PNAS

Supporting Information for

Elucidation of tropane alkaloid biosynthesis in *Erythroxylum coca* using a microbial pathway discovery platform

Benjamin G. Chavez, Prashanth Srinivasan, Neill Kim, Olga Estrada, Kayla Glockzin, Jan Jirschitzka, Gage Rowden, Jonathan Shao, Lyndel Meinhardt, Christina D. Smolke*, John C. D'Auria*

*Correspondence to: John C. D'Auria & Christina D. Smolke

Email: dauria@ipk-gatersleben.de & csmolke@stanford.edu

This PDF file includes:

SI Materials and Methods
Figures S1 to S30
Tables S1 to S4
Legends for Dataset S1
SI References

Other supporting materials for this manuscript include the following:

Dataset S1

Supplementary Materials and Methods

Controlled substances

All work involving Schedule II controlled substances was performed under the authorization of Stanford University's institutional research registration with the U.S. Drug Enforcement Agency (DEA). Work with coca plants and controlled substances in Germany were performed under license authorization by the Federal Institute for Drugs and Medical Devices (BfArM). Additional biosecurity measures included background screening for researchers handling Schedule II compounds and yeast strains capable of producing them; restrictions on culture volume for yeast metabolite production (all performed in deep-well 96-well plates, 300 µL culture volume); protocols for sample storage and/or disposal after experiments; and explicit reporting of all experiments involving Schedule II controlled substances at concentrations at which they were deemed retrievable from stored samples.

Chemical compounds and standards

Putrescine dihydrochloride (sc-202786), *N*-methylputrescine dihydrochloride (sc-212242), and hygrine (sc-488263) were purchased from Santa Cruz Biotechnology (Dallas, TX). Spermidine (S2626-5G) was purchased from Sigma (St. Louis, MO). A DEA-exempt (non-Schedule II) stock solution of methylecgonine in acetonitrile (E-001-1ML; 1 mL of 1.0 mg/mL) was purchased from Sigma. *N*-methylpyrrolinium (NMPy) was prepared from γ-Methylaminobutyraldehyde (4MAB) diethyl acetal (M285740) purchased from Toronto Research Chemicals (Toronto, ON), as described previously (1). MPMOB was prepared as described previously (2). All other chemicals were purchased from Sigma.

NMR spectra

All NMR spectra were collected using a JEOL ECS 400 MHz NMR spectrometer.

Synthesis of 2-(3-((4-(1,3-dioxoisoindolin-2-yl) butyl) amino) propyl) isoindoline-1,3-dione (Compound 1)

Spermidine (2.75mmol, 400mg) and phthalic anhydride (5.50mmol, 816mg) were dissolved in 20mL glacial acetic acid, and the mixture was heated until reflux overnight. The solution was concentrated with a rotary evaporator and then dried under a Schlenk line for an additional 1.5 hrs to remove residual acetic acid. The crude product was a light orange powder. Crude yield = 1.41g. **^{1}H-NMR** (ClCD3) of 1: = 1.73 (4H, q), 2.06 (2H, d), 2.91 (4H, q), 3.69 (2H, t), 3.79 (2H, t), 7.70 (4H, m), 7.81 (4H, m) ppm. **^{13}C-NMR** (ClCD3) of 1: = 24.33 (s, 1C), 25.99 (s, 1C), 26.47 (s, 1C), 35.08 (s, 1C), 37.24 (s, 1C), 45.48 (s, 1C), 47.68 (s, 1C), 123.39 (s, 2C), 123.55 (s, 2C), 132.07 (s, 2C), 132.18 (s, 2C), 134.08 (s, 2C), 134.26 (s, 2C), 168.48 (s, 2C), 168.66 (s, 2C) ppm. **FTIR (ATR, cm^{-1})**: 2979 (m, Csp^3-H), 1771 (s, Csp^2=O) 1694 (s, Csp^2-Csp^2), 1594 (w, Csp^2–Csp^2), 1262 (s, Csp^3-N).

Synthesis of 2-(3-((4-(1,3-dioxoisoindolin-2-yl)butyl)(methyl)amino)propyl)isoindoline-1,3-dione (Compound 2)

2-(3-((4-(1,3-dioxoisoindolin-2-yl)butyl)amino)propyl)isoindoline-1,3-dione (1) with traces of acetic acid (362.3mg, 0.8mmol) and formalin (0.2ml, 7.1mmol) were added to 4mL of acetic acid. The mixture was stirred overnight at 100°C and then cooled to room temperature. The solution was adjusted with 3N NaOH to a pH of 10, where precipitation began, and the solid (product) was collected by centrifugation. The crude product was a dark brown oil residue. Crude product yield = 271.02mg. **^{1}H-NMR** spectra were collected. **^{1}H-NMR** (ClCD3) of 2: 1.47 (2H, m), 1.67 (2H, m), 1.84 (2H, m), 2.18 (3H, s), 2.36, 2.42 (4H, m), 3.76 (4H, m), 7.70 (4H, m), 7.83 (4H, m) ppm. **^{13}C-NMR** (ClCD3) of 2: = 23.21 (s, 1C), 25.35 (s, 1C), 26.47 (s, 1C), 36.20 (s, 1C), 37.63 (s, 1C), 40.94 (s, 1C), 54.43 (s, 1C), 56.26(s,1C) 123.38 (s, 5C), 132.22 (s, 2C), 134.08 (s, 3C), 168.51 (s, 2C), 168.56 (s, 2C) ppm. **FTIR (ATR, cm^{-1})**: 2979 (s, Csp^3-H), 1761 (s, Csp^2=O) 1699 (s, Csp^2-Csp^2), 1588 (w, aromatic Csp^2–Csp^2), 1262 (s, Csp^3-N).

Synthesis of *N*-methylspermidine (Compound 3)

Biochemical characterization of SMT enzymes in *E. coca* necessitated synthesis of *N*-methylspermidine (**3**); the general chemical workflow is described here and depicted in Fig. S16. Synthesis started with protection of both primary amino groups of spermidine with phthalic anhydride to yield compound **1**. This reaction is well described in literature. After protection of the primary amines, the central amine (a

secondary amine) was methylated using formaldehyde to yield compound **2**, after which concentrated hydrochloric acid was used to deprotect the primary amines to yield compound **3**. (Note: removal of phthalic anhydride is typically performed with hydrazine; however, as deprotection with hydrazine monohydrate takes the same time as with hydrochloric acid and requires highly hazardous substances, hydrochloric acid was chosen instead). Spectrometric analysis of the isolated products confirmed the formation of: compounds **1**, **2**, and **3** (Fig. S17-S25). The full synthesis protocol is provided below.

2-(3-((4-(1,3-dioxoisoindolin-2-yl) butyl) (methyl)amino)propyl)isoindoline-1,3-dione with traces of water and acetic acid (200mg, 0.47mmol) and 10mL of concentrated hydrochloric acid were added to a reaction vessel. The mixture was refluxed overnight at 128°C, and the solution was cooled to room temperature, causing a precipitate to crash out of solution. The precipitate was removed via filtration, and the solution was concentrated with a rotary evaporator. The crude product was rinsed with acetone, and the acetone was removed by decanting. A brown powder was collected as pure *N*-methylspermidine. Crude product yield = 44.18mg. NMR spectra were collected. **^{1}H-NMR** (D2O) of compound 3: 1.74 (4H, m), 2.13 (2H, m), 2.89 (3H, s), 3.08 (4H, m), 3.28 (4H, m) ppm. **^{13}C-NMR** (D2O spiked with acetone-d6) of 3: = 22.33 (s, 1C), 26.27 (s, 1C), 31.21 (s, 1C), 37.21 (s, 1C), 43.95 (s, 1C), 51.90 (s, 1C), 56.74 (s, 1C), 59.71 (s, 1C) ppm. **FTIR (ATR, cm^{-1})**: 3371(w, N-H), 2979 (s, Csp^3-H), 1262 (s, Csp^3-N). Images of all ^{1}H, ^{13}C NMR, and FTIR spectra are available in Figs. S17-S25.

Plasmid construction

DNA oligonucleotides used in this study were synthesized by Integrated DNA Technologies (Coralville, IA). Endogenous yeast genes were amplified from *Saccharomyces cerevisiae* CEN.PK2-1D (3) genomic DNA via colony PCR (4). Gene sequences encoding heterologous enzymes were codon-optimized for expression in *S. cerevisiae* using GeneArt GeneOptimizer software (ThermoFisher Scientific) and synthesized as double-stranded gene fragments (Twist Bioscience). Plasmids used in this study are listed in Table S3. Three types of plasmids were used in this work: yeast expression plasmids, yeast integration plasmids, and *Agrobacterium tumefaciens* binary vectors.

Yeast expression plasmids harbored a gene of interest flanked by a constitutive promoter and terminator, an auxotrophic selection marker, and a low-copy CEN6/ARS4 yeast origin of replication.

These plasmids were constructed via Gibson assembly of insert fragments synthesized with 5' (AGTCAAGGATCCACTAGTCACAAAA) and 3' (TAATAGCCAGTTCTCGAGTGTCAG) adapter sequences with linearized yeast vector backbones pAG414GPD-ccdB, pAG415GPD-ccdB, pAG416GPD-ccdB, pAG424GPD-ccdB, pAG425GPD-ccdB, or pAG426GPD-ccdB (5) with identical terminal adapters added using PCR. Yeast expression plasmids expressing fusions of multiple proteins or enzymes (e.g., fluorescent protein fusions) were prepared by PCR-amplification of each gene of interest with 15-25 bp of overlap to adjacent fragments and Gibson assembly into the selected expression vector.

Yeast integration plasmids comprised a gene of interest flanked by a constitutive promoter and terminator, but lacked a selection marker and origin of replication for yeast expression. These plasmids were constructed by PCR linearization of the empty holding vectors pCS2656, pCS2657, pCS2658, pCS2661, or pCS2663 (6) using primers complementary to the 3' and 5' ends of the promoter and terminator, respectively. Genes intended for yeast genomic integration were PCR-amplified to append 5' and 3' overhangs with 35–40 bp of homology to the termini of the linearized holding vectors and then assembled using Gibson assembly.

For transient expression of *E. coca* pathway genes in *Nicotiana benthamiana*, *A. tumefaciens* pHREAC binary vectors (7) harboring a transfer-DNA (T-DNA) region comprising a gene of interest flanked by the constitutive Cauliflower Mosaic Virus (CaMV) 35S promoter/ 5S0 synthetic 5'UTR and a nopaline synthase terminator, as well as an analogous expression cassette for the p19 RNAi-suppressor protein. Golden Gate assembly was used to clone each gene of interest (with flanking 5' and 3' BsaI sites added via PCR) into the vector. All primers designed for cloning *E. coca* pathway genes into the pHREAC vector were designed in Benchling using the Golden Gate Assembly program and subsequent primers were ordered from metabion (metabion GmbH). All PCRs were conducted with vector specific primers pHREAC FWD 5'-(CAACCACAACGCTCTAACGC)-3' and pHREAC REV 5'-(AAAATTAAAATCTTTTTGTGTCCTTGCT)-3' for sequencing and colony PCR.

All PCR amplification was performed using Q5 DNA polymerase (NEB) and linear DNA was purified using the DNA Clean and Concentrator-5 kit (Zymo Research). Assembled plasmids were transformed into chemically competent *Escherichia coli* (TOP10, Thermo Fisher Scientific) via heat-shock and propagated with selection in Luria-Bertani (LB) broth or on LB-agar plates with either carbenicillin

(100 μg/mL) or kanamycin (50 μg/mL) selection. *E. coli* plasmid DNA was isolated by alkaline lysis from overnight cultures grown at 37 °C and 250 rpm in selective LB media using Econospin columns (Epoch Life Science) or QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. Plasmid sequences were verified by Sanger sequencing (Quintara Biosciences).

Yeast strain construction

Yeast strains used in this study (Table S2) were derived from the parental strain CEN.PK2-1D (3). Strains were grown non-selectively in yeast-peptone media supplemented with 2% w/v dextrose (YPD media); in yeast nitrogen base (YNB) defined media (Becton, Dickinson and Company (BD)) supplemented with synthetic complete amino acid mixture (YNB-SC; Clontech), 2% w/v dextrose, and 5% v/v glycerol to support heterologous protein expression and NADPH availability (6); or on agar plates of the aforementioned media. Strains transformed with plasmids bearing auxotrophic selection markers (*URA3*, *TRP1*, and/or *LEU2*) were grown selectively in YNB media supplemented with 2% w/v dextrose, 5% v/v glycerol, and the appropriate dropout solution (YNB-DO; Clontech) or on YNB-DO agar plates. Yeast strains lacking the *ALD6* gene were deficient in acetate metabolism and were therefore grown on the aforementioned media supplemented with 0.1% w/v potassium acetate (i.e., YPAD or YNBA).

Yeast genomic modifications were performed using the CRISPRm method (8). CRISPRm plasmids expressing *Streptococcus pyogenes* Cas9 (*Sp*Cas9) and a single guide RNA (sgRNA) targeting a genomic locus were constructed by assembly PCR and Gibson assembly of DNA fragments encoding *Sp*Cas9 (pCS3410), tRNA promoter and HDV ribozyme (pCS3411), a 20-nt guide RNA sequence oligonucleotide, and tracrRNA and terminator (pCS3414) (6). Integration fragments comprising one or more genes of interest flanked by unique promoters and terminators were PCR-amplified from yeast integration plasmids using Q5 DNA polymerase (NEB) with flanking 40 bp microhomology regions to adjacent fragments and/or to the yeast genome at the integration site. Approximately 0.5-1 μg of each integration fragment was co-transformed with 500 ng of CRISPRm plasmid targeting the desired genomic site. Positive integrants were identified by yeast colony PCR, Sanger sequencing, and functional screening by liquid chromatography and tandem mass spectrometry (LC-MS/MS).

Yeast transformations

Yeast strains were chemically transformed using the Frozen-EZ Yeast Transformation II Kit (Zymo Research) as per the manufacturer's instructions, with the following modifications. For competent cell preparation, individual colonies were inoculated into YP(A)D media and grown overnight at 30 °C and 460 rpm. Saturated cultures (~14-18 h) were back-diluted between 1:10 and 1:50 in fresh YP(A)D media and grown to exponential phase (~5–7 h). Cultures were pelleted by centrifugation at 500 × g for 4 min, washed twice with EZ1 solution, and then resuspended in 20-50 µL of EZ2 solution per transformation. For transformation, competent cells were mixed with 250–1000 ng of total DNA and 200-500 µL of EZ3 solution. Cell suspensions were incubated at 30 °C with slow rotation for 1-1.5 h. For plasmid transformations, 50-250 µL of cell suspension was directly plated onto YNB(A)-DO agar plates (Teknova). For CRISPRm genomic modifications, yeast suspensions were instead mixed with 1mL YP(A)D media, pelleted by centrifugation at 500 × g for 4 min, and then resuspended in 300-500 µL of fresh YP(A)D media. Suspensions were incubated at 30 °C with gentle rotation for 3-4 h to allow expression of geneticin resistance and then spread on YP(A)D plates supplemented with 200-400 mg/L G418 (geneticin) sulfate (Teknova). Plates were incubated at 30 °C for 72 h to allow sufficient colony formation before downstream applications.

Growth conditions for metabolite assays

Small-scale metabolite production assays were performed in YNB(A)-SC or YNB(A)-DO media supplemented with 2% dextrose and 5% glycerol in at least three biological replicates. Our previous work showed that accumulation of tropane alkaloid intermediates in yeast is significantly enhanced by higher starting cell densities (1, 6). Therefore, yeast colonies were initially inoculated in triplicate into 1.5 mL YP(A)D or YNB(A)-DO and grown to saturation (~18-22 h) at 30 °C and 460 rpm, pelleted by centrifugation at 500 × g for 4 min and 3,000 × g for 1 min, resuspended in 1 mL of fresh selective or non-selective media, and then 300 µL transferred into 2 mL deep-well 96-well plates sealed with AeraSeal gas-permeable film (Excel Scientific). Cultures were grown for 72-96 h at 25 °C, 460 rpm, and 80% relative humidity in a Lab-Therm LX-T shaker (Adolf Kuhner).

RNA extraction and cDNA synthesis on *E. coca* Tissue

Approximately 100 mg of fresh *E. coca* tissue was used to extract total RNA using a Qiagen plant RNA extraction kit, and genomic DNA was degraded via an in-column RNase-free DNase I treatment (Qiagen, Germany). The quality and quantity of RNA were determined using a micro-volume spectrophotometer DeNovix DS-11 (DeNovix, Wilmington, USA). cDNA was synthesized with 2 g of total RNA using a SuperScript II First Strand Kit (Invitrogen, Germany) according to the manufacturer's protocol.

Quantitative real-time PCR (qRT-PCR) of *E. coca* tissue samples

Gene-specific qRT-PCR primers were synthesized and purified by Integrated DNA Technologies (IDT, USA). To identify single amplicon products, legacy PCR was performed on all primer sets with each reaction consisting of 2 µL Taq polymerase, 10 µL 5X GoTaq buffer (Promega, USA), 1 µL of 1mM dNTP mix, 1 µL of each primer and 1 µL cDNA with the reaction volume brought up to 50 µL using sterile DDI water. PCR cycling conditions were as follows: 94°C for 45 sec, 30 cycles of denaturation at 94°C for 45 sec, annealing at 60°C for 30 sec, extension at 72°C for 30 sec, and a final extension at 72°C for 7 min. Amplified products were visualized on a 4% agarose gel with a white/dual UV Transilluminator photoimager (VWR, US). The baseline reference genes Ec6409 and Ec10131 were selected based on experimental evidence from previous studies (9).

All qPCR experiments were performed on a QuantStudio3 instrument (Thermo Fisher Scientific, USA) with each reaction consisting of 5 µl PowerUp SYBR green Master Mix (Thermo Fisher Scientific, USA), 0.5 µl of each primer, and 2 µl of cDNA brought up to a final volume of 10 µl. The thermocycling conditions are as follows: denaturation at 95°C for 10 min, 40 cycles of denaturation at 95°C for 15 sec and annealing/extension at 60°C for 1 min. Experiments were performed using three biological replicates and three technical replicates. Primer efficiencies were determined by a standard curve based on seven different two-fold dilutions of cDNA cloned amplicon for each primer set. At the end of each run, melting curve analysis was performed to rule out the presence of primer dimers or non-specific amplicons. RNA-only controls and non-template controls were included for each primer set to ensure the absence of contamination. The raw qPCR data was exported to Microsoft Excel using the qBase v1.3.5 macro (10). Primer efficiencies were analyzed by linear regression within the qBase software. All relative quantities

were calculated in the qBase software accounting for each primer efficiency and normalizing to reference genes Ec10131 and Ec6409.

Cloning, heterologous expression and purification of *Ec*SMT, *Ec*SPMT, and *Ec*SPDS from yeast

The pEPStrepGW expression vector was modified from the pPICHOLI Shuttle Vector System designed for heterologous expression in *K. phaffii* KM71 cells as well as in *E. coli* and contains a strep-tag located on the N-terminus of the gene sequence. The pEPStrepGW vector contains a yeast inducible alcohol oxidase (AOX) promoter and an *E. coli* T7 promoter. The expression vector was then introduced into *K. phaffii* KM71, and a starter culture was grown overnight in YPD medium supplemented with 100 μg mL^{-1} of Zeocin (InvivoGen) at 28 °C with shaking at 250 rpm. Fresh BMMY (Buffered Methanol-complex Medium) medium supplemented with 100 μg mL^{-1} of Zeocin (InvivoGen) was then inoculated with the overnight culture (10% final concentration of cell suspension) and grown at 28 °C with shaking at 250 rpm to an OD_{600} of 1.0. Protein expression was induced with the addition of 1% methanol (v/v) twice a day (morning and evening) and grown at 28 °C with shaking at 250 rpm for 2-3 days. The cells were then harvested at 2000 g at 4 °C for 10 min and then stored at -20 °C until ready for purification.

For protein purification, cells were resuspended in 100mM Tris-HCl buffer, pH 8, supplemented with 150mM NaCl, 1mM EDTA, and 5mM dithiothreitol then lysed using a pressure cell homogenizer (Avestin EmulsiFlex-C3). The lysate was then centrifuged at 15,000 g at 4 °C for 15 min. The soluble protein was then loaded onto a Strep-Tactin superflow column (IBA Lifesciences) using an Äkta Pure protein purification system (Cytiva). The recombinant protein was eluted using 100 mM Tris-HCl, pH 8, supplemented with 150 mM NaCl, 1 mM EDTA, and 2.5 mM desthiobiotin. Fractions containing the protein of interest were then loaded onto a HiTrap Desalting column and eluted using 100mM Tris-HCl pH 8.4 containing 15% glycerol and 5mM dithiothreitol. Concentration of total protein was measured using the Bradford protein assay (Bio-Rad) according to the manufacturer's protocol. The putative size and purity of the protein was determined via SDS-PAGE. Approximately 100 ng of total protein was loaded into each well and stained with Colloidal Coomassie staining solution.

Polyamine extraction of *E. coca* tissues

Polyamines were extracted from various tissues of *E. coca* plants. Plant tissue was cut from the plant and immediately frozen in liquid nitrogen. The frozen tissue was ground to a fine powder with a mortar and pestle. Approximately 100mg of tissue powder was briefly vortexed with extraction buffer (1:5 w/v ratio of plant material:extraction buffer). The extraction buffer is a solution of 80% methanol spiked with 5% perchloric acid (5mL of 80% methanol and 100uL of 5% PCA spiked with a final concentration of 10mM piperidine). The mixture is stored on ice for 1 hour and briefly vortexed every 15-30 minutes. Tubes are then centrifuged for 15 minutes at 21,000 x g at 4°C to remove plant material. The supernatant was then used immediately for analysis or stored at -20°C.

Derivatization of polyamines

To detect polyamines (PAs) using fluorescence, they first must be derivatized using the derivative agents o-phthalaldehyde-ethanethiol (OPA-ET) and fluorenylmethyloxycarbonyl carbamate (FMOC) as described by Hanczkó et al. (2007) (11) (Fig. S26). A 50uL aliquot of sample was derivatized with 50uL of OPA-ET stock and left to react for 1.5 minutes, then 2uL of the FMOC stock was added and allowed to sit for 2 minutes before injection.

Polyamine detection by HPLC-Fluorescence

For the detection of the polyamines (Fig. S27-S30), an Ultimate-3000 HPLC (Thermo Scientific) with a fluorescence detector (337nm Emission/ 454nm Excitation) with a Nucleodur Sphinx column (Macherey Nagel) was used. The buffers used for this method of separation are 0.2% formic acid (Buffer A) and 100% acetonitrile (Buffer B). Flow rate: 1mL/min. Injection volume: 5uL, Gradient is as follows: 0 min, 10% B; 0-8 min, 10-100% B; 8-15.17 min, 100% B; 15.17-15.34 min, 10% B; and 15.34-19.34 min, 10% B for column equilibration.

***In vitro* enzyme assays for SPDS and SMT activity**

Enzyme assays were performed as follows: SMT assays contained 2μL of putrescine from 50mM stock final concentration 1mM, 1μL dcSAM from 50mM stock final concentration 0.5mM, 20μL of purified enzyme, Total volume up to 100μL in 50mM potassium phosphate buffer (pH 7.6). All SPDS assays

contained 2µL of spermidine from 50mM stock final concentration 1mM, 1µL SAM from 50mM stock final concentration 0.5mM, 20µL of purified enzyme, Total volume up to 100µL in 50mM potassium phosphate buffer (pH 7.6). Let the enzyme assays stand overnight at RT. Quench the assays by heating at 95°C for 10 minutes. Briefly vortex and centrifuge mixture. Derivatize the PAs with OPA-ET and FMOC.

Analysis of metabolite production in engineered yeast cultures

Yeast cultures were pelleted by centrifugation at 3500 × g for 5 min at 12-16 °C and 150 µL aliquots of supernatant were removed for analysis. Metabolites were analyzed by LC-MS/MS using an Agilent 1260 Infinity Binary HPLC and an Agilent 6420 Triple Quadrupole or Agilent 6490 Triple Quadrupole mass spectrometer. Chromatography was performed using a Zorbax EclipsePlus C18 column (2.1 × 50 mm, 1.8 µm; Agilent Technologies) with 0.1% v/v formic acid in water as mobile phase solvent A and 0.1% v/v formic acid in acetonitrile as solvent B. The column was operated with a constant flow rate of 0.4 mL/min at 40 °C and a sample injection volume of 10 µL. Chromatographic separation was performed using the following gradient[1]: 0.00–0.75 min, 1% B; 0.75–1.33 min, 1–25% B; 1.33–2.70 min, 25–40% B; 2.70–3.70 min, 40–60% B; 3.70–3.71 min, 60–95% B; 3.71–4.33 min, 95% B; 4.33–4.34 min, 95–1% B; 4.34–5.00 min, equilibration with 1% B. The LC eluent was directed to the MS from 0.01–5.00 min operating with electrospray ionization (ESI) in positive mode, source gas temperature 350 °C, gas flow rate 11 L/min, and nebulizer pressure 40 psi. Metabolites were identified using MassHunter Workstation software (Agilent) based on retention time and the mass fragment/transition parameters in Table S4. Multiple reaction monitoring (MRM) transitions were identified by analysis of 0.1–1 mM aqueous standards and/or yeast culture control samples using MassHunter Optimizer software (Agilent) and corroborated against published mass transitions if available, and/or against predicted transitions determined using the CFM-ID fragment prediction utility (12) and the METLIN database (13).

Protein expression in *N. benthamiana*

Agrobacterium tumefaciens strain GV2260 cells were transformed with pHREAC binary transformation vectors (7) containing *E. coca* genes of interest using the freeze-thaw method (14) and confirmed via PCR. A single *Agrobacterium* colony containing the construct of interest was cultured in liquid YEB (5 g/L

beef extract, 1 g/L yeast extract, 5 g/L peptone, 5 g/L sucrose, 0.5 g/L $MgCl_2$) medium overnight at 28°C with appropriate antibiotics. The bacteria were pelleted at 5000 g for 30 min, resuspended in infiltration media (10 mM MES pH 5.6, 10 mM $MgCl_2$), and adjusted to an OD_{600} between 0.6 and 1.0. *Agrobacterium* cultures transformed with the tobacco etch potyvirus helper component protein (HCPro) silencing suppressor (15) were resuspended in infiltration media to the same OD_{600}. An equal volume of cultures containing the construct of interest and HCPro were mixed and transiently infiltrated into 4-week-old *N. benthamiana* leaves by using a vacuum system for 90 seconds. After infiltration, the plants were kept in the greenhouse for six days in the same condition as above. The plant leaves were collected on day six and stored at -80 °C for further analysis. Metabolites were extracted from approximately 100 mg of powdered leaves at a ratio of 100 mg mL^{-1} of extraction solvent; 20% methanol containing 0.1% formic acid, and 1000 ng/mL of deuterated *N*-methylpyrolinium (D_3NMPy) as an internal standard. Extracts were incubated on an orbital shaker for approximately overnight at room temperature. *N. benthamiana* leaves were extracted and analyzed immediately after extraction by liquid chromatography-tandem mass spectrometry.

Metabolite analysis of *N. benthamiana* extracts

Secondary metabolites were analyzed by LC-MS/MS using an Agilent 1260 Infinity Binary HPLC and an Agilent 6490 Triple Quadrupole mass spectrometer. Chromatography was performed using an Eclipse Plus C18 column (2.1 × 50 mm, 1.8 µm; Agilent Technologies) with 0.1% v/v formic acid in water as mobile phase solvent A and 0.1% v/v formic acid in acetonitrile as solvent B. The column was operated with a constant flow rate of 0.4 mL/min at 40 °C and a sample injection volume of 1 µL. Chromatographic separation was performed using the following gradient: 0.00–0.75 min, 1% B; 0.75–1.33 min, 1–25% B; 1.33–2.70 min, 25–40% B; 2.70–3.70 min, 40–60% B; 3.70–3.71 min, 60–95% B; 3.71–4.33 min, 95% B; 4.33–4.34 min, 95–1% B; 4.34–5.00 min, equilibration with 1% B. The LC eluent was directed to the MS from 0.01–5.00 min operating with iFunnel electrospray ionization (ESI) in positive mode, source gas temperature 250 °C, sheath gas temperature 300°C, gas flow rate 12 L/min, and nebulizer pressure 30 psi. Metabolites were identified using MassHunter Workstation software ver B.07.01 (Agilent, Germany) based on retention time and the mass fragment/transition parameters in Table S4. Multiple reaction

monitoring (MRM) transitions and quantification were performed by the analysis of 0.1ng/mL-0.1mg/mL aqueous standards using MassHunter Optimizer software (Agilent) and corroborated against published mass transitions if available, and/or against predicted transitions determined using the CFM-ID fragment prediction utility and the METLIN database.

Fluorescence microscopy

Individual colonies of yeast strains transformed with plasmids encoding biosynthetic enzymes fused to fluorescent protein reporters were inoculated into 1 mL selective media and grown overnight (~14-18 h) at 30 °C and 460 rpm. Overnight cultures were back-diluted between 1:2 and 1:4 into fresh media and grown to exponential phase at 30 °C and 460 rpm for an additional 6–8 h to allow slow-maturing fluorescent proteins to fold before imaging. Approximately 5–10 µL of cell suspension was spotted onto a glass microscope slide and covered with a glass coverslip (Thermo Fisher) and then imaged using an upright Zeiss AxioImager Epifluorescence/Widefield microscope with a ×64 oil immersion objective. Fluorescence excitation was performed using an EXFO X-Cite 120 illumination source and the following Semrock Brightline filter settings: GFP, 472/30 excitation and 520/35 emission; mCherry/DsRed/Cy3/TexasRed, 562/40 excitation and 624/40 emission. Emitted light was captured with a Zeiss Axiocam 503 mono camera and Zen Pro software, and subsequent image analysis was performed in ImageJ/Fiji (NIH). Images were converted to pseudocolor using the 'Merge Channels' and 'Split Channels' functions (Image→ Color→Merge/Split Channels). For each sample, linear histogram stretching was applied across all images for a given channel to improve contrast.

Hierarchical clustering analysis of *Erythroxylum* transcriptomes

First, transcriptome datasets for five *E. coca* cultivars (specimen no. 48, 113, 124, 209, and 228) and one non-cocaine-producing species, *E. hondense*, were processed to retain only the transcript ID, closest BLAST hit and E-value, fold change, length, and abundance (in reads per kilobase of transcript per million mapped reads, RPKM) columns. Next, each dataset was further filtered to retain only BLAST-annotated CYP450, 2-ODD, and polyketide-related transcript candidates, as well as any transcripts manually annotated as encoding known *E. coca* TA pathway genes (i.e., SPDS, SPMT, SMT, AOF1, AOC1, AOC2,

PKS1/OGAS1, PKS2/OGAS2, MPOBMT, MecgoR, cocaine synthase) as positive control or 'bait' genes. At this stage, any transcripts without any close BLAST matches (E-value less than or equal to 1E-80) were removed, as such transcripts are typically incomplete or too short for further characterization.

After appending unique identifiers to eliminate duplicate BLAST hits, RPKM values for each transcript were normalized to the maximum RPKM value present within each dataset to correct for global differences in transcript abundance across biological samples. For each transcript/gene in each of the five *E. coca* datasets, the ratio of normalized RPKM in *E. coca* to normalized RPKM of its ortholog in *E. hondense* was computed. If a transcript/gene was not present in *E. hondense*, it was assigned a normalized RPKM in *E. hondense* equal to that of the lowest-abundance transcript to avoid divide-by-zero errors, and the *E. coca*:*E. hondense* RPKM ratio was rounded to the nearest whole number so that these outlier transcripts could be easily identified by manual annotation. Candidates in each dataset were then sorted by the *E. coca* : *E. hondense* abundance ratio to facilitate manual curation of ortholog matches across cultivars based on their BLAST hits.

Finally, expression profiles for each candidate gene were generated in the form of a matrix containing either $\log_2$ RPKM values or max-normalized RPKM values for each gene across each *E. coca* cultivar or *E. hondense* sample. Note that two separate matrices were generated, one with $\log_2$ RPKM values and one with max-normalized RPKM values. Hierarchical clustering heatmaps were generated using the `pheatmap` command with row (gene candidate) clustering but not column (*E. coca* cultivar or *E. hondense*) clustering for both matrices. As the heatmap generated using $\log_2$ RPKM values showed stronger clustering of known TA pathway genes, we used this heatmap for identification of methylecgonone synthase candidates for testing. A separate R script (`Find_hit_locus_IDs.R`) was used to collect transcript locus IDs for each candidate gene in each of the five *E. coca* samples; the translated open reading frames were then aligned and converted into a consensus sequence for each candidate.

Supplementary Figures

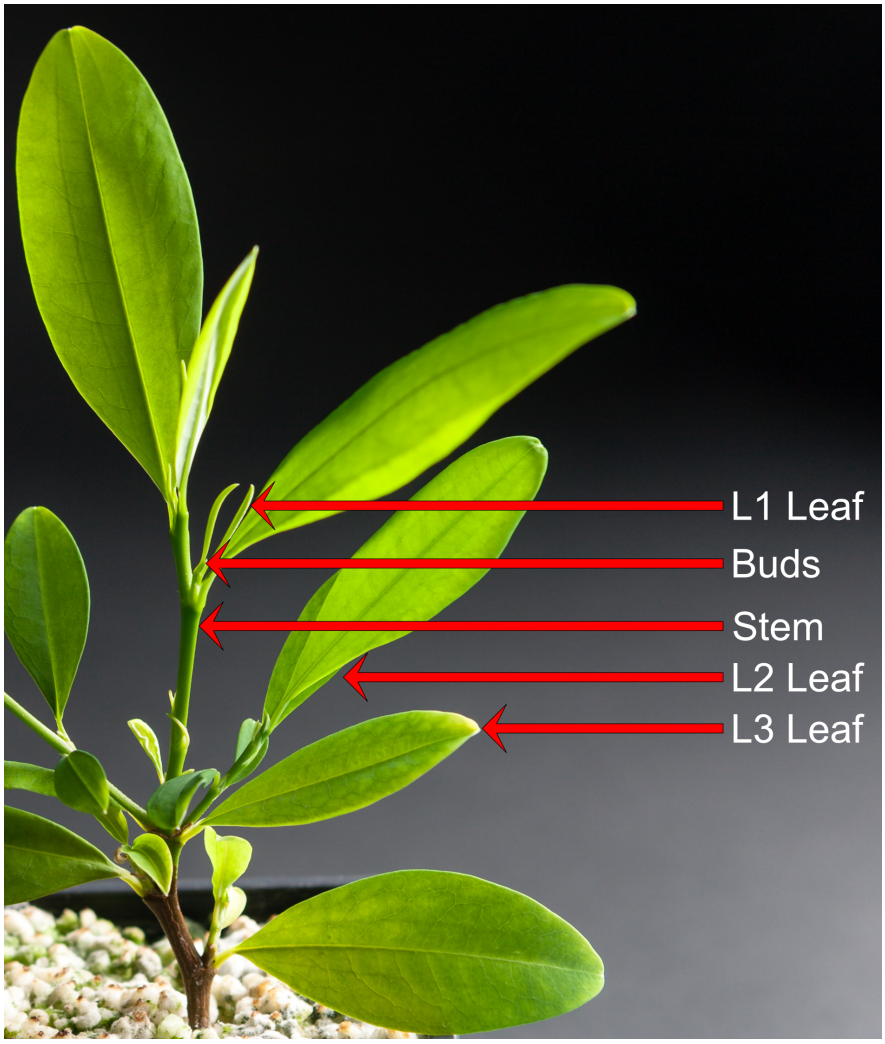
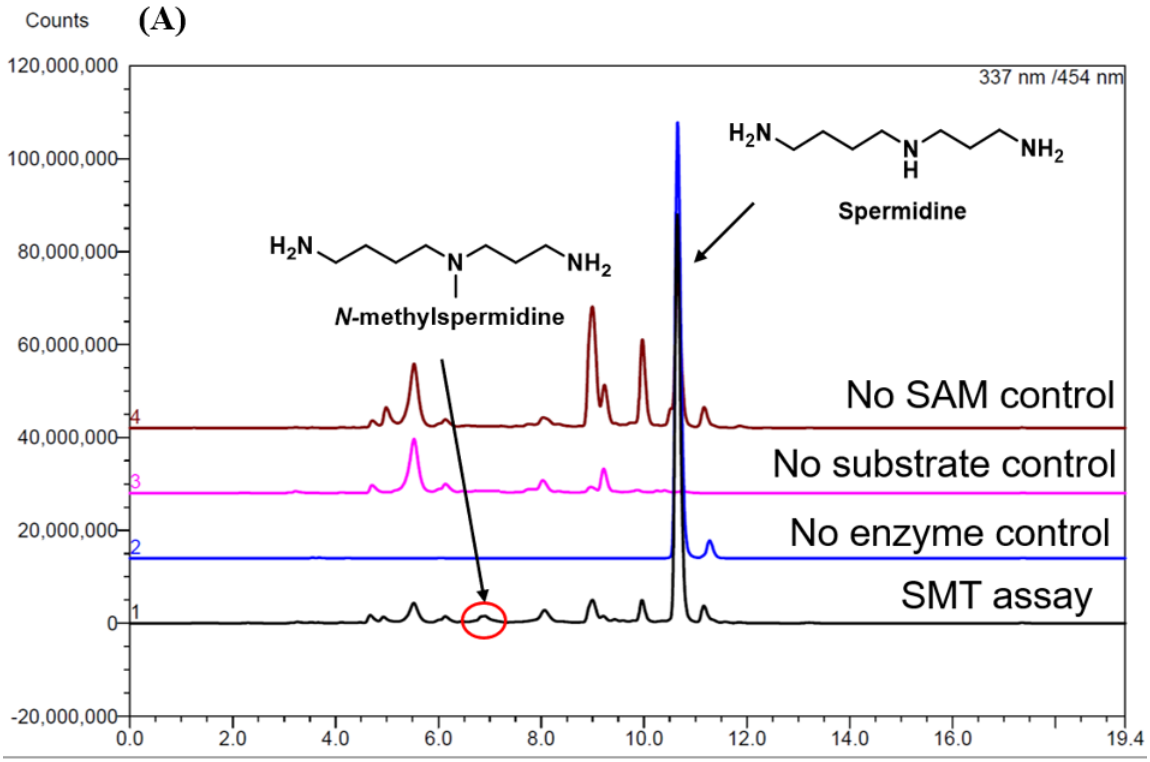
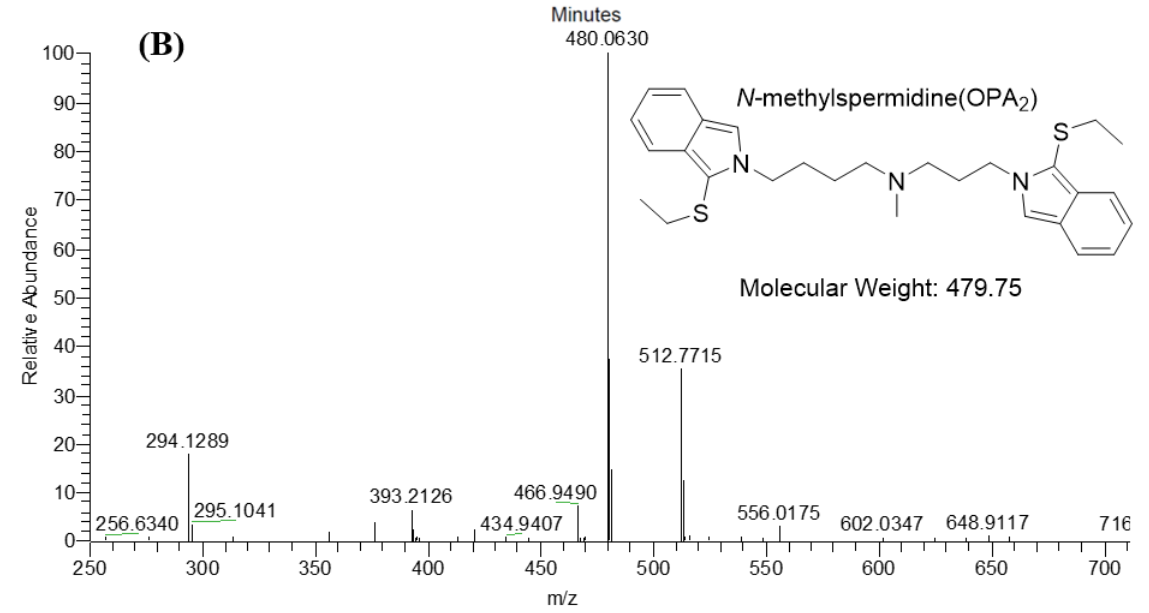


Figure S1. Photograph of *E. coca* plant showing the different developmental stages of the leaves. L1 stage leaf is a rolled leaf, L2 stage leaf is an unrolled leaf, and L3 stage leaf is a mature leaf. Stems and buds are also indicated.


(A)
Counts
337 nm /454 nm
Spermidine
N-methylspermidine
No SAM control
No substrate control
No enzyme control
SMT assay
Minutes


(B)
480.0630
N-methylspermidine(OPA2)
Molecular Weight: 479.75
512.7715
294.1289
393.2126
466.9490
256.6340
295.1041
434.9407
556.0175
602.0347
648.9117
Relative Abundance
m/z

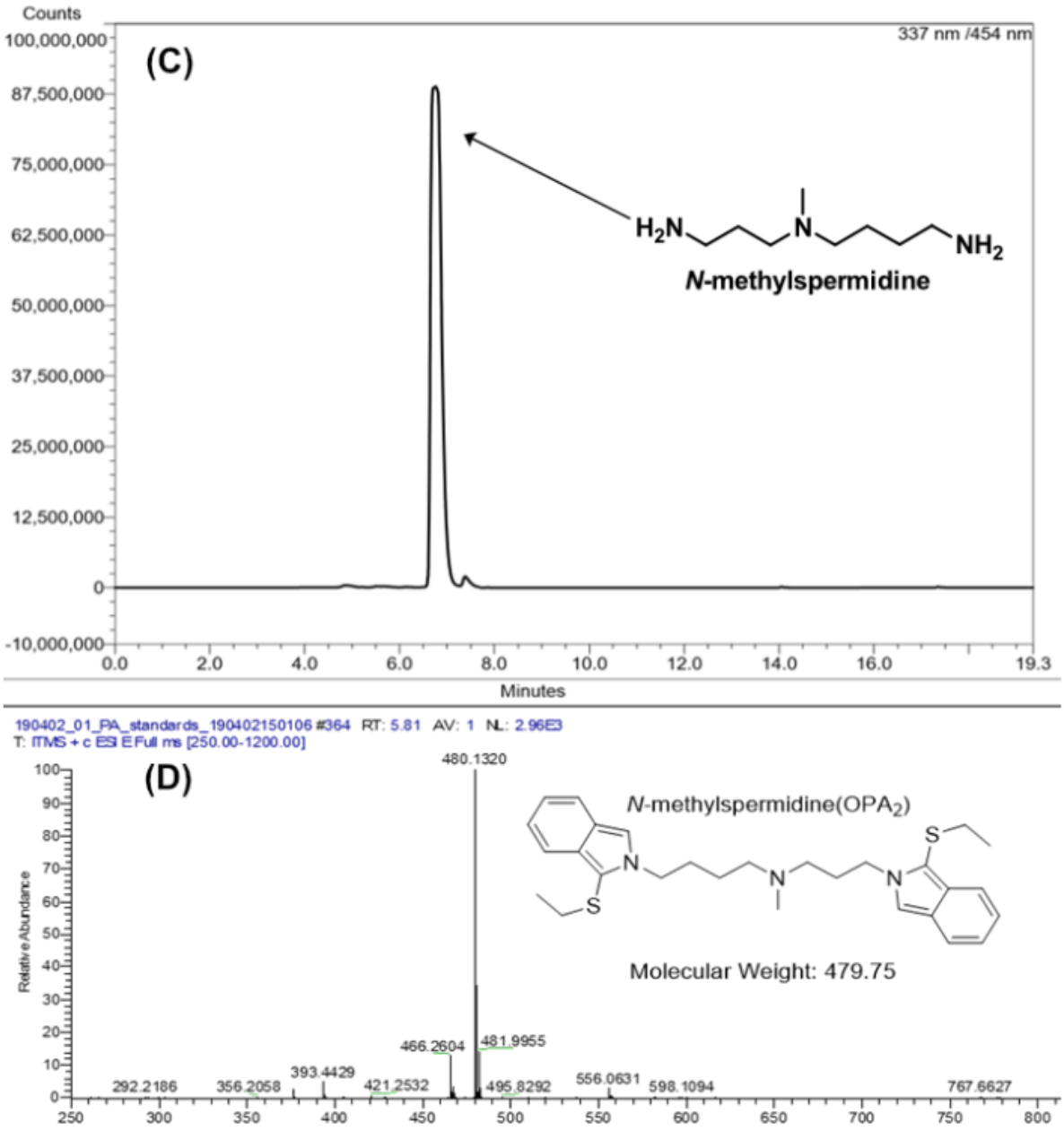


Figure S2. L3 leaves crude SMT activity assay. (**A**) HPLC-FLD chromatogram shows crude protein extracts from *E. coca* L3 leaves have SMT activity. Before injection the assay was derivatized with OPA and FMOC following the procedure reported in Hanczkó et. al. SMT assay (black), no enzyme control (blue), no substrate control (pink), no SAM control (brown). (**B**) HPLC-MS TIC (total ion chromatogram) of the E. coca crude SMT assay shows the peak present has a m/z of 480.06, which is consistent with derivatized MeSpd (MW: 479.75g/mol). The injection volume was increased to 10uL from 5uL for samples run on the LC-MS due to limitations of the autosampler. **(C)** HPLC-FLD chromatogram of *N*-methylspermidine standard (1mM) shows a retention times of approximately 6.8 minutes. Before injection, *N*-methylspermidine was derivatized with OPA-ET and FMOC following the procedure reported in Hanczkó et. al 2007. **(D)** HPLC-MS spectrum of *N*-methylspermidine standard (1mM) derivatized with OPA-ET shows a m/z ratio of 480.13 and *N*-methylspermidine(OPA2) has a molecular weight of 479.8g/mole and has retention time at 5.81 minutes.

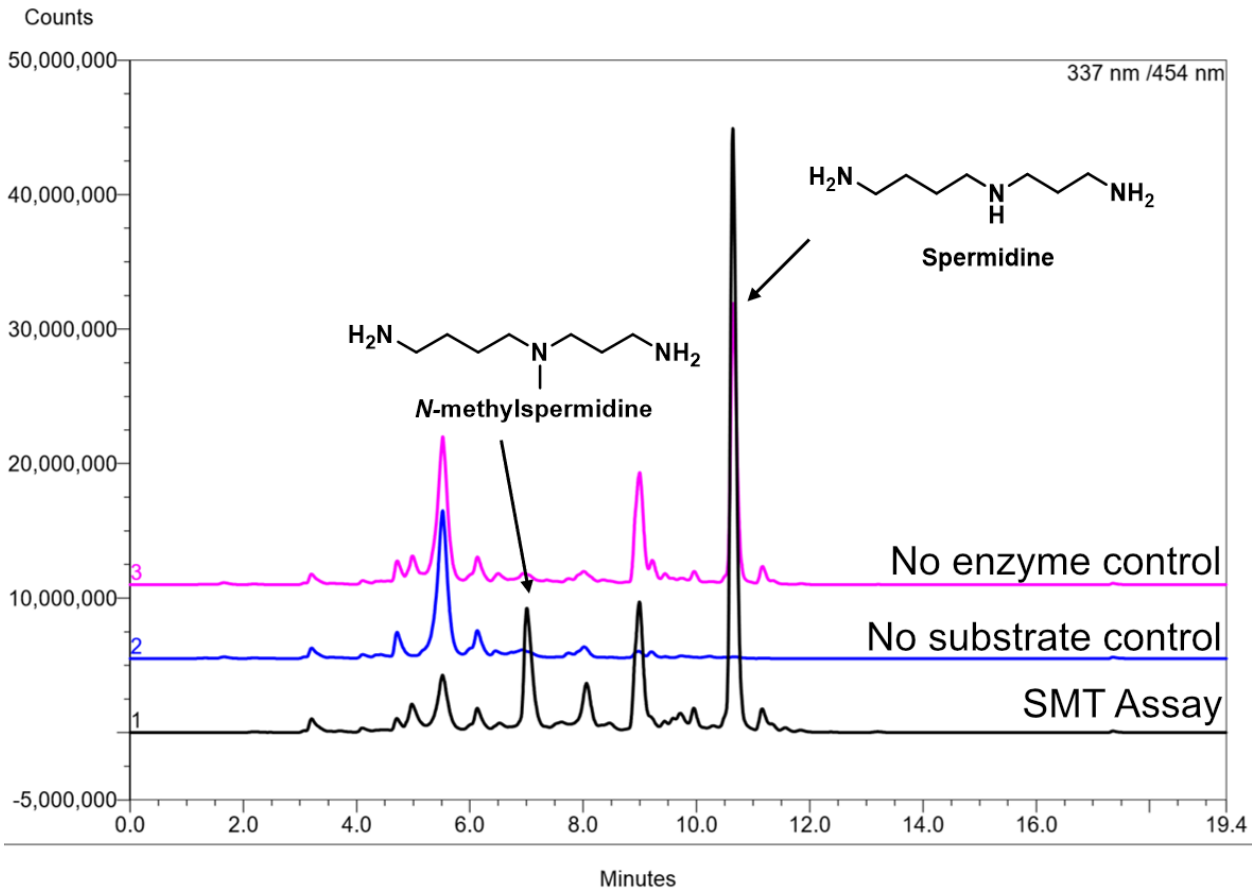


Figure S3. **L1 leaves crude SMT activity assay.** HPLC-FLD chromatogram with crude protein extracts from *E. coca* L1 leaves have SMT activity. Before injection the assay was derivatized with OPA and FMOC following the procedure reported in Hanczkó et. al. SMT assay (black), no enzyme control (blue), no substrate control (pink).

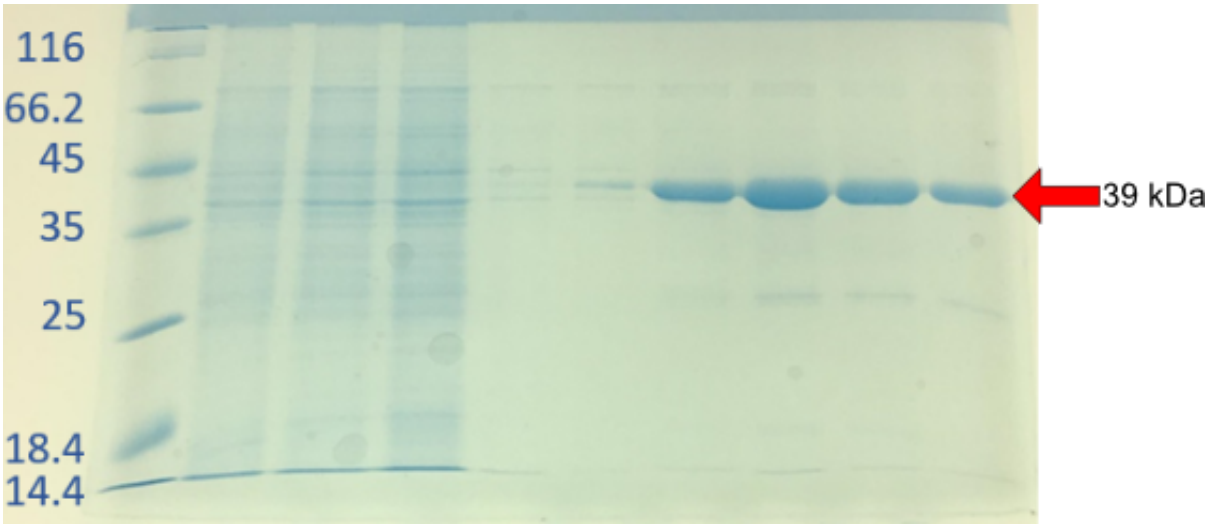


Figure S4. SDS-PAGE gel of *Ec*SMT (formerly known as *Ec*SPDS1) pEPStrep (Molecular Weight: 39 kDa). Order of lanes: MW (KDa) ladder, crude, wash, flow-through, fraction 19, 20, 21, 22, 23, 24. Fractions of EcSPDS1 pEPStrepGW are collected after a Strep-tag affinity purification. The purified protein is depicted using a red arrow along with the molecular weight in kDa.

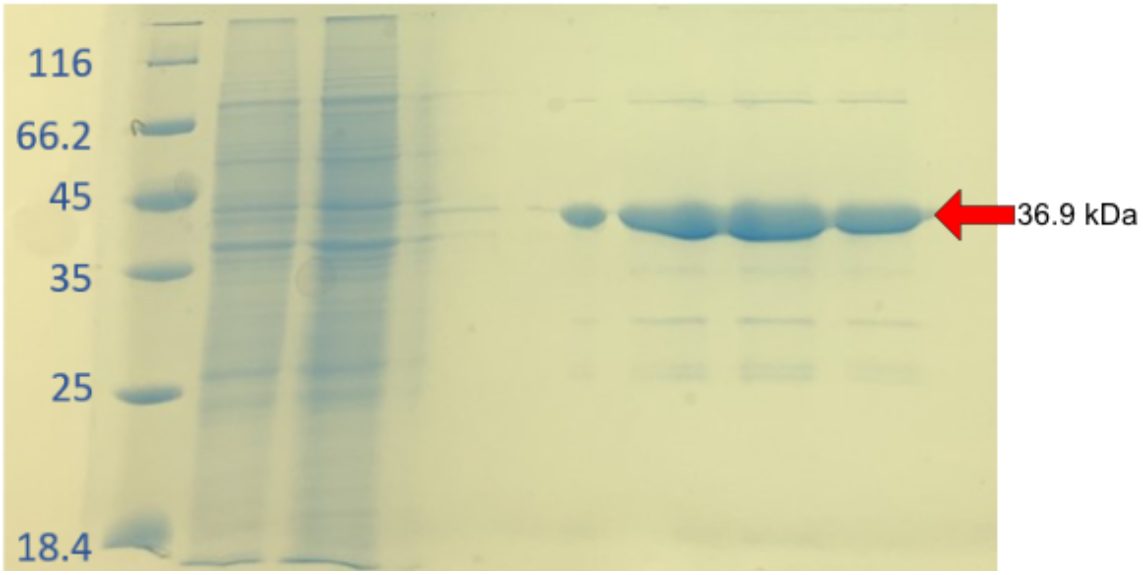


Figure S5. SDS-PAGE gel of *Ec*SPMT (formerly known as *Ec*SPDS2) pEPStrep (Molecular Weight: 36.9 kDa). Order of lanes: MW (kDa) ladder, crude, flow-through, wash, fraction 23, 24, 25, 26, 27. Fractions of EcSPMT pEPStrepGW are collected after a strep-tag affinity purification. The purified protein is depicted using a red arrow along with the molecular weight in kDa.

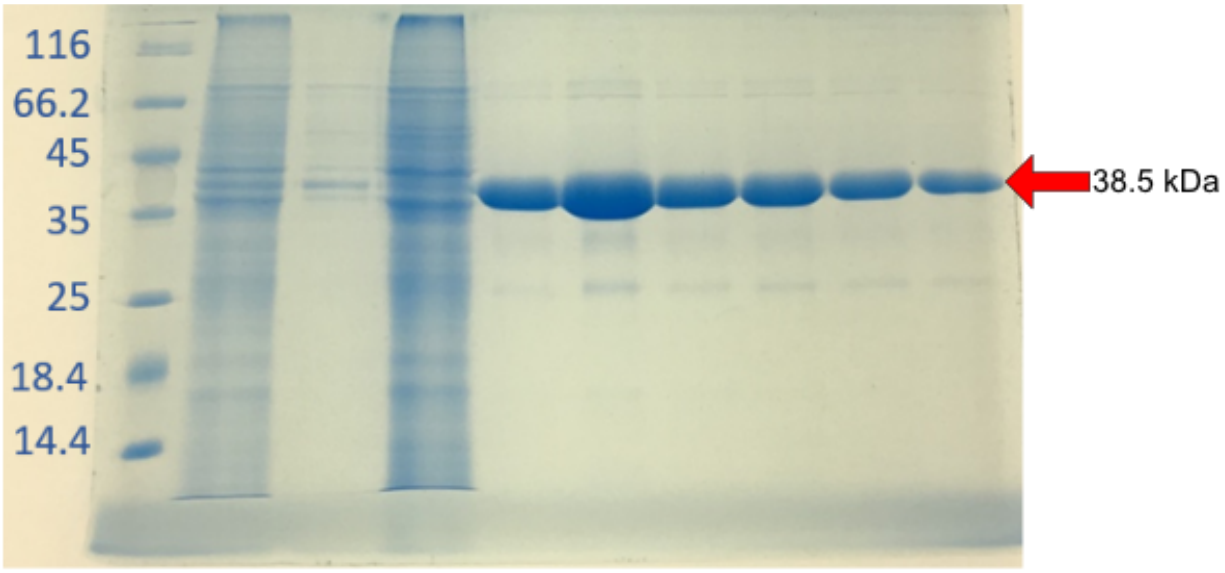


Figure S6. SDS-PAGE gel of *Ec*SPDS (formerly known as *Ec*SPDS3) pEPStrep (Molecular Weight: 38.5 kDa). Order of lanes: MW (kDa) ladder, Crude, flow-through, wash, fraction 22, 23, 24, 25, 26, 27. Fractions of EcSPDS pEPStrepGW are collected after a strep-tag affinity purification. The purified protein is depicted using a red arrow along with the molecular weight in kDa.

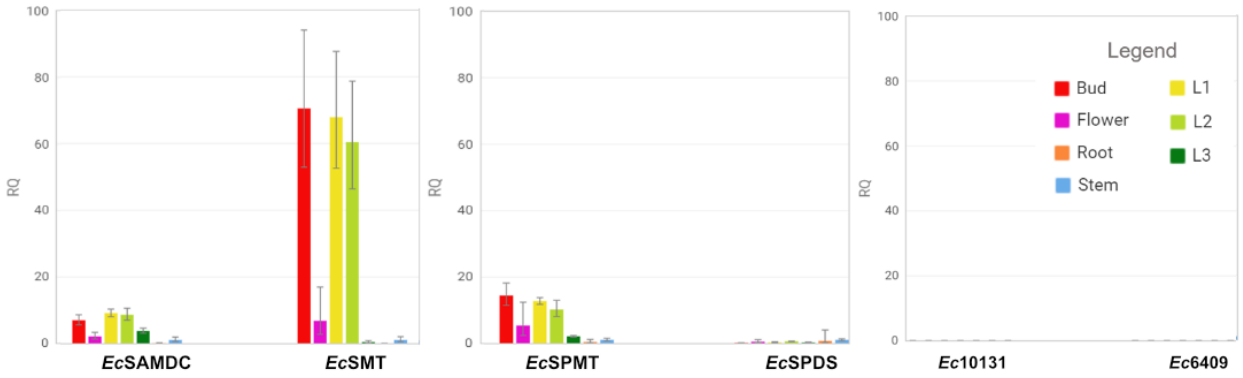


Figure S7. Quantitative PCR analysis of SPDS-like candidate genes within *E. coca* tissues. Relative quantification of transcript profiles of EcSAMDC, EcSMT, EcSPMT and EcSPDS as well as quantification of endogenous controls Ec10131 and Ec6409. There was a total of three biological replicates, each with three technical replicates. A comparative Ct (ΔΔCt) experiment was performed to generate this figure. Raw data was imported into the Applied Biosystems™ analysis software, relative quantification analysis module, version 3.9. Ec10131 and Ec6409 were designated as endogenous controls and 'stem' was assigned as the reference sample. Primer efficiency values were assigned to each target. P values greater than 0.05 were determined to have no significance. Error bars represent the standard deviation.

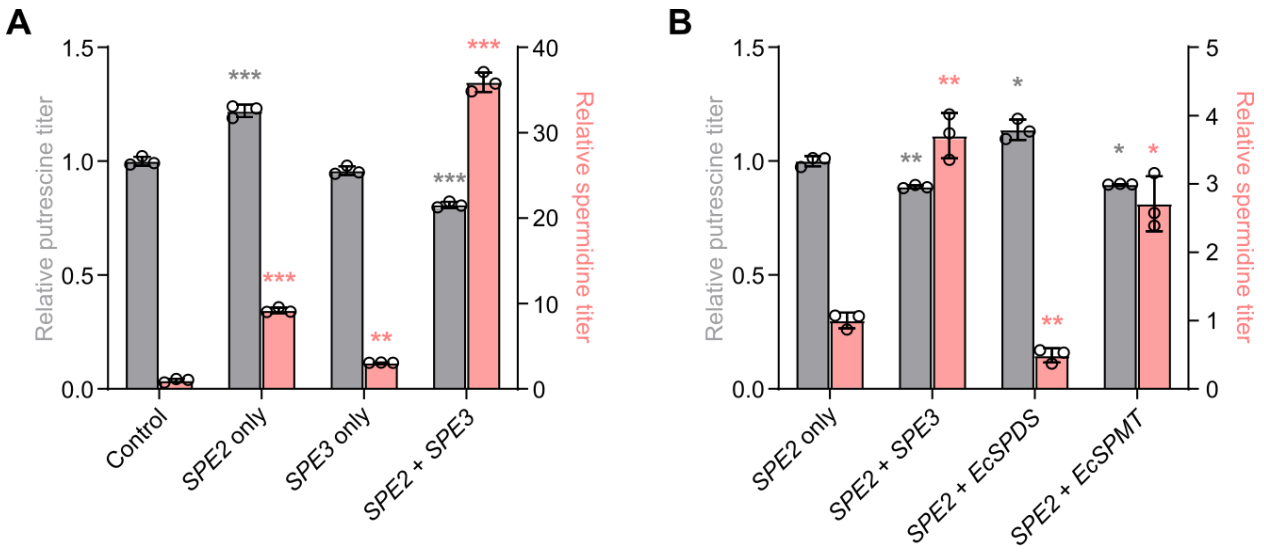


Figure S8. Evaluation of (A) yeast and (B) *E. coca* spermidine synthase activities in engineered yeast. For both (A) and (B), spermidine synthases or a BFP control were expressed from low-copy plasmids in CSY1242, and metabolites in supernatant were quantified by LC–MS/MS after 72-h growth in selective media. Data represent the mean of $n = 3$ biologically independent samples (open circles), and error bars show the SD. Student's two-tailed t test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

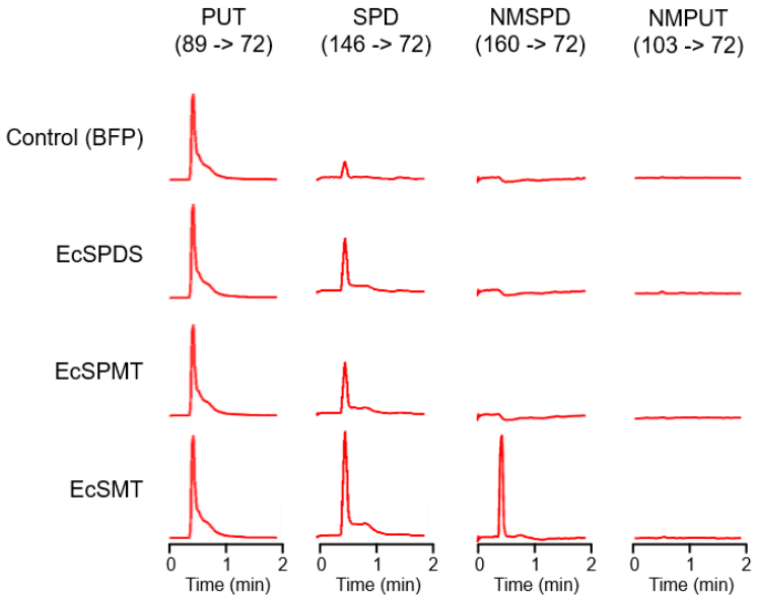


Figure S9. Characterization of *E. coca* spermidine *N*-methyltransferase activities in engineered yeast. *N*-methyltransferase candidates or a BFP control were expressed from low-copy plasmids in CSY1340, and metabolites in supernatant were analyzed by LC–MS/MS after 72-h growth in selective media. MS/MS multiple reaction monitoring (MRM) transitions for putrescine (PUT), spermidine (SPD), *N*-methylspermidine (NMSPD), and *N*-methylputrescine (NMPUT) are indicated in parentheses below the compound names. Chromatogram traces are representative of $n = 3$ biologically independent samples and peak heights are shown to scale with respect to all other peaks for the same compound.

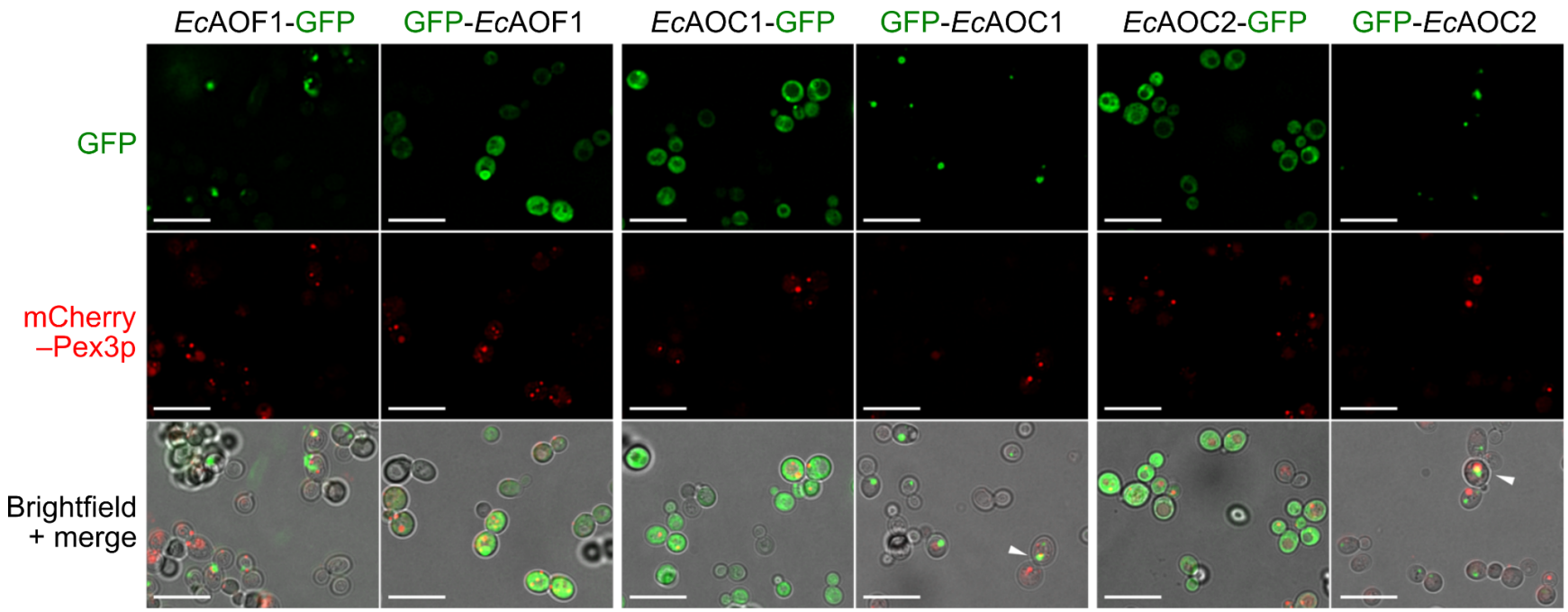
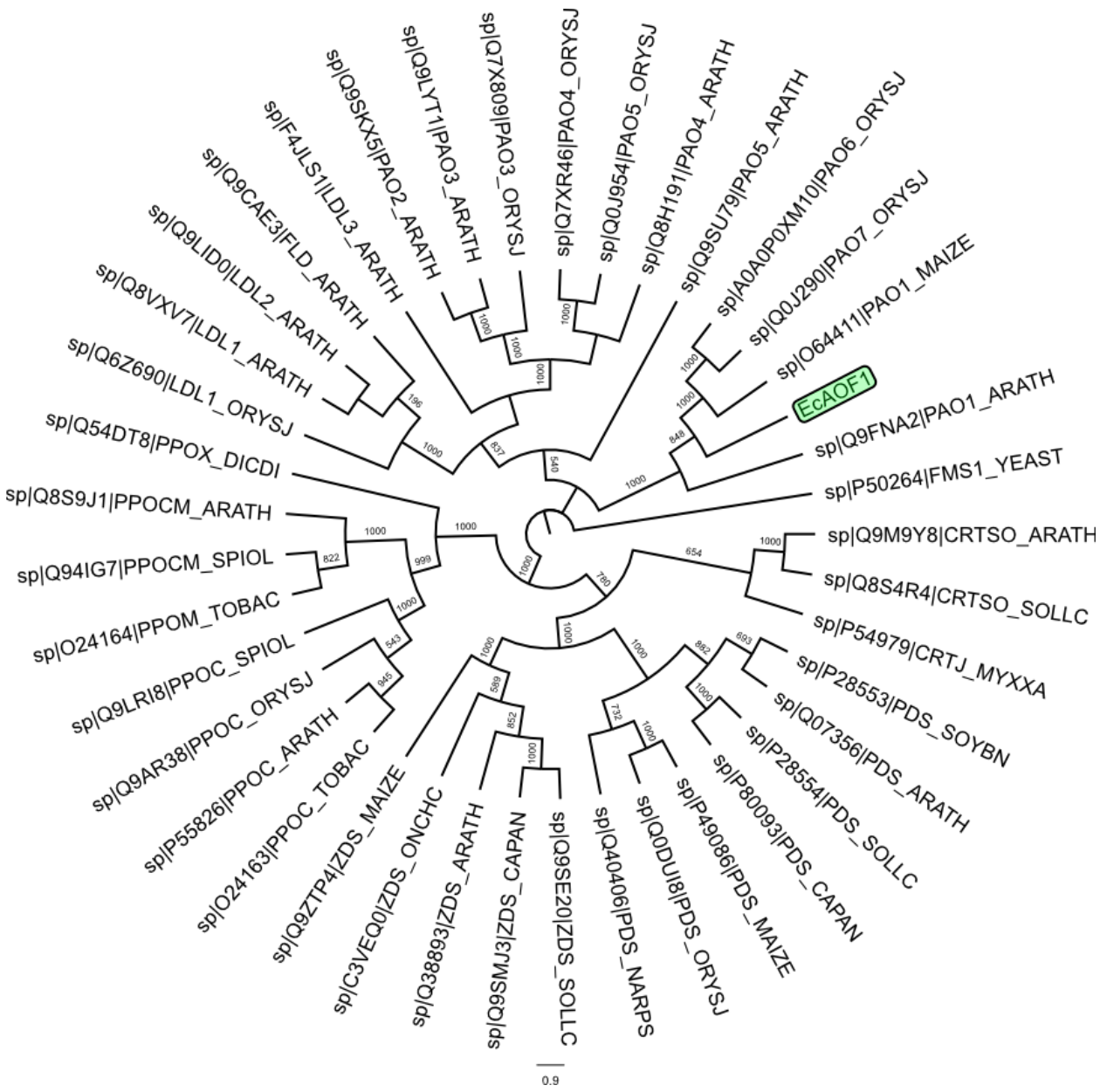


Figure S10. Identification of *E. coca* amine oxidase subcellular localization in yeast via fluorescence microscopy. N- and C-terminal GFP fusions of *EcAOF1*, *EcAOC1*, and *EcAOC2* were co-expressed with a yeast peroxisomal localization marker gene (mCherry-*PEX3*) from low-copy plasmids in wild-type yeast (CEN.PK2). Microscopy and image analysis were performed as described in SI Materials and Methods. Scale bar, 10 µm.

A


sp|Q7XR46|PAO4_ORYSJ
sp|Q0J954|PAO5_ORYSJ
sp|Q8H191|PAO4_ARATH
sp|Q9SU79|PAO5_ARATH
sp|A0A0P0XM10|PAO6_ORYSJ
sp|Q0J290|PAO7_ORYSJ
sp|O64411|PAO1_MAIZE
EcAOF1
sp|Q9FNA2|PAO1_ARATH
sp|P50264|FMS1_YEAST
sp|Q9M9Y8|CRTSO_ARATH
sp|Q8S4R4|CRTSO_SOLLC
sp|P54979|CRTJ_MYXXA
sp|P28553|PDS_SOYBN
sp|Q07356|PDS_ARATH
sp|P28554|PDS_SOLLC
sp|P80093|PDS_CAPAN
sp|P49086|PDS_MAIZE
sp|Q0DUI8|PDS_ORYSJ
sp|Q40406|PDS_NARPS
sp|Q9SE20|ZDS_SOLLC
sp|Q9SMJ3|ZDS_CAPAN
sp|Q38893|ZDS_ARATH
sp|C3VEQ0|ZDS_ONCHC
sp|Q9ZTP4|ZDS_MAIZE
sp|O24163|PPOC_TOBAC
sp|P55826|PPOC_ARATH
sp|Q9AR38|PPOC_ORYSJ
sp|Q9LRI8|PPOC_SPIOL
sp|O24164|PPOM_TOBAC
sp|Q94IG7|PPOCM_SPIOL
sp|Q8S9J1|PPOCM_ARATH
sp|Q54DT8|PPOX_DICDI
sp|Q6Z690|LDL1_ORYSJ
sp|Q8VXV7|LDL1_ARATH
sp|Q9LID0|LDL2_ARATH
sp|Q9CAE3|FLD_ARATH
sp|F4JLS1|LDL3_ARATH
sp|Q9SKX5|PAO2_ARATH
sp|Q9LYT1|PAO3_ARATH
sp|Q7XX80|PAO3_ORYSJ
0.9

B

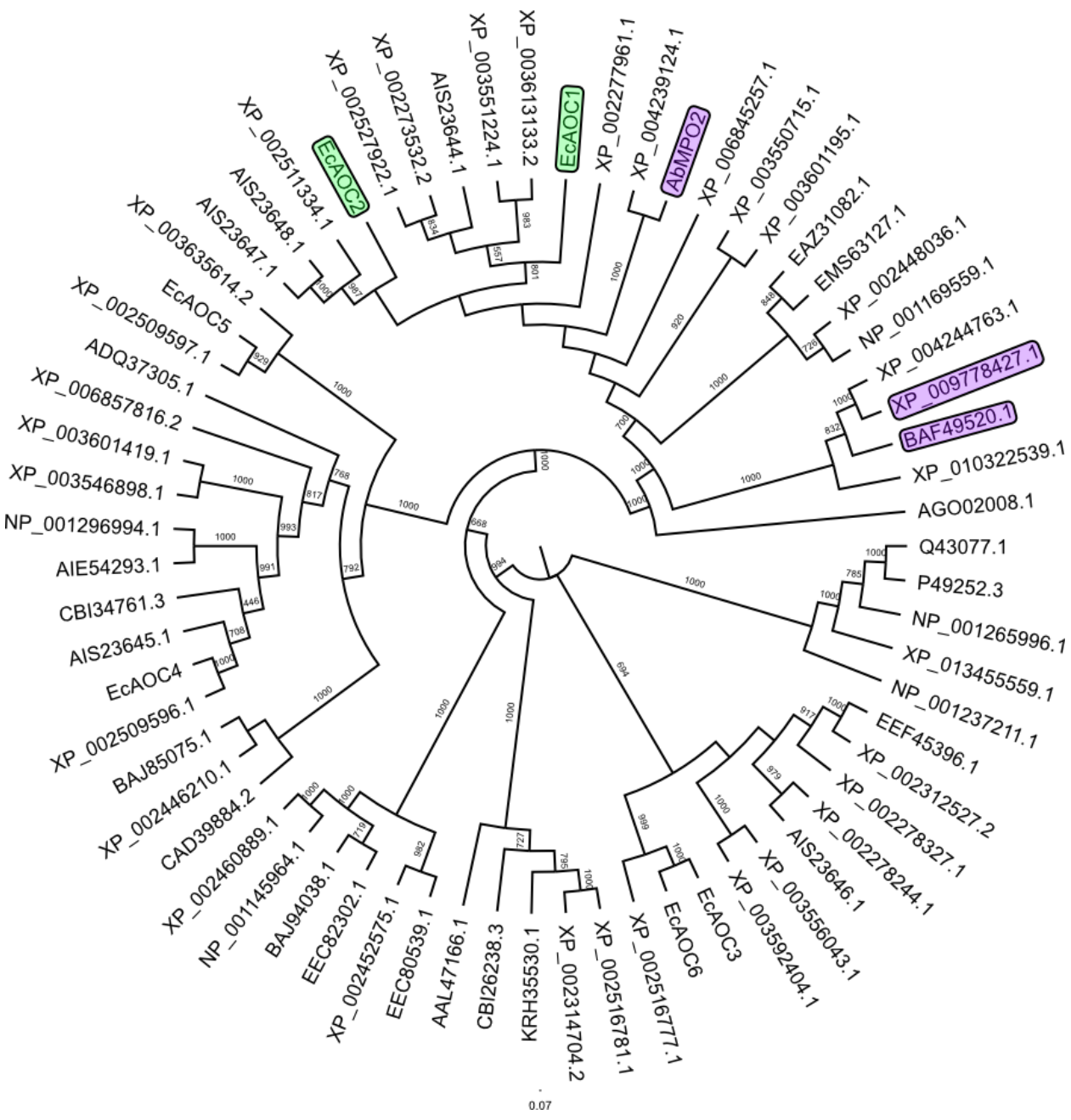


Figure S11. Phylogenetic analysis of novel polyamine oxidases from *E. coca*. (A) Phylogenetic tree showing evolutionary sequence relationships between *Ec*AOF1 and other flavin-dependent polyamine oxidases. *E. coca Ec*AOF1 identified in this study is highlighted in green. (B) Phylogenetic tree showing evolutionary sequence relationships between *Ec*AOC1 and *Ec*AOC2 in comparison with other copper-dependent polyamine oxidases. *E. coca Ec*AOC1 and *Ec*AOC2 identified in this study are highlighted in green and solanaceous copper dependent oxidases highlighted in magenta. Refer to Dataset S1 for identities and accession numbers of all enzymes in the dendrogram. Values at each junction signify the number of bootstraps out of *n* = 1,000 iterations. Scale bar, number of substitutions per amino acid site. Note that bootstrap values below 500 were not reported.

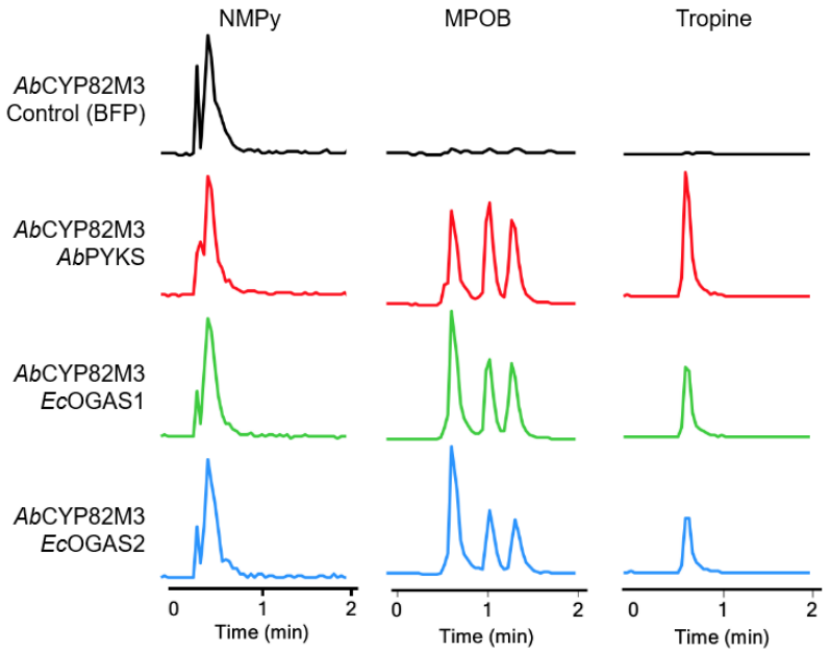


Figure S12. Evaluation of Solanaceae PYKS and Erythroxylaceae OGAS activities in engineered yeast. *A. belladonna* PYKS, each *E. coca* OGAS, or a BFP control were co-expressed with *A. belladonna* tropinone synthase (*Ab*CYP82M3) from low-copy plasmids in CSY1246, and metabolites in supernatant were analyzed by LC–MS/MS after 72-h growth in selective media. Chromatogram traces are representative of $n = 3$ biologically independent samples and peak heights are shown to scale with respect to all other peaks for the same compound.

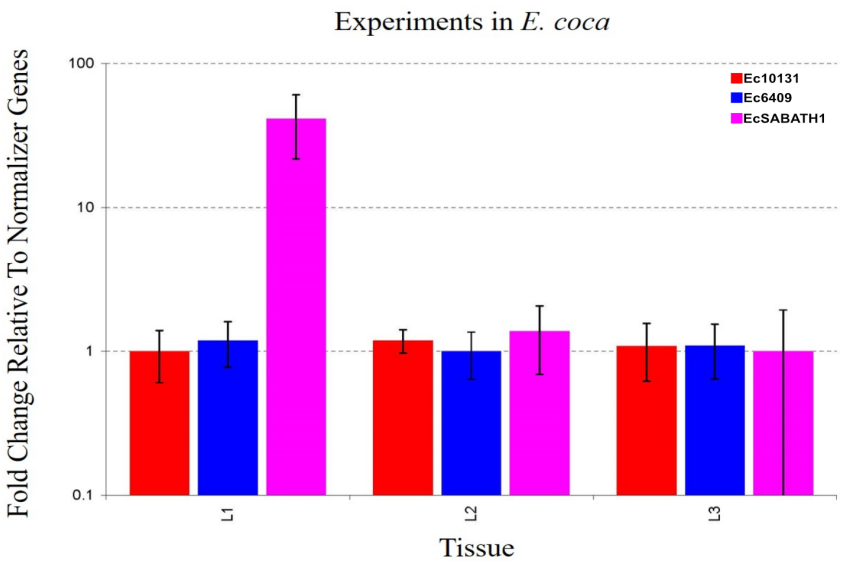


Figure S13. Relative Quantification of EcSABATH1 (EcMPOBMT) in *E. coca* leaf tissues. To determine significance among Ct values obtained from qPCR analysis, t-tests were performed for all comparisons made. P values greater than .05 were determined to have no significance. Significant differences were determined for EcSABATH1 in *E. coca* between tissues L1 and L2 and between tissues L1 and L3. The relative expression levels of EcSABATH1 in *E. coca* were about 50 times greater in L1 than in L2 and L3. No significant differences in relative expression were determined between L2 tissue and L3 tissue.

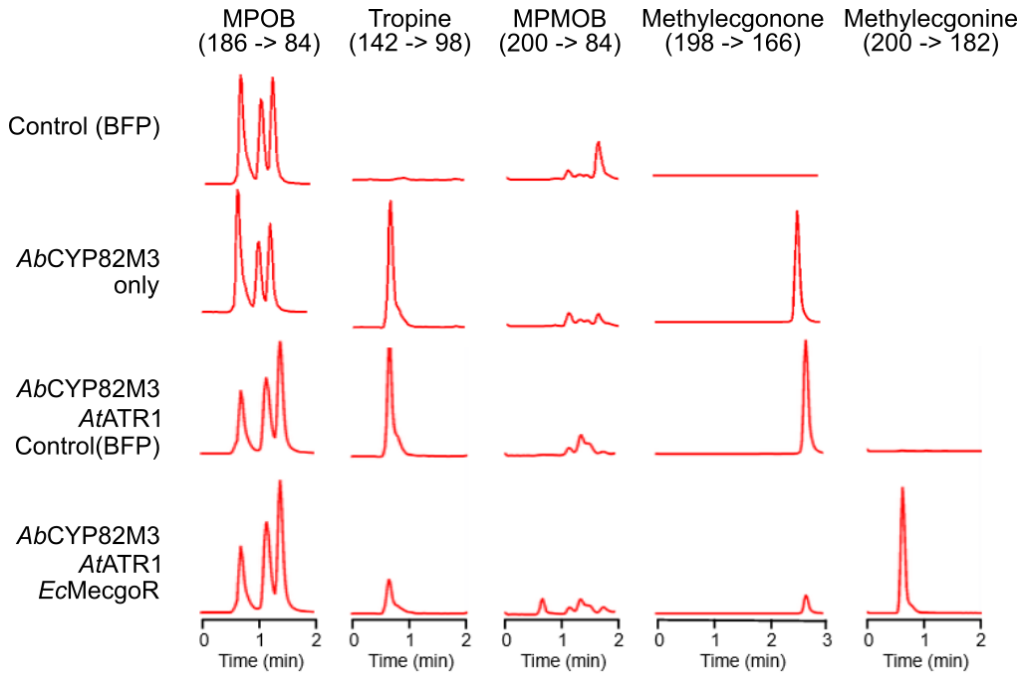


Figure S14. Evaluation of enzyme promiscuity for Solanaceae tropinone synthase and tropinone reductase. *A. belladonna* tropinone synthase (*Ab*CYP82M3) was expressed alone or co-expressed with *A. thaliana* NADPH:cytochrome P450 reductase (*At*ATR1) and *E. coca* methylecgonone reductase (*Ec*MecgoR) or a BFP control from low-copy plasmids in CSY1341 (a strain which also possesses tropinone reductase activity via *D. stramonium* tropinone reductase I, *Ds*TR1). Metabolites in supernatant were analyzed by LC–MS/MS after 72-h growth in selective media. Chromatogram traces are representative of n = 3 biologically independent samples and peak heights are shown to scale with respect to all other peaks for the same compound.

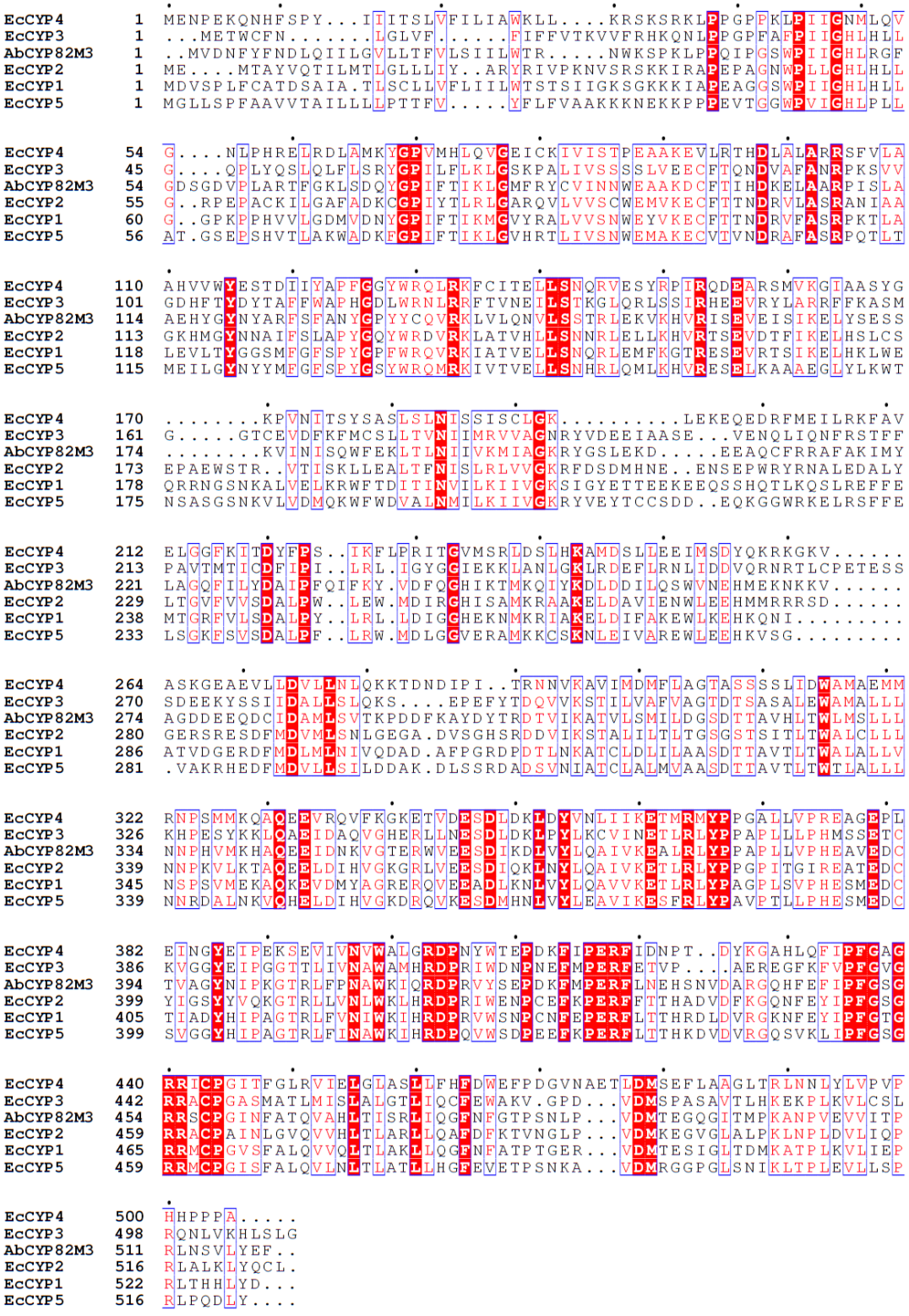


Figure S15. Sequence alignment of *E. coca* CYP82M3-like sequences (*Ec*CYP1-5) identified via BLAST search. The sequence of CYP82M3 from *A. belladonna* (*Ab*CYP82M3) is shown as a reference. Solid red columns indicate perfect residue identity; red text columns indicate residue similarity. Contiguous blocks or identical or similar residues are outlined in blue.

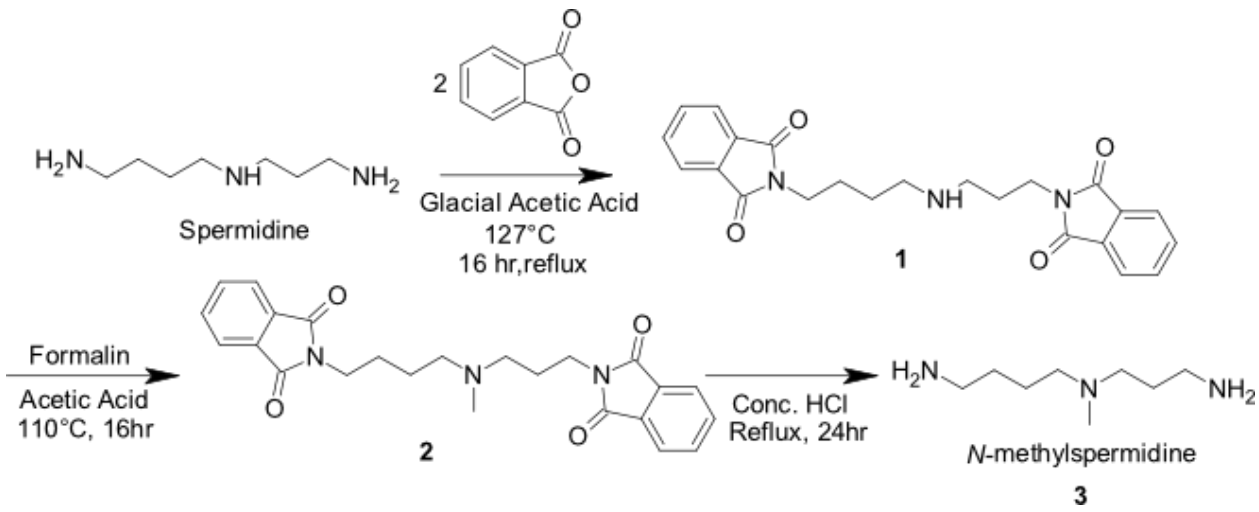


Figure S16. Scheme for synthesizing *N*-methylspermidine from spermidine.

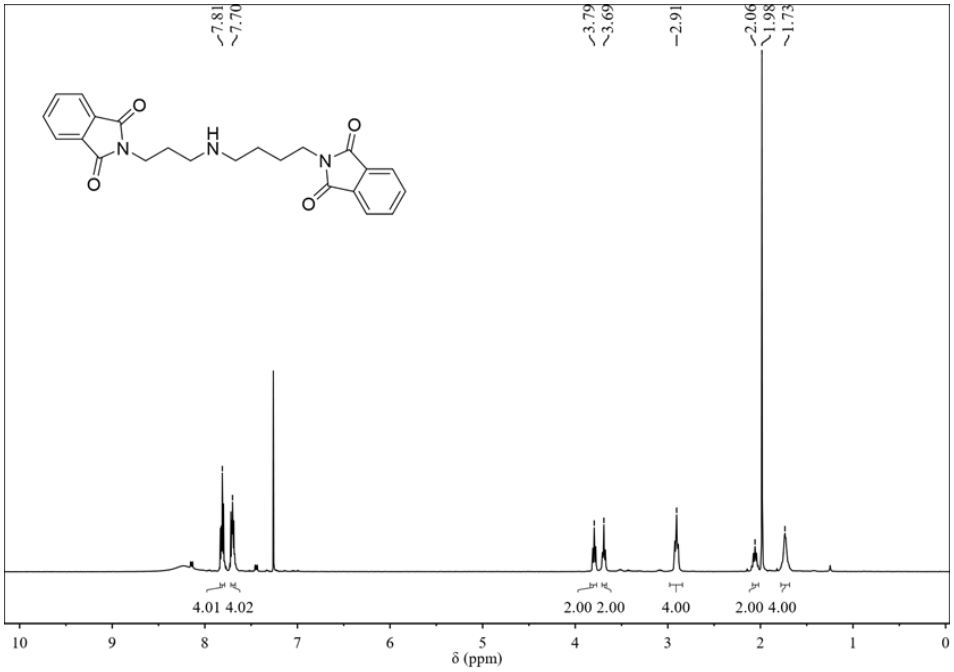


Figure S17. ^{1}H-NMR (ClCD3) of 2-(3-((4-(1,3-dioxoisoindolin-2-yl) butyl) amino) propyl) isoindoline-1,3-dione (compound 1) with traces of acetic acid.

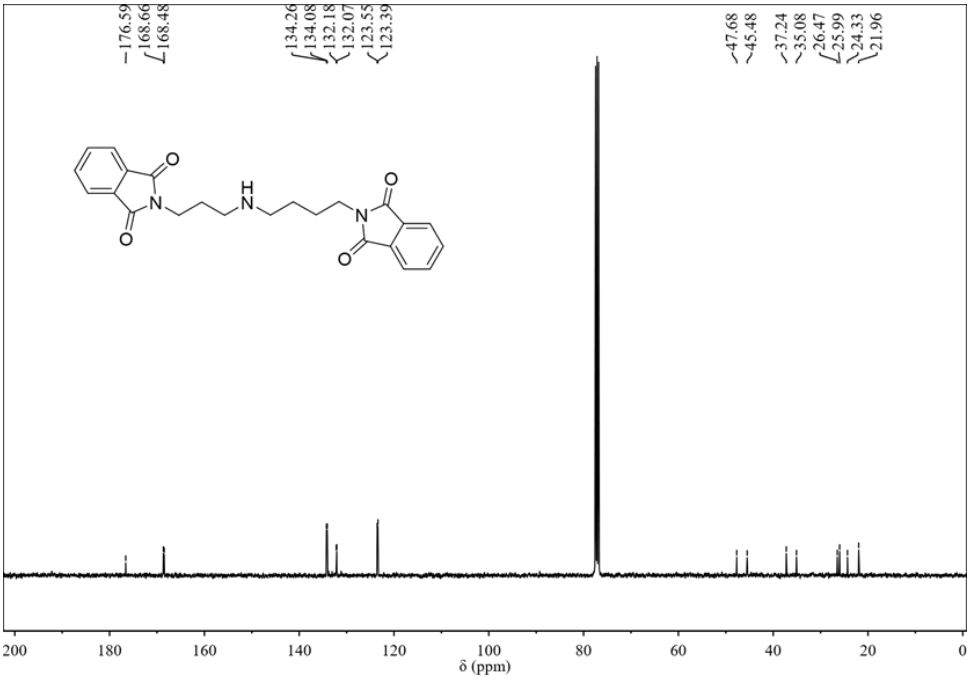


Figure S18. ^{13}C-NMR (ClCD3) of 2-(3-((4-(1,3-dioxoisoindolin-2-yl) butyl) amino) propyl) isoindoline-1,3-dione (compound 1) with traces of acetic acid.

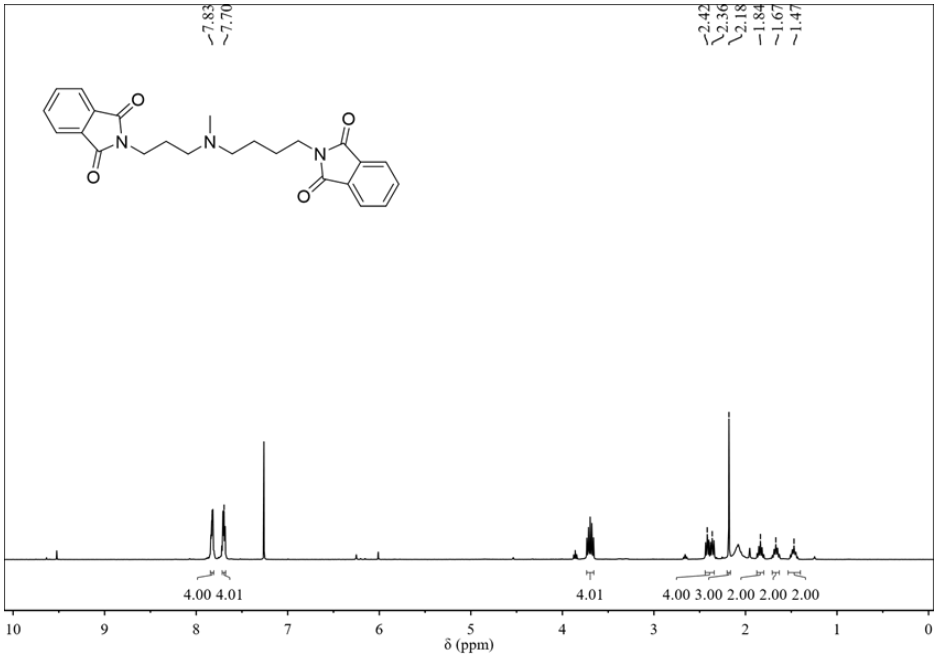


Figure S19. ^{1}H-NMR (ClCD3) of 2-(3-((4-(1,3-dioxoisoindolin-2-yl)butyl)(methyl)amino)propyl)isoindoline-1,3-dione (compound 2).

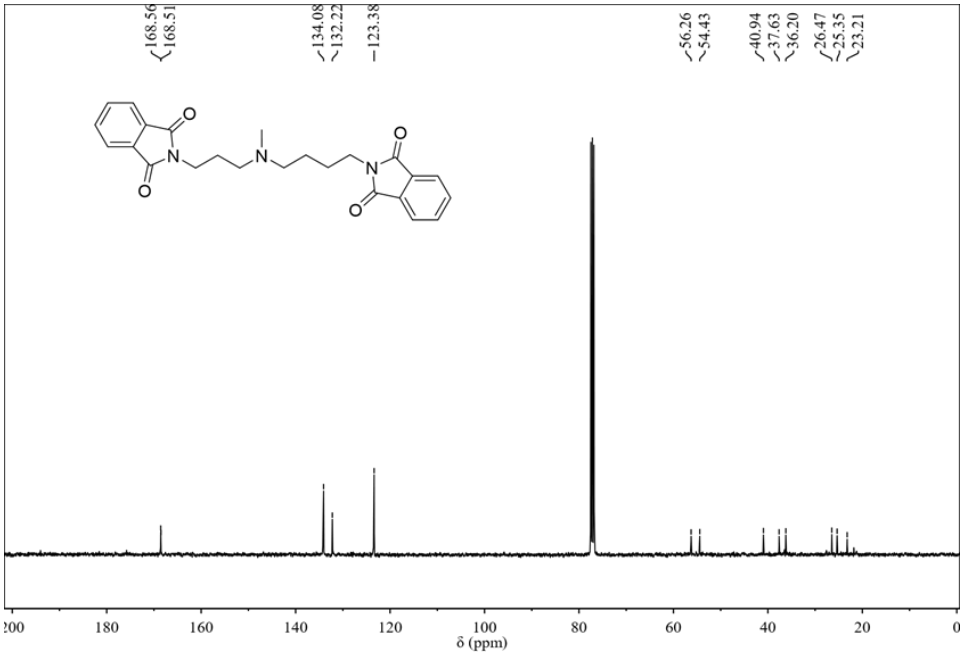


Figure S20. ^{13}C-NMR (ClCD3) of 2-(3-((4-(1,3-dioxoisoindolin-2-yl)butyl)(methyl)amino)propyl)isoindoline-1,3-dione (compound 2).

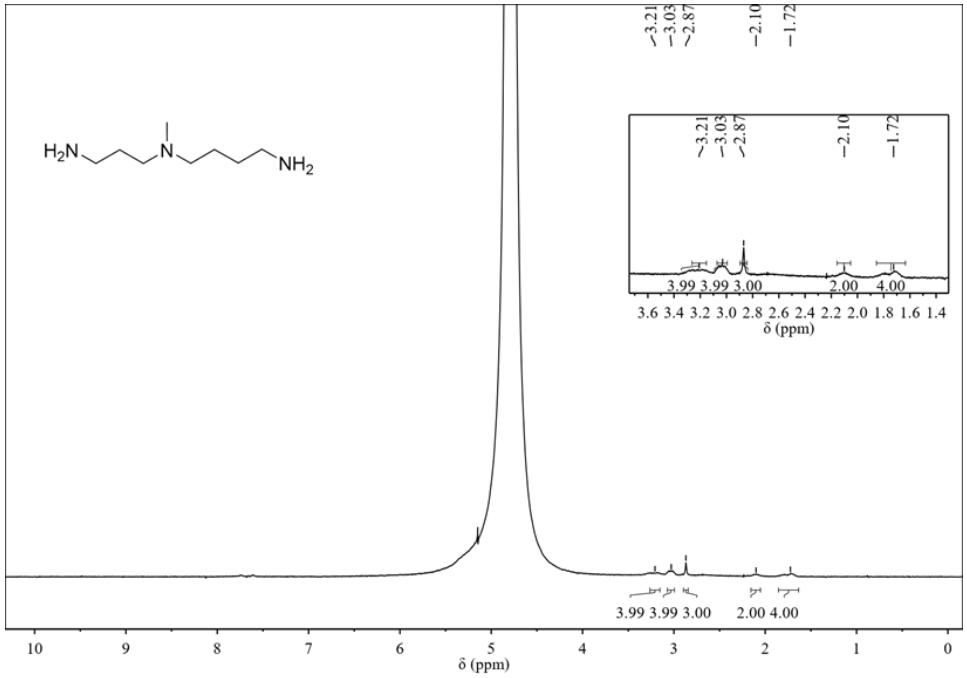


Figure S21. ^{1}H-NMR (D2O) of *N*-methylspermidine (compound 3).

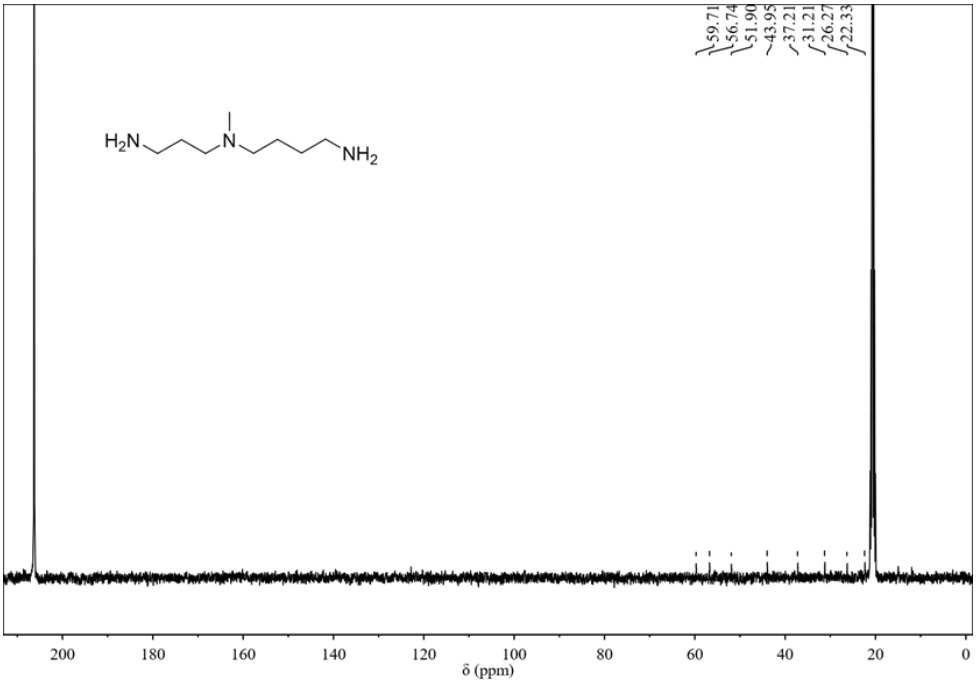


Figure S22. ^{13}C-NMR (D2O spiked with acetone-d6) of *N*-methylspermidine (compound 3).

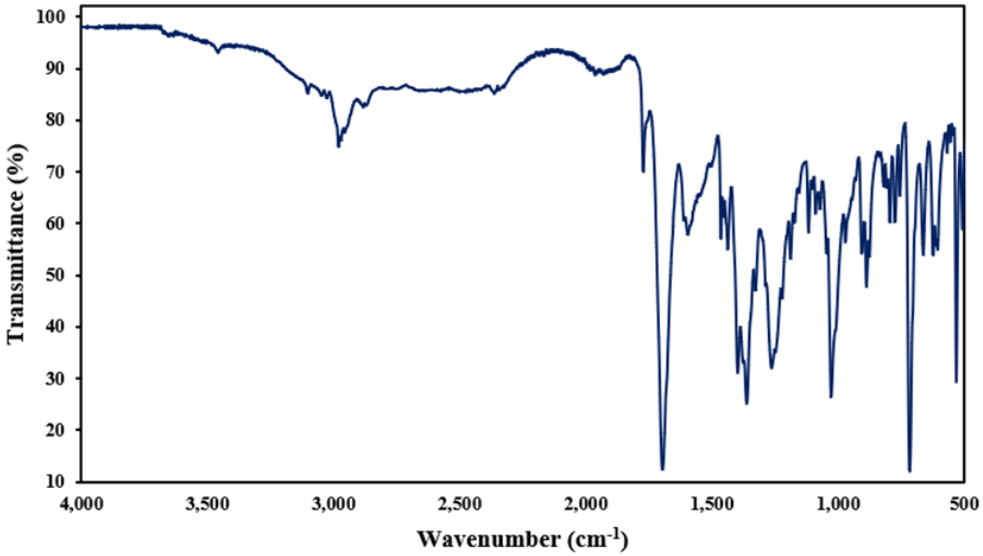


Figure S23. Di-ATR FTIR of 2-(3-((4-(1,3-dioxoisoindolin-2-yl) butyl) amino) propyl) isoindoline-1,3-dione (compound 1).

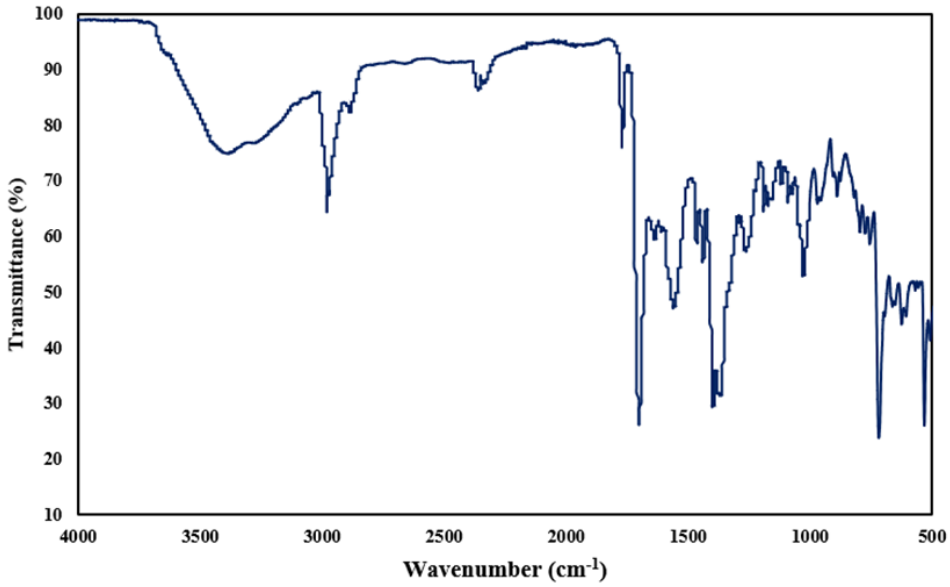


Figure S24. Di-ATR FTIR of 2-(3-((4-(1,3-dioxoisoindolin-2-yl)butyl)(methyl)amino)propyl)isoindoline-1,3-dione (compound 2).

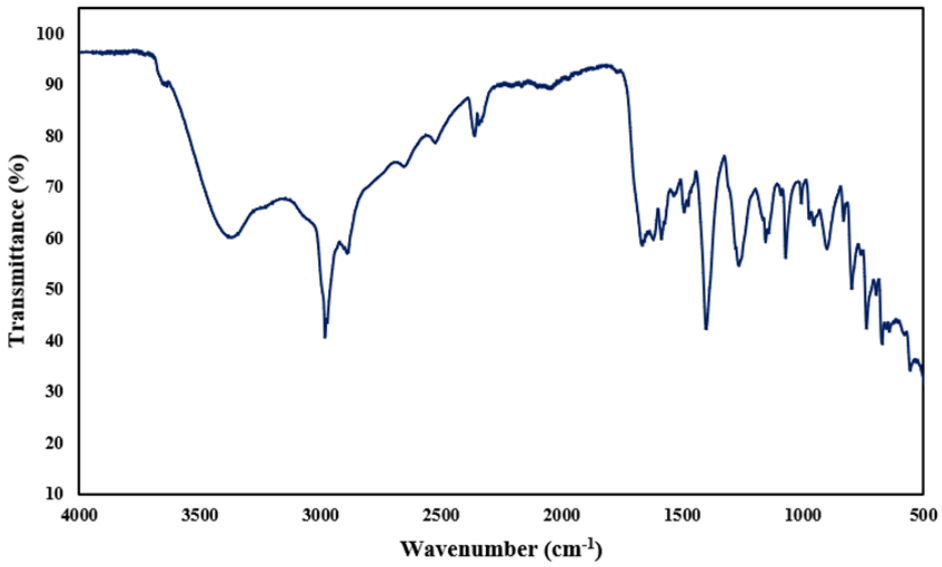


Figure S25. Di-ATR FTIR of *N*-methylspermidine (compound 3).

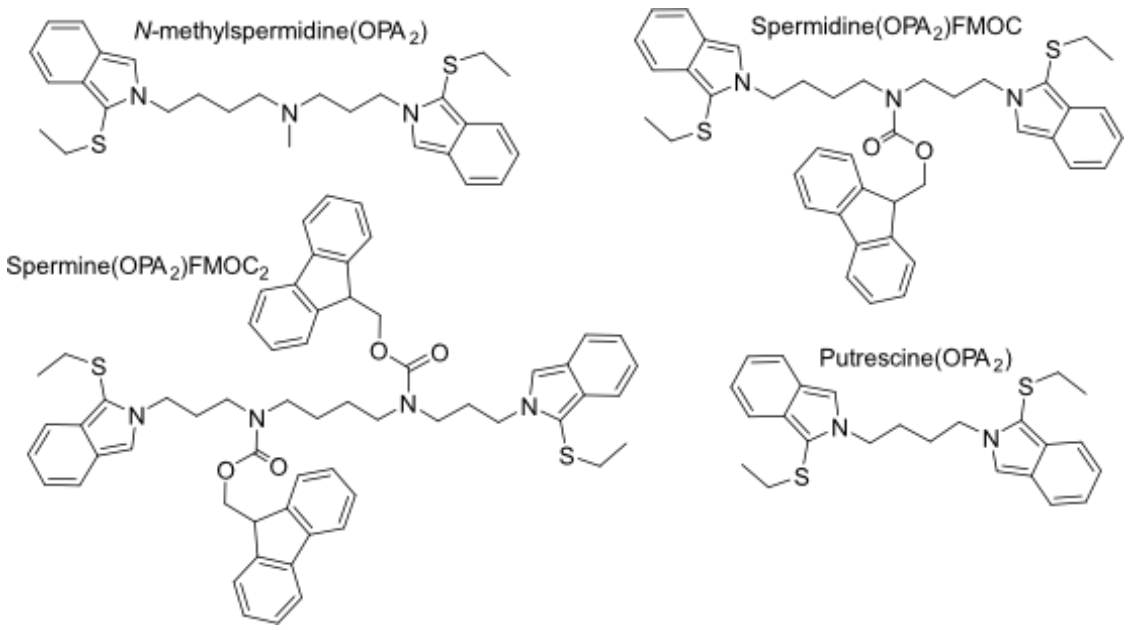


Figure S26. Structure of polyamines with OPA-Et and FMOC derivatizing agents.

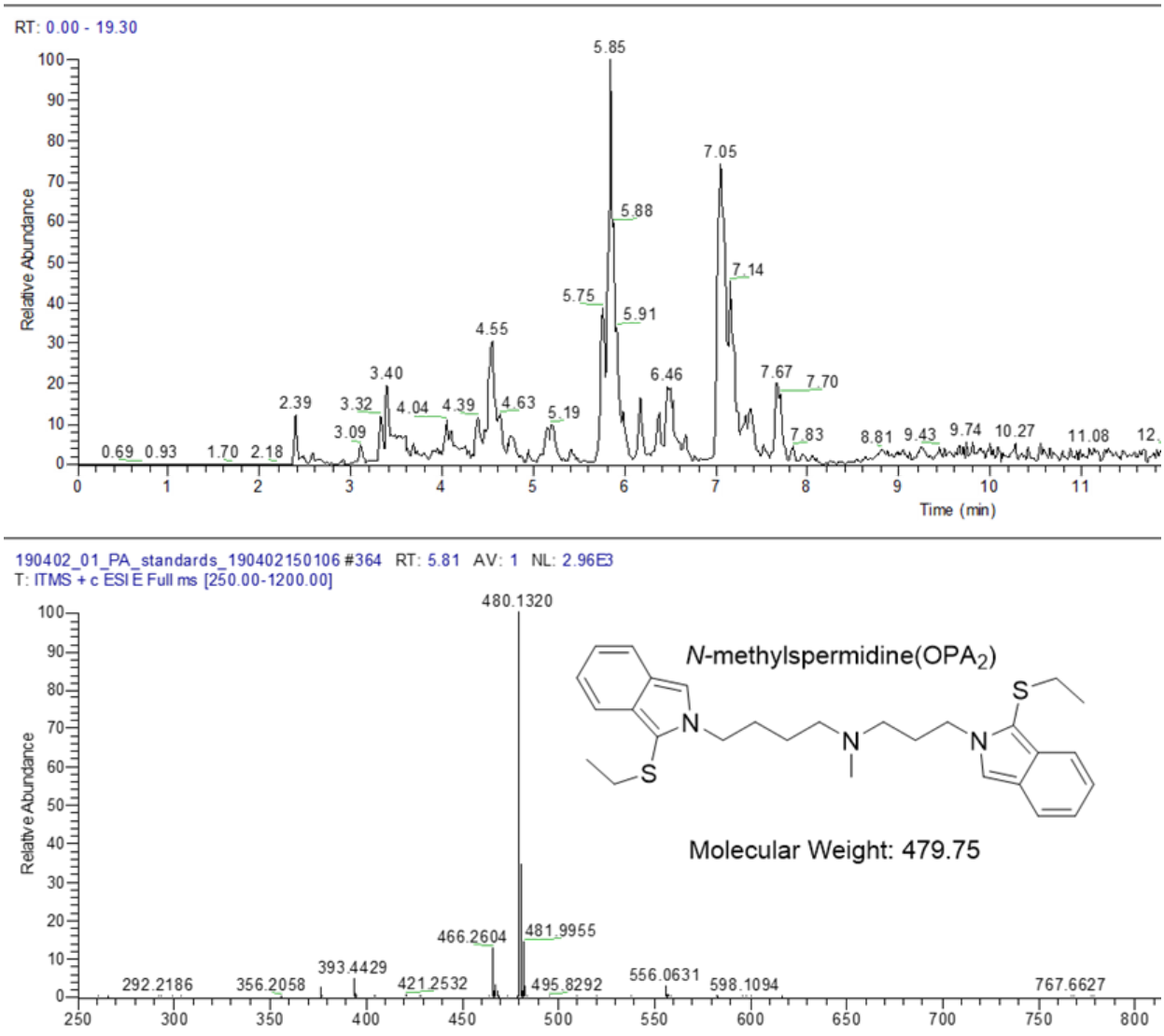


Figure S27. HPLC-MS TIC of derivatized *N*-methylspermidine. Shows N-methylspermidine(OPA2) has a retention time of approximately 5.81 minutes and a m/z ratio of 480.13. The molecular weight of N-methylspermidine(OPA2) is 479.75 g/mole.

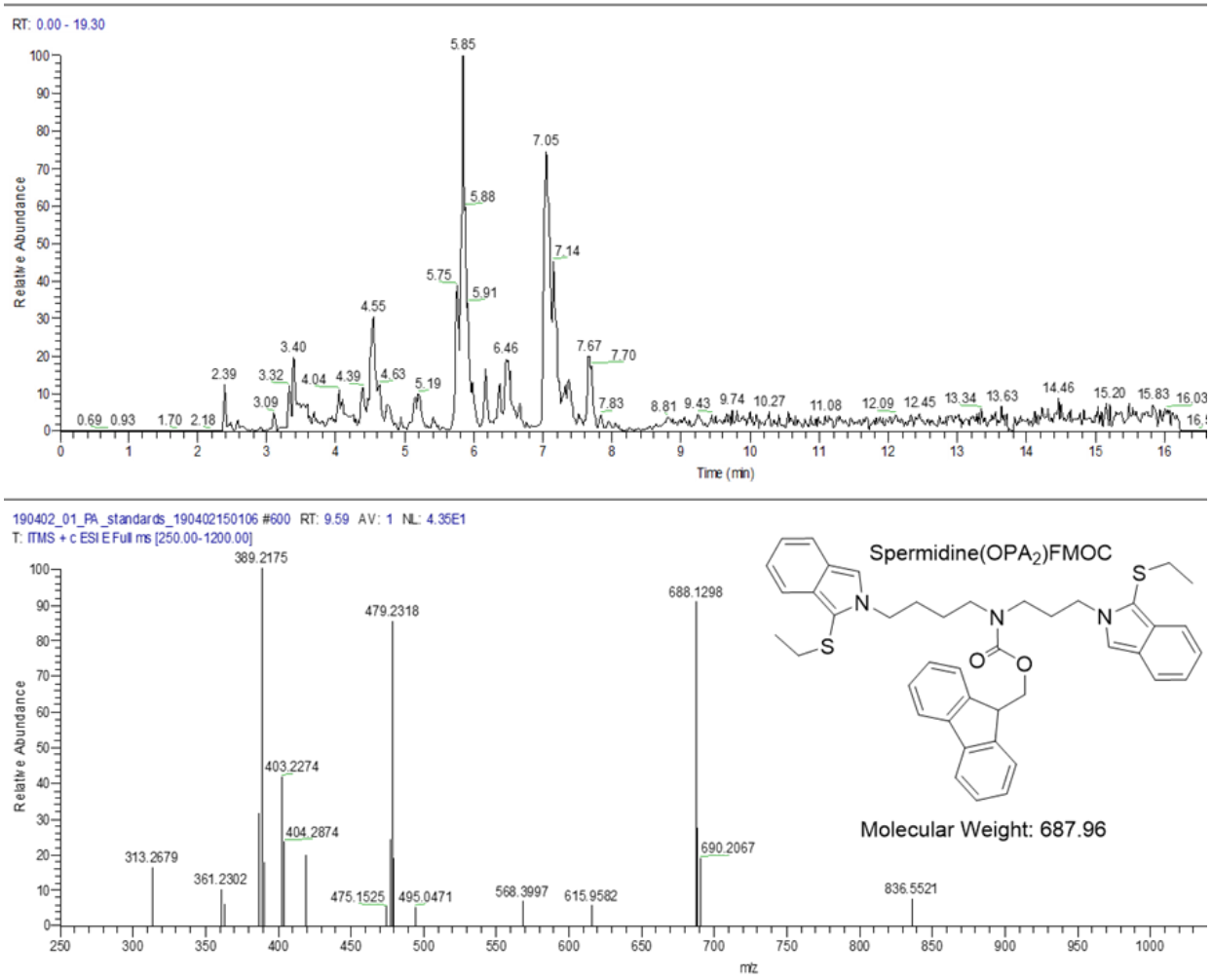


Figure S28. HPLC-MS TIC of derivatized spermidine. Shows spermidine(OPA2)FMOC has a retention time of approximately 9.59 minutes and a m/z ratio of 688.13. The molecular weight of spermidine(OPA2)FMOC is 687.96 g/mole.

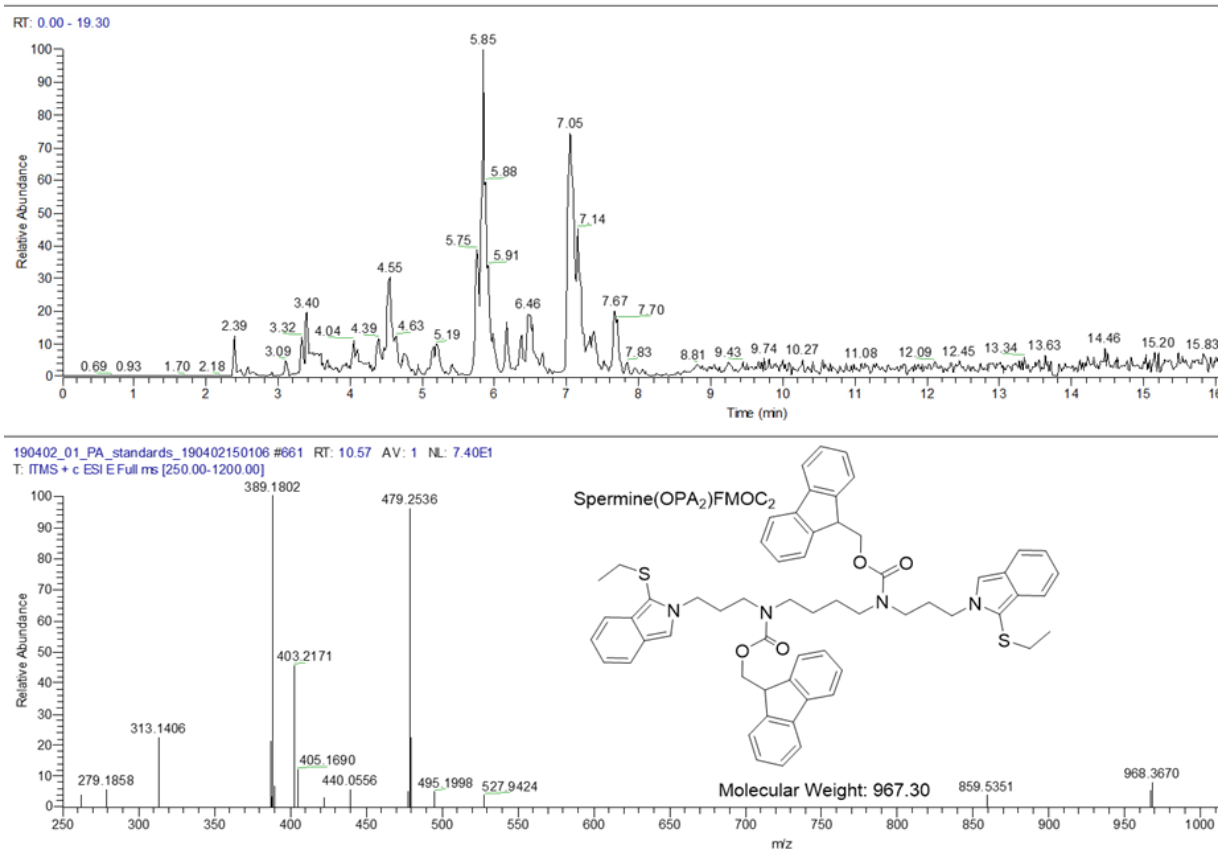


Figure S29. HPLC-MS TIC of derivatized spermine. Shows spermine(OPA2)FMOC has a retention time of approximately 10.57 minutes and a m/z ratio of 968.37. The molecular weight of spermine(OPA2)FMOC is 967.30 g/mole.

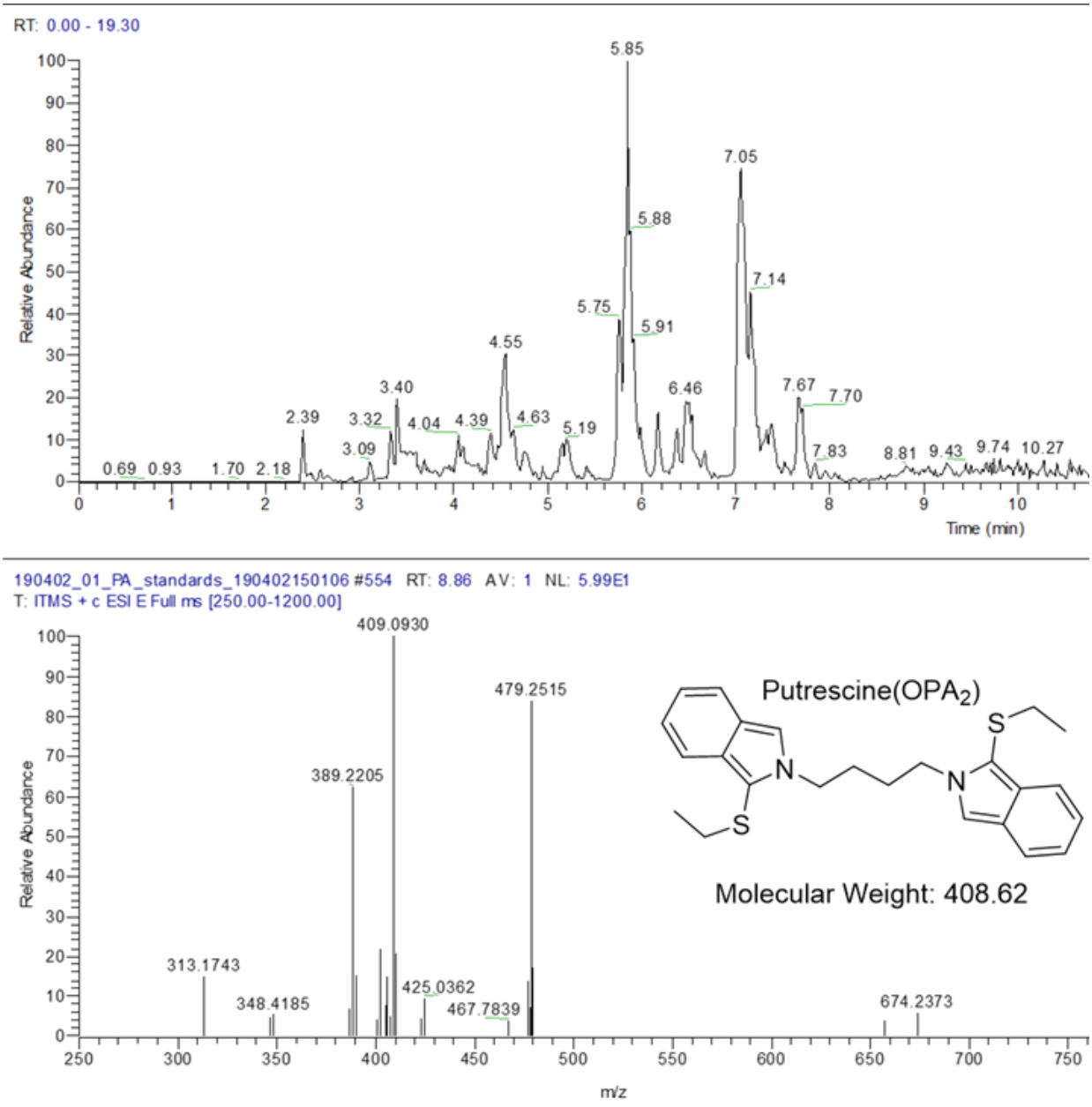


Figure S30. HPLC-MS TIC of derivatized putrescine. Shows putrescine(OPA2) has a retention time of approximately 8.86 minutes and a m/z ratio of 409.09. The molecular weight of putrescine(OPA2) is 408.62 g/mole.

Supplementary Tables

Table S1. Sequences and performance parameters of primers used for qPCR analysis of *E. coca* genes of interest. Primer efficiency values in E. coca. Primer efficiency was calculated from the slope of a standard curve generated from a series of seven, two-fold dilutions of *E. coca* L3 cDNA.

Gene of Interest	Primer Sequence (5'→3')	Efficiency	R^2	SD
Ec10131	Fwd: TGGAAGGGTAGTGGGGTAACAATG	100.74%	0.98	0.18
	Rev: GAGCGTAGTCGTCAGAGAAGGC			
Ec6409	Fwd: GAAGAGACAAGTGGTGGGGTGAG	104.38%	0.99	0.07
	Rev: AGAAGAGAGCAAAGAGGAAGAGTGG			
EcSMT	Fwd: AAGATATTCAGCGGCTCTGTCAA	90.23%	0.98	0.074
	Rev: TTTGGCTACCCCGTAGTTATCTG			
EcSPMT	Fwd: CACCTTCCACTGTGCTCAATTC	100.45%	0.98	0.031
	Rev: CTTTGAAACATCAACCACCATCTT			
EcSPDS	Fwd: CAACCTCCCACCTGGTACGA	108.50%	0.98	0.008
	Rev: CCCATCAATCACCAGGACCTT			
EcSAMDC	Fwd: ACTCTCGGAATCCAGCCTCTT	100.65%	0.98	0.029
	Rev: CTTCCACGAGTATAACGCACAGA			
EcMPOBMT	Fwd: TTTTCCCTAGCAACAGTCTGCAT	107.51%	0.99	0.045
	Rev: CTCCAGCCCTTCAGGAACCT			

Table S2. Genotypes of yeast strains used in this study. '>' symbol denotes promoter ('p') or terminator ('t') sequence associated with a gene of interest

Strain	Genotype
CEN.PK2-1D	MATα; *ura3-52*; *trp1-289*; *leu2-3/112*; *his3Δ1*; *MAL2-8C*; *SUC2*
CSY1242	CEN.PK2-1D; *meu1-14*; *oaz1-16*; *his2*:: p*TPI1>ARG2*>t*STE2*, p*PGK1>CAR1*>t*PHO5*, p*TDH3>FMS1*>t*ADH1*; *hfd1-45*; *ald4-48*; *ald5-53*; *ald6-55*; *ALD2Δ0-ALD3Δ0*:: *HIS3*; *trp1*:: p*PGK1>AsADC*>t*ADH1*, p*TEF1>SPE1*>t*CYC1*, p*TDH3>speB*>t*PHO5*
CSY1246	CSY1242; *ura3*:: p*PGK1>AbPMT1*>t*PHO5*, p*TDH3>DmMPO1*$^{\Delta C\text{-}PTS1}$ >t*ADH1*, p*TEF1>DsTR1*>t*CYC1*
CSY1297	CSY1246; *leu2*:: p*PGK1>AbCYP82M3*>t*PHO5*, p*TEF1>AtATR1*>t*CYC1*, p*HXT7-AbPYKS*-t*PGK1*; *ald6-55*:: *ALD6*; *pad1*:: p*PGK1>AbPMT1*>t*PHO5*, p*TDH3>DsPMT1*>t*CYC1*, p*HXT7>AbPYKS*>t*PGK1*; *spe4*:: p*TEF1>DsH6H*>t*CYC1*, p*TDH3>AbCYP80F1*>t*ADH1*; *ald4*:: p*PGK1>AbUGT84A27*>t*PHO5*, p*TDH3>WfPPR*>t*ADH1*, p*TEF1>DsHDH*>t*CYC1*, p*HXT7>DsH6H*>t*PGK1*; *egh1*:: p*PGK1>UGP1*>t*PHO5*, p*TDH3>DsRed-AbLS*>t*CYC1*, p*TEF1>WfPPR*>t*ADH1*; *exg1*:: p*PGK1>DsH6H*>t*PHO5*, p*TDH3>WfPPR*>t*ADH1*, p*TEF1>NtJAT1*>t*CYC1*
CSY1340	CSY1242; *spe4*:: p*TEF1>SPE2*>t*CYC1*, p*TDH3>SPE3*>t*ADH1*, p*PGK1>EcSPMT*>t*PHO5*
CSY1341	CSY1246; *leu2*:: p*PGK1>EcOGAS1*>t*PHO5*, p*TEF1>EcMPOBMT*>t*CYC1*, p*TDH3>AbPYKS*>t*ADH1*
CSY1343	CSY1340; *leu2*:: p*PGK1>EcOGAS1*>t*PHO5*, p*TPI1>EcAOC2*>t*STE2*, p*TDH3-EcAOF1*>t*CYC1*, p*TEF1>EcSMT*>t*ADH1*; pCS4859 (EcMPOBMT), pCS4200 (AtATR1), pCS4892 (EcCYP81AN15 + EcMecgoR)

Table S3. List of plasmids used in this study.

Plasmid no.	Description / genotype	Organism	Origin of replication	Selection marker
pCS4208	pAG416-PTDH3-BFP-TCYC1	Yeast	CEN/ARS	URA3
pCS4212	pAG414-PTDH3-BFP-TCYC1	Yeast	CEN/ARS	TRP1
pCS4213	pAG415-PTDH3-BFP-TCYC1	Yeast	CEN/ARS	LEU2
pCS4238	pAG414-PTDH3-DmMPO1ΔC-PTS1-TCYC1	Yeast	CEN/ARS	TRP1
pCS4217	pAG415-PTDH3-mCherry-PEX3-TCYC1	Yeast	CEN/ARS	LEU2
pCS4246	pAG416-PTDH3-AbPYKS-TCYC1	Yeast	CEN/ARS	URA3
pCS4247	pAG415-PTDH3-AbCYP82M3-TCYC1	Yeast	CEN/ARS	LEU2
pCS4200	pAG414-PTEF1-AtATR1-TCYC1	Yeast	CEN/ARS	TRP1
pCS4831	pAG416-PTDH3-SPE2-TCYC1	Yeast	CEN/ARS	URA3
pCS4832	pAG415-PTDH3-SPE3-TCYC1	Yeast	CEN/ARS	LEU2
pCS4833	pAG415-PTDH3-EcSPDS-TCYC1	Yeast	CEN/ARS	LEU2
pCS4834	pAG415-PTDH3-EcSPMT-TCYC1	Yeast	CEN/ARS	LEU2
pCS4835	pAG415-PTDH3-EcSMT-TCYC1	Yeast	CEN/ARS	LEU2
pCS4836	pCS2657-PTEF1-SPE2-TCYC1	Yeast	--	--
pCS4837	pCS2656-PTDH3-SPE3-TADH1	Yeast	--	--
pCS4838	pCS2663-PPGK1-EcSPMT-TPHO5	Yeast	--	--
pCS4839	pAG416-PTDH3-EcAOF1-TCYC1	Yeast	CEN/ARS	URA3
pCS4840	pAG416-PTDH3-EcAOF2-TCYC1	Yeast	CEN/ARS	URA3
pCS4841	pAG416-PTDH3-EcAOF3-TCYC1	Yeast	CEN/ARS	URA3
pCS4842	pAG414-PTDH3-EcAOF2-TCYC1	Yeast	CEN/ARS	TRP1
pCS4843	pAG414-PTDH3-EcAOF3-TCYC1	Yeast	CEN/ARS	TRP1
pCS4844	pAG414-PTDH3-EcAOC1-TCYC1	Yeast	CEN/ARS	TRP1
pCS4845	pAG414-PTDH3-EcAOC2-TCYC1	Yeast	CEN/ARS	TRP1
pCS4846	pAG414-PTDH3-EcAOC3-TCYC1	Yeast	CEN/ARS	TRP1
pCS4847	pAG414-PTDH3-EcAOC4-TCYC1	Yeast	CEN/ARS	TRP1
pCS4848	pAG414-PTDH3-EcAOC5-TCYC1	Yeast	CEN/ARS	TRP1
pCS4849	pAG414-PTDH3-EcAOC6-TCYC1	Yeast	CEN/ARS	TRP1
pCS4850	pAG416-PTDH3-EcAOF1-TCYC1-PTEF1-EcSMT-TADH1	Yeast	CEN/ARS	URA3

pCS4851	pAG416-PTDH3-EcAOF1-GFP-TCYC1	Yeast	CEN/ARS	URA3
pCS4852	pAG416-PTDH3-GFP-EcAOF1-TCYC1	Yeast	CEN/ARS	URA3
pCS4853	pAG416-PTDH3-EcAOC1-GFP-TCYC1	Yeast	CEN/ARS	URA3
pCS4854	pAG416-PTDH3-GFP-EcAOC1-TCYC1	Yeast	CEN/ARS	URA3
pCS4855	pAG416-PTDH3-EcAOC2-GFP-TCYC1	Yeast	CEN/ARS	URA3
pCS4856	pAG416-PTDH3-GFP-EcAOC2-TCYC1	Yeast	CEN/ARS	URA3
pCS4857	pAG416-PTDH3-EcOGAS1-TCYC1	Yeast	CEN/ARS	URA3
pCS4858	pAG416-PTDH3-EcOGAS2-TCYC1	Yeast	CEN/ARS	URA3
pCS4859	pAG415-PTDH3-EcMPOBMT-TCYC1	Yeast	CEN/ARS	LEU2
pCS4860	pAG415-PTDH3-EcSABATH2-TCYC1	Yeast	CEN/ARS	LEU2
pCS4861	pAG415-pTDH3-EcSABATH3-TCYC1	Yeast	CEN/ARS	LEU2
pCS4862	pCS2663-PPGK1-EcOGAS1-TCYC1	Yeast	--	--
pCS4863	pCS2657-PTEF1-EcMPOBMT-TCYC1	Yeast	--	--
pCS4864	pCS2656-PTDH3-AbPYKS-TADH1	Yeast	--	--
pCS4865	pAG416-PTDH3-EcMecgoR-TCYC1	Yeast	CEN/ARS	URA3
pCS4866	pAG416-PTDH3-EcODD1-TCYC1	Yeast	CEN/ARS	URA3
pCS4867	pAG416-PTDH3-EcODD2-TCYC1	Yeast	CEN/ARS	URA3
pCS4868	pAG416-PTDH3-EcODD3-TCYC1	Yeast	CEN/ARS	URA3
pCS4869	pAG416-PTDH3-EcODD4-TCYC1	Yeast	CEN/ARS	URA3
pCS4870	pAG416-PTDH3-EcODD5-TCYC1	Yeast	CEN/ARS	URA3
pCS4871	pAG416-PTDH3-EcP4H1-TCYC1	Yeast	CEN/ARS	URA3
pCS4872	pAG416-PTDH3-EcP4H2-TCYC1	Yeast	CEN/ARS	URA3
pCS4873	pAG416-PTDH3-EcCYP1-TCYC1	Yeast	CEN/ARS	URA3
pCS4874	pAG416-PTDH3-EcCYP2-TCYC1	Yeast	CEN/ARS	URA3
pCS4875	pAG416-PTDH3-EcCYP3-TCYC1	Yeast	CEN/ARS	URA3
pCS4876	pAG416-PTDH3-EcCYP4-TCYC1	Yeast	CEN/ARS	URA3
pCS4877	pAG416-PTDH3-EcCYP5-TCYC1	Yeast	CEN/ARS	URA3
pCS4878	pAG416-PTDH3-EcCYP6-TCYC1	Yeast	CEN/ARS	URA3
pCS4879	pAG416-PTDH3-EcCYP7-TCYC1	Yeast	CEN/ARS	URA3
pCS4880	pAG416-PTDH3-EcCYP8-TCYC1	Yeast	CEN/ARS	URA3

pCS4881	pAG416-PTDH3-EcCYP81AN15 (EcCYP9)-TCYC1	Yeast	CEN/ARS	URA3
pCS4882	pAG416-PTDH3-EcCYP10-TCYC1	Yeast	CEN/ARS	URA3
pCS4883	pAG416-PTDH3-EcCYP11-TCYC1	Yeast	CEN/ARS	URA3
pCS4884	pAG416-PTDH3-EcCYP12-TCYC1	Yeast	CEN/ARS	URA3
pCS4885	pAG416-PTDH3-EcCYP13-TCYC1	Yeast	CEN/ARS	URA3
pCS4886	pAG416-PTDH3-EcCYP14-TCYC1	Yeast	CEN/ARS	URA3
pCS4887	pAG416-PTDH3-EcCYP15-TCYC1	Yeast	CEN/ARS	URA3
pCS4888	pAG416-PTDH3-EcCYP16-TCYC1	Yeast	CEN/ARS	URA3
pCS4889	pAG416-PTDH3-EcCYP17-TCYC1	Yeast	CEN/ARS	URA3
pCS4890	pAG415-PTDH3-EcPKC1-TADH1-PTPI1-EcPKC2-TCYC1	Yeast	CEN/ARS	LEU2
pCS4891	pCS2661-PTPI1-EcAOC2-TSTE2	Yeast	--	--
pCS4892	pAG416-PTDH3-EcCYP81AN15-TCYC1-TSTE2-EcMecgoR-PTPI1	Yeast	CEN/ARS	URA3
pJD001	pHREAC-m35S-5S0-EcODC-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD002	pHREAC-m35S-5S0-EcSPMT-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD003	pHREAC-minip35S-5S0-EcSMT-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD004	pHREAC-minip35S-5S0-EcAOF1-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD005	pHREAC-minip35S-5S0-EcAOC1-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD006	pHREAC-minip35S-5S0-EcAOC2-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD007	pHREAC-minip35S-5S0-EcOGAS1-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)
pJD008	pHREAC-minip35S-5S0-EcOGAS2-3CPMV-NosT-p35S-NS-35Sterm	Agro-bacterium	OriV	Kan(R)

Table S4. Mass fragment transitions and parameters used for compound identification by LC-MS/MS in this study.

Compound	MRM transition (m/z [M+H]+)	Fragmentor	Collision energy	Reference
Putrescine	89 → 72	18	9	(1)
Spermidine	146 → 72	81	17	This study
N-methylspermidine	160 → 72	81	17	This study
N-methylputrescine	103 → 72	20	9	(1)
N-methylpyrrolinium*	84 → 57	50	35	(1)
D_3 *N*-methylpyrrolinium*	87 → 60	380	20	This study
Hygrine*	142 → 84	74	13	(1)
MPOB	186 → 84	97	25	(1)
MPMOB	200 → 84	100	17	This study
Tropine	142 → 98	50	21	(1)
Methylecgonone	198 →166	50	21	(16)
Methylecgonine (CMO**-tropine)	200 → 182 200 → 82	50 50	21 30	(16) (17)
Littorine / hyoscyamine	290 → 124	100	25	(6)
Anisodamine	306 → 140	50	25	(18)
Scopolamine	304 → 138	64	17	(6)
CMO-littorine / CMO-hyoscyamine	348 → 182	100	25	This study
CMO-anisodamine	364 → 198	50	25	This study
CMO-scopolamine	362 → 196	64	17	This study

*****Note that for experiments where *N*-methylpyrrolinium was the expected terminal metabolite produced, hygrine was used as a proxy due to superior MS/MS sensitivity (and greater confidence in positive detection), as observed in (1). For experiments in which NMPy was not a key/terminal metabolite, NMPy was analyzed directly.

**CMO = 2-carbomethoxy-.

Additional Supplementary Information

Dataset S1 (separate file, “Supplemental Dataset 1.xlsx”). Species origins, gene names, and public database accession numbers of all enzymes used in phylogenetic analyses and dendrograms in this study.

Supplementary References

1. P. Srinivasan, C. D. Smolke, Engineering a microbial biosynthesis platform for de novo production of tropane alkaloids. *Nat. Commun.* **10**, 1–15 (2019).
2. G. Li, J. D'Auria, N. Kumar Katakam, C. W. Seifert, Efficient Synthesis of Methyl (S)-4-(1-Methylpyrrolidin-2-yl)-3-oxobutanoate as the Key Intermediate for Tropane Alkaloid Biosynthesis with Optically Active Form. *HETEROCYCLES* **99**, 604 (2019).
3. K. D. Entian, P. Kötter, 25 Yeast Genetic Strain and Plasmid Collections. *Methods Microbiol.* **36**, 629–666 (2007).
4. T. J. Kwiatkowski, H. Y. Zoghbi, S. A. Ledbetter, K. A. Ellison, A. C. Chinault, Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR. *Nucleic Acids Res.* **18**, 7191 (1990).
5. S. Alberti, A. D. Gitler, S. Lindquist, A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. *Yeast Chichester Engl.* **24**, 913–9 (2007).
6. P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* **585**, 614–619 (2020).
7. H. Peyret, J. K. M. Brown, G. P. Lomonossoff, Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions. *Plant Methods* **15**, 108 (2019).
8. O. W. Ryan, *et al.*, Selection of chromosomal DNA libraries using a multiplex CRISPR system. *eLife* **3**, 1–15 (2014).
9. T. Docimo, G. W. Schmidt, K. Luck, S. K. Delaney, J. C. D'Auria, Selection and validation of reference genes for quantitative gene expression studies in *Erythroxylum coca* (2013) https:/doi.org/10.12688/f1000research.2-37.v1 (September 7, 2022).
10. J. Hellemans, G. Mortier, A. De Paepe, F. Speleman, J. Vandesompele, qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **8**, R19 (2007).
11. R. Hanczkó, A. Jámbor, A. Perl, I. Molnár-Perl, Advances in the o-phthalaldehyde derivatizations. Comeback to the o-phthalaldehyde-ethanethiol reagent. *J. Chromatogr. A* **1163**, 25–42 (2007).
12. F. Allen, A. Pon, M. Wilson, R. Greiner, D. Wishart, CFM-ID: A web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res.* **42**, 94–99 (2014).
13. C. Guijas, *et al.*, METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* **90**, 3156–3164 (2018).
14. R. Höfgen, L. Willmitzer, Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Res.* **16**, 9877 (1988).
15. L. Ballut, *et al.*, HcPro, a multifunctional protein encoded by a plant RNA virus, targets the 20S proteasome and affects its enzymic activities. *J. Gen. Virol.* **86**, 2595–2603 (2005).
16. J. Jirschitzka, *et al.*, Plant tropane alkaloid biosynthesis evolved independently in the Solanaceae and Erythroxylaceae. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 10304–10309 (2012).
17. L. de Lima Feltraco Lizot, *et al.*, Simultaneous determination of cocaine, ecgonine methyl ester, benzoylecgonine, cocaethylene and norcocaine in dried blood spots by ultra-performance liquid chromatography coupled to tandem mass spectrometry. *Forensic Sci. Int.* **298**, 408–416 (2019).
18. Z. Jandrić, *et al.*, Development of a liquid chromatography-tandem mass spectrometric method for the simultaneous determination of tropane alkaloids and glycoalkaloids in crops. *Food Addit. Contam. Part Chem. Anal. Control Expo. Risk Assess.* **28**, 1205–1219 (2011).