Molecular Plant
Research article

CellPress
Partner Journal

The biosynthetic pathway of the hallucinogen mescaline and its heterologous reconstruction

Paula Berman[1], Luis Alejandro de Haro[1], Ana-Rita Cavaco[1], Sayantan Panda[1,4], Younghui Dong[2], Nikolay Kuzmich[3], Gabriel Lichtenstein[1,5], Yoav Peleg[2], Hila Harat[1], Adam Jozwiak[1,6], Jianghua Cai[1,7], Uwe Heinig[2], Sagit Meir[1], Ilana Rogachev[1] and Asaph Aharoni[1,*]

[1]Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel

[2]Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot 7610001, Israel

[3]The Maurice and Vivienne Wohl Institute for Drug Discovery, Weizmann Institute of Science, Rehovot 7610001, Israel

[4]Present address: Department of Cell and Metabolic Biology, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany

[5]Present address: Department of Molecular & Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL 33136, USA

[6]Present address: Department of Botany and Plant Sciences, University of California Riverside, Riverside, CA 92521, USA

[7]Present address: Laboratory of Plant Hormones and Development Regulation of Chongqing, School of Life Sciences, Chongqing University, Chongqing 401331, China

***Correspondence: Asaph Aharoni (asaph.aharoni@weizmann.ac.il)**

https://doi.org/10.1016/j.molp.2024.05.012

ABSTRACT

Mescaline, among the earliest identified natural hallucinogens, holds great potential in psychotherapy treatment. Nonetheless, despite the existence of a postulated biosynthetic pathway for more than half a century, the specific enzymes involved in this process are yet to be identified. In this study, we investigated the cactus *Lophophora williamsii* (Peyote), the largest known natural producer of the phenethylamine mescaline. We employed a multi-faceted approach, combining *de novo* whole-genome and transcriptome sequencing with comprehensive chemical profiling, enzymatic assays, molecular modeling, and pathway engineering for pathway elucidation. We identified four groups of enzymes responsible for the six catalytic steps in the mescaline biosynthetic pathway, and an *N*-methyltransferase enzyme that *N*-methylates all phenethylamine intermediates, likely modulating mescaline levels in Peyote. Finally, we reconstructed the mescaline biosynthetic pathway in both *Nicotiana benthamiana* plants and yeast cells, providing novel insights into several challenges hindering complete heterologous mescaline production. Taken together, our study opens up avenues for exploration of sustainable production approaches and responsible utilization of mescaline, safeguarding this valuable natural resource for future generations.

Key words: mescaline, *Lophophora williamsii*, Peyote, biosynthetic pathway, *de novo* genome sequencing, molecular modeling

Berman P., de Haro L.A., Cavaco A.-R., Panda S., Dong Y., Kuzmich N., Lichtenstein G., Peleg Y., Harat H., Jozwiak A., Cai J., Heinig U., Meir S., Rogachev I., and Aharoni A. (2024). The biosynthetic pathway of the hallucinogen mescaline and its heterologous reconstruction. Mol. Plant. **17**, 1129–1150.

INTRODUCTION

In recent years, there has been a resurgence of interest in psychedelics owing to their reported efficacy in producing long-lasting symptom reduction, offering potentially transformative effects on neuroscience and psychiatry as we currently understand them (Nichols, 2016; Kwan et al., 2022; McClure-Begley and Roth, 2022). The use of psychedelics in healing and divinatory religious rituals stretches back thousands of years. Some important examples include the use of psilocybin mushrooms by ancient Aztec shamans, Ayahuasca by the Indigenous tribes of South America, and psychedelic cacti by Native North Americans, dating as far back as 5700 years (Nichols, 2016).

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

Mescaline, the primary hallucinogenic compound in psychedelic cacti, was first isolated and identified by Arthur Heffter in 1896 (Heffter, 1896, 1898). This serotonin receptor (5-HT_{2A}) agonist with psychotherapeutic potential exerts effects 1000–3000 times less potent than those of lysergic acid diethylamide and 30 times less potent than those of the fungal compound psilocybin, two other classical psychedelics, owing to its lower binding affinity (Rickli et al., 2016; Ley et al., 2023). Nevertheless, mescaline is classified as a Schedule I controlled substance, and its consumption is prohibited in most countries. Remarkably, *Lophophora williamsii* (here termed Peyote), which serves as the primary natural reservoir of mescaline, has obtained legal sanction for ceremonial use by the Native American Church as a medium for spiritual connection with the Great Spirit (God) (Nichols, 2016).

The biosynthetic pathway of mescaline **1** in Peyote has been postulated, relying mainly on radioisotope labeling studies and extensive metabolic profiling (Lundström and Agurell, 1969; Lundström, 1971). This predicted pathway can be divided into three modules, with some metabolites potentially biosynthesized through multiple routes (Figure 1A). First, the amino acid L-tyrosine **2** is hydroxylated via a hydroxylase (cytochrome P450 [CYP], polyphenol oxidase [PPO], or other) and decarboxylated by an L-tyrosine/L-DOPA decarboxylase enzyme (TyDC) to dopamine **3** through tyramine **4** or L-3,4-dihydroxyphenylalanine (L-DOPA **5**). Next, dopamine **3** is *O*-methylated by an *O*-methyltransferase (OMT) and hydroxylated by a hydroxylase to 3,4-dihydroxy-5-methoxy-phenethylamine (3,4-diOH-5-MeO-phenethylamine **6**) via 3-methoxy-tyramine (3-MeO-tyramine **7**) or 3,4,5-trihydroxy-phenethylamine (3,4,5-triOH-phenethylamine **8**). Lastly, two consecutive *O*-methylations catalyzed by one or more OMTs lead to mescaline **1** through 4-hydroxy-3,5-dimethoxy-phenethylamine (4-OH-3,5-diMeO-phenethylamine **9**). 3,4-diOH-5-MeO-phenethylamine **6** was also predicted to be *O*-methylated, probably by an OMT, to 3-hydroxy-4,5-dimethoxy-phenethylamine (3-OH-4,5-diMeO-phenethylamine **10**), which is thought to be the substrate of tetrahydroisoquinolines (THIQs) in Peyote (Kapadia and Fayez, 1970; Lundström, 1971) (Figures 1A and 1B). Although dopamine **3** and 3,4-diOH-5-MeO-phenethylamine **6** can potentially be biosynthesized via two routes, early studies have pointed toward a predominant biosynthetic pathway (highlighted by colored arrows in Figure 1A), with the other two routes suggested to be of minor importance (Lundström, 1971) (Figure 1A). *N*-Methylated metabolites have also been identified and are predicted to be biosynthesized by one or several *N*-methyltransferases (NMTs; Figure 1A). This pathway exhibits resemblances (in terms of enzyme types and intermediates) to those of various other groups of specialized plant metabolites, including betalains, catecholamines, hydroxycinnamic acid amides (HCAAs), benzylisoquinoline alkaloids, and other compounds that also originate from L-tyrosine **2** (Hagel and Facchini, 2013; Schenck and Maeda, 2018).

Despite the long-standing existence of this prediction, only one transcriptomic study, lacking functional analysis, has been reported (Ibarra-Laclette et al., 2015). Here, we confirmed the predicted biosynthetic pathway by identifying four groups of enzymes responsible for the six steps required for mescaline **1** biosynthesis. Notably, we identified a novel CYP76AD enzyme, which, to the best of our knowledge, is the first tyramine hydroxylase identified from plants. We also identified two groups of regiospecific OMTs and a highly similar methyltransferase (MT) ortholog capable of non-specifically *N*-methylating all the phenethyl intermediates within this pathway. Furthermore, we reconstructed the mescaline **1** pathway in both *Nicotiana benthamiana* plants and *Saccharomyces cerevisiae* yeast. Despite the obstacles posed by endogenous detoxification pathways in these organisms, we successfully produced all the intermediates from the pathway in *N. benthamiana*, albeit not reaching the final step.

RESULTS

Taxonomically distant cacti produce mescaline

To explore the mescaline **1** pathway, we first analyzed the metabolic profile of different tissues from Peyote cacti using high-resolution liquid chromatography–tandem mass spectrometry (LC–MS/MS) on the basis of previously identified metabolites in this cactus (Kapadia and Fayez, 1970; Lundström, 1971) (Figure 1). We observed most of the metabolites from the predicted pathway (metabolites **1**–**4**, **6**, **7**, **9**, and **10**; Supplemental Figure S1), as well as some other related molecules (Supplemental Figures S2 and S3). For instance, we identified a small peak with a mass and MS/MS spectrum similar to those of 3-MeO-tyramine **7**, which we attributed to 4-methoxy-tyramine (4-MeO-tyramine **11**; Figure 1A and Supplemental Figure S2D). We further putatively identified *N*-methylated products of most of these phenethylamines, probably produced via one or more NMTs (metabolites **12**–**25** Figure 1A and Supplemental Figure S2), and several major THIQs (metabolites **26**–**31**) previously identified in Peyote (Kapadia and Fayez, 1970; Lundström, 1971) (Figure 1B and Supplemental Figure S3). However, we did not observe peaks of L-DOPA **5,** 3,4,5-triOH-phenethylamine **8**, or their *N*-methylated derivatives in any of the tissues. The identified metabolites **1**–**31** are summarized in Supplemental Table S1.

In accordance with previous reports (Cassels and Sáez-Briones, 2018), we observed that mescaline **1** primarily accumulates in the Peyote top aerial part (button), particularly in the skin (Figures 2A and 2B). Using matrix-assisted laser desorption/ionization–mass spectrometry imaging we further spatially localized mescaline **1** in horizontal and lateral cryo-sections of Peyote. Consistent with our findings from extracted tissues (Figure 2A), mescaline **1** and structurally related phenethylamines (mainly *N*-methyl-4-hydroxy-3,5-dimethoxy-phenethylamine **20** and *N*-methyl-3-hydroxy-4,5-dimethoxy-phenethylamine **22**; Figures 2C–2G) with *m/z* 212.1281 Da ($[M + H]^+$) localized specifically to the button sub-epidermal green tissue (skin), which also contained a high content of chlorophyll b (Figures 2H and 2I). As previously reported by Cassels and Sáez-Briones (2018), hordenine **13** was most abundant in the lower parts of the stem and roots compared with the button (Figures 2J and 2K), exhibiting a unique localization pattern compared with other metabolites in these sections (Supplemental Figure S4 and S5).

We also analyzed the chemical repertoire of two other cacti: *Echinopsis pachanoi* (also known as *Trichocereus pachanoi*, San Pedro, or Wachuma), which is taxonomically distant from Peyote (Hernández-Hernández et al., 2011) but also produces large amounts of mescaline **1** (Figure 2A); and *Lophophora diffusa*, a

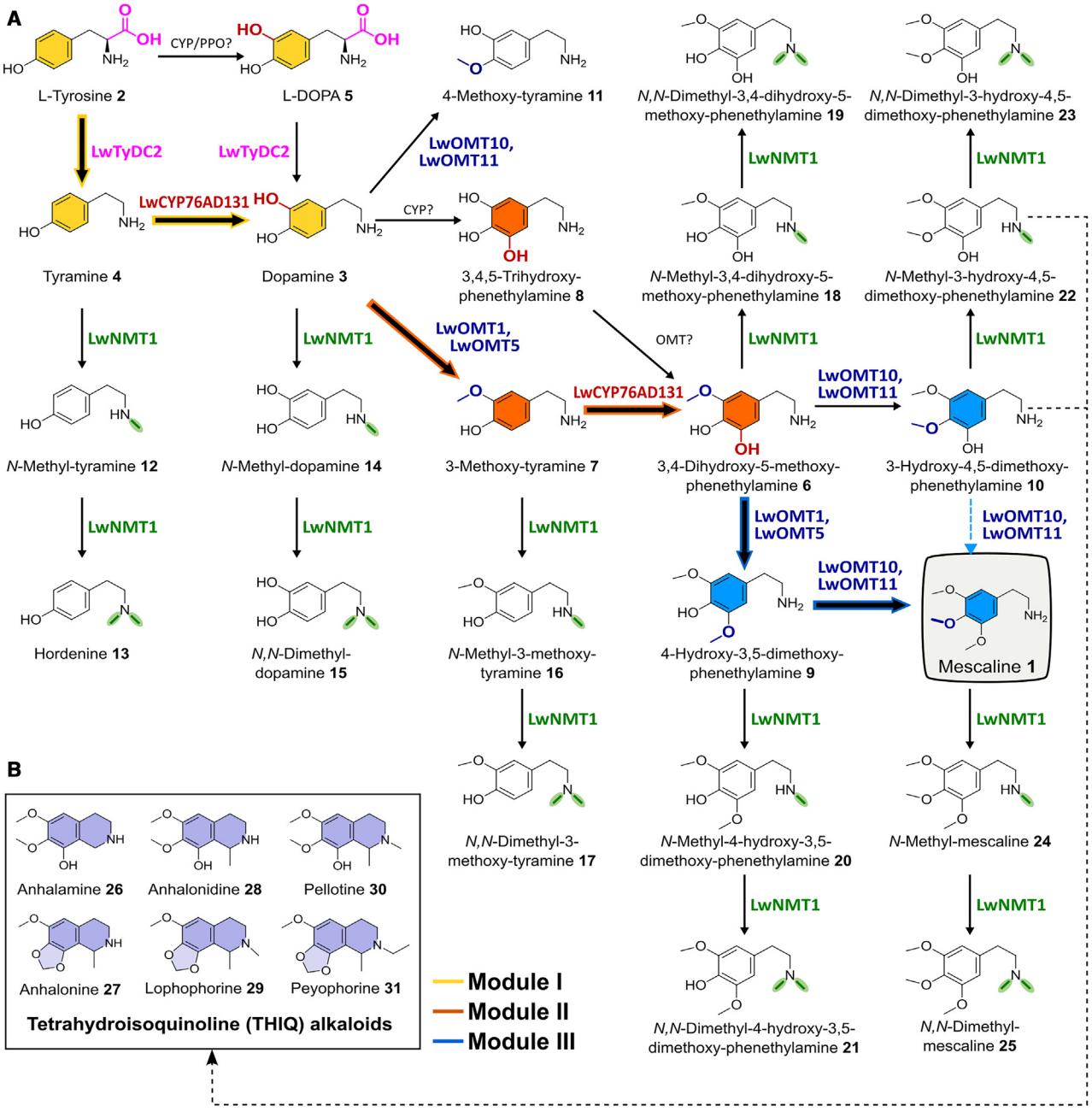

Figure 1. Predicted biosynthetic pathway of mescaline and related phenethylamines in Peyote (*L. williamsii*).
(A) Enzymes identified as being part of the pathway in this study are indicated by their names. TyDC, L-tyrosine/L-DOPA decarboxylase; CYP, cytochrome P450; PPO, polyphenol oxidase; OMT, *O*-methyltransferase; NMT, *N*-methyltransferase. Putative types of enzymes appear in black. Certain intermediates may be synthesized through multiple routes, as depicted by the main metabolites marked in yellow, orange, and blue (modules I–III, respectively). However, a primary pathway has been proposed, as outlined by colored arrows (Lundström, 1971). Decarboxylation, hydroxylation, *O*-methylation, and *N*-methylation reactions are marked in pink, red, blue, and green, respectively. **(B)** Chemical structures of several known tetrahydroisoquinolines (THIQs, highlighted in purple) from Peyote putatively assigned according to LC–MS/MS fragmentation (Supplemental Figure S3). The broken lines represent potential biosynthetic routes (Lundström, 1971) not studied in this work. The identified metabolites **1**–**31** are summarized in Supplemental Table S1.

taxonomically close relative of Peyote that was reported to lack mescaline **1** (Figure 2L). Each species exhibited a unique chemical profile, suggesting possible differences in biosynthetic pathways (Supplemental Figure S6). For example, we observed that San Pedro accumulates similar metabolites to Peyote for the production of mescaline **1**, but we did not detect *N*-methylated (apart from *N*-methylmescaline **24**; Supplemental Figure S6) or THIQ (apart from the putatively identified anhalamine **26**, Supplemental Figure S6) metabolites in this cactus. In *L. diffusa*, on the other hand, we identified several THIQs from Peyote, yet we could not detect mescaline **1** or metabolites **3**–**9** but rather their *N*-methylated forms (Supplemental

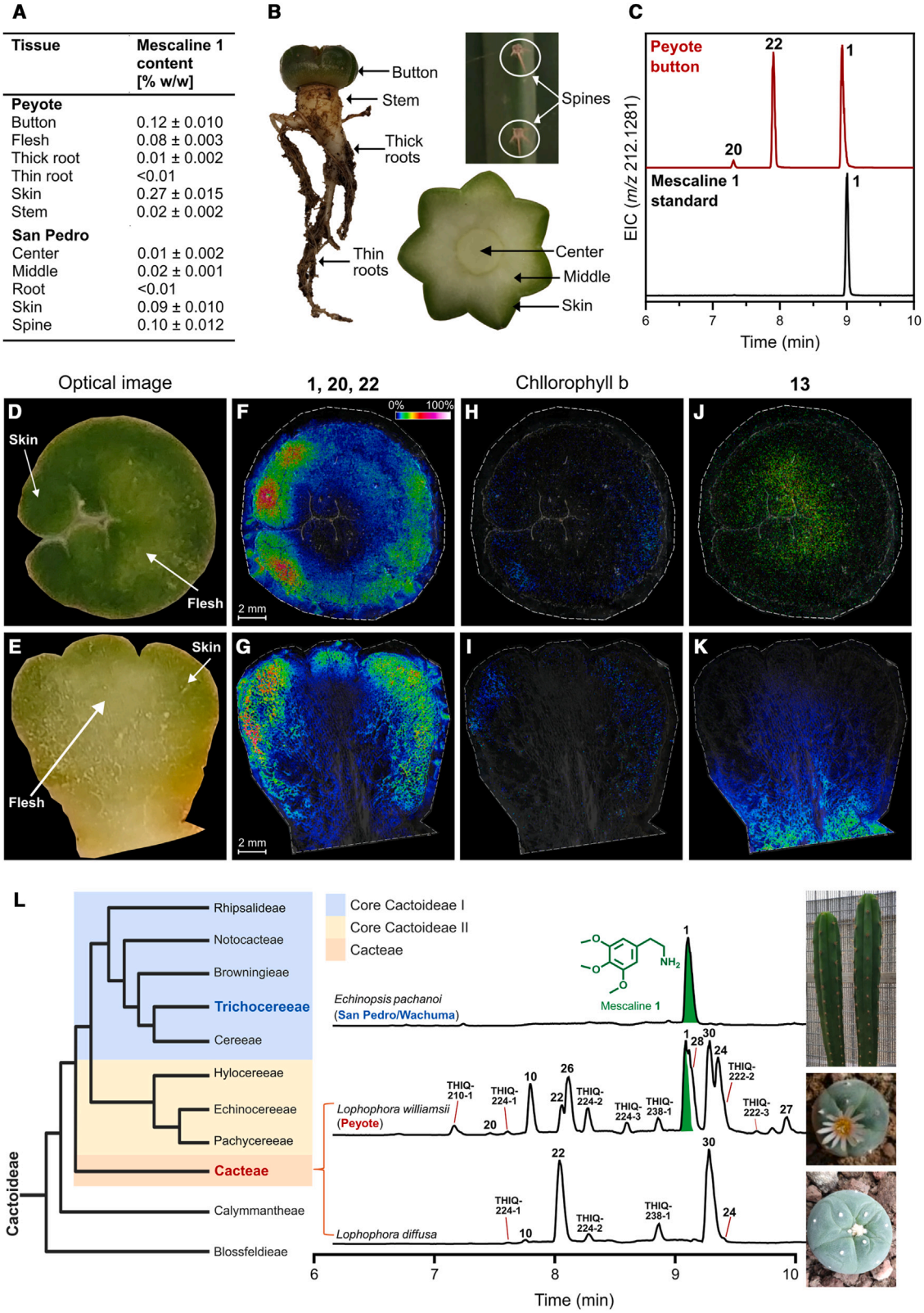

Figure 2. Taxonomically distant cacti produce mescaline.

(A) Absolute quantification of mescaline **1** in different tissues of Peyote (*L. williamsii*) and San Pedro (*E. pachanoi*) cacti (%w/w per fresh weight, mean ± SD; $n = 3$). The tissues used for the analysis are marked in **(B).(C)** Extracted ion current chromatograms from a Peyote button extract versus mescaline **1** standard.**(D–K)** Matrix-assisted laser desorption/ionization–mass spectrometry imaging (MALDI–MSI) of Peyote sections ($n = 2$). Optical images of

(legend continued on next page)

Figure S6). This may suggest very rapid *N*-methylation activity in *L. diffusa* that impedes the accumulation of non-*N*-methylated metabolites.

Possible biosynthetic pathways of San Pedro and *L. diffusa* constructed on the basis of the main observed metabolites are shown in Supplemental Figures S7 and S8, respectively. Notably, in all the species, greater accumulation of most metabolites from the early steps of the biosynthetic pathway (**2–4** and their *N*-methylated forms) was observed in the roots, whereas the other metabolites were more abundant in the skin and spines (the flesh area around the spines in San Pedro; Figure 2B). Interestingly, Peyote was the only species of the three that accumulated metabolites with a methylenedioxy group, as in anhalonine **27**, lophophorine **29**, peyophorine **31** (Supplemental Figure S6), and other types reported by others (Kapadia and Fayez, 1970; Lundström, 1971; Cassels and Sáez-Briones, 2018). This reaction, commonly catalyzed by CYPs (Hagel and Facchini, 2013), suggests the presence of a unique enzyme in Peyote compared with the other two cacti.

De novo sequencing and annotation of the Peyote genome

To identify the enzymes responsible for mescaline **1** biosynthesis in Peyote, we obtained an 11× PacBio HiFi-based assembly. Approximately 37 000 contigs were generated (11 × 3.2 Gb genome size, N50 127 kb), and genes were annotated using PacBio Iso-Seq and Illumina paired-end-based transcriptomes from different tissues (Supplemental Figure S9A; Supplemental Table S2–S4). The BUSCO completeness values were 94.1% for the primary genome assembly, 97.6% for the Illumina-based transcriptome, and 88.9% for the PacBio Iso-Seq transcriptome (Supplemental Figure S9B). After collapsing all transcripts into the genome, we determined that ∼300 000 transcripts were transcribed from 35 000 different loci. The ability to differentiate between paralogs and isoforms of genes enabled us to significantly narrow down the candidate gene lists for functional testing. Interestingly, two groups of MTs were organized in tandem gene clusters (Supplemental Figure S9C; Supplemental Table S5).

On the basis of Figure 1A, we focused our study on four enzyme families, namely TyDC, CYP, PPO, and MTs. Approximately five *TyDC*, 270 *CYP*, three *PPO*, and 350 *MT* genes were annotated in the Peyote genome (Supplemental Table S6). Out of these gene sets, we selected candidates on the basis of their relative expression levels in Peyote tissues, prioritizing those highly expressed in the skin and button and showed low expression in thin or thick roots (Supplemental Table S6). We also considered protein size, the presence of protein family (Pfam) motifs, and similarity to enzymes functionally characterized as involved in secondary metabolite biosynthesis (Supplemental Table S6).

A single TyDC from Peyote catalyzes the decarboxylation of both L-tyrosine and L-DOPA

TyDC enzymes catalyze the decarboxylation of L-tyrosine **2** into tyramine **4** and/or L-DOPA **5** into dopamine **3** (Figure 1A). These enzymes are categorized as aromatic-L-amino-acid decarboxylases and depend on pyridoxal 5-phosphate (PLP). Among the five *TyDC* genes identified in the genome, *TyDC2* exhibited the highest expression in all tissues, particularly in thick roots (Supplemental Table S7), consistent with the higher concentrations of metabolites from the early stages of the pathway (Figure 1A and Supplemental Figure S6). We first recombinantly expressed *TyDC2*, *TyDC4*, and *TyDC5* in *Escherichia coli* and tested their activities using lysates supplemented with PLP and either L-tyrosine **2** or L-DOPA **5** as substrates. We could not amplify the additional two candidates, likely owing to their very low expression in all tissues. Only TyDC2 (LwTyDC2) was functional, showing activity for the production of both tyramine **4** from L-tyrosine **2** and dopamine **3** from L-DOPA **5** (Figure 3A). Phylogenetic analysis revealed that LwTyDC2 clusters with the *Papaver somniferum* and *Thalictrum flavum* TyDCs, which exhibit similar activities toward both substrates (Facchini et al., 2000) (Supplemental Figure S10A; Supplemental Table S8). Partial purification of the enzyme revealed significantly better *in vitro* kinetics of LwTyDC2 with L-tyrosine **2** than with L-DOPA **5** (Michaelis–Menten *K*m (μM) = 229 ± 77 versus 1575 ± 503, respectively; Figures 3B–3D and Supplemental Figure S10B).

We next transiently expressed LwTyDC2 by agro-infiltration in *N. benthamiana* leaves. Capitalizing on the existing pool of L-tyrosine **2** in the plant, we observed a significant increase in the peak areas of the naturally occurring tyramine **4**, dopamine **3**, and 3-MeO-tyramine **7** along with the appearance of small peaks of L-DOPA **5** (Figure 3E). Co-expression of this enzyme with BvCYP76AD6, a betalain-related red-beet enzyme that hydroxylates L-tyrosine **2** to L-DOPA **5** (Polturak et al., 2016), led to significant accumulation of L-DOPA **5** and a substantial increase in dopamine **3** and its *O*-methylated product 3-MeO-tyramine **7** (Figure 3E), providing further evidence for LwTyDC2 activity with both substrates *in planta*.

A member of the yet uncharacterized CYP76AD family γ clade hydroxylates tyramine to dopamine and 3-MeO-tyramine to 3,4-diOH-5-MeO-phenethylamine

We next mined our sequencing data for potential hydroxylating enzymes active in modules I and II (Figure 1A). We first selected

Peyote **(D)** horizontal and **(E)** lateral sections. **(F** and **G)** MALDI–MSI of *m/z* 212.128 ± 0.01 Da, which corresponds mainly to the protonated masses of mescaline **1**, *N*-methyl-4-hydroxy-3,5-dimethoxy-phenethylamine **20**, and *N*-methyl-3-hydroxy-4,5-dimethoxy-phenethylamine **22** (as shown in **C**). **(H** and **I)** MALDI–MSI of *m/z* 893.545 ± 0.01 Da, which corresponds to chlorophyll b. **(J** and **K)** MALDI–MSI of *m/z* 166.123 ± 0.01 Da which corresponds to hordenine **13**. Skin (the sub-epidermal green area) and flesh tissues are marked on the optical images **(D–E)** of Peyote sections.**(L)** Phylogenetic relationships of the Cactoideae subfamily and total ion current chromatograms of *L. diffusa* skin, Peyote skin, and San Pedro flesh surrounding the spines (spine) show that Peyote and San Pedro, two taxonomically distant cacti (from the Cacteae and Trichocereae subfamilies, respectively) produce mescaline **1**, whereas *L. diffusa*, a taxonomically close relative of Peyote (both Cacteae), contains none. Representative images of each species appear next to the corresponding chromatogram. The major metabolites identified are marked on each chromatogram (full details appear in Supplemental Table S1). Chromatograms were normalized to the highest value. Owing to the limited availability of molecular data, this proposed phylogenetic hypothesis relies on a synthesis of existing literature (Cassels and Sáez-Briones, 2018) and represents a preliminary interpretation of the relationships among the taxa under study.

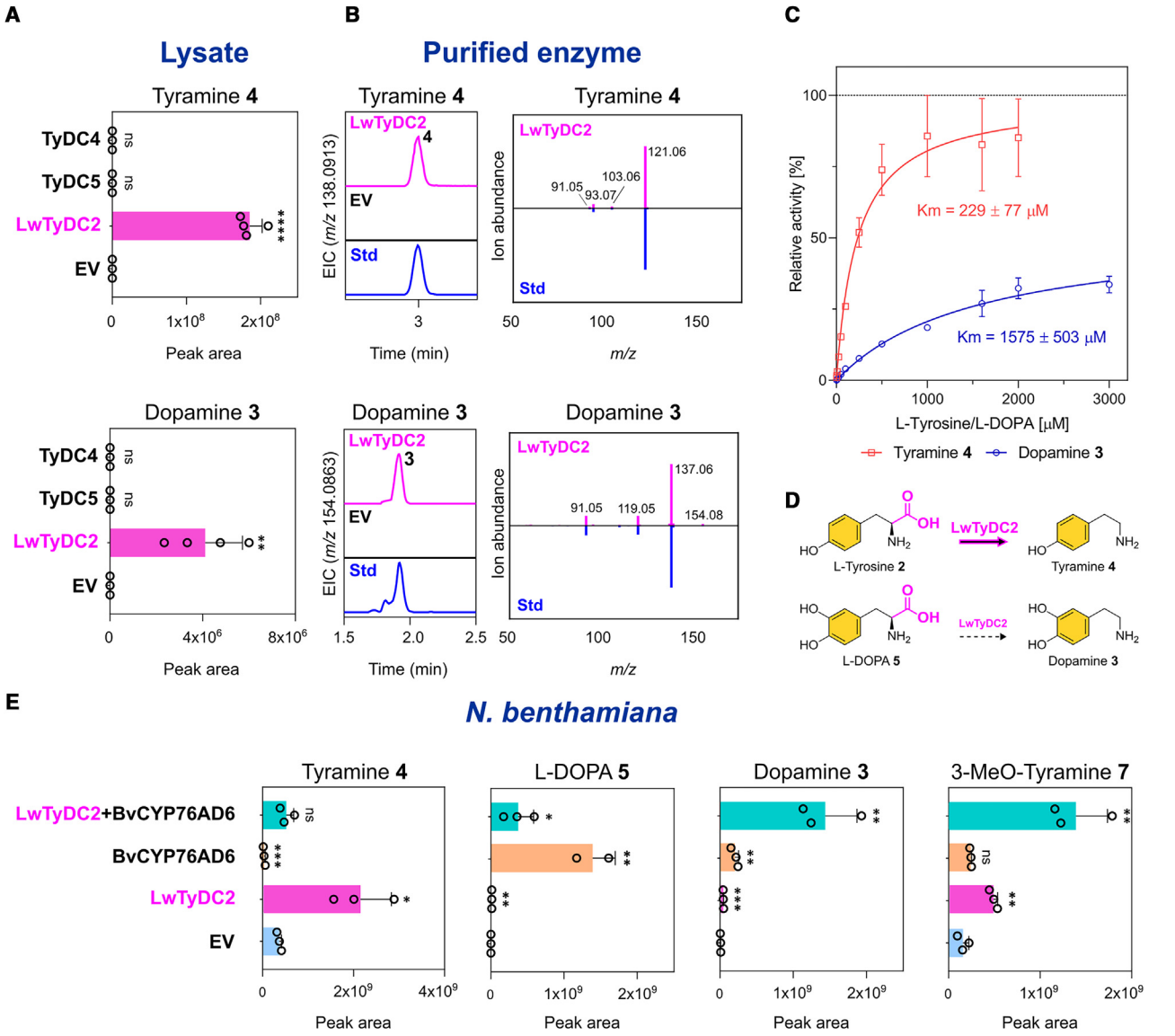

Figure 3. Characterization of Peyote TyDC.

(A) Comparison of ion abundance peak areas (mean ± SD; n = 3–4 biological replicates) following *in vitro* assays using lysates from *E. coli* cells transformed with TyDC candidates, using either L-tyrosine **2** or L-DOPA **5** as a substrate in the presence of pyridoxal 5-phosphate (PLP). Statistical significance was assessed using a two-tailed Student's t-test with an assumption of unequal variance; $^{**}p < 0.01$, $^{****}p < 0.0001$.**(B)** Extracted ion chromatograms and MS/MS spectra of tyramine **4** and dopamine **3** following *in vitro* assays with purified LwTyDC2 and either L-tyrosine **2** or L-DOPA **5** as a substrate in the presence of PLP (n = 3 biological replicates). Chromatograms were normalized to the highest value.**(C and D) (C)** Relative activity of LwTyDC2 with L-tyrosine **2** or L-DOPA **5** in the presence of PLP shows better *in vitro* kinetics of LwTyDC2 with L-tyrosine **2** than with L-DOPA **5 (D)**. The Michaelis–Menten K_m values were calculated from absolute concentrations of tyramine **4** or dopamine **3** acquired using various (2.5 μM to 3 mM) concentrations of each substrate (mean ± SD; n = 3 technically independent experiments).**(E)** Comparison of ion abundance peak areas (mean ± SD; n = 3 biological replicates) following transient expression of LwTyDC2, BvCYP76AD6, and their combinations in *N. benthamiana* leaves. Statistical significance was assessed using a two-tailed Student's t-test with an assumption of unequal variance; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All metabolites were identified by exact mass, retention time, and MS/MS spectra. EV, empty vector.

two candidates that were highly expressed in Peyote aerial tissues (Supplemental Table S7) and exhibited 51% and 62% protein sequence similarity to the red-beet betalain-producing enzyme BvCYP76AD6 (i.e., CYP76F270 and CYP76AD131; Supplemental Figure S11A). These CYPs were individually expressed in *N. benthamiana* leaves, taking advantage of the available pool of L-tyrosine **2** and tyramine **4** in the plant. Notably, neither of the two candidates produced L-DOPA **5** (Supplemental Figure S11B). However, the expression of CYP76AD131 resulted in a decrease in tyramine **4**, accompanied by an increase in dopamine **3** and 3-MeO-tyramine **7**, and in the *de novo* production of 3,4-diOH-5-MeO-phenethylamine **6**, 4-OH-3,5-diMeO-phenethylamine **9**, and 3-OH-4,5-diMeO-phenethylamine **10** (Figure 4A). On the basis of these results, we concluded that

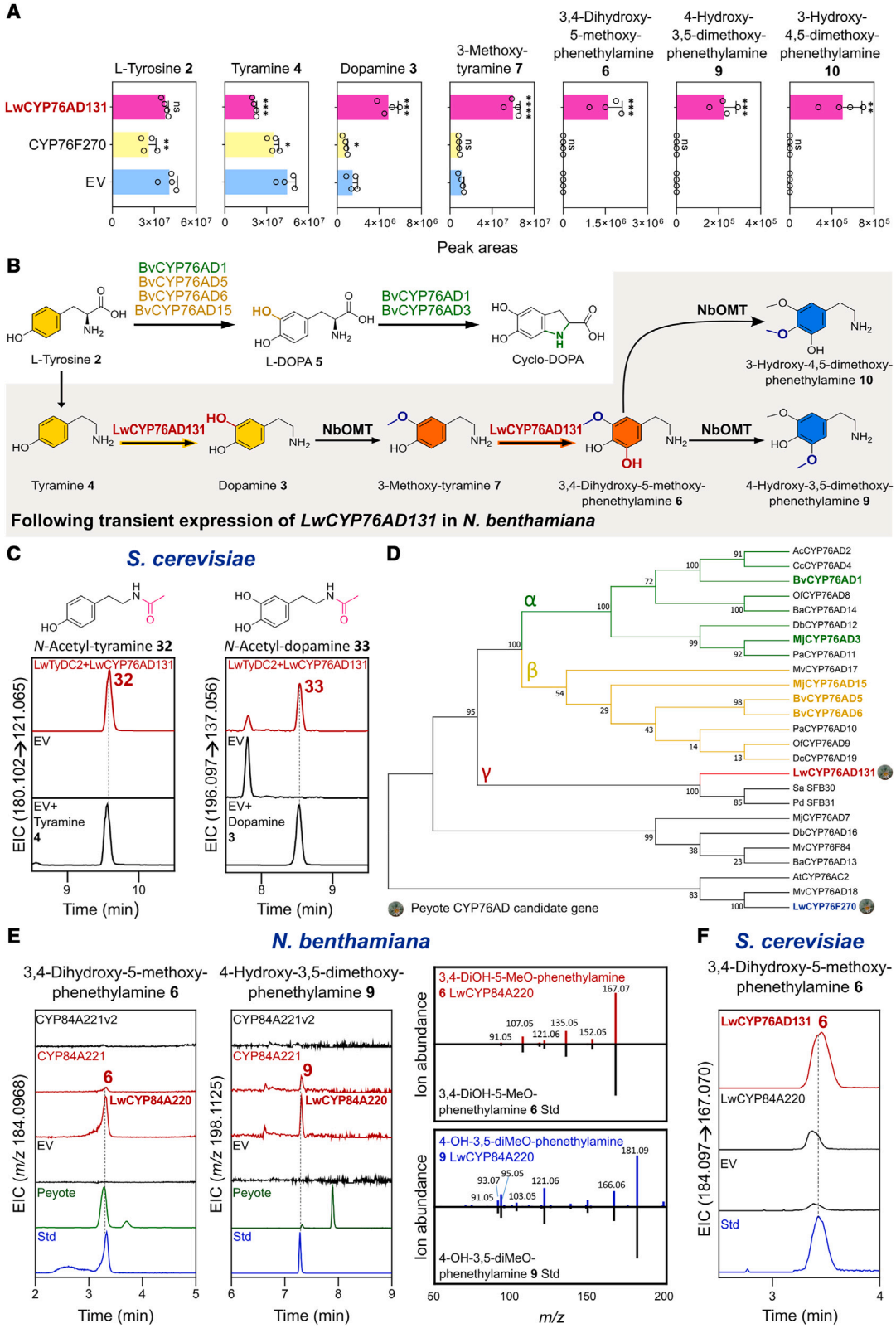

Figure 4. Characterization of Peyote CYP enzymes.

(A) Ion abundance peak areas (mean ± SD; *n* = 4 biological replicates) following transient expression of CYP76F270 and LwCYP76AD131 in *N. benthamiana* leaves. LwCYP76AD131 catalyzes two hydroxylation reactions (marked in red on the structures) of tyramine **4** to dopamine **3** and 3-methoxy-tyramine **7** to 3,4-dihydroxy-5-methoxy-phenethylamine **6**. Statistical significance was assessed using a two-tailed Student's *t*-test with an

(legend continued on next page)

LwCYP76AD131 regioselectively catalyzes the two meta hydroxylations (marked in red in Figure 4B) of both tyramine **4** to dopamine **3** and 3-MeO-tyramine **7** to 3,4-diOH-5-MeO-phenethylamine **6** (Figure 4B). The observed increase in 3-MeO-tyramine **7**, 4-OH-3,5-diMeO-phenethylamine **9**, and 3-OH-4,5-diMeO-phenethylamine **10** in these experiments was most likely mediated by one or several endogenous OMTs (termed here NbOMTs) present in *N. benthamiana* (Figure 4B). CYP76F270 did not show a significant increase in dopamine **3** or 3,4-diOH-5-MeO-phenethylamine **6** and was not further characterized in this study (Figure 4A).

To validate the enzyme activity observed in *N. benthamiana*, we co-expressed LwTyDC2 and LwCYP76AD131 in *S. cerevisiae* yeast and noted the production of tyramine **4** and a minor peak of dopamine **3** (Supplemental Figure S11C), accompanied by their putative *N*-acetylated forms (Figure 4C and Supplemental Figure S11D). Consistent with the results in *N. benthamiana*, we did not observe L-DOPA **5** in these samples or after expression of LwCYP76AD131 alone (Supplemental Figure S11E). Incubation of empty vector (EV)-transformed yeast with either tyramine **4** or dopamine **3** resulted in the appearance of peaks with masses and retention times similar to those of *N*-acetyl-tyramine **32** and *N*-acetyl-dopamine **33** (Figure 4C). Identification of these *N*-acetylated metabolites was confirmed using analytical standards (Supplemental Figure S11D). The detection of dopamine **3** and *N*-acetyl-dopamine **33** confirms the tyramine **4** hydroxylating activity of LwCYP76AD131 and suggests competition for the newly formed phenethylamines between the expressed Peyote enzymes and an endogenous *N*-acetylating enzyme in yeast.

CYP76AD subfamily members in the betalain-producing Caryophyllales species can be categorized into three distinct clades (Polturak et al., 2016; Sunnadeniya et al., 2016). Functionally characterized enzymes in the β clade (CYP76AD5/6/15) specialize in catalyzing the hydroxylation of L-tyrosine **2**, whereas enzymes in the α clade (CYP76AD1/3) catalyze both L-tyrosine **2** hydroxylation and L-DOPA **5** oxidative cyclization (Figures 4B and 4D). Notably, LwCYP76AD131 is a member of the γ clade (Figures 4B and 4D), which, until its discovery in this study, lacked any functionally characterized members (Sunnadeniya et al., 2016).

To test other potential hydroxylating enzymes active in this pathway, we selected nine additional CYP candidates on the basis of their relative expression levels and similarity to previously reported CYPs with known functions (Plant cytochrome P450 database, 2020) (Supplemental Figure S11A; Supplemental Tables S7 and S8). We also selected a PPO homolog that was highly expressed in the aerial parts of the plant (Supplemental Table S7). None of the new candidates produced statistically significant changes in dopamine **3**, tyramine **4**, or 3-MeO-tyramine **7** levels (Supplemental Figure S12A), nor did they produce L-DOPA **5** following transient expression in *N. benthamiana* leaves (Supplemental Figure S12B). Interestingly, the expression of CYP84A220 led to reduced concentrations of dopamine **3** and 3-MeO-tyramine **7** compared with the EV control (Supplemental Figure S12A). Supplementation with 50 μg ml^{-1} of 3-MeO-tyramine **7** showed that CYP84A220 (LwCYP84A220), and to a lesser extent CYP84A221, produced 3,4-diOH-5-MeO-phenethylamine **6** (Supplemental Figure S12C). None of the infiltrated enzymes produced peaks of 3,4,5-triOH-phenetethylamine **8**. Increasing the substrate concentration with LwCYP84A220, CYP84A221, and CYP84A221v2 (CYP84A221v2 has 99% sequence similarity to clusters with CYP84A221; Supplemental Figure S11A) resulted in a higher intensity of 3,4-diOH-5-MeO-phenethylamine **6** and the appearance of 4-OH-5-diMeO-phenethylamine **9**, further validating the activity of LwCYP84A220 and CYP84A221 but not of CYP84A221v2 (Figure 4E). The appearance of 4-OH-3,5-diMeO-phenethylamine **9** was likely mediated by an NbOMT, as suggested previously (Figure 4B). LwCYP84A220 and CYP84A221 belong to the same clade as CYP84A1 (Supplemental Figure S11A), which has previously been identified as a ferulic acid hydroxylase (Meyer et al., 1996). This may explain the comparable regioselectivity and catalytic functions observed with these enzymes.

We next individually expressed LwCYP76AD131 and LwCYP84A220 in yeast and supplemented them with 3-MeO-tyramine **7** as the substrate. After 2 days of galactose induction, we observed peaks of 3,4-diOH-5-MeO-phenethylamine **6** in the samples expressing LwCYP76AD131 (Figure 4F). Because the peaks observed with LwCYP84A220 could not be confirmed by MS/MS owing to their low abundance, we inferred that LwCYP76AD131 exhibited superior catalytic activity for the production of 3,4-diOH-5-MeO-phenethylamine **6**. This confirmed

assumption of unequal variance; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. All the metabolites were identified by exact mass, retention time, and MS/MS spectra.**(B)** Scheme showing the reactions catalyzed by LwCYP76AD131 and an unknown *O*-methyltransferase from *N. benthamiana* (NbOMT). Reactions catalyzed by other functionally characterized members of the CYP76AD subfamily are shown for reference.**(C)** Extracted ion chromatograms and chemical structures of *N*-acetyl tyramine **32** and *N*-acetyl dopamine **33** obtained after co-expression of LwTyDC2 and LwCYP76AD131 in yeast ($n = 4$ biological replicates). Metabolites were analyzed using the acquired data-dependent MS/MS (dd-MS^2) data. Chromatograms were normalized to the highest value. *N*-Acetylation occurs in yeast by an endogenous acetylating enzyme, as shown for samples transformed with the EV and supplemented with each intermediate (blue chromatograms).**(D)** Phylogenetic analysis of the CYP76AD subfamily within betalain-producing Caryophyllales species, showing three distinct clades (α, β, and γ; clades were assigned according to Sunnadeniya et al. (2016); a full list of GenBank accession numbers appears in Supplemental Table S8). Functionally characterized enzymes in the α (CYP76AD1/3), β (CYP76AD5/6/15), and γ clades (LwCYP76AD131) appear in **(B)** and **(C)** in green, yellow, and red, respectively; the Peyote CYP76AD homolog CYP76F270 appears in blue. The maximum likelihood tree was constructed with 1000 bootstrap tests on the basis of a MUSCLE multiple alignment using MEGA11 software.**(E)** LC–MS extracted ion chromatograms of 3,4-dihydroxy-5-methoxy-phenethylamine **6** and 4-hydroxy-3,5-dimethoxy-phenethylamine **9** following transient expression of LwCYP84A220, CYP84A221, and CYP84A221v2 in *N. benthamiana* leaves with 100 μg ml^{-1} 3-methoxy-tyramine **7** supplementation. All the metabolites were identified by exact mass, retention time, and MS/MS spectra. Metabolites were analyzed using the acquired dd-MS^2 data.**(F)** Extracted ion chromatograms of 3,4-dihydroxy-5-methoxy-phenethylamine **6** following expression of LwCYP76AD131 or LwCYP84A220 in yeast ($n = 3$ biological replicates). Metabolites were analyzed using the acquired dd-MS^2 data. Chromatograms were normalized to the highest value.

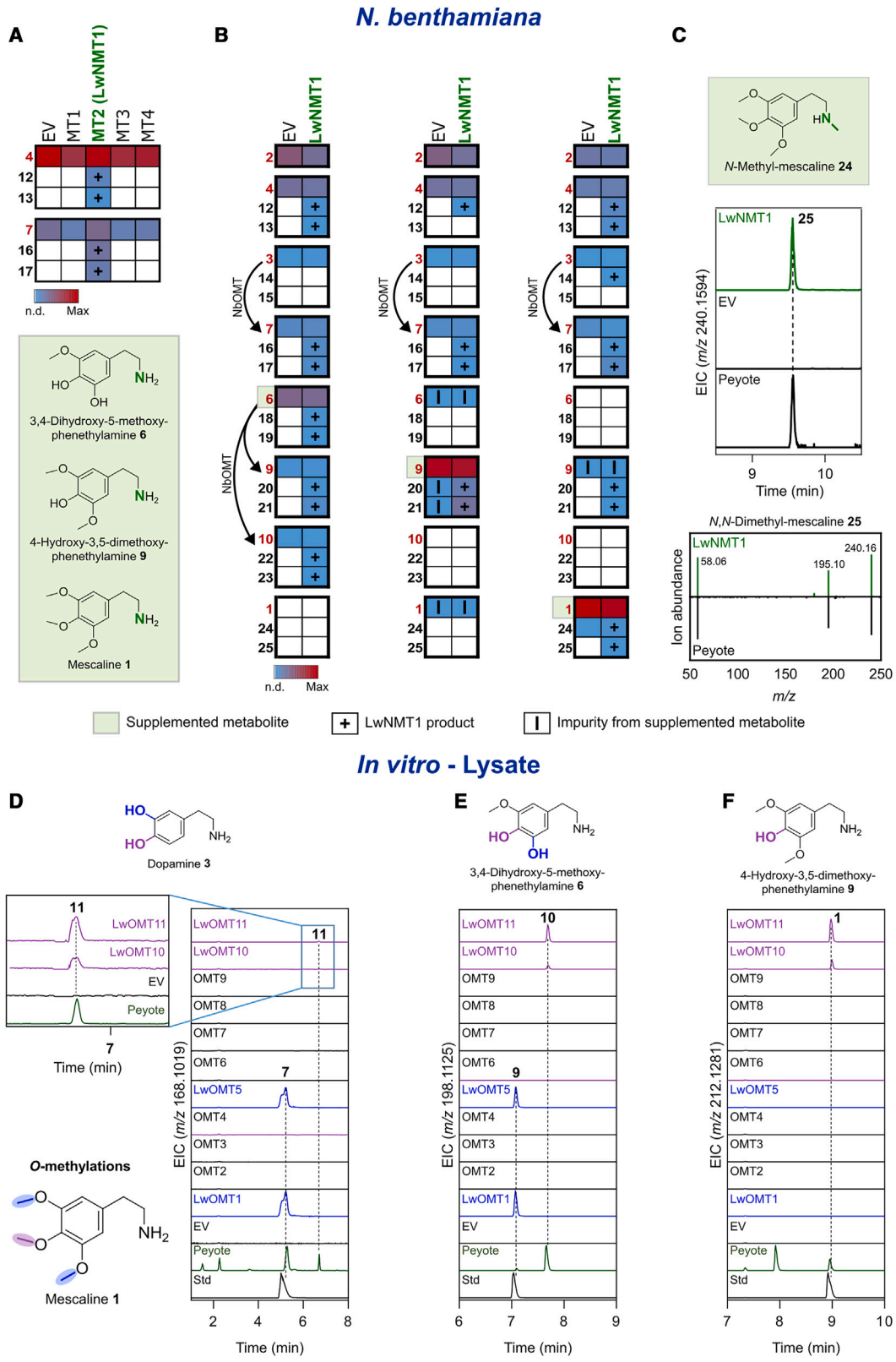

Figure 5. Functional characterization of Peyote MTs.

(A and **B)** Heatmaps of ion abundance peak areas (mean, *n* = 3 biological replicates) following transient expression in *N. benthamiana* leaves of **(A)** individual MTs and **(B)** only MT2 (i.e., LwNMT1) following supplementation with 3,4-dihydroxy-5-methoxy-phenethylamine **6,** 4-hydroxy-3,5-dimethoxy-phenethylamine **9**, or mescaline **1**. Main phenethylamines from the mescaline **1** biosynthetic pathway are marked in red, and

(legend continued on next page)

the second hydroxylating activity of LwCYP76AD131 toward 3,4-diOH-5-MeO-phenethylamine **6**.

A single NMT enzyme in Peyote catalyzes multiple *N*-methylation reactions

To identify the OMTs in modules II and III, we first selected four candidates with protein sequence similarities to other MTs from red beet (MT1–MT4; Supplemental Figure S13A; Supplemental Tables S7 and S8) and transiently co-expressed them (in one group) in *N. benthamiana* leaves with and without 3,4-diOH-5-MeO-phenethylamine **6** supplementation (supplementation was performed to provide the substrate for the first methylation reaction in module III; Figure 1A). Although our aim was to identify OMT enzymes, one or more of the MT candidates in the infiltrated group demonstrated *N*-methylation activity on several of the intermediates (Supplemental Figure S13B). We then individually infiltrated each of the MT candidates and identified MT2 as the active NMT (termed here LwNMT1; Figure 5A). Furthermore, supplementation with 3,4-diOH-5-MeO-phenethylamine **6**, 4-OH-3,5-diMeO-phenethylamine **9**, or mescaline **1** resulted in the production of numerous *N*-phenethylamine products, demonstrating *N*-methylation and *N*,*N*-dimethylation activity by this single enzyme (products **12**–**14** and **16**–**25** in Figure 5B). We next supplemented plants with *N*-methyl-mescaline **24** and confirmed that the second *N*-methylation is catalyzed by LwNMT1 and not by the plant (Figure 5C). Lastly, we expressed LwNMT1 recombinantly in *E. coli* and, using lysates in the presence of S-adenosyl L-methionine (SAM) and dopamine **3**, 3,4-diOH-5-MeO-phenethylamine **6**, or 4-OH-3,5-diMeO-phenethylamine **9** as substrates, confirmed its broad *N*-methylation and *N*,*N*-dimethylation activity *in vitro* (products **14**, **15**, **18**, **20**, and **21** in Supplemental Figure S13C–S13E). Notably, the *in vitro* activity of LwNMT1 with 3,4-diOH-5-MeO-phenethylamine **6** was relatively low, yielding only a small peak of *N*-methyl-3,4-diOH-5-MeO-phenethylamine **18** and no *N*,*N*-dimethyl-3,4-diOH-5-MeO-phenethylamine **19** (Supplemental Figure S13D), although both products were detected in the *N. benthamiana* assays (Figure 5B). Overall, from the expression of LwNMT1 both in *N. benthamiana* and *in vitro*, we concluded that this enzyme broadly catalyzes the *N*-methylation and *N*,*N*-dimethylation of all intermediates from the mescaline **1** pathway, as illustrated in Figure 1A.

Two groups of OMTs participate in mescaline production

Considering the endogenous *O*-methylation activity mediated by the NbOMTs (**3** to**7**, **6** to –**9**, and **6** to **10**; Figure 5B), we recombinantly expressed 11 candidate enzymes in *E. coli* to evaluate their activities *in vitro* (OMT1–OMT7 and OMT10–OMT11; Figure 5B; Supplemental Table S7). We also included MT3 and MT4 from our first MT screen among these candidates (termed OMT8 and OMT9, respectively; Supplemental Figure S13A). These 11 candidates contained conserved amino acid sequence regions characteristic of OMTs (Ibrahim et al., 1998) (Supplemental Figure S14A). *OMT10*, *OMT11*, and *OMT12* (an ortholog of *OMT11* with 99.2% nucleotide similarity and an identical protein sequence; Supplemental Table S7) were found in tandem in the genome (Supplemental Figure S9A). *OMT10* shared 98.9% similarity with *OMT11* and *OMT12* (Supplemental Table S7).

We next tested lysates from *E. coli* expressing the OMT1–OMT11 candidate enzymes in the presence of SAM and dopamine **3**, 3,4-diOH-5-MeO-phenethylamine **6**, or 4-OH-3,5-diMeO-phenethylamine **9** as the substrates in modules II and III (Figure 1A). Two distinct groups of enzymes were identified: OMT1 and OMT5 (i.e., LwOMT1 and LwOMT5), which *O*-methylated the two meta hydroxyl groups found in mescaline **1** (**3** to **7** and **6** to **9**; Figure 1A), and OMT10 and OMT11 (i.e., LwOMT10 and LwOMT11), which targeted the para hydroxyl (**3** to **11**, **6** to **10**, and **9** to **1**; Figure 1A) (Figures 5D–5F for substrates **3**, **6**, and **9**, respectively). Curiously, LwOMT10 and LwOMT11 enzymes also generated small quantities of mescaline **1** with 3,4-diOH-5-MeO-phenethylamine **6** as the substrate (Figure 1A). This is likely due to secondary activity of these enzymes on the resulting 3-OH-4,5-diMeO-phenethylamine **10**. We further purified LwOMT1 and LwOMT11 and repeated the assays, obtaining similar results (Supplemental Figure S14B and S14C).

According to the phylogeny, LwOMT1, LwOMT5, and LwNMT1 belong to the hydroxycinnamic acid (HCA) clade and exhibit the closest resemblance to the caffeic acid 3-OMT from red beet (BvHCA1 in Supplemental Figure S13A). By contrast, LwOMT10–LwOMT12 were similar to phenol OMTs (Lam et al., 2007) (Supplemental Figure S13A; Supplemental Table S8). This phylogenetic grouping aligns with the closely related chemical structures of the intermediates within this pathway and other simple phenols.

Regioselectivities of LwOMT1, LwOMT11, and LwNMT1 are related to spatial proximity of SAM to specific target sites

Functional characterization of MTs demonstrated their regioselectivity toward specific hydroxyl or amine groups. For instance,

their *N*-methylated products appear below on the same square. Metabolites used for feeding are highlighted in green, and their structures are shown for reference. A single NMT from Peyote (i.e., LwNMT1) was identified with broad substrate specificity (*N*-methylated products produced by LwNMT1 are marked with "+"). *In planta* *O*-methylation activities (**3** to **7**, **6** to **9**, and **6** to **10**) were attributed to one and/or several endogenous OMTs from *N. benthamiana* (NbOMTs) because these products were also observed in plants infiltrated with an EV. Low-abundance impurities from the standards used for supplementation are marked with "I." Peaks that were not detected (n.d.) appear in white.**(C–F)** **(C)** Extracted ion chromatogram and MS/MS spectra of *N*,*N*-dimethyl-mescaline **25** produced in *N. benthamiana* following transient expression of LwNMT1 and supplementation with *N*-methyl-mescaline **24**. Chromatograms were normalized to the highest value. Extracted ion chromatograms of lysates from *E. coli* transformed with OMT1 to OMT11 candidate enzymes in the presence of SAM and **(D)** dopamine **3**, **(E)** 3,4-dihydroxy-5-methoxy-phenethylamine **6**, or **(F)** 4-hydroxy-3,5-dimethoxy-phenethylamine **9**. The active enzymes, LwOMT1 and LwOMT5, *O*-methylate the meta hydroxyls in mescaline **1** (marked in blue), whereas LwOMT10 and LwOMT11 *O*-methylate the para hydroxyl (marked in purple). Substrate structures are provided for reference, and the site of *O*-methylation is indicated accordingly. Products in each assay were identified by comparison with analytical standards and/or an extract from a Peyote skin sample. Chromatograms were normalized to the highest value.

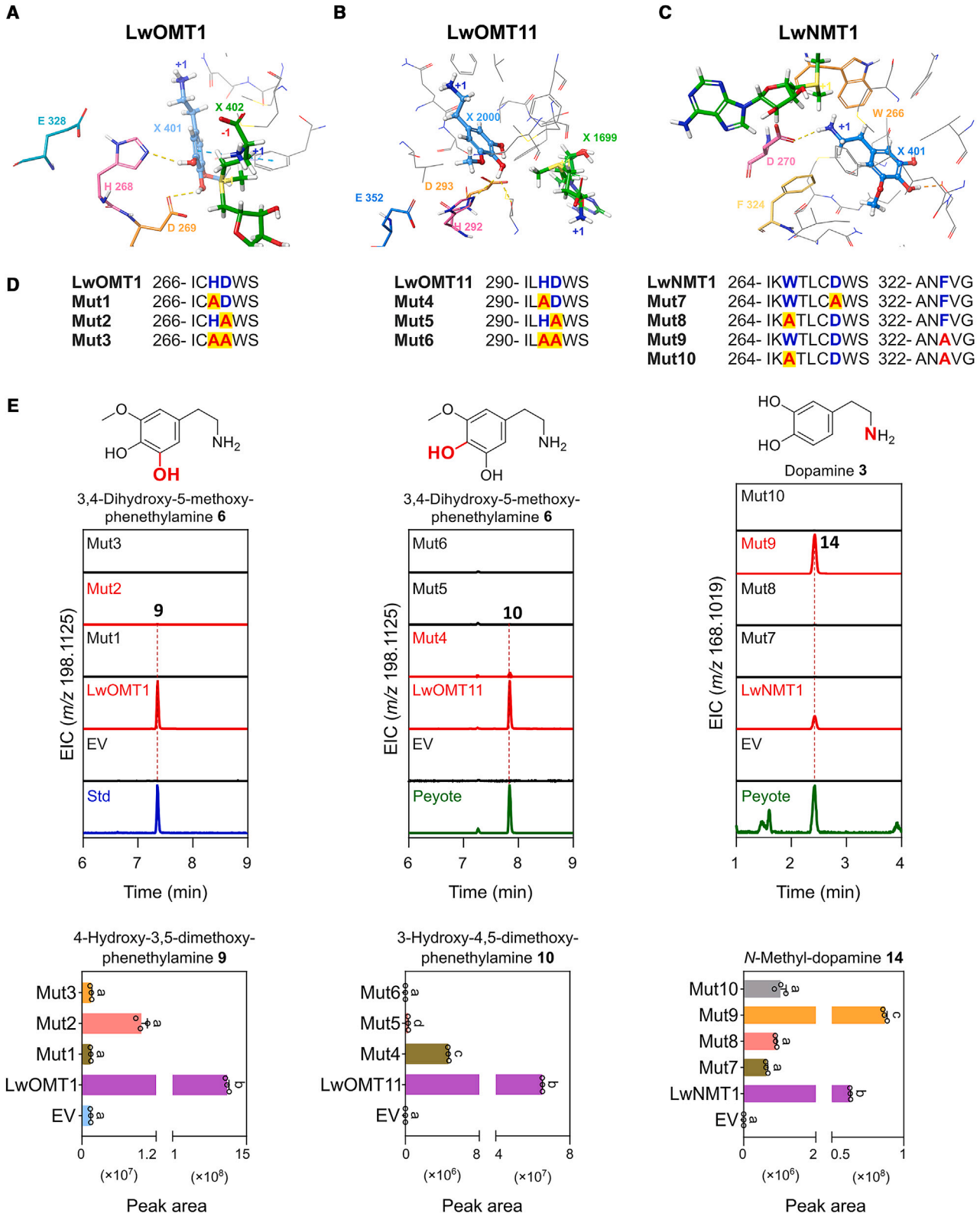

Figure 6. Molecular modeling of 3,4-dihydroxy-5-methoxy-phenethylamine in LwOMT1, LwOMT11, and LwNMT1 reveals the specificity of methylation sites.

(A–C) Reagents in the active site of **(A)** LwOMT1, **(B)** LwOMT11, and **(C)** LwNMT1 according to our predicted models. S-adenosyl-L-methionine (SAM) appears in green (thick tubes), and 3,4-dihydroxy-5-methoxy-phenethylamine **6** is shown in cyan. The protein non-polar hydrogens are not shown.

(legend continued on next page)

although LwOMT1, LwOMT11, and LwNMT1 enzymes were all capable of methylating 3,4-diOH-5-MeO-phenethylamine **6**, each enzyme targeted the methylation of a distinct group (Figure 1A). In addition, LwNMT1 did not cluster with other functionally characterized NMTs from the literature and showed closer sequence similarity to LwOMT1 and LwOMT5 (Supplemental Figure S13A). To understand the structural basis of this regioselectivity, we performed molecular modeling studies for each of the MT enzymes.

Structural models of LwNMT1, LwOMT1, and LwOMT11 were obtained using AlphaFold2 (Jumper et al., 2021) with root-mean-square deviations of backbone atomic positions were 3.20, 1.65, and 2.14 Å, respectively. For LwNMT1, the proximity of at least one Asp/Glu was necessary as a base for proton withdrawal and to facilitate nucleophilic attack of the substrate on the SAM methyl group. For OMTs, the presence of His (to be exact, His with adjacent Glu separated by an anion–hydrogen bond) was found to be essential for basic catalysis. The SAM-binding pockets of the predicted structures coincided spatially with those of the template MTs and accommodated SAM without major steric clashes, as did the substrates. This was also supported by retaining interactions observed from experimentally determined structures of SAM-dependent MTs, such as those with PDB ID 2AN4, 6LFE, and others. These also include an Asp/Glu side-chain carboxylate interaction with the ribose residue of SAM and a carboxylate group interaction with the 6-aminoadenine group.

The crystal structures of flavonoid OMTs (PDB: 6I6K, 1FP1, 1FP2, 5ICE, 6I72, and 3P9I) reveal the presence of both a His/Glu dyad and Asp near the substrate's phenolic oxygen intended for methylation. This structural arrangement was also predicted by the AlphaFold2 model for the placement of 3,4-diOH-5-MeO-phenethylamine **6** in LwOMT1 (H268/E328 and D269) and LwOMT11 (H292/E352 and D293), as illustrated in Figures 6A and 6B. In the case of LwNMT1, the initial placement of 3,4-diOH-5-MeO-phenethylamine **6** according to the alignment was rejected because the side-chain carboxylate anion essential for catalytic activity was not in proximity to the S-methyl and ammonium groups. Considering the sequence similarity between LwNMT1 and LwOMT1, we hypothesized that LwNMT1's substrate pocket was homologous to that of LwOMT1. Therefore, 3,4-diOH-5-MeO-phenethylamine **6** was moved to that site and oriented accordingly (Figure 6C).

Site-directed mutagenesis targeting the predicted active-site residues was performed to validate these predictions (His and Asp, as illustrated in Figure 6D). In addition, for LwNMT1, we performed mutagenesis on residues W266 and F324, which bear bulky aromatic side chains (Figure 6D), as we predicted that mutating these residues to alanine would induce significant changes in the substrate binding-site topology affecting substrate binding and stabilization before modification. Following mutagenesis and expression in *E. coli*, a comparative analysis of enzymatic activities was conducted, with testing against various substrates. Given the relatively low *in vitro* activity of LwNMT1 observed with 3,4-diOH-5-MeO-phenethylamine **6** (Supplemental Figure S13D), we tested the activity of purified LwNMT1 using dopamine **3**.

For all the enzymes, single- or double-point mutations in the active site led to complete or substantial loss of activity compared with the wild type (LwOMT1 vs. Mut1–Mut3, LwOMT11 vs. Mut4–Mut6, and LwNMT1 vs. Mut7; Figures 6D, 6E, and Supplemental Figures S15–S17). In the case of LwNMT1, the D270A mutation confirmed the hypothesized substrate pocket and catalytic site (Mut7; Figure 6E and Supplemental Figure S17). Interestingly, loss of F324 in the active site of LwNMT1 (i.e., Mut9) did not lead to a substantial loss of activity, unlike the W266A mutation (Mut8 and Mut10; Figure 6E and Supplemental Figure S17). This observation underscores the critical role of W266 in stabilizing or delocalizing the positive charge of the ammonium group more effectively, owing to a more electron-rich indolic aromatic system, thus likely reducing the reaction's activation barrier. A more detailed description of the mutation results is provided in Supplemental Data S1.

In vivo reconstruction of the mescaline biosynthetic pathway

We next attempted to reconstruct the mescaline **1** biosynthetic pathway by co-expressing LwTyDC2, LwCYP76AD131, LwOMT1, and LwOMT11 enzymes in yeast. After 2 days of galactose induction, we observed peaks of tyramine **4** and dopamine **3**, in addition to their respective *N*-acetylated forms, in both the pellet and the supernatant of the transformed yeast (Supplemental Figure S18A). However, no other intermediates from the pathway were detected. After the detection of *N*-acetylated metabolites, we incubated EV-transformed yeast with individual intermediates from the pathway to account for potential *N*-acetylation by a yeast *N*-acetyltransferase. After 2 days, we could detect all *N*-acetylated intermediates from the pathway (metabolites **32**–**36**; Supplemental Figure S18B; Supplemental Table S9). Consistent with these findings, an arylalkyl amine *N*-acetyltransferase from *S. cerevisiae* (scAANAT, GenBank: Q12447) has been reported to be functional on several phenethylamines, including mescaline **1**, dopamine **3**, tyramine **4**, and 3-MeO-tyramine **7** (Ganguly et al., 2001), suggesting that it may be responsible for the observed activity.

To confirm the *in vivo* activity of the enzymes, we supplemented yeast co-expressing the four enzymes from Peyote with each

(D) Sequence alignment of the methyltransferases and their mutants. Mutated residues are marked in red. Residues highlighted in yellow were found to be essential for enzyme activity according to the activity assays.**(E)** Extracted ion chromatograms and ion abundance peak areas (mean ± SD; $n = 3$ biological replicates) of 4-hydroxy-3,5-dimethoxy-phenethylamine **9**, 3-hydroxy-4,5-dimethoxy-phenethylamine **10**, and *N*-methyl-dopamine **14** following *in vitro* enzyme assays with the purified enzymes in the presence of SAM and either 3,4-dihydroxy-5-methoxy-phenethylamine **6** or dopamine **3**. Statistical significance was assessed using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons as a *post hoc* test; letters denote statistically significant differences ($p < 0.05$). Chromatograms were normalized to the highest value. Products in each assay were identified by comparison with analytical standards and/or an extract from a Peyote skin sample. The chemical structures of substrates are shown for reference. The site of methylation by the specific enzyme is marked in red.

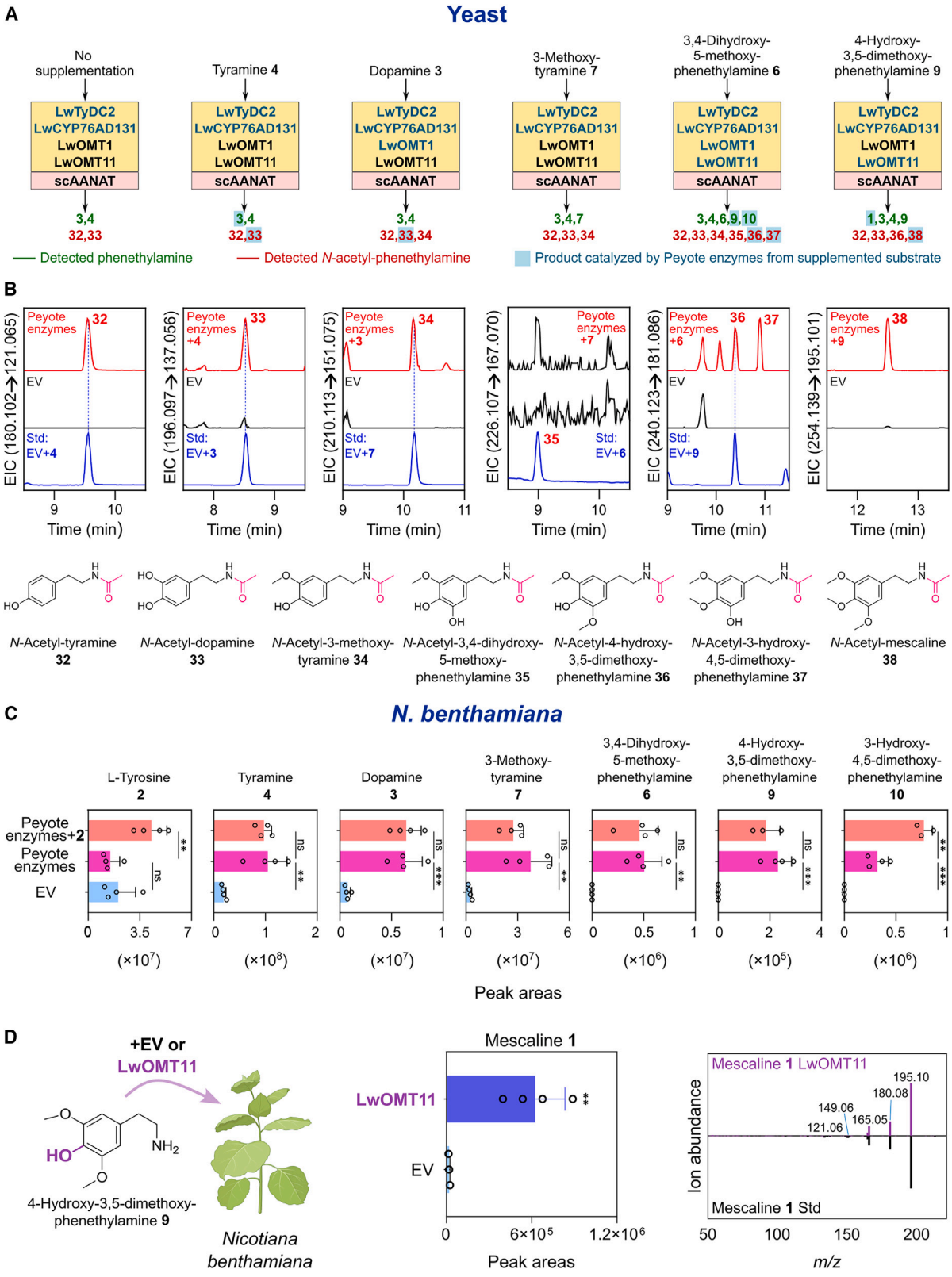

Figure 7. *In vivo* reconstruction of the mescaline biosynthetic pathway in yeast and *N. benthamiana*.

(A) Schematic representation of the experimental approach and observed products in yeast following co-expression of LwTyDC2, LwCYP76AD131, LwOMT1, and LwOMT11 and incubation with each intermediate from the mescaline **1** pathway individually (n = 2 biological replicates). *N*-Acetylation occurs in yeast by an endogenous *N*-acetylating enzyme (scAANAT) as presented for samples transformed with the

(legend continued on next page)

pathway intermediate individually (Figure 7A). Depending on the supplemented substrate, we detected the phenethylamine products tyramine **4**, dopamine **3**, 4-OH-3,5-diMeO-phenethylamine **9**, 3-OH-4,5-diMeO-phenethylamine **10**, and mescaline **1** (Figure 7A and Supplemental Figure S18C). In addition, we observed the appearance of *N*-acetylated products, including *N*-acetyl tyramine **32**, *N*-acetyl-dopamine **33**, *N*-acetyl-3-MeO-tyramine **34**, *N*-acetyl-4-OH-3,5-diMeO-phenethylamine **36**, and two newly observed metabolites putatively identified as *N*-acetyl-3-OH-4,5-diMeO-phenethylamine **37** and *N*-acetyl-mescaline **38** (Figures 7A and 7B, Supplemental Figures S18D and S18E; Supplemental Table S9). After incubation with 3-MeO-tyramine **7**, only a low-abundance peak of *N*-acetyl-3,4-diOH-5-MeO-phenethylamine **35** was detected, which could not be verified via MS/MS (Figure 7B). Overall, the production of phenethylamine intermediates and their *N*-acetylated forms by yeast co-expressing the Peyote enzymes confirms the enzymes' activity. However, the presence of *N*-acetylation competition, possibly via scAANAT, and the potentially low catalytic activity of specific enzymes prevented us from achieving *in vivo* synthesis of mescaline **1**. This highlights the complexity in future engineering efforts targeting the mescaline **1** pathway in yeast, suggesting that knocking down scAANAT may be necessary to enable its production.

We subsequently attempted to reconstitute the mescaline **1** biosynthetic pathway in *N. benthamiana* plant leaves. Co-expression of the four core enzymes led to a significant increase in tyramine **4**, dopamine **3**, and 3-MeO-tyramine **7** and *de novo* production of 3,4-diOH-5-MeO-phenethylamine **6,** 4-OH-3,5-diMeO-phenethylamine **9**, and 3-OH-4,5-diMeO-phenethylamine **10** compared with the EV control, albeit without production of mescaline **1** (Figure 7C). Supplementation with L-tyrosine **2** did not lead to increased concentrations of the intermediates, apart from 3-OH-4,5-diMeO-phenethylamine **10**, suggesting a more efficient catalytic activity of LwOMT11 with 3,4-diOH-5-MeO-phenethylamine **6** compared with LwOMT1 (Figure 7C). Additional 4-OH-3,5-diMeO-phenethylamine **9** supplementation to plants expressing only LwOMT11 resulted in the production of mescaline **1**, demonstrating the *in planta* activity of this enzyme (Figure 7D). Interestingly, *N. benthamiana* did not show substantial 4-*O*-methylation activity on 3,4-diOH-phenethylamine **6**, suggesting that the plant does not contain an LwOMT11 ortholog exhibiting this activity.

We next co-expressed the four enzymes stepwise in *N. benthamiana* and observed significant changes in the concentrations of some intermediates. For example, expression of LwTyDC2 resulted in a significant increase in tyramine **4**, and co-expression of LwTyDC2 and LwCYP76AD131 led to a significant increase in dopamine **3** and subsequent intermediates from the pathway owing to the promiscuous activity of NbOMT and the second hydroxylation activity of LwCYP76AD131 (Supplemental Figure S19A and Figure 4B). Co-expression of LwOMT1 and LwOMT11 alongside LwTyDC2 and LwCYP76AD131 did not yield significant increases in 4-OH-3,5-diMeO-phenethylamine **9** or 3-OH-4,5-diMeO-phenethylamine **10** (Supplemental Figure S19A). Upon revisiting our sequencing data using OrthoFinder, we identified an ortholog of LwOMT1 and LwOMT5 (termed LwOMT13, with 94% and 93.5% protein similarity, respectively) that were highly expressed in all Peyote tissues (Supplemental Table S7). Repeating the stepwise co-expression of enzymes in *N. benthamiana* using LwOMT13 instead of LwOMT1 resulted in similar trends (Supplemental Figure S19B).

Given the reactivity of the amine group in phenethylamines (observed *N*-methylation in Peyote and *N*-acetylation in yeast), we hypothesized that the absence of phenethylamine intermediate accumulation in *N. benthamiana* might be attributed to the plant's redirection of these metabolites to other pathways. We therefore screened the LC–MS data from the infiltrated *N. benthamiana* samples and identified several HCAAs (metabolites **39**–**48**, identified according to Li et al. 2018) whose levels changed as a result of the expression of the Peyote enzymes (for example, HCAA derivatives of tyramine **4** and dopamine **3** increased significantly as a result of LwTyDC2 expression; Supplemental Figure S20; Supplemental Table S10). Interestingly, we could not detect any of the identified HCAAs in Peyote (Supplemental Figure S20A).

Orthologous analyses provide clues about the mescaline biosynthetic pathway in two other cacti

Having identified the active enzymes responsible for mescaline **1** biosynthesis in Peyote, we next looked for orthologs in San Pedro and *L. diffusa* in order to shed light on their respective pathways. We sequenced and assembled the transcriptomes of the two cacti using Illumina paired-end sequencing (BUSCO completeness values of 96.9% and 97.4%, respectively; Supplemental Figure S21A). After protein annotation, whole-proteome phylogeny was performed to quantify the evolutionary distances among the three cacti (Supplemental Figure S21B). Using OrthoFinder, we identified orthologs of LwTyDC2, LwOMT1, LwOMT5, LwOMT10, LwOMT11, LwCYP76AD131, and LwCYP84A220 (Supplemental Figure S21C–S21E; Supplemental Table S11) in

EV and supplemented with each intermediate (Supplemental Figure S18B). Detected phenethylamines and *N*-acetyl-phenethylamines in each sample appear in green and red, respectively. The Peyote enzymes that exhibit activity with the supplemented intermediate and the respective products are highlighted in blue.

(B) Extracted ion chromatograms and chemical structures of *N*-acetylated products from **(A)**. Identification of *N*-acetylated phenethylamines was according to yeast transformed with the EV and following incubation with each intermediate as standard (blue chromatogram; full details appear in Supplemental Figure S18D, and S18E). Metabolites were analyzed using the acquired data-dependent MS/MS (dd-MS^2) data. Chromatograms were normalized to the highest value.

(C and **D)** Transient co-expression in *N. benthamiana* of **(C)** LwTyDC2, LwCYP76AD131, LwOMT1, and LwOMT11 with and without L-tyrosine **2** supplementation and **(D)** only LwOMT11 with 4-hydroxy-3,5-dimethoxy-phenethylamine **9** supplementation. Peak areas of ion abundance were used for the comparisons (mean ± SD; $n = 3$–4 biological replicates). Metabolites **6**, **9**, and **10** were analyzed on the same instrument using the acquired dd-MS^2 data. Statistical significance was assessed using a two-tailed Student's *t*-test with an assumption of unequal variance; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. All metabolites were identified by exact mass, retention time, and MS/MS spectra.

both *L. diffusa* and San Pedro. Interestingly, we identified an orthologous enzyme of LwNMT1 (LdNMT1) in *L. diffusa* that is missing in San Pedro (Supplemental Figure S21D).

DISCUSSION

Hallucinogenic cacti such as Peyote have played a significant role in the cultural and religious practices of Indigenous communities for thousands of years. However, the increasing recreational use, unsustainable harvesting, and slow growth rates of these cacti pose a potential risk of depleting natural sources of mescaline **1**. Exploration of sustainable production methods is imperative to safeguard its future development as a potential next-generation neuropsychiatric drug, ensuring responsible utilization while preserving the ecological balance for generations to come. Elucidation of the mescaline **1** biosynthetic pathway is a crucial initial step in this endeavor.

Biosynthesis of dopamine **3** from L-tyrosine **2** in plants proceeds via two routes: either decarboxylation followed by hydroxylation or vice versa. Whereas TyDCs frequently decarboxylate both L-tyrosine **2** and L-DOPA **5**, hydroxylases are much more specific to either L-tyrosine **2** or tyramine **4**. In Caryophyllales, several CYP76AD enzymes have been reported to hydroxylate L-tyrosine **2** to L-DOPA **5** (Polturak et al., 2016; Sunnadeniya et al., 2016). On the other hand, biosynthesis of catecholamines is thought to occur via tyramine **4**; however, no hydroxylases with this activity have been characterized (Schenck and Maeda, 2018). Metabolic profiling, transcriptomics, and functional characterization of Peyote enzymes provide evidence that dopamine **3** is biosynthesized from tyramine **4** and 3,4-diOH-5-MeO-phenethylamine **6** from 3-MeO-tyramine **7** by LwCYP 76AD131. While our manuscript was undergoing revision, another study on the mescaline **1** biosynthetic pathway in Peyote was published (Watkins et al., 2023), identifying a CYP76AD homolog (LwCYP76AD94) that produces L-DOPA **5** from L-tyrosine **2**. Surprisingly, we could not detect the *LwCYP76AD94* gene in our transcriptomics data or L-DOPA **5** in any of the tissues analyzed. Reverse transcription-quantitative PCR (RT-qPCR) of *LwCYP76AD131* and *LwCYP76AD94* confirmed expression of the two genes in skin and flesh of Peyote buttons, with higher expression of *LwCYP76AD131* than *LwCYP76AD94*, especially in the flesh tissue (Supplemental Figure S22).

Production of dopamine **3** from L-DOPA **5** in Peyote by LwCYP76AD94 was supported by a study by Lundström and Agurell (1969), which showed that feeding Peyote with L-DOPA **5** resulted in better incorporation into the mescaline **1** pathway compared with feeding tyramine **4**. However, this could be explained by superior catalytic activity of the LwTyDC2 enzyme for the decarboxylation of L-DOPA **5** versus the hydroxylation activity of LwCYP76AD131 on tyramine **4**, both of which lead to dopamine **3**. In a subsequent study, the same author concluded, on the basis of inverse isotope dilution experiments, that the production of dopamine **3** from L-DOPA **5** and the production of 3,4-diOH-5-MeO-phenethylamine **6** from 3,4,5-triOH-phenethylamine **8** in Peyote are likely of minor importance (Lundström, 1971). The latter findings align with the LwCYP76AD131 activities observed in this study. Nevertheless, given the functional activities of the two enzymes, it is possible that dopamine **3** biosynthesis in Peyote proceeds via both routes (Supplemental Figure S23). In addition, it is possible that LwCYP76AD94 is involved in the production of L-DOPA-derived betalain pigments (Supplemental Figure S23).

The presence of mescaline **1** and other pathway intermediates has been reported in various cacti, especially from the Cacteae and Trichocereae subfamilies (Smith, 1977). These metabolites may therefore confer an evolutionary advantage to plants, possibly by serving as a defense mechanism against herbivores, as indicated by the extreme bitterness of Peyote buttons (Cassels and Sáez-Briones, 2018). Notably, several early intermediates of this pathway, including L-tyrosine **2**, tyramine **4**, dopamine **3**, and 3-MeO-tyramine **7**, are also present in various other plant species (Hagel and Facchini, 2013; Schenck and Maeda, 2018), including *N. benthamiana*, and share structural similarities with common metabolites and orthologous enzymes found in other pathways such as HCA, HCAA, and phenylpropanoids. However, mescaline **1** has not been identified in any other plants outside the Cactaceae family. In this regard, the meta-hydroxylation of 3-MeO-tyramine **7** and the para-*O*-methylation of 4-OH-3,5-diMeO-phenethylamine **9** were not observed in *N. benthamiana*, suggesting that it lacks LwCYP76AD131 and LwOMT11 homologs. This may explain the lack of mescaline **1** in *N. benthamiana* despite the observed accumulation of 3-MeO-tyramine **7**.

In vivo reconstruction of the mescaline **1** biosynthetic pathway in *N. benthamiana* leaves produced several intermediates up to one step prior to mescaline **1**. In yeast, we observed activities of all the enzymes after feeding of the specific substrates. However, the pathway starting from L-tyrosine **2** stopped at dopamine **3** (Supplemental Figure S18A). Halting of this pathway in host organisms probably stems from the toxicity and reactivity of the phenethylamine intermediates. This conclusion is strongly supported by the observed production of *N*-acetylated (**32**–**38**) or HCAA (**39**–**48**) metabolites in yeast or *N. benthamiana*, respectively, which diverts intermediates away from the mescaline **1** pathway. The absence of HCAA metabolites (**39**–**48**) in Peyote (Supplemental Figure S21B) suggests that this type of metabolic competition is absent in the cactus. Alternatively, we found that Peyote *N*-methylates all pathway phenethylamines through the promiscuous activity of LwNMT1, which possibly functions as a detoxification mechanism, as previously suggested for benzylisoquinolines in *P. somniferum* (Brochmann-Hanssen et al., 1975). Consequently, LwNMT1 emerges as a key regulator of this pathway, steering intermediates away from the route to mescaline **1** (Figure 1A).

Interruption of the pathway in heterologous hosts may also arise from the engagement of a single enzyme across multiple steps in the pathway that exhibits diverse catalytic efficiencies with different intermediates. Similar to our study, Watkins et al. (2023) also identified LwOMT1, LwOMT13, and LwOMT11 (OMT6, OMT2, and OMT10 according to their numbering system) as the active OMTs in the pathway. In their study, LwOMT13 exhibited lower catalytic efficiency with 3,4-diOH-5-MeO-phenethylamine **6** compared with LwOMT11. This suggests that these two enzymes may compete for the same substrate, leading to more efficient production of 3-OH-4,5-diMeO-phenethylamine **10** versus 4-OH-3,5-diMeO-phenethylamine **9**. This observation

may also explain the increase in 3-OH-4,5-diMeO-phenethylamine **10**, but not 4-OH-3,5-diMeO-phenethylamine **9,** after supplementation with L-tyrosine **2** (Figure 7C). Given the large number of paralogs observed for this enzyme, it is likely that Peyote expresses several similar enzymes that catalyze the same reaction and thus compensate for lower kinetic rates. Alternatively, the high abundance of mescaline **1** in the cactus may result from accumulation over long time periods, as mescaline **1** concentration in Peyote is known to depend on the maturity and size of the cactus (Kalam et al., 2013).

Particularly notable is the sequence similarity between LwNMT1 and the active LwOMT1. According to our molecular modeling, the two enzymes exhibit regioselective methylation patterns determined by the proximity of the SAM methyl donor to the targeted functional group. In light of our findings, the presence of an LwNMT1 ortholog in *L. diffusa* aligns with the occurrence of *N*-methylated metabolites in this cactus and their general absence in San Pedro (Supplemental Figure S6). Consequently, San Pedro may have evolved an alternative mechanism to address the toxicity evoked by phenethylamines. It is therefore plausible that although *L. diffusa* synthesizes mescaline **1**, along with other pathway intermediates, these metabolites fail to accumulate owing to a highly active NMT enzyme (Supplemental Figure S8). Further exploration of LwNMT1 and its counterpart from *L. diffusa*, as well as functional characterization of enzymes in the mescaline **1** and related pathways in the two other cacti, stand to offer valuable insights into the evolutionary dynamics across these species.

By connecting these findings, the present study advances our understanding of mescaline **1** biosynthesis and offers promising avenues for sustainable production and responsible utilization of this valuable natural resource. To gain a deeper understanding of the evolutionary history of phenethylamines in various cacti, particularly the accumulation of mescaline **1** in *Lophophora* and *Echinopsis* genera, further investigations involving genetic and metabolic profiling of intermediate members within the Cactoideae family are needed. These studies will contribute significantly to a comprehensive grasp of the biosynthetic pathways of mescaline **1** and related compounds. The preservation of hallucinogenic cacti and their biosynthetic potential is critical for the development of mescaline-based treatments and the protection of cultural and ecological heritage.

METHODS

Chemicals and plant material

All the analytical metabolites were ≥95% pure. Mescaline (hydrochloride) **1**, dopamine (hydrochloride) **3**, 3-MeO-tyramine (hydrochloride) **7**, hordenine **13**, and *N*-methyl-mescaline **24** (hydrochloride) were purchased from Cayman Chemical (Ann Arbor, MI, USA), L-DOPA **5** and L-tyrosine **2** were purchased from Sigma-Aldrich (Rehovot, Israel), 3,4-diOH-5-MeO-phenethylamine (hydrochloride) **6** was purchased from Rare Chemicals (Kiel, Germany), 4-OH-3,5-diMeO-phenethylamine (hydrochloride) **9** was purchased from Ambinter (Orléans, France), and tyramine **4** was purchased from Holland Moran (Yehud, Israel). *N-p-trans*-Coumaroyltyramine **39**, *N-trans*-caffeoyltyramine **40**, *N-trans*-feruloyltyramine **41**, and *N-trans*-feruloyl-3-methoxytyramine **42** were purchased from BioBioPha (Yunnan, China). *N*-Acetyl-tyramine **32** and *N*-acetyl-dopamine **33** were purchased from MedChem Express (USA). All cacti were acquired from a local nursery in Israel. Validation of the assignment of Peyote and *L. diffusa* to the appropriate species was performed by a cactus taxonomist and growing expert and complemented by chemical profiling.

LC–MS chemical analysis

Unless otherwise stated, 100 mg of frozen powdered plant tissue was extracted with 300 μl of 80% methanol and 0.1% formic acid, sonicated for 15 min, agitated for 30 min, and centrifuged at 14 000 *g* for 10 min. The supernatant was filtered through a 0.22-μm syringe filter. Samples were analyzed using an ultra-performance liquid chromatograph (Waters Acquity) with a diode array detector connected to a quadrupole time-of-flight system (XEVO G2-S or Synapt HDMS qTOF systems, Waters), an Orbitrap IQ-X Tribrid MS, or an Orbitrap Exploris 240 MS (Thermo Scientific, Bremen, Germany). Chromatographic separation was performed on an ultra-performance liquid chromatography BEH-T3 C18 column (100 × 2.1 mm, 1.7 μm internal diameter, Waters Acquity). The mobile phase consisted of 0.1% formic acid in acetonitrile:water (1:99, v/v; phase A) and 0.1% formic acid in acetonitrile (phase B). The multistep program was as follows: initial conditions were 0% B for 4 min, raised to 75% B until 12 min, raised to 100% B in 0.2 min, held at 100% B until 15 min, decreased to 0% B in 0.2 min, and held at 0% B until 18 min for re-equilibration of the system. The flow rate was 0.3 ml min^{-1}, and the column temperature was kept at 35°C.

Electrospray ionization was used in positive ionization mode at an *m/z* range of 50–1000 Da. With the qTOF system, masses were detected with the following settings: capillary 1 kV, source temperature 140°C, desolvation temperature 450°C, and desolvation gas flow 800 L h^{-1}. MS/MS experiments were performed according to the specific protonated masses with the following settings: capillary spray of 1 kV, cone voltage of 30 eV, and collision energy ramp of 10–45 eV. Argon was used as the collision gas. On the Orbitrap MS systems, the source parameters were sheath gas flow rate, auxiliary gas flow rate, and sweep gas flow rate, 45, 10, and 1 arbitrary units, respectively; vaporizer temperature, 275°C; ion transfer tube temperature, 275°C; and spray voltage, 2.3 kV. The instrument was operated in full MS^1 mode with data-dependent MS/MS (MS-dd-MS^2). Data acquisition in full MS^1 mode was with 60 000 resolution, a normalized automatic gain control (AGC) target of 25%, and a maximum injection time of 50 ms. Data acquisition in dd-MS^2 mode was with 15 000 resolution, a normalized AGC target of 20%, a maximum IT of 150 ms, an isolation window of 1.5 m/z, and a normalized collision energy of 30. Identification of metabolites **1**–**7**, **9**, **13**, and **24** was performed according to analytical standards. All other metabolites in the cacti were putatively identified according to LC–MS/MS as detailed in Supplemental Figures S1–S3. Identification of some metabolites was validated in the *in vitro* and/or *in planta* assays (Supplemental Table S1). Some differences in fragmentation spectra, retention times, and ion absorbance were observed between instruments. Metabolites were

therefore compared in each run versus a standard and/or an extract analyzed under similar conditions.

Absolute quantification of mescaline and differential analysis of phenethylamines and THIQs

Fresh samples of whole buttons, stems, and thick and thin roots were collected from medium-sized Peyote and *L. diffusa* cacti (Figure 2B). The green sub-epidermal and flesh tissues of buttons were separated using a scalpel and analyzed separately. San Pedro flesh tissues were dissected with a scalpel into skin (green sub-epidermal area), middle, and center tissues. Spines correspond to the flesh tissue surrounding the spines (Figure 2B). For extraction, 100 mg of frozen powder was accurately weighed in triplicate, extracted with 1 ml 80% methanol and 0.1% formic acid, and prepared as described previously. Absolute quantification of mescaline **1** was performed by external calibration prepared in a similar diluent. Samples were injected in several dilutions to fit into the linear range of the calibration curve. Injections were performed on the IQ-X Tribrid MS system in full MS^1 mode as described previously. Samples were also used for differential analysis, in which peak areas were used for comparisons.

MALDI imaging

For localization of metabolites in Peyote tissues, two young fresh cacti were cryo-sectioned horizontally and laterally (n = 2 biological replicates each), and matrix was sprayed using a TM-sprayer (HTX Technologies, USA) with 2,5-dihydroxybenzoic acid (DHB; 40 mg ml^{-1} in water/methanol, v/v 30/70, containing 0.2% trifluoroacetic acid). The nozzle temperature was set to 70°C, and the DHB matrix solution was sprayed in 16 passes over the tissue sections at a linear velocity of 120 cm min^{-1} with a flow rate of 50 μl min^{-1}. Sections were imaged with a Nikon DS-Ri2 microscope. MALDI imaging was performed using a 7 T Solarix Fourier transform ion cyclotron resonance mass spectrometer (Bruker Daltonics). Data acquisition and analysis were performed using flexImaging V4.1 (Bruker). The datasets were collected in positive ionization mode using lock mass calibration (DHB matrix peak: $[3DHB+H-3H_2O]^+$, m/z 409.055408 Da) at a frequency of 1 kHz and a laser power of 40%, with 200 laser shots per pixel and 50-μm pixel size. Each mass spectrum was recorded in the range of m/z 120–2000 in broadband mode with a time domain for acquisition of 1M, providing an estimated resolving power of 115 000 at m/z 400. The spectra were normalized to root-mean-square intensity, and MALDI images were plotted at theoretical m/z ± 0.005% with pixel interpolation on.

Genome assembly, RNA sequencing, and annotation

High-molecular-weight DNA was extracted from frozen young green tissue and shipped to Edinburgh Genomics for sequencing. Sequencing was performed on the PacBio Sequel II platform with ~8-kb DNA SMRTbell library preparation. Two different SMRT 8M cells were used, yielding 35.5 Gb of HiFi data. Hifiasm software (Cheng et al., 2021) was used to assemble the genome to produce an 11 × 3.3-Gb genome assembly.

RNA was extracted from six tissues of Peyote, *L. diffusa*, and San Pedro (Supplemental Table S3). RNA integrity was checked using a TapeStation instrument. Paired-end Illumina libraries were prepared and sequenced on an Illumina HiSeq 3000 instrument (PE 2 × 150, ~40 million reads per sample) and processed following the guidelines in Freedman and Weeks (2022). In brief, random sequencing errors were corrected using Rcorrector (Song and Florea, 2015), and uncorrectable reads were removed. Adaptor and quality trimming were performed using TrimGalore! (Krueger, 2021). Ribosomal RNA was filtered by discarding reads that mapped to the SILVA_132_LSURef and SILVA_138_SSURef non-redundant databases using bowtie 2 (Langmead and Salzberg, 2012). Fastq quality checks on each of the steps were performed using MultiQC (Ewels et al., 2016). The remaining reads were pooled and used for genome-independent *de novo* transcriptome assembly using Trinity (Grabherr et al., 2011).

The Iso-Seq data were obtained from a pool of six tissues from *L. williamsii* (Supplemental Table S3) and processed with isoseq3 (Pacific Biosciences, 2019). Fused and unspliced transcripts were removed, and only poly(A)-positive transcripts were retained for a unique set of high-quality isoforms. Iso-Seq and Trinity transcripts were aligned to the assembly using minimap2 (Li, 2018), and the BAM files were used to produce RNA-based gene models using Iso-Seq collapse (Pacific Biosciences, 2019). Gene functional annotation was performed for the predicted mature transcripts using TransDecoder (Haas, 2017), which takes into account HMMER (Finn et al., 2011) hits against Pfam (Finn et al., 2014) and BLASTP (Altschul et al., 1990) hits against UniProt (UniProt Consortium, 2008) databases for similarity retention criteria. Further annotation of protein-coding transcripts was performed by taking the best hits of BLASTP searches against other plant protein databases (UniProt protein FASTA files of arabidopsis id UP000006548_3702, tomato id UP000004994_4081, rice id UP000059680_39947, sugar beet id Bvulgaris_548_EL10_1.0, and spinach id UP000054095_3562). Signal peptides were predicted with SignalP (Almagro Armenteros et al., 2019), transmembrane domains were predicted with TMHMM (Krogh et al., 2001), and Gene Ontology and Kyoto Encyclopedia of Genes and Genomes terms were obtained with Trinotate (Bryant et al., 2017). BUSCO (Simão et al., 2015) was used at multiple stages of the analysis to assess the completeness of different transcriptome and genome versions. Circos and gene cluster plots were generated from the gff file. The circos plot was made with the R circlize package (Gu et al., 2014), and the gene cluster plots were made with the gggenes package (Wilkox, 2023). The full script used for protein functional annotation can be found at https://github.com/Luisitox/Cacti_paper.

Phylogenetic analyses

The selection of proteins for each of the families analyzed in this study was based on functionally tested enzymes according to the studies referenced in each figure. The full list of IDs can be found in Supplemental Table S8. The maximum likelihood trees were constructed with 300 or 1000 bootstrap tests based on a MUSCLE multiple alignment using MEGA11 software. The evolutionary distances were computed using the JTT matrix-based method. The phylogenetic relationships of the Cactoideae subfamily were drawn according to a synthesis of existing

literature (Hernández-Hernández et al., 2011) and represent a preliminary interpretation of the relationships among the taxa in this study.

Orthology analyses

Proteomes of the three cacti were obtained from each genome/transcriptome annotation. In the case of Trinity assemblies, CD-HIT (Fu et al., 2012) was run to cluster closely related transcripts with a tolerance of 90%. For the PacBio-based assembly, the longest functional protein was used. Orthogroups and their phylogenetic relationships were inferred with OrthoFinder (Emms and Kelly, 2019). Proteins that belonged to the same orthogroups as LwTyDC2, LwCYP76AD131, LwCYP84A220, LwOMT1, LwOMT5, LwOMT10, LwOMT11, and LwNMT1 were identified and realigned, and their phylogenetic relationships were inferred with MEGA11 as described previously.

Recombinant protein expression in *E. coli*

TyDC2 , *TyDC4* , *TyDC5*, *OMT1-11*, and *NMT1* coding sequences from Peyote were individually cloned into the pET28b vector digested with EcoRI using the ClonExpress II One Step Cloning Kit (Vazyme, Germany). *Mut1–Mut10* coding sequences of mutants were synthesized by the company Twist Biosciences. All constructs along with the EV as the control were sequenced and expressed in *E. coli* BL21(DE3)pLysS competent cells (a complete list of the primers appears in Supplemental Table S12). Bacterial starters were grown overnight in Luria–Bertani (LB) medium at 37°C, diluted in fresh LB 1:100, and reincubated at 37°C. When cultures reached A600 = 0.6, protein expression was induced with 500 μM of isopropyl-1-thio-β-d-galactopyranoside overnight at 16°C. Bacterial cells were lysed by sonication in 50 mM Tris–HCl (pH 8.0), 0.5 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich) solution in isopropanol, 10% glycerol, protease inhibitor cocktail (Sigma-Aldrich), and 0.1 mg ml^{-1} lysozyme (Sigma-Aldrich). The whole-cell extract was either kept for functional activity or used for protein purification. Purification of hexahistidine-tagged proteins was performed using a Capturem His-tagged purification 24-well plate (Takara Bio, USA) according to the manufacturer's protocol. The proteins were eluted with 2 ml of 500 mM imidazole (Fluka) in buffer containing 20 mM NaH_2PO_4 (pH 7.6) and 500 mM NaCl. Desalting was performed by repeated centrifugation at 4°C and 4300 rpm using an Amicon Ultra 10 K column (Merck Millipore, Tullagreen, Carrigtwohill, Ireland) in either 50 mM HEPES (pH 7.2) or the reaction buffer. Protein concentrations of the eluted fractions were measured with the Pierce 660-nm protein assay reagent (Thermo Scientific). Protein purity was assessed by gel electrophoresis using a 10% (w/v) polyacrylamide gel. Protein purification and enzymatic assays were performed consecutively on the same day.

TyDC enzyme assays

Individual TyDC assays were carried out as described by Lehmann and Pollmann (2009) with some modifications. Lysate enzyme assays (n = 3 biological replicates) were performed in 50 μl with 50 mM sodium phosphate buffer (pH 7.5), 1 mM PLP, 2 mM L-tyrosine **2**, or L-DOPA **5**, and 20 μl of the lysate enzyme solution. The assay mixtures were incubated at 30°C for 2 h, after which the reactions were terminated by addition of 70 μl ice-cold methanol with 0.1% formic acid, vortexing, and centrifugation at 15 000 *g* for 10 min. The supernatant was filtered and analyzed by LC–MS. Products in each assay were identified by comparison with analytical standards. Activity of LwTyDC2 was evaluated in 50-μl reactions (n = 3 biological replicates) with 2 μl of the partially purified enzyme (measured concentration corresponded to 1.5 μg of enzyme), 50 mM Tris–HCl buffer (pH 8), 1 mM PLP, and various concentrations (2.5 μM to 3 mM) of L-tyrosine **2** or L-DOPA **5**. Samples were incubated at 30°C for 15 min, after which the reactions were terminated by addition of 70 μl of ice-cold methanol with 0.1% formic acid or 0.1 M HCl for extraction of tyramine **4** or dopamine **3**, respectively. Samples were further extracted and analyzed by LC–MS as described previously. Absolute quantification of tyramine **4** and dopamine **3** was performed using external calibration curves prepared in the respective diluents. After numerous attempts to purify the protein, which resulted in partial purification, we proceeded to compute the relative activity and determine the Michaelis–Menten *Km* value to assess the enzyme's activity toward both substrates for comparison.

Methyltransferase enzyme assays

Individual MT enzyme assays (n = 3 biological replicates) were performed in 50 μl with 50 mM potassium phosphate buffer (pH 7.5), 1 mM SAM, 25 mM ascorbic acid, 20 μl of the lysate enzyme solution or 2 μg of pure enzyme, and 1 mM of dopamine **3**, 3,4-diOH-5-MeO-phenethylamine **6,** or 4-OH-3,5-diMeO-phenethylamine **9**. The assay mixtures were incubated at 30°C for 2 h, after which the reactions were terminated by addition of 100 μl of ice-cold methanol with 0.1% formic acid and further extracted and analyzed by LC–MS as described previously. Identification of products for which no analytical standards were available was performed by comparison to a Peyote skin extract. Activity testing of LwOMT1 versus Mut1–Mut3 mutants and LwOMT11 versus Mut4–Mut6 mutants was performed under similar conditions with 15 μg of pure enzymes and 3,4-diOH-5-MeO-phenethylamine **6** as the substrate. Activity testing of LwNMT1 versus Mut7–Mut10 mutants was performed under similar conditions with 50 mM Tris–HCl buffer (pH 8), 20 μg of pure enzymes, and dopamine **3** as the substrate.

Transient expression of selected genes in *N. benthamiana*

LwTyDC2, *CYP71AT375*, *LwCYP84A220*, *CYP84A221*, *CYP84A221v2*, *CYP81CJ19*, *CYP81CL5*, *CYP89A380*, *CYP73A424*, *CYP81B310*, *PPO1*, *MT1–MT4*, *LwOMT1*, *LwOMT11*, and *LwOMT13* coding sequences were individually cloned into the pAlpha2-NPT II-Ubq10-CCD-Ter10 vector digested with BsaI using the ClonExpress II One Step Cloning kit (Vazyme). The full list of oligonucleotides used for cloning appears in Supplemental Table S12. For BvCYP76AD6, we used stocks of previously transformed bacteria containing the specific gene (Polturak et al., 2016). *CYP76F270* and *LwCYP76AD131* coding sequences were synthesized by Twist Biosciences and similarly cloned into the same vector. All plasmids were sequenced and transformed into *Agrobacterium tumefaciens* (*A. tumefaciens*) strain GV3101 by electroporation. *A. tumefaciens* harboring the overexpression

constructs were grown overnight at 28°C in LB medium in the presence of kanamycin and gentamycin. After overnight incubation, 100 μl of each culture was grown for an additional 24 h in 10 ml of LB solution containing kanamycin, gentamycin, 10 mM MES, and 20 μM acetosyringone. Bacterial cells were collected by centrifugation, washed, and resuspended in infiltration buffer (10 mM MES, 10 mM $MgCl_2$, and 10 μM acetosyringone) to optical density 600 = 0.8. Individual and/or equal volumes of *A. tumefaciens* suspensions with different expression vectors were combined to obtain the desired gene combinations. The solutions were infiltrated into 4- or 5-week-old *N. benthamiana* leaves from the abaxial side using a 1-ml needleless syringe. Unless otherwise stated, substrates (50 μg ml^{-1} each) were infiltrated into the same leaf areas 2 days after initial infiltration, and all leaves were collected for metabolite analysis 60–72 h after substrate infiltration (samples without substrate supplementation were collected at similar time points). Leaf samples were flash frozen and extracted as described previously with 300 μl of 80% methanol with 0.1% formic acid and analyzed via LC–MS. Metabolites were identified using analytical standards and/or compared with similar peaks in a Peyote skin extract. Where specified, metabolites **6**, **9**, and **10** were quantified on the same instrument using the acquired dd-MS^2 data for increased sensitivity: 3,4-diOH-5-MeO-phenethylamine **6**, 184.097 → 167.070; 4-OH-3,5-diMeO-phenethylamine **9** and 3-OH-4,5-diMeO-phenethylamine **10**, 198.113 → 181.086. HCAAs were analyzed using conditions, column, and mobile phase similar to those described previously with the following multistep method: initial conditions were 5% B raised to 33% B until 22 min, raised to 100% B in 0.5 min, held at 100% B until 26 min, decreased to 5% B in 0.5 min, and held at 5% B until 30 min for re-equilibration of the system.

Heterologous expression in yeast

LwCYP76AD131, *LwTyDC2*, *LwCYP84A22O*, *LwOMT1*, and *LwOMT11* coding sequences were individually cloned into pESC plasmids, enabling simultaneous expression of up to two genes in one plasmid. *LwCYP76AD131* was cloned into pESC-Trp, *LwCYP84A22O* and *LwOMT11* were cloned into pESC-Leu, and *LwTyDC2* and *LwOMT1* were cloned into pESC-His. The cloning was performed by in-fusion reactions, and the pESC vectors were linearized by PCR. The full list of primers appears in Supplemental Table S12. The pESC constructs were sequenced and transformed into *S. cerevisiae* WAT21 using the Yeastmaker yeast transformation system (Clontech) as follows: (1) LwTyDC2-pESC-His and LwCYP76AD131-pESC-Trp; (2) LwCYP76AD131-pESC-Trp; (3) LwCYP84A220-pESC-Leu; and (4) LwTyDC2-LwOMT1-pESC-His, LwCYP76AD131-pESC-Trp, and LwOMT1-pESC-Leu. Each strain was additionally transformed with the appropriate EV as a control. Transformed yeast were grown on synthetic defined (SD) minimal medium supplemented with the appropriate amino acids and 2% glucose. Colonies were screened, and the presence of the transgene was confirmed by colony PCR. For induction of gene expression, transformed cells were grown in 5 ml of SD minimal medium with 2% glucose; after 24 h, they were transferred to SD minimal medium with 2% galactose and without additional supplementation or with 1 mM of tyramine **4**, dopamine **3**, 3-MeO-tyramine **7**, 3,4-diOH-5-MeO-phenethylamine **6**, or 4-OH-3,5-diMeO-phenethylamine **9** and grown for an additional 48 h at 30°C.

Cultures were centrifuged at 8000 *g* for 1 min, and the pellet and supernatant were collected separately. The supernatant was lyophilized, 1 ml of methanol with 0.1% formic acid was added, the mixture was sonicated for 30 min and centrifuged at 4300 rpm for 10 min, and the clear supernatant was collected and dried under a stream of N_2. Dry residues were dissolved in 200 μl methanol, sonicated for 20 min, centrifuged at 15 000 *g* for 10 min, and filtered through a 0.22-μm filter for LC–MS analysis. The cell pellet was weighed, double the amount of glass beads (diameter 500 μm) and 500 μl methanol with 0.1% formic acid were added, and the mixture was vortexed at maximum speed for 10 min. Lysed cells were centrifuged at 21 000 *g* for 5 min, and the clear supernatant was collected and dried under a stream of N_2. Dry residues were dissolved in 100 μl methanol, sonicated for 20 min, centrifuged at 15 000 *g* for 10 min, and filtered through a 0.22-μm filter for LC–MS analysis. Samples were analyzed using the acquired dd-MS^2 data according to the following transitions: L-tyrosine **2**, 182.081 → 165.055; dopamine **3**, 154.086 → 137.056; tyramine **4**, 138.091 → 121.065; 3,4-diOH-5-MeO-phenethylamine **6**, 184.097 → 167.070; 3-MeO-tyramine **7**, 168.102 → 151.075; 4-OH-3,5-diMeO-phenethylamine **9** and 3-OH-4,5-diMeO-phenethylamine **10**, 198.113 → 181.086; mescaline **1**, 212.128 → 195.101; *N*-acetyl-tyramine **32**, 180.102 → 121.065; *N*-acetyl-dopamine **33**, 196.097 → 137.056; *N*-acetyl-3-MeO-tyramine **34**, 210.113 → 151.075; *N*-acetyl-3,4-diOH-5-MeO-phenethylamine **35**, 226.107 → 167.070; *N*-acetyl-4-OH-3,5-diMeO-phenethylamine **36** and *N*-acetyl-3-OH-4,5-diMeO-phenethylamine **37**, 240.123 → 181.086; and *N*-acetyl-mescaline **1**, 254.139 → 195.101.

Molecular modeling and docking of LwOMT1, LwOMT11, and LwNMT1

All molecular modeling procedures were performed using Schrodinger Suite 2022-2 (Jacobson et al., 2004; Johnston et al., 2023). The protein structures were prepared using Protein Preparation Wizard (Sastry et al., 2013). Hydrogens were added, bond orders were assigned, and missing side chains were built using Prime (Jacobson et al., 2002; Jacobson et al., 2004; Johnston et al., 2023). Histidine side chains were left neutral, aspartates and glutamate side chains were deprotonated, and lysine and arginine were in cationic form. Finally, the protein structures and the final complexes underwent restrained minimization (root-mean-square deviation = 0.5 Å) using OPLS4 forcefield (Lu et al., 2021). Structural models for all three enzymes (LwNMT1, LwOMT1, and LwOMT11) were built *de novo* using AlphaFold2 software (Jumper et al., 2021). Each prediction run resulted in five structures, and the model with the highest confidence score (pLDDT) was chosen for modeling as follows: 2AN4-1, human phenylethylamine NMT, 6I6K (*P. somniferum* OMT), and 1FP2 (isoflavone OMT from *Medicago sativa*). To place the non-methylated substrate and SAM structures into the obtained enzyme 3D models, backbone alignment of known enzyme PDB structures with corresponding regioselectivity to *N*-methylation, 3-*O*-hydroxyphenyl, and 4-*O*-hydroxyphenyl was performed. When necessary, manual ligand structure editing was performed in the workspace. In the resulting complexes,

steric clashes of the side chains with SAM or substrates and products were resolved by side-chain local minimizations.

Gene expression analysis by RT–qPCR

Total RNA was isolated from skin and flesh tissues (n = 3) using the Spectrum Plant total RNA kit (Sigma-Aldrich). DNase I (Sigma-Aldrich)-treated RNA was reverse transcribed using a high-capacity cDNA reverse transcription kit (Applied Biosystems). Gene-specific primer pairs (Supplemental Table S13) were designed using Benchling software (Benchling, USA). Reactions were performed with SYBR Green PCR Master Mix (Applied Biosystems) on a QuantStudio 1 RT–PCR system (Thermo Fisher Scientific). Actin was used as the reference to normalize cDNA content, and the $2^{-\Delta\Delta Ct}$ method was used to compare gene expression.

DATA AND CODE AVAILABILITY

All raw next-generation sequencing data and the primary genome assembly version Lwill_v1 have been deposited in the European nucleotide archive (ENA) under accession number PRJEB74428. The sequences of the active genes reported in this article have been deposited at NCBI GenBank (accessions PP839057–PP839065). The genomic positions, functional annotations, CDSs, and protein sequences of these genes are available in Supplemental Table S7. The protein databases used in this study were UniProt Swiss-Prot release-2022_02, arabidopsis UP000006548_3702, tomato UP000004994_4081, rice UP000059680_39947, sugar beet Bvulgaris_548_EL10_1.0, and spinach UP000054095_3562. The Pfam database was Pfam-A.hmm release 34.0. All the code used in this study, along with a description of the scripts, can be found at https://github.com/Luisitox/Cacti_paper.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

AUTHOR CONTRIBUTIONS

P.B., L.A.d.H., and A.A. designed the study. P.B. performed the chemical profiling with the aid of I.R. L.A.d.H. performed the Peyote genome and transcriptome assembly and annotation, and G.L. performed the Illumina-based transcriptome assembly and annotation. P.B., L.A.d.H., A.-R.C., S.P., and H.H. performed all the molecular biology work, *in vitro* enzyme assays, protein purification, and transient expression of genes in *N. benthamiana*. A.-R.C. cloned the genes for expression in yeast, Y.P. transformed yeast, and P.B. grew the cells and performed the activity assays with the help of J.C. Y.D. prepared samples and performed the MALDI imaging experiments. N.K. performed the molecular modeling, analyzed the data, and designed the site-directed mutants. P.B. cloned, expressed, purified, and tested the site-directed mutants with the help of H.H. P.B. and L.A.d.H. analyzed the data, and P.B. wrote the manuscript with the assistance and input of all coauthors. S.M., U.H., A.J., and J.C. contributed materials to the study. A.A. supervised the study. All the authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

A.A. is the incumbent of the Peter J. Cohn Professorial Chair. L.A.d.H. was partially supported by fellowships from the Israel Ministry of Absorption and the Dean of the Plant Science Department in the Weizmann Institute. We thank Prof. David Nelson (University of Tennessee, USA) for the systematic naming of the cytochrome P450 enzymes characterized in this study; Shmuel Regev and The Regev Nursery (Beit Elazari, Israel) for help in differentiating among *Lophophora* species; Dr. Ziv Spiegelman for providing several *N. benthamiana* plants; XINTEZA (https://xinteza.com/) for funding this research; and the Adelis Foundation, the Leona M. and Harry B. Helmsley Charitable Trust, the Jeanne and Joseph Nissim Foundation for Life Sciences, Tom and Sondra Rykoff Family Foundation Research, Ron Sklare and the Raymond Burton Plant Genome Research Fund for supporting the A.A. laboratory.

Received: August 18, 2023
Revised: April 2, 2024
Accepted: May 30, 2024
Published: June 3, 2024

REFERENCES

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotech. **37**:420–423.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

Brochmann-Hanssen, E., Chen, C.H., Chen, C.R., Chiang, H.C., Leung, A.Y., and McMurtrey, K. (1975). Opium alkaloids. Part XVI. The biosynthesis of 1-benzylisoquinolines in Papaver somniferum. Preferred and secondary pathways; stereochemical aspects. J. Chem. Soc. Perkin 1 **1**:1531–1537. https://doi.org/10.1039/P19750001531.

Bryant, D.M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., Lee, T.J., Leigh, N.D., Kuo, T.-H., Davis, F.G., et al. (2017). A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. Cell Rep. **18**:762–776.

Cassels, B.K., and Sáez-Briones, P. (2018). Dark classics in chemical neuroscience: mescaline. ACS Chem. Neurosci. **9**:2448–2458. https://doi.org/10.1021/acschemneuro.8b00215.

Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods **18**:170–175.

Emms, D.M., and Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biol. **20**:238.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics **32**:3047–3048.

Facchini, P.J., Huber-Allanach, K.L., and Tari, L.W. (2000). Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. Phytochemistry **54**:121–138. https://doi.org/10.1016/S0031-9422(00)00050-9.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., et al. (2014). Pfam: the protein families database. Nucleic Acids Res. **42**:D222–D230.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. **39**:W29–W37.

Freedman, A., and Weeks, N. (2022). https://informatics.fas.harvard.edu/best-practices-for-de-novo-transcriptome-assembly-with-trinity.html.

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

Ganguly, S., Mummaneni, P., Steinbach, P.J., Klein, D.C., and Coon, S.L. (2001). Characterization of the *Saccharomyces cerevisiae* homolog of the melatonin rhythm enzyme arylalkylamine *N*-acetyltransferase (EC 2.3.1.87). J. Biol. Chem. **276**:47239–47247. https://doi.org/10.1074/jbc.M107222200.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al.

(2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. **29**:644–652.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). *circlize* implements and enhances circular visualization in R. Bioinformatics **30**:2811–2812.

Haas, B. (2017). https://github.com/TransDecoder/TransDecoder/.

Hagel, J.M., and Facchini, P.J. (2013). Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. Plant Cell Physiol. **54**:647–672. https://doi.org/10.1093/pcp/pct020.

Heffter, A. (1896). Ueber Cacteenalkaloïde. Chem. Ges. **29**:216–227.

Heffter, A. (1898). Ueber Pellote. Beiträge zur chemischen und pharmakologischen Kenntniss der Cacteen. II. Mittheilung. Naunyn-Schmiedeberg's Arch. Pharmacol. **40**:385–429.

Hernández-Hernández, T., Hernández, H.M., De-Nova, J.A., Puente, R., Eguiarte, L.E., and Magallón, S. (2011). Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). Am. J. Bot. **98**:44–61. https://doi.org/10.3732/ajb.1000129.

Ibarra-Laclette, E., Zamudio-Hernández, F., Pérez-Torres, C.A., Albert, V.A., Ramírez-Chávez, E., Molina-Torres, J., Fernández-Cortes, A., Calderón-Vázquez, C., Olivares-Romero, J.L., Herrera-Estrella, A., et al. (2015). De novo sequencing and analysis of Lophophora williamsii transcriptome, and searching for putative genes involved in mescaline biosynthesis. BMC Genom. **16**:1–14. https://doi.org/10.1186/s12864-015-1821-9.

Ibrahim, R.K., Bruneau, A., and Bantignies, B. (1998). Plant O-methyltransferases: molecular analysis, common signature and classification. Plant Mol. Biol. **36**:1–10. https://doi.org/10.1023/A:1005939803300.

Jacobson, M.P., Friesner, R.A., Xiang, Z., and Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. J. Mol. Biol. **320**:597–608.

Jacobson, M.P., Pincus, D.L., Rapp, C.S., Day, T.J., Honig, B., Shaw, D.E., and Friesner, R.A. (2004). A hierarchical approach to all-atom protein loop prediction. Proteins: Struct. Funct. Bioinf. **55**:351–367.

Johnston, R.C., Yao, K., Kaplan, Z., Chelliah, M., Leswing, K., Seekins, S., Watts, S., Calkins, D., Chief Elk, J., Jerome, S.V., Repasky, M.P., and Shelley, J.C. (2003). Epik: pKa and protonation state prediction through machine learning. J. Chem. Theory Comput. **19**:2380–2388.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature **596**:583–589. https://doi.org/10.1038/s41586-021-03819-2.

Kalam, M.A., Trout, K., Klein, M.T., Daley, P., Hulsey, D., and Terry, M. (2013). A preliminary report of mescaline concentrations in small regrowth crowns vs. mature crowns of *Lophophora williamsii* (Cactaceae): Cultural, economic, and conservation implications. J. Bot. Res. Inst. Tex. **27**:435–440. https://www.jstor.org/sTable/24621093.

Kapadia, G.J., and Fayez, M.B. (1970). Peyote constituents: chemistry, biogenesis, and biological effects. J. Pharm. Sci. **59**:1699–1727. https://doi.org/10.1002/jps.2600591202.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305**:567–580.

Krueger, F. (2021). https://github.com/FelixKrueger/TrimGalore.

Kwan, A.C., Olson, D.E., Preller, K.H., and Roth, B.L. (2022). The neural basis of psychedelic action. Nat. Neurosci. **25**:1407–1419. https://doi.org/10.1038/s41593-022-01177-4.

Lam, K.C., Ibrahim, R.K., Behdad, B., and Dayanandan, S. (2007). Structure, function, and evolution of plant O-methyltransferases. Genome **50**:1001–1013. https://doi.org/10.1139/G07-077.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods **9**:357–359.

Lehmann, T., and Pollmann, S. (2009). Gene expression and characterization of a stress-induced tyrosine decarboxylase from Arabidopsis thaliana. FEBS Lett. **583**:1895–1900.

Ley, L., Holze, F., Arikci, D., Becker, A.M., Straumann, I., Klaiber, A., Coviello, F., Dierbach, S., Thomann, J., Duthaler, U., et al. (2023). Comparative acute effects of mescaline, lysergic acid diethylamide, and psilocybin in a randomized, double-blind, placebo-controlled cross-over study in healthy participants. Neuropsychopharmacology **48**:1659–1667. https://doi.org/10.1038/s41386-023-01607-2.

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics **34**:3094–3100.

Li, Z., Zhao, C., Zhao, X., Xia, Y., Sun, X., Xie, W., Ye, Y., Lu, X., and Xu, G. (2018). Deep annotation of hydroxycinnamic acid amides in plants based on ultra-high-performance liquid chromatography–high-resolution mass spectrometry and its in silico database. Anal. Chem. **90**:14321–14330. https://doi.org/10.1021/acs.analchem.8b03654.

Lu, C., Wu, C., Ghoreishi, D., Chen, W., Wang, L., Damm, W., Ross, G.A., Dahlgren, M.K., Russell, E., Von Bargen, C.D., et al. (2021). OPLS4: Improving force field accuracy on challenging regimes of chemical space. J. Chem. Theory Comput. **17**:4291–4300. https://doi.org/10.1021/acs.jctc.1c00302.

Lundström, J. (1971). Biosynthesis of mescaline and tetrahydroisoquinoline alkaloids in *Lophophora williamsii* (Lem.) Coult. Occurrence and biosynthesis of catecholamine and other intermediates. Acta Chem. Scand. **25**:3489–3499. https://doi.org/10.3891/acta.chem.scand.25-3489.

Lundström, J., and Agurell, S. (1969). A complete biosynthetic sequence from tyrosine to mescaline in two cactus species. Tetrahedron Lett. **10**:3371–3374. https://doi.org/10.1016/S0040-4039(00)99766-1.

McClure-Begley, T.D., and Roth, B.L. (2022). The promises and perils of psychedelic pharmacology for psychiatry. Nat. Rev. Drug Discov. **21**:463–473. https://doi.org/10.1038/s41573-022-00421-7.

Meyer, K., Cusumano, J.C., Somerville, C., and Chapple, C.C. (1996). Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases. Proc. Natl. Acad. Sci. USA **93**:6869–6874. https://doi.org/10.1073/pnas.93.14.6869.

Nichols, D.E. (2016). Psychedelics. Pharmacol. Rev. **68**:264–355. https://doi.org/10.1124/pr.115.011478.

Pacific Biosciences. (2019). https://github.com/PacificBiosciences/IsoSeq.

Plant cytochrome P450 database. (2020). https://p450.biodesign.ac.cn/Tables.

Polturak, G., Breitel, D., Grossman, N., Sarrion-Perdigones, A., Weithorn, E., Pliner, M., Orzaez, D., Granell, A., Rogachev, I., and Aharoni, A. (2016). Elucidation of the first committed step in betalain biosynthesis enables the heterologous engineering of betalain pigments in plants. New Phytol. **210**:269–283. https://doi.org/10.1111/nph.13796.

Rickli, A., Moning, O.D., Hoener, M.C., and Liechti, M.E. (2016). Receptor interaction profiles of novel psychoactive tryptamines compared with classic hallucinogens. Eur. Neuropsychopharmacol **26**:1327–1337. https://doi.org/10.1016/j.euroneuro.2016.05.001.

Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. J. Comput. Aided Mol. Des.Aide **27**:221–234.

Schenck, C.A., and Maeda, H.A. (2018). Tyrosine biosynthesis, metabolism, and catabolism in plants. Phytochemistry **149**:82–102. https://doi.org/10.1016/j.phytochem.2018.02.003.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**:3210–3212.

Smith, T.A. (1977). Phenethylamine and related compounds in plants. Phytochemistry **16**:9–18. https://doi.org/10.1016/0031-9422(77)83004-5.

Song, L., and Florea, L. (2015). Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. GigaScience **4**:48.

Sunnadeniya, R., Bean, A., Brown, M., Akhavan, N., Hatlestad, G., Gonzalez, A., Symonds, V.V., and Lloyd, A. (2016). Tyrosine hydroxylation in betalain pigment biosynthesis is performed by cytochrome P450 enzymes in beets (Beta vulgaris). PLoS One **11**:e0149417. https://doi.org/10.1371/journal.pone.0149417.

UniProt Consortium. (2008). The universal protein resource (UniProt). Nucleic Acids Res. **36**:D190–D195.

Watkins, J.L., Li, Q., Yeaman, S., and Facchini, P.J. (2023). Elucidation of the mescaline biosynthetic pathway in Peyote (*Lophophora williamsii*). Plant J. **116**:635–649. https://doi.org/10.1111/tpj.16447.

Wilkox. (2023). https://github.com/wilkox/gggenes.