.

Functional Characterization of Mutated *En*CYP81AN15 in *N. benthamiana*. The PCR products of *En*CYP81AN15 mutants were amplified from the wild *En*CYP81AN15–pEAQ-HT plasmid (17). These PCR products were assembled into a linearized pEAQ-HT plasmid, which was digested by *Age*I and *Xho*I restriction enzymes. The assembly mixtures were transformed into chemically competent *E. coli* DH5α. The resulting constructs were transformed into *Agrobacterium tumefaciens* LBA4404 using the freeze−thaw method. The obtained strains were cultured on LB plates for 3 d at 30 °C. After the strains were suspended in MMA buffer (10 mM MES, 10 mM $MgCl_2$, 150 μM acetosyringone, pH = 5.6), these suspensions (OD_{600} = 0.3) were incubated at room temperature for 40 min and then infiltrated into the leaves of 4 to 6-week-old *N. benthamiana*. After 3 d, *Agrobacterium*-infiltrated leaves were infiltrated with 500 μM deuterated MPOA solution. The resulting leaves were harvested 1 d later and were lyophilized to determine the dry weight of the leaves. After being ground in liquid nitrogen, the pulverized leaves were extracted with CH_2Cl_2. The resulting organic layers were extracted with hydrochloric acid aqueous solution (pH = 2.0). The obtained aqueous layers were adjusted to pH 9.0 with 1 M NaOH and then reextracted with CH_2Cl_2. The organic layers were evaporated in vacuo and dissolved in acetonitrile (2 μL acetonitrile per milligram of dry tissues). The resulting solutions were analyzed by LC-MS (conditions were consistent with the assay of mutated *En*MT4 with ecgonone in vitro). Each mutated *En*CYP81AN15 experiment was carried out in triplicate, and the specific activities of mutated *En*CYP81AN15 were calculated based on the peak area ratios of deuterated tropinone between mutated and wild *En*CYP81AN15.

In Vitro Characterization of Wild and Mutated *Ab*CYP82M3 with MPOA. The assays were carried out in a 100-μL mixture containing potassium phosphate buffer (50 mM K_2HPO_4/KH_2PO_4, pH 7.5), 100 μM MPOA, 1 mM NADPH, and 10 μL microsomal proteins. After incubation at 30 °C for 30 min, the reactions were quenched by adding 100 μL acetonitrile and were then analyzed by LC-MS (conditions were consistent with the assay of mutated *En*MT4 with ecgonone in vitro). The assays were carried out in triplicate, and the specific activities of mutated *Ab*CYP82M3 were calculated based on the peak area ratios of tropinone between mutated and wild *Ab*CYP82M3.

Data, Materials, and Software Availability. The *E. novogranatense* and *A. acutangulus* genome projects have been deposited at the NCBI under the BioProject numbers PRJNA749480 (45) and PRJNA752913 (46), respectively. The fifteen transcriptome datasets of *A. acutangulus* have accession numbers SRR15559830 (47), SRR15559831 (48), SRR15559832 (49), SRR15559833 (50), SRR15559834 (51), SRR15559835 (52), SRR15559836 (53), SRR15559837 (54), SRR15559838 (55), SRR15559839 (56), SRR15559840 (57), SRR15559841 (58), SRR15559842 (59), SRR15559843 (60), and SRR15559844 (61). The GenBank accession numbers for *Aa*PYKS2 are MZ819700 (62). All data that support the findings of this study are available in the main text and the *SI Appendix*.

ACKNOWLEDGMENTS. We thank Prof. Zhenhua Liu (Shanghai Jiao Tong University) for informative discussions. We thank Profs. Yi Shang (Yunnan Normal University) and Qinggang Liao (Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences) for processing the gene annotation data. This research was supported by the National Key R&D Program of China (2018YFA0900600), the National Natural Science Foundation of China (Nos. 82225043, U1902212, and 32000239), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB27020205), and Yunnan Revitalization Talent Support Program "Yunling Scholar" Project.

Author affiliations: [a]State Key Laboratory of Phytochemistry and Plant Resources in West China and Yunnan Key Laboratory of Natural Medicinal Chemistry, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; [b]University of the Chinese Academy of Sciences, Beijing 100049, China; [c]Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming 650223, China; and [d]Department of Molecular Genetics, Leibniz Institute for Plant Genetics and Crop Plant Research Ortsteil Gatersleben, Seeland D-06466, Germany

1. K. Springob, T. M. Kutchan, "Introduction to the different classes of natural products" in *Plant-derived Natural Products*, A. Osbourn, V. Lanzotti, Eds. (Springer, 2009), pp. 3–50.
2. W. J. Griffin, G. D. Lin, Chemotaxonomy and geographical distribution of tropane alkaloids. *Phytochemistry* **53**, 623–637 (2000).
3. B. Christ *et al.*, Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants. *Nat. Commun.* **10**, 3206 (2019).
4. B. R. Lichman *et al.*, The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* **6**, eaba0721 (2020).

5. T. Tian *et al.*, Catalytic innovation underlies independent recruitment of polyketide synthases in cocaine and hyoscyamine biosynthesis. *Nat. Commun.* **13**, 4994 (2022).
6. J. Jirschitzka *et al.*, Plant tropane alkaloid biosynthesis evolved independently in the Solanaceae and Erythroxylaceae. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10304–10309 (2012).
7. J.-K. Weng, The evolutionary paths towards complexity: A metabolic perspective. *New Phytol.* **201**, 1141–1149 (2014).
8. J. Kroymann, Natural diversity and adaptation in plant secondary metabolism. *Curr. Opin. Plant Biol.* **14**, 246–251 (2011).
9. G. C. Conant, K. H. Wolfe, Turning a hobby into a job: How duplicated genes find new functions. *Nat. Rev. Genet.* **9**, 938–950 (2008).
10. A. Junker, J. Fischer, Y. Sichhart, W. Brandt, B. Drager, Evolution of the key alkaloid enzyme putrescine *N*-methyltransferase from spermidine synthase. *Front. Plant Sci.* **4**, 260 (2013).
11. Q. Zhao *et al.*, The reference genome sequence of *Scutellaria baicalensis* provides insights into the evolution of wogonin biosynthesis. *Mol. Plant.* **12**, 935–950 (2019).
12. P. K. Sahu, S. P. Pradhan, P. S. Kumar, "Isolation, elucidation, and structure–activity relationships of phytoalkaloids from Solanaceae" in *Studies in Natural Products Chemistry*, Atta-ur-Rahman, Eds. (Elsevier, 2022). pp. 371–389.
13. M. Lounasmaa, T. Tamminen, "The tropane alkaloids" in *The Alkaloids: Chemistry and Pharmacology*, G. A. Cordell, Eds. (Elsevier, 1993). pp. 1–114.
14. G. Grynkiewicz, M. Gadzikowska, Tropane alkaloids as medicinally useful natural products and their synthetic derivatives as new drugs. *Pharmacol. Rep.* **60**, 439–463 (2008).
15. P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* **585**, 614–619 (2020).
16. J.-P. Huang *et al.*, Tropane alkaloid biosynthesis: A centennial review. *Nat. Prod. Rep.* **38**, 1634–1658 (2021).
17. Y.-J. Wang *et al.*, Discovery and engineering of the cocaine biosynthetic pathway. *J. Am. Chem. Soc.* **144**, 22000–22007 (2022).
18. B. G. Chavez *et al.*, Elucidation of tropane alkaloid biosynthesis in *Erythroxylum coca* using a microbial pathway discovery platform. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2215372119 (2022).
19. G. W. Schmidt *et al.*, The last step in cocaine biosynthesis is catalyzed by a BAHD acyltransferase. *Plant Physiol.* **167**, 89–101 (2015).
20. K. Nakajima, T. Hashimoto, Y. Yamada, Two tropinone reductases with different stereospecificities are short-chain dehydrogenases evolved from a common ancestor. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9591–9595 (1993).
21. F. Qiu *et al.*, Functional genomics analysis reveals two novel genes required for littorine biosynthesis. *New Phytol.* **225**, 1906–1914 (2020).
22. C. S. P. Foster *et al.*, Evaluating the impact of genomic data and priors on bayesian estimates of the angiosperm evolutionary timescale. *Syst. Biol.* **66**, 338–351 (2017).
23. S. Sato *et al.*, The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).
24. S. Xu *et al.*, Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6133–6138 (2017).
25. O. Jaillon *et al.*, The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467 (2007).
26. M. Naconsie, K. Kato, T. Shoji, T. Hashimoto, Molecular evolution of *N*-methylputrescine oxidase in tobacco. *Plant Cell Physiol.* **55**, 436–444 (2014).
27. J.-P. Huang *et al.*, Tropane alkaloids biosynthesis involves an unusual type III polyketide synthase and non-enzymatic condensation. *Nat. Commun.* **10**, 4036 (2019).
28. M. A. Bedewitz, A. D. Jones, J. C. D'Auria, C. S. Barry, Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. *Nat. Commun.* **9**, 5281 (2018).
29. I. M. De-la-Cruz *et al.*, Genomic signatures of the evolution of defence against its natural enemies in the poisonous and medicinal plant *Datura stramonium* (Solanaceae). *Sci. Rep.* **11**, 882 (2021).
30. X.-M. Han *et al.*, Evolution and function of the *Populus* SABATH family reveal that a single amino acid change results in a substrate switch. *Plant Cell Physiol.* **59**, 392–403 (2018).
31. C. Zubieta *et al.*, Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. *Plant Cell* **15**, 1704–1716 (2003).
32. N. Zhao *et al.*, Structural, biochemical, and phylogenetic analyses suggest that indole-3-acetic acid methyltransferase is an evolutionarily ancient member of the SABATH family. *Plant Physiol.* **146**, 455–467 (2008).
33. N. Petronikolou, S. K. Nair, Biochemical studies of mycobacterial fatty acid methyltransferase: A catalyst for the enzymatic production of biodiesel. *Chem. Biol.* **22**, 1480–1490 (2015).
34. N. Petronikolou, A. J. Hollatz, M. A. Schuler, S. K. Nair, Loganic acid methyltransferase: Insights into the specificity of methylation on an iridoid glycoside. *Chembiochem* **19**, 784–788 (2018).
35. M. Suekawa, Y. Fujikawa, S. Inada, A. Murano, M. Esaka, Gene expression and promoter analysis of a novel tomato aldo-keto reductase in response to environmental stresses. *J. Plant Physiol.* **200**, 35–44 (2016).
36. S. Porebski, L. G. Bailey, B. R. Baum, Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rep.* **15**, 8–15 (1997).
37. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
38. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
39. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).
40. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
41. Z. H. Yang, PAML: A program package for phylogenetic analysis by Maximum Likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
42. Y. Wang *et al.*, MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
43. C. Chen *et al.*, TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* **13**, 1194–1202 (2020).
44. L. Guo *et al.*, The opium poppy genome and morphinan production. *Science* **362**, 343–347 (2018).
45. Y.-J. Wang. *Erythroxylum novogranatense* genome sequencing and assembly. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA749480. Deposited 25 July 2021.
46. Y.-J. Wang. *Anisodus acutangulus* genome sequencing and assembly. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA752913. Deposited 9 August 2021.
47. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* main root1. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559830. Deposited 22 August 2021.
48. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* stem3. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559831. Deposited 22 August 2021.
49. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* stem2. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559832. Deposited 22 August 2021.
50. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* stem1. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559833. Deposited 22 August 2021.
51. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* leaf3. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559834. Deposited 22 August 2021.
52. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* leaf2. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559835. Deposited 22 August 2021.
53. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* leaf1. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559836. Deposited 22 August 2021.
54. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* bud3. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559837. Deposited 22 August 2021.
55. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* lateral root3. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559838. Deposited 22 August 2021.
56. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* lateral root2. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559839. Deposited 22 August 2021.
57. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* lateral root1. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559840. Deposited 22 August 2021.
58. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* main root3. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559841. Deposited 22 August 2021.
59. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* main root2. NCBI SRA *Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559842. Deposited 22 August 2021.
60. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* bud2. *NCBI SRA Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559843. Deposited 22 August 2021.
61. Y.-J. Wang. *RNA*-seq of *Anisodus acutangulus* bud1. *NCBI SRA Database*. https://www.ncbi.nlm.nih.gov/sra/SRR15559844. Deposited 22 August 2021.
62. Y.-J. Wang *et al.*, *Anisodus acutangulus*. NCBI Nucleotide *Database*. https://www.ncbi.nlm.nih.gov/nuccore/MZ819700. Deposited 13 August 2021.