



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/236519/>

Version: Accepted Version

Article:

GROGAN, GIDEON JAMES, WOOD, CATHARINE, Jiang, Zhouqian et al. (2026) Parallel evolution of plant alkaloid biosynthesis from bacterial-like decarboxylases. *New Phytologist*. ISSN: 1469-8137

<https://doi.org/10.1111/nph.70884>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Parallel evolution of plant alkaloid biosynthesis from bacterial-like decarboxylases

Catharine X. Wood^{1,2,4°}, Zhouqian Jiang^{1,2°}, Inesh Amarnath¹, Lachlan J. N. Waddell^{1,2}, Uma Sophia Batey¹, Oriana Serna Daza¹, Katherine Newling³, Sally James³, Gideon Grogan², William P. Unsworth² and Benjamin R. Lichman^{1*}

¹Centre for Novel Agricultural Products, Department of Biology, University of York, York, YO10 5DD UK

²Department of Chemistry, University of York, York, YO10 5DD UK

³Biosciences Technology Facility, Department of Biology, University of York, York, UK.

⁴Current address: University of Wisconsin, Madison, Department of Botany, 430 Lincoln Drive, WI 53706 USA

[°]These authors contributed equally to manuscript.

*Correspondence to: benjamin.lichman@york.ac.uk

Abstract

Alkaloids are nitrogen-containing natural products derived from amino acids¹. The basic amino acids lysine and ornithine are precursors to a wide range of alkaloids including the bioactive compounds nicotine, hyoscyamine and securinine. Isotope feeding experiments have shown that the amino acids can be incorporated into alkaloids in a symmetric or nonsymmetric manner². The symmetric pathway is catalysed by two enzymes, a decarboxylase and oxidase^{3,4}, forming a cyclic iminium which acts as the electrophile in the scaffold forming step¹. Here, we describe the ornithine/lysine/arginine decarboxylase-oxidases (OLADOs), PLP-dependent enzymes responsible for the nonsymmetric pathway, catalysing the single step decarboxylative oxidative deamination of lysine, ornithine or arginine. These enzymes are part of the group III ornithine/lysine/arginine decarboxylase-like family (OLADLs), previously exclusively associated with prokaryotes⁵. We reveal OLADLs to be widespread in plants and show that OLADOs have repeatedly emerged through parallel evolution from OLADLs, *via* similar active site substitutions. This investigation introduces a new class of eukaryotic decarboxylases, and describes enzymes involved in multiple alkaloid biosynthesis pathways. It furthermore demonstrates how the principle of parallel evolution⁶ at a genomic and enzymatic level can be leveraged for gene discovery across multiple lineages.

Introduction

Plant alkaloids are nitrogen-containing specialised metabolites derived from amino acids. They are notable for their complex polycyclic structures, significant bioactivities, and extensive roles as pharmaceuticals, components of traditional remedies and intoxicants. The basic amino acid lysine gives rise to various alkaloid classes, including piperidine alkaloids (anabasine, *Nicotiana sp*)⁷, *Lycopodium* alkaloids (huperzine A, *Huperzia serrata*)⁸, and quinolizidines (lupinine, *Lupinus sp*) (**Fig. 1A**)⁹. The non-proteogenic basic amino acid ornithine contributes to pyrrolidine alkaloids including nicotine and the tropane alkaloids (e.g. hyoscyamine) found in Solanaceae species (**Fig. 1B**). Elucidating the biosynthesis of alkaloids can enable the formation of valuable pharmaceuticals in heterologous systems^{10,11} or the elimination of alkaloid biosynthesis in host tissues^{9,12}.

Typical alkaloid biosynthesis pathways involve a scaffold-forming step wherein an iminium electrophile, derived from an amino acid, reacts with a nucleophile *via* a Mannich-like reaction¹. For lysine and ornithine derived alkaloids the electrophiles are Δ^1 -piperideine or 1-pyrroline (or methylated derivatives) which form from the amino acids *via* decarboxylation, oxidation and cyclisation (**Fig. 1C** and **D**). The feeding of isotopically labelled precursors to plants revealed that the formation of Δ^1 -piperideine or 1-pyrroline could occur either in a symmetric or nonsymmetric manner, with the former indicating the presence of a free intermediate (e.g. cadaverine or putrescine), whereas the latter was proposed to proceed *via* an enzyme bound intermediate¹³.

The enzymatic origins of symmetric Δ^1 -piperideine from the *Lycopodium*⁴ and quinolizidine¹⁴ alkaloid families have been determined, and involve two enzymes: a PLP-dependent lysine decarboxylase (LDC) catalysing the formation of cadaverine from lysine, followed by a copper amine oxidase (CAO) catalysing the oxidation of cadaverine into 5-aminopentanal, which spontaneously cyclises into Δ^1 -piperideine (**Fig. 1C**). Analogous pathways from ornithine to symmetric pyrrolidine or tropane alkaloids in Solanaceae including *Atropa belladonna*¹⁵ and *Nicotiana*¹⁶ have been proposed, *via* putrescine, though include a methylation step prior to oxidation (**Fig. 1D**)^{17,18}.

Piperidine alkaloids anabasine¹⁹ and sedamine (*Sedum acre*)¹³ are known to have nonsymmetric origins, as does the *Securinega* alkaloid securinine, a GABA-receptor antagonist from the medicinal plant *Flueggea suffruticosa* (**Fig. 1A**)²⁰. There is also evidence of nonsymmetric labelling in tropane alkaloids from *Datura stramonium*²¹. In 1973, Leistner and Spenser hypothesised that a PLP-dependent oxidative enzyme could be responsible for the nonsymmetric pathway to Δ^1 -piperideine¹³, but to date no enzymes responsible for nonsymmetric alkaloid biosynthesis have been described.

Parallel evolution, a subset of convergent evolution in which new activities arise independently from within the same enzyme family, is a common phenomenon in plant specialised metabolism, including alkaloid biosynthesis^{6,22}. For example, piperidine

alkaloid related LDCs have evolved multiple times independently from ornithine decarboxylases (ODCs), members of the group IV PLP-dependent decarboxylase family⁴. Homospermidine synthase, a key enzyme in pyrrolizidine alkaloid biosynthesis, has evolved on multiple occasions through duplication of the essential enzyme deoxyhypusine synthase²³. Nonsymmetric lysine/ornithine derived alkaloid biosynthesis is present in the *Sedum*¹³, *Flueggea*²⁰ and *Nicotiana*¹⁹ genera (**Fig. 1A** and **B**), all of which are in different orders (Saxifragales, Rosids and Asterids respectively), so a common origin of nonsymmetric biosynthesis is unlikely. Therefore, if one enzyme responsible for nonsymmetric Δ^1 -piperidine or 1-pyrrolidine formation were to be characterised, it may be possible to use the principle of parallel evolution to identify functionally equivalent enzymes from other lineages.

Our investigation began with the goal of identifying the early steps of securinine biosynthesis in *F. suffruticosa*, as part of a broader effort to elucidate the pathway of the *Securinega* alkaloids²⁴. In doing so, we discovered an enzyme (*Fs1864*) capable of catalysing the nonsymmetric oxidative formation of Δ^1 -piperidine from lysine (**Fig. 1C**). This enzyme is a member of the PLP-dependent group III ornithine/lysine/arginine decarboxylase-like (OLADL) family, which is typically associated with prokaryotes. We showed that *Fs1864* has ornithine/lysine/arginine decarboxylating deaminative oxidase (OLADO) activity whereas a paralog (*Fs3498*) has low activity for both decarboxylation and oxidation. Using mutagenesis, we identify key amino acids contributing to the substrate acceptance and oxidation activity. The expectation that other nonsymmetric alkaloid pathways emerged *via* parallel evolution of OLADO enzymes led to the identification and characterisation of equivalent pairs of OLADL/OLADO enzymes in the alkaloid producing *Nicotiana tabacum* and in *Artemisia annua*, which is not yet known to produce alkaloids.

This work resolves the longstanding question of nonsymmetric piperidine alkaloid biosynthesis, and in doing so also (i) describes key enzymes in securinine and anabasine biosynthesis, (ii) reveals a new class of eukaryotic bacterial-like group III decarboxylases (OLADLs), (iii) demonstrates novel OLADO activities, and (iv) highlights how the expectation of parallel evolution can be used for gene discovery.

Results

Identification of Δ^1 -piperideine forming enzyme

We set out to discover a lysine decarboxylase (LDC) from *F. suffruticosa* that could catalyse the first step in the formation of Δ^1 -piperideine that contributes the piperidine moiety to securinine. First, we generated a *F. suffruticosa* transcriptomic resource from 15 different tissues and developmental stages (Table S1), combining short and long-read sequencing to assemble a *de novo* transcriptome (Table S2, Fig. S1). Using functional annotation, we identified five transcript candidates identified as members of ornithine decarboxylase families (Table S3). This term was selected as plant LDCs are known to be closely related to ornithine decarboxylase⁴. Two candidates were filtered out due to apparent transcriptome misassembly (gene 25968, Fig. S2) and low expression (gene 3498, Fig. S3).

We cloned the three remaining candidates from *F. suffruticosa* (Table S4) and, using *Nicotiana benthamiana* transient expression leaf disc assays, we tested their activity with lysine. Using LC-MS, we observed cadaverine formation in the positive control (*Lupinus angustifolius* LDC, LaLDC)⁴ and gene 4984 (Fig. S4). For gene 1864, we saw consumption of lysine, no accumulation of cadaverine, but accumulation of Δ^1 -piperideine, suggesting direct conversion of lysine to Δ^1 -piperideine.

In vitro validation of non-symmetric Δ^1 -piperidiene formation from lysine

We aimed to verify the activity of the enzyme encoded by *F. suffruticosa* gene 1864 (*Fs1864*) through *in vitro* assays, using recombinant protein produced in *Escherichia coli*. Initial attempts to express the full-length protein did not result in soluble protein. The AlphaFold structural model revealed an unstructured N-terminus that may have been problematic for expression or stability (Fig. S5A). Consequently, we truncated 74 amino acids from the N-terminus and were able to achieve soluble expression and purify the protein (Fig. S5B, Table S5). Incubation of the purified enzyme with lysine and PLP led to the formation of cadaverine as well as Δ^1 -piperideine (Fig. 2A). Δ^1 -Piperideine is a reactive compound, capable of forming dimers and trimers within solution in a pH and solvent dependent manner that are not completely interchangeable^{1,25}. LC-MS analysis of the enzyme product reveals monomer, dimer and trimer Δ^1 -piperideine peaks whose proportions do not correspond directly to those observed in the chemically synthesised standard (Fig. 2B). This means that quantification of Δ^1 -piperideine in solution is not possible; instead, we report the largest Δ^1 -piperideine monomer peak area unless noted otherwise.

Coupling the incubation of lysine, PLP and *Fs1864* to a peroxidase assay showed H₂O₂ accumulated over the course of the reaction (Fig. S5C), which was enhanced by supplemental PLP co-factor (Fig. S5D). Production of H₂O₂ is a typical byproduct of oxidative PLP-dependent enzymes that use molecular oxygen as a co-substrate²⁶. Further verification of PLP-dependence was obtained by incubating the enzyme with

the PLP-inhibitor semicarbazide prior to the reaction which led to reduction of both cadaverine and Δ^1 -piperideine production and was partially rescued by PLP cofactor supplementation (**Fig. S6**).

Isotope labelling experiments in *F. suffruticosa* had shown that Δ^1 -piperideine in securinine originates from lysine in a nonsymmetric manner, not proceeding via a free cadaverine intermediate²⁷. Having identified *Fs1864* as the candidate enzyme for this process, we were able to examine this phenomenon *in vitro*. First, we determined that the enzyme cannot use cadaverine as a substrate to form Δ^1 -piperideine (**Fig. S7**). Next, we examined the nonsymmetric formation of Δ^1 -piperideine using a lysine substrate with the side chain-nitrogen and ϵ -carbon labelled with heavy isotopes ($\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine). Using this substrate, we could follow the two possible routes to Δ^1 -piperideine from lysine (**Fig. 2C**). In the step-wise route, where the reaction occurs *via* a free symmetric cadaverine intermediate and either amine group could be converted to an aldehyde, we would expect to observe both singly and doubly labelled products, which would provide a mixture of two Δ^1 -piperideine monomer masses, $[\text{M}+\text{H}]^+ = 85$ and 86. In a concerted route, where the decarboxylated intermediate is enzyme bound, only the nitrogen derived from the lysine α -amine would be expected to be converted into the aldehyde, leading to only the double labelled product: $[\text{M}+\text{H}]^+ = 86$. When *Fs1864* was supplied with $\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine, only the doubly labelled product was observed, indicating a concerted reaction (**Fig. 2D**).

To further verify the reaction mechanism, we analysed the isotopically labelled enzyme product using ^{13}C -NMR (**Fig. 2E**). The spectra were obtained directly from enzyme assays conducted in water; hence the product is present as a mixture of Δ^1 -piperideine forms. The fully ^{13}C -labelled $^{13}\text{C}_6$ -lysine was used as a control to establish the chemical shifts (ppm) of all carbon atoms of the enzyme product in water. These were classified into four groups based on chemical shifts: aldehyde (a), iminium (b), sp^3 -hybridised adjacent to nitrogen (c) and sp^3 -hybridised adjacent to carbon (d). This spectrum was compared to the spectrum of the enzyme product derived from $\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine, where only the carbons derived from the $\epsilon^{13}\text{C}$ -lysine would show NMR signals. In the ^{13}C -NMR of this product, we could only see peaks of carbons in group c (sp^3 -hybridised adjacent to nitrogen) and none in groups a or b (**Fig. 2E**). This pattern of labelling is in line with the concerted pathway; for the step-wise pathway we would expect signals from groups a or b as well as c.

This *in vitro* assessment of *Fs1864* supports its ability to producing Δ^1 -piperideine from lysine *via* a PLP-dependent decarboxylation-oxidative deamination reaction²⁶. The accumulating cadaverine is not an intermediate to Δ^1 -piperideine but an alternative product. The labelling study supports, using both mass-spectrometry and ^{13}C -NMR, a concerted reaction mechanism that proceeds without a free intermediate and provides an enzymatic explanation for the nonsymmetric labelling observed in the plant.

Ornithine/lysine/arginine decarboxylase-like family (OLADLs)

LDCs are not known to be universally found in plants: characterised LDCs operating in plant specialised metabolism have evolved from ODCs on multiple independent occasions²⁰. The tendency of plant LDCs to evolve convergently alongside the unprecedented nonsymmetric oxidative deamination of lysine catalysed by *Fs1864* prompted an investigation into the enzyme's evolutionary origin.

PLP-dependent enzymes are categorized overall by fold type²⁸, and PLP-dependent amino acid decarboxylases are categorized by groups⁵. Based on annotation and sequence similarity, *Fs1864* was revealed to be a member of the group III decarboxylase family^{5,29}, which is part of the PLP fold I superfamily. The group II decarboxylases, which include plant amino acid decarboxylases and aldehyde synthases^{30,31}, are also in the PLP fold I superfamily. Notably *Fs1864* is from a different fold type and decarboxylase group to typical plant LDC/ODCs, which are group IV decarboxylases, with PLP fold type III. Structures of representatives of these sequences including *Fs1864* confirm the sequence-based relationships (**Fig. 3A**). The decarboxylase group III is known to contain prokaryotic type ornithine/lysine/arginine (Orn/Lys/Arg) decarboxylase-like sequences (OLADLs)⁵. To our knowledge, there is only a single report mentioning eukaryotic OLADLs³², and none has been characterised.

We inferred a phylogenetic tree of decarboxylase group III OLADL proteins based on an amino acid alignment of both prokaryotic and eukaryotic sequences, and used the plant group II decarboxylases as an outgroup (**Fig. 3B, Table S6, Dataset S1**). The eukaryotic OLADL lineage leading to *Fs1864* originates from within a clade of cyanobacterial enzymes. This lineage includes representatives from major algal clades (e.g. rhodophyta, chlorophyta), as well as across green plants (viridiplantae), land plants (Embryophyta) and flowering plants (Angiospermae), including monocots (e.g. rice, maize and wheat), asterids and rosids. However, only a single homolog was identified in gymnosperms (*Taxus chinensis*) and none in Brassicaceae species including *Arabidopsis*. The gene tree accurately reflects the evolution of the chloroplast in photosynthetic organisms from its cyanobacterial origins, even recovering proposed secondary endosymbiotic events giving rise to ochrophyte and haptophyte chloroplasts from the rhodophyte chloroplast. Therefore, this phylogeny indicates that the origin of the prokaryotic-like OLADLs in plants was the photosynthetic endosymbiosis event, when a cyanobacteria was engulfed by a eukaryotic cell and was transformed into the chloroplast. Examination into the origin of *Fs1864* appears to have revealed a hitherto neglected family of plant enzymes, the OLADLs, which may have wider roles in metabolism such as polyamine biosynthesis.

***F. suffruticosa* OLADL and OLADO**

To examine *Fs1864* evolution more closely, we constructed a phylogenetic tree of OLADLs from the rosid clade of flowering plants (**Fig. 3C, Table S6, Dataset S2**). This analysis revealed that *Fs3498*, which was previously excluded from screening due to low expression, is a close homolog of *Fs1864*. *Fs1864* and *Fs3498* are sister

sequences yet share only 54% sequence identity. Furthermore, *Fs1864* has <60% sequence identity with orthologs from Malpighiales, whereas *Fs3498* has >65% identity with the same sequences. Altogether this indicates a gene duplication and activity divergence event in the *Flueggea* lineage, where *Fs3498* maintained ancestral OLADL activity and *Fs1864* evolved new oxidative activity related to alkaloid biosynthesis. Furthermore, we examined the genome assembly of *F. virosa*, which also produces securinine-type alkaloids. We identified orthologs of *Fs1864* and *Fs3498* present as tandem duplicates on *F. virosa* chromosome 5 (05G000661 and 05G000660 respectively), a genomic signature of lineage specific neofunctionalisation (**Fig. 3D**)³³.

Based on this hypothesis, we elected to test the activity of *Fs1864* and *Fs3498* on lysine, ornithine and arginine. Based on *in vitro* end-point assays, all three amino acids were substrates for the enzymes, with *Fs3498* performing decarboxylation and only trace oxidation. *Fs3498* was renamed *FsOLADL* for its annotation as an ornithine/lysine/arginine decarboxylase-like enzyme. *Fs1864* was able to perform both decarboxylation and oxidative deamination on all three substrates (**Fig. S8**), so was renamed ornithine/lysine/arginine decarboxyl-oxidase (*FsOLADO*). The ability of *FsOLADO* to make both Δ^1 -piperidine and 1-pyrroline is notable as this mirrors the presence in *F. suffruticosa* of securinine and its pyrrolidine analog norsecurinine. Furthermore, *FsOLADO* is the first reported plant enzyme catalysing formation of γ -guanidinobutyraldehyde (GBAL) from arginine.

To assess whether these enzymes could convert the substrates at biologically relevant rates, we measured their specific activities on lysine and ornithine, comparing them to *LaLDC* (**Fig. 4A** and **Table S7**)³. *FsOLADO* demonstrated greater rates of lysine or ornithine consumption than *LaLDC*. We can infer that *FsOLADO* substrate consumption was mostly contributing to the Δ^1 -piperidine or 1-pyrroline formation, as those products accumulated over the time course and whereas there were very low rates of cadaverine or putrescine formation. Therefore, *FsOLADO* catalyses concerted decarboxylative-oxidation of ornithine and lysine at biologically relevant rates (compared to *LaLDC*) (**Table S7**). Both *FsOLADO* and *FsOLADL* have extremely low decarboxylation activity on ornithine or lysine, indicating that these are unlikely to be biologically relevant substrates for *FsOLADL*/OLADO decarboxylation, despite the product accumulating over longer reaction times (**Fig. S8**). The relatively low rates for *FsOLADL* across both substrate consumption and product formation suggests that neither lysine nor ornithine are likely to be its native substrate.

Subcellular localisation

The subcellular localisation of metabolism in plants has important implications for metabolic flux, control and potential engineering of the pathway³⁴. For example, *Lupinus* LDC is chloroplastic³³ whereas the subsequent oxidation step by CAOs is peroxisomal³⁵. We suspected that the unstructured N-terminal extensions of *FsOLADO* was a signal peptide (**Fig. S5A**), and computational analysis of the

FsOLADO and *FsOLADL* sequences gave strong prediction of a chloroplastic transit peptide and plastidal localisation (**Table S8**). To examine this further, we performed localisation experiments by transiently expressing fluorescent fusion proteins in *N. benthamiana* (**Table S9**). Transient expression of a full-length *FsOLADO*-GFP fusion failed to show green fluorescence, perhaps as the GFP disrupted the enzyme dimer and led to instability and degradation. However, fusion of the first 67 N-terminal residues of *FsOLADL* to GFP led to a protein fusion that co-localised with the chloroplast (**Fig. S9**). Curiously the fluorescent signal appears to be concentrated on the chloroplast periphery, suggesting incomplete processing of the signal peptide. Whilst further examination of the OLADL signal peptides would be desirable, this result, together with the computation prediction and likely origin of this gene from the cyanobacterial chloroplast progenitor lends support to chloroplastic localisation.

Active site variants

Next, we aimed to understand the structural basis of the different activities of *FsOLADL* and *FsOLADO*. To do this, we generated protein models with AlphaFold³⁶, added the PLP cofactor with AlphaFill³⁷ and inspected the active site (**Fig. S10**). We identified residues near the PLP cofactor and lining a binding pocket. We also identified residues by comparing the sequence alignments to characterised homologs (**Fig. S11**). Three regions of interest were identified where residues *FsOLADO* appeared to provide greater space in the active site than those in *FsOLADL*: N383-N384 (OLADO) compared to S330-Y331 (OLADL), T169-G170 compared to P116-E117, and F560 compared to Y508 (**Fig. 4B** and **Fig. S10**). In the *FsOLADL* model, E117 and Y508 side-chains interact through an H-bond. This interaction is absent in the *FsOLADO* model which appears to result in a notable increase in active site volume. The N384/Y331 position aligns with the residue found to be involved in the transition between decarboxylation (tyrosine) and aldehyde synthase (phenylalanine) activities in the distantly related group II decarboxylases (**Fig. S11**)³⁸. In these enzymes, when the tyrosine is present, its hydroxyl group is proposed to protonate the C α carbanion of the PLP-bound intermediate after decarboxylation, leading to release of the monoamine after formation of the internal aldimine. When the phenylalanine is present the protonation is impaired, which allows molecular oxygen to be attacked by the C α carbanion to produce a peroxide intermediate which decomposes into the aldehyde alongside ammonia and hydrogen peroxide³⁸.

To examine the role of active site residues we substituted them to the corresponding residue in the paralog (**Table S5** and **Table S10**), purified the enzymes and tested the variants on the three substrates (**Fig. 4C** and **Figs. S12** and **S13**). For *FsOLADO*, T169P and G170E were disruptive to both decarboxylation and oxidation activities. The N384Y substitution caused no change to decarboxylation and oxidation activities. The F560Y substitution significantly increased decarboxylation and decreased oxidation of ornithine (forming putrescine and pyrroline respectively) and arginine (forming agmatine and GBAL respectively). Such changes were not significant for lysine in the

WT background. However, introducing F560Y to the quad mutant background (T169P, G170E, N383S, N384Y) resulted in significant increase in the decarboxylation product (cadaverine) and significant decrease in oxidation product (Δ^1 -piperidine), suggesting interactions between these residues.

For *FsOLADL*, E117G significantly boosted decarboxylation activity, both alone and in combination with other substitutions. The substitution most influential for determining decarboxylation or oxidation activity in *OLADL* was Y508F, which significantly decreased decarboxylation products of arginine, lysine, ornithine and arginine (cadaverine, putrescine, and agmatine respectively), and increased oxidation products of ornithine and arginine (1-pyrroline and GBAL respectively). Significant increases in the lysine oxidation product (Δ^1 -piperidine) were evident through addition of Y508F into the quad mutant background (P116T/E117G/S330N/Y331). Overall, across *FsOLADL* and *FsOLADO*, the key residue for determining decarboxylation or oxidation activity appears to be Y508/F560, analogous to the tyrosine/phenylalanine of the group II decarboxylases described above. However, the E117/G170 residue also appears to be highly influential, with the glutamate side chain potentially activating the tyrosine hydroxyl group.

Based on these results, we speculate that in *FsOLADL* Y508 donates a proton to the intermediate-PLP complex to trigger imine formation and the decarboxylation product (**Fig. S14**). Therefore, in *FsOLADO* F560Y decarboxylation is enhanced. We also propose that in *FsOLADL* E117 is H-bonded to Y508 and restricts the active site volume, preventing effective substrate binding: the *OLADL*-E117G substitution enables improved substrate binding and boosts decarboxylation. In *OLADL*, the E117-Y508 H-bond interaction may be impacted by P116 which influences the loop structure on which E117 is situated. This H-bond interaction may be crucial for the positioning of the tyrosine residue, as in the *OLADO* model where it is absent the equivalent residue side chain (F560) is pointing away from the active site. Together, E117 and Y508 could precisely position a proton, and increase its acidity, allowing for facile protonation of a reaction intermediate, leading to the decarboxylation product (**Fig. S14**).

The impact of Y508F on *OLADL* oxidation is only clear in the presence of E117G as the substrate is required to bind. This interpretation also implies that the natural substrate of *FsOLADL* may not be lysine, ornithine or arginine but a smaller molecule, as supported by the low measured activities with these substrates (**Fig. 4A** and **Table S7**).

Parallel evolution of OLADO from OLADL in Solanaceae

Convergent evolution of enzyme activity or gene function is not a rare phenomenon in plant specialised metabolism^{6,37}. In particular, parallel evolution, which is the independent emergence of activity in separate lineages but from similar ancestors (i.e. same enzyme class) appears to be commonplace^{6,39}. Two relevant examples are the

parallel evolution of canonical LDCs from ODCs³⁶, and the development of aldehyde synthase activities from group II decarboxylases³⁸. Parallel evolution may even involve convergence of amino acid sequences, especially for active site residues³⁹. We predicted that multiple lineage specific nonsymmetric alkaloid biosynthesis pathways were controlled by OLADO enzymes that had evolved, in parallel, from the OLADL family. As in *Nicotiana* sp. Δ^1 -piperidine is incorporated into anabasine in a non-symmetric manner (**Fig. 1A**)¹⁹, we predicted an OLADO would be responsible for Δ^1 -piperidine formation in these species. To investigate this, we inferred a focussed phylogeny of OLADLs from the asterid clade and examined the active site residues for similarities to FsOLADO (**Fig. 5, Table S6, Dataset S3**). This investigation revealed two clades in the asterids showing OLADO-like residues, notably phenylalanines in the FsOLADO-F560 position and no tyrosine in the FsOLADO-N384 position (**Fig. 5**). One of these OLADO-like clades featured solely Solanaceae species, an alkaloid-rich family which includes *Nicotiana* and tropane alkaloid producing species such as *Datura stamonium*. Another clade with OLADO features contained Asteraceae species including *Artemisia annua*, which is not known to produce alkaloids. For both taxa (Solanaceae and Asteraceae) there were also clades that lack OLADO sequence features, which were presumably OLADL paralogs.

We elected to investigate OLADL/OLADO activities in *N. tabacum*, which produces anabasine nonsymmetrically as a minor alkaloid component³⁸. Closer investigation into the *Nicotiana tabacum* genome assembly revealed five OLADL paralogs in tandem arrays, with two on chromosome 9 and three on chromosome 18 (**Fig. 6A**). The published gene models failed to accurately define the gene boundaries so these were refined by cross reference with other datasets (**Fig. S15**). Each array has at least one gene from each of the Solanaceae OLADL and OLADO clades (**Fig. 5**), suggesting the two clades arose through a tandem gene duplication in a common ancestor of the Solanaceae. Two arrays are present due to whole genome duplication, with the homeologous chromosomes 9 and 18 originating from *N. tomentosiformis* and *N. sylvestris* respectively, the parents of the allotetraploid *N. tabacum*⁴⁰.

Due to the high similarity between the paralogs within each clade, we elected to test the activity of one representative from each clade, both from chromosome nine: A0A1S3ZLK7 from the OLADL clade and A0A1S3ZKS9 from the proposed OLADO clade (**Table S5**). As hypothesised, the activities closely resembled the FsOLADL/FsOLADO pair, with A0A1S3ZLK7 demonstrating decarboxylation of lysine, ornithine and arginine but minimal oxidation and A0A1S3ZKS9 showing oxidative deamination of lysine, ornithine and arginine, alongside the decarboxylation (**Fig. 6B**). Based on this activity we renamed the enzymes *NtOLADL* and *NtOLADO* respectively. We would expect other enzymes (paralogs or orthologs) in the same Solanaceae clades and sharing the key active site residues, to share similar distinct OLADL and OLADO activities (**Fig. 5**).

Gene expression patterns would provide further insight into the role of these sequences. Due to the inaccurate gene models we were unable to assess *N. tabacum* expression using existing resources but instead used *N. benthamiana* data, which also produces nicotine and anabasine^{7,41}. The two *N. benthamiana* OLADL genes showed balanced expression across multiple tissues, contrasting with OLADOs, which were more varied across tissues and included pseudogenes (i.e. fragmented or incomplete open reading frames) with near zero expression (**Fig. S16**). In *N. benthamiana*, two OLADO paralogs (Scf02530g05010 and Ctg12894g00002) had high expression in root and stem samples. These genes had very high expression correlation with orthologs from the nicotine biosynthesis pathway (top 30 correlating genes, PCC = 0.75–0.93, **Table S11**). This assessment suggests that OLADOs are involved in *Nicotiana* alkaloid biosynthesis.

Parallel evolution of Asteraceae OLADO

Through our examination of OLADL genes in the asterid clade, we noticed a clade of Asteraceae sequences that had distinctly OLADO-like residues: phenylalanines in the FsOLADO-F560 position and no tyrosine in the FsOLADO-N384 position (**Fig. 5**). We elected to examine these enzymes further, focussing on *A. annua*, despite there being limited reports of alkaloids in this species⁴². Analysis of the genome assembly revealed six genes across two chromosomes, in arrays of three (**Fig. 6C**). Based on the phylogeny, these appear to be the outcome of Asteraceae-specific tandem and whole genome duplications (**Fig. 5**). We expressed, purified and characterised four of these enzymes, two hypothesised to be OLADL and two OLADO based on active site residues (**Fig. 5, Table S5**). The hypothesised OLADLs chr7g04669011 and chr4g02112611 indeed showed decarboxylation activity and minimal oxidation activity with lysine, ornithine and arginine, though notably chr4g02112611 had superior activity across all substrates (**Fig. 6D**). This may be because it has a threonine present in the FsOLADO-E383 position (**Fig. 5**), which based on the mutation analysis (**Fig. 4**) would provide more space in the active site for these substrates. The predicted OLADOs demonstrated decarboxylation and oxidation activities, with chr7g04669441 turning over all three substrates tested but chr7g04669261 only showing comparable activity on lysine.

Discussion

The nonsymmetric incorporation of basic amino acids into alkaloids was noted in early biosynthetic investigations employing radioactively labelled precursors⁴³. Later, Leistner and Spenser proposed that a PLP-dependent oxidase could be responsible for the direct nonsymmetric conversion of lysine into Δ^1 -piperidine^{8,42}. Despite recent advances in elucidating the elaborate pathways of alkaloid biosynthesis^{8,44}, the nature of this crucial early step has been unresolved until now.

In this work we have revealed the plant enzymes that are capable of the nonsymmetric formation of lysine and ornithine derived alkaloids: the OLADOs. We first identified an

OLADO in *F. suffruticosa*, showing it to be a PLP-dependent enzyme capable of producing Δ^1 -piperideine from lysine *via* decarboxylation and oxidation, validating the previous hypothesis⁴³. In the putative *Securinega* alkaloid pathway, OLADO's products Δ^1 -piperideine and 1-pyrroline are precursors to alkaloids such as securinine and norsecurinine. Recently, a sulfotransferase was identified in *F. suffruticosa* based on co-expression with OLADO and was revealed to be involved in the scaffold remodelling of *Securinega* alkaloids⁴⁵. This supports OLADO's role in alkaloid formation. This work therefore describes the first enzyme of *Securinega* alkaloid biosynthesis, enabling future elucidation of the full pathway towards these polycyclic bioactive compounds.

Guided by the principle of parallel evolution, we identified analogous enzymes from the alkaloid-rich Solanaceae, characterising an OLADO from *N. tabacum*. The activity and alkaloid gene expression correlation of *Nicotiana* OLADO suggests it has a role in alkaloid biosynthesis, especially of anabasine, which is derived nonsymmetrically from lysine¹⁹. OLADOs may also contribute to other Solanaceae pathways, including nicotine formation in *N. benthamiana* (see **Table S11**) the tropane alkaloid biosynthesis in *D. stramonium*, which are derived nonsymmetrically from ornithine²¹, or the granatane alkaloids that are derived from Δ^1 -piperideine⁴⁰. It is curious that in Solanaceae species there may be two pathways that contribute cyclic iminiums to alkaloids. *N*-methylpyrrolinium is known to be derived from a symmetric pathway via ODC and an oxidase in the peroxisome, whereas the newly discovered OLADOs may contribute Δ^1 -piperideine or 1-pyrroline from the chloroplast. Further investigation could reveal the exact contribution of each pathway, and whether there are species-specific or tissue-specific differences in their contributions. Again, using parallel evolution as a gene discovery tool, we identified OLADOs in *A. annua*, a species not reported to produce alkaloids. However, alkaloids likely derived from basic amino acids are known in the *Artemisia* genus, including flavoalkaloids and long-chain alkyl derivatives, hinting that there are similar compounds present in *A. annua*⁴².

Piperidine derived alkaloids are known outside the Solanaceae, notably in *Sedum* (Crassulaceae, Saxifragales) and pomegranate (*Punica granatum*, Lythraceae, Rosids)⁴⁴, yet the early steps of these pathways are not known despite investigation⁴⁶. It is possible that these pathways employ lysine active enzymes from the OLADL family, either catalysing decarboxylation or perhaps demonstrating yet another example of OLADO parallel evolution.

PLP-dependent enzymes capable of decarboxylative oxidative deamination are well known, including for ornithine, which can be converted to 1-pyrroline by bacterial ODCs^{32,40}. However, the plant OLADOs characterised here are also capable of accepting lysine or arginine as substrates in an equivalent reaction. The product from arginine, γ -guanidinobutyraldehyde (GBAL), is a known plant natural product and potential precursor to GABA⁴⁷, and has been found as a component in some alkaloid structures⁴⁸.

The discovery of the plant OLADOs adds an enzyme class to our biocatalytic and synthetic biology toolkit. For example, heterologous reconstructions of tropane alkaloid synthesis tend to use the classical separate decarboxylase and oxidase enzymes, the latter requiring peroxisomal localisation⁴⁹. It is possible that OLADOs could be used in place of decarboxylase/oxidase pairs as part of *in vitro* or heterologous reconstructions of alkaloid biosynthesis. They could also be targets for genome engineering (e.g. in *Nicotiana*) to modify alkaloid content.

The OLADOs are part of the PLP-dependent group III ornithine/lysine/arginine (Orn/Lys/Arg) decarboxylase-like family (OLADLs), and we identified OLADO paralogs which retain typical OLADL active site residues and decarboxylation activities. Previously, OLADLs were known as prokaryotic enzymes, though here we have found representatives across major clades of green plants, including major crop species. They appear to have been transferred into plants via the cyanobacterial precursor of the chloroplast (**Fig. 3A**). Their absence from most animal clades and Arabidopsis likely contributed to them being overlooked. Whilst the OLADOs appear to contribute to alkaloid biosynthesis in a lineage specific manner, the conservation of the OLADLs indicates a possible role in plant primary metabolism. However, the low activity of OLADLs on the basic amino acid substrates tested here, together with the increase in activity provided by single amino acid substitutions, is strong indication the basic amino acids are not the enzymes' natural substrates. However, our evidence suggests that they are likely to be responsible for decarboxylation of an amino acid within chloroplasts under certain conditions or developmental stages. We will assess OLADL expression patterns and perform mutations and substrate screens to reveal the identity of this amino acid.

Prior to this work, there were multiple examples of PLP-dependent decarboxylases gaining oxidative activities: for example, disruption of metabolic pathways in *Escherichia coli* led to the evolution of oxidative activity in the ODC SpeC in a laboratory timescale⁵⁰. Another example are the plant aromatic amino acid aldehyde synthases which have emerged repeatedly from decarboxylases³⁸. This repeated evolution may occur due to the catalytic promiscuity of enzymes with PLP co-factors⁵¹. By comparing predicted structures of *F. suffruticosa* OLADL and OLADO we identified and validated key active site substitutions. These signatures enabled us to trace independently evolved OLADOs across multiple lineages. Furthermore, the OLADL/OLADO pairs were found on the genomes as tandem duplications, which is characteristic of lineage specific neofunctionalisation³³. Therefore, alongside the OLADL/OLADO gene discovery, which has implications across plant primary and specialised metabolism, this work highlights the predictive power of parallel evolution, which, empowered by sequence information, structural prediction, and phylogenetic inference, can serve as a tool for gene discovery.

506 **Methods**

507 **General**

508 Chemicals and consumables used for molecular biology and biochemistry were
509 obtained from Millipore, Invitrogen, or Merck. Substrates used in enzyme assays were
510 obtained from Sigma-Aldrich. Lysine isotopologs were obtained from CKisotopes (UK
511 distributor for Cambridge Isotope Laboratories). Sanger sequence verification was
512 obtained from Eurofins Genomics or SourceBioscience. Primers were ordered from
513 Integrated DNA technologies. Illumina sequencing was performed by Novogene, ONT
514 sequencing by University of York Bioscience Technology Facility. Sequencing data
515 was analysed using Geneious Prime. Figures were compiled using R-Studio, Microsoft
516 Excel, ChemDraw, Geneious Prime, Chimera, iTOL⁵² and Adobe Illustrator.

517 **Plant samples**

518 *Flueggea suffruticosa* plants (3-5 years old) were obtained from Brighton Plants
519 (Sussex, UK), grown from seed purchased from Chiltern Seeds (Oxford, UK) in
520 January 2014. These seeds originated from a plant in Poland, where it is a common
521 garden ornamental. Prior to sampling, plants were grown in spring-like conditions (20
522 °C 16 h day, 15 °C 8 h night).

523 **RNA extraction**

524 Samples (i.e. root, stem, old leaves >3 months old, young leaves < 3 months old) were
525 isolated from plants, immediately flash-frozen in liquid nitrogen and ground in liquid
526 nitrogen using a pestle and mortar. Roots were washed thoroughly in water prior to
527 freezing. Frozen ground tissue (100 mg) was extracted using a standard CTAB
528 protocol⁴⁴ including a gDNA removal step (Invitrogen™ TURBO DNA-free Kit). The
529 quality of the RNA was assessed using an Agilent 2100 BioAnalyzer, with all samples
530 achieving RNA Integrity Number >8.4.

531 Illumina short read RNA-sequencing was performed by Novogene (UK) Ltd, enriching
532 for polyA mRNA transcripts, and generating paired end 150 base sequencing reads
533 on an Illumina NovaSeq 6000. Full length cDNA sequencing libraries were prepared
534 using the Oxford Nanopore Technologies (ONT) cDNA PCR barcoding kit (SQK-
535 PCB109), and pooled barcoded libraries were sequenced on ONT MinION and
536 PromethION R9.4.1 flowcells (FLO-MIN106 and FLO-PRO004).

537 **Transcriptome assembly**

538 The ONT reads were base-called using Guppy v5.0.11 and subjected to quality control
539 using PycoQC to ensure the integrity and reliability of the sequencing data⁵³.
540 Concurrently, the Illumina reads underwent quality assessment with FastQC, were
541 trimmed of adapters with Cutadapt⁵⁴, and ribosomal RNA was removed using BBSplit.

The long-read sequences from all samples were pooled and assembled *de novo* to construct the transcriptome using Rattle⁵⁵, a tool optimised for long-read assembly. This initial assembly was then polished with Medaka, employing the ONT reads to correct sequencing errors. To enhance the accuracy of the assembled transcriptome further, three rounds of polishing were conducted using Pilon with the combined Illumina short reads. Post-assembly, the transcripts were filtered through Evigene using the tr2aacds.pl Perl script to remove redundancy and low-quality sequences.

For annotation, the assembled transcripts were first processed with Transdecoder to identify open reading frames (ORFs). Annotation was then performed using BLASTP and BLASTX against the UniProt Swiss-Prot database to identify known protein functions. Protein domains were delineated using InterProScan⁵⁶.

Transcriptome quantification

For the quantification of RNA-seq reads, we employed Kallisto, a tool designed for the fast and accurate estimation of transcript abundances through pseudoalignment⁵⁷. Initially, a Kallisto index was generated using our assembled transcriptome as a reference. Reads were then quantified against this index, employing Kallisto's quant command with parameters tailored to our data, including 100 bootstrap samples for uncertainty estimation. The resulting output provided both estimated counts and transcripts per million (TPM) values for each transcript, facilitating downstream analyses such as differential expression.

Transient expression in *Nicotiana benthamiana*

Nicotiana benthamiana plants were grown in F2+S and Calypso insecticide-treated soil in a growth room before and after infiltration (25°C on a 16/8 h day/night cycle). Plants were infiltrated with LBA4404 *Agrobacterium tumefaciens* at 5 weeks old. For preparation of LBA4404 *A. tumefaciens*, chemically competent cells (50 µL) were transformed with plasmid (200 ng) using a heat shock protocol: after mixing cells and plasmids, cells were flash frozen in liquid nitrogen, then defrosted (5 min, 37 °C), and then incubated on ice (30 min). LB media (250 µL) was added, cells were incubated (120 rpm, 2h, 28 °C), and then plated on LB agar (50 µg/mL streptomycin, 50 µg/mL kanamycin). Colonies from plates (after 2–3 days growing at 28 °C) were inoculated into LB media and incubated (50 µg/mL streptomycin, 50 µg/mL kanamycin; 28 °C, 220 rpm, 24 hours). Cells were pelleted (20 min, 4000 xg, RT) and resuspended in infiltration media (10 mM MES, 10 mM MgCl₂, 100 µM acetosyringone, OD600 = 0.5) and incubated (RT, 50 rpm, 2 h). Cultures were infiltrated onto the abaxial side of leaves using a 1 mL syringe. One leaf per plant and three plant replicates per plasmid construct were used. For controls we used GFP vector strains (infiltration positive control) and empty vector strain (metabolite negative control). All plants were grown in 16/8 h light/dark, 25 °C, for 5 days before harvest for analysis.

Leaf discs (8 mm diameter) were excised and placed into wells (24 well plate) with buffer (400 μ L, 50 mM HEPES, pH 7.5), epidermis facing upwards. Discs were gently pressed to sink to the bottom of the well, and then substrate was added (50 μ L, 10 mM lysine). The wells were sealed and placed in a growth chamber for one light cycle (16 h light, 8 h dark, 25 °C). For metabolite extraction, discs were removed from the wells, dried on tissue paper, placed in 2 mL tubes and the tubes frozen in liquid nitrogen. For metabolite isolation, three tungsten carbide beads were added to each tube, and samples were homogenised to a fine powder using a TissueLyser (2x 40 s, 25 Hz) whilst frozen. Solvent was added (200 μ L, 80:20 MeOH:H₂O), tubes were vortexed (1 h, RT), and samples centrifuged (10,000 $\times g$, 5 min). Extracts were diluted with MeOH (600 μ L), filtered through 0.45 μ m PTFE filters and transferred to LC-MS glass vials for subsequent analysis.

Protein expression and purification

For protein expression, *E. coli* Rosetta™(DE3) or SoluBL21 (DE3) (Genlantis) were used. Single colonies transformed with plasmids containing genes of interest were used to inoculate overnight cultures with antibiotic selection (2xYT media, 37 °C, 200 rpm). These were diluted with 2xYT with antibiotics (OD600 = 0.1) and grown to OD600 = 0.7 for Rosetta cells or OD600 = 0.4-0.6 for SoluBL21 (37 °C, 200 rpm). Then the cells were induced with IPTG (1 mM) and incubated overnight (~14 h, 200 rpm, 18 °C). Cells were harvested by centrifugation (3200 $\times g$, 10 min, 4 °C), the supernatant removed, and the pellet resuspended in lysis buffer (10% v/v compared to original culture; 50 mM Tris-HCl, 50 mM glycine, 5% v/v glycerol, 0.5 M NaCl, 20 mM imidazole, 0.1 mM PLP, pH 8; per 50 mL, 1x EDTA-free protease inhibitor and 10 mg lysozyme) and incubated (30 min, 4 °C). The cells were lysed through a cell disruptor (26 kPsi) and clarified by centrifugation (20 min, 35,000 $\times g$, 4 °C). Purification was achieved through nickel affinity purification using an ÄKTA start FPLC. Clarified lysate was loaded onto a Ni-NTA column (5 mL) and then the column washed with binding buffer (50 mM Tris-HCl, 50 mM glycine, 5% v/v glycerol, 0.5 M NaCl, 40 mM imidazole, 0.1 mM PLP, pH 8) and the protein eluted with elution buffer (50 mM Tris-HCl, 50 mM glycine, 5% v/v glycerol, 0.5 M NaCl, 500 mM imidazole, 0.1 mM PLP, pH 8). A protein purification outcome was inspected by SDS-PAGE, and if a pure was protein obtained, samples were buffer exchanged into storage buffer (50 mM HEPES, 250 mM NaCl, 0.1 mM PLP, pH 7.5) using a PD10 column, using a gravity method as per manufacturer instructions. The protein was concentrated by centrifugation (Amicon® Ultra-2 mL Centrifugal Filters, 30 kDa MWCO) prior to addition of glycerol (10% v/v), flash freezing and storage at -70 °C.

Cloning

Initially, genes of interest for plant expression (4074, 4984, 1864) were identified using the transcriptome assembly and cloned from the plant source. RNA isolated from *Flueggea suffruticosa* (see above) was reverse transcribed to cDNA using Invitrogen™ SuperScript IV kit as per the manufacturer's protocol. The cDNA was

immediately used for PCR amplification. Primers were designed that contained an overhang for cloning into vectors (see supplementary information for primers). PCR was performed using SuperFi™ polymerase enzyme and amplified genes were isolated using a gel extraction kit. For LaLDC (Uniprot: F5HPR7_LUPAN)⁵⁵ and truncated OLADL/OLADO sequences for expression in *E. coli*, we first selected protein sequences based on the transcriptome or genome sequences, and then ordered them as synthetic codon-optimised genes cloned into pET-28a(+) (Twist Bioscience) (sequences are available in the supplementary information). For LaLDC plant-based expression, the synthesised gene from pET-28a(+) was subcloned into pHREAC (see below).

Purified pHREAC vector plasmid⁵⁸ was digested with BsaI (50 µL reactions; 1 µg DNA, 10 U of BsaI in CutSmart™ Buffer, 37 °C, 1 h) and processed using a PCR clean-up kit (Promega Wizard™). Amplified genes with pHREAC overhangs, deriving either from the cDNA or a plasmid (see above), were inserted into linear pHREAC using In-Fusion HD cloning (with insert:plasmid 2:1). A colony PCR using pHREAC primers (see supp table) was carried out to confirm plasmid transformation and successful cloning. After overnight incubation of transformed *E. coli*, plasmids were minipreped and sequence verified before transformation into LBA4404 *Agrobacterium tumefaciens*.

Mutagenesis

Forward and reverse site directed mutagenesis primers were designed to overlap by 8-15 bases at the 5' end, with the mutation present in the overlapping region⁵⁹. The templates were the codon optimised genes on pET-28a(+). Primers were synthesised by Integrated DNA Technologies UK LTD. Plasmids were propagated in DH5α *E. coli* cells and extracted using the Zymopure™ plasmid miniprep kit. PCR was performed using SuperFi™ polymerase enzyme. Each PCR reaction (50 µL) mix contained template DNA (0.2 ng/µL), DMSO (1.5% v/v), Platinum™ SuperFi II PCR Master Mix (50% v/v), and forward and reverse primers (0.5 µM). PCR cycles were initiated at 98 °C for 1.5 min; followed by 30 amplification cycles: 98 °C for 10 sec, 60 °C for 10 sec and 72 °C for 3.5 mins; then an elongation step at 72 °C for 5 min. DpnI (1 µL) restriction enzyme was added directly to the PCR mix and left to incubate overnight at room temperature. PCR mix was used to transform into DH5α *E. coli* cells to propagate plasmid and extracted using the Zymopure™ plasmid miniprep kit. Plasmids were sequence verified before transformation into SoluBL21 (DE3) (Genlantis) for protein expression and purification.

In vitro end-point enzyme assays

Reaction components were mixed (100 nM, 10 ng/µL or 20 ng/µL purified protein, 0.1 mM PLP, 50 mM NaCl, 2 mM DTT, 50 mM HEPES pH 7.5, 5 or 10 mM substrate, 100 µL final volume). Substrate was added last and the reactions incubated (16 h, 30 °C) prior to quenching with MeCN (100 or 400 µL) and LC-MS analysis. For protein

controls, proteins were pre-incubated (5 min, 99 °C) prior to substrate addition. Reactions were performed in triplicate, with similar outcomes. One representative chromatogram is presented. For inhibitor assays, the protein was purified without the addition of PLP in the buffers and was incubated with 1 mM of semicarbazide (30 min, 25 °C) prior to its addition into the reaction mixture (final semicarbazide concentration in assays was 50 µM). In a subset of reactions PLP was not added to the assays.

Specific activity enzyme assays

For specific activity measurements, *LaLDC*, *FsOLADO* and *FsOLADL* were expressed in *E. coli* and purified as described above. Reactions contained 10 mM lysine or 5 mM ornithine, 0.1 mM PLP, 50 mM NaCl, 2 mM DTT, 50 mM HEPES pH 7.5 with 100 µL final volume. To identify a linear rate, different enzyme concentrations and time points were tested. Enzyme concentrations tested were 50 ng/µL, 0.2 mg/mL or 0.5 mg/mL for lysine and 0.2 mg/mL for ornithine; time points were 20, 45 and 65 minutes for lysine and 10 and 20 minutes for ornithine. Reactions were quenched with MeCN containing leucine internal standard and analysed by LC-MS (see below). Peak areas, relative to internal standard, were calculated and plotted to identify and select linear activity regimes, prior to calculation of specific activity. Reactions were performed in triplicate. For relative Δ^1 -piperidine or 1-pyrroline activity, peak area of the monomers was normalised to the maximum (*FsOLADO*). For inferred relative Δ^1 -piperidine or 1-pyrroline activity rate, the rate of cadaverine/putrescine formation was subtracted from the rate of substrate consumption.

LC-MS analysis

Enzyme activity analysis was performed using hydrophilic interaction liquid chromatography (HILIC). Chromatography was performed using a HPLC instrument Dionex UltiMate 3000 RSLC nano system or a Thermo Fisher Vanquish UHPLC. The column was an XBridge® BEH Amide column (Waters, 5 µm, 4.6mm x 100 mm). Phase A [H₂O 0.125% v/v formic acid, 10 mM ammonium formate] and phase B [95:5 MeCN:Phase A] were used as the mobile phase components. For the Dionex, chromatography was performed with 0.6 mL/min flow rate and a gradient: 0-4 min, 100% B; 4-24 min, 100-60% B; 24-25 min, 60-100% B; 25-35 min, 100% B. For the Thermo Fisher, chromatography was performed at 0.6 mL/min flow rate and 30 °C with gradient: 0-2 min, 100% B; 2-10 min, 100-60% B; 10-11 min, 60-100% B; 11-12 min, 100% B. Mass spectrometry was performed with an MS Bruker HCT ultra-ETD II (50-500 m/z collected). Standards were dissolved in either phase A or B depending on solubility. Data was processed in Bruker Compass DataAnalysis Version 4.4 (Dionex) or Freestyle (Thermo). For statistical analysis of variants, the peak areas were compared with each substrate/product combination using ANOVA with post-hoc Tukey HSD, assigning statistically significantly different groupings ($p < 0.05$) with letters (R studio).

H₂O₂ Assays

Peroxide formation was detected using phenol and 4-aminoantipyrine (4-AAP) to produce colorimetric readout at 505 nm. Reactions contained: 0.5 mg/mL purified enzyme, PBS pH 8.0, 20 mM lysine, 1 mM DTT, 4 mM phenol, 6 mM 4-AAP, 100 U peroxidase and PLP (0.0125-0.2 mM). Negative control assays were performed without enzyme, and positive controls by the addition of H₂O₂. Spectrophotometric measurements were taken using a CLARIOstar® Plus UV-Visible spectrophotometer (BMG Labtech) at 505 nm in continuous plate mode. Runs were performed in triplicate.

¹³C-NMR Assays

Enzyme assays performed as previously described but with ε¹⁵N, ε¹³C-lysine or ¹³C₆-lysine. After incubation (16 h, 30 °C), samples were pooled, diluted in 10:90 D₂O:H₂O and placed in an NMR tube (S400). The spectrum was measured at 700 MHz using a Bruker NEO spectrometer equipped with a cryoprobe prodigy TCI. ¹³C spectra used 14,000 scans with a 2s relaxation delay. Data analysis was carried out using MestRe Nova V14.

Synthesis of Δ¹-piperideine

To a solution of *N*-chlorosuccinimide (8.34 g, 62.5 mmol) in dry diethyl ether (60 mL) under argon was added piperidine (5.90 mL, 58.7 mmol) dropwise and the resulting mixture was stirred for 2 hours at room temperature. The reaction mixture was then filtered and the filtrate was washed with water (2 x 35 mL), dried (Na₂SO₄), filtered and concentrated *in vacuo* to give an oily residue. A solution of NaOMe in MeOH, freshly prepared by adding sodium metal portionwise (1.53 g, 66.7 mmol) to MeOH (35 mL), was added to the oily residue at room temperature. The reaction mixture was stirred at reflux for 45 mins. After cooling to room temperature, water (50 mL) was added and the mixture was extracted with Et₂O (3 x 50 mL). The combined organic layers were dried (MgSO₄), filtered and concentrated *in vacuo* to give Δ¹-piperideine (2.72 g, 56%). NMR data obtained were consistent with the literature. NMR data obtained matched those previously described⁶⁰.

Synthesis of 1-pyrroline

To form 1-pyrroline *in situ* for LC-MS analysis, 4-aminobutyraldehyde diethyl acetal (SigmaAldrich) was dissolved in 0.1 M HCl and incubated at 30 °C for 30 mins prior to analysis.

Synthesis of 4-guanidinobutyraldehyde diethylacetal hemisulfate

To a solution of *S*-methylisothioureia hemisulfate (0.278 g, 2.00 mmol) in water (4 mL) and ethanol (4 mL) was added 4-aminobutyraldehyde diethyl acetal (0.345 mL, 2.00 mmol) and the reaction mixture was stirred at room temperature for 18 hours. The reaction mixture was then concentrated *in vacuo*. Purification by trituration with diethyl ether gave 4-guanidinobutyraldehyde diethylacetal hemisulfate (0.423 g, 84%) as a white solid. Mp 95–97 °C; $\nu_{\text{max}}/\text{cm}^{-1}$ (neat) 3139, 2973, 2881, 1689, 1634, 1372, 1062,

596; δ_H (400 MHz, DMSO- d_6) 1.18 (6H, t, J 7.1 Hz), 1.59–1.84 (4H, m), 3.12–3.23 (2H, m), 3.47–3.57 (2H, m), 3.61–3.72 (2H, m), 4.53 (1H, t, J 5.2 Hz); δ_C (101 MHz, DMSO- d_6) 15.7 (2 $\times$ CH₃), 25.3 (CH₂), 31.9 (CH₂), 42.1 (CH₂), 62.8 (2 $\times$ CH₂), 104.2 (CH), 158.7 (C); m/z (ESI) 204.1708 (MH⁺. C₉H₂₂N₃O₂ requires 204.1707). To form 4-guanidinobutyraldehyde *in situ* for LC-MS analysis, the diethylacetal hemisulfate was dissolved in 0.1 M HCl and incubated at 30 °C for 30 mins prior to analysis.

Protein structural analysis

Structural models of enzymes were constructed using AlphaFold2 and AlphaFold2-multimer on ColabFold³⁶. PLP cofactors were added with AlphaFill³⁷. The best scoring model was selected for display. For dimers of PS, the unstructured N-terminal sequence was removed prior to calculation. The structural models, along with X-ray structures were depicted using Chimera (UCSF). For EcLDCI (PDB: 3N75) selected subunits of the homo 10-mer were depicted to enable visual comparison.

Phylogenetic analysis

For phylogenetic analysis of the group III decarboxylase proteins, due to differences in InterPro annotations and density of sampling required, prokaryotic and eukaryotic sequences were obtained differently. For prokaryotes, amino acid sequences of proteins in the Orn/Lys/Arg decarboxylase family were obtained from InterPro (IPR052357). These sequences were clustered at a 70% sequence identity threshold using CD-HIT⁶¹, with a single representative from each cluster maintained for the next step. To obtain eukaryotic group III decarboxylase amino acid sequences, eukaryotic sequences of proteins containing the Orn/Lys/Arg decarboxylase, major domain were obtained from InterPro (IPR000310). Then duplicate and obsolete sequences were removed, as were sequences with <400 residues and those without Pfam PF01276 OKR_DC_1 annotation (the major domain of the OLADs). After alignment with MAFFT⁶², sequences were further curated by removing those missing the PLP binding lysine and the conserved glycine (570 on OLADO) (which are likely to be misannotated or pseudogenes), selecting a single sequence from those with >95% sequence identity with others, and removing sequences with large alignment gaps with respect to OLADO residues 98-580.

For the combined prokaryotic and eukaryotic alignment and tree, the selected sequences along with the group II decarboxylase outgroup sequences were aligned with MAFFT. To curate the alignment: columns containing 99% gaps or more were removed, then sequences that aligned across less than 800 columns (i.e. amino acids including gaps) or had below 400 ungapped amino acids (i.e. sequence length) were removed. Taxon-id and species corresponded to each gene accession was obtained from UniProtKB and the corresponding phyla were obtained through NCBI taxonomy. An approximately-maximum likelihood tree was inferred using FastTree 2.1.11 with three site rate categories and an optimised gamma distribution⁶³.

For the rosoid phylogeny, a subset of the eukaryotic sequences were obtained, aligned with MAFFT, then columns with >95% gaps were removed prior to tree inference. A maximum likelihood tree was then inferred using W-IQ-TREE⁶⁴ with ModelFinder⁶⁵, ultrafast bootstraps (UFBoot2, X1000)⁶⁶, and SH-aLRT supports (X1000)⁶⁷. Trees were depicted using iTOL⁵².

For the focussed asterid phylogeny, we accessed published genomes to obtain homologs from: *Nicotiana tabacum*⁶⁸, *Artemisia annua*⁶⁹, *Atropa belladonna*⁷⁰, *Datura stramonium*⁷⁰, *Solanum lycopersicum* (ITAG4.1)⁷¹, *Capsicum annuum* (UCD10X, Ref)⁷², *N. benthamiana* (Niben1.0.1)⁷¹ and *N. attenuata* (NIATTr2)⁷³. The nucleotide sequences from *Nicotiana tabacum* homologs identified from the InterPro domain were used as queries for blastn or Megablast with default settings, with the genome extracted coding sequences (CDS) as the databases. The top hit genes according to blastn or MegaBlast, as judged by Bit-Score were obtained, filtered for length (>1000 nucleotides), translated to peptides and then used in the alignment. The alignments were performed with MAFFT and the tree inferred with W-IQ-TREE. For *N. tabacum* and *A. annua* the genomic location of the genes of interest were viewed using the genome fasta and gff file in Geneious Prime 2024. Alignment statistics (e.g. length, pairwise id etc) were calculated using the automatic statistics in Geneious Prime. Amino acid saturation metrics were calculated using both alignment and tree with PhyKit's saturation function⁷⁴.

***N. benthamiana* gene expression**

For *N. benthamiana* the gene expression was assessed using the Plant Gene Expression Omnibus⁷⁵. For gene expression boxplots the TPM values across all 476 samples in the omnibus were downloaded for each OLAD/OLADO paralog and plotted in R-Studio, grouping by tissue. For expression correlations, we used the inbuilt co-expression calculation (PCC value) with the exception of the ODC genes, for which we calculated PCCs manually. *N. benthamiana* alkaloid biosynthesis gene orthologs were annotated using *N. tabacum* derived query sequences and obtaining top hits from tblastn on the extracted coding sequences (CDS) from the *N. benthamiana* genome. UniProt accession values for queries: PMT1_TOBAC (PMT, putrescine methyltransferase), AO2A_TOBAC (AO, aspartate oxidase), IFRH_TOBAC (A622), ODC1B_TOBAC (ODC, ornithine decarboxylase [group IV]), QPT2B_TOBAC (QPT, quinolinate phosphoribosyltransferase), QS1_TOBAC (QS, quinolinate synthase), F1T160_TOBAC (BBL, berberine bridge enzyme), and MPO1_TOBAC (MPO, N-methylputrescine oxidase).

Subcellular localisation

Subcellular localisation was predicted using DeepLoc 2.1⁷⁶. For experimental examination of subcellular localisation, the full-length open reading frame for *F. suffruticosa* OLADO and OLADL genes was amplified from the *F. suffruticosa* cDNA using the primers detailed in (Table S9) and inserted into the pJET clone vector using

818 the CloneJET PCR cloning Kit (Thermo Fisher Scientific, UK). OLADO, OLADL-N67
819 and GFP fragments were amplified with overhangs which were then assembled into a
820 BamHI and EcoRI digested pHREAC expression vector by using In-Fusion cloning
821 tools (TaKaRa Bio Inc., Kusatsu, Japan). GFP was fused with the C-terminal of gene
822 of interest by a (GS)₃ linker. The constructs were transformed into Stellar competent
823 cells and confirmed by Plasmidsaurus sequencing. Confirmed constructs were
824 transformed into GV3101 agrobacterium chemical competent cells. Agrobacteria
825 carrying genes of interest and plastid located mCherry⁷⁷ were co-expressed in *N.*
826 *benthamiana* leaves for three days and then florescence was visualized under ZEISS
827 LSM980 confocal microscope.

828 **References**

- 829 1. Lichman, B. R. The scaffold-forming steps of plant alkaloid biosynthesis. *Nat. Prod.*
830 *Rep.* **38**, 103–129 (2021).
- 831 2. Leete, E. Recent developments in the biosynthesis of the tropane alkaloids. *Planta*
832 *Med.* **56**, 339–352 (1990).
- 833 3. Bunsupa, S. *et al.* Lysine decarboxylase catalyzes the first step of quinolizidine alkaloid
834 biosynthesis and coevolved with alkaloid production in leguminosae. *Plant Cell* **24**,
835 1202–1216 (2012).
- 836 4. Bunsupa, S. *et al.* Molecular Evolution and Functional Characterization of a Bifunctional
837 Decarboxylase Involved in Lycopodium Alkaloid Biosynthesis. *Plant Physiol.* **171**,
838 2432–2444 (2016).
- 839 5. Sandmeier, E., Hale, T. I. & Christen, P. Multiple evolutionary origin of pyridoxal-5'-
840 phosphate-dependent amino acid decarboxylases. *Eur. J. Biochem.* **221**, 997–1002
841 (1994).
- 842 6. Weng, J.-K. The evolutionary paths towards complexity: a metabolic perspective. *New*
843 *Phytol.* **201**, 1141–1149 (2014).
- 844 7. Saitoh, F., Noma, M. & Kawashima, N. The alkaloid contents of sixty *Nicotiana* species.
845 *Phytochemistry* **24**, 477–480 (1985).
- 846 8. Nett, R. S. *et al.* Plant carbonic anhydrase-like enzymes in neuroactive alkaloid
847 biosynthesis. *Nature* **624**, 182–191 (2023).
- 848 9. Mancinotti, D. *et al.* The causal mutation leading to sweetness in modern white lupin
849 cultivars. *Sci Adv* **9**, eadg8866 (2023).
- 850 10. Eljounaidi, K. & Lichman, B. R. Nature's Chemists: The Discovery and Engineering of
851 Phytochemical Biosynthesis. *Front Chem* **8**, 596479 (2020).
- 852 11. Trenchard, I. J. & Smolke, C. D. Engineering strategies for the fermentative production
853 of plant alkaloids in yeast. *Metab. Eng.* **30**, 96–104 (2015).
- 854 12. Frick, K. M., Kamphuis, L. G., Siddique, K. H. M., Singh, K. B. & Foley, R. C.
855 Quinolizidine Alkaloid Biosynthesis in Lupins and Prospects for Grain Quality
856 Improvement. *Front. Plant Sci.* **8**, 87 (2017).
- 857 13. Leistner, E. & Spenser, I. D. Biosynthesis of the piperidine nucleus. Incorporation of
858 chirally labeled cadaverine-1-3H. *J. Am. Chem. Soc.* **95**, 4715–4725 (1973).
- 859 14. Castillo, M., Gupta, R. N., MacLean, D. B. & Spenser, I. D. Biosynthesis of lycopodine
860 from lysine and acetate. The pelletierine hypothesis. *Canadian Journal of Chemistry* **48**,
861 1893–1903 (1970).
- 862 15. Zhao, T. *et al.* Engineering tropane alkaloid production based on metabolic

- 863 characterization of ornithine decarboxylase in *Atropa belladonna*. *ACS Synth. Biol.* **9**,
864 437–448 (2020).
- 865 16. Dalton, H. L. *et al.* Effects of down-regulating ornithine decarboxylase upon putrescine-
866 associated metabolism and growth in *Nicotiana tabacum* L. *J. Exp. Bot.* **67**, 3367–3381
867 (2016).
- 868 17. Junker, A., Fischer, J., Sichhart, Y., Brandt, W. & Dräger, B. Evolution of the key
869 alkaloid enzyme putrescine *N*-methyltransferase from spermidine synthase. *Front. Plant*
870 *Sci.* **4**, 260 (2013).
- 871 18. Kim, N., Estrada, O., Chavez, B., Stewart, C. & D'Auria, J. C. Tropane and granatane
872 alkaloid biosynthesis: A systematic analysis. *Molecules* **21**, 1510 (2016).
- 873 19. Leete, E. Biosynthesis of the *nicotiana* alkaloids. XIV. The incorporation of delta 1-
874 piperidine-6-14C into the piperidine ring of anabasine. *J. Am. Chem. Soc.* **91**, 1697–
875 1700 (1969).
- 876 20. Beutler, J. A. *et al.* Securinine alkaloids: a new class of GABA receptor antagonist.
877 *Brain Res.* **330**, 135–140 (1985).
- 878 21. Leete, E. The stereospecific incorporation of ornithine into the tropine moiety of
879 hyoscyamine. *J. Am. Chem. Soc.* **84**, 55–57 (1962).
- 880 22. Lou, Y.-R., Pichersky, E. & Last, R. L. Deep roots and many branches: Origins of plant-
881 specialized metabolic enzymes in general metabolism. *Curr. Opin. Plant Biol.* **66**,
882 102192 (2022).
- 883 23. Reimann, A., Nurhayati, N., Backenköhler, A. & Ober, D. Repeated evolution of the
884 pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages. *Plant*
885 *Cell* **16**, 2772–2784 (2004).
- 886 24. Chirkin, E., Atkatiian, W. & Porée, F.-H. The *Securinega* alkaloids. *Alkaloids Chem.*
887 *Biol.* **74**, 1–120 (2015).
- 888 25. Sato, H., Uchiyama, M., Saito, K. & Yamazaki, M. The Energetic Viability of Δ^1 -
889 Piperidine Dimerization in Lysine-derived Alkaloid Biosynthesis. *Metabolites* **8**, (2018).
- 890 26. Hoffarth, E. R., Rothchild, K. W. & Ryan, K. S. Emergence of oxygen- and pyridoxal
891 phosphate-dependent reactions. *FEBS J.* **287**, 1403–1428 (2020).
- 892 27. Marek Golebiewski, W., Horsewood, P. & Spenser, I. D. Biosynthesis of securinine: the
893 mode of incorporation of lysine. *J. Chem. Soc. Chem. Commun.* 217–218 (1976).
- 894 28. Schneider, G., Käck, H. & Lindqvist, Y. The manifold of vitamin B6 dependent enzymes.
895 *Structure* **8**, R1-6 (2000).
- 896 29. Blum, M. *et al.* The InterPro protein families and domains database: 20 years on.
897 *Nucleic Acids Res.* **49**, D344–D354 (2021).

- 898 30. Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. Modeling of the spatial structure of
899 eukaryotic ornithine decarboxylases. *Protein Sci.* **4**, 1291–1304 (1995).
- 900 31. Liang, J., Han, Q., Tan, Y., Ding, H. & Li, J. Current advances on structure-function
901 relationships of pyridoxal 5'-phosphate-dependent enzymes. *Front. Mol. Biosci.* **6**, 4
902 (2019).
- 903 32. Li, B., Liang, J., Hanfrey, C. C., Phillips, M. A. & Michael, A. J. Discovery of ancestral L-
904 ornithine and L-lysine decarboxylases reveals parallel, pseudoconvergent evolution of
905 polyamine biosynthesis. *J. Biol. Chem.* **297**, 101219 (2021).
- 906 33. Panchy, N., Lehti-Shiu, M. & Shiu, S.-H. Evolution of gene duplication in plants. *Plant*
907 *Physiol.* **171**, 2294–2316 (2016).
- 908 34. Facchini, P. J. & St-Pierre, B. Synthesis and trafficking of alkaloid biosynthetic
909 enzymes. *Curr. Opin. Plant Biol.* **8**, 657–666 (2005).
- 910 35. Yang, T. *et al.* Transcript profiling of a bitter variety of narrow-leaved lupin to discover
911 alkaloid biosynthetic genes. *J. Exp. Bot.* **68**, 5527–5537 (2017).
- 912 36. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**,
913 679–682 (2022).
- 914 37. Hekkelman, M. L., de Vries, I., Joosten, R. P. & Perrakis, A. AlphaFill: enriching
915 AlphaFold models with ligands and cofactors. *Nat. Methods* **20**, 205–213 (2023).
- 916 38. Torrens-Spence, M. P. *et al.* Structural basis for divergent and convergent evolution of
917 catalytic machineries in plant aromatic amino acid decarboxylase proteins. *Proc. Natl.*
918 *Acad. Sci. U. S. A.* **117**, 10806–10817 (2020).
- 919 39. Gherardini, P. F., Wass, M. N., Helmer-Citterich, M. & Sternberg, M. J. E. Convergent
920 evolution of enzyme active sites is not a rare phenomenon. *J. Mol. Biol.* **372**, 817–845
921 (2007).
- 922 40. Zan, Y. *et al.* The genome and GeneBank genomics of allotetraploid *Nicotiana tabacum*
923 provide insights into genome evolution and complex trait regulation. *Nat. Genet.* **57**,
924 986–996 (2025).
- 925 41. Vollheyde, K. *et al.* An improved *Nicotiana benthamiana* bioproduction chassis provides
926 novel insights into nicotine biosynthesis. *New Phytol.* **240**, 302–317 (2023).
- 927 42. Ur Rashid, M. *et al.* The chemistry and pharmacology of alkaloids and allied nitrogen
928 compounds from *Artemisia* species: A review. *Phytother. Res.* **33**, 2661–2684 (2019).
- 929 43. James, W. O. Biosynthesis of the belladonna alkaloids. *Nature* **158**, 654–656 (1946).
- 930 44. Bedewitz, M. A., Jones, A. D., D'Auria, J. C. & Barry, C. S. Tropinone synthesis via an
931 atypical polyketide synthase and P450-mediated cyclization. *Nat. Commun.* **9**, 5281
932 (2018).

- 933 45. Kim, S.-G. *et al.* Chemically guided single-cell transcriptomics reveals sulfotransferase-
934 mediated scaffold remodeling in securinine biosynthesis. *Research Square* (2025)
935 doi:10.21203/rs.3.rs-6626700/v1.
- 936 46. Zam, T., D'Auria, J. C. & Jaisi, A. Advancements in biotechnological approaches for the
937 production of high-value metabolites in *Punica granatum*. *ACS Food Sci. Technol.* **5**,
938 2086–2097 (2025).
- 939 47. Kopečný, D. *et al.* Plant ALDH10 family: identifying critical residues for substrate
940 specificity and trapping a thiohemiacetal intermediate. *J. Biol. Chem.* **288**, 9491–9507
941 (2013).
- 942 48. Gödecke, T. *et al.* Guanidine alkaloids and Pictet-Spengler adducts from black cohosh
943 (*Cimicifuga racemosa*). *J. Nat. Prod.* **72**, 433–437 (2009).
- 944 49. Srinivasan, P. & Smolke, C. D. Biosynthesis of medicinal tropane alkaloids in yeast.
945 *Nature* **585**, 614–619 (2020).
- 946 50. Pontrelli, S. *et al.* Metabolic repair through emergence of new pathways in *Escherichia*
947 *coli*. *Nat. Chem. Biol.* **14**, 1005–1009 (2018).
- 948 51. Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent
949 enzymes. *EMBO Rep.* **4**, 850–854 (2003).
- 950 52. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic
951 tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
- 952 53. Leger, A. & Leonardi, T. pycoQC, interactive quality control for Oxford Nanopore
953 Sequencing. *J. Open Source Softw.* **4**, 1236 (2019).
- 954 54. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing
955 reads. *EMBnet.journal* **17**, 10–12 (2011).
- 956 55. de la Rubia, I. *et al.* RATTLE: reference-free reconstruction and quantification of
957 transcriptomes from Nanopore sequencing. *Genome Biol.* **23**, 153 (2022).
- 958 56. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification.
959 *Bioinformatics* **30**, 1236–1240 (2014).
- 960 57. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq
961 quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
- 962 58. Peyret, H., Brown, J. K. M. & Lomonosoff, G. P. Improving plant transient expression
963 through the rational design of synthetic 5' and 3' untranslated regions. *Plant Methods*
964 **15**, 108 (2019).
- 965 59. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single
966 and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
- 967 60. Barker, G. *et al.* Enantioselective, palladium-catalyzed α -arylation of N-Boc pyrrolidine:

968 in situ react IR spectroscopic monitoring, scope, and synthetic applications. *J. Org.*
969 *Chem.* **76**, 5936–5953 (2011).

970 61. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-
971 generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).

972 62. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
973 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

974 63. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood
975 trees for large alignments. *PLoS One* **5**, e9490 (2010).

976 64. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast
977 online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44**, W232-
978 5 (2016).

979 65. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S.
980 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods*
981 **14**, 587–589 (2017).

982 66. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2:
983 Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

984 67. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood
985 phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).

986 68. Sierro, N., Auberson, M., Dulize, R. & Ivanov, N. V. Chromosome-level genome
987 assemblies of *Nicotiana tabacum*, *Nicotiana sylvestris*, and *Nicotiana tomentosiformis*.
988 *Sci. Data* **11**, 135 (2024).

989 69. Liao, B. *et al.* Allele-aware chromosome-level genome assembly of *Artemisia annua*
990 reveals the correlation between ADS expansion and artemisinin yield. *Mol. Plant* **15**,
991 1310–1328 (2022).

992 70. Zhang, F. *et al.* Revealing evolution of tropane alkaloid biosynthesis by analyzing two
993 genomes in the Solanaceae family. *Nat. Commun.* **14**, 1446 (2023).

994 71. Fernandez-Pozo, N. *et al.* The Sol Genomics Network (SGN)--from genotype to
995 phenotype to breeding. *Nucleic Acids Res.* **43**, D1036-41 (2015).

996 72. Hulse-Kemp, A. M. *et al.* Reference quality assembly of the 3.5-Gb genome of
997 *Capsicum annuum* from a single linked-read library. *Hortic. Res.* **5**, 4 (2018).

998 73. Xu, S. *et al.* Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proc.*
999 *Natl. Acad. Sci. U. S. A.* **114**, 6133–6138 (2017).

1000 74. Steenwyk, J. L. *et al.* PhyKIT: A multitool for phylogenomics. *Curr. Protoc.* **4**, e70016
1001 (2024).

1002 75. Koh, E., Goh, W., Julca, I., Villanueva, E. & Mutwil, M. PEO: Plant Expression Omnibus

- 1003 - a comparative transcriptomic database for 103 Archaeplastida. *Plant J.* **117**, 1592–
1004 1603 (2024).
- 1005 76. Nielsen, H. Practical applications of language models in protein sorting prediction:
1006 SignalP 6.0, DeepLoc 2.1, and DeepLocPro 1.0. *Methods Mol. Biol.* **2941**, 153–175
1007 (2025).
- 1008 77. Nagano, M. Generation of *Arabidopsis* lines with a red fluorescent marker for
1009 endoplasmic reticulum using a tail-anchored protein cytochrome b(5) -B. *Plant Signal*
1010 *Behav.* **15**, 1790196 (2020).

1011 **Acknowledgements**

1012 We thank Matthew Davy for running samples on the Jeol-700 NMR. We also thank Dr
1013 Ed Bergstrom for his training and guidance with the LC-MS instrumentation, and the
1014 Centre for Excellence in Mass Spectrometry for equipment. The plasmid pHREAC was
1015 a gift from George Lomonossoff (Addgene plasmid #134908). We thank the
1016 Biosciences Technology Facility including Grant Calder for support with fluorescence
1017 microscopy. We are grateful for computational support from the University of York
1018 High-Performance Computing service, Viking, and the Research Computing team.
1019 B.R.L. acknowledges support from UKRI (MR/S01862X/1) and BBSRC
1020 (BB/Y003586/1).

1021 **Author Contributions**

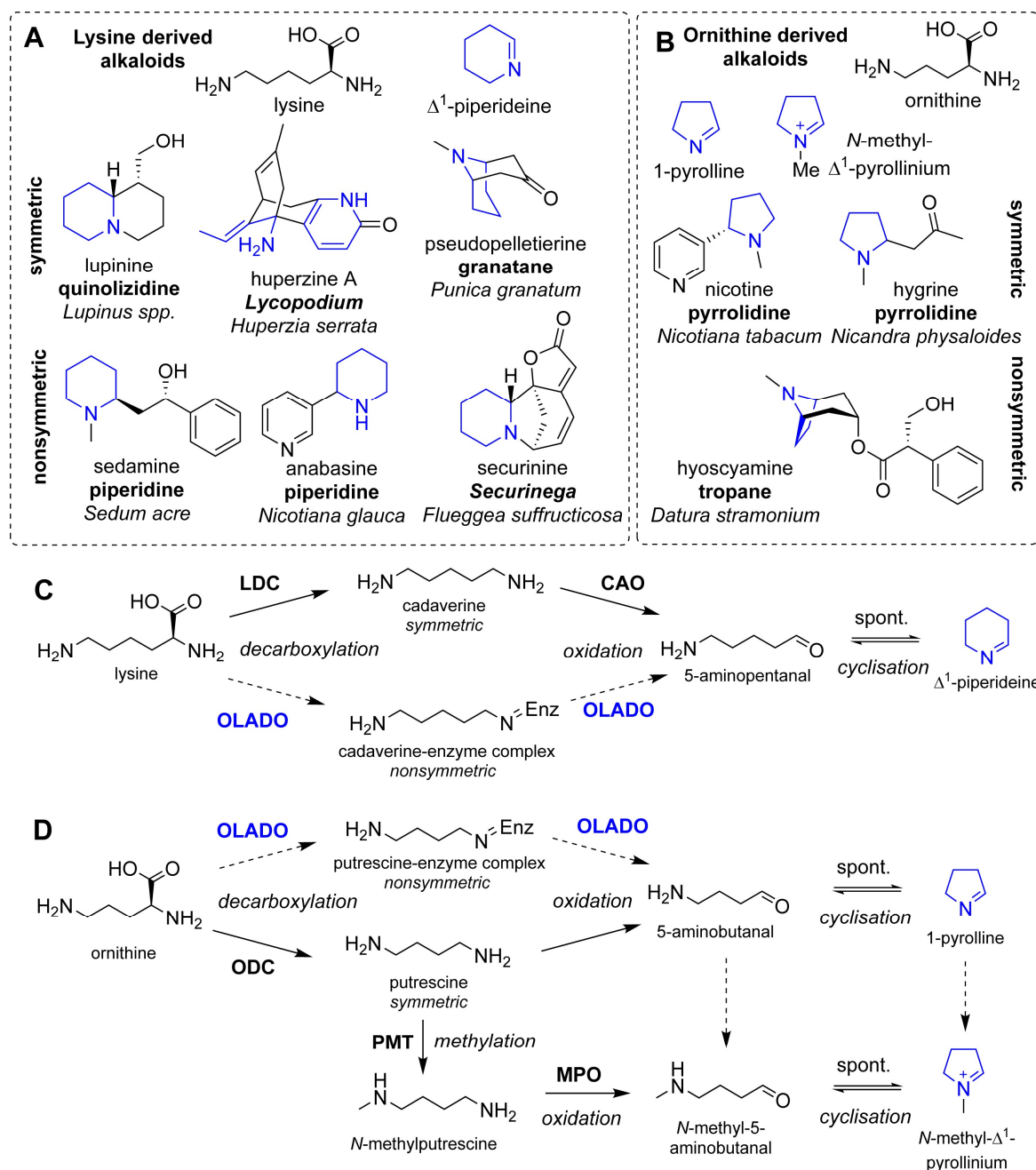
1022 C.X.W. carried out the nucleic acid extraction, transcriptome analysis, cloning, enzyme
1023 purification/characterisation and chemical synthesis. Z.J. performed enzyme
1024 characterisation including substrate and inhibitor tests and performed subcellular
1025 localisation experiments. I. A. performed mutagenesis, variant assays and specific
1026 activity assays. U.S.B. characterised *Artemisia* enzymes. L.J.N.W. performed
1027 chemical synthesis. O.S.D. performed the peroxidase assays. S.J. prepared
1028 sequencing libraries and ran ONT sequencing. K.N. assembled the transcriptome.
1029 W.P.U. assisted in the chemical synthesis and interpretation of NMR data. G.G.
1030 assisted in the interpretation of data. B.R.L. designed and supervised the project, and
1031 conducted the phylogenetic and structural analyses. C.X.W. and B.R.L. wrote the first
1032 draft of the manuscript. B.R.L. wrote the final manuscript with input from all authors.
1033 C.X.W. and Z.J. contributed equally.

1034 **Competing interests**

1035 The authors declare no competing interests.

1036 **Data availability**

1037 Raw read sequences for all generated data in this study are available in the National
1038 Center for Biotechnology Information Sequence Read Archive under BioProject ID
1039 PRJNA1106657. Sequences of coding sequences of genes are available on
1040 GenBank: gene 1864 (*FsOLADO*) (PP748646), gene 3498 (*FsOLADL*) (PP748647),
1041 gene 4984 (PP748648) and gene 4074 (PP748649). All other data is available upon
1042 reasonable request to the corresponding author.



1044

1045 **Figure 1. Basic amino acid derived alkaloids. A.** Lysine and derived alkaloids formed via
 1046 Δ^1 -piperideine (blue bonds) or methylated derivatives. Alkaloid class in bold, example
 1047 producing species in italics. **B.** Ornithine derived alkaloids formed via 1-pyrroline or
 1048 methylated derivatives. **C.** Enzymatic formation of Δ^1 -piperideine from lysine. The symmetric
 1049 route proceeds via a free symmetric cadaverine intermediate. Enzymes: lysine decarboxylase
 1050 (LDC) and copper amine oxidase (CAO). The previously unresolved nonsymmetric route is
 1051 thought to proceed via a nonsymmetric cadaverine-enzyme complex. **D.** Enzymatic formation
 1052 of 1-pyrroline and *N*-methyl- Δ^1 -pyrrolinium from ornithine. Enzymes: ornithine decarboxylase
 1053 (ODC), putrescine methyltransferase (PMT) and *N*-methylputrescine oxidase (MPO). Dashed
 1054 arrows represent proposed but unverified steps. Chemicals are drawn as neutral forms though
 1055 may be charged/zwitterionic in physiological conditions.

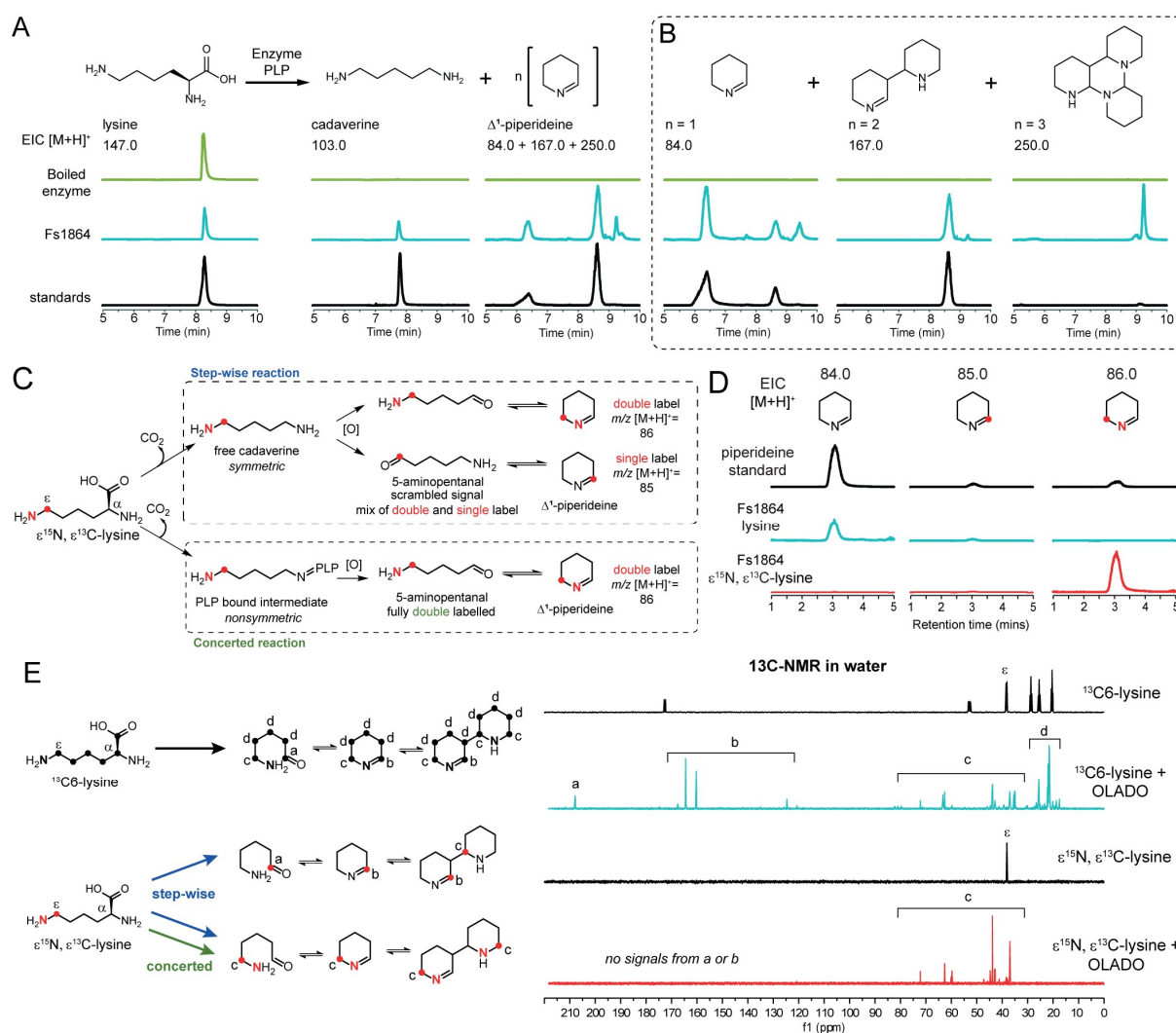


Figure 2. Identification of *Fs1864* catalysed non-symmetric Δ^1 -piperidine formation. A. *In vitro* assay of recombinant *Fs1864* (20 ng/ μ L) showing formation of cadaverine and Δ^1 -piperidine from lysine (10 mM) with the addition of PLP. Plots are extracted ion chromatograms (EICs) **B.** Δ^1 -Piperidine takes multiple forms in solution (monomer, dimer, trimers) as shown by EICs of the different masses from the reaction in panel A. **C.** Expected isotope labelling outcomes of stepwise (i.e. free cadaverine intermediate) or concerted (i.e. bound intermediate) formation of Δ^1 -piperidine using $\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine. **D.** LC-MS results of *Fs1864* (20 ng/ μ L) reaction with unlabelled lysine or $\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine (5 mM). **E.** ^{13}C -NMR spectra of lysine isotopologues and corresponding enzyme products. The $^{13}\text{C}_6$ -lysine derived product serves as a standard, showing all ^{13}C product shifts categorised as groups a-d based on ppm range. The $\epsilon^{15}\text{N}$, $\epsilon^{13}\text{C}$ -lysine derived product shows ^{13}C product shifts at specific group c ppm range, supporting the concerted mechanism.

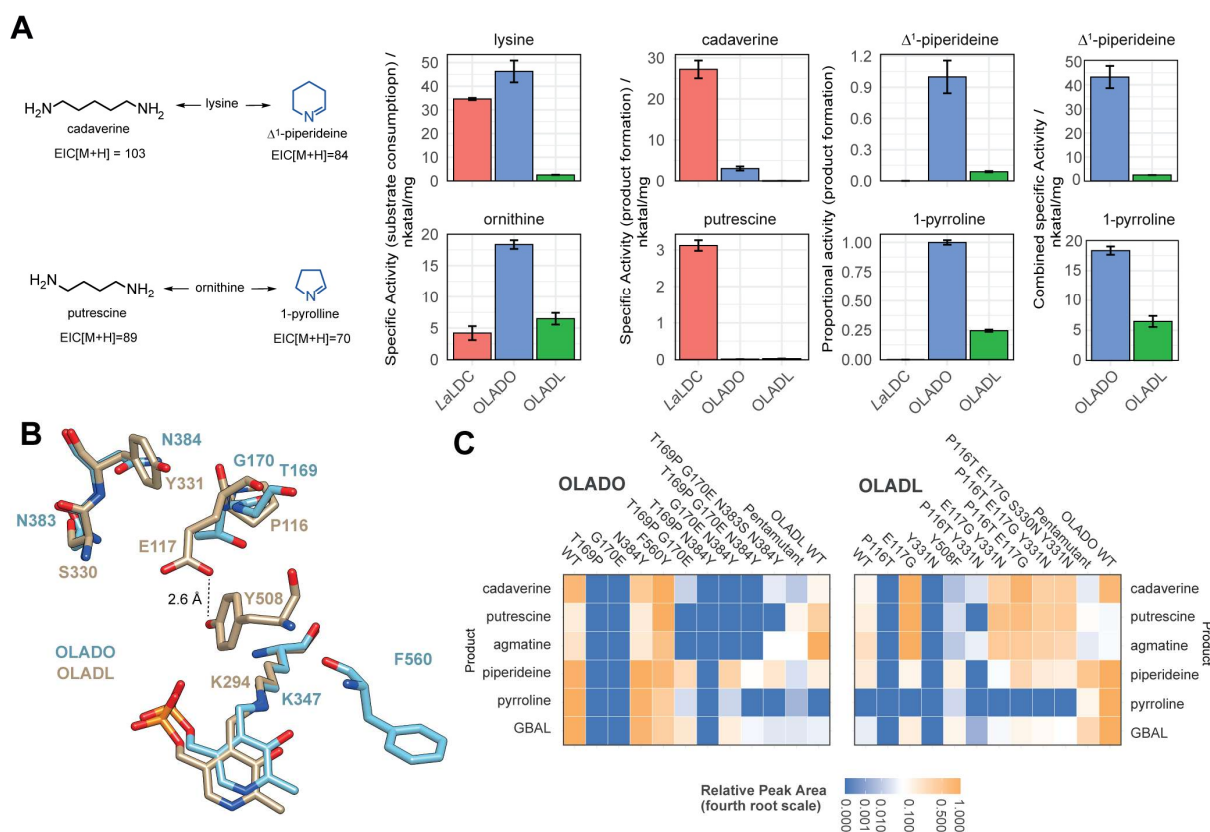


Figure 4. Activities of *FsOLADL* and *FsOLADO*. **A.** Specific activity of purified enzymes: substrate consumption (lysine/ornithine) or product formation. For proportional Δ^1 -piperidine or 1-pyrroline activity, peak area of the product monomers was normalised to the maximum (*FsOLADO*). For inferred Δ^1 -piperidine or 1-pyrroline specific activity, the rate of cadaverine/putrescine formation was subtracted from the rate of substrate consumption (combined specific activity). See **Table S7** for values. Lysine starting concentration was 10 mM, ornithine was 5 mM. Enzyme concentrations tested were 50 ng/ μ L, 0.2 mg/mL or 0.5 mg/mL, depending on activities. Bar charts show means of triplicate measurements; error bars show standard error of the mean. **B.** Active site comparison of AlphaFold models of *FsOLAD* (beige) and *FsOLADO* (blue) showing PLP-binding lysine alongside residues proposed to influence activity. Residues were selected based on proximity to PLP-cofactor and protein sequence alignments. **C.** Heat map showing *in vitro* product formation in purified enzyme variants, with variants shown in columns and products shown in rows, with colours relative to maximum reported conversions in each row. Pentamutants have all residues (as depicted in panel B) swapped. Substrate concentration 5 mM, enzyme concentration 20 ng/ μ L. Values are the mean of at least three assays, shown in full with statistical analysis in **Fig. S12** and **S13**. The fourth root scale transformation ($x^{0.25}$) enables differences between small values to be observed more clearly.

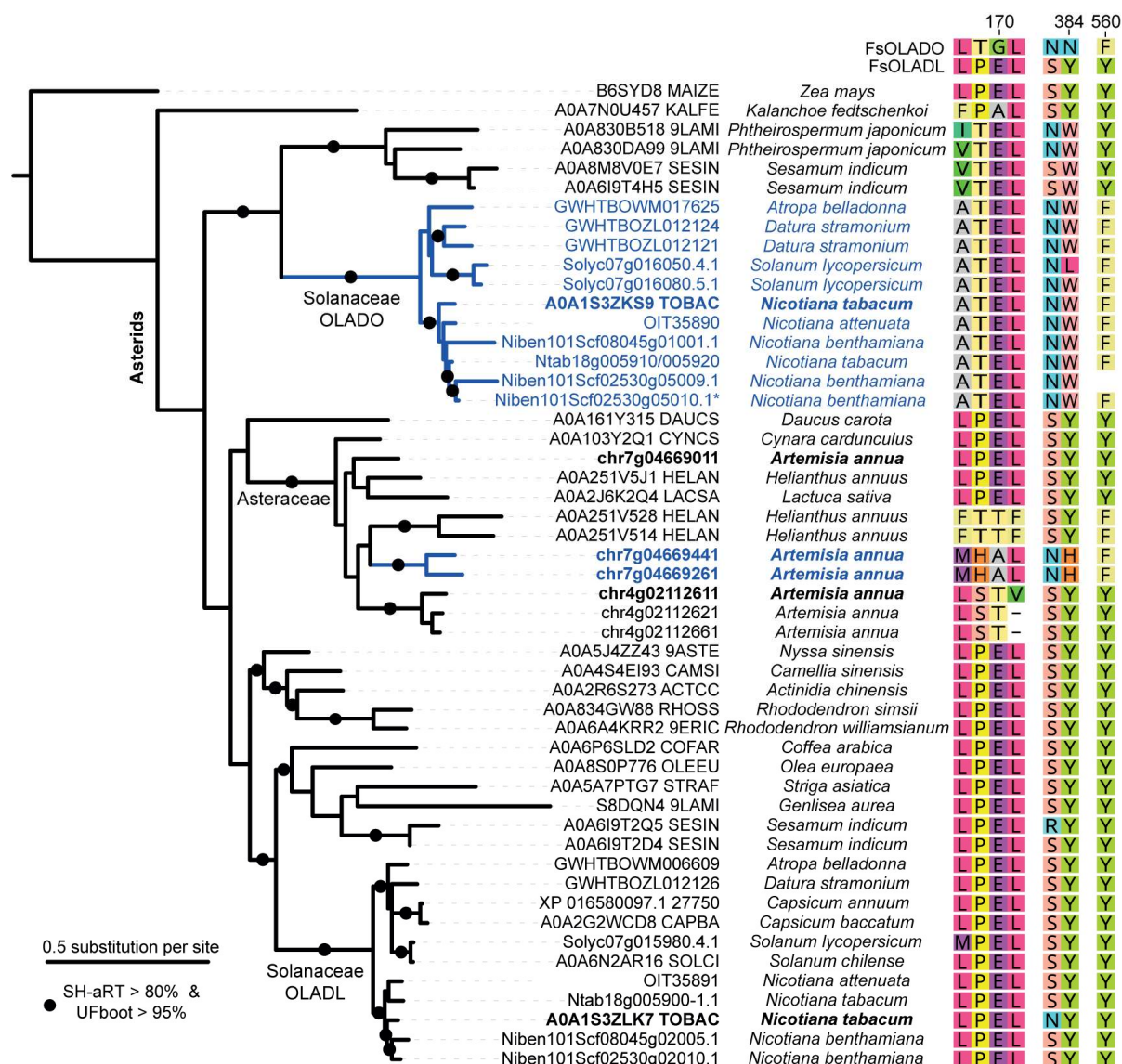


Figure 5. Parallel evolution of OLADO in Asterids. Focussed Asterid phylogeny of selected group III decarboxylase (OLADL) protein sequences. Black dots show highly supported nodes (SH-aLRT $\geq 80\%$ and UFboot $\geq 95\%$). Bold sequences are those that were selected for testing (Fig. 6). Blue clades are OLADO-like due to key residue identities. Selected alignment columns show identity of key residues in selected sequences compared to FsOLADL/OLADO. Residue numbering is based on FsOLADO. Starred *N. benthamiana* OLADO sequence was used for co-expression analysis (Fig. S13, Table S11).

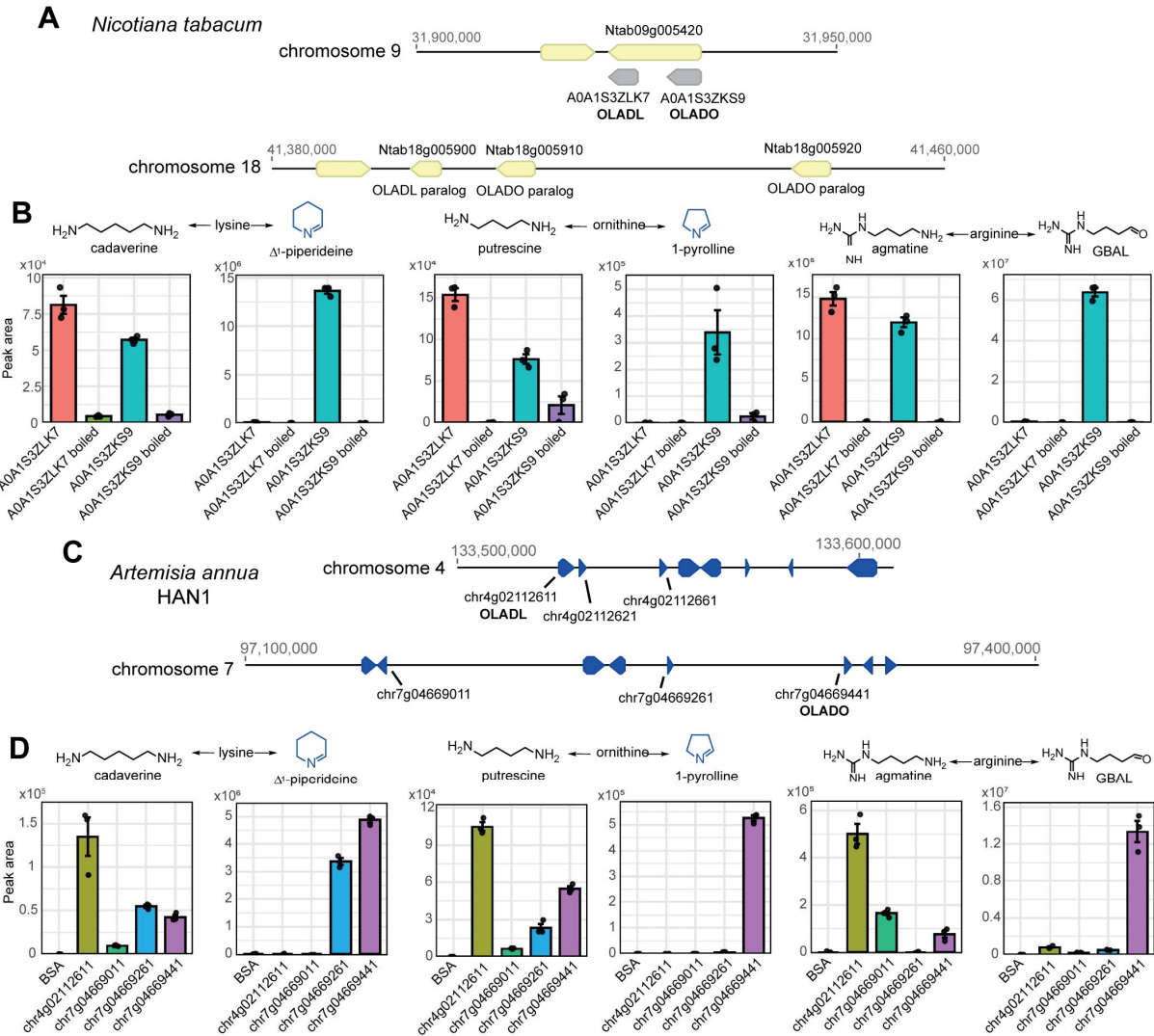


Figure 6. Activities of Asterid OLADL/OLADOs. **A.** Genomic location of *N. tabacum* OLADL paralogs. Refined gene model of Ntab09g005420 shown in **Fig. S15**. **B.** Activity of purified enzymes with lysine, ornithine and arginine. Bar graphs show LC-MS derived EIC peak areas; bars show mean of triplicate reactions, points from each reaction (n = 3), error bars are SE. **C.** Genomic location of *A. annua* OLADL paralogs. **D.** Activity of purified enzymes with lysine, ornithine and arginine. Bar graphs show LC-MS derived EIC peak areas; bars show mean of triplicate reactions, points from each reaction (n = 3), error bars are SE, substrate concentration 10 mM, enzyme concentration 10 ng/μL.