Unraveling Tropane Alkaloid Biosynthesis in *Erythroxylum coca*

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I

- Biowissenschaften -

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt

von

Herrn **Benjamin Gabriel Chavez**

Eröffnet am 11.04.2024

Verteidigt am 22.01.2025

Gutachter: Prof. Dr. Thomas Altmann

Prof. Dr. Steffen Abel

Prof. Dr. Alisdair Fernie

Parts of this thesis have been published in the Proceedings of the National Academy of Sciences

Chavez, B. G., Srinivasan, P., Glockzin, K., Kim, N., Montero Estrada, O., Jirschitzka, J., Rowden, G., Shao, J., Meinhardt, L., Smolke, C. D., & D'Auria, J. C. (2022). Elucidation of tropane alkaloid biosynthesis in *Erythroxylum coca* using a microbial pathway discovery platform. Proceedings of the National Academy of Sciences, 119(49), e2215372119. https://doi.org/10.1073/pnas.2215372119

TABLE OF CONTENTS

TABLE OF CONTENTS iii

I. LIST OF FIGURES vii

II. LIST OF TABLES ix

III. ABBREVIATIONS x

IV. SUMMARY xii

1 INTRODUCTION 1

1.1 Tropane Alkaloids 4

1.2 Background of *Erythroxylum coca* 8

1.3 Diverse Roles of Polyamines in Plants 10

1.3.1 Putrescine Biosynthesis 11

1.3.2 *S*-adenosyl-L-methionine Dependent Methyltransferases in Plants 14

1.3.3 dcSAM is an Essential Cofactor for Polyamine Biosynthesis 14

1.3.4 Polyamine Elongation via Aminopropyltransferases 16

1.4 Putrescine *N*-methyltransferases 17

1.5 Polyamine Oxidases 19

1.6 First Ring Closure of Tropane Alkaloids 21

1.7 Second Ring Closure of Tropane Alkaloids 21

1.8 Tropinone Reductases 22

1.9 Acyltransferase Enzymes Involved in TA Biosynthesis 24

1.10 Scope of the Thesis 26

2 MATERIALS & METHODS 27

2.1 Codon Optimization of Synthetic DNA 27

2.2 DNA Primers and Plasmid Sequencing 27

2.3 DNA Amplification and PCR validation 27

2.4 DNA Agarose and SDS-PAGE Protein Gel Imagining 28

2.5 Modular Cloning (MoClo)-Based Method of Cloning 28

2.6 YPD and BMMY Media for *Komagataella phaffii* Protein Expression 28

2.7 Cloning, Heterologous Expression, and Protein Purification of EcSMT, EcSPMT, and EcSPDS from *Komagataella phaffii* 28

2.8 Construction of *Saccharomyces cerevisiae* strains for *in vivo* Screening of Candidate Genes 29

2.9 Preparation of Chemically Competent *E. coli* Cells 29

2.10 Transformation of Chemically Competent *E. coli* Cells 30

2.11 Preparation of Chemically Competent *Agrobacterium tumefaciens* Cells 30

2.12 Protein Sequence Alignments 30

2.13 Generation of Phylogenetic Trees 30
2.14 Transformation of Plasmids into *Agrobacterium tumefaciens* cells 31
2.15 Transformation of Plasmids into *K. pastoris* cells 31
2.16 Fluorescent Microscopy of Organelle Targeting in *Saccharomyces cerevisiae* 31
2.17 Identification of TA Genes from *E. coca* Transcriptome 32
2.18 RNA Extraction from *E. coca* Tissue and cDNA Synthesis 32
2.19 qPCR Primers Design Parameters 33
2.20 Polyamine Extraction of *E. coca* tissues 33
2.21 Derivatization of Polyamines 33
2.22 Polyamine Detection by HPLC-Fluorescence 33
2.23 Polyamine Detection by UPLC-Fluorescence 34
2.24 *In vitro* Enzyme Assays for SPDS and SMT Activity via Fluorescence Detection 34
2.25 Plant Material 34
2.26 Plasmid Construction for Transient Expression in *Nicotiana benthamiana* 35
2.27 Transient Protein Expression in *Nicotiana benthamiana* 35
2.28 Metabolite Analysis of *Nicotiana benthamiana* Extracts 36
2.29 LC-MS Separation of Polyamines and Methionine Cycle-related Metabolites 36
2.30 Hierarchical Clustering Analysis of Erythroxylum Transcriptomes 37
2.31 Statistics 37
2.32 Chemical Standards 37
2.33 Genbank Accession Numbers of Identified Genes from *E. coca* 38
3 RESULTS 38
3.1 Detecting Derivatized Polyamines from *E. coca* Tissue Extracts via HPLC-FLD 38
3.2 The First Committed Step of TA Biosynthesis in *E. coca* Begins with Spermidine *N*-methylation 40
3.2.1 Homology Search of PMT-like Candidates 40
3.2.2 Heterologous Expression in *Komagataella phaffii* and Purification of EcSPDS1-3 Candidates 40
3.2.3 *In vitro* Enzymatic Activities Assays of SPDS-like Candidates 40
3.2.4 Phylogenetic Analysis of EcSMT, EcSPMT, and EcSPDS 43
3.2.5 *In vivo* Screening of EcSMT, EcSPMT, and EcSPDS Enzymatic Activities in *Saccharomyces cerevisiae* 45
3.3 A Flavin-Dependent Amine Oxidase mediates the Formation of *N*-methylputrescine 46
3.3.1 Homology Search of Flavin-Dependent Amine Oxidases in *E. coca* 46
3.3.2 Phylogenetic Analysis of Flavin-Dependent Oxidases from *E. coca* 46

3.3.3 *In vivo* Screening of EcAOF1-3 Candidates for *N*-methylspermidine Oxidase Activities in *Saccharomyces cerevisiae* 47
3.4 Two Copper-Dependent Amine Oxidases mediate First Ring Closure of Tropane Alkaloids in *E. coca* 48
3.4.1 Homology Search of Copper-Dependent Amine Oxidases in *E. coca* 48
3.4.2 Phylogenetic Analysis of Copper-Dependent Oxidases from *E. coca* 49
3.4.3 *In vivo* Screening of EcAOC1-6 Candidates for *N*-methylputrescine Oxidase Activity in *Saccharomyces cerevisiae* 51
3.4.4 Subcellular Localization of EcAOF1, EcAOC1, and EcAOC2 in *Saccharomyces cerevisiae* 52
3.5 First Ring Closure Pathway Reconstruction of *E. coca* Genes in *Nicotiana benthamiana* 53
3.5.1 Transformation of *Agrobacterium tumefaciens* GV2260 for Transient Expression in *Nicotiana benthamiana* 53
3.5.2 Agroinfiltration and Gene Stacking Techniques in *Nicotiana benthamiana* 54
3.6 A Member of the SABATH Methyltransferase Family Mediates the 2-CMO Group Retention in TA present in *E. coca* 56
3.6.1 Homology Search of SABATH-like Methyltransferases in *E. coca* 56
3.6.2 Phylogenetic Analysis of SABATH-like Methyltransferases from *E. coca* 57
3.6.3 *In vivo* Screening of EcAOC1-6 for *N*-methylputrescine Oxidase Activity in *Saccharomyces cerevisiae* 58
3.7 A CYP81-Family Monooxygenase Catalyzes Methylecgonone Ring Closure in *E. coca* 59
3.7.1 Homology Search of Tropinone Synthase-like Candidates in *E. coca* 59
3.7.2 *In vivo* Screening of Hydroxylase and Cyclase Candidates for Methylecgonone Synthase Activity in *Saccharomyces cerevisiae* 60
3.7.3 Phylogenetic Analysis of Methylecgonone Synthase from *E. coca* 60
3.8 RNA Isolation and cDNA Synthesis from *Erythroxylum coca* leaf tissue 61
3.8.1 Total RNA Extraction from *Erythroxylum coca* 61
3.8.2 Primer Optimization for qPCR Amplicon Targets 63
3.9 Utilizing an LC-MS Method without Derivatizing Polyamines for Detection 66
3.9.1 HPLC-FLD Detection of Derivatized Polyamine Standards 66
3.9.2 UPLC Separation and Detection of Derivatized Polyamine Standards 67
3.9.3 LC-MS/MS Separation and Detection of Polyamine Standards without Derivatization Techniques 68
4 DISCUSSION 75
4.1 Identification of Missing Steps of the Tropane Alkaloid Pathway in *E. coca* 75
4.2 Spermidine is Essential for Growth and Development in Eukaryotes 77
4.3 Evolutionary Insights on EcSPMT Versus PMT Steps in TA Biosynthesis 77

4.4 EcSPMT Possesses Both Spermidine Synthase and Spermidine *N*-methyltransferase Activities 78
4.5 Both Flavin-Dependent and Copper-Dependent Oxidases Are Required for First Ring Closure in TA Biosynthesis in *E. coca* 80
4.6 MPOBMT is Essential for Formation of the 2-carboxymethyl Ester Present in TAs in *Erythroxylum coca* 82
4.7 Methylecgonone Synthase (EcCYP81AN15) Facilitates Second Ring Closure of TAs in *E. coca* 83
4.8 Tissue-specific Expression of Tropane Alkaloid Biosynthesis Genes 83
4.9 LC-MS/MS Separation and Detection of Polyamine and SAM-Related Metabolites 84
5 CONCLUSION 87
6 OUTLOOK 88
7 BIBLIOGRAPHY 89
8 APPENDIX 106
8.1.1 Protein Alignment of SPDS-like Candidates with PMTs and SPDSs Enzymes 106
8.1.2 Heterologous Expression in *Komagataella phaffii* and Purification of SPDS Candidates 107
8.1.3 SPDS Sequence Accession for Phylogenetic Analysis 109
8.1.4 Flavin-Dependent Oxidases Sequences Accession for Phylogenetic Analysis 111
8.1.5 Copper-Dependent Oxidases Sequence Accession for Phylogenetic Analysis 112
8.1.6 SABATH Methyltransferases Sequence Accession for Phylogenetic Analysis 114
8.1.7 P450 Cyclases Sequence Accession for Phylogenetic Analysis 115
8.1.8 Hierarchical clustering heatmap showing expression profiles for cyclase 116
8.1.9 Primers for Cloning Synthetic DNA into pHREAC Vector 117
8.1.10 Nanodrop Quantification of RNA Samples from *Erythroxylum coca* 118
8.1.11 Bioanalyzer Quantification of RNA Samples from *Erythroxylum coca* 119
8.1.12 qPCR Primer Table 121
8.1.13 Mass fragment transitions and parameters for compound identification by LC-MS/MS 122
9 CURRICULUM VITAE 123
10 DECLARATION (ERLÄRUNG) 125
11 ACKNOWLEDGEMENTS 126

I. LIST OF FIGURES

Figure 1: Specialized metabolites and their influences and interactions with hormones and primary metabolism within plants. 2
Figure 2: The core structure of tropane alkaloids. 4
Figure 3: *Erythroxylum coca, Atropa belladonna*, and *Hyoscyamus niger,* along with their respective tropane alkaloid cocaine, atropine, and scopolamine..... 5
Figure 4: The diversity of tropane alkaloids in plants derived from the tropane core structure. 6
Figure 5: Simplified overview of the chemical synthesis of tropinone..... 7
Figure 6: A timeline overview of tropane alkaloid pathway discovery over a century from the 1900s until the modern day. 8
Figure 7: *Erythroxylum coca*. 9
Figure 8: Chemical structure of polyamines observed in plants. 11
Figure 9: General polyamine biosynthetic pathway in plants. 13
Figure 10: Chemical structure of SAM and dcSAM..... 16
Figure 11: Putrescine *N*-methyltransferase reaction 18
Figure 12: Simplified view of back conversion pathway (A) via a flavin-dependent oxidase and terminal catabolic pathway (B) via a copper-dependent polyamine oxidase..... 20
Figure 13: Depiction of methylputrescine oxidase (MPO) activity leads to *N*-methylpyrrolinium cation formation. 21
Figure 14: Second ring closure of tropane rings in *Atropa belladonna*. 22
Figure 15: Tropinone reductase and methylecgonone reductase catalyzing the reduction of the C3 position of the tropane ring of either tropinone or methylecgonone, respectively..... 24
Figure 16: The acylation of tropane alkaloids mediated by either a BAHD acyltransferase or SCPL acyltransferase..... 26
Figure 17: Polyamine content in *E. coca* tissues. 39
Figure 18: *In vitro* enzymatic characterization of purified EcSMT. 41
Figure 19: *In vitro* enzymatic characterization of purified EcSPMT. 42
Figure 20: *In vitro* enzymatic characterization of purified EcSPDS..... 42
Figure 21: A dendrogram showing the sequence relationships based on amino acids between EcSPDS, EcSMT, and EcSPMT..... 44
Figure 22: Characterization of *E. coca* spermidine *N*-methyltransferase activities in engineered yeast. 45
Figure 23: Phylogenetic analysis of flavin-dependent oxidases from *E. coca*..... 47
Figure 24: LC-MS/MS multiple reaction monitoring (MRM) chromatogram traces showing the screening of *E. coca* flavin-dependent amine oxidase (AOF) candidates in yeast for activity on *N*-methylspermidine..... 48
Figure 25: Phylogenetic analysis of copper-dependent polyamine oxidases from *E. coca*. 50
Figure 26: The biosynthetic pathway leading to *N*-methylpyrrolinium formation from *N*-methylputrescine. 52
Figure 27: Subcellular localization of *E. coca* amine oxidases in yeast via fluorescence microscopy. 53
Figure 28: Colony PCR of transformed *Agrobacterium tumefaciens* GV2260 harboring pHREAC expression vector. 54
Figure 29: Validation of the biosynthetic sequence for the first tropane ring formation in *E. coca* via transient co-expression in *N. benthamiana*. 56
Figure 30: Dendrogram showing phylogenetic analysis of three SABATH methyltransferase candidates in green identified from *E. coca* transcriptome. 57
Figure 31: LC-MS/MS MRM chromatogram traces showing the screening of *E. coca* SABATH methyltransferase candidates for activity on MPOB in engineered yeast..... 59
Figure 32: Methylecgonone production in yeast strains engineered for co-expression of hydroxylase/cyclase candidates. 60
Figure 33: Dendrogram showing phylogenetic analysis of methylecgonone synthase 61
Figure 34: Photograph of *E. coca* plant showing the different developmental stages of the leaves. 63

Figure 35: 4% Agarose DNA gel of qPCR primer optimization of L3 stage cDNA from *Erythroxylum coca* generated by NEB Lunascript Reverse Transcriptase. 64
Figure 36: 4% Agarose gel of qPCR primer optimization of L3 stage cDNA from *Erythroxylum coca* generated by SuperScript IV Reverse Transcriptase ThermoScientific. 65
Figure 37: 4% Agarose DNA gel of qPCR primer optimization of L1 stage cDNA from *Erythroxylum coca* generated by ThermoScientific SuperScript IV Reverse Transcriptase. 66
Figure 38: HPLC-FLD chromatogram of *N*-methylspermidine standard compared to OPA-FMOC derivatization blank. 67
Figure 39: UPLC-FLD chromatogram of *N*-methylspermidine and spermidine standard compared to OPA-FMOC derivatization blank. 68
Figure 40: UPLC-FLD chromatogram of acetonitrile wash after injecting the derivatized polyamine standards 68
Figure 41: LC-MS/MS detection of polyamine standard mixture containing putrescine (red peak) and spermidine (green peak) at 10µM concentration along with 100µM *N*-methylspermidine (purple peak). 69
Figure 42: Linear regression curve of the peak area of spermidine standard MRM detection transitions at 146.1→72.1 m/z. 70
Figure 43: LC-MS/MS detection of MRM transitions at 146.1→72.1 m/z corresponds to the spermidine standard at varying concentrations from 2.5µM-100µM. 71
Figure 44: Linear regression curve of the peak area of *N*-methylspermidine standard MRM detection transitions at 160→72.1 m/z. 71
Figure 45: LC-MS/MS detection of MRM transitions at 160.0→72.1 m/z corresponds to the *N*-methylspermidine standard at varying concentrations from 25µM-0.5mM. 72
Figure 46: LC-MS/MS MRM separation of the standards *S*-adenosyl-L-homocysteine (orange peak) and *S*-adenosyl-L-methionine (blue peak) at 10µM concentration. 73
Figure 47: LC-MS/MS MRM detection of substrates and products generated from EcSMT enzyme assay. SAH (orange peak), *N*-methylspermidine (purple peak), and spermidine (green peak). ... 74
Figure 48: A comparison of tropane alkaloid pathway in Erythroxylaceae (green color) and Solanaceae (purple). 76
Figure 49: Depicting the differences between the beginning steps of tropane alkaloid biosynthesis in Solanaceous plants in purple and *Erythroxylum coca* in green color. 79

Figure S1: Protein alignment of EcSPDS1-3 candidates, putrescine *N*-methyltransferases from *Datura stramonium* and *Atropa belladonna*, and spermidine synthase 1 from *Arabidopsis thaliana*. ... 106
Figure S2: SDS-PAGE of streptag purification of EcSMT from *K. phaffii* KM71. 107
Figure S3: SDS-PAGE of streptag purification of EcSPMT from *K. phaffii* KM71. 107
Figure S4: EcSPMT densitometric analysis of lane 6 band percentage from the SDS-PAGE. 108
Figure S5: SDS-PAGE of streptag purification of EcSPDS from *K. phaffii* KM71. 108
Figure S6: Hierarchical clustering heatmap showing expression profiles for cyclase 117
Figure S7: An electrophoresis run of RNA isolated from *Erythroxylum coca* leaf tissues L1, L2, and L3 leaf stages. 119
Figure S8: Agilent Bioanalyzer assessment of RNA quality from *E. coca* leaf tissues. L1 is rolled young leaves, L2 is unrolled young leaves, and L3 is mature leaves. 119

II. LIST OF TABLES

Table 1: Gene stacking matrix combinations of Agrobacterium-mediated transformation of *Nicotiana benthamiana*..... 55
Table 2: Reaction setup for EcSPMT enzyme assay to evaluate SMT activity. 74

Table S1: Key for the abbreviations of SPDS-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. 109
Table S2: A key for the abbreviations of AOF-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. 111
Table S3: A key for the abbreviations of AOC-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. 112
Table S4: This is the key for the abbreviations of SABATH-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. 114
Table S5: A key for the abbreviations of Cytochrome P450-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. 115
Table S6: Primer sequences for cloning into pHREAC vector via Golden Gate Cloning. 117
Table S7: Nanodrop readings of RNA quality from *E. coca* leaf tissues. 118
Table S8: Preparation of RNA samples from *E. coca* for cDNA synthesis 118
Table S9: Agilent Bioanalyzer assessment of RNA quality from *E. coca* leaf tissues. 120
Table S10: Primers used to optimize qPCR cDNA targets in *E. coca*. 121
Table S11: Multiple Reaction Monitoring (MRM) transition list for LC-MS/MS detection. 122

III. ABBREVIATIONS

Table 1: Abbreviations used in this doctoral thesis.

Abbreviation	Name
ACL5	Thermospermine synthase
ADC	Arginine decarboxylase
AOC	Copper-dependent amine oxidase
AOF	Flavin-dependent amine oxidase
ARGAH	Arginase
Cad	Cadaverine
CaMV	Cauliflower Mosaic Virus
cDNA	Complementary DNA
CMO	Carbomethoxy ester
CS	Cocaine Synthase
DAP	1,3-diaminopropane
dcSAM	Decarboxylated *S*-adenosyl-L-methionine
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray Ionization
FLD	Fluorescent Detection
GFP	Green fluorescent protein
HFBA	Heptafluorobutyric acid
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High Performance Liquid Chromatography
HPLC-FLD	High Performance Liquid Chromatography Fluorescent Detection
LC-MS	Liquid Chromatography Mass Spectrometry
MAT	Methionine Adenosyltransferase
MecgoR	Methylecgonone Reductase
MES	2-(N-Morpholino) Ethanesulfonic acid
MPOB	4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate
MRM	Multiple Reaction Monitoring
mRNA	Messenger Ribonucleic acid
MTA	5′-Deoxy-5′-Methylthioadenosine
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCPAH	N-Carbamoylputrescine Amidohydrolase
NLS	Nuclear Localization Signal
ODC	Ornithine Decarboxylase
OGAS	3-Oxoglutaric Acid Synthase
PAO	Polyamine Oxidase
PCA	Perchloric Acid
PCR	Polymerase Chain Reaction
PKS	Polyketide Synthase
PLP	Pyridoxal phosphate
PTM	Posttranslational Modification
PTS	Peroxisomal Targeting Signal
Put	Putrescine
PYKS	Pyrrolidine Ketide Synthase
ROS	Reactive Oxygen Species

SAM	*S*-adenosyl-L-methionine
SAMDC	*S*-adenosyl-L-methionine Decarboxylase
SD	Standard Deviation
SDR	Short-Chain Dehydrogenase/Reductase
SMT	Spermidine *N*-methyltransferase
SOD	Superoxide Dismutase
Spd	Spermidine
SPDS	Spermidine Synthase
Spm	Spermine
SPMT	Spermidine Synthase/Spermidine *N*-methyltransferase
SPS	Spermine Synthase
TA	Tropane Alkaloid
TBON	8-thiabicyclo[3.2.1]octan-3-one
TR-I	Tropinone reductase 1
TR-II	Tropinone reductase 2
TRs	Tropinone reductases
Tspm	Thermospermine
UPLC	Ultra Performance Liquid Chromatography
UPLC-FLD	Ultra Performance Liquid Chromatography Fluorescent Detection
UTR	Untranslated Region
4MAB	γ-Methylaminobutyraldehyde diethyl acetal

IV. SUMMARY

Tropane alkaloids (TAs) are heterocyclic nitrogenous metabolites found across seven orders of angiosperms, including Malpighiales (Erythroxylaceae) and Solanales (Solanaceae). Despite cocaine being an infamous tropane alkaloid from *Erythroxylum coca,* the biosynthetic pathway has remained incomplete for the past couple of decades. Using yeast as a screening platform, the missing enzymatic steps of TA biosynthesis in *Erythroxylum coca*. This work characterized a polyamine synthase along with amine oxidase-like enzymes *in vitro*, in yeast, and *in planta,* revealing that the first ring closure of TAs in *E. coca* occurs via bifunctional spermidine synthase/*N*-methyltransferases and both flavin- and copper-dependent amine oxidases. Identification of a SABATH family methyltransferase is responsible for the 2-carbomethoxy moiety characteristic of TAs from the Erythroxylaceae family. Finally, a cytochrome P450 of the CYP81A family was identified as responsible for the second tropane ring closure in *Erythroxylum coca*. These results demonstrate that tropane alkaloid biosynthesis in Erythroxylaceae and Solanaceae is polyphyletic in origin, further revealing that independent recruitment of unique biosynthetic mechanisms and enzyme classes occurred at nearly every step in the evolution of this pathway.

1 INTRODUCTION

Due to their sessile nature, plants have established themselves as some of the most sophisticated chemists, developing diverse specialized metabolites in response to abiotic and biotic stresses. Angiosperms are well known for several key innovations, including the diversity of flower development, double fertilization, fruits, and specialized metabolites (Benton et al., 2022; Wink, 2019). These specialized metabolites enabled flowering plants to defend themselves against predators, subsequently driving the diversification of numerous animal groups to evolve alongside angiosperms. After hundreds of millions of years of evolution, plants, especially flowering plants, provide a critical source of bioactive compounds and medicines (Veeresham, 2012). These innovations have played a significant role in the success of angiosperms, enabling them to colonize diverse habitats, adapt to a wide range of environmental conditions, and outcompete other plant groups. Their dominance has shaped the ecosystems of Earth, providing food sources, habitats, and essential resources for numerous organisms.

Secondary metabolites are a vast and diverse group of organic compounds produced by many organisms, including bacteria, fungi, animals, and plants. The term secondary metabolite is outdated; therefore, specialized metabolite will be used instead to refer to secondary metabolites. These specialized metabolites play crucial roles in various aspects of biological and ecological systems. Specifically, these specialized metabolites are not essential for primary growth and development in plants but serve specialized functions that enhance the plant's survival and adaptation to environmental conditions. Generally, secondary metabolites in plants can be divided into four metabolite classes: terpenes, phenylpropanoids, sulfur-containing, and amine-containing compounds. Secondary metabolites are essential to plants as they possess many biological effects, which play a pivotal role in defense, signaling, and photoprotection. Their diverse biological functions contribute to the success and adaptation of plants in various environments through interactions with both primary and phytohormone metabolism [

Figure 1] (Erb & Kliebenstein, 2020).

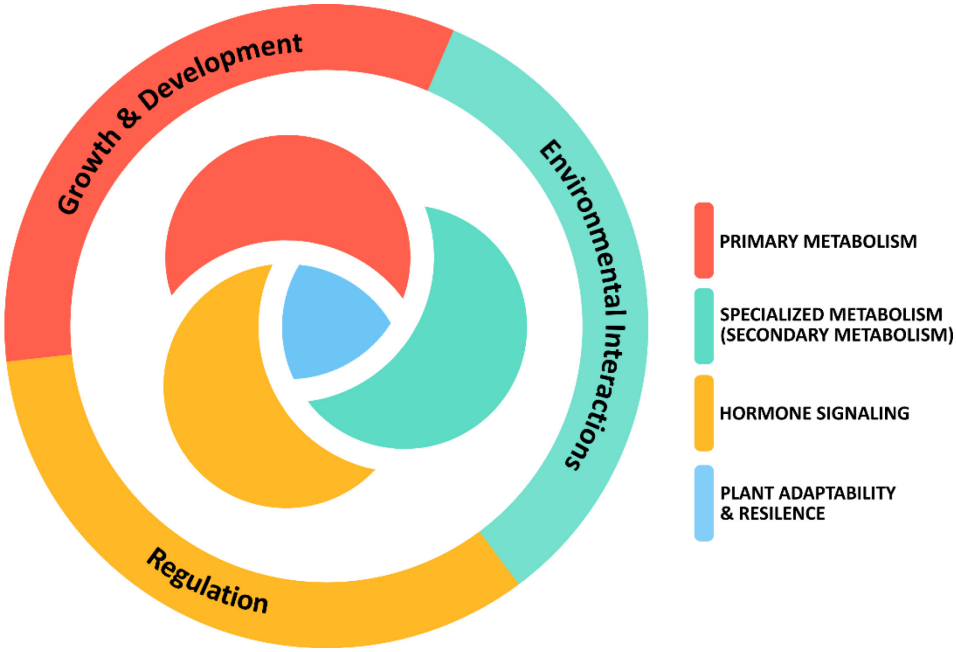

Figure 1: Specialized metabolites and their influences and interactions with hormones and primary metabolism within plants.

Terpenoids are the largest group of plant-specialized metabolites (Divekar et al., 2022). Terpenoids, sometimes referred to as terpenes, are divided into six groups: hemiterpenoids, monoterpenoids, diterpenoids, triterpenoids, tetraterpenoids, and sesquiterpenoids. These complex chemical structures, derived from isoprene units, serve various essential functions, including defense against biotic stresses, allelopathic interactions, photoprotection, signaling, and communication. Terpenoids can serve as potent defense mechanisms against herbivorous insects, microorganisms, and pathogens (Pichersky & Raguso, 2018). They can act as insect repellents, antimicrobial agents, and antifungal compounds, protecting the plant from damage and infection. Terpenoids act as chemical signals, mediating interactions between plants, insects, and even within the plant itself (Wei et al., 2007). They can serve as attractants or repellents, signaling the presence of food, mates, or potential threats. For instance, volatile terpenoids emitted by damaged plant tissues attract predators of herbivores, providing a defense mechanism against herbivory. Furthermore, terpenoids can attract beneficial organisms such as mycorrhizal fungal networks and pollinators (Sharma et al., 2017). Terpenoids are involved in various biochemical pathways, contributing to the synthesis of other essential compounds, such as chlorophyll, carotenoids, and phytohormones. They play a critical role in plant secondary metabolism, growth, and development.

Phenylpropanoids are a diverse group of specialized metabolites that play a wide range of crucial roles in the lives of plants. Phenylpropanoids are phenolic compounds that are

derived from amino acids, either the amino acid phenylalanine or tyrosine. These compounds are involved in various biological processes, including antioxidants, UV protection, pigmentation, and photoprotection (Ortiz & Sansinenea, 2023). Phenylpropanoids can be divided into five groups: flavonoids, coumarins, phenolic acids, stilbenes, and monolignols. Phenylpropanoids are essential components of plant cell walls, providing structural support and resistance to mechanical stress and decay. Lignin, the most abundant phenylpropanoid, forms a strong network within cell walls, contributing to the rigidity and durability of plant tissues (Boerjan et al., 2003). Phenylpropanoids are involved in various physiological processes, including signaling, hormone synthesis, and adaptation to environmental changes (Dong & Lin, 2021). Phenylpropanoids regulate the opening and closing of stomata, influence photosynthesis, and modulate plant growth and development. Phenylpropanoids' involvement in various biological processes, including structural support, defense mechanisms, and signaling, highlights their multifaceted importance in the plant kingdom.

Sulfur-containing specialized metabolites are mainly amino acid-derived and are most commonly found in the Brassicaceae and Capperales, giving rise to the indicative tastes and smells of cruciferous and allium vegetables (Divekar et al., 2022; Hill et al., 2023). One of the primary functions of sulfur-containing specialized metabolites is to protect plants from herbivores and pathogens. These compounds act as deterrents to herbivores, making plant tissues less palatable and even toxic. For instance, the pungent smell of garlic (*Allium sativum*) and onions (*Allium cepa*) contain thiosulfinate esters that can serve as a defensive response (Reiter et al., 2020). The formation of allicin only occurs when the plant suffers mechanical damage, such as cutting or crushing, allowing for a rapid response to herbivory. Additionally, these sulfur-based compounds possess antimicrobial properties, inhibiting the growth of microorganisms and protecting plants from diseases caused by viruses, bacteria, and fungi (Blume et al., 2023). These sulfur-containing metabolites play roles in defense against herbivores and pathogens, adaptation to environmental stresses, and promotion of plant growth and development.

Nitrogen-containing compounds, which consist of cyanogenic glucosides, non-proteinogenic amino acids, and alkaloids, are abundant in plants. Alkaloids are a particular class of specialized metabolites known for their medicinal properties. Alkaloids are low molecular weight nitrogen-containing compounds with a heterocyclic ring containing a nitrogen atom. These nitrogen-containing compounds are found in various plants, including coffee, tea, and tobacco, and possess a range of physiological and pharmacological effects.

Alkaloids represent a diverse class of secondary metabolites widely distributed in the plant kingdom, especially among angiosperm plants, where over 20% of all species produce alkaloids (Wink, 2007). These plant-derived alkaloids can be divided into differing classes depending on their precursors, and to date, more than 20 different classes of alkaloids have been reported (Matsuura & Fett-Neto, 2015; L. Yang & Stöckigt, 2010). Some examples of plant-derived alkaloids are the following classes: lycopodium, quinolizidine, benzylisoquinoline, tetrahydroisoquinolines, pyrrolizidine, monoterpene indoles, piperidine, colchicine, nicotine, and tropane alkaloids (Lichman, 2021).

1.1 Tropane Alkaloids

Human beings have cultivated plants for medicinal properties for several thousands of years. Tropane alkaloids (TAs) are specialized secondary metabolites found in multiple plant families with a scattered distribution, suggesting that these pathways evolved independently (J.-P. Huang et al., 2021; Jirschitzka et al., 2012). Tropane alkaloids share the characteristic 8-methyl-8-azabicyclo[3.2.1]octane nucleus, widely referred to as a tropane ring [Figure 2] (Docimo et al., 2012).

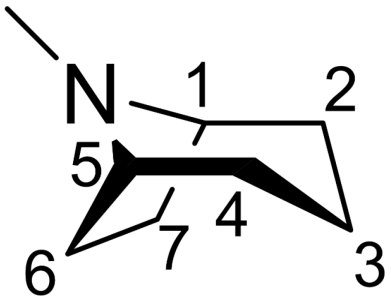

Figure 2: The core structure of tropane alkaloids. The core scaffold of tropane alkaloids is defined as an 8-methyl-8-azabicyclo[3.2.1]octane nucleus. The numbering indicates the carbon positions on the tropane core.Tropane alkaloids are known to interact with the central nervous systems and are thought to act as a deterrent against herbivores and insects. TAs occur naturally in many members of the plant family Solanaceae, including deadly nightshade (*Atropa belladonna*), jimson weed (*Datura stramonium*), black henbane (*Hyoscyamus niger*), and mandrake (*Mandragora officinarum*). Although the Solanaceae family is prominently known to produce tropane

alkaloids, they can also be found in members of the Erythroxylaceae family, specifically in coca (*Erythroxylum coca*) [Figure 3].

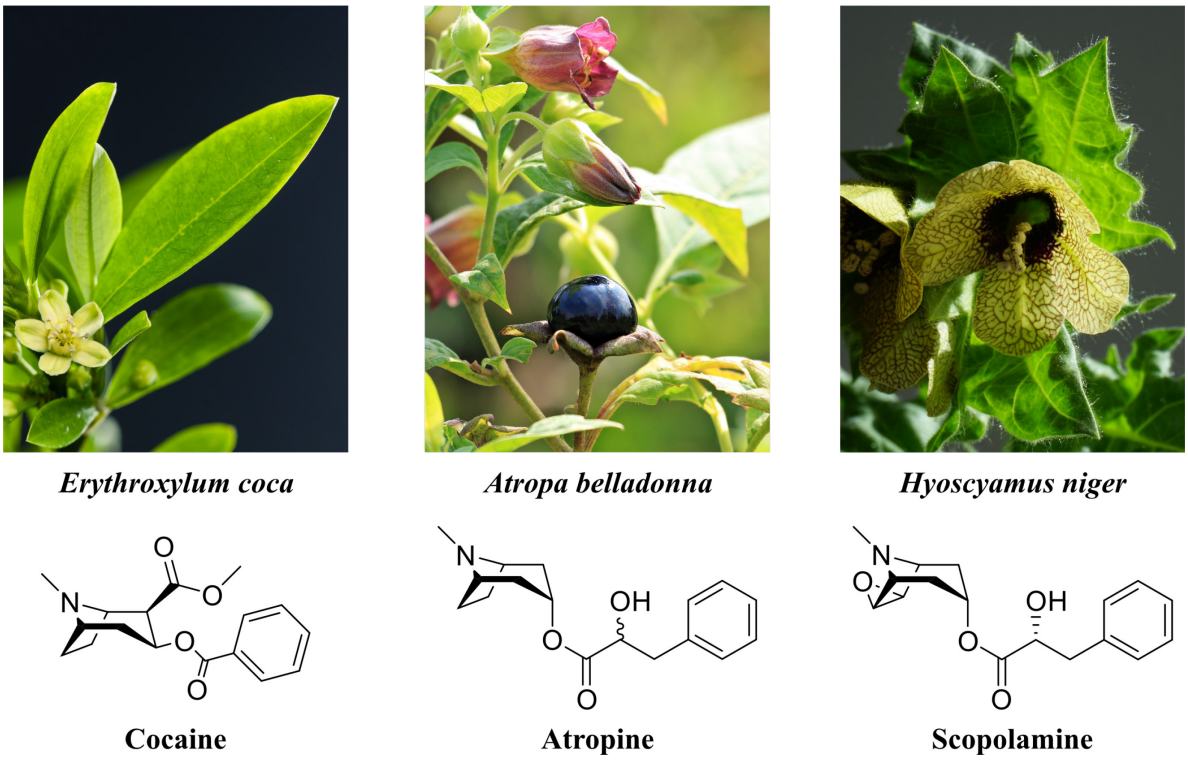

Figure 3: *Erythroxylum coca, Atropa belladonna*, and *Hyoscyamus niger,* along with their respective tropane alkaloid cocaine, atropine, and scopolamine.The *Atropa belladonna* plant image is from https://pixabay.com/photos/great-cherry-atropa-belladonna-plant-1547692/ under a free-use license. The *Hyoscyamus niger* image source is from https://pixabay.com/photos/hyoscyamus-niger-henbane-2306579/ under a free-use license, and Dr. Danny Kessler from the Max Planck Institute for Chemical Ecology in Jena, Germany provided the *Erythroxylum coca* image.

Other TAs, like calystegines, are also present in solanaceous plants and are mainly hydroxylated derivatives of tropane alkaloids (Bekkouche et al., 2001; Drager et al., 1995). To date, there have been eleven plant families found to produce TAs, such as Convolvulaceae, Proteaceae, Rhizophoraceae, Apocynaceae, Brassicaceae, Olacaceae, Phyllanthaceae, Moraceae, Fabaceae, Solanaceae, Erythroxylaceae families (J.-P. Huang et al., 2021).

Tropane alkaloids are valuable in traditional medicine and modern pharmaceutical applications. Through interactions with muscarinic acetylcholine receptors, TAs can possess anticholinergic, deliriant, and stimulant properties (Dey et al., 2020). They are primarily used to treat conditions involving excessive smooth muscle contractions, such as irritable bowel syndrome, urinary tract spasms, and motion sickness (*WHO Model List of Essential Medicines - 23rd List, 2023*, 2023). Atropine, a well-known tropane alkaloid, is commonly used as an eye drop to dilate pupils during eye examinations. Scopolamine, another important tropane

alkaloid, is employed as an antispasmodic agent and a sedative, particularly for motion sickness and insomnia. The medicinal effects of tropane alkaloids come from the diversity of functional modifications of the tropane ring, which can be classified into six different TA derivatives (J.-P. Huang et al., 2021) [Figure 4].

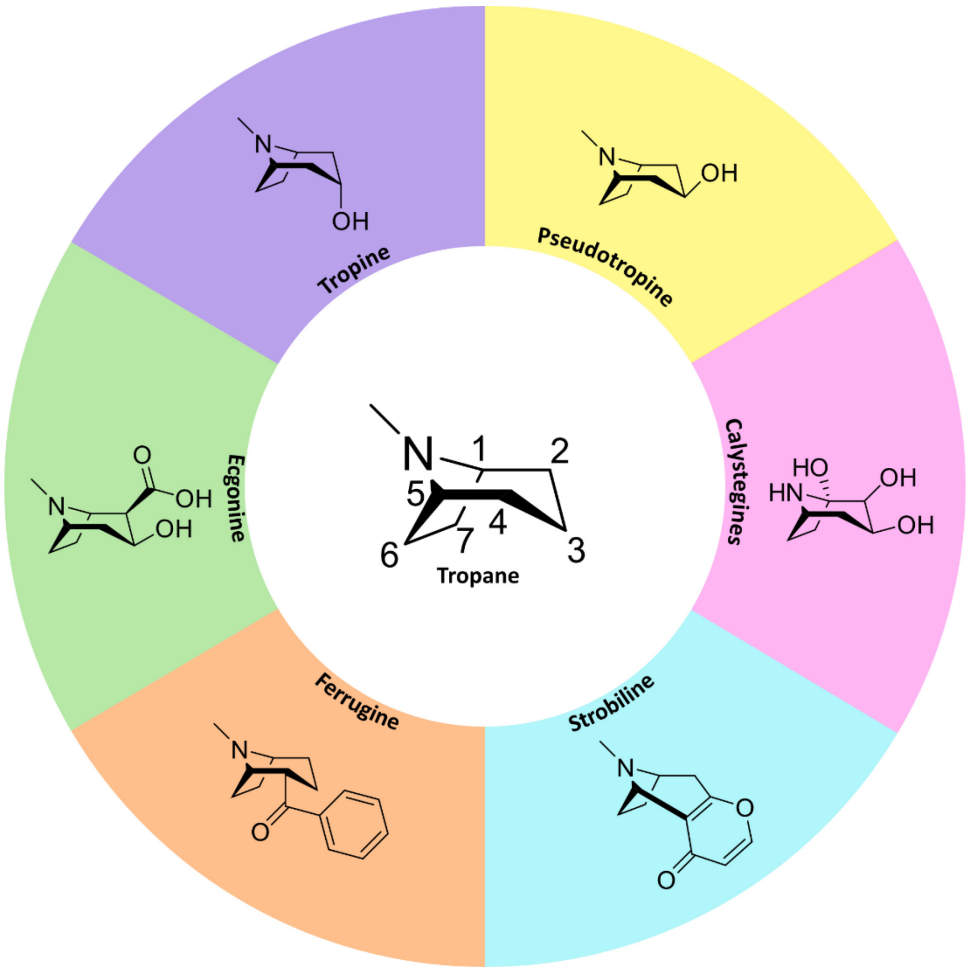

Figure 4: The diversity of tropane alkaloids in plants derived from the tropane core structure.

During the 19th century, chemists were particularly interested in isolating compounds from plants, and tropane alkaloids were of particular interest due to their medicinal properties [Figure 4]. Atropine, from *Atropa belladonna,* was first isolated by Geiger and Hesse in 1833 (Geiger & Hesse, 1833a, 1833b). At the beginning of the 20th century, Richard Willstätter synthesized tropidine, which he also converted into tropine (Willstätter, 1901a, 1901b). In 1917, Robert Robinson developed a successful chemical synthesis method to produce tropinone, a precursor to most tropane alkaloids, opening avenues for making high-value TAs for medicinal applications (Robinson, 1917). Robinson's one-pot synthesis used succinaldehyde, acetonedicarboxylic acid (3-oxoglutaric acid), and methylamine as reagents to

produce tropinone at an astonishing yield of 42%, along with hygrine and cuscohygrine as side products [Figure 5].

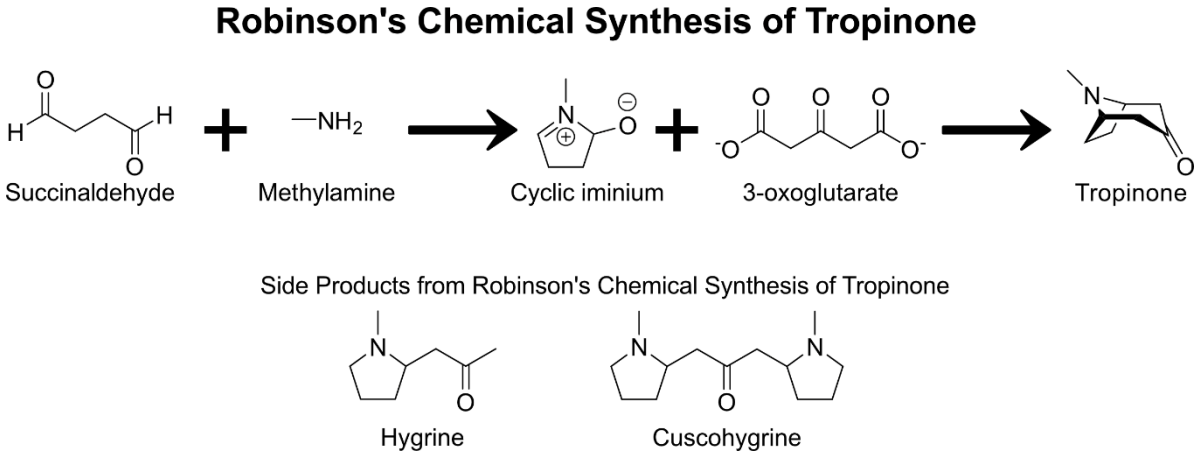

Figure 5: Simplified overview of the chemical synthesis of tropinone as described by Robert Robinson in 1917.

The formation of hygrine and cuscohygrine is not enzymatically controlled but rather a nonenzymatic chemical reaction that occurs when samples are either in long-term storage or exposure to UV light. Hygrine and cuscohygrine are commonly found in plants that produce tropane alkaloids, providing evidence that Robinson's mechanisms of chemical synthesis may share similar reactions that might occur in tropane alkaloid biosynthesis within plants (Johnson, 1995; O'Donovan & Keogh, 1969; Rubio et al., 2016). Robinson's hypothesis provided chemists and biochemists with a foundational building block to help unlock the critical steps of TA biosynthesis in plants (J.-P. Huang et al., 2021).

Uncovering tropane alkaloid biosynthesis during the 20th century relied on two eras of pathway discovery to help understand the biological precursor molecules involved. The first era was the precursor feeding and isotopic labeling of key metabolites behind tropane alkaloid biosynthesis. Scientists learned the necessary steps of TA biosynthesis by observing where these labeled metabolite precursors would accumulate in the plants. These feeding studies discovered that tropane alkaloids are derived from the amino acid arginine or the nonstandard amino acid ornithine. The molecular biology era allowed gene discovery and protein expression to validate critical candidates in the tropane alkaloid pathway. The characterized proteins involved in the TA biosynthetic pathway could be studied further to understand protein structure and function *in vitro*. The molecular biology era allowed genes to be identified in tropane alkaloid biosynthesis [Figure 6].

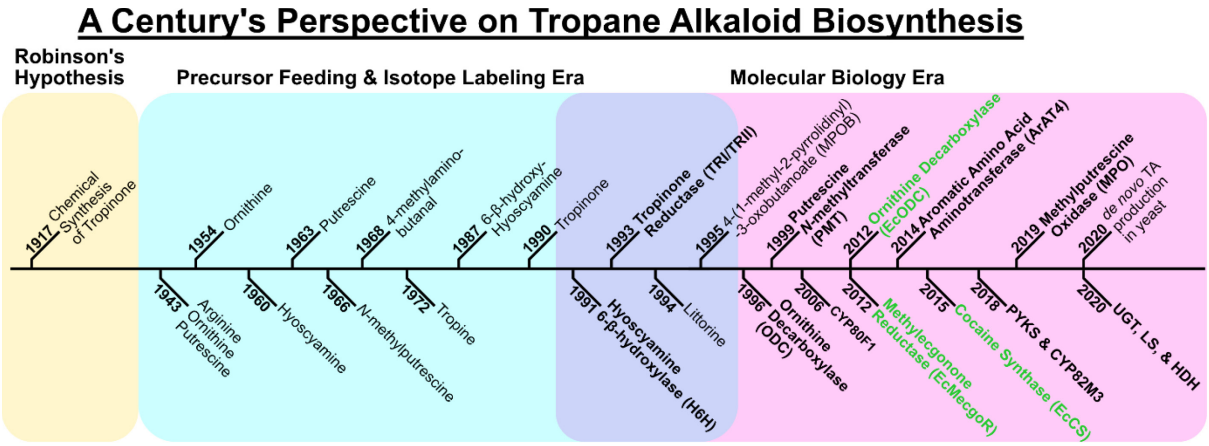

Figure 6: A timeline overview of tropane alkaloid pathway discovery over a century from the 1900s until the modern day. In the following eras, the chemical synthesis of tropinone in yellow color, precursor feeding and isotope labeling studies in light blue, the molecular biology era in pink, and the overlap of precursor feeding and isotope labeling studies along with the molecular biology era in light purple. Enzymes discovered are bold, and metabolites are unbolded. Enzymes isolated and characterized from *E. coca* are colored green. Modified with permission from Huang et al., 2021; Page 3; Figure 2.

Tropane alkaloid biosynthetic pathways in the Solanaceae family have been well explored and characterized. In comparison, tropane alkaloid biosynthesis in Erythroxylaceae has not been as well characterized as the Solanaceae plant family (Paul et al., 2021). However, several key advancements have been recently made in biosynthetic pathway discovery over the last decade. The assembly of the first publicly available *Erythroxylum coca var. novogranatense* genome has provided further insights into TA biosynthesis (Kim, 2020; Tian et al., 2022; Y.-J. Wang et al., 2022, 2023). Our work uncovered the missing steps of TA biosynthesis in *Erythroxylum coca* var. *novogranatense*. Our pathway discoveries in *E.* coca occurred before genomic datasets on coca were publicly available, relying entirely on transcriptome mining of genes of interest (Chavez et al., 2022). These discoveries helped distinguish tropane alkaloid biosynthesis's evolutionary divergence between the Solanaceae and Erythroxylaceae plant families.

1.2 Background of *Erythroxylum coca*

The tropical, flowering shrubs and trees in the *Erythroxylum* genus contain around 230 species throughout Central America, South America, and the Caribbean. *Erythroxylum coca* var. *coca* is the primary cultivar used in the illicit production of cocaine (Restrepo et al., 2019) [Figure 7]. *Erythroxylum coca* has been cultivated for at least 8,000 years, mainly in South America (Dillehay et al., 2010; Plowman, 1984). *E. coca,* colloquially referred to as coca originates from the Aymaran word "Khoka," which means "the tree" (Biondich & Joslin, 2016). The Aymara people are an indigenous population native to the Andes and the Altiplano

regions of South America. Coca leaves serve the Aymaran culture's medicinal and spiritual roles (Eisenberg, 2013). Coca is commonly used for medicinal properties to provide relief from cold, exhaustion, pain, and even hunger (Stolberg, 2011). In Western culture, cocaine became a prominently abused narcotic due to its addictive properties (Trifilieff & Martinez, 2014).

Figure 7: *Erythroxylum coca*. Dr. Danny Kessler provided the image from the Max Planck Institute for Chemical Ecology in Jena, Germany.

Tropane alkaloid biosynthesis in *E. coca* is known to be localized in the aerial tissues, which differs from the TA biosynthesis in Solanaceae, which occurs mainly in root tissues. The development of the leaf stages of *E. coca* proceeded in the following chronological order: buds, L1 rolled, L2 unrolled, and L3 as mature leaves (Docimo et al., 2013). The highest cocaine content in the buds and young leaves can be as high as 0.77% of the total dry weight of the leaf, depending on the coca cultivar. Interestingly, no significant differences were observed between laboratory and field-grown Amazonian coca plants, suggesting that TA biosynthesis is regulated at the genetic level in *E. coca* (Restrepo et al., 2019). Traditionally, in TA-producing plants, the influence of plant hormones such as salicylic acid or jasmonic acid causes an upregulation in the production of tropane alkaloids. However, these hormones seem to have little influence on the upregulation of TAs in *Erythroxylum coca* (Docimo et al., 2015).

Coca-leaf chewing allows cocaine to be absorbed by forming a quid, an alkaline mixture to help extract the alkaloids, in the cheek. A quid uses lime or any other alkaline substances in combination with the plant material in which little chewing of the coca leaves is involved. This process of ingesting coca leaves allows for the extraction of the alkaloids along the mucosal membrane of the mouth; ultimately, these tropane alkaloids cross into the brain to enact their

pharmacological effects. The use of coca leaves significantly reduced the effects of high-altitude sickness. In fact, Incan messengers used coca to increase endurance during long-distance running at high altitudes (Stolberg, 2011). In ancient Incan culture, coca leaves were considered sacred and only given to the elite members. Incan priest decreed that coca was a divine gift from the Sun God, Inti (Gay et al., 1975). Chewing of the coca leaves was restricted to only the highest members of society, mainly the ruling class and priests.

Two species of coca, *Erythroxylum coca* and *Erythroxylum novogranatense,* are widely known to produce the notorious narcotic cocaine and are observed in Colombia, Bolivia, Ecuador, and Peru (Restrepo et al., 2019). Both *Erythroxylum coca* and *Erythroxylum novogranatense* are referred to as coca, and these plants grow at altitudes between 500 and 1500 meters above sea level. *E. novogranatense* is well known for its methyl salicylate, which gives the plant a minty aroma when crushing the leaves (Plowman, 1984). In the 19th century, cocaine was most notably present in products like Coca-Cola, although nowadays, the plant extracts from coca remove cocaine before being added to products like Coca-Cola due to their toxic and addictive effects (Gootenberg, 2002). Coca plants remain culturally and economically significant despite the stigmatized nature around these plants (M. Islam, 2011).

1.3 Diverse Roles of Polyamines in Plants

Polyamines (PAs) are a group of ubiquitous compounds containing two or more amines that play crucial roles in various aspects of plant biology. These small, aliphatic, and positively charged molecules are found in all living organisms, including bacteria, fungi, animals, and plants. Polyamines exert their biological effects by interacting with negatively charged cellular components, such as DNA, RNA, lipids, and proteins. They are implicated in numerous cellular processes, including membrane stabilization, DNA replication, buffering against reactive oxygen species (ROS), and stress-related responses. PAs are divided into diamines, triamines, and tetramines based on the number of amino groups in the molecule. The three common polyamines found in plants are putrescine (Put), spermidine (Spd), and spermine (Spm). These polyamines are grouped into diamine, triamines, and tetraamines [Figure 8]. Polyamines can exist in several forms, such as unbound, covalently attached to other small molecules, or bound to cell walls (Hura et al., 2015; Luo et al., 2009; Macoy et al., 2015).

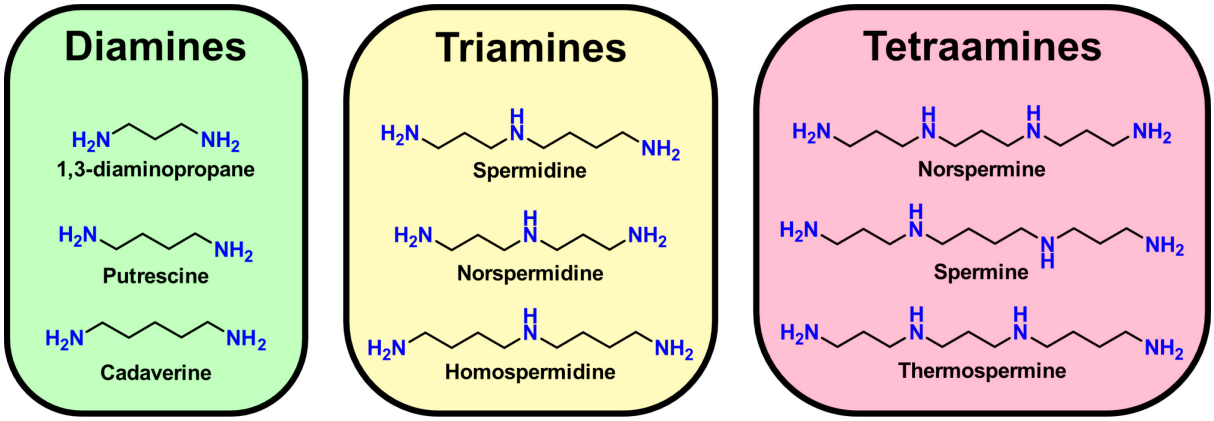

Figure 8: Chemical structure of polyamines observed in plants. These polyamines are divided into diamine, triamine, and tetraamine groups.

PAs additionally have an interaction with phytohormones from plants in which polyamines appear to have a positive influence on cytokinins and gibberellic acids while having a negative feedback on ethylene and jasmonic acid hormones (Napieraj et al., 2023). These interactions can alter the conformation and activity of these macromolecules, leading to various cellular responses. Other specialized polyamines, like thermospermine (Tspm) and cadaverine (Cad), are also found in plants and play important roles in plant growth and development (Gibbs et al., 2021; Jancewicz et al., 2016; Takano et al., 2012). Polyamines regulate growth, development, stress tolerance, and defense against pathogens, contributing to the overall resilience and adaptability of plants in various environmental conditions (Alcázar et al., 2020; Tiburcio et al., 2014). Polyamines, while essential in primary metabolism, also play a vital role in the secondary metabolism of plants.

1.3.1 Putrescine Biosynthesis

Polyamine biosynthesis begins with the formation of the metabolite putrescine, which is a core precursor in classical polyamines. Putrescine can be formed from arginine or ornithine (Hanfrey et al., 2001; Majumdar et al., 2013, 2016). Both arginine and ornithine are intermediates involved in the urea cycle, although some plants prefer either arginine or ornithine as precursors for putrescine formation. The biosynthetic pathway of arginine into putrescine requires a three-step enzymatic process. First, arginine must undergo a decarboxylation reaction mediated by arginine decarboxylase (ADC), forming agmatine as a product (Hashimoto et al., 1989). Next, agmatine iminohydrolase (AIH) converts the agmatine into *N*-carbamoylputrescine (Janowitz et al., 2003; Sekula & Dauter, 2019b). Finally, *N*-carbamoylputrescine amidohydrolase (NCPAH) converts the *N*-carbamoylputrescine into putrescine (Piotrowski et al., 2003).

ADC utilizes a pyridoxal-5'-phosphate (PLP) cofactor to facilitate the decarboxylation reaction of arginine to form agmatine [Figure 9]. Arginine decarboxylase is not as prevalent in mammals and some eukaryotes, although some tissue-specific expression of ADC in mammals is known to occur (Facchini, 2001; X. Wang et al., 2014; Zhu et al., 2004). In *Arabidopsis thaliana*, ADC is essential for normal seed development (Urano et al., 2005). Arginases (ARGAH), which participate in the urea cycle, were also found to utilize agmatine as a substrate, demonstrating agmatinase activity. The substrate flexibility of ARGAH creates an additional route to form putrescine in plants (Patel et al., 2017). The enzyme arginase can also catalyze the removal of the guanidinium group on arginine to form ornithine in a one-step reaction. The biosynthetic pathway of putrescine from arginine is complex, especially when compared to the putrescine derived from ornithine.

The conversion of putrescine from ornithine is a one-step reaction with an ornithine decarboxylase (ODC) to form putrescine directly (Docimo et al., 2012; Kim et al., 2021) [Figure 10]. ODC requires PLP as a cofactor to facilitate the decarboxylation of ornithine. Putrescine derived from ornithine is crucial in providing polyamines for cellular division, differentiation, and development, while putrescine derived from arginine is more critical for cell expansion and environmental stress responses in plants (Docimo et al., 2012). Some organisms do not encode ornithine decarboxylases but instead rely on putrescine derived from arginine rather than ornithine (Hanfrey et al., 2001; Joshi et al., 2022; Upadhyay et al., 2021).

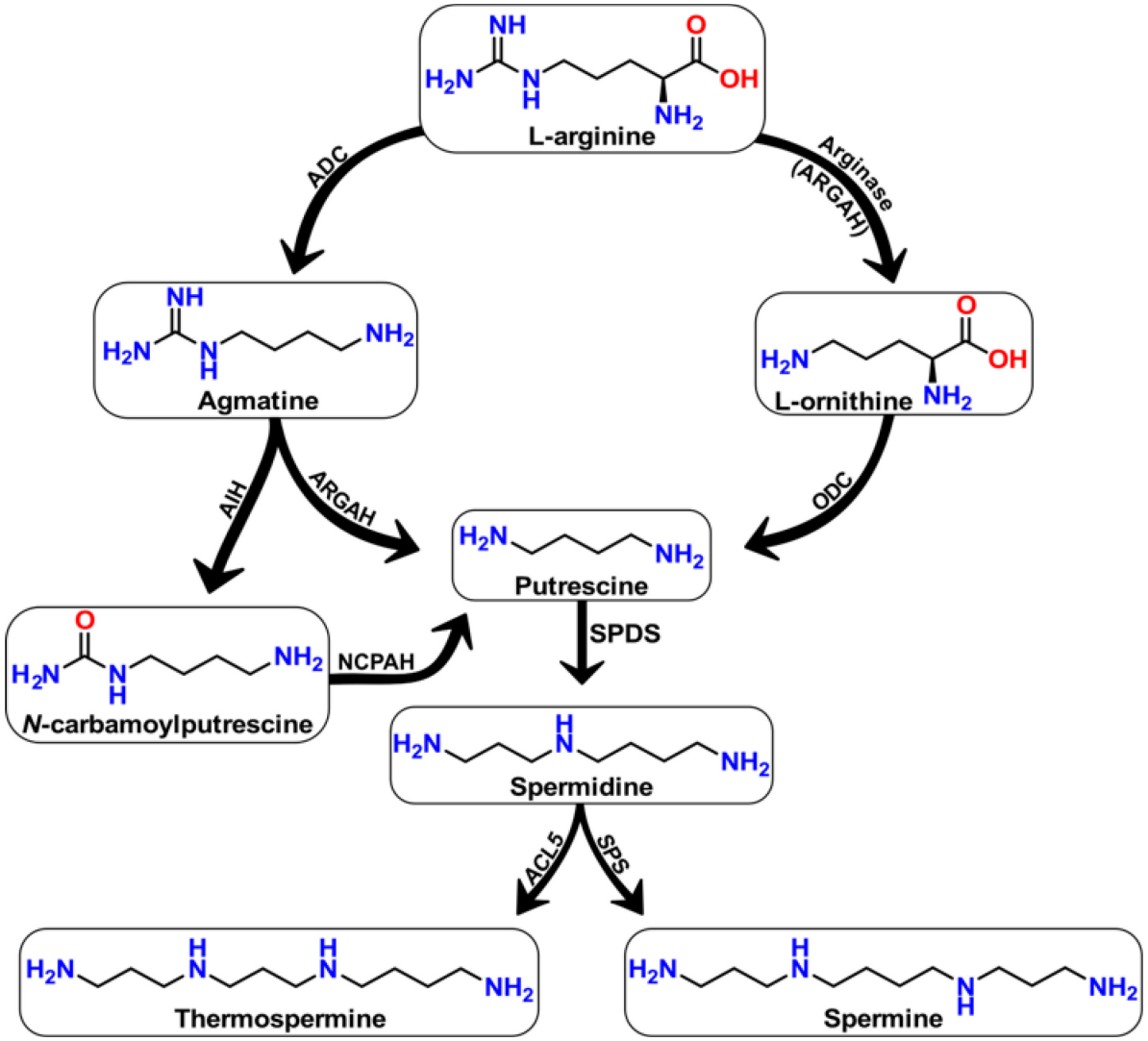

Figure 9: General polyamine biosynthetic pathway in plants. The following abbreviations are labeled as arginine decarboxylase (ADC), arginase (ARGAH), ornithine decarboxylase (ODC), agmatine iminohydrolase (AIH), *N*-carbamoylputrescine amidohydrolase (NCPAH), spermidine synthase (SPDS), spermine synthase (SPS), and thermospermine synthase (ACL5).

ODCs are known to be regulated with an antizyme that binds to ornithine decarboxylases, causing decreased enzyme activity, increasing degradation of ODC, and additionally represses polyamine uptake in fungi and metazoan cells, but this antizyme regulation is not observed in ODCs derived from plants (Chattopadhyay et al., 2011; Matsufuji et al., 1995; Porat et al., 2008). Polyamines, including putrescine, spermidine, and agmatine, had little effect on the overall activity of ODC expressed in tobacco cells (Illingworth & Michael, 2012). Plant ODCs seem to be missing this antizyme orthologue and, therefore, need to regulate polyamine biosynthesis differently than fungi and animals. In plants, putrescine production is rather intricate due to the biosynthetic pathways and organelle-specific

biosynthetic routes that can create putrescine pools that are likely destined for different cellular processes during growth and development.

Putrescine biosynthesis in plants is reported to be localized to the cytosol, mitochondria, and chloroplast (González-Hernández et al., 2022). A recent preprint challenges this traditional view on putrescine biosynthesis in plants (Joshi et al., 2022). The authors propose that putrescine biosynthesis is not present in the cytosol but instead localized to the chloroplast, mitochondria, and endoplasmic reticulum (ER) as the source of putrescine production. However, peer review is needed since this is currently an unreviewed preprint.

1.3.2 *S*-adenosyl-L-methionine Dependent Methyltransferases in Plants

S-adenosyl-L-methionine (SAM), also known as AdoMet, is a cofactor commonly observed in methyltransferase reactions. SAM is formed from L-methionine via the SAM cycle. Methionine first undergoes an ATP-dependent reaction with SAM synthetase known as methionine adenosyltransferase (MAT), which generates *S*-adenosyl-L-methionine, which contains a positively charged sulfur atom. The positive charge of the sulfur atom gives the methyl group a strong electrophilic character that participates in the nucleophilic reaction within the active site of SAM methyltransferase enzymes. *S*-adenosyl-L-methionine is used by three major families of SAM-dependent methyltransferases, such as *O*-, *N*-, *S*-, and *C*-methyltransferases, and are involved in creating a chemical diversity observed in both primary and specialized metabolism (Fontecave et al., 2004; Roje, 2006). SAM-dependent methyltransferases exhibit great flexibility, even having five different structural folds defined as classes I-V, and are known to methylate a diverse set of substrates (Schubert et al., 2003). Examples of this methyltransferase substrate flexibility include DNA methylation, protein methylation, and metabolite methylation, to name a few (D'Auria et al., 2003; Gallego-Bartolomé, 2020; S. Kumar & Mohapatra, 2021; Serre et al., 2018, 2020). *S*-adenosyl-L-methionine is the second most commonly used enzyme substrate, following adenosine triphosphate (ATP), the most prominent substrate in the cell (Schubert et al., 2003). In addition to SAM being a key player in methyltransferase reactions, it is also used to form decarboxylated S-adenosyl-L-methionine (dcSAM), an integral cofactor involved in polyamine biosynthesis reactions.

1.3.3 dcSAM is an Essential Cofactor for Polyamine Biosynthesis

An essential step for polyamine biosynthesis begins with the decarboxylation of *S*-adenosyl-L-methionine via *S*-adenosyl-L-methionine decarboxylase (SAMDC). SAMDC

enzymes have a unique posttranslational modification in which the enzyme self-cleaves the peptide backbone at the SESS peptide motif. SAMDC, in the uncleaved form, is an inactive enzyme known as a proenzyme. Upon the autocatalytic cleavage of SAMDC, it forms two subunits: the α-subunit, which retains the original *N*-terminus, and the β-subunit, which retains the original *C*-terminus. During this autocatalytic cleavage, a pyruvate group is covalently attached to a serine residue on the α-subunit and acts as a cofactor that facilitates the decarboxylation reactions of SAM to dcSAM. SAMDC is a critical rate-limiting step of polyamine biosynthesis (Carbonell & Blázquez, 2009). The mechanism of SAMDC proenzyme to a fully active SAMDC appears to be widely conserved when comparing mammalian and plant SAMDCs (Xiong et al., 1997). Interestingly, SAMDC enzymes in plants are monomeric, unlike the dimeric forms of SAMDC enzymes found in mammals. Plant SAMDCs do not seem to be as tightly regulated by the presence of putrescine compared to SAMDCs found in yeast and humans (Bennett et al., 2002).

SAMDCs found in plants are highly regulated at the transcriptomic and posttranscriptional levels (W.-W. Hu et al., 2005). Overexpression of the SAMDC gene was lethal to potato plants, most likely due to the depletion of SAM (A. Kumar et al., 1996). In tomato plants, overexpression of SAM synthetase led to polyamine accumulation and a more robust root system under alkali stress conditions (Gong et al., 2014). Knockdown of both BUD2 and SAMDC1 genes, which encode SAMDC in *Arabidopsis thaliana,* were embryo-lethal (Ge et al., 2006). The adverse growth defects indicate that a delicate balance of dcSAM levels must be maintained for proper growth and development. Once the SAMDC enzyme is mature, it can form decarboxylated SAM, a crucial co-substrate for polyamine chain elongation reactions mediated by the aminopropyltransferase enzyme family [Figure 10].

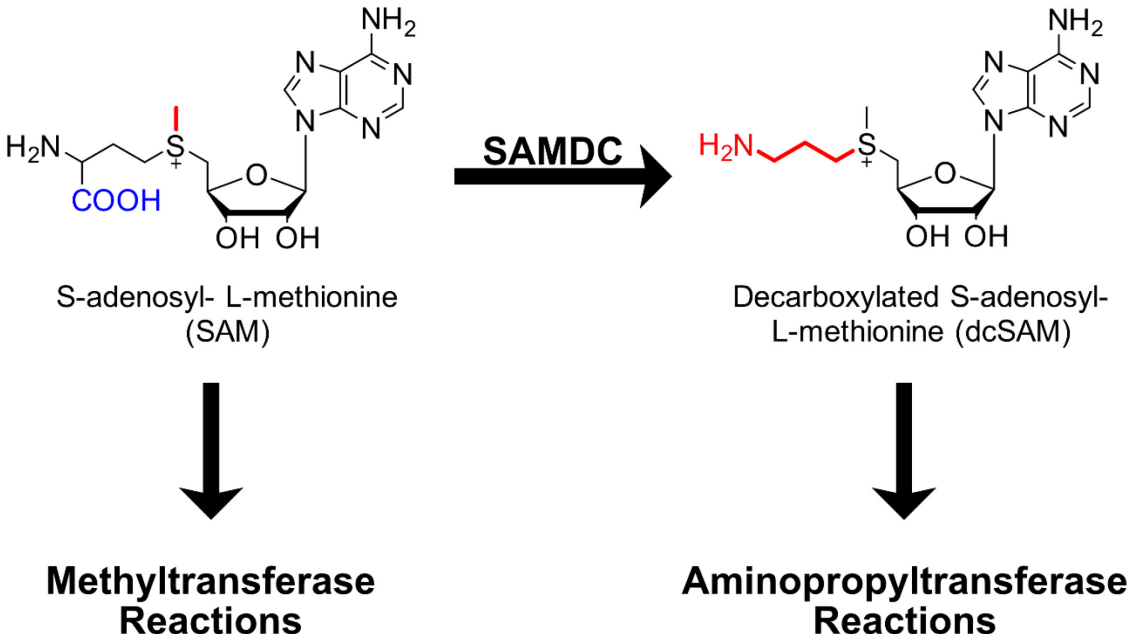

Figure 10: Chemical structure of SAM and dcSAM. The red coloration indicates the functional group transferred to a substrate, either a methyl or an aminopropyl group, depending on the cofactor.

1.3.4 Polyamine Elongation via Aminopropyltransferases

Spermidine synthases (SPDS) and spermine synthases (SPS) are members of the aminopropyltransferase (APT) family of enzymes that utilize decarboxylated *S*-adenosyl-L-methionine (dcSAM) cofactor to transfer an aminopropyl group to a polyamine substrate causing a chain elongation of the polyamine. Spermidine synthase requires putrescine and dcSAM to form spermidine and 5′-deoxy-5′-methylthioadenosine (MTA). Spermidine can undergo another round of chain elongation with the enzyme spermine synthase, producing spermine as a byproduct. Specialized polyamines, such as thermospermine, a constitutional isomer of spermine, are essential for plant growth and development (Kakehi et al., 2008; Takano et al., 2012). All SPDS enzymes evolved from a common ancestor before the split between prokaryotes and eukaryotes. Unlike spermidine synthases, spermine synthases evolved approximately three times independently in fungi, animals, and plants (Minguet et al., 2008). Aminopropyltransferases have an important dcSAM binding motif consisting of a GGGDG peptide sequence. The aspartate residue in this peptide motif is reported to create steric clashes with the carboxyl group on SAM, only allowing dcSAM to bind (Ikeguchi et al., 2006).

Localization of spermidine synthases was initially thought to be cytosolic and chloroplastic. Transient expression of the two SPDS isoforms from *Arabidopsis thaliana* revealed that AtSPDS2 was localized primarily in the nucleus while AtSPDS1 was localized

to both the cytoplasm and the nucleus in *Nicotiana benthamiana* (Belda-Palazón et al., 2012). The authors state that AtSPDS1 and AtSPDS2 possess no obvious nuclear localization signal (NLS); both isoforms can interact to form heteromultimers in *Arabidopsis thaliana* when localized in the nucleus (Panicot et al., 2002; Sekula & Dauter, 2019a). Further investigation is needed to understand how the nuclear import of SPDSs is operating in plants. Deletion of both isoforms of SPDS genes in *Arabidopsis* led to severe growth defects, abnormally shrunken seeds, and embryos arrested morphologically at the heart-torpedo transition stage of development (Imai, Matsuyama, et al., 2004). Mutants in *Arabidopsis thaliana* showed that a double mutant knockout of SPS and ACL5 genes could generate viable plants, indicating that spermine is not essential for the survival of *Arabidopsis*. A knockout of ACL5, known as thermospermine synthase, showed a mutant phenotype of reduced stem growth (Imai, Akiyama, et al., 2004). This reduced stem phenotype is known to occur with disruption of thermospermine biosynthesis, as thermospermine is involved with stem elongation in plants (Kakehi et al., 2008; Takano et al., 2012).

Polyamine biosynthesis is highly regulated across many organisms, including bacteria, yeast, mammals, and plants. Growth defects are commonly observed when polyamine metabolism is disrupted by gene knockout or overexpression. In *E. coli*, excess spermidine is cytotoxic when the spermidine acetyltransferase (SAT) gene is disrupted, indicating that SAT is crucial to control and prevent spermidine levels from being in excess amounts that could become toxic for the cell (V. Kumar et al., 2022; Sakamoto et al., 2020). In *Saccharomyces cerevisiae,* spermidine deficiency increases ribosomal frameshifting efficiency while inhibiting retrotransposition activity (Balasundaram, Dinman, et al., 1994). Overexpression of SAMDC in yeast mutants that lack a functional ODC enzyme displayed polyamine deficiency along with inhibited growth, indicating that excess dcSAM may have a toxic effect on amine-deficient cells (Balasundaram, Xie, et al., 1994). An intricate balance of polyamine levels is essential to maintain and regulate cellular homeostasis regardless of prokaryotic or eukaryotic cell origins.

1.4 Putrescine *N*-methyltransferases

Putrescine *N*-methyltransferases are the first committed step of nicotine and tropane alkaloid biosynthesis in Solanaceae and Convolvulaceae (Biastoff, Brandt, et al., 2009; Teuber et al., 2007). Putrescine is directly methylated via a putrescine *N*-methyltransferase (PMT) enzyme [Figure 11]. PMTs do not share sequence identity with other plant methyltransferases

except for the conserved SAM binding motif (D'Auria et al., 2003). PMT enzymes are related to spermidine synthases, but instead of using dcSAM as a cofactor, PMT enzymes have neo-functionalized a SAM binding domain, allowing for the selection of methyltransferase activity instead of aminopropyltransferase activity that SPDS enzymes exhibit.

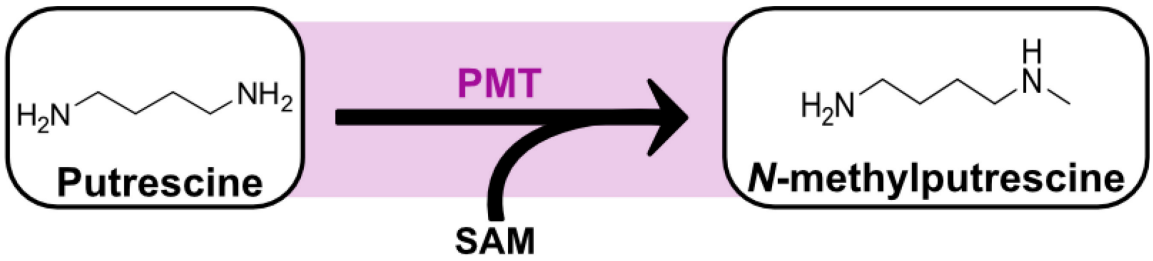

Figure 11: Putrescine *N*-methyltransferase reaction using putrescine and *S*-adenosyl-L-methionine (SAM) to generate *N*-methylputrescine as a product.

Tissue-specific expression of secondary metabolism is widely observed in plants. PMT enzymes were first isolated in the root tissue of tobacco due to mutant lines of tobacco plants that were deficient in nicotine (Hibi et al., 1994). For tropane alkaloid biosynthesis in Solanaceous plants, putrescine *N*-methyltransferase has higher levels of expression in the roots of *Atropa belladonna*, *Datura stramonium*, *Hyoscyamus niger*, and *Scopolia lurida*. Some exceptions exist, like *Anisodus acutangulus,* where lower levels of PMT expression are present in leaves and stem tissues (Kai et al., 2009; Paul et al., 2021). PMT gene expression is also inducible with the hormone methyl jasmonate (MeJA). Further investigations into this tissue-specific expression were studied using hairy root cultures from *Agrobacterium rhizogenes*-mediated transformation. Attempts to overexpress PMT in *A. belladonna* by itself increased *N*-methylputrescine, although it did not increase the tropane alkaloid content compared with the control (Rothe et al., 2003).

The *N*-terminal domain of both PMT and SPDS from *Datura stramonium* is involved in the catalytic activity and substrate specificity between these two enzymes (Biastoff, Reinhardt, et al., 2009). PMT enzymes possess a GGGIG motif that participates in the binding of SAM, which differs slightly from the GGGDG motif involved in dcSAM, which is commonly found in aminopropyltransferase enzymes like spermidine or spermine synthases. The change of an aspartate residue to an isoleucine residue in this motif is thought to stabilize the binding of SAM over the dcSAM, creating a change from aminopropyltransferase to methyltransferase activity. Junker et al. demonstrated that a singular change of D103I in AtSPDS1 from *Arabidopsis thaliana* abolished spermidine synthase activity and subsequently detected minimal putrescine *N*-methyltransferase activity. The presence of PMT activity shows

that it is possible to gain PMT activity from SPDS. However, the PMT activity was barely detectable, indicating that a single amino acid substitution is insufficient to create an efficient PMT enzyme from SPDS. This study continued to modify the AtSPDS1 enzyme using additional site-directed mutagenesis, leading to a mutant enzyme that displayed spermidine synthase and putrescine *N*-methyltransferase activity, giving credence to the idea of neofunctionalization of an SPDS into a PMT enzyme from a couple of point mutations within the active site (Junker et al., 2013). To date, no functional PMT enzymes have been isolated and characterized in *E. coca,* suggesting that an alternate step must exist in place of the missing PMT activity.

1.5 Polyamine Oxidases

Polyamine oxidases (PAOs) are a group of enzymes that catalyze the oxidative degradation of polyamines, a group of positively charged organic molecules with essential roles in plant growth, development, and stress responses. In plants, PAOs are involved in various physiological processes, including regulation of polyamine levels, defense against pathogens, photoprotection, and stress tolerance. There are two main types of PAOs in plants: back-conversion type PAOs (BC-type) and terminal catabolic type PAOs (TC-type). BC-type polyamine oxidases catalyze the conversion of spermidine and spermine to putrescine, the simplest polyamine. This back conversion of spermidine into putrescine via flavin-dependent oxidase activity produces putrescine, 4-aminopropanal, hydrogen peroxide (H_2O_2), and ammonia as products [Figure 12A]. TC-type PAOs can degrade 1,3-diaminopropane (DAP) to 3-aminopropanal, hydrogen peroxide, and ammonia [Figure 12B]. Polyamine oxidases have been implicated in plant tissue differentiation and organ development processes (Fincato et al., 2012; Tavladoraki et al., 2016).

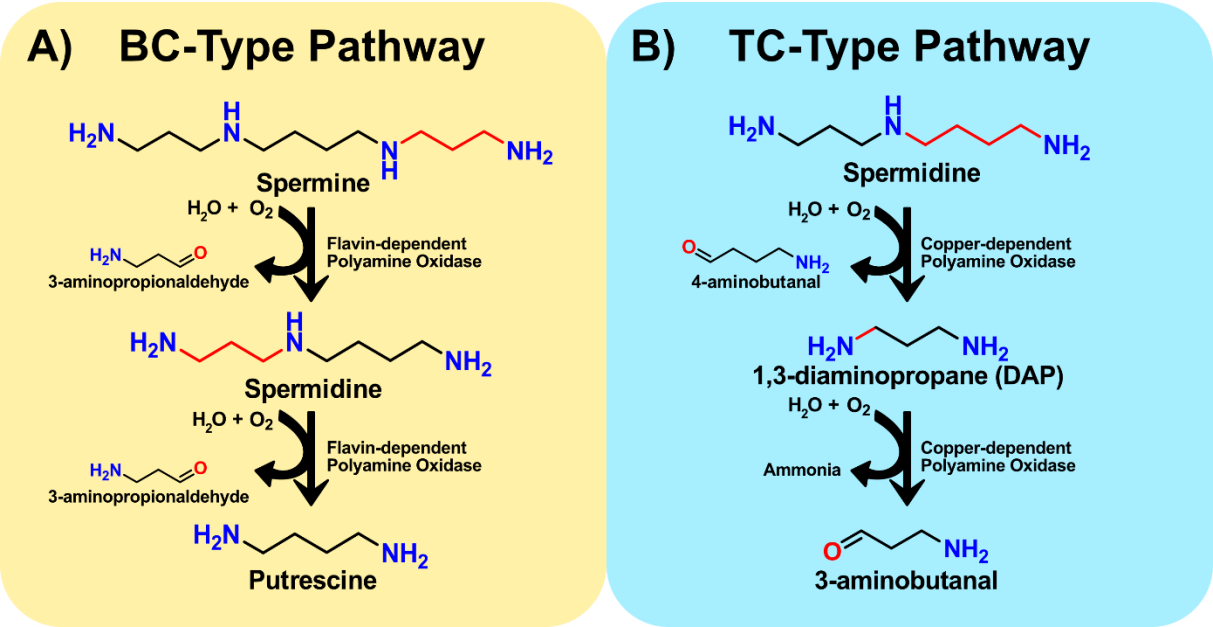

Figure 12: Simplified view of back conversion pathway (A) via a flavin-dependent oxidase and terminal catabolic pathway (B) via a copper-dependent polyamine oxidase . The red color on the carbon chains indicates the functional group removed by the flavin-dependent or copper-dependent polyamine oxidase.

Copper-dependent amine oxidases (AOCs) participate in polyamines' terminal catabolic (TC-type) pathway, meaning the polyamine's terminal end is oxidized. Flavin-dependent amine oxidases (AOFs) participate in the back conversion (BC-type) pathway and are usually involved in shortening polyamines (Yu et al., 2019). AOF enzymes are generally involved in the oxidation of secondary amino groups, while AOC enzymes oxidize primary amines. Plant AOF enzymes are known to localize in the cytosol, apoplast, peroxisome, and vacuole. Apoplastic AOFs generally oxidize the carbon at the *endo*-side of the N4 atom of spermidine and spermine substrates (Cervelli et al., 2001; T. Liu et al., 2014). When the oxidation on the *endo*-side of the N4 atom occurs, it makes products that cannot be directly converted back into polyamines like putrescine or spermidine.

An example is the OsPAO7 from *Oryza sativa,* which takes spermidine or spermine and generates 1,3-diaminopropane (DAP) (T. Liu et al., 2014). DAP cannot be used as a substrate for standard polyamine biosynthesis and can even act as a suicide substrate against the activity of copper-dependent amine oxidase enzymes (Awal & Hirasawa, 1995). However, when the carbon atom oxidizes on the *exo*-side of spermidine or spermine, it generates products that can participate in polyamine biosynthesis, such as putrescine or spermidine, respectively.

Polyamine oxidases are found in several compartments of plant cells, such as the peroxisome, apoplast, vacuole, and cytosol (Benkő et al., 2022; Tavladoraki et al., 2016).

Various environmental factors, including light, nutrient availability, and stress conditions, regulate the expression of PAOs. In response to stress, the expression of PAOs is often increased, leading to a decrease in polyamine levels and an increase in ROS production. *N*-methylputrescine oxidase (MPO), a key enzyme involved in both nicotine and tropane alkaloid biosynthesis, is known to cluster in clade III of other copper-dependent amine oxidases (Naconsie et al., 2014; Tavladoraki et al., 2016). AOC enzymes that cluster in clade III generally prefer shorter polyamines like putrescine and cadaverine over longer polyamines such as spermidine and spermine.

1.6 First Ring Closure of Tropane Alkaloids

In tropane alkaloid biosynthesis, the first ring closure is when *N*-methylputrescine is cyclized into *N*-methyl-Δ^1-pyrrolinium cation (NMPy) through an amine oxidation reaction. The family of enzymes performing these reactions is a copper amine oxidases (AOC) member called *N*-methylputrescine oxidases (MPO) in Solanaceae. The methyl group present in *N*-methylputrescine is speculated to act as a protecting group from the MPO enzyme oxidation, allowing only the primary amine end to be efficiently oxidized. After the oxidation of *N*-methylputrescine, the reactive aldehyde product subsequently forms NMPy by non-enzymatic intramolecular cyclization through a retro-Michael addition (Chavez et al., 2022; Kim et al., 2021) [Figure 13]. This spontaneous cyclization reaction is also observed in *Atropa belladonna*, *Hyoscyamus niger*, *Datura stramonium*, and other solanaceous plants that produce tropane alkaloids. *N*-methylputrescine oxidases are known to cluster in Clade III of copper-dependent polyamine oxidases (Tavladoraki et al., 2016). AOC enzymes can oxidize important intermediates of nicotine and tropane alkaloid biosynthesis like *N*-methylputrescine but at a much lower efficiency than the methylputrescine oxidase (MPO), which has a higher specificity for *N*-methylputrescine over other polyamines. MPO are crucial enzymes as they mediate the steps that lead to first ring closure in tropane and nicotine biosynthesis.

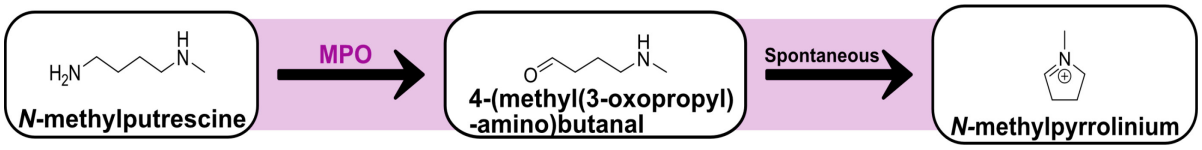

Figure 13: Depiction of methylputrescine oxidase (MPO) activity leads to *N*-methylpyrrolinium cation formation.

1.7 Second Ring Closure of Tropane Alkaloids

The second ring closure of tropane alkaloids occurs in a series of enzymatic steps to take NMPy and form the second ring closure present in tropane alkaloids. The first step after

forming NMPy is a spontaneous reaction with 3-oxoglutarate, also known as beta-ketoglutarate. The source of this 3-oxoglutarate is pyrrolidine ketide synthase (PYKS), also known as 3-oxoglutaric acid synthase (OGAS), which uses two rounds of malonyl-CoA to form 3-oxoglutarate. Spontaneously, 3-oxoglutarate and NMPy react with each other to form 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate (MPOB) [Figure 14] (Bedewitz et al., 2018; J.-P. Huang et al., 2019; Kim, 2020; Tian et al., 2022). MPOB can also spontaneously decarboxylate into hygrine, indicating that this molecule is unstable, a characteristic observed in other β-keto acids (Lohman et al., 2019).

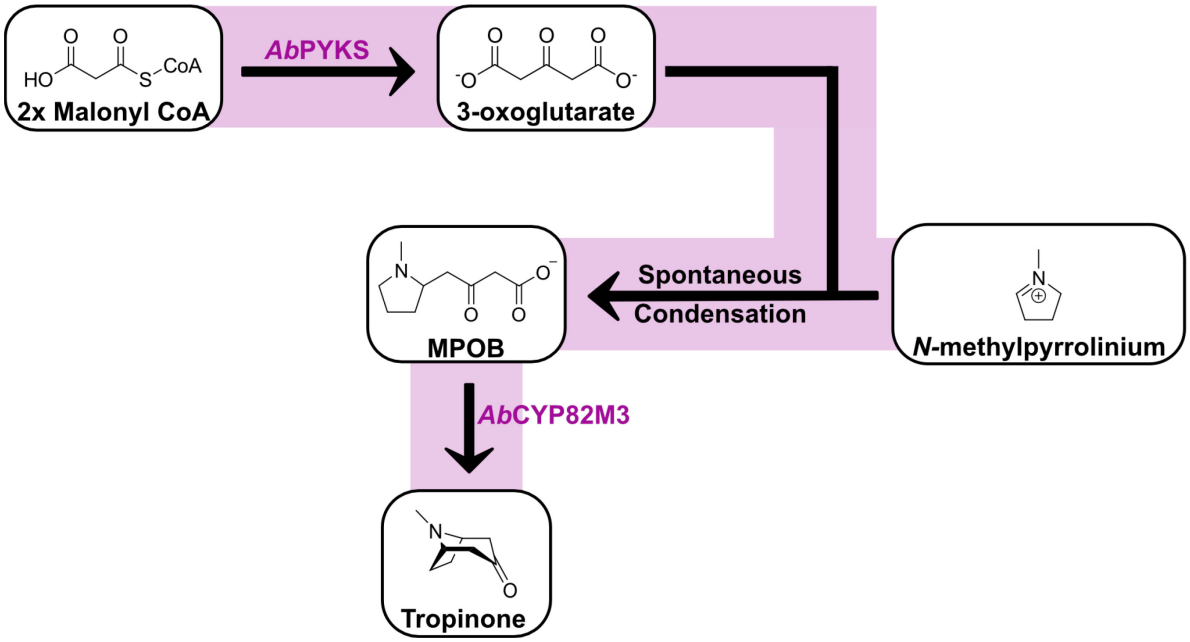

Figure 14: Second ring closure of tropane rings in *Atropa belladonna*. A spontaneous condensation of *N*-methylpyrrolinium cation and 3-oxoglutarate yield 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate (MPOB), which is cyclized via CYP82M3 mediated reaction to form the tropinone.

MPOB becomes a substrate for a cytochrome P450 enzyme that facilitates the second ring closure of MPOB to form tropinone. Cytochrome P450s are one of the most prevalent and diverse enzyme families known. In general, CYP genes represent approximately ~1% of the protein-encoding genes in the genomes of plants (Chakraborty et al., 2023). In *A. belladonna,* tropinone synthase (AbCYP82M3) performs the second ring closure of MPOB into tropinone (Bedewitz et al., 2018).

1.8 Tropinone Reductases

The diversity of tropane alkaloids is due to the functional groups that decorate the tropane ring. The addition of an ester functional group at the C3 carbonyl group on the tropane ring is observed in both Solanaceae and Erythroxylaceae species. Tropinone reductase (TRs)

enzymes control the stereospecificity of the C3 hydroxyl group, forming either α-tropine or β-tropine mediated by TR-I and TR-II, respectively. The tropinone reductase I enzyme generates α-tropine (tropine), while the tropinone reductase II enzyme generates β-tropine (pseudotropine) at the C3 position on the tropane ring. TR-I and TR-II are found in members of the Solanaceae and Convolvulaceae family. Crystal structure analysis of TR-I and TR-II enzymes from *Datura stramonium* shows that the amino acid identity between the two reductases is 64% identical on an amino acid level, but the tertiary structure between the enzymes is almost identical. The orientation of the tropinone substrate binding within the active site mediates the stereospecificity between these two tropinone reductases (Kim et al., 2021; Nakajima et al., 1998).

TR-I and TR-II use nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor in reducing tropinone to α-tropine or β-tropine, respectively. TRs are members of the short-chain dehydrogenase/reductase (SDR) enzyme family, which contains the NSTK amino acid motif, also known as a catalytic tetrad (Kavanagh et al., 2008). An interesting difference between these tropinone reductases is that the reductase reaction mediated by TR-I is reversible, while TR-II reductase activity is irreversible. Using 8-thiabicyclo[3.2.1]octan-3-one (TBON), a tropinone mimic in which the nitrogen atom in tropinone is substituted for sulfur atom, TR-I would catalyze the reduction of TBON, but TR-II was unable to catalyze the reduction of TBON (Hashimoto et al., 1992). Generally, tropane alkaloids found in the Solanaceous plants are esterified at the C3 hydroxyl position, and the stereochemistry is commonly in the α-hydroxy conformation in common TAs like atropine and scopolamine (X. Hu et al., 2023).

Erythroxylum coca is known to produce tropane alkaloids that contain a β-hydroxy group at the C3 position almost exclusively. The enzyme performing this reductase activity is methylecgonone reductase (EcMecgoR) (Jirschitzka et al., 2012). Although EcMecgoR performs similar reactions to tropinone reductase, it belongs to a different enzyme family known as aldo-keto reductases (AKRs) [Figure 15]. AKRs that have been characterized contain a common α/β-barrel motif and use NADH or NADPH cofactors. Enzyme assays with tropinone reductases isolated from Solanaceous plants failed to have significant activity toward methylecgonone as a substrate (Couladis et al., 1991; Hashimoto et al., 1992). It appears that the only enzyme known to date that can efficiently reduce the keto group at the C3 position on methylecgonone is mediated by methylecgonone reductase. Additionally, EcMecgoR can catalyze the reverse reaction, leading to the oxidation of methylecgonine. However, the

conditions for doing this are not biologically relevant as it requires a pH optimum of 9.8, and EcMecgoR is thought to be localized in the cytosol where the pH is around physiological pH between 7.1-7.5 (Cosse & Seidel, 2021; Jirschitzka et al., 2012).

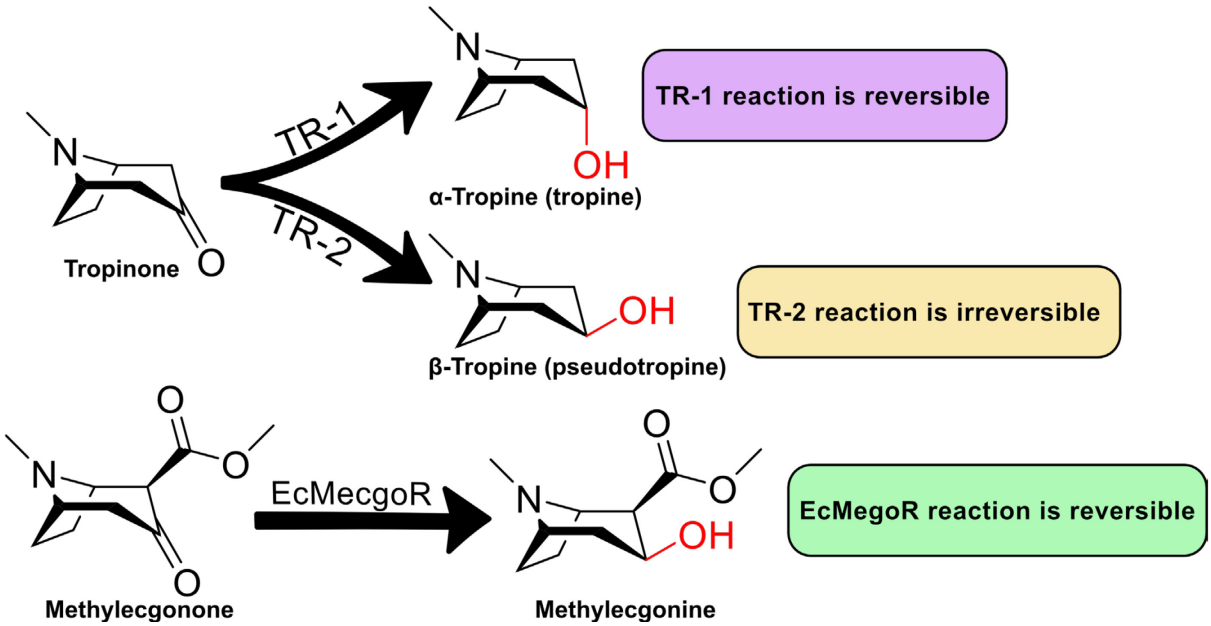

Figure 15: Tropinone reductase and methylecgonone reductase catalyzing the reduction of the C3 position of the tropane ring of either tropinone or methylecgonone, respectively. The stereochemistry of the reductase activity is highlighted in red. Acronyms are defined as tropinone reductase 1 (TR-1), tropinone reductase 2 (TR-2), and methylecgonone reductase (EcMecgoR).

1.9 Acyltransferase Enzymes Involved in TA Biosynthesis

The diversity of tropane alkaloids is due to the functional groups that decorate the tropane ring and give different bioactive properties. A common modification observed amongst TAs in Solanaceae and Erthroxylaceae is the addition of an acyl group, usually at the C3 hydroxyl group on the tropane ring. Acyltransferases are enzymes that use an energy-rich donor as a cofactor to transfer an acyl group to an acyl receptor substrate. Two enzyme families in plants are known to participate in acyltransferase reactions. Serine carboxy peptidase-like (SCPL) acyltransferases and BAHD acyltransferases can synthesize various plant metabolites (Bontpart et al., 2015).

SCPL acyltransferases belong to the α/β hydrolases superfamily, which utilizes an energy-rich donor, such as β-acetal glucose esters, also known as 1-*O*-β-D-glucose esters), as substrate in SCPL-mediated acyltransferase reactions (Ciarkowska et al., 2019). SCPL acyltransferases possess a catalytic triad of serine, histidine, and aspartate, similar to catalytic triads in the protease chymotrypsin. Despite this similarity of the active site to known proteases, SCPL enzymes lack functional peptidase activity. SCPL acyltransferases are involved in

numerous metabolic pathways, including the biosynthesis of phytohormones and antimicrobial secondary metabolites (Lehfeldt et al., 2000; H. Liu et al., 2008; Mugford et al., 2009).

In Solanaceae, a serine carboxy peptidase-like acyltransferases family catalyzes these acyltransferase reactions at the C3 position of the α-tropine. Recently, in *Atropa belladonna,* two genes were found to be involved in the biosynthesis of littorine, an essential intermediate for tropane alkaloid biosynthesis in solanaceous plants (Qiu et al., 2019). The two genes, phenyllactate UDP-glycosyltransferase (UGT1) and littorine synthase (LS), are essential to produce littorine. UGT1 provides the critical cofactor, phenyllactylglucose, which acts as the energy-rich acyl donor in the enzymatic conversion of tropine into littorine via the LS. Littorine synthase, a key enzyme in TA biosynthesis in solanaceous plants, is a member of the serine carboxy peptidase-like acyltransferase family (SCPL-ATs).

The acronym BAHD is derived from the first letter of each of the first four biochemically characterized acyltransferases (D'Auria, 2006; Moghe et al., 2023; St-Pierre & De Luca, 2000). BAHD acyltransferases utilize acyl-CoA thioesters for the energy-rich acyl donor instead of the β-acetal glucose esters, as seen with SCPL acyltransferases. Acyl Coenzyme A thioesters are synthesized in various plant cell compartments but can be exported to the cytosol (Agrimi et al., 2012; Shockey & Browse, 2011). BAHD acyltransferases share several conserved domains. The HXXXD catalytic motif is integral for a general base-catalyzed acyl transfer mechanism (Shaw, 1992). The DFGWG domain near the carboxyl terminus is not involved in the catalytic function, even though this motif is highly conserved across BAHD enzymes. This DFGWG motif is thought to play a structural role rather than a catalytic role in BAHD acyltransferase reactions (Morales-Quintana et al., 2015).

In Erythroxylaceae, the acylation of tropane alkaloid differs from the SCPL acyltransferase-mediated biosynthesis of littorine observed in Solanaceae TA biosynthesis. Instead, the last step in the biosynthesis of cocaine in *E. coca* is mediated by a BAHD acyltransferase known as cocaine synthase (EcCS) [Figure 16]. Cocaine synthase can use either benzoyl or cinnamoyl CoA thioesters in combination with methylecgonine as a substrate to form the products cocaine or cinnamoylcocaine, respectively (Schmidt et al., 2015). Cocaine synthase was reported to be the first BAHD enzyme involved in tropane alkaloid biosynthesis in plants. The discrepancy between BAHD and SCPL involvement in modifying tropane alkaloids between the Erythroxylaceae and Solanaceae families suggests a polyphyletic

relationship in which tropane alkaloid biosynthesis has evolved multiple times in different plant families.

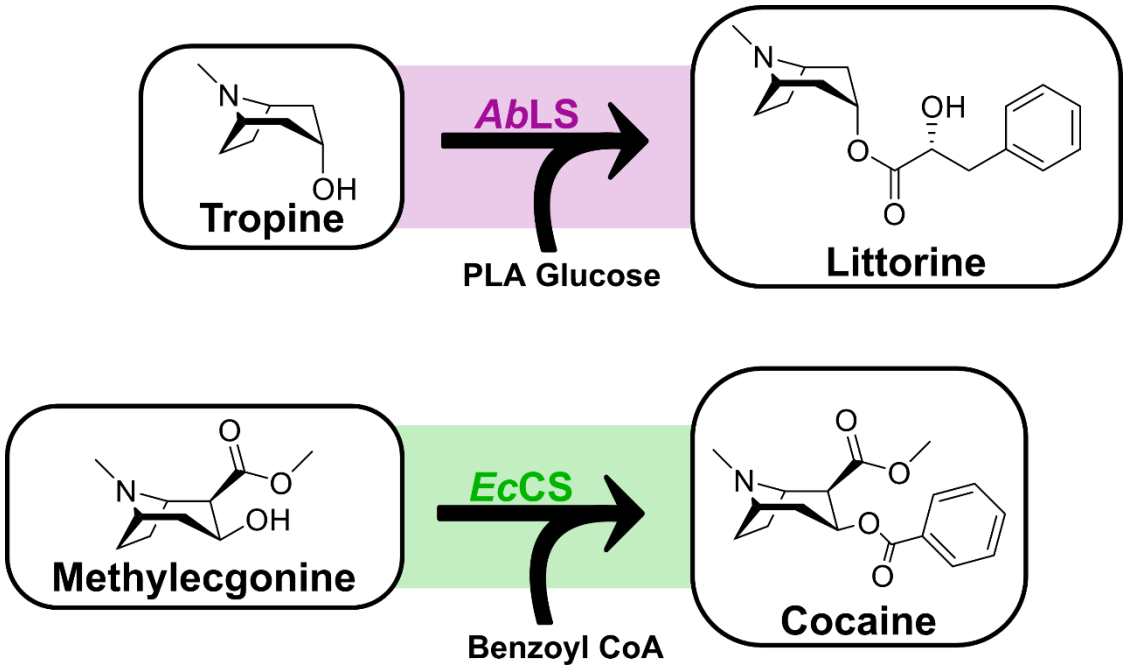

Figure 16: The acylation of tropane alkaloids mediated by either a BAHD acyltransferase or SCPL acyltransferase. Cocaine synthase (EcCS) is a member of the BAHD acyltransferase family (green color), while littorine synthase (AbLS) is a member of the SCPL acyltransferase family (purple color).

1.10 Scope of the Thesis

Synthetic biology is a multidisciplinary field that applies engineering principles to biology. The field of synthetic biology is a conglomerate of several scientific disciplines that seek to develop molecular tools, metabolic engineering, and organism engineering to facilitate the formation of biologically relevant products like medicines, biofuels, and sustainable biomaterials. The principle behind synthetic biology is to deconstruct biology at the genetic level, allowing modular genetic parts to be recombined in a design, build, test, and learn principles that can be applied to any living organism. It aims to design and build new biological systems that do not exist in nature. A recent publication created a yeast strain that contains >50% of its genome as synthetic or artificial chromosomes, demonstrating the potential synthetic biology has to offer in terms of the genome engineering of microorganisms (Zhao et al., 2023). Synthetic biology could be used to manipulate microorganisms through genetic engineering, producing valuable metabolites or developing bioengineered plants resistant to disease or pests.

In the last couple of years, synthetic biology approaches towards tropane alkaloid biosynthesis have generally focused on using microorganisms such as *Saccharomyces*

cerevisiae to synthesize tropane alkaloids *de novo* without relying on the cultivation of plants in order to produce tropane alkaloids for medicinal purposes (Ping, Li, Xu, et al., 2019; Ping, Li, You, et al., 2019; Srinivasan & Smolke, 2019, 2020). These advances rely upon the solid foundation of molecular biology and the identification of genes involved in tropane alkaloid biosynthesis. Our work focuses on uncovering tropane alkaloid biosynthesis in *Erythroxylum coca.* In collaboration with Dr. Christine Smolke and Dr. Prashanth Srinivasan at Stanford University, we leveraged metabolic engineering and synthetic biology approaches to rapidly identify and functionally validate key steps of tropane alkaloid biosynthesis within *E. coca*.

2 MATERIALS & METHODS

2.1 Codon Optimization of Synthetic DNA

Synthetic DNAs were synthesized either through Twist Biosciences or GENEWIZ. All coding sequences were codon optimized in Benchling using *Arabidopsis thaliana* or *Pichia pastoris* codon optimization parameters and removed the following type IIS restriction enzymes recognition sites of BsaI, BbsI, BsmBI, and SapI for maximum compatibility with GoldenGate, GoldenBraid, and MoClo assembly. GENEWIZ Germany GmbH synthesized the following genes for one-step cloning into the pHREAC binary expression vector using Golden Gate assembly: EcODC, EcSPMT, EcOGAS1, and EcOGAS2.

2.2 DNA Primers and Plasmid Sequencing

Golden Gate or MoClo assembly primers were designed using the Benchling Assembly Wizard tool using standard recommended settings. Primers were ordered from Metabion GmbH, and LGC Genomics GmbH performed Sanger sequencing of plasmids.

2.3 DNA Amplification and PCR validation

All PCR amplifications were done with either NEB Q5® High-Fidelity 2X Master Mix (Catalog #M0492S) for cloning applications, NEB OneTaq® 2X Master Mix with standard buffer (Catalog #M0482S) or Thermo Scientific DreamTaq Green PCR Master Mix 2X (Catalog #K1081) were used for colony PCRs applications. To determine the optimal annealing temperature, either New England Biolabs (NEB) Tm Calculator or Thermofisher Tm Calculator was used, respectively.

2.4 DNA Agarose and SDS-PAGE Protein Gel Imagining

All DNA agarose gels were visualized using the GelDoc Go Imaging System from Biorad using the Ethidium Bromide standard settings. SDS-PAGE protein gels were visualized using the GelDoc Go Imaging System from Biorad using the Coomassie Blue standard settings for densitometric analysis of protein bands of interest [Figure S4].

2.5 Modular Cloning (MoClo)-Based Method of Cloning

All MoClo-type IIS restriction enzyme ligation assemblies were performed following the protocol designed for type IIS DNA assembly, with the main difference being that 40fmol of DNA was used in all reactions instead of the standard 20fmol suggested in the protocol (Marillonnet & Grützner, 2020). The fmol amount of DNA was calculated using the following formula: 40*(fmol) × DNA size (in bp) / DNA concentration (ng/µl)/1520. When calculating DNA fragments or plasmids, it is essential to include the complete size of the DNA before restriction digestion to keep an equimolar concentration between plasmids and DNA fragments.

2.6 YPD and BMMY Media for *Komagataella phaffii* Protein Expression

For *K. phaffii* pre-starter and starter cultures, YPD media was used (10g/L yeast extract, 20g/L peptone, and 2% glucose per liter). The concentration of antibiotic zeocin in all liquid media and agar plates was 100µg/mL. For the expression cultures, BMMY media was used (10g/L yeast extract, 20 g/L peptone, 100 mL 1M potassium phosphate buffer pH = 6.0, 100mL of 10X YNB, 2mL of 500X Biotin, 5mL of 100% methanol, and 100µg/mL of zeocin per liter).

2.7 Cloning, Heterologous Expression, and Protein Purification of EcSMT, EcSPMT, and EcSPDS from *Komagataella phaffii*

Biochemically characterized PMT sequences from the Solanaceae family were used as a reference when performing a tBLASTn against the transcriptome of *Erythroxylum coca*. The candidates were named EcSMT, EcSPMT, and EcSPDS, and primers were designed to amplify these candidates from *E. coca* cDNA. Total RNA was extracted from *E. coca* leaves with a Qiagen kit. mRNA was reverse transcribed into cDNA with reverse transcriptase II (Invitrogen). EcSMT, EcSPMT, and EcSPDS were amplified from cDNA using designed gene-specific primers. After amplification, EcSPDS candidates were cloned using the Gateway system into pDONR207. A subsequent LR reaction into a pEPStrepGW vector yielded the construct used during heterologous expression.

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 II 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 100µg/mL of Zeocin (InvivoGen) at 28 °C with shaking at 250 rpm. Fresh BMMY (Buffered Methanol-complex Medium) medium supplemented with 100µg/mL 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 OD600 of 1.0. Protein expression was induced by adding 1% methanol (v/v) twice a day (morning and evening) and grown at 28 °C with shaking at 220-250 rpm for 2-3 days. The cells were harvested at 6000 x g at 4 °C for 10 min and then stored at -20 °C until 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 Stansted FPG12800 pressure cell homogenizer equipped with a 10mL pressure cell. The lysate was centrifuged at 6000 x 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 100mM Tris-HCl, pH 8, supplemented with 150mM NaCl, 1mM EDTA, and 2.5mM 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. Following the manufacturer's protocol, the total protein concentration was measured using the Bradford protein assay (Carl Roth 5x ROTI®Quant). The putative size and purity of the protein were determined via SDS-PAGE. Approximately 100ng of total protein was loaded into each well and stained with a Colloidal Coomassie or Serva Quick Coomassie® stain staining solution.

2.8 Construction of *Saccharomyces cerevisiae* strains for *in vivo* Screening of Candidate Genes

The method of constructing the *S. cerevisiae* strains and plasmid vectors for transformation has been described in detail in the supplemental data of Chavez et al., 2022.

2.9 Preparation of Chemically Competent *E. coli* Cells

E. coli TOP10 or DH5α were used to prepare chemically competent cells using the Inoue method for generating ultra-competent *E. coli* cells (Green & Sambrook, 2020).

2.10 Transformation of Chemically Competent *E. coli* Cells

E. coli TOP10 or DH5α competent cells were defrosted on ice for 10 minutes. Approximately 100-150ng of plasmid was added to the cells. After 10 minutes on ice, the mix was heat shocked at 42°C in a water bath for 45 seconds and placed back on ice for 2 minutes. Afterward, 250µL of SOC media was added, and the cells were shaken at 250 rpm for one hour at 37°C. Cells were spread out on LB agar plates with the corresponding antibiotic.

2.11 Preparation of Chemically Competent *Agrobacterium tumefaciens* Cells

Agrobacterium tumefaciens GV2260 strain was prepared following the protocol with minor adaptations (Höfgen & Willmitzer, 1988). An overnight culture of *Agrobacterium tumefaciens* was diluted in 200 mL YEB media. After 3-4 hours of growth, *Agrobacterium* cells were centrifuged at 3000 x g for 20 minutes at 4°C. The pellet was washed once in 10 ml precooled TE buffer (10mM Tris-HCI, pH 7.5; 1mM EDTA) and resuspended in 20mL fresh YEB-medium. Aliquots of 100µL and 200µL are frozen in liquid nitrogen and stored at -80°C until ready for transformation into *Agrobacterium tumefaciens*.

2.12 Protein Sequence Alignments

FASTA files containing the proteins of interest were first aligned using MEGA11 software using the MUSCLE alignment algorithm. The parameter settings for the MUSCLE protein alignment are the following settings: Gap Open -2.90, Gap Extend 0.00, Hydrophobicity Multiplier 1.20, Max Memory in MB was 2048, Max Iterations was 16, Cluster methods were both Neighbor Joining and Minimum Diagonal Length (lambda) at a value of 24. After performing the protein alignment in MUSCLE, it was exported as a FASTA alignment file for drawing the phylogenetic tree.

2.13 Generation of Phylogenetic Trees

Phylogenetic trees were generated using TreeViewer phylogenetic tree software via direct import of the protein alignment generated previously by MEGA11(Bianchini & Sánchez-Baracaldo, 2024; Tamura et al., 2021). Using Neighbor Joining algorithm to create an unrooted tree was used. The distance model parameter was BLOSUM62 and bootstrap replicates were set at 1000. After the generation of the tree, adjustments were made to display the trees in a circular or linear format for optimal graphical viewing. Scale bars representing the proportion of the total length from the root of the tree to the farthest tip were added to the tree. Finally, a filter was applied to remove all bootstrap values from the nodes that were less than 500. The

phylogenetic trees were then exported as a scalable vector graphics file using the embed full font export selection for future work in a vector graphics program. Additional modifications were performed using Affinity Designer V2 vector graphics software to highlight parts of the tree and generate a figure legend.

2.14 Transformation of Plasmids into *Agrobacterium tumefaciens* cells

Stored *Agrobacterium tumefaciens* GV2260 competent cells are thawed on ice prior to transformation. Competent cells are mixed with approximately 500ng of plasmid DNA. The cells are incubated successively for 5 minutes on ice, 5 minutes in liquid nitrogen, and 5 minutes in a 37°C water bath. After dilution in 1ml of YEB-medium, the cells are shaken for 2-4 hours at room temperature or 28°C. 100µL of transformed cells are plated on YEB agar plates containing rifampicin at a concentration of 25µg/mL and kanamycin at a concentration of 50µg/mL antibiotics and incubated for 2-3 days at 28°C. Single colonies were validated with colony PCR for pHREAC expression vector and subsequently used for transient expression in *Nicotiana benthamiana*.

2.15 Transformation of Plasmids into *K. pastoris* cells

The electro-competent cells of *K. pastoris* strain KM71 were taken from the -80°C freezer. First, 40µl of 1M sorbitol was added to the competent cells. A 38µL aliquot of the mixture was placed in a new Eppendorf tube. Next, add 1-5µl of plasmid to the new Eppendorf (final concentration of 150ng/µl) total. Let the tube rest on ice for 5 mins. Fill the mixture in the 1mm cuvette and pulse cells using the *Picha* method on GenePulser/Bio-Rad (Set to 1.5kV, 200Ω, 2.5µF). Immediately add 1ml of cold 1M sorbitol and transfer into sterile 1.5ml Eppendorf. Cells were regenerated for 1-2 hours at 28ºC to recover after electroporation. Spread aliquots onto YPD agar plates containing 100µg/mL Zeocin. Incubate for 2-4 days at 28ºC.

2.16 Fluorescent Microscopy of Organelle Targeting in *Saccharomyces cerevisiae*

The method of analyzing *S. cerevisiae* cells via fluorescent microscopy was described in the publication Chavez et al., 2022. 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 5-10µL of cell suspension was spotted onto a glass microscope slide, covered with a glass coverslip (Thermo Fisher), and then imaged using an upright Zeiss Axio Imager 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.

2.17 Identification of TA Genes from *E. coca* Transcriptome

Candidate genes involved in tropane alkaloid biosynthesis were manually screened using a combination of transcriptomic datasets from the United States Department of Agriculture (USDA), the 1kp transcriptomic database (Carpenter et al., 2019), as an in-house transcriptome generated by the D'Auria lab. Previous work in the D'Auria lab identified tropane alkaloid-related genes that were involved in the biosynthesis of cocaine (Anderson, 2019; Chavez et al., 2022; Estrada, 2017; Kim, 2020). Utilizing tBLASTn to screen for gene candidates in *Erythroxylum coca* transcriptomic datasets using TA biosynthetic genes from solanaceous plants. After the gene candidates were identified from the tBLASTn search, other tools like Interproscan, pHMMR, and NCBI tBLASTn/BLASTp were used in combination to identify known protein motifs for each candidate to ensure full-length protein sequences.

2.18 RNA Extraction from *E. coca* Tissue and cDNA Synthesis

Approximately 100mg of fresh *E. coca* tissue was used to extract total RNA using the NucleoSpin RNA Plant mini kit, and genomic DNA was removed via an in-column RNase-free DNase I treatment (Machery Nagel, Germany). RNA concentrations and quality were determined using a NanoDrop One Microvolume UV-Vis Spectrophotometer (ThermoScientific Germany) and an Agilent 2100 Bioanalyzer Instrument (Agilent Germany). According to the manufacturer's protocol, cDNA was synthesized with 300ng total RNA using either NEB LunaScript® RT SuperMix Kit (Catalog # E3010S) or SuperScript™ IV First-Strand Synthesis System (Invitrogen, Germany) for cDNA synthesis.

2.19 qPCR Primers Design Parameters

qPCR primers were designed in Benchling using the qPCR (Intercalating Dyes) parameter. Primers were optimized at around 50% GC content. The melting temperature for the primers was designed using the parameters between 58°C minimum, 60°C optimum, and 65°C maximum. Primer lengths were optimized between 17bp minimum, 22bp optimum, and 26bp maximum. Amplicon sizes were selected to be approximately 70bp minimum, 100bp optimum, and 150bp maximum as a DNA amplicon size for qPCR. The best-scoring primer pairs from Benchling were selected for PCR targeting genes of interest within tissue-specific cDNA libraries of *E. coca*.

2.20 Polyamine Extraction of *E. coca* tissues

Polyamines were extracted from buds, L1-L3 stage leaf, stem, and root 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 100µL 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.

2.21 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). A 50µL aliquot of sample was derivatized with 50µL of OPA-ET stock and left to react for 1.5 minutes, then 2µL of the FMOC stock was added and let the FMOC react with the sample for at least 2 minutes before injection.

2.22 Polyamine Detection by HPLC-Fluorescence

For the detection of the polyamines, an Ultimate-3000 HPLC (Thermo Scientific) with a fluorescence detector (337nm Emission/ 454nm Excitation) with a Nucleodur Sphinx RP 5µm, 4.6mm x 250mm column (Macherey Nagel) was used. The buffers used for this separation method are 0.2% formic acid (Buffer A) and 100% acetonitrile (Buffer B). Flow

rate: 1mL/min. Injection volume: 5µL, 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.

2.23 Polyamine Detection by UPLC-Fluorescence

An Acquity UPLC system (Waters Corporation, Milford, MA, USA) coupled with an Acquity Fluorescence Detector (337nm Emission/ 454nm Excitation) to detect the derivatized polyamines. An Acquity UPLC BEH C18 130A, 1.7µm, 2.1mm x 50mm column was used. The separation method was combined with the polyamine derivatization protocol (Hanczkó et al., 2007; Zierer et al., 2016). Flow rate of 0.4 mL/min for 4.5 min. The column temperature was set to 40°C. HPLC-grade water (A) and LCMS-grade acetonitrile (B) were used as solvents. The binary gradient used was as follows: 0 to 0.06 min at 90% B, 0.06 to 1.7 min at 58% B, 1.7 to 3.17 min at 0% B, 3.17 to 3.29 min at 58% B, and 3.29 to 4.50 min at 90% B for column equilibration.

2.24 *In vitro* Enzyme Assays for SPDS and SMT Activity via Fluorescence Detection

Enzyme assays were performed as follows: SMT assays contained 2µL of putrescine from 50mM stock final concentration of 1mM, 1µL dcSAM from 50mM stock final concentration of 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 with a final concentration of 1mM, 1µL SAM from 50mM stock with a final concentration of 0.5mM, 20µL of purified enzyme, and a 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.

2.25 Plant Material

Nicotiana benthamiana plants (4–6 weeks old) were grown in a greenhouse at 55% humidity, 22°C, and 16-h-light/8-h-dark photoperiod. *E. coca* v. *coca* plants were grown at 22°C under a 12-h-light/12-h-dark photoperiod, with humidity of 65% and 70%, respectively, and were fertilized weekly with Ferty 3 (15-10-15) and Wuxal Top N (Planta). A voucher specimen was deposited at the Herbarium Haussknecht (JE) at Friedrich Schiller University in Jena, Germany.

2.26 Plasmid Construction for Transient Expression in *Nicotiana benthamiana*

For transient expression of *E. coca* pathway genes in *Nicotiana benthamiana*, A. tumefaciens pHREAC binary vectors 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 [Table S6]. 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 GmbH. Vector confirmation was conducted with vector-specific primers pHREAC FWD 5'-(CAACCACAACGCTCTAACGC)-3' and pHREAC REV 5'-(AAAATTAAAATCTTTTTGTGTCCTTGCT)-3' for sequencing and colony PCR.

2.27 Transient Protein Expression in *Nicotiana benthamiana*

Agrobacterium tumefaciens strain GV2260 cells were transformed with pHREAC binary transformation vectors (Peyret et al., 2019) containing *E. coca* genes of interest using the freeze-thaw method (Höfgen & Willmitzer, 1988) and confirmed via colony PCR. A single agrobacterium colony containing the construct of interest was cultured in liquid YEB (5 g/L, beef extract, 1g/L yeast extract, 5g/L peptone, 5g/L sucrose, 0.5g/L $MgCl_2$) medium overnight at 28°C with appropriate antibiotics. The overnight cultures were pelleted at 5000 x g for 30 min, resuspended in infiltration media (10mM MES pH 5.6, 10mM $MgCl_2$), 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 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 using a vacuum infiltration 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 100mg/mL of extraction solvent. The extraction solvent was 20% methanol containing 0.1% formic acid and 1000 ng/mL of deuterated *N*-methylpyrrolinium (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.

2.28 Metabolite Analysis of *Nicotiana 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 version B.07.01 (Agilent, Germany) based on retention time and the mass fragment/transition parameters in [Table S11]. 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 or against predicted transitions determined using the CFM-ID fragment prediction utility and the METLIN database.

2.29 LC-MS Separation of Polyamines and Methionine Cycle-related Metabolites

LC-MS method was modified from Su et al. 2021, which separates polyamines and methionine cycle-related metabolites without the need for derivatization (Su et al., 2021). Targeted analyses of polyamines and methionine cycle-related metabolites were carried out by Agilent 1260 Infinity Binary HPLC coupled to an Agilent 6495 triple-quadrupole mass spectrometer. The ACQUITY UPLC BEH Amide column (2.1 × 100 mm, 1.7 µm, Waters, USA) was used for metabolite separation at 30 °C. The mobile phases were water (A) and ACN (B), both containing 0.1% formic acid. The LC gradient program was optimized as follows: 0 min, 85% B; 9 min, 45% B; 10 min, 30% B; 11 min, 85% B; 14 min, 85% B. The injection volume was 10µL, and the flow rate was 0.600mL/min. The data was collected using multiple reaction monitoring (MRM) mode. 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 version B.07.01 (Agilent, Germany) based on retention time and the mass fragment/transition parameters in [Table S11]. Multiple reaction monitoring (MRM) transitions and quantification was performed by the analysis of 1µM-1mM standards diluted in 85% acetonitrile and 15% water using MassHunter Optimizer software (Agilent) and corroborated against published mass transitions if available, or against predicted transitions determined using the CFM-ID fragment prediction utility and the METLIN database.

2.30 Hierarchical Clustering Analysis of Erythroxylum Transcriptomes

Hierarchical Clustering Analysis of Erythroxylum Transcriptomes. Hierarchical clustering of methylecgonone synthase candidates from the Erythroxylum transcriptome datasets (generated as described in the previous section) was performed using a custom R script (coca_analysis_v1.R), available on GitHub at the following link https://github.com/smolkelab/coca_coexpression_analysis; the entire computational pipeline was described in Chavez et al., 2022.

2.31 Statistics

Biological replicates are defined as independent cultures inoculated from separate yeast colonies or streaks and cultivated in separate containers for yeast experiments. For *Nicotiana benthamiana* experiments, one biological replicate is defined as all infiltrated leaves from a single plant. Four biological replicates for each gene stacking for all performed transient expression matrices.

2.32 Chemical Standards

Chemical compounds and standards Putrescine dihydrochloride (Duchefa Biochemie P0927.0005), Spermidine trihydrochloride (AppliChem A0673.0001), and hygrine (sc-488263) were purchased from Carl Roth. *N*-methylspermidine was chemically synthesized following protocols described in (Anderson, 2019). *N*-methylputrescine dihydrochloride (sc-212242) and hygrine (sc-488263) were purchased from Santa Cruz Biotechnology (Dallas, TX). *N*-methylpyrrolinium (NMPy) was prepared from γ-Methylaminobutyraldehyde (4MAB) diethyl acetal (M285740) purchased from Toronto Research Chemicals (Toronto, ON), as

described previously (Srinivasan & Smolke, 2019). MPMOB was prepared as described previously (Katakam et al., 2019). All other chemicals were purchased from Sigma.

2.33 GenBank Accession Numbers of Identified Genes from *E. coca*

Complete DNA sequences of *E. coca* genes newly identified in this work are hosted in GenBank with the following accession numbers: EcSPMT, OP382839; EcSMT, OP382840; EcSPDS, OP382841; EcAOF1, OP382842; EcAOC1, OP382843; EcAOC2, OP382844; EcOGAS1, OP382845; EcOGAS2, OP382846; EcMPOBMT, OP382847; EcCYP81AN15, OP382848.

3 RESULTS

3.1 Detecting Derivatized Polyamines from *E. coca* Tissue Extracts via HPLC-FLD

Polyamine profiles were characterized in *E. coca* tissues, specifically buds, roots, stems, and L1 through L3 stage leaves, using high-performance liquid chromatography coupled to a fluorescence detector (HPLC-FLD) to detect derivatized polyamine profiles within the different tissues of *Erythroxylum coca*. Additional LC-MS/MS profiling of the derivatized polyamines from *E. coca* identified the presence of putrescine, spermidine, and spermine. Additionally, we observed the presence of a novel polyamine called *N*-methylspermidine [Figure 17]. Interestingly, we could not detect any presence of *N*-methylputrescine (NMPUT) in any tissues from *E. coca*.

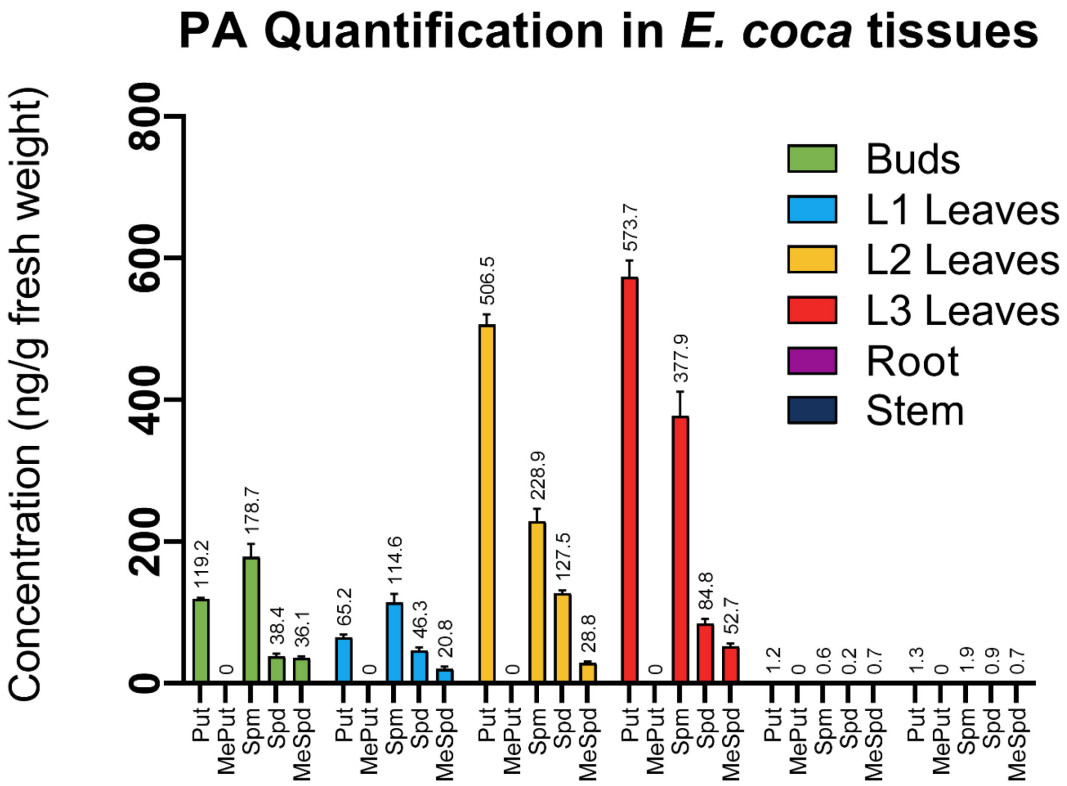

Figure 17: Polyamine content in *E. coca* tissues. Polyamines were detected from acidic methanol extracts and are labeled as putrescine (Put), *N*-methylputrescine (MePut), spermidine (Spd), spermine (Spm), and *N*-methylspermidine (MeSpd). Bars and error bars represent the mean and SD of three tissue biological replicates, each with three technical replicates. This figure was modified from Chavez et al., 2022; Page 3; Figure 2A.

N-methylspermidine was detected in higher amounts in the aerial leaf tissues, where tropane alkaloid biosynthesis is known to occur in *E. coca* (Docimo et al., 2012; Jirschitzka et al., 2012; Schmidt et al., 2015). Coincidentally, this contrasts with known observations of putrescine *N*-methyltransferase expression and the accumulation of *N*-methylputrescine within the root tissues of plants in the Solanaceae and Convolvulaceae families. The presence of *N*-methylspermidine and the absence of putrescine *N*-methyltransferase in aerial tissues where tropane alkaloid biosynthesis is known to occur in *E. coca* provides evidence that the well-established TA biosynthesis in Solanaceae may not be operating identically to the TA biosynthesis in *Erythroxylum coca*.

3.2 The First Committed Step of TA Biosynthesis in *E. coca* Begins with Spermidine *N*-methylation

3.2.1 Homology Search of PMT-like Candidates

A homology-based search for putrescine *N*-methyltransferase-like sequences failed to yield close matches. The absence of *N*-methylputrescine in polyamine extracts is inconsistent with tropane alkaloid biosynthesis starting from the formation of putrescine *N*-methylation as reported within TA-producing plants in the Solanaceae family. The presence of *N*-methylspermidine within leaves of *E. coca* suggests an alternate route in Erythroxylaceae, which uses spermidine rather than putrescine in the first committed step of tropane alkaloid biosynthesis.

Since no spermidine *N*-methyltransferases (SMTs) were reported before, we decided to use sequences from biochemically characterized solanaceous putrescine *N*-methyltransferases (PMTs), which were predicted to be the closest potential sequence match for SMT. Utilizing these PMTs as a query, we performed a tBLASTn search of our in-house transcriptome database of *E. coca* leaf tissues for potential SMT candidate genes. We identified three full-length open reading frames (ORFs) that more closely resembled spermidine synthases (SPDSs) rather than putrescine *N*-methyltransferases [Figure S1]. We initially annotated these putative candidates as EcSPDS1, EcSPDS2, and EcSPDS3.

3.2.2 Heterologous Expression in *Komagataella phaffii* and Purification of EcSPDS1-3 Candidates

EcSPDS1-3 candidates were cloned into the pDONR207® Gateway entry vector and finally cloned into a pEPStrepGW expression vector. The genes of interest cloned into the pEPStrepGW vector were used for heterologous expression and protein purification using *Komagataella phaffii* (formerly known as *Pichia pastoris*) KM71 cells. Following StrepTag II affinity column purification, proteins were analyzed by SDS-PAGE, which indicated molecular weights of 39.0, 36.8, and 38.5 kDa for EcSPDS1, 2, and 3, respectively [Figure S2, Figure S3, and Figure S5].

3.2.3 *In vitro* Enzymatic Activities Assays of SPDS-like Candidates

Enzyme assays with purified recombinant EcSPDS1-3 failed to show PMT activity, which is consistent with our observations on the absence of the metabolite *N*-methylputrescine

in *E. coca* tissues. Instead, these enzymes possessed spermidine synthase (SPDS) or spermidine *N*-methyltransferase (SMT) activities. EcSPDS1 showed SMT activity in which spermidine is methylated at the (internal) N4 position using SAM as a cofactor [Figure 18]. EcSPDS3 exhibited spermidine synthase activity using decarboxylated SAM (dcSAM) as a cofactor [Figure 20]. EcSPDS2 demonstrated both SMT and SPDS activity. EcSPDS2 alone converted putrescine to spermidine and *N*-methylspermidine when dcSAM and SAM cofactors were present in the same enzyme assay [Figure 19]. Based on these observations, we have renamed EcSPDS1, EcSPDS2, and EcSPDS3 to EcSMT, EcSPMT, and EcSPDS, respectively.

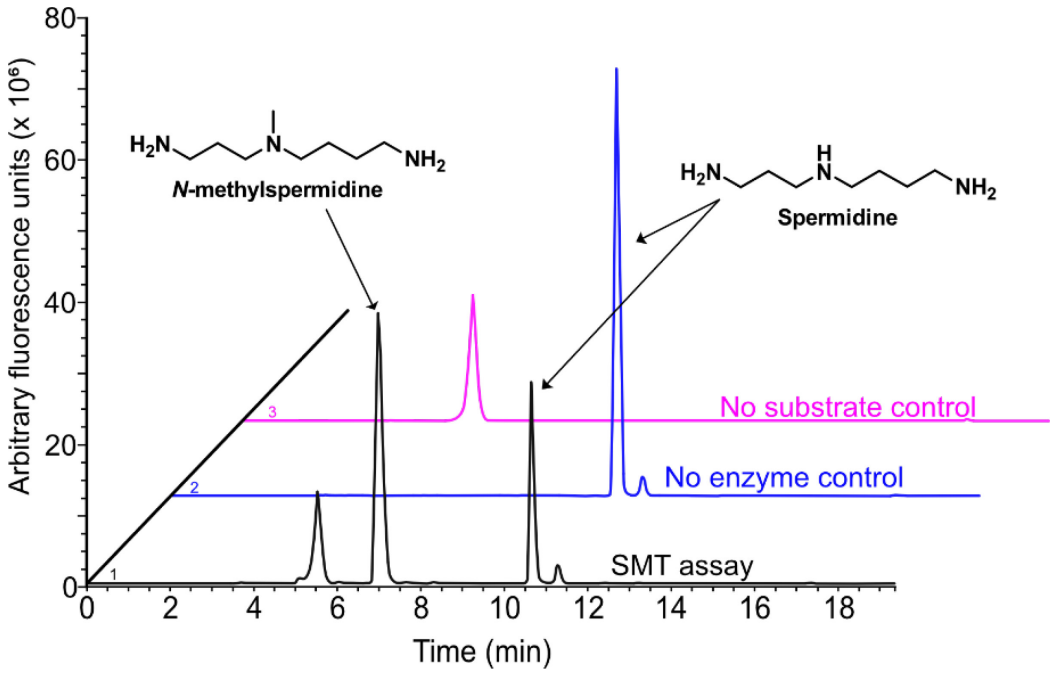

Figure 18: *In vitro* enzymatic characterization of purified EcSMT. Before injection, the in vitro enzyme assay samples were derivatized with o-phthalaldehyde (OPA) and fluorenylmethyl chloroformate (FMOC). Chromatograms represent n = 5 independent enzyme assays, and peak heights are shown to scale with respect to all other peaks for the same compound. All peak heights are shown to scale with respect to all other peaks. HPLC chromatography data was rendered using Serif Affinity Designer. This figure is from Chavez et al., 2022; Page 3; Figure 2C; The data was adapted from Anderson et al., 2019.

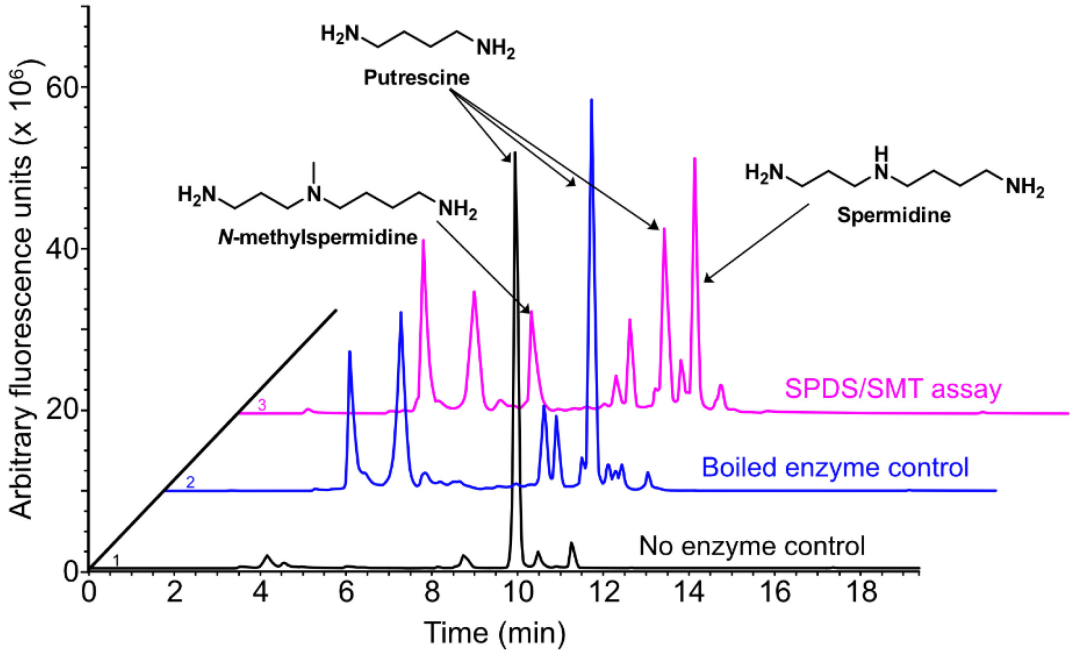

Figure 19: *In vitro* enzymatic characterization of purified EcSPMT. Before injection, the in vitro enzyme assay samples were derivatized with o-phthalaldehyde (OPA) and fluorenylmethyl chloroformate (FMOC). Chromatograms represent n = 5 independent enzyme assays, and peak heights are shown to scale with respect to all other peaks for the same compound. HPLC chromatography data was rendered using Serif Affinity Designer. This figure is from Chavez et al., 2022; Page 3; Figure 2D. The data was adapted from Anderson et al., 2019.

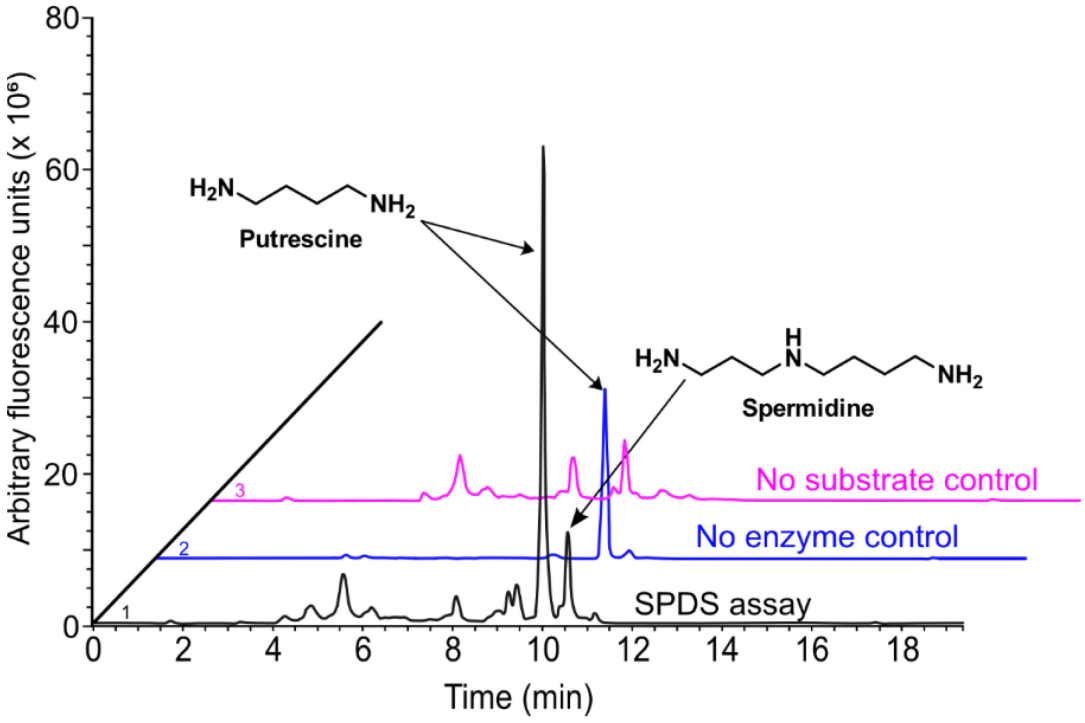

Figure 20: *In vitro* enzymatic characterization of purified EcSPDS. Before injection, the in vitro enzyme assay samples were derivatized with o-phthalaldehyde (OPA) and fluorenylmethyl chloroformate (FMOC). Chromatograms represent n = 5 independent enzyme assays, and peak heights are shown to scale with respect to

all other peaks for the same compound. HPLC chromatography data was rendered using Serif Affinity Designer. This figure is from Chavez et al., 2022; Page 3; Figure 2B. The data was adapted from Anderson et al., 2019.

3.2.4 Phylogenetic Analysis of EcSMT, EcSPMT, and EcSPDS

The phylogenetic relationships between EcSMT, EcSPMT, and EcSPDS enzymes were evaluated through sequence alignments of known spermidine synthases (SPDSs), spermine synthases (SPSs), and putrescine *N*-methyltransferase (PMTs) protein sequences and evaluated clustering patterns in the resulting dendrogram [Figure 21]. The sequences used in this phylogenetic tree are available in [Table S1]. All characterized PMTs cluster into a separate clade from SPDSs and SPSs, consistent with the proposed early divergence of PMT from the ancestral SPDS family (Biastoff, Reinhardt, et al., 2009; Junker et al., 2013). In contrast, there appears to be no apparent clustering between other spermidine and spermine synthases. EcSMT, EcSPMT, and EcSPDS are scattered throughout the SPDSs and SPSs protein families.

Phylogenetic analysis of SPDS-related enzymes from *Erythroxylum coca* was compared, and the following differences were observed. The gene candidates EcSMT, EcSPMT, and EcSPDS, when compared against other spermidine synthases (SPDS), spermine synthases (SPS), and putrescine *N*-methyltransferases (PMT) enzymes, these candidates did not cluster closely with themselves but instead clustered with other SPDS and SPS enzymes. This lack of clustering revealed that EcSMT and EcSPMT only shared ~49% amino acid identity. EcSMT, compared to EcSPDS, only shared ~25% identity at the amino acid level. EcSPMT and EcSPDS share only ~27% amino acid identity. None of these SPDS-related enzymes from *E. coca* clustered with known PMTs from the Solanaceae and Convolvulaceae plant families [Figure 21].

Upon further analysis, we found that EcSMT clusters closely with the spermidine synthase, TcSPDS1, and TcSPDS2, from *Theobroma cacao* and CcSPDS4 from *Cajunus cajan*. These protein sequences share a percent identity between 53-57% compared to EcSMT based on amino acid percent identity. EcSPMT clusters closely with CsSPDS3 (AAT66041.1) from *Cucumis sativus*, TcSPDS3 from *Theobroma cacao*, CaSPDS from *Coffea arabica*, and PgSPDS from *Panax ginseng*. These protein sequences share a percent identity of greater than 82% compared to EcSPMT based on amino acid percent identity. EcSPDS clusters closely with CuSPDS from *Citrus unshiu* and ACL5 from *Arabidopsis thaliana*. CuSPDS is currently an uncharacterized putative spermidine synthase to which the protein annotations are inferred from homology. ACL5, known as thermospermine synthase, was originally mischaracterized as a spermidine synthase from *Arabidopsis thaliana* (Hanzawa et al., 2000; Knott et al., 2007).

These protein sequences share a percent identity between ~69-75% compared to EcSPDS based on amino acid percent identity. Interestingly, the enzymes that cluster with each of the candidates from *E. coca* are all members of the rosids clade of plants, unlike the PMT cluster from Solanaceae and Convolvulaceae families, which are members of the asterid clade of plants, suggesting a polyphyletic origin of the first committed step of tropane alkaloid biosynthesis.

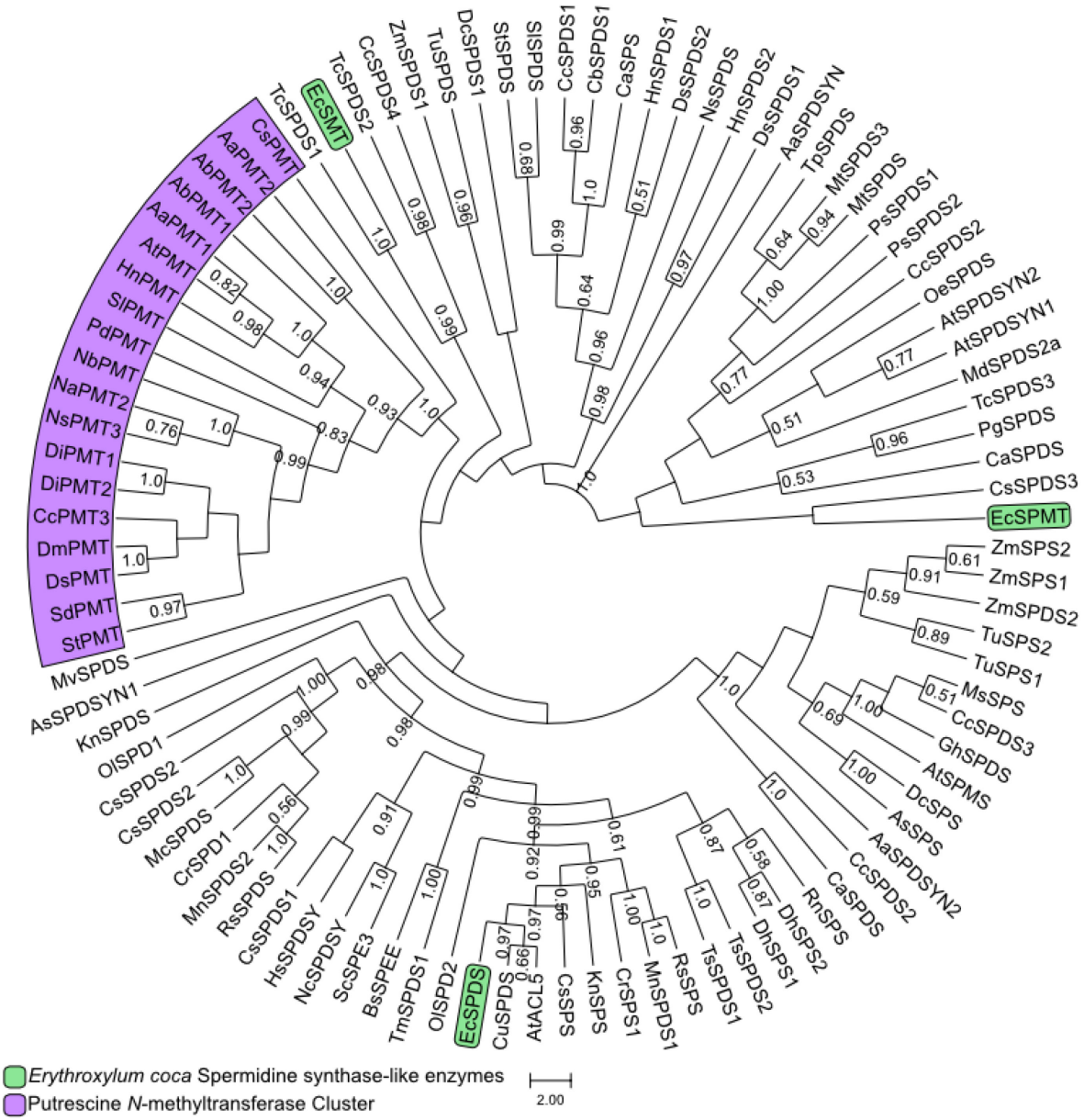

Figure 21: A dendrogram showing the sequence relationships based on amino acids between EcSPDS, EcSMT, and EcSPMT. PMT clade is indicated in purple, and the identified *E. coca* SPDS-related enzymes are highlighted in green. Refer to Table S1 for identities and accession numbers of all enzymes in the dendrogram. Values at each

junction signify the number of bootstraps out of n = 1,000 iterations. Bootstrap values below 500 were not reported. The scale bar represents the proportion of the total length from the root of the tree to the farthest tip.

3.2.5 *In vivo* Screening of EcSMT, EcSPMT, and EcSPDS Enzymatic Activities in *Saccharomyces cerevisiae*

We then characterized the *in vitro* activities of the EcSPDS and EcSPMT enzymes with spermidine synthase activity in *Saccharomyces cerevisiae*. The SAMDC gene from yeast (SPE2) is known to enhance SPDS activity by providing excess dcSAM cofactor. We co-expressed EcSPDS or EcSPMT, native yeast SPE3, a BFP control, and SPE2 from low-copy plasmids. LC-MS/MS analysis of culture supernatants after 72 hours growth of transformed strains in selective media indicated that EcSPMT, like native Spe3p, can function as an SPDS when paired with Spe2p in yeast. We constructed a spermidine overproduction strain, CSY1340, by integrating expression cassettes for SPE2, SPE3, and EcSPMT and disrupting the endogenous spermine synthase gene SPE4 in the genome of the previously engineered putrescine overproduction strain CSY1242 (Srinivasan & Smolke, 2019). We then demonstrated that, unlike our previous observations of *in vitro* enzyme assays, EcSMT, not EcSPMT, exhibits spermidine *N*-methyltransferase activity when expressed from a low-copy plasmid in the CSY1340 yeast strain, as indicated by LC-MS/MS analysis of culture supernatants following 72-hours growth of transformed strains in selective media [Figure 22] (Chavez et al., 2022).

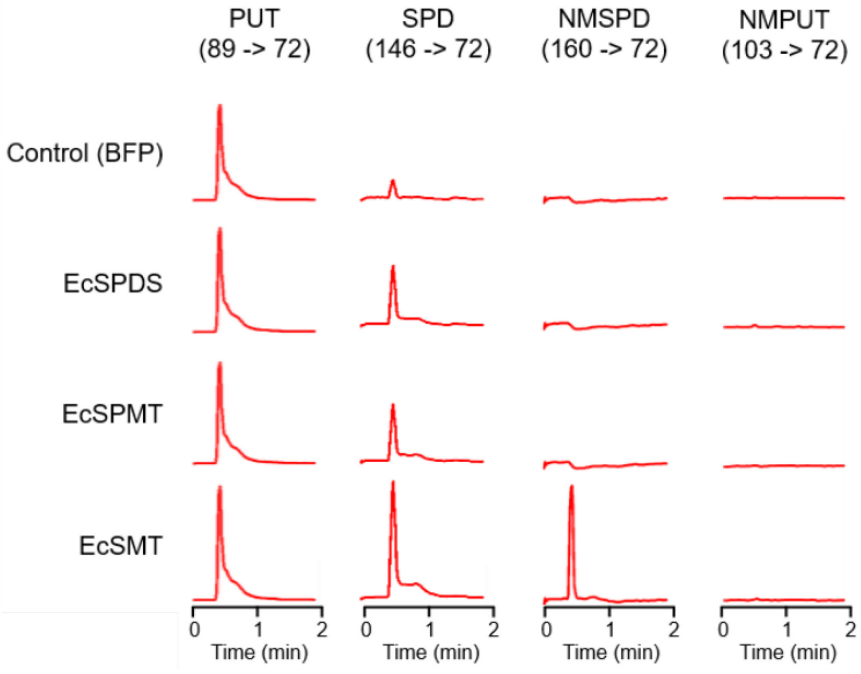

Figure 22: Characterization of *E. coca* spermidine *N*-methyltransferase activities in engineered yeast. *N*-methyltransferase candidates or a blue fluorescent protein (BFP) control were expressed from low-copy plasmids

in strain CSY1340, and metabolites in the supernatant were analyzed by LC-MS/MS after 72 hours of 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. Published in Chavez et al., 2022; Supplemental Material Page 24; Figure S9.

3.3 A Flavin-Dependent Amine Oxidase mediates the Formation of *N*-methylputrescine

3.3.1 Homology Search of Flavin-Dependent Amine Oxidases in *E. coca*

Initial homology search attempts using solanaceous TA pathway genes failed to reveal flavin-dependent amine oxidases (AOFs), so the decision was made to use other known plant AOF genes. The search for a potential enzyme that could perform the oxidation of the *N*-methylspermidine step was performed by using flavin-dependent amine oxidase protein sequence Q9SU79 from *Arabidopsis thaliana* for the tBLASTn query Q9SU79 (M. S. Islam et al., 2017). We identified three possible flavin-dependent enzymes that are associated with polyamine catabolism, like spermine and spermidine oxidases. We annotated these flavin-dependent oxidase candidates as EcAOF1, EcAOF2, and EcAOF3.

3.3.2 Phylogenetic Analysis of Flavin-Dependent Oxidases from *E. coca*

To further investigate phylogenetic relationships of the flavin-dependent amine oxidases potentially involved in TA biosynthesis in *E. coca*. We performed sequence alignments of EcAOF1-3 oxidase candidates against other flavin-dependent oxidases. We evaluated clustering patterns in the resulting dendrograms [Figure 23]. The EcAOF1-3 candidates do not cluster among themselves but rather with other known polyamine oxidases from plants. EcAOF1 clustered with other known plant polyamine oxidases such as spermine oxidases isolated from *Arabidopsis thaliana* (48% identity to Q9FNA2), *Oryza sativa* (50% identity to A0A0P0XM10, 51% identity to Q0J290), and a spermidine oxidase from *Zea mays* (51% identity to O64411). EcAOF2 clustered other plant polyamine oxidases, such as spermine oxidase AtPAO4 from *Arabidopsis thaliana* (58% identity to Q9SKX5). EcAOF3 also clustered with another spermine oxidase, AtPAO5, from *Arabidopsis thaliana* (61% identity to Q9SU79).

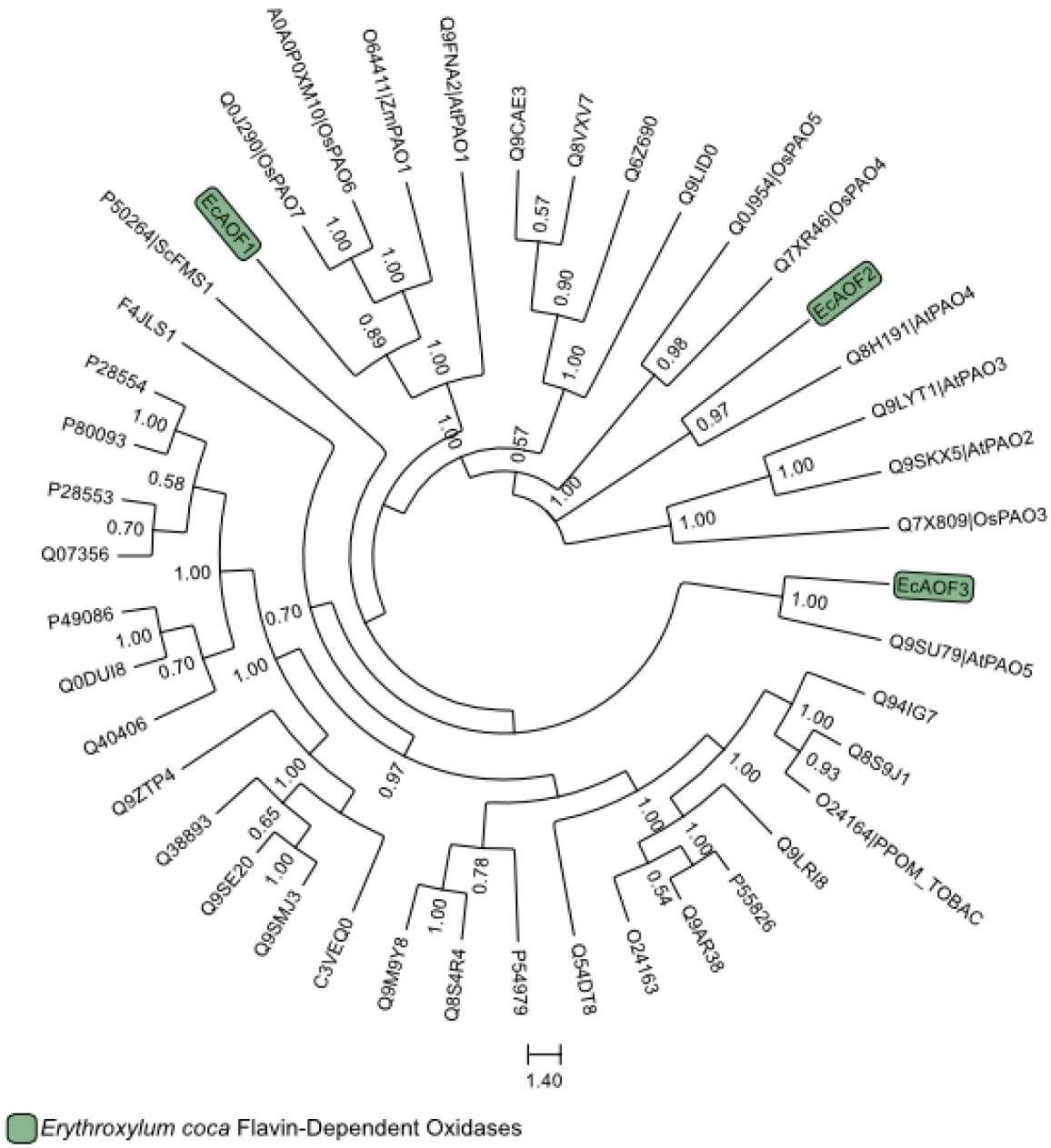

Figure 23: Phylogenetic analysis of flavin-dependent oxidases from *E. coca*. Phylogenetic tree showing evolutionary sequence relationships between EcAOF1 and other flavin-dependent polyamine oxidases. EcAOF1 from *E. coca* is highlighted in green. Refer to Table S2 for identities and accession numbers of all enzymes in the dendrogram. Values at each junction signify the number of bootstraps out of n = 1,000 iterations. Note that bootstrap values below 500 were not reported. The scale bar represents the proportion of the total length from the root of the tree to the farthest tip.

3.3.3 *In vivo* Screening of EcAOF1-3 Candidates for *N*-methylspermidine Oxidase Activities in *Saccharomyces cerevisiae*

To further characterize the flavin-dependent oxidase candidates from *E. coca* for functional oxidase activity towards *N*-methylspermidine, a *Saccharomyces cerevisiae* strain that overproduces *N*-methylspermidine was developed (Chavez et al., 2022). This *N*-methylspermidine overproducing yeast strain was used to screen the AOF candidates by co-

expressing each *E. coca* flavin-dependent oxidase candidate along with EcSMT from low-copy plasmids in the spermidine overproduction yeast strain (Chavez et al., 2022). EcAOF1 was the only flavin-dependent oxidase candidate capable of oxidizing *N*-methylspermidine into *N*-methylputrescine. In contrast, the remaining EcAOF2 and EcAOF3 candidates displayed no functional activity toward converting *N*-methylspermidine to *N*-methylputrescine in the yeast expression strain [Figure 24] (Chavez et al., 2022).

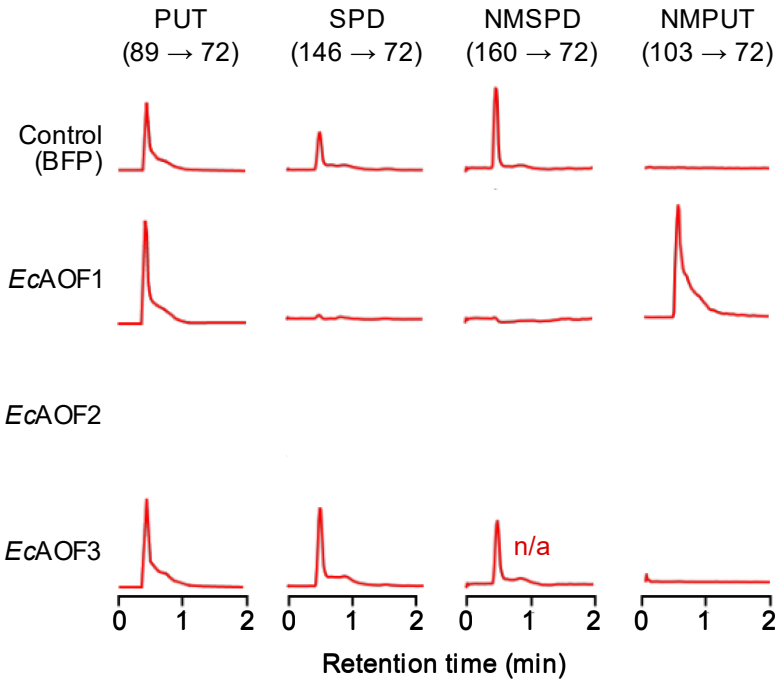

Figure 24: LC-MS/MS multiple reaction monitoring (MRM) chromatogram traces showing the screening of *E. coca* flavin-dependent amine oxidase (AOF) candidates in yeast for activity on *N*-methylspermidine. MS/MS mass transitions are used to identify putrescine (PUT), spermidine (SPD), *N*-methylspermidine (NMSPD), and *N*-methylputrescine (NMPUT). Chromatogram traces represent n = 3 biologically independent samples, and peak heights are shown to scale with respect to all other peaks for the same compound. Published in Chavez et al., 2022; Page 3; Figure 2F.

3.4 Two Copper-Dependent Amine Oxidases mediate First Ring Closure of Tropane Alkaloids in *E. coca*

3.4.1 Homology Search of Copper-Dependent Amine Oxidases in *E. coca*

The homology search of copper-dependent amine oxidases (AOC) was performed by using a copper-dependent oxidase from *Pisum sativum* (Q43077.1) along with the characterized *N*-methylputrescine oxidase (MPO) from *Nicotiana tabacum* (Heim et al., 2007; V. Kumar et al., 1996). Initial evaluation of the gene candidates revealed that several AOC transcripts from *E. coca* were not full-length but contained copper-binding motifs and peroxisomal targeting signals on the C-terminus. Utilizing the 1KP transcriptomic database,

United States Department of Agriculture (USDA) transcriptomic database, and our in-house transcriptome of *Erythroxylum coca*, these genes were reconstructed through additional amino acid alignments, leading to full open reading frames (ORFs) of these AOC gene candidates (Carpenter et al., 2019).

3.4.2 Phylogenetic Analysis of Copper-Dependent Oxidases from *E. coca*

Phylogenetic analysis of EcAOC1-6 candidates from *Erythroxylum coca* shows clustering amongst other reported AOC enzymes from plants (Tavladoraki et al., 2016). These copper-dependent amine oxidases from plants generally cluster into three main clades: Clades I, II, and III. When evaluating our sequences from *E. coca*, we noticed EcAOC1-6 were dispersed throughout each of the AOC clades. EcAOC1 and EcAOC2 cluster within known copper-dependent amine oxidases in Clade III of plant copper-dependent oxidases [Figure 25]. EcAOC1 clusters closely with AOCs from *M. domestica* (~83% AA identity to AIS23648.1, ~84% AA identity to AIS23647.1) and *R. communis* (84% AA identity to XP_002511334.1). EcAOC2 clusters more closely with AOCs from *Malus domestica* (~83% AA identity to AIS23647.1), *Vitis vinifera* (~86% AA identity to XP_002273532.2), *Ricinus communis* (~85% AA identity to XP_002527922.1), *Glycine max* (~84% AA identity to XP_003551224.1), and *Medicago truncatula* (~86% AA identity to XP_003613133.2). EcAOC2, while it clusters in AOC Clade 3, does not cluster closely with known *N*-methylputrescine oxidases (MPOs) from *A. belladonna* (AbMPO2), *Nicotiana sylvestris* (XP_009778427.1), and *Nicotiana tabacum* (BAF49520.1) [Figure 27]. The remaining copper-dependent oxidase candidates cluster in AOC clades I and II. EcAOC3 and EcAOC6 cluster in AOC Clade I, while EcAOC4 and EcAOC5 cluster in AOC Clade II. The clustering of EcAOC3-6 candidates in AOC Clades I and II but not in Clade III suggests different copper-dependent oxidase activity.

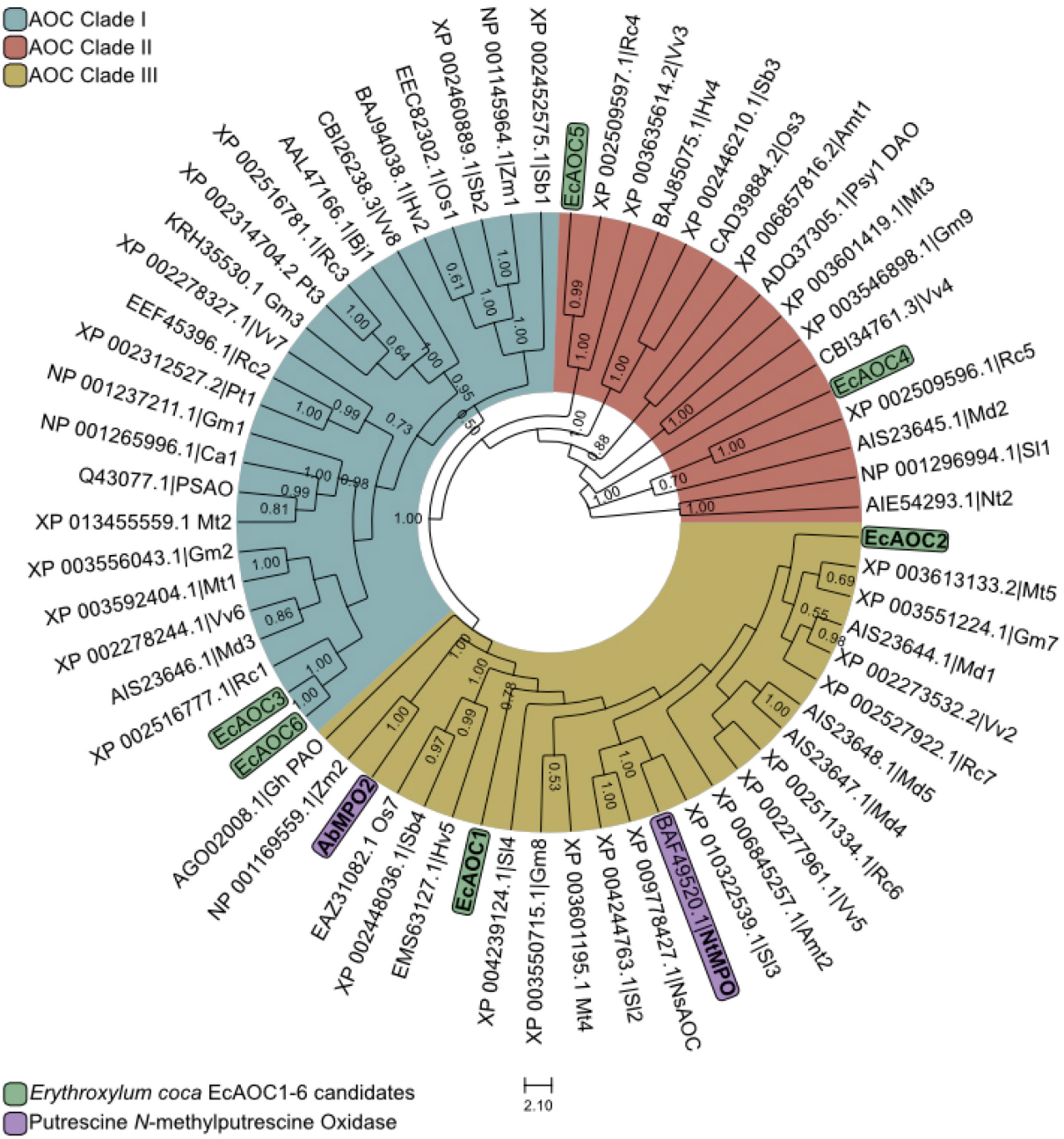

Figure 25: Phylogenetic analysis of copper-dependent polyamine oxidases from *E. coca*. The phylogenetic tree shows evolutionary sequence relationships between EcAOC1 and EcAOC2 compared to other copper-dependent polyamine oxidases. *E. coca* EcAOC1-6 identified in this study are highlighted in green, and the EcAOC1 and EcAOC2 are in bold, indicating validated *N*-methylputrescine oxidase activity. Solanaceous *N*-methylputrescine oxidases are highlighted in purple. The representation clades of copper-dependent oxidases are divided into three clades: AOC Clade I in blue, AOC Clade II in red, and AOC Clade III in yellow. Refer to Table S3 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. Note that bootstrap values below 500 were not reported. The scale bar represents the proportion of the total length from the root of the tree to the farthest tip.

3.4.3 *In vivo* Screening of EcAOC1-6 Candidates for *N*-methylputrescine Oxidase Activity in *Saccharomyces cerevisiae*

Six full-length AOC candidates, along with the remaining AOF candidates (EcAOF2 and EcAOF3) that were not involved in the oxidation of *N*-methylspermidine, were then screened in yeast for the ability to oxidize *N*-methylputrescine. Methylputrescine oxidase (MPO) from *Datura metel* was used as a positive control for MPO activity in yeast. Only two candidates from *E. coca*-derived copper-dependent amine oxidases, EcAOC1 and EcAOC2, possessed this MPO activity, consistent with our observations of clustering in AOC Clade III based on the phylogenetic tree analysis. These two candidates are EcAOC1 and EcAOC2 and were found to oxidize *N*-methylputrescine into the *N*-methyl-Δ1-pyrrolinium cation (NMPy), which the NMPy cation spontaneously reacted with endogenous polyketides to produce hygrine inside the yeast [Figure 26] (Chavez et al., 2022). Hygrine is a significant side product of the first tropane ring formation via spontaneous condensation of NMPy with endogenous yeast keto-metabolites such as acetate and acetoacetyl-CoA. Hygrine only accumulates in yeast engineered to have the required enzymes for NMPy biosynthesis. Hygrine was used as a proxy for LC-MS/MS detection over NMPy due to its vastly improved MS/MS detectability (Srinivasan & Smolke, 2020).

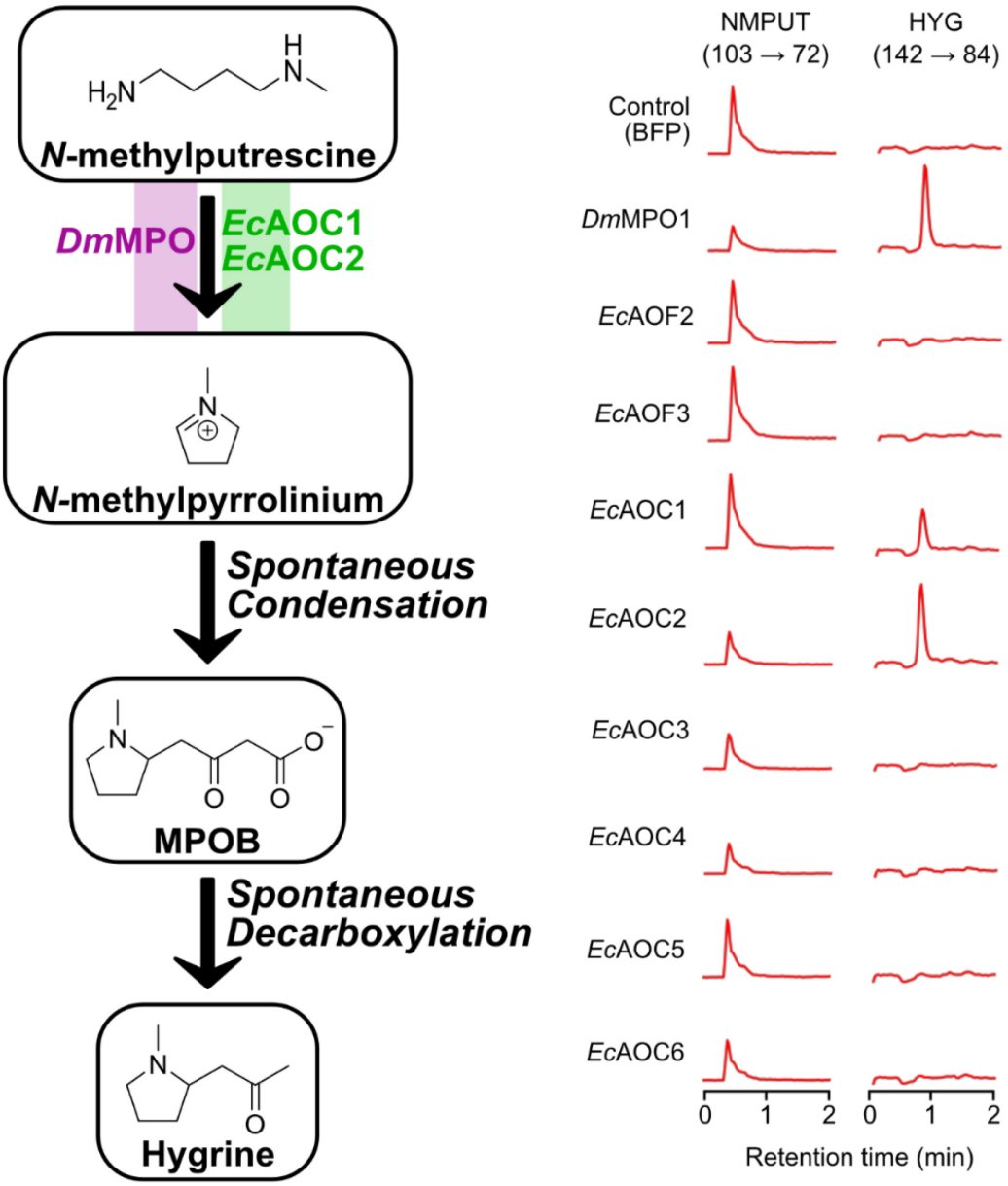

Figure 26: The biosynthetic pathway leading to *N*-methylpyrrolinium formation from *N*-methylputrescine. The green color denotes the biosynthetic pathway of *Erythroxylum coca,* and purple denotes the MPO biosynthetic step from Solanaceae. LC-MS/MS multiple reaction monitoring (MRM) chromatogram traces showing the screening of *E. coca* flavin-dependent amine oxidase (AOF) and copper-dependent amine oxidase (AOC) candidates for activity on *N*-methylputrescine in engineered yeast. MS/MS mass transitions used for identifying *N*-methylputrescine (NMPUT) and hygrine (HYG; a proxy for *N*-methylpyrrolinium formation in yeast) are indicated in parentheses. Chromatogram traces represent n = 3 biologically independent samples, and peak heights are shown to scale with respect to all other peaks for the same compound. Modified from Chavez et al., 2022; Page 3, Figure 2G.

3.4.4 Subcellular Localization of EcAOF1, EcAOC1, and EcAOC2 in *Saccharomyces cerevisiae*

Subcellular localization can provide insight into enzymes' functions and, by extension, phylogenetic origins. As plant AOFs and AOCs differ in their canonical subcellular location,

the localization of the newly identified *E. coca* amine oxidases was verified via fluorescence microscopy of wild-type yeast (CEN.PK2) co-expressing N- or C-terminal GFP fusions of EcAOF1, EcAOC1, or EcAOC2 with a peroxisomal marker (mCherry-Pex3p) from low copy plasmids. Consistent with its lack of detectable localization signals, EcAOF1-GFP appears to form cytosolic inclusion bodies, while GFP-EcAOF1 is soluble and cytosolic [Figure 27]. Consistent with the peroxisomal localization of AOCs like the analogous MPOs in Solanaceae and with the presence of C-terminal peroxisome targeting signals (PTSs), both GFP-EcAOC1 and GFP-EcAOC2 co-localize with mCherry-Pex3p to yeast peroxisomes, whereas EcAOC1-GFP and EcAOC2-GFP with masked C-terminal PTSs remain in the cytosol. These results suggest that the newly identified *E. coca* amine oxidases follow the expected subcellular localization pattern for their associated enzyme families.

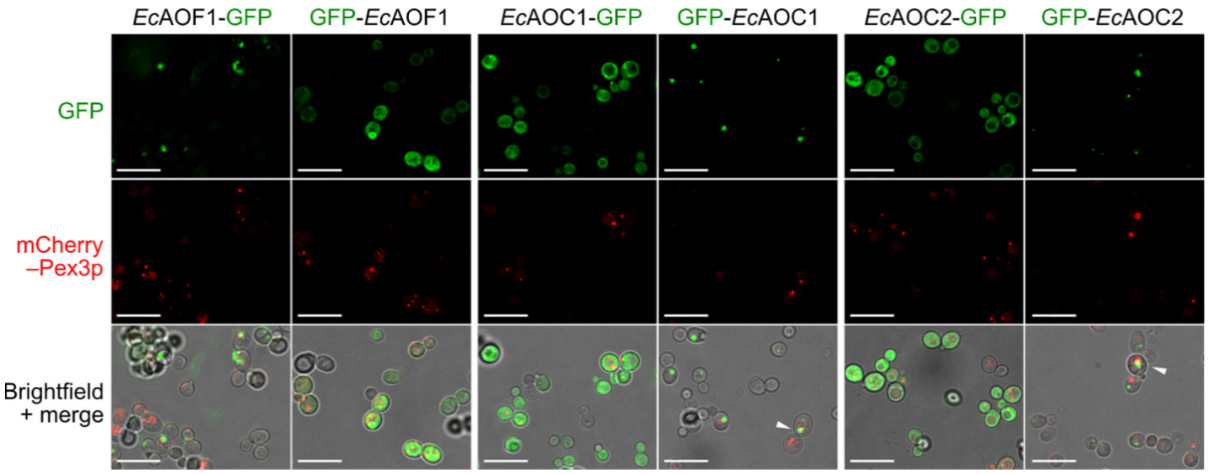

Figure 27: Subcellular localization of *E. coca* amine oxidases in yeast via fluorescence microscopy. N- and C-terminal GFP fusions of EcAOF1, EcAOC1, and EcAOC2 were expressed along with a yeast peroxisomal localization marker gene (mCherry-PEX3) from low-copy plasmids in wild-type yeast (CEN.PK2). Microscopy and image analysis were performed as described in Materials and Methods. The scale bar is 10µm. Published in Chavez et al., 2022; Supplemental Material Page 25; Figure S10.

3.5 First Ring Closure Pathway Reconstruction of *E. coca* Genes in *Nicotiana benthamiana*

3.5.1 Transformation of *Agrobacterium tumefaciens* GV2260 for Transient Expression in *Nicotiana benthamiana*

After identifying the missing steps leading to the formation of NMPy in *E. coca* and functionally validated in yeast, expression *in planta* was the next step. Tropane alkaloid pathway reconstruction in the heterologous host, *Nicotiana benthamiana.* This pathway reconstruction led to rapid validation *in planta* of biosynthetic genes involved in the first ring closure of tropane alkaloids in *E. coca*. The transformation method of *N. benthamiana* was

performed through gene stacking and agroinfiltration techniques (Chuang & Franke, 2022). Colonies of *Agrobacterium tumefaciens* GV2260 transformed with the pHREAC vector containing our genes of interest were screened via colony PCR to ensure that the ORF of the expected size was present in the pHREAC vector backbone before proceeding with agroinfiltration into *Nicotiana benthamiana* plants [Figure 28].

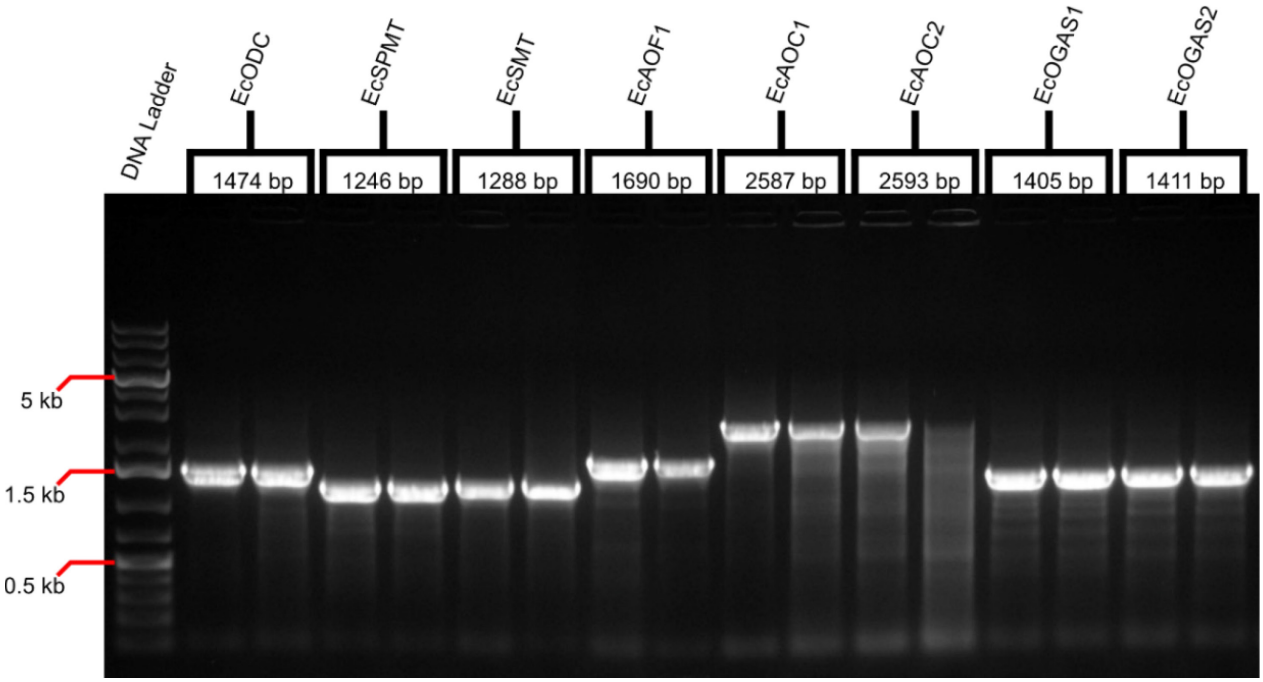

Figure 28: Colony PCR of transformed *Agrobacterium tumefaciens* GV2260 harboring pHREAC expression vector. The expected amplicon size is listed below the name of each gene. Each gene was screened with two colonies. DNA ladder was GeneRuler 1 kb Plus DNA Ladder (Catalog number SM1331).

3.5.2 Agroinfiltration and Gene Stacking Techniques in *Nicotiana benthamiana*

The *N. benthamiana* plants were vacuum infiltrated using gene stacking approaches to allow for a quick and scalable matrix of genes to be evaluated simultaneously. All agroinfiltration experiments included wild-type *Nicotiana benthamiana* plants as an untransformed control, a mock control with infiltration buffer only, and an empty pHREAC vector (EV) control. All *Agrobacterium*-mediated vacuum infiltration expression matrices are defined in [Table 1].

Matrix Name	Gene Stacking Matrix Combinations
M1	EcODC + EcSPMT
M2	EcODC + EcSPMT + EcSMT
M3	EcODC + EcSPMT + EcSMT + EcAOF1
M4	EcODC + EcSPMT + EcSMT + EcAOF1 + EcAOC1/2
M5	EcODC + EcSPMT + EcSMT + EcAOF1 + EcAOC1/2 + EcOGAS1/2
EV	Empty pHREAC vector (EV) control
Mock	Mock only with infiltration buffer and without *Agrobacterium tumefaciens*
WT	Wildtype untransformed *Nicotiana benthamiana* plants

Table 1: Gene stacking matrix combinations of Agrobacterium-mediated transformation of *Nicotiana benthamiana* for functional validation of genes involved in the first ring closure of TAs.The following gene names are defined as ornithine decarboxylase (EcODC), spermidine synthase/ spermidine *N*-methyltransferase (EcSPMT), spermidine *N*-methyltransferase (EcSMT), flavin-dependent amine oxidase (EcAOF1), copper-dependent amine oxidase (EcAOC1/EcAOC2), 3-oxoglutaric acid synthase (EcOGAS1/EcOGAS2).Ornithine decarboxylase (EcODC) was present in all transient expression matrices except for the controls to ensure that the putrescine pool was sufficient to support the beginning steps of tropane alkaloid biosynthesis. EcODC and EcSPMT were tested together (Matrix 1) since the *in vivo* expression in yeast failed to demonstrate the dual functionality of EcSPMT and only functioned effectively as a spermidine synthase when expressed in *Saccharomyces cerevisiae*. (Chavez et al., 2022). However, when EcSPMT was transiently expressed in *N. benthamiana,* the production of *N*-methylspermidine was detected via LC-MS/MS analysis [Figure 29]. The inclusion of EcSMT with EcSPMT also generated *N*-methylspermidine. Upon the addition of flavin-dependent amine oxidase (EcAOF1) (Matrix 3), the *N*-methylspermidine pool observed in Matrices 1 and 2 were converted into *N*-methylputrescine [Figure 29]. The addition of two copper-dependent amine oxidases (EcAOC1 and EcAOC2) (Matrix 4) was confirmed to catalyze the conversion of *N*-methylputrescine to *N*-methyl-Δ^1-pyrrolinium cation (NMPy) through detection of hygrine as a proxy. The addition of 3-oxoglutaric acid synthases (EcOGAS1 and EcOGAS2) to the final experimental gene stacking matrix (Matrix 5) generated the precursor 3-oxoglutaric acid in *Nicotiana benthamiana* leaves. This precursor can spontaneously react with NMPy cation, forming 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate (MPOB) as a product [Figure 31]. MPOB was detected using the metabolite hygrine as a proxy, a more stable intermediate that can be measured more reliably than MPOB, which is prone to spontaneous decarboxylation reaction. These results demonstrate that the beginning steps of TA biosynthesis using gene candidates isolated from *Erythroxylum coca* and first validated in

Saccharomyces cerevisiae are also functional in an orthologous expression system such as *Nicotiana benthamiana*.

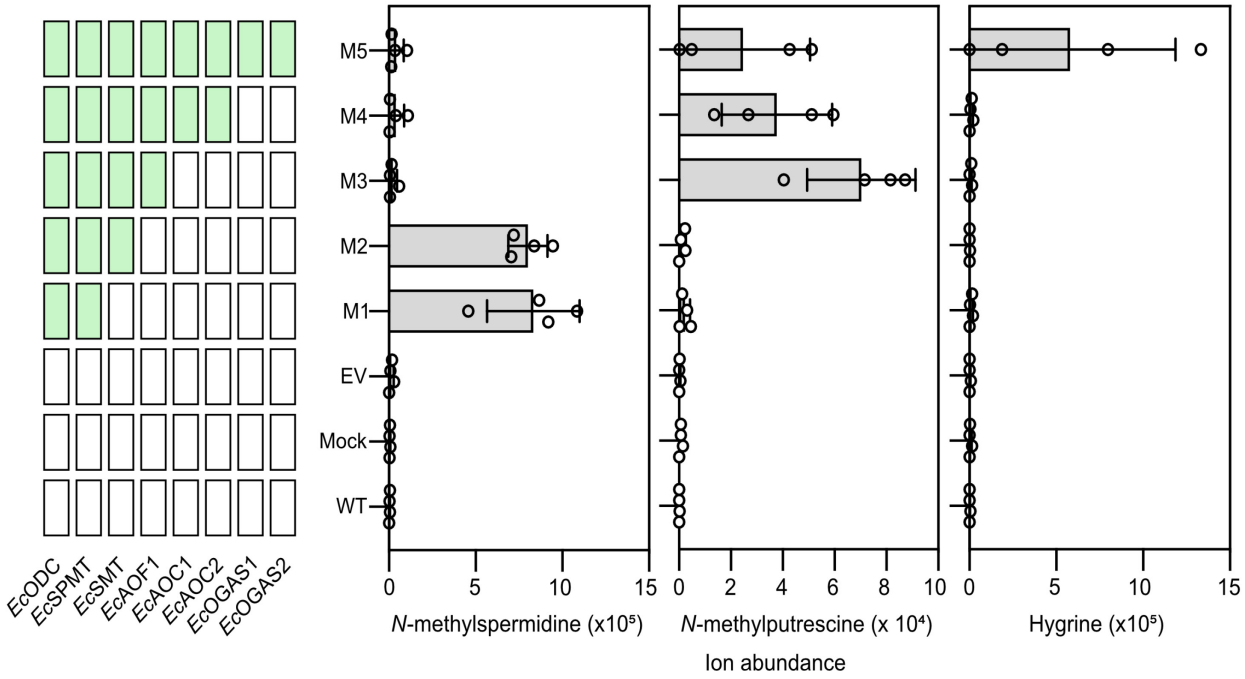

Figure 29: Validation of the biosynthetic sequence for the first tropane ring formation in *E. coca* via transient co-expression in *N. benthamiana*. Multiple genes were stacked via A. tumefaciens vacuum infiltration and analyzed via MRM LC-MS/MS detection of *N*-methylspermidine, *N*-methylputrescine, and hygrine, the spontaneous decarboxylation product of MPOB. WT, wild-type *N. benthamiana* with no agroinfiltration; Mock, control infiltration with infiltration buffer only; EV, empty pHREAC vector control; M1–M5, transient co-expression matrices 1–5 (each matrix is a unique combination of transiently co-expressed genes). The grid on the left side indicates the presence (green) or absence (white) of binary expression vectors for the indicated *E. coca* gene in the Agrobacterium co-infiltration matrices. Data indicate the mean of n = 4 biologically independent samples (open circles), and error bars show standard deviation. Published in Chavez et al., 2022; Page 3; Figure 2H.

3.6 A Member of the SABATH Methyltransferase Family Mediates the 2-CMO Group Retention in TA present in *E. coca*

3.6.1 Homology Search of SABATH-like Methyltransferases in *E. coca*

The tropane alkaloid, cocaine found in *Erythroxylum coca,* possesses a carbomethoxy (CMO) group at the C2 position of the tropane ring. The 2-carbomethoxy ester (2-CMO) group is distinctive to the Erythroxylaceae family and not observed in TAs from the Solanaceae and Convolvulaceae family. Previous work in the D'Auria lab identified members of the SABATH family of methyltransferases in *E. coca* (Estrada, 2017). We searched for SABATH methyltransferases in the transcriptome of *E. coca* using a tBLASTn search with a salicylic acid methyltransferase from *Clarkia breweri* (CbSAMT) and a jasmonic acid

methyltransferase from *A. thaliana* (AtJAMT) as sequence queries. We identified three candidates: EcSABATH1, EcSABATH2, and EcSABATH3.

3.6.2 Phylogenetic Analysis of SABATH-like Methyltransferases from *E. coca*

Sequence alignments against known SABATH methyltransferases and dendrogram analysis showed that EcSABATH1 clusters closely with previously characterized salicylic acid methyltransferases from multiple species as well as the jasmonic acid methyltransferase AtJAMT (46% AA identity) (Zubieta et al., 2003). The candidates, EcSABATH2 and EcSABATH3, cluster with biochemically characterized paraxanthine methyltransferases from *A. thaliana,* sharing approximately 37% identity [Figure 30].

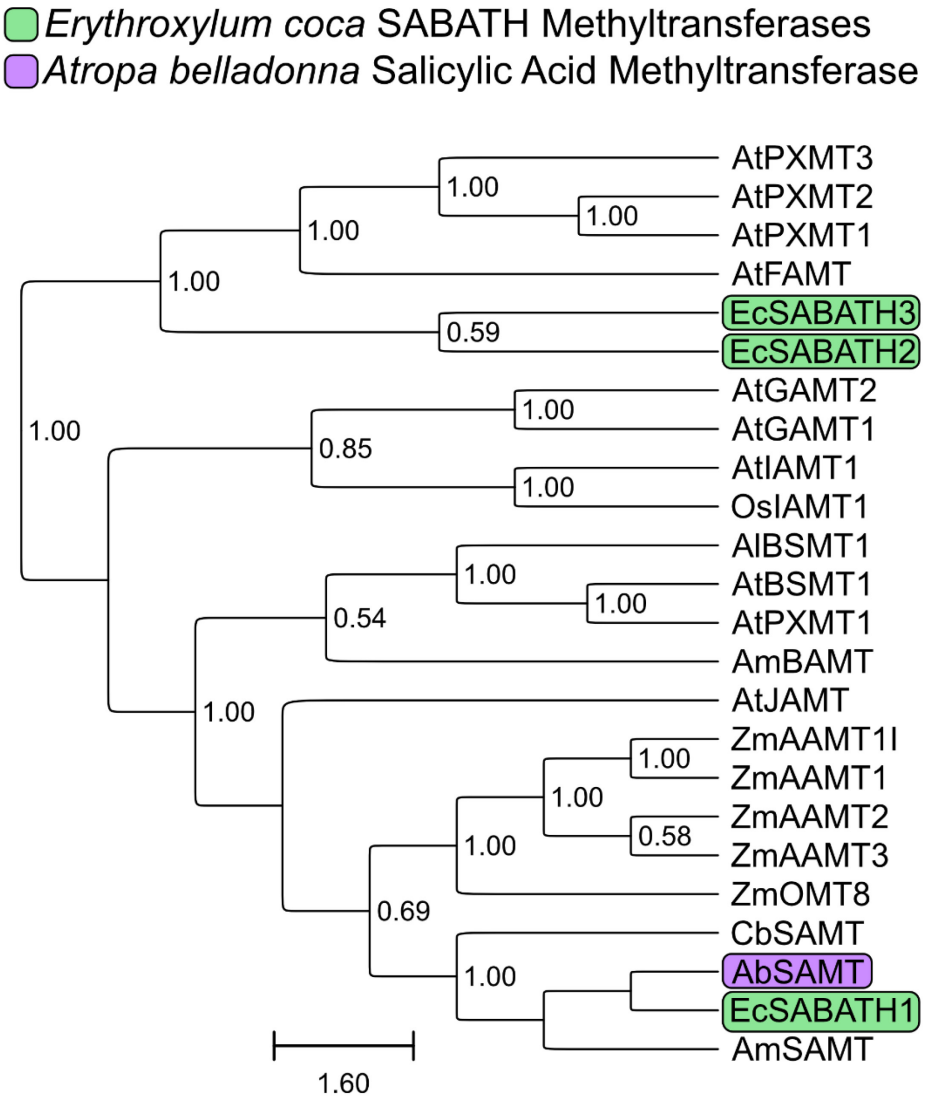

Figure 30: Dendrogram showing phylogenetic analysis of three SABATH methyltransferase candidates in green identified from *E. coca* transcriptome. Refer to Table S4 for identities and accession numbers of all enzymes in

the dendrogram. Values at each junction signify the number of bootstraps out of n = 1,000 iterations. The scale bar represents the proportion of the total length from the root of the tree to the farthest tip.

3.6.3 *In vivo* Screening of EcAOC1-6 for *N*-methylputrescine Oxidase Activity in *Saccharomyces cerevisiae*

EcSABATH1-3 candidates were evaluated in the yeast expression system to screen for the ability to methylate 4-(1-methyl-2-pyrrolidinyl)3-oxobutanoate (MPOB) producing the metabolite 4-(1-methyl-2-pyrrolidinyl)3-methyloxobutanoate (MPMOB). LC-MS/MS analysis of culture supernatants following 72 h growth of transformed yeast strains indicated that while all three PYKS/OGAS enzymes are functional in yeast, AbPYKS from *Atropa belladonna* enabled increased flux through MPOB to tropine and was used to support the screening of SABATH methyltransferase candidates for MPMOB production (Chavez et al., 2022). Coexpression of the three SABATH methyltransferase candidates and additionally, we coexpressed a blue fluorescent protein (BFP) control with AbPYKS from low-copy plasmids in yeast strain CSY1246 and analyzed the accumulation of NMPy, MPOB, and MPOB methyl ester (MPMOB) in the culture medium via LC-MS/MS after 72 h growth of transformed strains [

Figure 31]. Only EcSABATH1 was found to methylate MPOB, creating MPMOB in the yeast expression system. Therefore, we renamed EcSABATH1 to MPOB methyltransferase (EcMPOBMT) (Chavez et al., 2022).

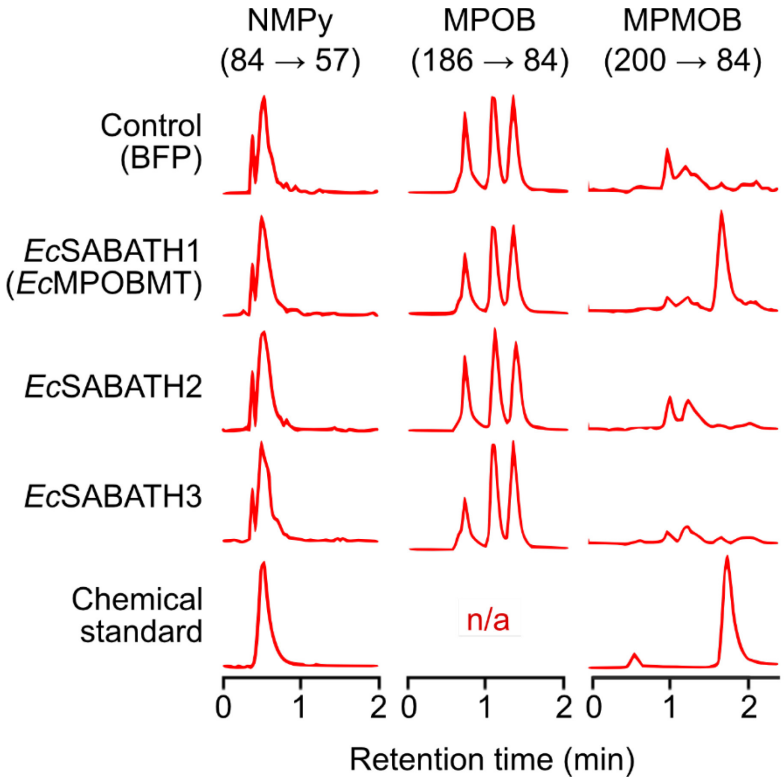

Figure 31: LC-MS/MS MRM chromatogram traces showing the screening of *E. coca* SABATH methyltransferase candidates for activity on MPOB in engineered yeast. Published in Chavez et al., 2022; Page 6; Figure 3B.

3.7 A CYP81-Family Monooxygenase Catalyzes Methylecgonone Ring Closure in *E. coca*

3.7.1 Homology Search of Tropinone Synthase-like Candidates in *E. coca*

Using tropinone synthase (AbCYP82M3) identified from *Atropa belladonna,* we performed a homology-based search against the transcriptome of *E. coca* and identified five candidate CYP450s (denoted EcCYP1-5). Hydroxylation of secondary metabolites in plants via 2-oxoglutarate-dependent dioxygenases (2ODDs) may also be involved (Kawai et al., 2014). Additionally, we expanded our search to include more diverse CYP450 families and 2ODDs. Using a hierarchical clustering approach, we searched the transcriptomes of three *E. coca* plants (denoted 48, 113, and 124), two *E. novogranatense* plants (denoted 209 and 228), and low-cocaine-producing Erythroxylum species, *Erythroxylum hondense* (Bieri et al., 2006), for CYP450 and 2ODD candidates exhibiting similar patterns of expression across the three species as previously identified Erythroxylaceae TA genes [Figure S6].

3.7.2 *In vivo* Screening of Hydroxylase and Cyclase Candidates for Methylecgonone Synthase Activity in *Saccharomyces cerevisiae*

The candidate hydroxylase and cyclase enzymes were screened in the yeast platform. We co-expressed each of the 24 hydroxylase candidates (EcODD1-5, EcCYP1-17, EcP4H1-2) or a BFP control with the known cytochrome P450 reductase AtATR1 from *Arabidopsis thaliana* in the MPMOB-producing yeast strain CSY1341 and evaluated accumulation of methylecgonone, via LC-MS/MS analysis of culture supernatants after 72 hours growth of transformed strains in selective media (Urban et al., 1997). Only one of the 24 candidates, EcCYP9 (denoted methylecgonone synthase), resulted in the accumulation of a mass transition ([M+H]+ m/z of 198→166), which corresponds to that of methylecgonone (Chavez et al., 2022; Jirschitzka et al., 2012), and whose retention time matches that of the same mass transition peak produced via expression of tropinone synthase (AbCYP82M3) in the same strain (Chavez et al., 2022) [Figure 32].

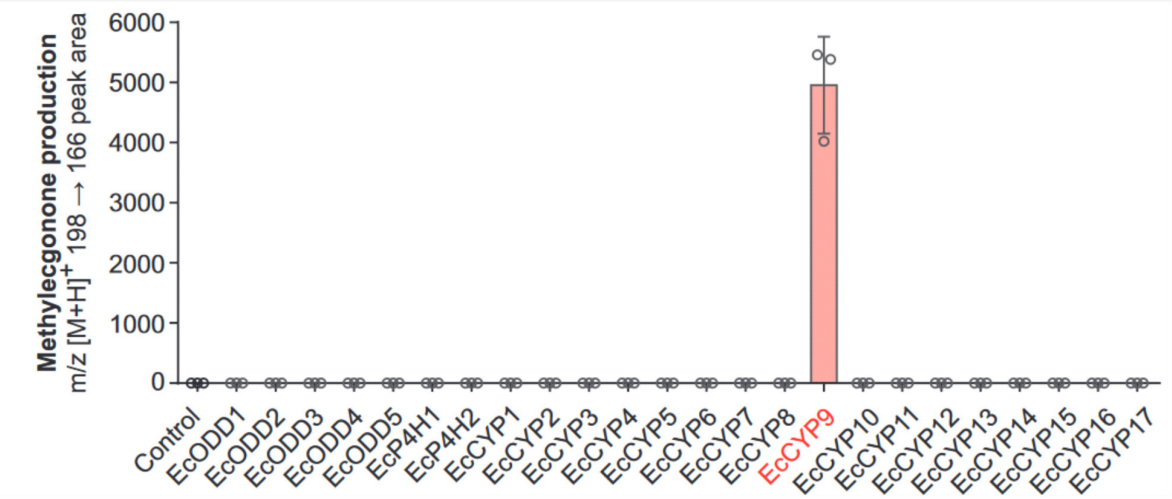

Figure 32: Methylecgonone production in yeast strains engineered for co-expression of hydroxylase/cyclase candidates. Strains were cultured in selective media for 72 hours at 25°C prior to LC-MS/MS analysis of supernatants. Data indicates the mean of n = 3 biologically independent samples (open circles), and error bars show standard deviation. Published in Chavez et al., 2022; Page 8; Figure 4D.

3.7.3 Phylogenetic Analysis of Methylecgonone Synthase from *E. coca*

We analyzed the phylogenetic relationship between the newly identified methylecgonone synthase (EcCYP9), tropinone synthase (CYP82M3), four previously studied *Erythroxylum* oxygenases of the CYP79D family (Luck et al., 2016), as well as other plant CYP450s [Figure 33]. Methylecgonone synthase, which we have designated as EcCYP81AN15 based on its clustering with members of the CYP81 P450 family. EcCYP81AN15 is closely related to flavonoid/isoflavone 2′-hydroxylases from *Jatropha*

curcas (60.4% identity to XP_012079324), *Manihot esculenta* (60% identity to XP_021631824 and 59% identity to XP_021632242), and *Prunus* species (59%, 58%, and 60% identity to XP_007206553, XP_034218022, and PQQ16710, respectively). EcCYP81AN15 clusters separately from the four previously characterized *Erythroxylum* CYP79D-family aldoxime synthases (CYP79D60-63) and is only distantly related to *A. belladonna* AbCYP82M3.

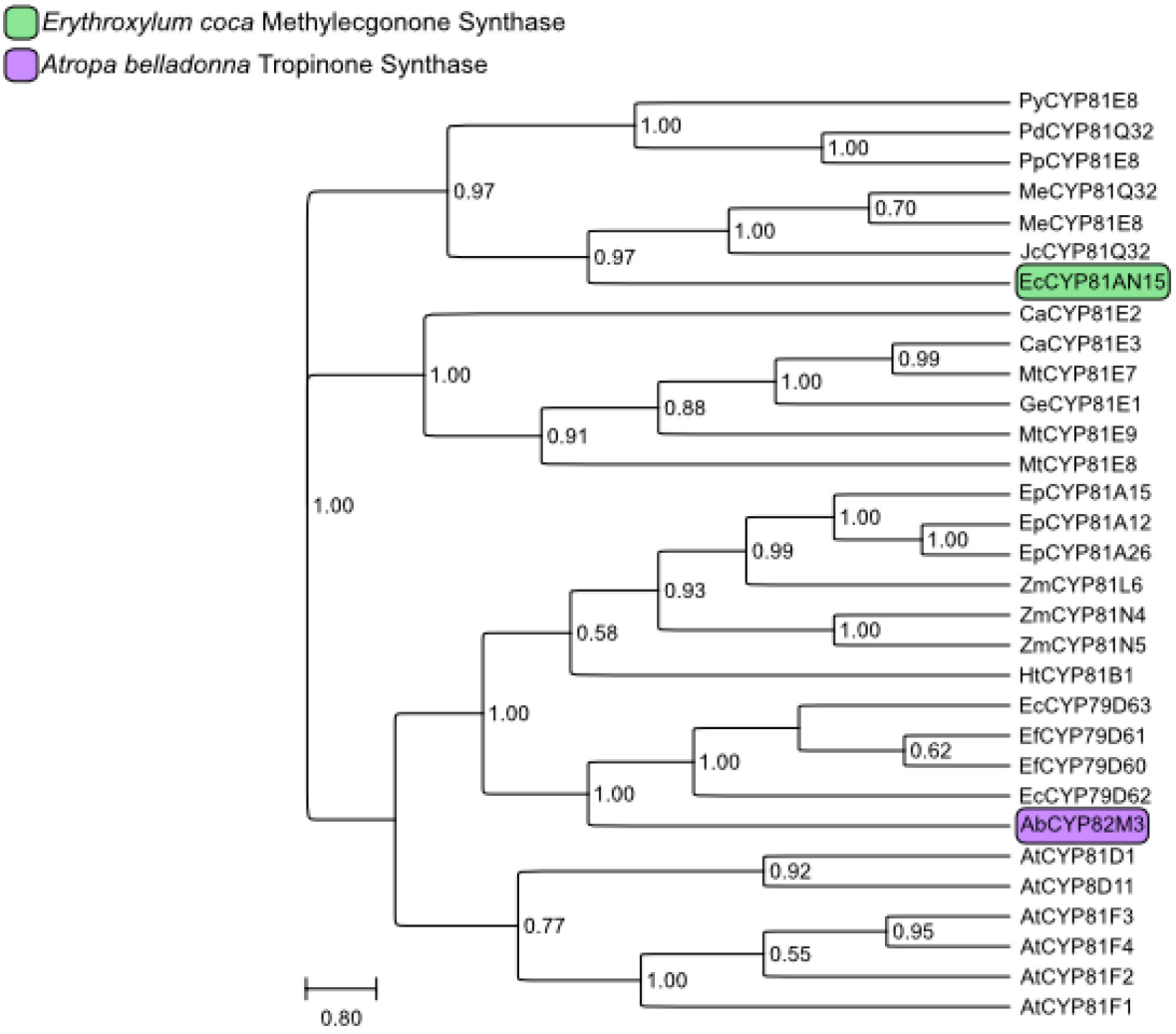

Figure 33: Dendrogram showing phylogenetic analysis of methylecgonone synthase (EcCYP81AN15; previously referred to as EcCYP9). Color scheme: green, EcCYP81AN15; black, previously identified CYP81-family enzymes; purple, *A. belladonna* tropinone synthase. Refer to Table S5 for identities and accession numbers of all enzymes in the dendrogram. Values at each junction signify the number of bootstraps out of n = 1,000 iterations. The scale bar represents the proportion of the total length from the root of the tree to the farthest tip.

3.8 RNA Isolation and cDNA Synthesis from *Erythroxylum coca* leaf tissue

3.8.1 Total RNA Extraction from *Erythroxylum coca*

Total RNA was extracted from L1-L3 leaf tissues of *Erythroxylum* coca. RNA concentrations were quantified using a nanodrop spectrometer [Table S7]. Subsequently, the

total RNA was evaluated on an Agilent 2100 Bioanalyzer for RNA quality [Figure S7, Figure S8, and Table S9]. The NucleoSpin RNA Plant mini kit was supplied with two different lysis buffers called RA1 and RAP, and the manufacturer's protocol suggested that the user figure out which lysis buffer is optimal for different tissue types and plant species. In order to determine the optimal buffer for the RNA extraction, a pilot experiment using L3 stage leaves from *E. coca* was used to evaluate the lysis buffer efficiency of RA1 and RAP. RA1 demonstrated higher RNA concentration from L3 stage leaves, with RNA concentration from the RA1 lysis buffer having approximately four to five-fold higher RNA concentration than RNA concentrations when using the RAP lysis buffer, as outlined in [Table S9]. After determining the optimal lysis buffer, we collected all leaf stages from L1-L3 [

Figure **34**] and extracted the RNA from these leaf stages using the RA1 buffer. cDNA was synthesized from *E. coca* leaf tissue using either NEB Lunascript RT Supermix or ThermoScientific SuperScript IV Reverse Transcriptase cDNA Synthesis of *Erythroxylum coca*. Following the manufacturer's protocol from NEB, cDNA from *E. coca* total RNA was generated using NEB LunaScript® RT SuperMix Kit (Catalog # E3010S). The LunaScript RT supermix includes random hexamer and poly-dT primers to capture the entire transcriptome. A total of 300ng of isolated total RNA was converted into cDNA [Table S8]. We also performed another cDNA synthesis using SuperScript IV Reverse Transcriptase, selecting for mRNA-specific transcripts by utilizing an oligo dT primer.

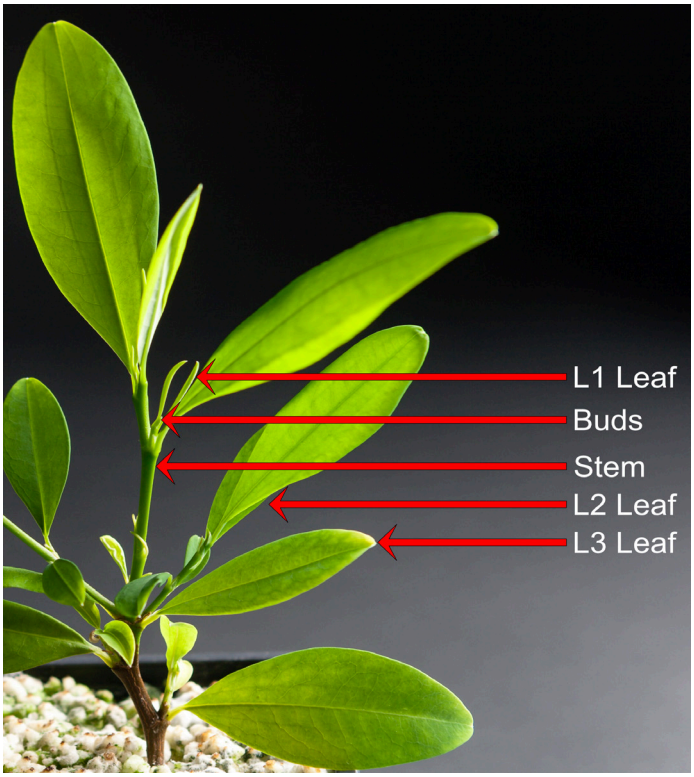

Figure 34: Photograph of *E. coca* plant showing the different developmental stages of the leaves. The L1 leaf stage is a rolled leaf, the L2 leaf stage is an unrolled leaf, and the L3 leaf stage is a mature leaf. Stems and buds are also indicated in this figure. This figure is from Chavez et al. 2022.

3.8.2 Primer Optimization for qPCR Amplicon Targets

qPCR primers were designed in Benchling using the qPCR intercalating dye parameters. Primers were tested in a PCR optimization test to ensure that the proper banding showed up with the expected amplicon size [Table S10]. The NEB SuperMix RT generated cDNA from L3 stage leaves from *E. coca* in a single-step process that contains hexamers and oligo dT for total transcriptome conversion. Evaluation of qPCR gene targets with gene-specific primers for single bands and amplicons of the predicted size. PCR products were evaluated via 4% agarose gel for analysis.

The DNA gel separation from the amplification gave a stronger band signal for the genes involved in the second ring closure steps of TA biosynthesis, including EcMPOBMT, EcCYP81AN15, and EcMecgoR [

Figure **35**]. Ec6409 and Ec10131 reference genes were selected because they are reported to be suitable reference genes across different tissue types of *Erythroxylum coca* (Docimo et al., 2013). The reference controls Ec6409 (lane 10), and Ec10131 (lane 11) gave

strong bands, although the Ec6409 reference gene has multiple banding, suggesting an off-target amplification has occurred.

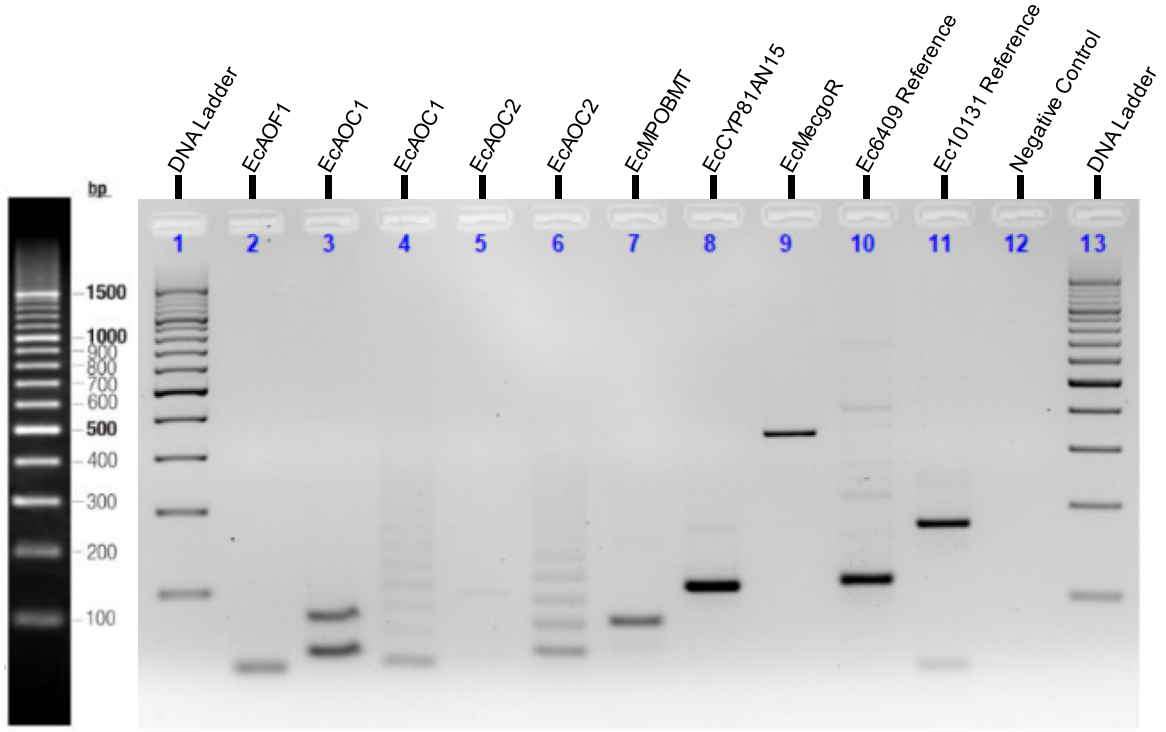

Figure 35: 4% Agarose DNA gel of qPCR primer optimization of L3 stage cDNA from *Erythroxylum coca* generated by NEB Lunascript Reverse Transcriptase. The DNA ladder is the O'RangeRuler 100 bp DNA Ladder (Catalog # SM0623).

The transcripts involved in the beginning steps of TA biosynthesis did not yield the proper amplicon size. EcAOC1 and EcAOC2 transcripts had multiple bands, indicating that the initial primers may have improperly targeted the genes of interest [

Figure **35**]. In parallel, we generated cDNA from L3 stage leaves from *E. coca* RNA via SuperScript IV reverse transcriptase and compared the PCR amplification between the two reverse transcriptase kits using the same gene-specific primers. cDNA generated from SuperScript IV (SSIV) reverse transcriptase did not utilize random hexamers but relied on oligo dT to convert messenger RNA (mRNA). PCR performed on cDNA from SSIV reverse transcriptase generally had a much cleaner amplicon band, although some off-target bands still exist. Moreover, the genes involved in the earlier stage of tropane alkaloid biosynthesis did not have the correct amplicon size, including EcAOF1, EcAOC1, and EcAOC2 gene targets [

Figure **36**]. The reference genes Ec6409 and Ec10131 gave single bands of the correct amplicon size.

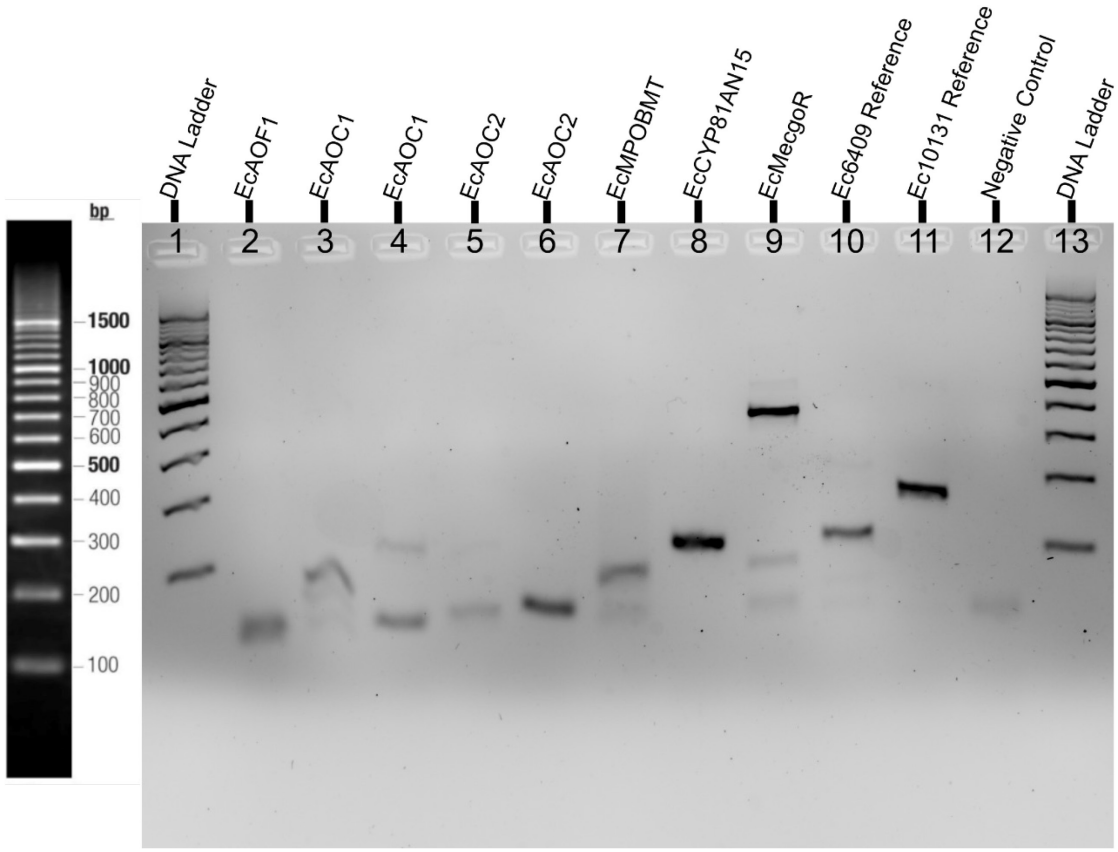

Figure 36: 4% Agarose gel of qPCR primer optimization of L3 stage cDNA from *Erythroxylum coca* generated by SuperScript IV Reverse Transcriptase ThermoScientific. The DNA ladder is the O'RangeRuler 100 bp DNA Ladder (Catalog # SM0623).

The decision to use cDNA from young leaf tissue was based on previous qPCR analysis on spermidine *N*-methyltransferase (EcSMT) genes, which were expressed highest in the buds and young leaf tissues of *Erythroxylum coca*, specifically L1 stage leaves (Anderson, 2019). Using RNA isolated from *E. coca* L1 leaves, cDNA was generated using SSIV reverse transcriptase using oligo dTs, and PCR amplification was performed and subsequently evaluated. The amplification of our genes of interest yielded bands of the correct size, specifically EcAOF1, EcAOC (lane 3), EcAOC2 (lane 5), EcMPOBMT (lane 7), EcCYP81AN15 (lane 8), and EcMecgoR (lane 9) [Figure 37].

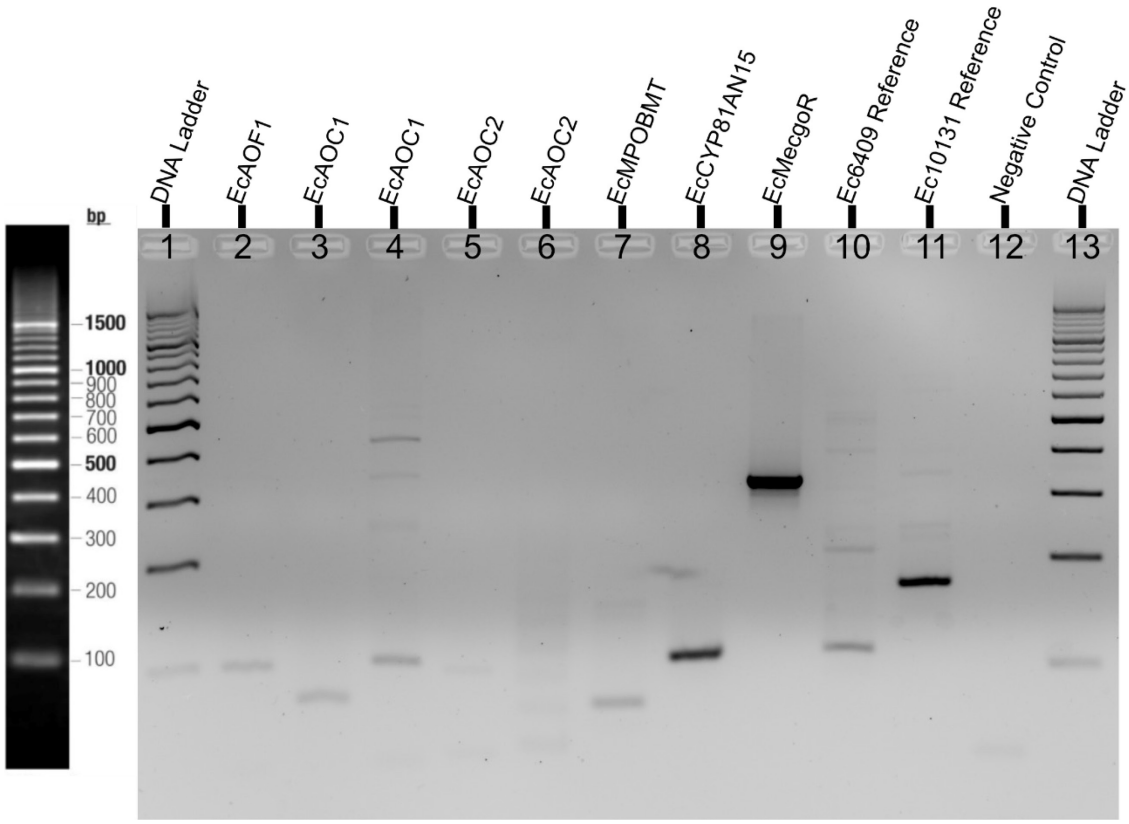

Figure 37: 4% Agarose DNA gel of qPCR primer optimization of L1 stage cDNA from *Erythroxylum coca* generated by ThermoScientific SuperScript IV Reverse Transcriptase. The DNA ladder is the O'RangeRuler 100 bp DNA Ladder (Catalog # SM0623).

EcAOC1 (lane 4) and EcAOC2 (lane 6) have the presence of multiple bands or a smearing effect, indicating that these primer pairs might not be optimal for future qPCR experiments. The reference gene Ec10131 was a clean single amplicon, while Ec6409 had multiple bands present, indicating a less than optimal targeting with the gene-specific primers. Further optimizations are necessary before proceeding with any qPCR detection of gene expression levels.

3.9 Utilizing an LC-MS Method without Derivatizing Polyamines for Detection

3.9.1 HPLC-FLD Detection of Derivatized Polyamine Standards

Previous work on the derivatization of polyamines was performed with an Ultimate-3000 HPLC (Thermo Scientific) with a fluorescence detector (337nm Emission/ 454nm Excitation) with a Nucleodur Sphinx column (Macherey Nagel) generated chromatograms peaks that were sharp and reproducible (Anderson, 2019; Hanczkó et al., 2007). However, moving this method to another HPLC instrument led to a complete change in the reliability of polyamine detection. When utilizing a Waters Alliance HPLC 2695 coupled to a Jasco FP-1520 fluorescent detector led to noisy fluorescent signal presence near the retention time of *N*-

methylspermidine, creating a significant background and limiting the detection range, in some cases, up to 10-15% of the dynamic range of the Jasco FP-1520 fluorescent detector which has a detection range of a thousand fluorescent units was taken up by noisy signal present in the OPA-FMOC blanks [Figure 38].

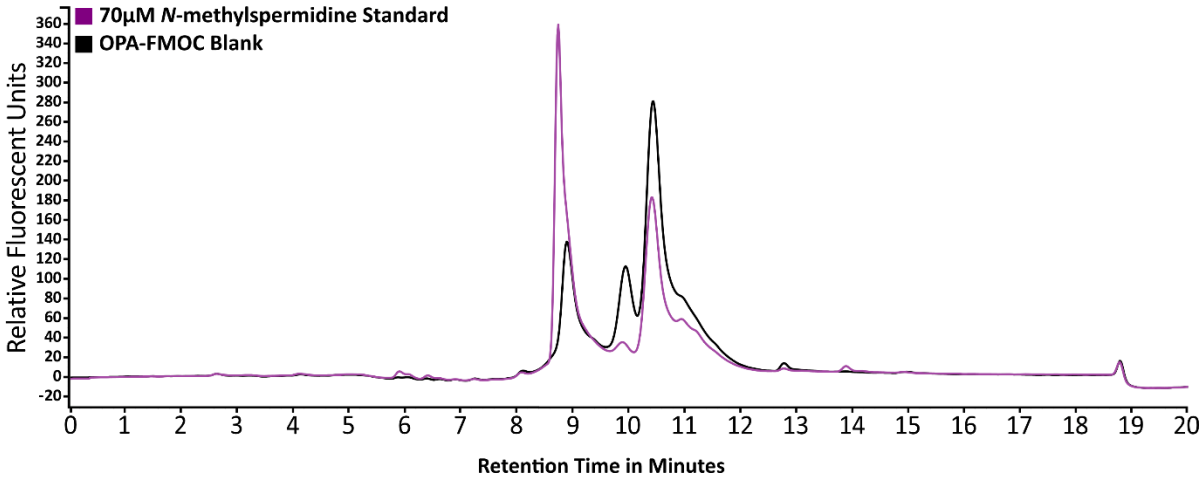

Figure 38: HPLC-FLD chromatogram of *N*-methylspermidine standard compared to OPA-FMOC derivatization blank. Issues quantifying *N*-methylspermidine standard (purple peak at ~8.6 minutes) with fluorescence detection due to high OPA-FMOC Blank background (black peak shapes). All peak heights are shown to scale with respect to all other peaks. HPLC chromatography data was visualized in Waters Empower 3 software and rendered using Serif Affinity Designer.

3.9.2 UPLC Separation and Detection of Derivatized Polyamine Standards

We adapted the polyamine separation method to an Acquity UPLC system (Waters Corporation, Milford, MA, USA) coupled with an Acquity Fluorescence Detector to see if we could overcome this noisy background observed with the HPLC method. An Acquity UPLC BEH C18 130 Å, 1.7 µM, 2.1 x 50 mm column was used, and a modified gradient was combined with the polyamine derivatization protocol (Hanczkó et al., 2007; Zierer et al., 2016). Although measurements of the spermidine and *N*-methylspermidine standards gave clean peak shapes and good separation [Figure 39], the derivatized standards were sticking to the column. The derivatized *N*-methylspermidine and spermidine standards created carryover from the previous injection [Figure 40]. This carryover created issues in determining if these derivatized polyamine peaks are genuinely present in the sample of interest or only carryover from the previous sample injection. Despite multiple washes with acetonitrile, *N*-methylspermidine and spermidine standards stuck to the BEH C18 column.

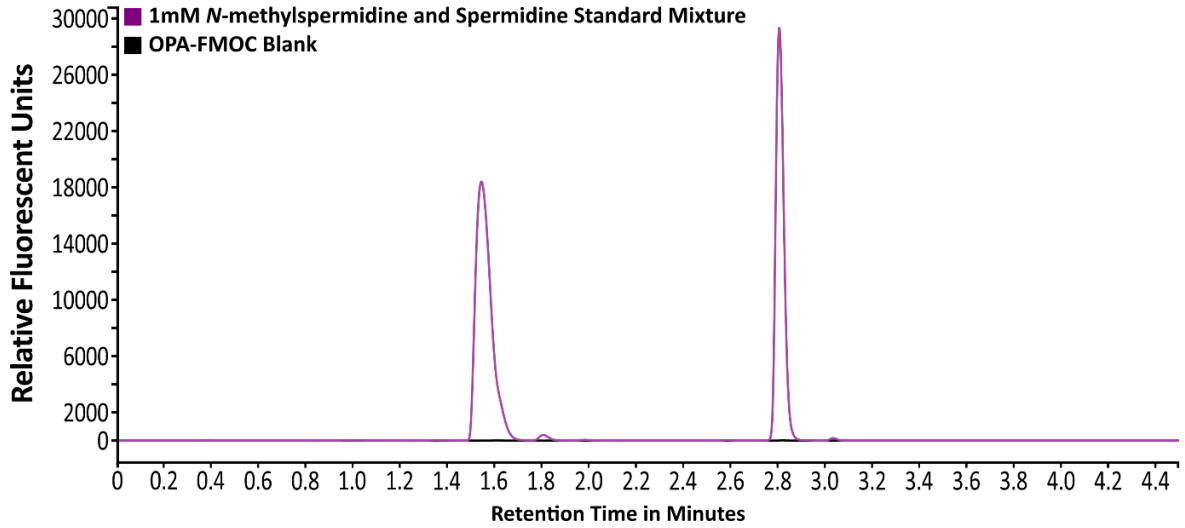

Figure 39: UPLC-FLD chromatogram of *N*-methylspermidine and spermidine standard compared to OPA-FMOC derivatization blank. Derivatization of polyamine standard mixture containing *N*-methylspermidine (first peak around ~1.6 minutes) and spermidine (retention time around 2.80 minutes) standards at a concentration of 1mM detected by UPLC fluorescence detection. All peak heights are shown to scale with respect to all other peaks. UPLC chromatography data was visualized in Waters Empower 3 software and rendered using Serif Affinity Designer.

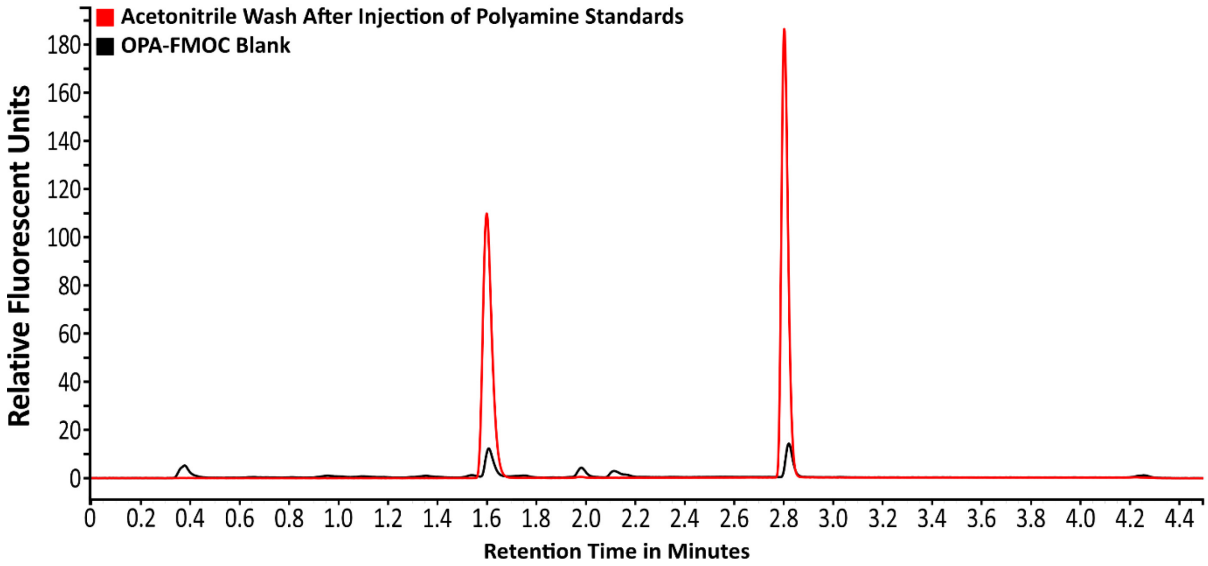

Figure 40: UPLC-FLD chromatogram of acetonitrile wash after injecting the derivatized polyamine standards of *N*-methylspermidine (first peak around ~1.6 minutes) and spermidine (retention time around 2.80 minutes) standards by UPLC fluorescence detection. All peak heights are shown to scale with respect to all other peaks. UPLC chromatography data was visualized in Waters Empower 3 software and rendered using Serif Affinity Designer.

3.9.3 LC-MS/MS Separation and Detection of Polyamine Standards without Derivatization Techniques

Due to this lack of reliability to quantitatively measure *N*-methylspermidine via fluorescent detection via both HPLC and UPLC C18-based column separation methods, we turned our attention to a new method of separating polyamines without requiring derivatization

reactions to visualize polyamines (Su et al., 2021). This LC-MS/MS method utilizes an Acquity BEH Amide column that uses Hydrophilic Interaction Liquid Chromatography (HILIC) chemistry to separate highly polar compounds like polyamines, amino acids, sugars, and organic acids. We evaluated the performance and detection of polyamine standards via multiple reaction monitoring (MRM) using an Agilent 1260 Infinity Binary HPLC coupled to an Agilent 6495 triple-quadrupole mass spectrometer for detecting underivatized polyamine standards. Initial testing with the BEH Amide column demonstrated that putrescine, spermidine, and *N*-methylspermidine standards had consistent retention times and suitable peak shapes [Figure 41]. Further analysis of *N*-methylspermidine using triple quad mass spectrometry revealed that *N*-methylspermidine is approximately ten times lower in the detection limit than spermidine [Figure 41].

Figure 41: LC-MS/MS detection of polyamine standard mixture containing putrescine (red peak) and spermidine (green peak) at 10μM concentration along with 100μM *N*-methylspermidine (purple peak). All peak heights are shown to scale with respect to all other peaks. Putrescine, *N*-methylspermidine, and spermidine were detected at MRM transitions of 89.1→72.1 m/z, 160.0→72.1 m/z, and 146.1→72.1 m/z, respectively. MRM Transitions

visualized in Agilent Qualitative Analysis Version B.06.00 and graphically rendered using Serif Affinity Designer.

Determining the limits of detection for spermidine and *N*-methylspermidine were performed using this HILIC approach. The limit of detection (LOD) for spermidine was approximately 2.5µM, while the LOD of *N*-methylspermidine was observed to be approximately 25µM. The linearity of spermidine was observed from 2.5µM to 100µM with an R^2 = 0,9923. *N*-methylspermidine linearity was observed from 25µM to 250µM with an R^2 = 0,9993 [Figure 42 and Figure 44]. While the retention time for *N*-methylspermidine and spermidine are close, 3.3 minutes and 3.5 minutes, respectively, the parent ions of spermidine ([M+H]+ at 146.1 m/z) and *N*-methylspermidine ([M+H]+ at 160.0 m/z) are different enough in mass to be distinguished by the triple quadrupole mass spectrometer [Figure 43 and Figure 45].

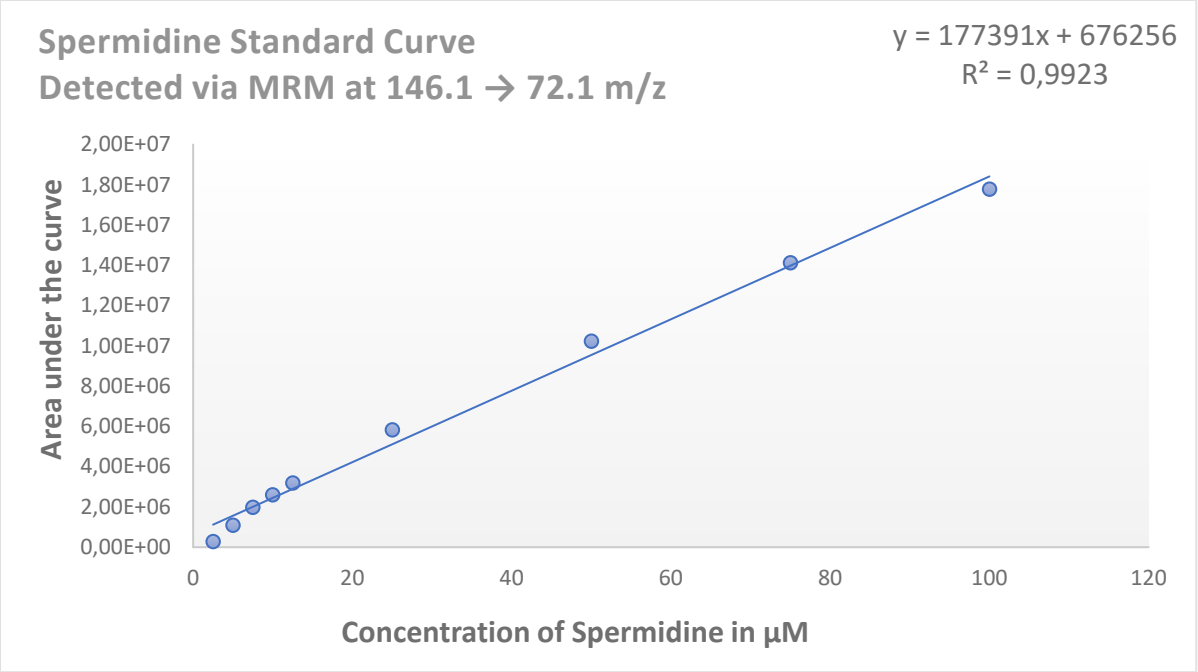

Figure 42: Linear regression curve of the peak area of spermidine standard MRM detection transitions at 146.1→72.1 m/z.

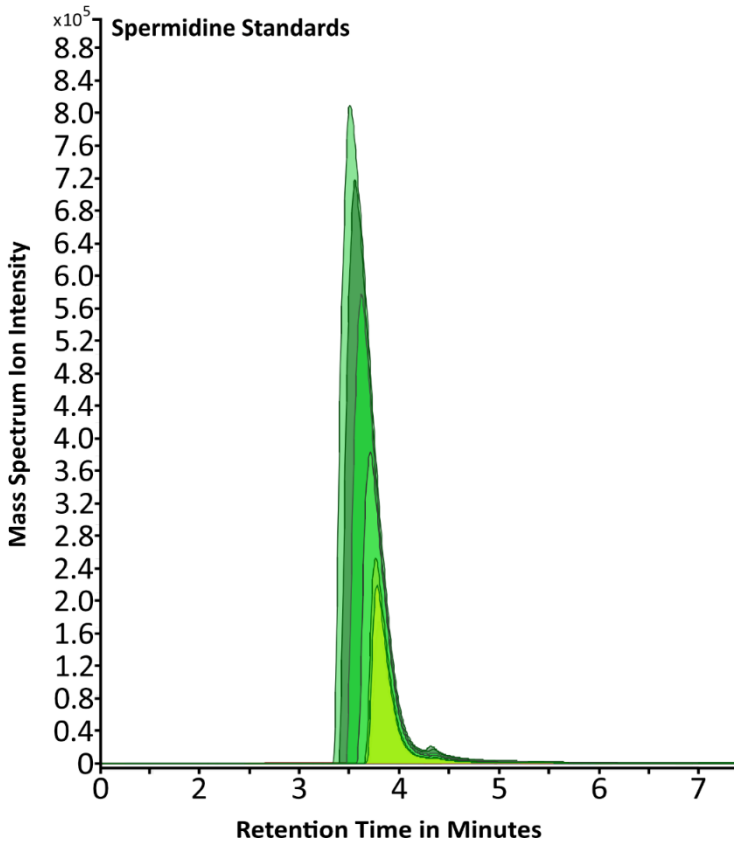

Figure 43: LC-MS/MS detection of MRM transitions at 146.1→72.1 m/z corresponds to the spermidine standard at varying concentrations from 2.5µM-100µM. All peak heights are shown to scale with respect to all other peaks. MRM Transitions visualized in Agilent Qualitative Analysis Version B.06.00 and graphically rendered using Serif Affinity Designer.

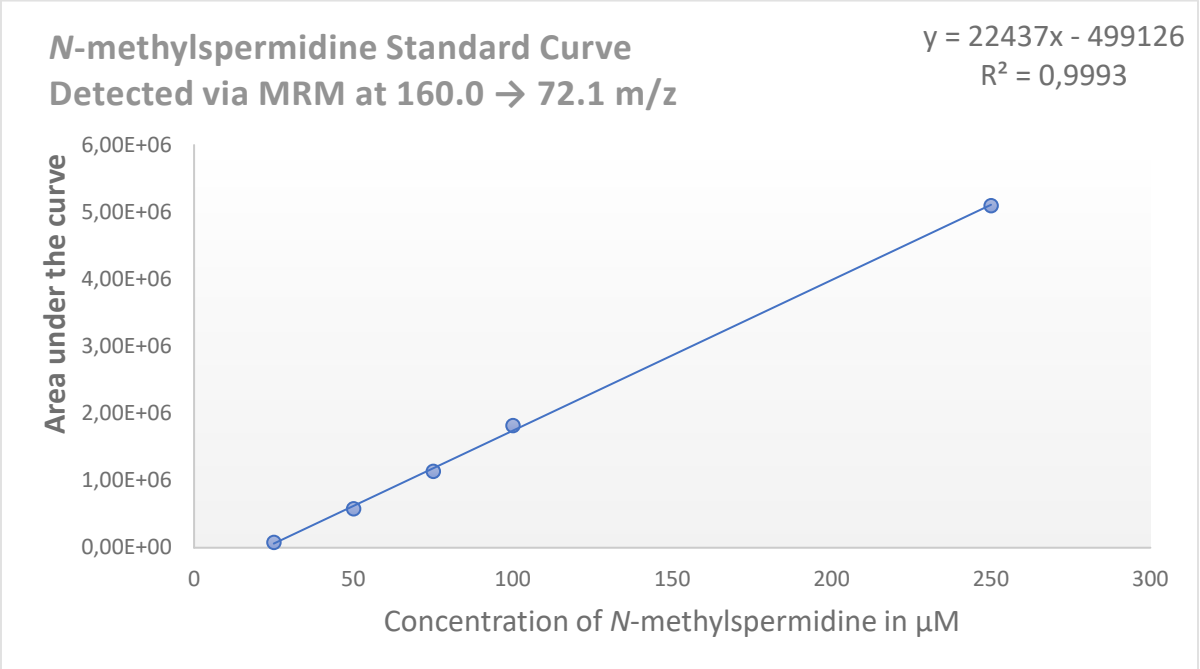

Figure 44: Linear regression curve of the peak area of *N*-methylspermidine standard MRM detection transitions at 160→72.1 m/z.

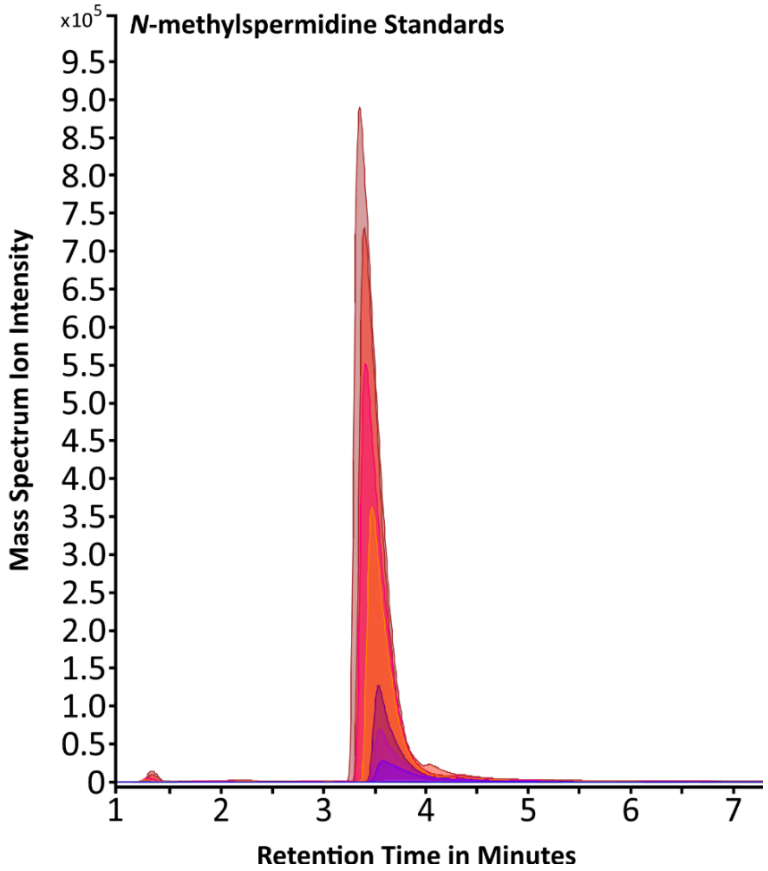

Figure 45: LC-MS/MS detection of MRM transitions at 160.0→72.1 m/z corresponds to the *N*-methylspermidine standard at varying concentrations from 25μM-0.5mM. All peak heights are shown to scale with respect to all other peaks. MRM Transitions visualized in Agilent Qualitative Analysis Version B.06.00 and graphically rendered using Serif Affinity Designer.

We also tested the separation and detection of *S*-adenosyl-L-methionine (SAM) and *S*-adenosyl-L-homocysteine (SAH) on this BEH amide HILIC column. We observed sharp peaks along with stable retention time of SAM and SAH [Figure 46]. The ability to separate SAM and SAH is crucial. SAM is the cofactor for methylation reaction, and SAH accumulates as a product after methyltransferase activity.

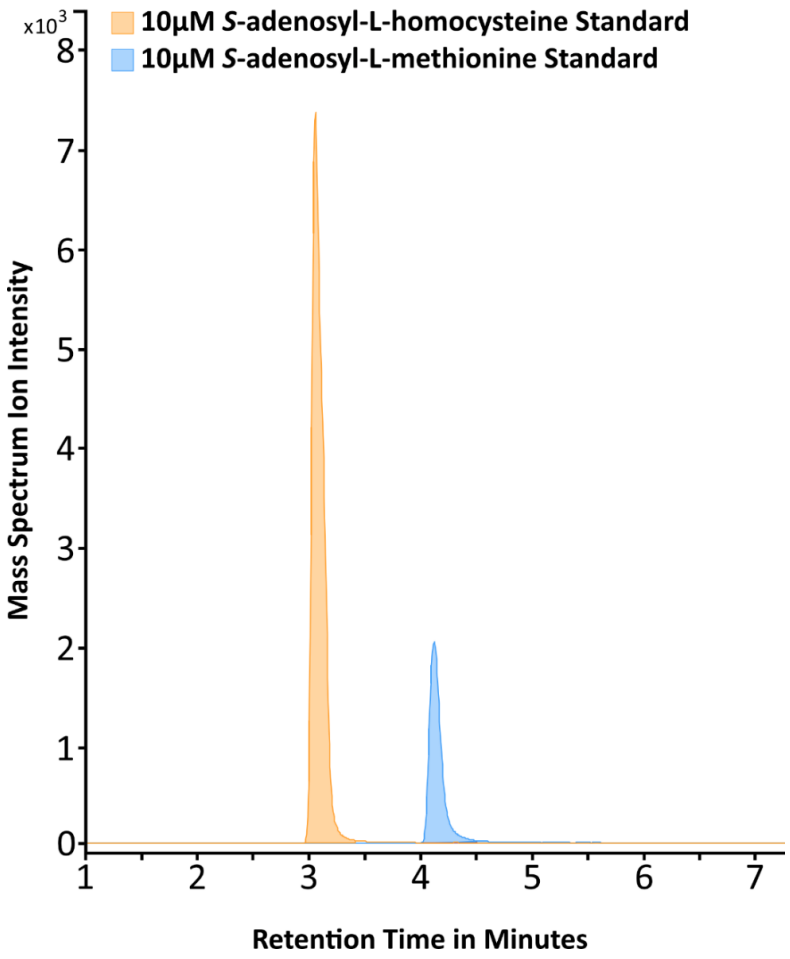

Figure 46: LC-MS/MS MRM separation of the standards *S*-adenosyl-L-homocysteine (orange peak) and *S*-adenosyl-L-methionine (blue peak) at 10µM concentration. All peak heights are shown to scale with respect to all other peaks. MRM Transitions visualized in Agilent Qualitative Analysis Version B.06.00 and graphically rendered using Serif Affinity Designer.

After protein expression and purification of EcSPMT from *K. phaffii* KM71 cells, an enzyme assay was performed to test the spermidine *N*-methyltransferase activity of EcSPMT. The cofactors and substrates separated on the BEH amide column detecting metabolites produced from the EcSPMT enzyme assay were evaluated. The SMT assay conditions are described below [Table 2].

Enzyme Assay Components	Amount in µL	Final Concentration
50mM Potassium Phosphate buffer pH = 7.6	69	34.5 mM
EcSPMT Purified Enzyme	7	1 µg/mL
50mM Ascorbic acid stock solution	2	1mM
100mM Dithiothreitol stock solution	2	2mM
100mM Spermidine stock solution	10	10mM
10mM SAM stock solution	10	1mM
Final Volume	100	

Table 2: Reaction setup for EcSPMT enzyme assay to evaluate SMT activity.

The SMT assay showed that the conversion of SAM cofactor into SAH was detected via LC-MS/MS, indicating successful methyltransferase activity and the formation of *N*-methylspermidine [Figure 47]. While the EcSPMT enzyme assay was successful, the peaks of spermidine and *N*-methylspermidine have broad shoulders, indicating that too much aqueous solution from the enzyme assay was present.

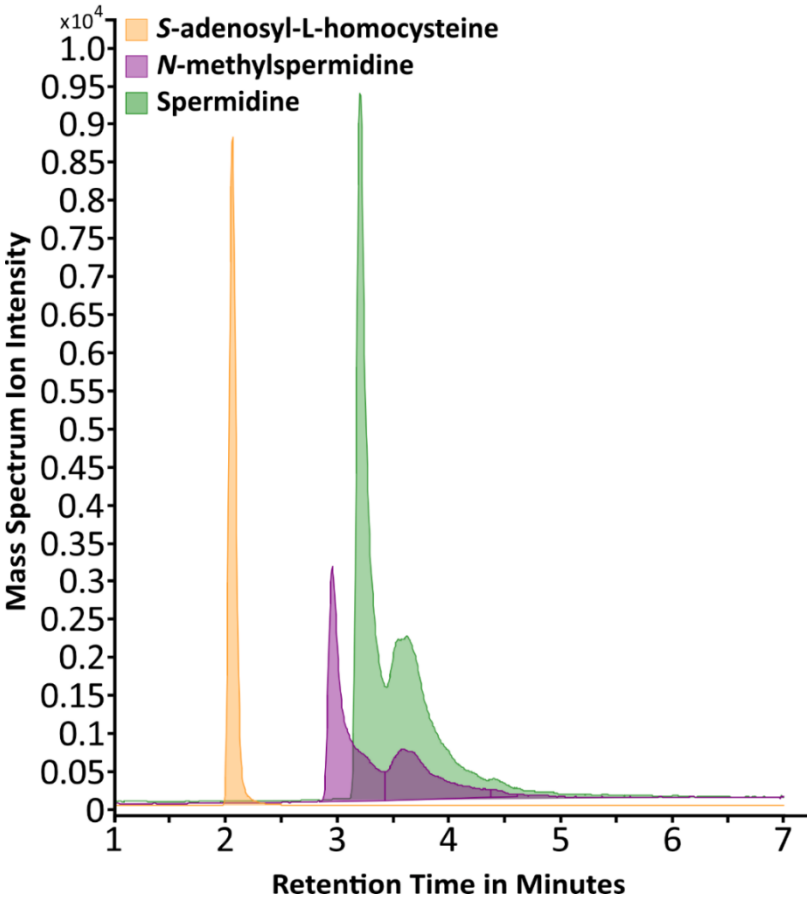

Figure 47: LC-MS/MS MRM detection of substrates and products generated from EcSMT enzyme assay. SAH (orange peak), *N*-methylspermidine (purple peak), and spermidine (green peak). All peak heights are shown to scale with respect to all other peaks. MRM Transitions visualized in Agilent Qualitative Analysis Version B.06.00 and graphically rendered using Serif Affinity Designer.

HILIC chromatography is sensitive to high amounts of water, which could cause the retention time to shift and shoulders to form on peaks like spermidine and *N*-methylspermidine. Additionally, the detection of SAM cofactor via LC-MS/MS was barely detectable even in the no enzyme control [Data Not Shown]. Heat inactivation was used to stop the enzyme assay, which likely contributed to the degradation of the SAM cofactor, as it is known to be unstable

in both liquid and solid forms (Morana et al., 2002). In preparation for enzyme kinetics, future enzyme assays will use alternative ways to stop enzyme activity, such as chemical inactivation.

4 DISCUSSION

4.1 Identification of Missing Steps of the Tropane Alkaloid Pathway in *E. coca*

Tropane alkaloids are specialized metabolites found in multiple plant families amongst angiosperms. Due to their accessibility, the Solanaceae and Convolvulaceae families have been the main focus of tropane alkaloid research. *Erythroxylum coca*, a member of the Erythroxylaceae family, also produces valuable TAs such as cocaine; despite having an infamous association in illicit drug markets, it still has value in medicinal applications in eye surgeries due to its vasoconstrictor and painkilling properties. Due to restrictions on *E. coca* plants, research into understanding how tropane alkaloid biosynthesis operates has progressed slower than other TA-producing plants. Our work aimed to resolve the missing steps of tropane alkaloid biosynthesis leading to the formation of cocaine in *Erythroxylum coca*. Utilizing a synthetic biology approach of engineering yeast platform in combination with *in planta* validation of key biosynthetic steps has uncovered the missing steps of cocaine biosynthesis (Chavez et al., 2022). The discovery of polyamine-related enzymes involved in the first committed steps of TA biosynthesis shows a clear divergence from the canonical putrescine *N*-methyltransferases steps operating in Solanaceae and Convolvulaceae plant families. Unlike TA biosynthesis in Solanaceae and Convolvulaceae, in which putrescine *N*-methyltransferase mediates, this PMT step is missing in *Erythroxylum coca*. Since spermidine *N*-methyltransferases catalyze the formation of *N*-methylspermidine in *E. coca* in place of a functional PMT enzyme, we sought to understand the roles that *N*-methylspermidine plays in tropane alkaloid biosynthesis in Erythroxylaceae. From bioinformatic identification of candidate genes to functional validation of the pathway in a yeast synthetic biology platform and complemented in planta validation, we finally identified the minimal set of genes involved in cocaine biosynthesis in *Erythroxylum coca*. A full tropane alkaloid biosynthetic pathway shows the differences and similarities between *Atropa belladonna* and *Erythroxylum coca* [Figure 48].

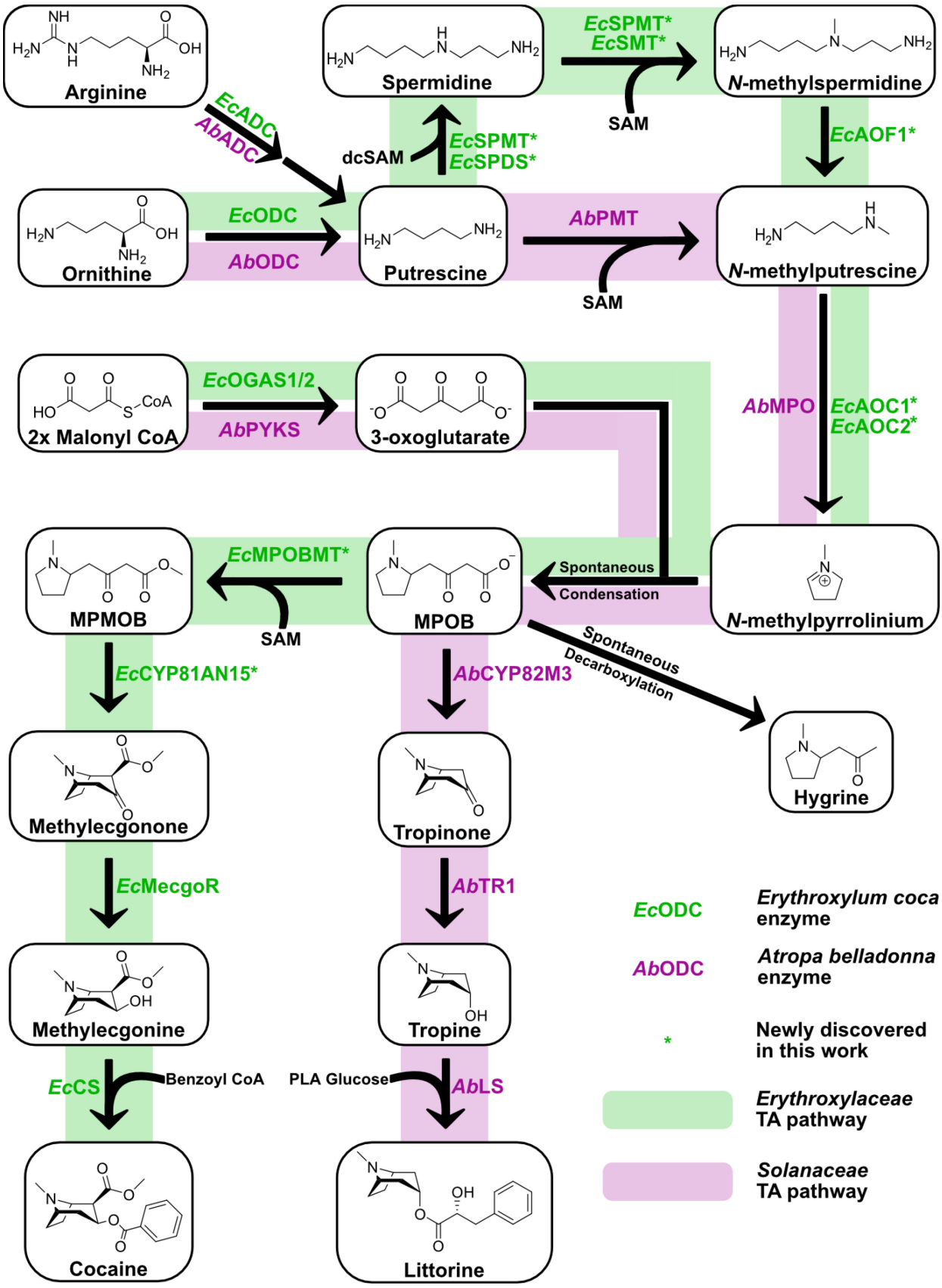

Figure 48: A comparison of tropane alkaloid pathway in Erythroxylaceae (green color) and Solanaceae (purple). Modified from Chavez et al., 2022; Page 2, Figure 1.

4.2 Spermidine is Essential for Growth and Development in Eukaryotes

We have established spermidine as an essential intermediate in tropane alkaloid biosynthesis in *Erythroxylum coca*. Spermidine is essential for plant growth as it regulates numerous biosynthetic processes. One particular biosynthetic process that relies on spermidine as the substrate is the formation of the rare amino acid hypusine. Hypusine is involved in a posttranslational modification known as hypusination. This modification is unusual because it is limited to the eukaryotic translation factor 5A (eIF5A), which is ubiquitous and highly conserved across eukaryotes. In plants, eIF5As are connected with multiple cellular processes, including regulation of protein synthesis, translation elongation, regulation of messenger RNA (mRNA) decay, seed yield, senescence, and programmed cell death (Pálfi et al., 2021). Spermidine's association with hypusination posttranslational modifications helps explain why significant growth defects are observed when spermidine biosynthesis is disrupted or impaired in eukaryotes.

Additionally, this supports previous observations of the biosynthesis of spermidine being essential, while the biosynthesis of spermine was not essential to the growth and development of *Arabidopsis thaliana* (Imai, Akiyama, et al., 2004; Imai, Matsuyama, et al., 2004). RNAi-mediated knockdown of deoxyhypusine synthase in *Arabidopsis thaliana* affected several plant biological mechanisms, such as flowering time and aerial and root tissue morphology alterations (Belda-Palazón et al., 2016). The role of spermidine in plant metabolism continues to be uncovered as research progresses on polyamines and their biological functions in plants' primary and secondary metabolism.

4.3 Evolutionary Insights on EcSPMT Versus PMT Steps in TA Biosynthesis

The dual activity of EcSPMT is the first ever reported spermidine *N*-methyltransferase and spermidine synthase (Chavez et al., 2022). However, there have been reports of aminopropyltransferase activity in the gymnosperm *Pinus sylvestris,* which can make spermidine and spermine depending on substrate availability, demonstrating substrate flexibility amongst SPDS enzymes in plants (Vuosku et al., 2018). The Erythroxylaceae and Solanaceae families are members of the rosids and asterids clades, respectively. The last common ancestor between Erythroxylaceae and Solanaceae was speculated to have lived approximately 120 million years ago (Magallon & Castillo, 2009; Y.-J. Wang et al., 2023). Several whole genome duplications and triplication events occurred after the asterid and the

rosids split from the core eudicots. In the case of the *Erythroxylum coca* var *novogranatense,* it appears that the whole genome duplication happened at least once (Y.-J. Wang et al., 2023).

Recently, genomic assembly of the *Erythroxylum coca* var. *novogranatense* provided insights into these novel spermidine *N*-methyltransferases. Wang et al. demonstrated that the SPMT gene identified in *E. novogranatense* was lost in the asterid while preserved in the rosids clade (Y.-J. Wang et al., 2023). *Erythroxylum coca*, a member of the rosids clade, preserved this bifunctional SPMT gene, while the asterid clade, which includes Solanaceae and Convolvulaceae families, have lost this SPMT gene. Following the loss of the SPMT gene in asterid led to the independent evolution of putrescine *N*-methyltransferase activity through gene duplication and neofunctionalization of spermidine synthase, providing additional evidence that TA biosynthesis in plants is polyphyletic. Further studies on plants within the Solanaceae family show loss/gain of function of critical TA pathway genes, supporting the idea of paraphyletic origins of tropane alkaloid biosynthesis even within the Solanaceae family (J. Yang et al., 2023).

To date, putrescine *N*-methyltransferase (PMT) enzymes have only been found in the Solanaceae and Convolvulaceae plant families, suggesting that other tropane-producing plants outside of Solanaceae and Convolvulaceae could be following a pathway similar to Erythroxylaceae, forming *N*-methylspermidine as the first committed step of tropane alkaloid biosynthesis. Future characterization of novel tropane alkaloid biosynthetic pathways should consider the possibility that spermidine *N*-methyltransferases (SMTs/SPMTs) might be involved in the first committed steps of tropane alkaloid biosynthesis from plants outside of Solanaceae and Convolvulaceae plant families.

4.4 EcSPMT Possesses Both Spermidine Synthase and Spermidine *N*-methyltransferase Activities

Spermidine *N*-methyltransferases (SMT) are a new class of *N*-methyltransferases and are the first committed step of TA biosynthesis in *Erythroxylum coca* (Anderson, 2019; Chavez et al., 2022). SMT enzymes utilize SAM as a cofactor to methylate spermidine directly. In *E. coca*, two enzymes, EcSMT and EcSPMT, possess this novel SMT activity [Figure 49]. EcSMT has spermidine *N*-methyltransferase activity exclusively, while EcSPMT has retained both spermidine synthase and spermidine *N*-methyltransferase activity. EcSPMT's dual activity is unique because it can use the cofactors SAM or dcSAM in conjunction with the substrate putrescine or spermidine, generating spermidine or *N*-methylspermidine as products.

This flexibility of substrate and cofactor is not currently observed with any reported spermidine synthase-related enzymes. Spermidine synthase enzymes display a preference for dcSAM as the optimal cofactor. Although dcSAM and SAM differ only by a carboxyl group, the absence of that carboxyl group seems to allow dcSAM to bind more efficiently, indicating that a structural domain is present in SPDS enzymes that mediate this cofactor binding preference. Unlike SPDS enzymes, putrescine *N*-methyltransferase (PMT) enzymes will bind SAM preferentially over dcSAM (Junker et al., 2013). EcSPMT, which can bind SAM and dcSAM cofactors, is a structure-function model for understanding this novel dual cofactor binding site mechanisms.

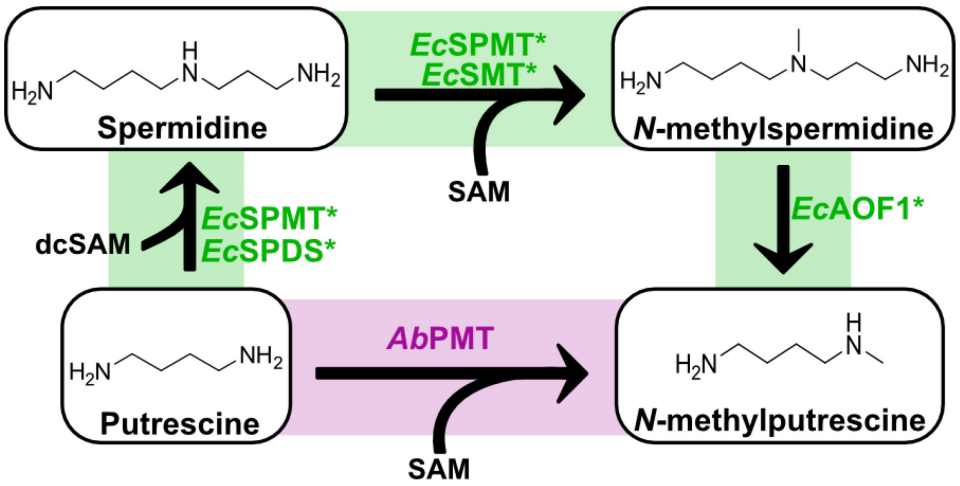

Figure 49: Depicting the differences between the beginning steps of tropane alkaloid biosynthesis in Solanaceous plants in purple and *Erythroxylum coca* in green color.

EcSPMT, when expressed in *Saccharomyces cerevisiae,* seemingly lost the spermidine *N*-methyltransferase activity previously observed during the *in vitro* enzyme activity assays. There can be several reasons why this functionality was lost. Despite being a eukaryotic host, the yeast expression system still does not possess the same posttranslational modifications that occur in plants. The reason for EcSPMT loss of dual activity is currently not understood but could be related to differences in substrate/cofactor availability or cellular environment, such as posttranslational modifications. AtSPDS1 shares a greater than 70% identity with EcSPMT. AtSPDS1 localizes within the nucleus of *Arabidopsis thaliana* despite not having a classical nuclear localization signal (Sekula & Dauter, 2019a). Subcellular localization prediction programs can help understand signal peptides in protein sequences responsible for protein localization. The LOCALIZER program focuses on signal peptides in plant and plant pathogen effector proteins commonly trafficked to the mitochondria, chloroplast, and nucleus (Sperschneider et al., 2017). LOCALIZER predicted an NLS sequence in both AtSPDS1 and

EcSPMT at the N-terminal region of the enzyme consisting of a KRXR motif in which X residue is either a proline or alanine residue in AtSPDS1 and EcSPMT, respectively. Future characterization *in planta* should focus on the subcellular localization of EcSPMT to understand whether nuclear localization is observed as seen with other spermidine synthases in *Arabidopsis thaliana*.

4.5 Both Flavin-Dependent and Copper-Dependent Oxidases Are Required for First Ring Closure in TA Biosynthesis in *E. coca*

The conversion of *N*-methylspermidine to *N*-methyl-Δ^1-pyrrolinium cation (NMPy) relies on two types of amine oxidases. The flavin-dependent amine oxidase, EcAOF1, mediates the oxidation of *N*-methylspermidine into *N*-methylputrescine. Expression in *Saccharomyces cerevisiae* and *Nicotiana benthamiana* demonstrated that EcAOF1 efficiently converts *N*-methylspermidine into *N*-methylputrescine. EcAOF1 is novel in that it acts on a tertiary amine by oxidizing the carbon on the exo-side of the methylated N4 atom to produce *N*-methylputrescine and 3-aminopropanal. The degradation of *N*-methylspermidine in *Erythroxylum coca* proceeds with EcAOF1, while in Solanaceae, this flavin-dependent oxidase step is not necessary due to the direct conversion of *N*-methylputrescine to NMPy cation via a copper-dependent methylputrescine oxidase. Since the SPMT gene is missing in plants belonging to the asterid clade, it is logical to assume that the AOF1 gene could also be absent. Investigative work performed by Wang and colleagues confirmed through whole genome sequencing of plants from both asterid and rosids clades that the AOF gene was found only in plants that belong to the rosids clade (Y.-J. Wang et al., 2023). These observations confirm the evolutionary divergence observed in the biosynthetic routes of tropane alkaloids, providing further evidence of polyphyletic origins of tropane alkaloids between Solanaceae and Erythroxylaceae families.

Two copper-dependent amine oxidases, EcAOC1 and EcAOC2, were found to convert *N*-methylputrescine into 4-(methyl(3-oxopropyl)amino)butanal as an intermediate in *Saccharomyces cerevisiae* and *Nicotiana benthamiana*. This intermediate then undergoes a spontaneous cyclization reaction, forming NMPy cation. Radiolabeled feeding studies long supported the first ring closure mechanism of tropane alkaloids in solanaceous plants, but these studies failed to understand that *N*-methylspermidine was the true precursor for TA biosynthesis in *E. coca* (Leete, 1982). Since spermidine can be reversibly converted into putrescine, relying on radiolabeled feeding studies alone to uncover these differences posed a significant challenge. Spermidine was proposed to be a precursor to NMPy in *Nicotiana*

glutinosa, although it was unclear how the reaction operated (Leete, 1985). An asymmetrically radiolabeled spermidine was used and noted that equal incorporation of the radiolabels at C-2′ and C-5′ of the pyrrolidine rings in nicotine and nornicotine, which did not validate his initial hypothesis of seeing incorporation at only the C-2' position of the NMPy molecule. This observation led to the hypothesis that spermidine could be methylated at the terminal end, then oxidized via an unknown oxidation step to *N*-methylputrescine then a methylputrescine oxidase step would lead to first ring closure. This hypothetical pathway proposed more than 40 years ago is strikingly close to the mechanism of first ring closure of tropane alkaloid biosynthesis we observed in *Erythroxylum coca*. In the 1980s, knowledge about the inner workings of TA biosynthesis in *E. coca* was relatively lacking compared to Solanaceae TA biosynthesis. Dr. Edward Leete was even quoted as saying, *"…cumulative failures caused me to question the validity of the proposed biosynthetic route to cocaine. An admittedly outlandish idea was that the tropane skeleton was being produced in the Erythroxylon genus by a biosynthetic process quite different from that operating in the Solanaceae family*" (Leete, 1982). Our discoveries on the inner workings of TA biosynthesis in *Erythroxylum coca* confirm these prior observations. Although EcAOC1 and EcAOC2 cluster in AOC clade III, they are dispersed throughout this clade despite sharing ~84% amino acid identity [Figure 25]. *In vivo* expression in yeast showed a relatively lower accumulation of hygrine from EcAOC1 compared to EcAOC2, suggesting that EcAOC2 may likely contribute to the primary production of NMPy cation [Figure 24]. However, future qPCR gene expression experiments across tissues of *E. coca* are required to address this observation.

Following the formation of the NMPy cation, a spontaneous condensation reaction with 3-oxoglutaric acid yields the intermediate 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate (MPOB). The 3-oxoglutaric acid is generated from two rounds of malonyl-CoA via oxoglutaric acid synthase (OGAS). MPOB formation is one of two existing nonenzymatic reactions in tropane alkaloid biosynthesis. Tian and coauthors reported on identifying pyrrolidine ketide synthases (PYKSs) from *Erythroxylum novogranatense* that produce 3-oxoglutaric acid (Tian et al., 2022). These PYKS enzymes matched 100% amino acid identity to the PYKS candidates, now renamed EcOGAS1 and EcOGAS2, that were isolated previously (Kim, 2020). Transient expression of EcOGAS1/2 demonstrated the functional activity towards the production of hygrine intermediate, indicating the production of 3-oxoglutaric acid was successful. The 3-oxoglutaric acid reacts nonenzymatically with NMPy cation within the tissues of *Nicotiana*

benthamiana, further confirming observations in Solanaceae related to first ring closure mechanisms of tropane alkaloid biosynthesis.

4.6 MPOBMT is Essential for Formation of the 2-carboxymethyl Ester Present in TAs in *Erythroxylum coca*

In *E. coca*, the process of second ring closure follows a pathway similar to Solanaceae, albeit with a slight deviation. After the formation of MPOB in *E. coca*, the MPOB undergoes an *O*-methylation via an MPOB methyltransferase (EcMPOBMT). EcMPOBMT, using SAM as a cofactor, transfers a methyl group to the oxygen atom on the carboxylic acid group of MPOB, forming 4-(1-methyl-2-pyrrolidinyl)-3-methyloxobutanoate (MPMOB). The formation and preservation of the 2-carbomethoxy ester group are distinctive to the tropane alkaloid produced by the Erythroxylaceae family. This methylation is speculated to act as a protecting group, preventing the spontaneous decarboxylation observed in tropinone formation in Solanaceae (Chavez et al., 2022). As Erythroxylaceae appear to have evolved one or more *N*-methyltransferases capable of using spermidine instead of putrescine, a novel enzymatic mechanism is required for the conversion of *N*-methylspermidine to NMPy, which condenses with 3-oxoglutarate produced by type III polyketide synthases (PYKS or OGAS) to form MPOB in both families (J.-P. Huang et al., 2019; Tian et al., 2022).

Reconstruction of the tropane alkaloid pathway in *Saccharomyces cerevisiae* showed direct methylation of MPOB before forming methylecgonone (Chavez et al., 2022). In addition to our discovery, Wang et al., 2022 reconstructed the tropane alkaloid pathway using a combination of Solanaceous and Erythroxylaceae biosynthetic genes and reported that ecgonone, not MPOB, as the substrate for EcMPOBMT led to the formation of methylecgonone. The order of the methylation of MPOB that occurs before or after the P450-mediated second ring closure is currently being debated (Chavez et al., 2022; Y.-J. Wang et al., 2022). While it is not fully understood which steps occur first, it must be acknowledged that *in vitro* and *in vivo* comparisons of enzymatic activities might not be so straightforward. A case in point is the alternative activities of cocaine synthase from *E. coca* when evaluated in both *in vivo* and *in vitro* environments. When first characterized, cocaine synthase was observed to acylate the C3 hydroxyl group of the tropane ring, specifically only when this hydroxyl group is in the β conformation (Schmidt et al., 2015). However, when expressed in *Saccharomyces cerevisiae*, cocaine synthase could also acylate α conformation of hydroxyl groups at the C3 position of the tropane ring, demonstrating a substrate flexibility not observed under *in vitro* conditions (Srinivasan & Smolke, 2019). This discrepancy in enzymatic functions may rely on

the environmental differences between living cells and *in vitro* assay conditions (García-Contreras et al., 2012; van Eunen & Bakker, 2014; Zotter et al., 2017).

4.7 Methylecgonone Synthase (EcCYP81AN15) Facilitates Second Ring Closure of TAs in *E. coca*

Methylecgonone synthase, also known by its P450 family classification EcCYP81AN15, was found to catalyze the second ring closure of tropane alkaloids in the yeast expression system around the same time of our publication; another group of scientists also confirmed our observations that a CYP81AN15 cyclase mediates the second ring closure in *Erythroxylum coca* var *novogranatense* (Y.-J. Wang et al., 2022). Interestingly, Tropinone synthase, AbCYP82M3, could accept MPMOB as a substrate to form methylecgonone, demonstrating the flexibility of substrates that these cytochrome P450s can utilize and may be further utilized in synthetic biology or metabolic engineering approaches (Chavez et al., 2022). Wang and colleagues used homology modeling along with mutational analysis of tropinone synthase and methylecgonone synthase, demonstrating essential residues involved in the binding of MPOB to the active site that contribute to overall enzymatic activity in both CYP82M3 and CYP81AN15 (Y.-J. Wang et al., 2023).

4.8 Tissue-specific Expression of Tropane Alkaloid Biosynthesis Genes

After identifying all the necessary steps for TA biosynthesis in *E. coca* using *Saccharomyces cerevisiae* and *Nicotiana benthamiana*, we explore gene expression of these biosynthetic genes in a heterologous host. Understanding the tissue-specific expression of genes involved in tropane alkaloid biosynthesis in the native plant is crucial. In solanaceous plants, tropane alkaloid pathway genes are mainly expressed in the roots. Although gene expression of tropane alkaloid-related enzymes in solanaceous plants is localized mainly to the roots, the tropane alkaloid metabolites produced are found throughout different tissues of the plant. *Erythroxylum coca* deviates from this observation as its tropane alkaloids are not widely distributed throughout different plant tissues. Instead, tropane alkaloid biosynthesis and metabolite storage occur in the leaf tissues of *E. coca* (Schmidt et al., 2015; Torre et al., 2013). Future investigations should identify gene regulation mechanisms, such as transcription factors and other regulatory elements, that underlie TA biosynthesis in *Erythroxylum* coca and other related TA-producing plants. Currently, qPCR primer tests via Taq-based PCR revealed a discrepancy in these cDNA amplicon targets being in higher abundance in L1 rolled leaf compared to L3 mature leaves in *E. coca* [

Figure **36** and Figure 37]. Future qPCR experiments will help to provide the gene expression map of all identified TA biosynthetic genes across the different tissue types in *E. coca*. Focusing on EcAOF1, EcAOC1, and EcAOC2 will help provide vital information about the expression of genes involved in tropane alkaloid biosynthesis leading to first ring closure.

4.9 LC-MS/MS Separation and Detection of Polyamine and SAM-Related Metabolites

Initial works using HPLC-FLD detection of derivatized polyamines were relatively simple and reproducible. A significant background between the derivatization blank and *N*-methylspermidine standards was observed when transitioning the established method to another HPLC-FLD system. This background is likely due to a lower detection range between different fluorescent detector manufacturers. In our quest to improve this HPLC method, we decided to utilize an Acquity UPLC system coupled with an Acquity fluorescent detector in hopes that we could adapt the established method and the known benefits of UPLC separation over traditional HPLC techniques.

UPLC techniques benefit from tighter packed columns and higher pressures, allowing for shorter run times and less solvent waste than HPLC techniques (Chawla, 2022). The UPLC method separated spermidine and *N*-methylspermidine standards via fluorescent detection. The fluorescent detector had a much higher dynamic range than the HPLC system, which was limited to a thousand fluorescence units. The separation between these two polyamine standards was better than the HPLC method, with a much lower background from the blank. However, we observed that in subsequent injections, there was a carryover, indicating that the polyamine standards were sticking to the column. This carryover problem complicated our goal of using the method for enzymatic assays. We aimed to have a simple, reproducible separation method that requires no derivatization. While fluorescence detection of polyamines can be sensitive, it does not provide information about the chromatographic peak other than retention time and fluorescent signal intensity. Additionally, derivatization procedures come at the cost of being time-sensitive since the derivatization reactions are only stable for less than a day, severely limiting the throughput of large amounts of samples. Due to this instability of derivatization for extended run times, it was logical to seek other methods to detect these polyamine compounds.

A liquid chromatography method to separate polyamines without following laborious derivatization procedures to visualize the polyamine peaks would be preferred. Liquid Chromatography with tandem mass spectrometry (LC-MS/MS) is a powerful analytical technique that combines the separating power of liquid chromatography with the highly sensitive and selective mass analysis capability of triple quadrupole mass spectrometry (Thomas et al., 2022). LC-MS/MS methods offer an alternative way to detect polyamine and related cofactors during the enzyme assay without using derivatization-based methods to detect small molecules effectively. Traditional LC-MS/MS methods use a reverse phase C18 column to retain nonpolar metabolites and elute throughout the LC-MS/MS run. Although other techniques are available that utilize LC-MS/MS detection, such as ion-pairing in which heptafluorobutyric acid (HFBA) acts as a stabilizer for difficult-to-retain compounds like polyamines, these techniques can be cumbersome and require lots of optimization (Häkkinen, 2011; Sánchez-López et al., 2009). Some mass spectrometry manufacturers warn that ion pairing is recommended when all other approaches are unsuccessful. They also suggest that separate columns and, even if possible, separate mass spectrometry instruments should be used when performing ion-pairing chromatography. Ion pairing reagents are reported to tend to stick to LC components, the column, and the MS source. This, in turn, can lead to signal suppression in future analyses and affect column performance. So, with this information, we decided to heed this warning and try alternative techniques that would allow for the separation of polyamines without using ion pairing and derivatization chromatographic techniques.

Hydrophilic Interaction Liquid Chromatography (HILIC) is a separation technique to retain and separate polar metabolites. Since amino acids, polyamine, and methionine cycle intermediates are generally polar, it might be reasonable to assume that HILIC might offer an advantage in our search for an optimal separation of polyamines. We used an Acquity BEH Amide HILIC column from Waters for our separation method, as it was reported to work for polyamines like putrescine, spermidine, and spermine (Su et al., 2021). Additionally, this method was to be able to separate *S*-adenosyl-L-methionine (SAM) and *S*-adenosyl-L-homocysteine (SAH), which could be advantageous in future enzyme assays to ascertain enzyme kinetics of EcSMT, EcSPMT, and EcSPDS isolated from *E. coca*. When comparing our LC-MS/MS results to our HPLC and UPLC fluorescent detection, the HILIC column provided great separation of polyamines without the need for derivatization, coupled with the sensitivity of a triple quadrupole mass spectrometry allowed for increased confidence of metabolite identification.

Enzyme assays for polyamine-related enzymes like putrescine *N*-methyltransferases (PMT) focus mainly on detecting polyamines that must be derivatized to visualize by photodiode array or fluorescent detector. Generally, these measurements exclude the detection of SAM and dcSAM, critical cofactors in these enzyme assays. The initial characterization of the thermospermine synthase (ACL5) gene from *Arabidopsis thaliana* was mischaracterized as a spermine synthase (SPS) (Hanzawa et al., 2000; Knott et al., 2007). Spermine and thermospermine are constitutional isomers of each other and would elute around similar retention times when derivatized due to structural similarities. This case is a prime example of having the resolving power of LC-MS/MS to identify structural differences between spermine and thermospermine and prevent misidentifications. Following this idea, having a separation method that allows the detection of substrates, cofactors, and products involved in polyamine biosynthesis via LC-MS/MS helps to increase the ease of use for performing enzyme kinetics while increasing the identification of metabolites with high confidence.

While the Acquity BEH Amide column delivered an excellent performance on separating polyamine standards, we must address some limitations of HILIC chromatography. One significant limitation is that HILIC columns are sensitive to aqueous solvents. The end user must be aware of these limitations and optimize accordingly. A prime example is when we compared the separation of polyamine standards to actual enzymatic assay conditions. The enzymatic assay conditions are performed under aqueous conditions along with the presence of polar compounds, such as phosphate buffers. During initial testing, the peak shapes for the enzyme assays were not as sharp, and the retention times were shifted compared to the standards. The broad shoulders observed on *N*-methylspermidine and spermidine can be alleviated by further diluting the enzyme assay samples in a less aqueous solvent, such as 85% acetonitrile and 15% water. Since the BEH Amide column separation method does not require any tedious derivatization reactions, we must consider the detection and measurement of *S*-adenosyl-L-methionine (SAM). SAM cofactor is sensitive to breaking down when heat is applied. Future considerations for enzyme assays should use another way to inactivate the enzyme besides protein denaturation. Other options may include using acids such as formic acid or hydrochloric acid to chemically inactivate the enzyme assay rather than the traditional heat inactivation to prevent the SAM cofactor from breaking apart (Morana et al., 2002). Additionally, more replicates are required for the polyamine standards for statistical confidence in the concentration curves for spermidine and *N*-methylspermidine. The BEH amide HILIC separation, along with LC-MS/MS, offers the capabilities to reliably separate and quantify

polyamines and other related compounds, providing increased confidence in both detection and confidence as further characterizations of enzymatic steps of TA biosynthesis are interrogated further using *in vitro* biochemical characterization techniques.

5 CONCLUSION

Tropane alkaloid biosynthesis in *Erythroxylum coca* consists of eleven enzymatic steps, beginning with ornithine and finishing with the formation of cocaine. This work uncovers seven previously unknown enzymatic steps and ten newly discovered enzymes participating in tropane alkaloid biosynthesis in *E. coca* (Chavez et al., 2022). Phylogenetic analysis of each enzymatic step revealed that the TA biosynthetic pathway is polyphyletic, and the formation of the tropane ring between Solanaceae and Erythroxylaceae is independent in origin. Convergent evolution of plant-specialized metabolism is established and is commonly observed in plants (Pichersky & Lewinsohn, 2011). A case in point would be the convergent evolution of purine alkaloids in plants. The methylation steps of the purine alkaloids, xanthine, produce caffeine and theobromine. This methylation is catalyzed by several distantly related SABATH methyltransferases whose recruitment occurred multiple times in different plant lineages, explaining the scattered distribution of these compounds among the angiosperms (R. Huang et al., 2016). Another well-established example of plant alkaloid convergent evolution is seen within the biosynthesis of pyrrolizidine alkaloids, for which the enzyme homospermidine synthase was independently recruited at least four times during the evolution of land plants (Ober & Kaltenegger, 2009).

The convergent evolution of plant-specialized metabolites is commonly observed and continues to be supported by increasing amounts of metabolomic data across many diverse plant species (Ono & Murata, 2023). The scattered distribution of TAs across the angiosperms suggests these pathways must have emerged after the split of the core eudicots around 120 million years ago (Chavez et al., 2022; Kim et al., 2016; Y.-J. Wang et al., 2023). Although the capacity for TA biosynthesis appears to be preserved across eleven plant families, the phylogenetic relationships between the tropane alkaloid pathways outside the Solanaceae or Erythroxylaceae families remain to be explored further.

6 OUTLOOK

Tropane alkaloid research over the past couple of years has accelerated due to technological innovations in genomic sequencing. Synthetic and systems biology have leveraged genomic information to resolve complex biosynthetic pathways. These approaches require detailed information about the genome and transcriptome. The future direction of tropane alkaloid research will likely be enhanced by four technologies: spatial transcriptomics, gene expression regulation via transcription factor binding sites, enzyme localization, and identifying essential metabolite transporters.

Spatial transcriptomics technologies can potentially reveal gene expression in specific cells in plant tissues, allowing a more comprehensive view of the TA gene expression at cellular resolution. Spatial transcriptomic technology has already been applied to germinating barley grain tissues to help understand grain formation at the transcriptomic level (Peirats-Llobet et al., 2023). Understanding where the mRNAs are localized across different tissue types, even at the subcellular level, will provide a clear expression map of genes involved in TA biosynthesis. TAs derived from Solanaceous plants are known to be highly expressed in the roots compared to the other tissues in the plants, such as *E. coca,* where the gene expression is observed within leaf tissues. Once spatial transcriptomics becomes widely available for plants, tropane alkaloid biosynthesis will be further dissected at the cellular or subcellular level.

In the presence of high-quality plant genome assemblies, future steps would involve dissecting the transcription factors and promoters driving the expression of TA biosynthetic genes. Techniques like DNA Affinity Purification Sequencing (DAP-Seq) and MNase-defined cistrome-Occupancy Analysis (MOA-seq) would allow for rapid identification of transcription factor binding sites throughout the genome, specifically focusing on promoter regions (Bartlett et al., 2017; Savadel et al., 2021). Identifying the interactions between the promoters and transcription factors would serve as a cistrome atlas to understanding the specific expression control of tropane alkaloid genes throughout the Solanaceae and Erythroxylaceae families.

Enzymatic localizations will be the next focus of these newly discovered tropane alkaloid enzymes, as some have been validated in yeast, but further work should focus on identifying the subcellular localization of these enzymes *in planta* especially since EcSPMT was shown not to possess bifunctional activity in yeast. Biosensors and fluorescent proteins fused to tropane alkaloid genes should be used to determine the subcellular localization of all enzymes involved in TA biosynthesis in *E. coca.* After identifying the subcellular

compartmentalization of TA enzymes, the following steps should focus on identifying metabolite transporters responsible for trafficking TA precursors throughout different subcellular compartments. Previous work on identifying alkaloid transporters for tropane alkaloids in yeast improves the production of TAs as they exhibit a cytotoxic effect on the cells when remaining in the cytosol (Srinivasan & Smolke, 2021). Identifying essential metabolite transporters will clarify the subcellular compartmentalization of different parts of the TA biosynthesis, allowing for better development in silico modeling of TA biosynthesis. All these technologies complementing each other will help unlock TA metabolism observed in other tropane alkaloid-producing plant families.

7 BIBLIOGRAPHY

Agrimi, G., Russo, A., Pierri, C. L., & Palmieri, F. (2012). The peroxisomal NAD+ carrier of *Arabidopsis thaliana* transports coenzyme A and its derivatives. *Journal of Bioenergetics and Biomembranes*, *44*(3), 333–340. https://doi.org/10.1007/s10863-012-9445-0

Alcázar, R., Bueno, M., & Tiburcio, A. F. (2020). Polyamines: Small Amines with Large Effects on Plant Abiotic Stress Tolerance. *Cells*, *9*(11), 2373. https://doi.org/10.3390/cells9112373

Anderson, K. N. (2019). *Elucidation of Novel N-methyltransferases in Tropane and Granatane Alkaloid Biosynthesis* [Thesis]. https://hdl.handle.net/2346/85378

Awal, H. M. A., & Hirasawa, E. (1995). 1,3-Diaminopropane is a suicide substrate for pea diamine oxidase. *Phytochemistry*, *39*(3), 489–490. https://doi.org/10.1016/0031-9422(95)00003-P

Balasundaram, D., Dinman, J. D., Wickner, R. B., Tabor, C. W., & Tabor, H. (1994). Spermidine deficiency increases +1 ribosomal frameshifting efficiency and inhibits Ty1 retrotransposition in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(1), 172–176.

Balasundaram, D., Xie, Q. W., Tabor, C. W., & Tabor, H. (1994). The presence of an active S-adenosylmethionine decarboxylase gene increases the growth defect observed in Saccharomyces cerevisiae mutants unable to synthesize putrescine, spermidine, and spermine. *Journal of Bacteriology*, *176*(20), 6407–6409. https://doi.org/10.1128/jb.176.20.6407-6409.1994

Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., & Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nature Protocols*, *12*(8), 1659–1672. https://doi.org/10.1038/nprot.2017.055

Bedewitz, M. A., Jones, A. D., D'Auria, J. C., & Barry, C. S. (2018). Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. *Nature Communications*, *9*(1), 5281. https://doi.org/10.1038/s41467-018-07671-3

Bekkouche, K., Daali, Y., Cherkaoui, S., Veuthey, J. L., & Christen, P. (2001). Calystegine distribution in some solanaceous species. *Phytochemistry*, *58*(3), 455–462.

Belda-Palazón, B., Almendáriz, C., Martí, E., Carbonell, J., & Ferrando, A. (2016). Relevance of the Axis Spermidine/eIF5A for Plant Growth and Development. *Frontiers in Plant Science*, *7*. https://www.frontiersin.org/articles/10.3389/fpls.2016.00245

Belda-Palazón, B., Ruiz, L., Martí, E., Tárraga, S., Tiburcio, A. F., Culiáñez, F., Farràs, R., Carrasco, P., & Ferrando, A. (2012). Aminopropyltransferases Involved in Polyamine Biosynthesis Localize Preferentially in the Nucleus of Plant Cells. *PLOS ONE*, *7*(10), e46907. https://doi.org/10.1371/journal.pone.0046907

Benkő, P., Gémes, K., & Fehér, A. (2022). Polyamine Oxidase-Generated Reactive Oxygen Species in Plant Development and Adaptation: The Polyamine Oxidase—NADPH Oxidase Nexus. *Antioxidants*, *11*(12), 2488. https://doi.org/10.3390/antiox11122488

Bennett, E. M., Ekstrom, J. L., Pegg, A. E., & Ealick, S. E. (2002). Monomeric S-adenosylmethionine decarboxylase from plants provides an alternative to putrescine stimulation. *Biochemistry*, *41*(49), 14509–14517.

Benton, M. J., Wilf, P., & Sauquet, H. (2022). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. *New Phytologist*, *233*(5), 2017–2035. https://doi.org/10.1111/nph.17822

Bianchini, G., & Sánchez-Baracaldo, P. (2024). TreeViewer: Flexible, modular software to visualise and manipulate phylogenetic trees. *Ecology and Evolution*, *14*(2), e10873. https://doi.org/10.1002/ece3.10873

Biastoff, S., Brandt, W., & Dräger, B. (2009). Putrescine *N*-methyltransferase—The start for alkaloids. *Phytochemistry*, *70*(15–16), 1708–1718. https://doi.org/10.1016/j.phytochem.2009.06.012

Biastoff, S., Reinhardt, N., Reva, V., Brandt, W., & Drager, B. (2009). Evolution of putrescine N-methyltransferase from spermidine synthase demanded alterations in substrate binding. *FEBS Letters*, *583*(20), 3367–3374. https://doi.org/10.1016/j.febslet.2009.09.043

Bieri, S., Brachet, A., Veuthey, J. L., & Christen, P. (2006). Cocaine distribution in wild *Erythroxylum* species. *Journal of Ethnopharmacology*, *103*(3), 439–447.

Biondich, A. S., & Joslin, J. D. (2016). Coca: The History and Medical Significance of an Ancient Andean Tradition. *Emergency Medicine International*, *2016*, 4048764. https://doi.org/10.1155/2016/4048764

Blume, L., Long, T. E., & Turos, E. (2023). Applications and Opportunities in Using Disulfides, Thiosulfinates, and Thiosulfonates as Antibacterials. *International Journal of Molecular Sciences*, *24*(10), Article 10. https://doi.org/10.3390/ijms24108659

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology*, *54*(1), 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Bontpart, T., Cheynier, V., Ageorges, A., & Terrier, N. (2015). BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. *New Phytologist*, *208*(3), 695–707.

Carbonell, J., & Blázquez, M. A. (2009). Regulatory mechanisms of polyamine biosynthesis in plants. *Genes & Genomics*, *31*(2), 107–118. https://doi.org/10.1007/BF03191144

Carpenter, E. J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J., Jimenez Vieira, F. R., Bowler, C., Dorrell, R. G., Gitzendanner, M. A., Li, L., Du, W., K. Ullrich, K.,

Wickett, N. J., Barkmann, T. J., Barker, M. S., Leebens-Mack, J. H., & Wong, G. K.-S. (2019). Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience*, *8*(10), giz126. https://doi.org/10.1093/gigascience/giz126

Cervelli, M., Cona, A., Angelini, R., Polticelli, F., Federico, R., & Mariottini, P. (2001). A barley polyamine oxidase isoform with distinct structural features and subcellular localization. *European Journal of Biochemistry*, *268*(13), 3816–3830. https://doi.org/10.1046/j.1432-1327.2001.02296.x

Chakraborty, P., Biswas, A., Dey, S., Bhattacharjee, T., & Chakrabarty, S. (2023). Cytochrome P450 Gene Families: Role in Plant Secondary Metabolites Production and Plant Defense. *Journal of Xenobiotics*, *13*(3), Article 3. https://doi.org/10.3390/jox13030026

Chattopadhyay, M. K., Fernandez, C., Sharma, D., McPhie, P., & Masison, D. C. (2011). Yeast ornithine decarboxylase and antizyme form a 1:1 complex in vitro: Purification and characterization of the inhibitory complex. *Biochemical and Biophysical Research Communications*, *406*(2), 177–182. https://doi.org/10.1016/j.bbrc.2011.01.113

Chavez, B. G., Srinivasan, P., Glockzin, K., Kim, N., Montero Estrada, O., Jirschitzka, J., Rowden, G., Shao, J., Meinhardt, L., Smolke, C. D., & D'Auria, J. C. (2022). Elucidation of tropane alkaloid biosynthesis in *Erythroxylum coca* using a microbial pathway discovery platform. *Proceedings of the National Academy of Sciences*, *119*(49), e2215372119. https://doi.org/10.1073/pnas.2215372119

Chawla, P. A. (2022). Prospects of UPLC in Pharmaceutical Analysis over HPLC. *Biomedical Journal of Scientific & Technical Research*, *45*(1). https://doi.org/10.26717/BJSTR.2022.45.007138

Chuang, L., & Franke, J. (2022). Rapid Combinatorial Coexpression of Biosynthetic Genes by Transient Expression in the Plant Host *Nicotiana benthamiana*. In E. Skellam (Ed.), *Engineering Natural Product Biosynthesis* (Vol. 2489, pp. 395–420). Springer US. https://doi.org/10.1007/978-1-0716-2273-5_20

Ciarkowska, A., Ostrowski, M., Starzyńska, E., & Jakubowska, A. (2019). Plant SCPL acyltransferases: Multiplicity of enzymes with various functions in secondary metabolism. *Phytochemistry Reviews*, *18*(1), 303–316. https://doi.org/10.1007/s11101-018-9593-x

Cosse, M., & Seidel, T. (2021). Plant Proton Pumps and Cytosolic pH-Homeostasis. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.672873

Couladis, M. M., Friesen, J. B., Landgrebe, M. E., & Leete, E. (1991). Enzymes catalysing the reduction of tropinone to tropine and ψ-tropine isolated from the roots of *Datura innoxia*. *Phytochemistry*, *30*(3), 801–805. https://doi.org/10.1016/0031-9422(91)85255-X

D'Auria, J. C. (2006). Acyltransferases in plants: A good time to be BAHD. *Current Opinion in Plant Biology*, *9*(3), 331–340. https://doi.org/10.1016/j.pbi.2006.03.016

D'Auria, J. C., Chen, F., & Pichersky, E. (2003). The SABATH family of MTS in Arabidopsis Thaliana and other plant species. In *Recent Advances in Phytochemistry* (Vol. 37, pp. 253–283). Elsevier. https://doi.org/10.1016/S0079-9920(03)80026-6

Dey, P., Kundu, A., Kumar, A., Gupta, M., Lee, B. M., Bhakta, T., Dash, S., & Kim, H. S. (2020). Analysis of alkaloids (indole alkaloids, isoquinoline alkaloids, tropane alkaloids). *Recent Advances in Natural Products Analysis*, 505–567. https://doi.org/10.1016/B978-0-12-816455-6.00015-9

Dillehay, T. D., Rossen, J., Ugent, D., Karathanasis, A., Vásquez, V., & Netherly, P. J. (2010). Early Holocene coca chewing in northern Peru. *Antiquity*, *84*(326), 939–953.

Divekar, P. A., Narayana, S., Divekar, B. A., Kumar, R., Gadratagi, B. G., Ray, A., Singh, A. K., Rani, V., Singh, V., Singh, A. K., Kumar, A., Singh, R. P., Meena, R. S., & Behera, T. K. (2022). Plant Secondary Metabolites as Defense Tools against Herbivores for Sustainable Crop Protection. *International Journal of Molecular Sciences*, *23*(5), 2690. https://doi.org/10.3390/ijms23052690

Docimo, T., Davis, A. J., Luck, K., Fellenberg, C., Reichelt, M., Phillips, M., Gershenzon, J., & D'Auria, J. C. (2015). Influence of medium and elicitors on the production of cocaine, amino acids and phytohormones by *Erythroxylum coca* calli. *Plant Cell Tissue and Organ Culture*, *120*(3), 1061–1075. https://doi.org/10.1007/s11240-014-0660-8

Docimo, T., Reichelt, M., Schneider, B., Kai, M., Kunert, G., Gershenzon, J., & D'Auria, J. C. (2012). The first step in the biosynthesis of cocaine in *Erythroxylum coca*: The characterization of arginine and ornithine decarboxylases. *Plant Molecular Biology*, *78*(6), 599–615. https://doi.org/10.1007/s11103-012-9886-1

Docimo, T., Schmidt, G. W., Luck, K., Delaney, S. K., & D'Auria, J. C. (2013). Selection and validation of reference genes for quantitative gene expression studies in *Erythroxylum coca*. *F1000Research*, *2*, 37. https://doi.org/10.12688/f1000research.2-37.v1

Dong, N.-Q., & Lin, H.-X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *Journal of Integrative Plant Biology*, *63*(1), 180–209. https://doi.org/10.1111/jipb.13054

Drager, B., Funck, C., Hrhler, A., Mrachatz, G., Nahrstedt, A., Portsteffen, A., Schaal, A., & Schmidt, R. (1995). *Calystegines as a new group of tropane alkaloids in Solanaceae*.

Eisenberg, A. (2013). *Aymara Indian perspectives on development in the Andes*. University of Alabama Press.

Erb, M., & Kliebenstein, D. J. (2020). Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. *Plant Physiology*, *184*(1), 39–52. https://doi.org/10.1104/pp.20.00433

Estrada, O. P. (2017). *Biochemical studies in the elucidation of genes involved in tropane alkaloid production in Erythroxylum coca and Erythroxylum novogranatense*. 119.

Facchini, P. J. (2001). Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annual Review of Plant Physiology and Plant Molecular Biology*, *52*, 29–66. https://doi.org/10.1146/annurev.arplant.52.1.29

Fincato, P., Moschou, P. N., Ahou, A., Angelini, R., Roubelakis-Angelakis, K. A., Federico, R., & Tavladoraki, P. (2012). The members of *Arabidopsis thaliana* PAO gene family exhibit distinct tissue- and organ-specific expression pattern during seedling growth and flower development. *Amino Acids*, *42*(2), 831–841. https://doi.org/10.1007/s00726-011-0999-7

Fontecave, M., Atta, M., & Mulliez, E. (2004). S-adenosylmethionine: Nothing goes to waste. *Trends in Biochemical Sciences*, *29*(5), 243–249. https://doi.org/10.1016/j.tibs.2004.03.007

Gallego-Bartolomé, J. (2020). DNA methylation in plants: Mechanisms and tools for targeted manipulation. *New Phytologist*, *227*(1), 38–44. https://doi.org/10.1111/nph.16529

García-Contreras, R., Vos, P., Westerhoff, H. V., & Boogerd, F. C. (2012). Why in vivo may not equal in vitro – new effectors revealed by measurement of enzymatic activities under the same in vivo-like assay conditions. *The FEBS Journal*, *279*(22), 4145–4159. https://doi.org/10.1111/febs.12007

Gay, G. R., Inaba, D. S., Sheppard, C. W., Newmeyer, J. A., & Rappolt, R. T. (1975). Cocaine: History, Epidemiology, Human Pharmacology, and Treatment. A Perspective on a New Debut for an Old Girl. *Clinical Toxicology*, *8*(2), 149–178. https://doi.org/10.3109/15563657508988061

Ge, C., Cui, X., Wang, Y., Hu, Y., Fu, Z., Zhang, D., Cheng, Z., & Li, J. (2006). BUD2, encoding an S-adenosylmethionine decarboxylase, is required for Arabidopsis growth and development. *Cell Research*, *16*(5), Article 5. https://doi.org/10.1038/sj.cr.7310056

Geiger, & Hesse. (1833a). Darstellung des Atropins. *Annalen Der Pharmacie*, *5*(1), 43–81. https://doi.org/10.1002/jlac.18330050108

Geiger & Hesse. (1833b). Fortgesetzte Versuche über Atropin. *Annalen Der Pharmacie*, *6*(1), 44–65. https://doi.org/10.1002/jlac.18330060107

Gibbs, N. M., Su, S.-H., Lopez-Nieves, S., Mann, S., Alban, C., Maeda, H. A., & Masson, P. H. (2021). Cadaverine regulates biotin synthesis to modulate primary root growth in Arabidopsis. *The Plant Journal*, *107*(5), 1283–1298. https://doi.org/10.1111/tpj.15417

Gong, B., Li, X., VandenLangenberg, K. M., Wen, D., Sun, S., Wei, M., Li, Y., Yang, F., Shi, Q., & Wang, X. (2014). Overexpression of S-adenosyl-l-methionine synthetase increased tomato tolerance to alkali stress through polyamine metabolism. *Plant Biotechnology Journal*, *12*(6), 694–708. https://doi.org/10.1111/pbi.12173

González-Hernández, A. I., Scalschi, L., Vicedo, B., Marcos-Barbero, E. L., Morcuende, R., & Camañes, G. (2022). Putrescine: A Key Metabolite Involved in Plant Development, Tolerance and Resistance Responses to Stress. *International Journal of Molecular Sciences*, *23*(6), 2971. https://doi.org/10.3390/ijms23062971

Gootenberg, P. (2002). *Cocaine: Global Histories*. Routledge.

Green, M., & Sambrook, J. (2020). The Inoue Method for Preparation and Transformation of Competent Escherichia coli: "Ultracompetent" Cells. *Cold Spring Harbor Protocols*, *2020*, pdb.prot101196. https://doi.org/10.1101/pdb.prot101196

Häkkinen, M. R. (2011). Polyamine Analysis by LC-MS. In A. E. Pegg & Jr. Casero Robert A. (Eds.), *Polyamines: Methods and Protocols* (pp. 505–518). Humana Press. https://doi.org/10.1007/978-1-61779-034-8_33

Hanczkó, R., Jámbor, A., Perl, A., & Molnár-Perl, I. (2007). Advances in the o-phthalaldehyde derivatizations. Comeback to the o-phthalaldehyde-ethanethiol reagent. *Journal of Chromatography. A*, *1163*(1–2), 25–42. https://doi.org/10.1016/j.chroma.2007.06.013

Hanfrey, C., Sommer, S., Mayer, M. J., Burtin, D., & Michael, A. J. (2001). Arabidopsis polyamine biosynthesis: Absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. *Plant Journal*, *27*(6), 551–560. https://doi.org/10.1046/j.1365-313x.2001.01100.x

Hanzawa, Y., Takahashi, T., Michael, A. J., Burtin, D., Long, D., Pineiro, M., Coupland, G., & Komeda, Y. (2000). ACAULIS5, an Arabidopsis gene required for stem elongation, encodes a spermine synthase. *The EMBO Journal*, *19*(16), 4248–4256. https://doi.org/10.1093/emboj/19.16.4248

Hashimoto, T., Nakajima, K., Ongena, G., & Yamada, Y. (1992). Two Tropinone Reductases with Distinct Stereospecificities from Cultured Roots of Hyoscyamus niger. *Plant Physiology*, *100*(2), 836–845. https://doi.org/10.1104/pp.100.2.836

Hashimoto, T., Yukimune, Y., & Yamada, Y. (1989). Putrescine and putrescine *N*-methyltransferase in the biosynthesis of tropane alkaloids in cultured roots of *Hyoscyamus albus* .2. Incorporation of labeled precursors. *Planta*, *178*(1), 131–137. https://doi.org/10.1007/bf00392536

Heim, W. G., Sykes, K. A., Hildreth, S. B., Sun, J., Lu, R. H., & Jelesko, J. G. (2007). Cloning and characterization of a *Nicotiana tabacum* methylputrescine oxidase transcript. *Phytochemistry*, *68*(4), 454–463. https://doi.org/10.1016/j.phytochem.2006.11.003

Hibi, N., Higashiguchi, S., Hashimoto, T., & Yamada, Y. (1994). Gene expression in tobacco low-nicotine mutants. *Plant Cell*, *6*(5), 723–735. https://doi.org/10.2307/3869875

Hill, C. R., Shafaei, A., Balmer, L., Lewis, J. R., Hodgson, J. M., Millar, A. H., & Blekkenhorst, L. C. (2023). Sulfur compounds: From plants to humans and their role in chronic disease prevention. *Critical Reviews in Food Science and Nutrition*, *63*(27), 8616–8638. https://doi.org/10.1080/10408398.2022.2057915

Höfgen, R., & Willmitzer, L. (1988). Storage of competent cells for Agrobacterium transformation. *Nucleic Acids Research*, *16*(20), 9877.

Hu, W.-W., Gong, H., & Pua, E. C. (2005). The Pivotal Roles of the Plant S-Adenosylmethionine Decarboxylase 5′ Untranslated Leader Sequence in Regulation of Gene Expression at the Transcriptional and Posttranscriptional Levels. *Plant Physiology*, *138*(1), 276–286. https://doi.org/10.1104/pp.104.056770

Hu, X., Liu, W., Yan, Y., Deng, H., & Cai, Y. (2023). Tropinone reductase: A comprehensive review on its role as the key enzyme in tropane alkaloids biosynthesis. *International Journal of Biological Macromolecules*, *253*, 127377. https://doi.org/10.1016/j.ijbiomac.2023.127377

Huang, J.-P., Fang, C., Ma, X., Wang, L., Yang, J., Luo, J., Yan, Y., Zhang, Y., & Huang, S.-X. (2019). Tropane alkaloids biosynthesis involves an unusual type III polyketide synthase and non-enzymatic condensation. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-11987-z

Huang, J.-P., Wang, Y.-J., Tian, T., Wang, L., Yan, Y., & Huang, S.-X. (2021). Tropane alkaloid biosynthesis: A centennial review. *Natural Product Reports*, *38*(9), 1634–1658. https://doi.org/10.1039/D0NP00076K

Huang, R., O'Donnell, A. J., Barboline, J. J., & Barkman, T. J. (2016). Convergent evolution of caffeine in plants by co-option of exapted ancestral enzymes. *Proceedings of the National Academy of Sciences*, *113*(38), 10613–10618. https://doi.org/10.1073/pnas.1602575113

Hura, T., Dziurka, M., Hura, K., Ostrowska, A., & Dziurka, K. (2015). Free and Cell Wall-Bound Polyamines under Long-Term Water Stress Applied at Different Growth Stages of ×Triticosecale Wittm. *PLOS ONE*, *10*(8), e0135002. https://doi.org/10.1371/journal.pone.0135002

Ikeguchi, Y., Bewley, M. C., & Pegg, A. E. (2006). Aminopropyltransferases: Function, Structure and Genetics. *The Journal of Biochemistry*, *139*(1), 1–9. https://doi.org/10.1093/jb/mvj019

Illingworth, C., & Michael, A. J. (2012). Plant ornithine decarboxylase is not post-transcriptionally feedback regulated by polyamines but can interact with a cytosolic ribosomal protein S15 polypeptide. *Amino Acids*, *42*(2–3), 519–527. https://doi.org/10.1007/s00726-011-1029-5

Imai, A., Akiyama, T., Kato, T., Sato, S., Tabata, S., Yamamoto, K. T., & Takahashi, T. (2004). Spermine is not essential for survival of Arabidopsis. *FEBS Letters*, *556*(1), 148–152. https://doi.org/10.1016/S0014-5793(03)01395-4

Imai, A., Matsuyama, T., Hanzawa, Y., Akiyama, T., Tamaoki, M., Saji, H., Shirano, Y., Kato, T., Hayashi, H., Shibata, D., Tabata, S., Komeda, Y., & Takahashi, T. (2004). Spermidine synthase genes are essential for survival of Arabidopsis. *Plant Physiology*, *135*(3), 1565–1573. https://doi.org/10.1104/pp.104.041699

Islam, M. (2011). *TRACING THE EVOLUTIONARY HISTORY OF COCA (*ERYTHROXYLUM*)*. University of Colorado.

Islam, M. S., Saito, J. A., Emdad, E. M., Ahmed, B., Islam, M. M., Halim, A., Hossen, Q. M. M., Hossain, M. Z., Ahmed, R., Hossain, M. S., Kabir, S. M. T., Khan, M. S. A., Khan, M. M., Hasan, R., Aktar, N., Honi, U., Islam, R., Rashid, M. M., Wan, X., … Alam, M. (2017). Comparative genomics of two jute species and insight into fibre biogenesis. *Nature Plants*, *3*, 16223. https://doi.org/10.1038/nplants.2016.223

Jancewicz, A. L., Gibbs, N. M., & Masson, P. H. (2016). Cadaverine's Functional Role in Plant Development and Environmental Response. *Frontiers in Plant Science*, *7*, 870. https://doi.org/10.3389/fpls.2016.00870

Janowitz, T., Kneifel, H., & Piotrowski, M. (2003). Identification and characterization of plant agmatine iminohydrolase, the last missing link in polyamine biosynthesis of plants. *FEBS Letters*, *544*(1–3), 258–261. https://doi.org/10.1016/S0014-5793(03)00515-5

Jirschitzka, J., Schmidt, G. W., Reichelt, M., Schneider, B., Gershenzon, J., & D'Auria, J. C. (2012). Plant tropane alkaloid biosynthesis evolved independently in the Solanaceae and Erythroxylaceae. *Proceedings of the National Academy of Sciences*, *109*(26), 10304–10309. https://doi.org/10.1073/pnas.1200473109

Johnson, E. L. (1995). Content and Distribution of Erythroxylum-Coca Leaf Alkaloids. *Annals of Botany*, *76*(4), 331–335.

Joshi, K., Ahmed, S., Ge, L., Phuntumart, V., Kalinoski, A., & Morris, P. F. (2022). *Compartmentation of Putrescine Synthesis in Plants* (p. 2022.09.03.506421). bioRxiv. https://doi.org/10.1101/2022.09.03.506421

Junker, A., Fischer, J., Sichhart, Y., Brandt, W., & Draeger, B. (2013). Evolution of the key alkaloid enzyme putrescine N-methyltransferase from spermidine synthase. *Frontiers in Plant Science*, *4*. https://www.frontiersin.org/articles/10.3389/fpls.2013.00260

Kai, G., Zhang, Y., Chen, J., Li, L., Yan, X., Zhang, R., Liao, P., Lu, X., Wang, W., & Zhou, G. (2009). Molecular characterization and expression analysis of two distinct putrescine *N*-methyltransferases from roots of *Anisodus acutangulus*. *Physiologia Plantarum*, *135*(2), 121–129. https://doi.org/10.1111/j.1399-3054.2008.01178.x

Kakehi, J., Kuwashiro, Y., Niitsu, M., & Takahashi, T. (2008). Thermospermine is required for stem elongation in *Arabidopsis thaliana*. *Plant & Cell Physiology*, *49*(9), 1342–1349. https://doi.org/10.1093/pcp/pcn109

Katakam, N. K., Seifert, C., D'Auria, J., & Li, G. (2019). 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*. https://doi.org/10.3987/COM-18-S(F)4

Kavanagh, K. L., Jörnvall, H., Persson, B., & Oppermann, U. (2008). Medium-and short-chain dehydrogenase/reductase gene and protein families. *Cellular and Molecular Life Sciences*, *65*(24), 3895.

Kawai, Y., Ono, E., & Mizutani, M. (2014). Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, *78*(2), 328–343. https://doi.org/10.1111/tpj.12479

Kim, N. (2020). *Tropane alkaloid biosynthesis in Erythroxylum coca involves an atypical type III polyketide synthase* [Thesis]. https://ttu-ir.tdl.org/handle/2346/85851

Kim, N., Chavez, B., Stewart, C., & D'Auria, J. C. (2021). Structure and Function of Enzymes Involved in the Biosynthesis of Tropane Alkaloids. In V. Srivastava, S. Mehrotra, & S. Mishra (Eds.), *Tropane Alkaloids: Pathways, Potential and Biotechnological Applications* (pp. 21–50). Springer. https://doi.org/10.1007/978-981-33-4535-5_2

Kim, N., Estrada, O., Chavez, B., Stewart, C., & D'Auria, J. C. (2016). Tropane and Granatane Alkaloid Biosynthesis: A Systematic Analysis. *Molecules*, *21*(11). https://doi.org/10.3390/molecules21111510

Knott, J. M., Römer, P., & Sumper, M. (2007). Putative spermine synthases from *Thalassiosira pseudonana* and *Arabidopsis thaliana* synthesize thermospermine rather than spermine. *FEBS Letters*, *581*(16), 3081–3086. https://doi.org/10.1016/j.febslet.2007.05.074

Kumar, A., Taylor, M. A., Arif, S. A. M., & Davies, H. V. (1996). Potato plants expressing antisense and sense S-adenosylmethionine decarboxylase (SAMDC) transgenes show altered levels of polyamines and ethylene: Antisense plants display abnormal phenotypes. *The Plant Journal*, *9*(2), 147–158. https://doi.org/10.1046/j.1365-313X.1996.09020147.x

Kumar, S., & Mohapatra, T. (2021). Dynamics of DNA Methylation and Its Functions in Plant Growth and Development. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.596236

Kumar, V., Dooley, D. M., Freeman, H. C., Guss, J. M., Harvey, I., McGuirl, M. A., Wilce, M. C., & Zubak, V. M. (1996). Crystal structure of a eukaryotic (pea seedling) copper-containing amine oxidase at 2.2 A resolution. *Structure (London, England*, *4*(8), 943–955. https://doi.org/10.1016/s0969-2126(96)00101-3

Kumar, V., Mishra, R. K., Ghose, D., Kalita, A., Dhiman, P., Prakash, A., Thakur, N., Mitra, G., Chaudhari, V. D., Arora, A., & Dutta, D. (2022). Free spermidine evokes superoxide radicals that manifest toxicity. *eLife*, *11*, e77704. https://doi.org/10.7554/eLife.77704

Leete, E. (1982). Biosynthesis of the pyrrolidine rings of cocaine and cuscohygrine from [5-14C]-ornithine via a symmetrical intermediate. *Journal of the American Chemical Society*, *104*(5), 1403–1408. https://doi.org/10.1021/ja00369a042

Leete, E. (1985). Spermidine: An indirect precursor of the pyrrolidine rings of nicotine and nornicotine in *Nicotiana glutinosa*. *Phytochemistry*, *24*(5), 957–960. https://doi.org/Doi: 10.1016/s0031-9422(00)83161-1

Lehfeldt, C., Shirley, A. M., Meyer, K., Ruegger, M. O., Cusumano, J. C., Viitanen, P. V., Strack, D., & Chapple, C. (2000). Cloning of the SNG1 gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *The Plant Cell*, *12*(8), 1295–1306. https://doi.org/10.1105/tpc.12.8.1295

Lichman, B. R. (2021). The scaffold-forming steps of plant alkaloid biosynthesis. *Natural Product Reports*, *38*(1), 103–129. https://doi.org/10.1039/D0NP00031K

Liu, H., Wang, X., Zhang, H., Yang, Y., Ge, X., & Song, F. (2008). A rice serine carboxypeptidase-like gene OsBISCPL1 is involved in regulation of defense responses against biotic and oxidative stress. *Gene*, *420*(1), 57–65. https://doi.org/10.1016/j.gene.2008.05.006

Liu, T., Kim, D. W., Niitsu, M., Maeda, S., Watanabe, M., Kamio, Y., Berberich, T., & Kusano, T. (2014). Polyamine Oxidase 7 is a Terminal Catabolism-Type Enzyme in *Oryza sativa* and is Specifically Expressed in Anthers. *Plant and Cell Physiology*, *55*(6), 1110–1122. https://doi.org/10.1093/pcp/pcu047

Lohman, J., Stunkard, L., Benjamin, A., & Boram, T. (2019). Exploring Enzymatic β-Keto Acid (De)Carboxylation with Malonyl-CoA Analogs. *The FASEB Journal*, *33*(S1), 633.17-633.17. https://doi.org/10.1096/fasebj.2019.33.1_supplement.633.17

Luck, K., Jirschitzka, J., Irmisch, S., Huber, M., Gershenzon, J., & Köllner, T. G. (2016). CYP79D enzymes contribute to jasmonic acid-induced formation of aldoximes and other nitrogenous volatiles in two Erythroxylum species. *BMC Plant Biology*, *16*(1), 215. https://doi.org/10.1186/s12870-016-0910-5

Luo, J., Fuell, C., Parr, A., Hill, L., Bailey, P., Elliott, K., Fairhurst, S. A., Martin, C., & Michael, A. J. (2009). A Novel Polyamine Acyltransferase Responsible for the Accumulation of Spermidine Conjugates in Arabidopsis Seed. *Plant Cell*, *21*(1), 318–333. https://doi.org/10.1105/tpc.108.063511

Macoy, D. M., Kim, W.-Y., Lee, S. Y., & Kim, M. G. (2015). Biosynthesis, physiology, and functions of hydroxycinnamic acid amides in plants. *Plant Biotechnology Reports*, *9*(5), 269–278. https://doi.org/10.1007/s11816-015-0368-1

Magallon, S., & Castillo, A. (2009). Angiosperm diversification through time. *American Journal of Botany*, *96*(1), 349–365. https://doi.org/10.3732/ajb.0800060

Majumdar, R., Barchi, B., Turlapati, S. A., Gagne, M., Minocha, R., Long, S., & Minocha, S. C. (2016). Glutamate, ornithine, arginine, proline, and polyamine metabolic interactions: The pathway is regulated at the post-transcriptional level. *Frontiers in Plant Science*, *7*, 78.

Majumdar, R., Shao, L., Minocha, R., Long, S., & Minocha, S. C. (2013). Ornithine: The overlooked molecule in the regulation of polyamine metabolism. *Plant and Cell Physiology*, *54*(6), 990–1004. https://doi.org/10.1093/pcp/pct053

Marillonnet, S., & Grützner, R. (2020). Synthetic DNA Assembly Using Golden Gate Cloning and the Hierarchical Modular Cloning Pipeline. *Current Protocols in Molecular Biology*, *130*(1), e115. https://doi.org/10.1002/cpmb.115

Matsufuji, S., Matsufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J. F., Gesteland, R. F., & Hayashi, S. (1995). Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell*, *80*(1), 51–60. https://doi.org/10.1016/0092-8674(95)90450-6

Matsuura, H. N., & Fett-Neto, A. G. (2015). Plant Alkaloids: Main Features, Toxicity, and Mechanisms of Action. In P. Gopalakrishnakone, C. R. Carlini, & R. Ligabue-Braun (Eds.), *Plant Toxins* (pp. 1–15). Springer Netherlands. https://doi.org/10.1007/978-94-007-6728-7_2-1

Minguet, E. G., Vera-Sirera, F., Marina, A., Carbonell, J., & Blázquez, M. A. (2008). Evolutionary Diversification in Polyamine Biosynthesis. *Molecular Biology and Evolution*, *25*(10), 2119–2128. https://doi.org/10.1093/molbev/msn161

Moghe, G., Kruse, L. H., Petersen, M., Scossa, F., Fernie, A. R., Gaquerel, E., & D'Auria, J. C. (2023). BAHD Company: The Ever-Expanding Roles of the BAHD Acyltransferase Gene Family in Plants. *Annual Review of Plant Biology*, *74*(1), null. https://doi.org/10.1146/annurev-arplant-062922-050122

Morales-Quintana, L., Moya-León, M. A., & Herrera, R. (2015). Computational study enlightens the structural role of the alcohol acyltransferase DFGWG motif. *Journal of Molecular Modeling*, *21*(8), 216. https://doi.org/10.1007/s00894-015-2762-6

Morana, A., Stiuso, P., Colonna, G., Lamberti, M., Cartenì, M., & De Rosa, M. (2002). Stabilization of S-adenosyl-L-methionine promoted by trehalose. *Biochimica Et Biophysica Acta*, *1573*(2), 105–108. https://doi.org/10.1016/s0304-4165(02)00333-1

Mugford, S. T., Qi, X., Bakht, S., Hill, L., Wegel, E., Hughes, R. K., Papadopoulou, K., Melton, R., Philo, M., Sainsbury, F., Lomonossoff, G. P., Roy, A. D., Goss, R. J. M., & Osbourn, A. (2009). A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *The Plant Cell*, *21*(8), 2473–2484. https://doi.org/10.1105/tpc.109.065870

Naconsie, M., Kato, K., Shoji, T., & Hashimoto, T. (2014). Molecular Evolution of N-Methylputrescine Oxidase in Tobacco. *Plant and Cell Physiology*, *55*(2), 436–444. https://doi.org/10.1093/pcp/pct179

Nakajima, K., Yamashita, A., Akama, H., Nakatsu, T., Kato, H., Hashimoto, T., Oda, J., & Yamada, Y. (1998). Crystal Structures of Two Tropinone Reductases: Different Reaction Stereospecificities in the Same Protein Fold. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(9), 4876–4881. https://doi.org/10.1073/pnas.95.9.4876

Napieraj, N., Janicka, M., & Reda, M. (2023). Interactions of Polyamines and Phytohormones in Plant Response to Abiotic Stress. *Plants*, *12*(5). https://doi.org/10.3390/plants12051159

Ober, D., & Kaltenegger, E. (2009). Pyrrolizidine alkaloid biosynthesis, evolution of a pathway in plant secondary metabolism. *Phytochemistry*, *70*(15), 1687–1695. https://doi.org/10.1016/j.phytochem.2009.05.017

O'Donovan, D. G., & Keogh, M. F. (1969). The role of hygrine in the biosynthesis of cuscohygrine and hyoscyamine. *Journal of the Chemical Society C: Organic*, *2*, 223–226. https://doi.org/10.1039/J39690000223

Ono, E., & Murata, J. (2023). Exploring the Evolvability of Plant Specialized Metabolism: Uniqueness Out Of Uniformity and Uniqueness Behind Uniformity. *Plant and Cell Physiology*, *64*(12), 1449–1465. https://doi.org/10.1093/pcp/pcad057

Ortiz, A., & Sansinenea, E. (2023). Phenylpropanoid Derivatives and Their Role in Plants' Health and as antimicrobials. *Current Microbiology*, *80*(12), 380. https://doi.org/10.1007/s00284-023-03502-x

Pálfi, P., Bakacsy, L., Kovács, H., & Szepesi, Á. (2021). Hypusination, a Metabolic Posttranslational Modification of eIF5A in Plants during Development and Environmental Stress Responses. *Plants*, *10*(7), Article 7. https://doi.org/10.3390/plants10071261

Panicot, M., Minguet, E. G., Ferrando, A., Alcázar, R., Blázquez, M. A., Carbonell, J., Altabella, T., Koncz, C., & Tiburcio, A. F. (2002). A Polyamine Metabolon Involving Aminopropyl Transferase Complexes in Arabidopsis. *The Plant Cell*, *14*(10), 2539–2551. https://doi.org/10.1105/tpc.004077

Patel, J., Ariyaratne, M., Ahmed, S., Ge, L., Phuntumart, V., Kalinoski, A., & Morris, P. F. (2017). Dual functioning of plant arginases provides a third route for putrescine synthesis. *Plant Science*, *262*, 62–73. https://doi.org/10.1016/j.plantsci.2017.05.011

Paul, A., Longchar, B., & Dkhar, J. (2021). Tropane Alkaloid Biosynthesis in Plants: Insights from Transcriptome Analysis. In V. Srivastava, S. Mehrotra, & S. Mishra (Eds.), *Tropane Alkaloids: Pathways, Potential and Biotechnological Applications* (pp. 133–156). Springer. https://doi.org/10.1007/978-981-33-4535-5_8

Peirats-Llobet, M., Yi, C., Liew, L. C., Berkowitz, O., Narsai, R., Lewsey, M. G., & Whelan, J. (2023). Spatially resolved transcriptomic analysis of the germinating barley grain. *Nucleic Acids Research*, *51*(15), 7798–7819. https://doi.org/10.1093/nar/gkad521

Peyret, H., Brown, J. K. M., & Lomonossoff, G. P. (2019). Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions. *Plant Methods*, *15*, 108. https://doi.org/10.1186/s13007-019-0494-9

Pichersky, E., & Lewinsohn, E. (2011). Convergent Evolution in Plant Specialized Metabolism. *Annual Review of Plant Biology*, *62*(1), 549–566. https://doi.org/10.1146/annurev-arplant-042110-103814

Pichersky, E., & Raguso, R. A. (2018). Why do plants produce so many terpenoid compounds? *New Phytologist*, *220*(3), 692–702. https://doi.org/10.1111/nph.14178

Ping, Y., Li, X. D., You, W. J., Li, G. Q., Yang, M. Q., Wei, W. P., Zhou, Z. H., & Xiao, Y. L. (2019). De Novo Production of the Plant-Derived Tropine and Pseudotropine in Yeast. *Acs Synthetic Biology*, *8*(6), 1257–1262. https://doi.org/10.1021/acssynbio.9b00152

Ping, Y., Li, X., Xu, B., Wei, W., Wei, W., Kai, G., Zhou, Z., & Xiao, Y. (2019). Building Microbial Hosts for Heterologous Production of *N*-Methylpyrrolinium. *ACS Synthetic Biology*, *8*(2), 257–263. https://doi.org/10.1021/acssynbio.8b00483

Piotrowski, M., Janowitz, T., & Kneifel, H. (2003). Plant CN hydrolases and the identification of a plant N-carbamoylputrescine amidohydrolase involved in polyamine biosynthesis. *Journal of Biological Chemistry*, *278*(3), 1708–1712.

Plowman, T. (1984). The ethnobotany of Coca (*Erythroxylum* spp., Erythroxylaceae). In G. T. Prance & J. A. Kallunki (Eds.), *Ethnobotany in the Neotropics* (pp. 62–111). New York Botanical Garden.

Porat, Z., Landau, G., Bercovich, Z., Krutauz, D., Glickman, M., & Kahana, C. (2008). Yeast Antizyme Mediates Degradation of Yeast Ornithine Decarboxylase by Yeast but Not by Mammalian Proteasome: NEW INSIGHTS ON YEAST ANTIZYME*. *Journal of Biological Chemistry*, *283*(8), 4528–4534. https://doi.org/10.1074/jbc.M708088200

Qiu, F., Zeng, J., Wang, J., Huang, J. P., Zhou, W., Yang, C., Lan, X., Chen, M., Huang, S. X., Kai, G., & Liao, Z. (2019). Functional genomics analysis reveals two novel genes required for littorine biosynthesis. *New Phytologist*. https://doi.org/10.1111/nph.16317

Reiter, J., Hübbers, A. M., Albrecht, F., Leichert, L. I. O., & Slusarenko, A. J. (2020). Allicin, a natural antimicrobial defence substance from garlic, inhibits DNA gyrase activity in bacteria. *International Journal of Medical Microbiology*, *310*(1), 151359. https://doi.org/10.1016/j.ijmm.2019.151359

Restrepo, D. A., Saenz, E., Jara-Muñoz, O. A., Calixto-Botía, I. F., Rodríguez-Suárez, S., Zuleta, P., Chavez, B. G., Sanchez, J. A., & D'Auria, J. C. (2019). Erythroxylum in Focus: An Interdisciplinary Review of an Overlooked Genus. *Molecules*, *24*(20), 3788. https://doi.org/10.3390/molecules24203788

Robinson, R. (1917). LXIII.-A synthesis of tropinone. *Journal of the Chemical Society, Transactions*, *111*(0), 762–768.

Roje, S. (2006). S-Adenosyl-l-methionine: Beyond the universal methyl group donor. *Phytochemistry*, *67*(15), 1686–1698. https://doi.org/10.1016/j.phytochem.2006.04.019

Rothe, G., Hachiya, A., Yamada, Y., Hashimoto, T., & Dräger, B. (2003). Alkaloids in plants and root cultures of Atropa belladonna overexpressing putrescine N-methyltransferase. *Journal of Experimental Botany*, *54*(390), 2065–2070. https://doi.org/10.1093/jxb/erg227

Rubio, N. C., Thurmann, D., Krumbiegel, F., & Pragst, F. (2016). Behaviour of hygrine and cuscohygrine in illicit cocaine production establishes their use as markers for chewing coca leaves in contrast with cocaine abuse. *Drug Testing and Analysis*, *9*, n/a-n/a. https://doi.org/10.1002/dta.1972

Sakamoto, A., Sahara, J., Kawai, G., Yamamoto, K., Ishihama, A., Uemura, T., Igarashi, K., Kashiwagi, K., & Terui, Y. (2020). Cytotoxic Mechanism of Excess Polyamines Functions through Translational Repression of Specific Proteins Encoded by Polyamine Modulon. *International Journal of Molecular Sciences*, *21*(7), 2406. https://doi.org/10.3390/ijms21072406

Sánchez-López, J., Camañes, G., Flors, V., Vicent, C., Pastor, V., Vicedo, B., Cerezo, M., & García-Agustín, P. (2009). Underivatized polyamine analysis in plant samples by ion pair LC coupled with electrospray tandem mass spectrometry. *Plant Physiology and Biochemistry*, *47*(7), 592–598. https://doi.org/10.1016/j.plaphy.2009.02.006

Savadel, S. D., Hartwig, T., Turpin, Z. M., Vera, D. L., Lung, P.-Y., Sui, X., Blank, M., Frommer, W. B., Dennis, J. H., Zhang, J., & Bass, H. W. (2021). The native cistrome and sequence motif families of the maize ear. *PLOS Genetics*, *17*(8), e1009689. https://doi.org/10.1371/journal.pgen.1009689

Schmidt, G. W., Jirschitzka, J., Porta, T., Reichelt, M., Luck, K., Torre, J. C. P., Dolke, F., Varesio, E., Hopfgartner, G., Gershenzon, J., & D'Auria, J. C. (2015). The Last Step in Cocaine Biosynthesis Is Catalyzed by a BAHD Acyltransferase. *Plant Physiology*, *167*(1), 89–101. https://doi.org/10.1104/pp.114.248187

Schubert, H. L., Blumenthal, R. M., & Cheng, X. (2003). Many paths to methyltransfer: A chronicle of convergence. *Trends in Biochemical Sciences*, *28*(6), 329–335. https://doi.org/10.1016/S0968-0004(03)00090-2

Sekula, B., & Dauter, Z. (2019a). Spermidine Synthase (SPDS) Undergoes Concerted Structural Rearrangements Upon Ligand Binding – A Case Study of the Two SPDS Isoforms From Arabidopsis thaliana. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00555

Sekula, B., & Dauter, Z. (2019b). Structural Study of Agmatine Iminohydrolase From Medicago truncatula, the Second Enzyme of the Agmatine Route of Putrescine Biosynthesis in Plants. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.00320

Serre, N. B. C., Alban, C., Bourguignon, J., & Ravanel, S. (2018). An outlook on lysine methylation of non-histone proteins in plants. *Journal of Experimental Botany*, *69*(19), 4569–4581. https://doi.org/10.1093/jxb/ery231

Serre, N. B. C., Sarthou, M., Gigarel, O., Figuet, S., Corso, M., Choulet, J., Rofidal, V., Alban, C., Santoni, V., Bourguignon, J., Verbruggen, N., & Ravanel, S. (2020). Protein lysine methylation contributes to modulating the response of sensitive and tolerant Arabidopsis species to cadmium stress. *Plant, Cell & Environment*, *43*(3), 760–774. https://doi.org/10.1111/pce.13692

Sharma, E., Anand, G., & Kapoor, R. (2017). Terpenoids in plant and arbuscular mycorrhiza-reinforced defence against herbivorous insects. *Annals of Botany*, *119*(5), 791–801. https://doi.org/10.1093/aob/mcw263

Shaw, W. V. (1992). Chemical anatomy of antibiotic resistance: Chloramphenicol acetyltransferase. *Science Progress*, *76*(301-302 Pt 3-4), 565–580.

Shockey, J., & Browse, J. (2011). Genome-level and biochemical diversity of the acyl-activating enzyme superfamily in plants. *The Plant Journal: For Cell and Molecular Biology*, *66*(1), 143–160. https://doi.org/10.1111/j.1365-313X.2011.04512.x

Sperschneider, J., Catanzariti, A.-M., DeBoer, K., Petre, B., Gardiner, D. M., Singh, K. B., Dodds, P. N., & Taylor, J. M. (2017). LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports*, *7*, 44598. https://doi.org/10.1038/srep44598

Srinivasan, P., & Smolke, C. D. (2019). Engineering a microbial biosynthesis platform for de novo production of tropane alkaloids. *Nature Communications*, *10*(1), 3634. https://doi.org/10.1038/s41467-019-11588-w

Srinivasan, P., & Smolke, C. D. (2020). Biosynthesis of medicinal tropane alkaloids in yeast. *Nature*, *585*(7826), 614–619. https://doi.org/10.1038/s41586-020-2650-9

Srinivasan, P., & Smolke, C. D. (2021). Engineering cellular metabolite transport for biosynthesis of computationally predicted tropane alkaloid derivatives in yeast. *Proceedings of the National Academy of Sciences*, *118*(25). https://doi.org/10.1073/pnas.2104460118

Stolberg, V. B. (2011). The use of coca: Prehistory, history, and ethnography. *J Ethn Subst Abuse*, *10*(2), 126–146. https://doi.org/10.1080/15332640.2011.573310

St-Pierre, B., & De Luca, V. (2000). Chapter Nine - Evolution of Acyltransferase Genes: Origin and Diversification of the BAHD Superfamily of Acyltransferases Involved in Secondary Metabolism. In J. T. Romeo, R. Ibrahim, L. Varin, & V. De Luca (Eds.), *Recent Advances in Phytochemistry* (Vol. 34, pp. 285–315). Elsevier. https://doi.org/10.1016/S0079-9920(00)80010-6

Su, X., Li, X., Wang, H., & Cai, Z. (2021). Simultaneous determination of methionine cycle metabolites, urea cycle intermediates and polyamines in serum, urine and intestinal tissue by using UHPLC-MS/MS. *Talanta*, *224*, 121868. https://doi.org/10.1016/j.talanta.2020.121868

Takano, A., Kakehi, J.-I., & Takahashi, T. (2012). Thermospermine is Not a Minor Polyamine in the Plant Kingdom. *Plant and Cell Physiology*, *53*(4), 606–616. https://doi.org/10.1093/pcp/pcs019

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Tavladoraki, P., Cona, A., & Angelini, R. (2016). Copper-Containing Amine Oxidases and FAD-Dependent Polyamine Oxidases Are Key Players in Plant Tissue Differentiation and Organ Development. *Frontiers in Plant Science*, *7*, 824. https://doi.org/10.3389/fpls.2016.00824

Teuber, M., Azemi, M. E., Namjoyan, F., Meier, A. C., Wodak, A., Brandt, W., & Drager, B. (2007). Putrescine N-methyltransferases—A structure-function analysis. *Plant Molecular Biology*, *63*(6), 787–801. https://doi.org/10.1007/s11103-006-9126-7

Thomas, S. N., French, D., Jannetto, P. J., Rappold, B. A., & Clarke, W. A. (2022). Liquid chromatography–tandem mass spectrometry for clinical diagnostics. *Nature Reviews Methods Primers*, *2*(1), Article 1. https://doi.org/10.1038/s43586-022-00175-x

Tian, T., Wang, Y.-J., Huang, J.-P., Li, J., Xu, B., Chen, Y., Wang, L., Yang, J., Yan, Y., & Huang, S.-X. (2022). Catalytic innovation underlies independent recruitment of polyketide synthases in cocaine and hyoscyamine biosynthesis. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-32776-1

Tiburcio, A. F., Altabella, T., Bitrián, M., & Alcázar, R. (2014). The roles of polyamines during the lifespan of plants: From development to stress. *Planta*, *240*(1), 1–18. https://doi.org/10.1007/s00425-014-2055-9

Torre, J. C. P., Schmidt, G. W., Paetz, C., Reichelt, M., Schneider, B., Gershenzon, J., & D'Auria, J. C. (2013). The biosynthesis of hydroxycinnamoyl quinate esters and their role in the storage of cocaine in Erythroxylum coca. *Phytochemistry*, *91*, 177–186. https://doi.org/10.1016/j.phytochem.2012.09.009

Trifilieff, P., & Martinez, D. (2014). Chapter Five - Cocaine: Mechanism and Effects in the Human Brain. In B. Madras & M. Kuhar (Eds.), *The Effects of Drug Abuse on the Human Nervous System* (pp. 103–133). Academic Press. https://doi.org/10.1016/B978-0-12-418679-8.00005-8

Upadhyay, R. K., Shao, J., & Mattoo, A. K. (2021). Genomic analysis of the polyamine biosynthesis pathway in duckweed Spirodela polyrhiza L.: Presence of the arginine decarboxylase pathway, absence of the ornithine decarboxylase pathway, and response to abiotic stresses. *Planta*, *254*(5), 108. https://doi.org/10.1007/s00425-021-03755-5

Urano, K., Hobo, T., & Shinozaki, K. (2005). Arabidopsis ADC genes involved in polyamine biosynthesis are essential for seed development. *FEBS Letters*, *579*(6), 1557–1564. https://doi.org/10.1016/j.febslet.2005.01.048

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., & Pompon, D. (1997). Cloning, yeast expression, and characterization of the coupling of two distantly related Arabidopsis thaliana NADPH-cytochrome P450 reductases with P450 CYP73A5. *The Journal of Biological Chemistry*, *272*(31), 19176–19186. https://doi.org/10.1074/jbc.272.31.19176

van Eunen, K., & Bakker, B. M. (2014). The importance and challenges of *in vivo*-like enzyme kinetics. *Perspectives in Science*, *1*(1), 126–130. https://doi.org/10.1016/j.pisc.2014.02.011

Veeresham, C. (2012). Natural products derived from plants as a source of drugs. *J Adv Pharm Technol Res*, *3*(4), 200–201. https://doi.org/10.4103/2231-4040.104709

Vuosku, J., Karppinen, K., Muilu-Mäkelä, R., Kusano, T., Sagor, G. H. M., Avia, K., Alakärppä, E., Kestilä, J., Suokas, M., Nickolov, K., Hamberg, L., Savolainen, O., Häggman, H., & Sarjala, T. (2018). Scots pine aminopropyltransferases shed new light on evolution of the polyamine biosynthesis pathway in seed plants. *Annals of Botany*, *121*(6), 1243–1256. https://doi.org/10.1093/aob/mcy012

Wang, X., Ying, W., Dunlap, K. A., Lin, G., Satterfield, M. C., Burghardt, R. C., Wu, G., & Bazer, F. W. (2014). Arginine decarboxylase and agmatinase: An alternative pathway for de novo biosynthesis of polyamines for development of mammalian conceptuses. *Biology of Reproduction*, *90*(4), 84. https://doi.org/10.1095/biolreprod.113.114637

Wang, Y.-J., Huang, J.-P., Tian, T., Yan, Y., Chen, Y., Yang, J., Chen, J., Gu, Y.-C., & Huang, S.-X. (2022). Discovery and Engineering of the Cocaine Biosynthetic Pathway. *Journal of the American Chemical Society*, jacs.2c09091. https://doi.org/10.1021/jacs.2c09091

Wang, Y.-J., Tain, T., Yu, J.-Y., Li, J., Xu, B., Chen, J., D'Auria, J. C., Huang, J.-P., & Huang, S.-X. (2023). Genomic and structural basis for evolution of tropane alkaloid

biosynthesis. *Proceedings of the National Academy of Sciences*, *120*(17), e2302448120. https://doi.org/10.1073/pnas.2302448120

Wei, J., Wang, L., Zhu, J., Zhang, S., Nandi, O. I., & Kang, L. (2007). Plants Attract Parasitic Wasps to Defend Themselves against Insect Pests by Releasing Hexenol. *PLOS ONE*, *2*(9), e852. https://doi.org/10.1371/journal.pone.0000852

WHO Model List of Essential Medicines—23rd list, 2023. (2023). https://www.who.int/publications-detail-redirect/WHO-MHP-HPS-EML-2023.02

Willstätter, R. (1901a). Synthese des Tropidins. *Berichte Der Deutschen Chemischen Gesellschaft*, *34*(1), 129–144. https://doi.org/10.1002/cber.19010340124

Willstätter, R. (1901b). Umwandlung von Tropidin in Tropin. *Berichte Der Deutschen Chemischen Gesellschaft*, *34*(2), 3163–3165. https://doi.org/10.1002/cber.190103402289

Wink, M. (2007). Ecological Roles of Alkaloids. In *Modern Alkaloids* (pp. 1–24). John Wiley & Sons, Ltd. https://doi.org/10.1002/9783527621071.ch1

Wink, M. (2019). Evolution of the Angiosperms and Co-evolution of Secondary Metabolites, Especially of Alkaloids. In J.-M. Merillon & K. G. Ramawat (Eds.), *Co-Evolution of Secondary Metabolites* (pp. 1–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-76887-8_22-1

Xiong, H., Stanley, B. A., Tekwani, B. L., & Pegg, A. E. (1997). Processing of mammalian and plant S-adenosylmethionine decarboxylase proenzymes. *The Journal of Biological Chemistry*, *272*(45), 28342–28348. https://doi.org/10.1074/jbc.272.45.28342

Yang, J., Wu, Y., Zhang, P., Ma, J., Yao, Y. J., Ma, Y. L., Zhang, L., Yang, Y., Zhao, C., Wu, J., Fang, X., & Liu, J. (2023). Multiple independent losses of the biosynthetic pathway for two tropane alkaloids in the Solanaceae family. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-44246-3

Yang, L., & Stöckigt, J. (2010). Trends for diverse production strategies of plant medicinal alkaloids. *Natural Product Reports*, *27*(10), 1469–1479. https://doi.org/10.1039/c005378c

Yu, Z., Jia, D., & Liu, T. (2019). Polyamine Oxidases Play Various Roles in Plant Development and Abiotic Stress Tolerance. *Plants*, *8*(6), Article 6. https://doi.org/10.3390/plants8060184

Zhao, Y., Coelho, C., Hughes, A. L., Lazar-Stefanita, L., Yang, S., Brooks, A. N., Walker, R. S. K., Zhang, W., Lauer, S., Hernandez, C., Cai, J., Mitchell, L. A., Agmon, N., Shen, Y., Sall, J., Fanfani, V., Jalan, A., Rivera, J., Liang, F.-X., … Boeke, J. D. (2023). Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions. *Cell*, S0092867423010796. https://doi.org/10.1016/j.cell.2023.09.025

Zhu, M.-Y., Iyo, A., Piletz, J. E., & Regunathan, S. (2004). Expression of human arginine decarboxylase, the biosynthetic enzyme for agmatine. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1670*(2), 156–164. https://doi.org/10.1016/j.bbagen.2003.11.006

Zierer, W., Hajirezaei, M. R., Eggert, K., Sauer, N., von Wirén, N., & Pommerrenig, B. (2016). Phloem-Specific Methionine Recycling Fuels Polyamine Biosynthesis in a

Sulfur-Dependent Manner and Promotes Flower and Seed Development. *Plant Physiology*, *170*(2), 790–806. https://doi.org/10.1104/pp.15.00786

Zotter, A., Bäuerle, F., Dey, D., Kiss, V., & Schreiber, G. (2017). Quantifying enzyme activity in living cells. *The Journal of Biological Chemistry*, *292*(38), 15838–15848. https://doi.org/10.1074/jbc.M117.792119

Zubieta, C., Ross, J. R., Koscheski, P., Yang, Y., Pichersky, E., & Noel, J. P. (2003). Structural Basis for Substrate Recognition in the Salicylic Acid Carboxyl Methyltransferase Family. *The Plant Cell*, *15*(8), 1704–1716. https://doi.org/10.1105/tpc.014548

8 APPENDIX

8.1.1 Protein Alignment of SPDS-like Candidates with PMTs and SPDSs Enzymes

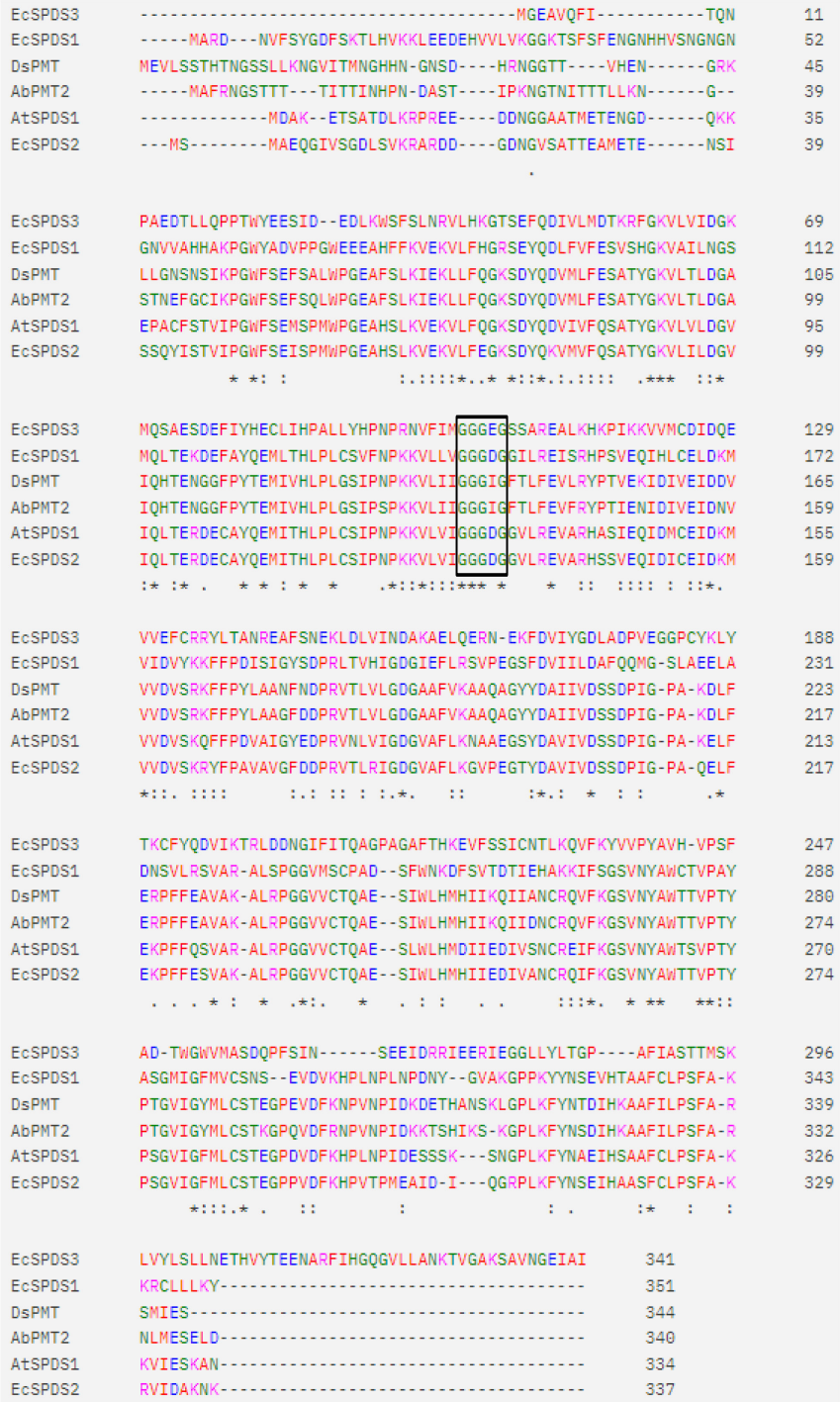

Figure S1: Protein alignment of EcSPDS1-3 candidates, putrescine *N*-methyltransferases from *Datura stramonium* and *Atropa belladonna*, and spermidine synthase 1 from *Arabidopsis thaliana*. The black box indicates the SAM and dcSAM binding domains between PMT and SPDS enzymes. Alignments were performed with Clustal Omega using standard settings: https://www.ebi.ac.uk/jdispatcher/msa/clustalo. The symbols for the protein alignment are defined as *(Asterix) positions with a single, fully conserved Residue, :(colon) positions

with a conservation between amino acid groups of similar properties, and .(period) positions with a conservation between amino acid groups of weakly similar properties.

8.1.2 Heterologous Expression in *Komagataella phaffii* and Purification of SPDS Candidates

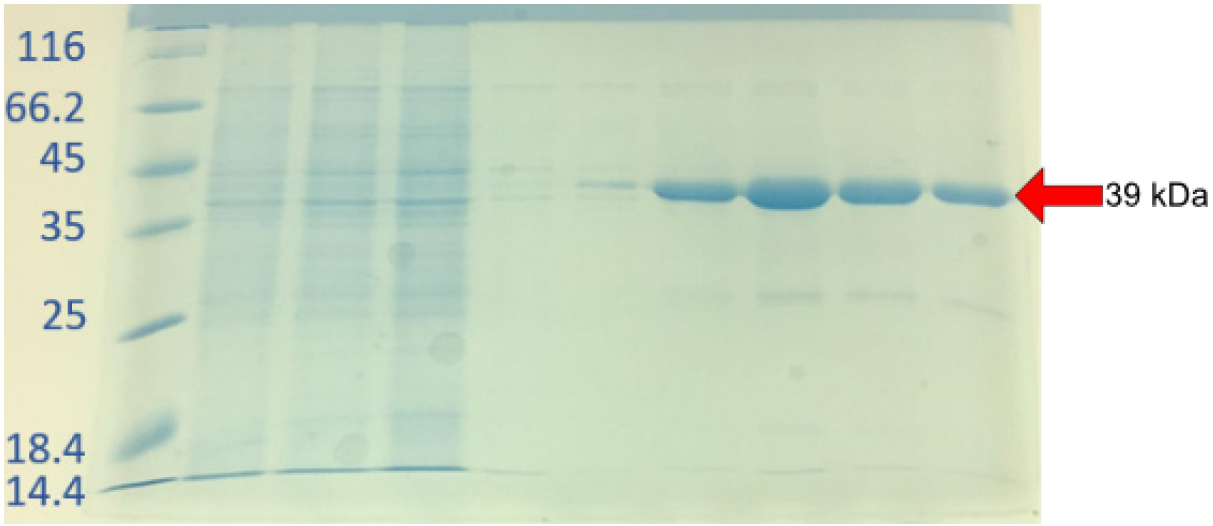

Figure S2: SDS-PAGE of streptag purification of EcSMT from *K. phaffii* KM71. Lane 1 is PageRuler Protein Ladder (Catalog # 26616), lane 2 *K. phaffii* is pellet after centrifugation, lane 3 is the crude supernatant, lane 4 is the flowthrough, lane 5 is wash, lane 6 is EcSPMT protein first fraction, and lane 7 is EcSPMT protein second fraction. Published in Chavez et al., 2022; Supplemental Material Page 19; Figure S4.

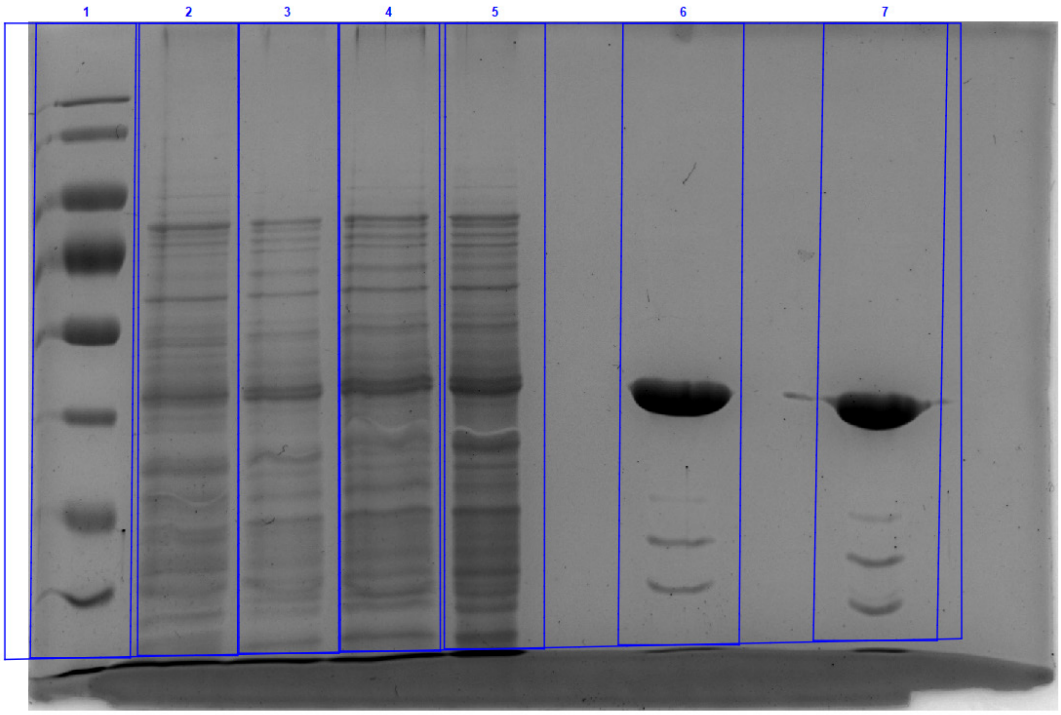

Figure S3: SDS-PAGE of streptag purification of EcSPMT from *K. phaffii* KM71. Lane 1 is PageRuler Protein Ladder (Catalog # 26616), lane 2 *K. phaffii* is pellet after centrifugation, lane 3 is the crude supernatant, lane 4 is the flowthrough, lane 5 is the wash, lane 6 is the 1st EcSPMT protein fraction, lane 7 is the EcSPMT 2nd protein

fraction. EcSPMT is the prominent band in lane 6, and lane 7 is at the predicted size of 36.8 kDa. SDS-PAGE stained with Serva Quick Coomassie® Stain for 2-3 hours.

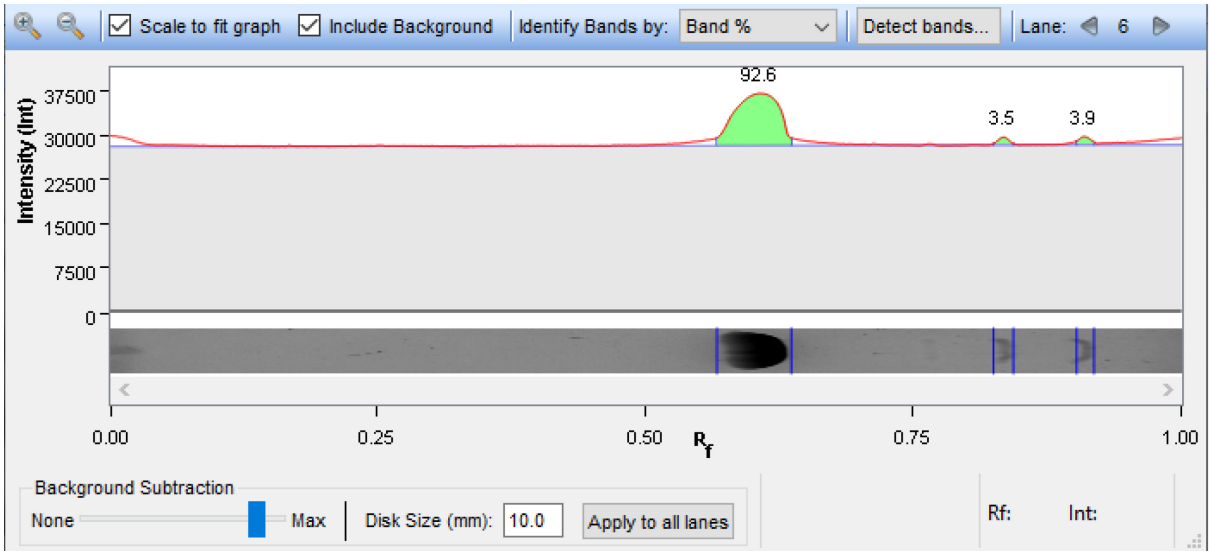

Figure S4: EcSPMT densitometric analysis of lane 6 band percentage from the SDS-PAGE. The prominent protein band is at the correct size of 36.8 kDa and makes up 92.6% of the total protein isolated in lane 6.

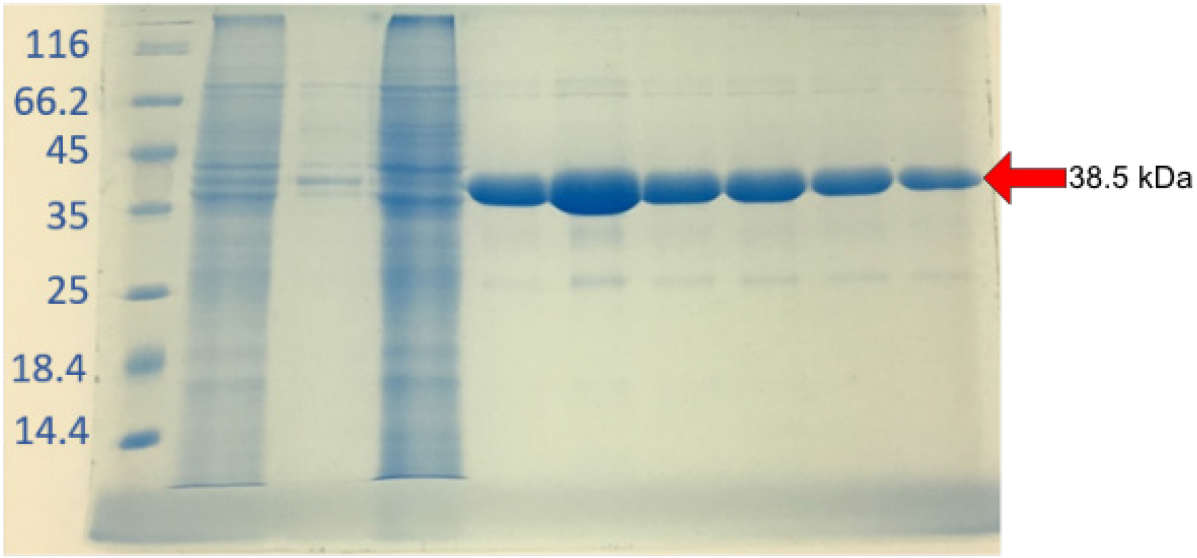

Figure S5: SDS-PAGE of streptag purification of EcSPDS from *K. phaffii* KM71. Lane 1 is PageRuler Protein Ladder (Catalog # 26616), lane 2 *K. phaffii* is pellet after centrifugation, lane 3 is the crude supernatant, and lane 4 is the flowthrough. Published in Chavez et al., 2022; Supplemental Material Page 21; Figure S6.

8.1.3 SPDS Sequence Accession for Phylogenetic Analysis

Table S1: Key for the abbreviations of SPDS-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. Published in Chavez et al., 2022; Supplemental Dataset; Table S1.

Species Name	Abbreviated name	Accession Number
Aa- Anisodus acutangulus	AaPMT2	ACF21006.1
Aa- Anisodus acutangulus	AaPMT1	ACF21005.1
Aa-Artemisia annua	AaSPDSYN1	PWA72973.1
Aa-Artemisia annua	AaSPDSYN2	PWA65721.1
Ab- Atropa belladonna	AbPMT1	BAA82264.1
Ab- Atropa belladonna	AbPMT2	BAA82262.1
As- Apostasia shenzhenica	AsSPS	PKA66679.1
As-Apostasia shenzhenica	AsSPDSYN1	PKA55102.1
At- Anisodus tanguticus	AtPMT	AAT99576.1
At- Arabidopsis thaliana	AtSPDSYN1	AEE30436.2
At- Arabidopsis thaliana	AtSPDSYN2	AEE35044.1
At- Arabidopsis thaliana	AtSPMS	AAM64782.1
At- Arabidopsis thaliana	AtACL5	BAB83652.1
Bs- Bacillus subtilis	BsSPEE	AOS00134.1
Ca- Capsicum annuum	CaSPS	PHT68126.1
Ca- Capsicum annuum	CaSPDS	PHT87459.1
Ca- Coffea arabica	CaSPDS	BAA29033.1
Cb- Capsicum baccatum	CbSPDS1	PHT34202.1
Cc- Cajanus cajan	CcSPDS2	KYP69872.1
Cc- Cajanus cajan	CcSPDS3	KYP41027.1
Cc- Cajanus cajan	CcSPDS4	KYP66172.1
Cc- Capsicum chinense	CcPMT3	PHU08967.1
Cc- Capsicum chinense	CcSPDS1	PHU02816.1
Cc- Capsicum chinense	CcSPDS2	PHU23234.1
Ce- Caenorhabditis elegans	CeSPDS1	CAC37332.1
Cr- Chlamydomonas reinhardtii	CrSPD1	XP_001702843.1
Cr- Chlamydomonas reinhardtii	CrSPS1	XP_001696651.1
Cs- Calystegia sepium	CsPMT	CAJ46252.1
Cs- Chlorella sorokiniana	CsSPDS1	PRW60069.1
Cs- Coccomyxa subellipsoidea	CsSPDS2	EIE22898.1
Cs- *Cucumis sativus*	CsSPDS3	AAT66041.1
Cs- Cucumis sativus	CsSPS	KGN56994.1
Cu- Citrus unshiu	CuSPDS	BAK61861.1
Dc- Dendrobium catenatum	DcSPDS1	PKU63394.1
Dc- Dendrobium catenatum	DcSPS	PKU68007.1
Dh- Dorcoceras hygrometricum	DhSPS1	KZV53164.1
Dh- Dorcoceras hygrometricum	DhSPS2	KZV50788.1
Di- Datura inoxia	DiPMT1	CAJ46253.1
Di- Datura inoxia	DiPMT2	CAJ46254.1
Dm- Datura metel	DmPMT	AAQ94738.1
Ds- Datura stramonium	DsPMT	CAE47481.1

Ds- Datura stramonium	DsSPDS1	CAA69420.1
Ds- Datura stramonium	DsSPDS2	CAA69421.1
Ec- Erythroxylum coca	EcSPDS	OP382841
Ec- Erythroxylum coca	EcSPMT	OP382839
Ec- Erythroxylum coca	EcSMT	OP382840
Gh- Gossypium hirsutum	GhSPDS	AHH02594.1
Hn- Hyoscyamus niger	HnPMT	BAA82263.1
Hn- Hyoscyamus niger	HnSPDS1	BAA24533.1
Hn- Hyoscyamus niger	HnSPDS2	BAA24534.1
Hs- Homo sapiens	HsSPDSY	NP_003123.2
Kn- Klebsormidium nitens	KnSPDS	GAQ81179.1
Kn- Klebsormidium nitens	KnSPS	GAQ87780.1
Mc- Micractinium conductrix	McSPDS	PSC73114.1
Md- Malus domestica	MdSPDS2a	BAC20171.1
Mn- Monoraphidium neglectum	MnSPDS1	KIZ00759.1
Mn- Monoraphidium neglectum	MnSPDS2	KIY96353.1
Ms- Malus sylvestris	MsSPS	BAE19758.1
Mt- Medicago truncatula	MtSPDS	KEH35054.1
Mt- Medicago truncatula	MtSPDS2	KEH27761.1
Mt- Medicago truncatula	MtSPDS3	AES96204.1
Mv- Marsilea vestita	MvSPDS	ADK89558.1
Na- Nicotiana attenuata	NaPMT2	AAK49871.1
Nb- Nicotiana benthamiana	NbPMT	ABY25273.1
Nc- Neurospora crassa	NcSPDSY	BAA81738.1
Ns- Nicotiana sylvestris	NsPMT3	BAA74544.1
Ns- Nicotiana sylvestris	NsSPDS	NP_001289514.1
Nt-Nicotiana tabacum	NtPMT4	AAF14881.1
Oe- ***Olea europaea***	OeSPDS	ACZ73829
Ol- Ostreococcus lucimarinus	OlSPD2	ABO96661.1
Ol- Ostreococcus lucimarinus	OlSPD1	ABO98745.1
Pd- Physalis divaricata	PdPMT	CAJ46255.1
Pg- Panax ginseng	PgSPDS	ACT21542.1
Ps- Pisum sativum	PsSPDS1	AAD02231.1
Ps- Pisum sativum	PsSPDS2	Q9ZTR0
Rn- Rattus norvegicus	RnSPS	NP_001029071.1
Rs- Raphidocelis subcapitata	RsSPDS	GBF99043.1
Rs- Raphidocelis subcapitata	RsSPS	GBF87456.1
Sc- Saccharomyces cerevisiae	ScSPE3	AAC17191.1
Sd- Solanum dulcamara	SdPMT	CAQ19733.1
Sl- Solanum lycopersicum	SlPMT	CAJ46251.1
Sl- Solanum lycopersicum	SlSPDS	NP_001234493.1
St- Solanum tuberosum	StPMT	CAE53633.1
St- Solanum tuberosum	StSPDS	CAC51027.1
Tc- Theobroma cacao	TcSPDS2	EOY20673.1
Tc- Theobroma cacao	TcSPDS1	EOY20671.1
Tc- Theobroma cacao	TcSPDS3	EOY00385.1
Tm- Thermotoga maritima	TmSPDS1	AKE26575.1

Tp- Trifolium pratense	TpSPDS	PNY16312.1
Ts- Tetrabaena socialis	TsSPDS1	PNH01615.1
Ts- Tetrabaena socialis	TsSPDS2	PNH00286.1
Tu- Triticum urartu	TuSPDS	EMS48579.1
Tu- Triticum urartu	TuSPS1	EMS58513.1
Tu- Triticum urartu	TuSPS2	EMS48760.1
Zm- Zea mays	ZmSPDS2	AAW57523.1
Zm- Zea mays	ZmSPDS1	NP_001149286.1
Zm- Zea mays	ZmSPS1	NP_001149310.1
Zm- Zea mays	ZmSPS2	PWZ27668.1

8.1.4 Flavin-Dependent Oxidases Sequences Accession for Phylogenetic Analysis

Table S2: A key for the abbreviations of AOF-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. Published in Chavez et al., 2022; Supplemental Dataset; Table S1.

Species Name	Description/Predicted function	Accession Number
Arabidopsis thaliana	Lysine-specific histone demethylase	Q8VXV7
Arabidopsis thaliana	Lysine-specific histone demethylase	Q9LID0
Arabidopsis thaliana	Lysine-specific histone demethylase	F4JLS1
Arabidopsis thaliana	FLD_ARATH Protein	Q9CAE3
Arabidopsis thaliana	PAO1_ARATH Polyamine oxidase	Q9FNA2
Arabidopsis thaliana	PAO2_ARATH Polyamine oxidase	Q9SKX5
Arabidopsis thaliana	PAO3_ARATH Polyamine oxidase	Q9LYT1
Arabidopsis thaliana	PAO4_ARATH Probable polyamine oxidase	Q8H191
Arabidopsis thaliana	PAO5_ARATH Probable polyamine oxidase	Q9SU79
Arabidopsis thaliana	CRTSO_ARATH Prolycopene isomerase	Q9M9Y8
Arabidopsis thaliana	PDS_ARATH 15-cis-phytoene desaturase	Q07356
Arabidopsis thaliana	ZDS_ARATH Zeta-carotene desaturase	Q38893
Arabidopsis thaliana	PPOC_ARATH Protoporphyrinogen oxidase	P55826
Arabidopsis thaliana	PPOCM_ARATH Protoporphyrinogen oxidase	Q8S9J1
Capsicum annuum	PDS_CAPAN 15-cis-phytoene desaturase	P80093
Capsicum annuum	ZDS_CAPAN Zeta-carotene desaturase	Q9SMJ3
Dictyostelium discoideum	PPOX_DICDI Protoporphyrinogen oxidase	Q54DT8
Erythroxylum coca	EcAOF1 methylspermidine oxidase	OP382842
Glycine max	PDS_SOYBN 15-cis-phytoene desaturase	P28553
Myxococcus xanthus	CRTJ_MYXXA zeta-carotene-forming phytoene desaturase	P54979
Narcissus pseudonarcissus	PDS_NARPS 15-cis-phytoene desaturase	Q40406
Nicotiana tabacum	PPOC_TOBAC Protoporphyrinogen oxidase	O24163
Nicotiana tabacum	PPOM_TOBAC Protoporphyrinogen oxidase	O24164
***Oncidium* hybrid cultivar**	ZDS_ONCHC Zeta-carotene desaturase	C3VEQ0
Oryza sativa subsp. japonica	Lysine-specific histone demethylase	Q6Z690

Oryza sativa subsp. japonica	PAO3_ORYSJ Polyamine oxidase	Q7X809
Oryza sativa subsp. japonica	PAO4_ORYSJ Polyamine oxidase	Q7XR46
Oryza sativa subsp. japonica	PAO5_ORYSJ Polyamine oxidase	Q0J954
Oryza sativa subsp. japonica	PAO6_ORYSJ Polyamine oxidase	A0A0P0XM10
Oryza sativa subsp. japonica	PAO7_ORYSJ Polyamine oxidase	Q0J290
Oryza sativa subsp. japonica	PDS_ORYSJ 15-cis-phytoene desaturase	Q0DUI8
Oryza sativa subsp. japonica	PPOC_ORYSJ Protoporphyrinogen oxidase	Q9AR38
Saccharomyces cerevisiae	FMS1_YEAST Polyamine oxidase	P50264
Solanum lycopersicum	PDS_SOLLC 15-cis-phytoene desaturase	P28554
Solanum lycopersicum	CRTSO_SOLLC Prolycopene isomerase	Q8S4R4
Solanum lycopersicum	ZDS_SOLLC Zeta-carotene desaturase	Q9SE20
Spinacia oleracea	PPOC_SPIOL Protoporphyrinogen oxidase	Q9LRI8
Spinacia oleracea	PPOCM_SPIOL Protoporphyrinogen oxidase	Q94IG7
Zea mays	PAO1_MAIZE Polyamine oxidase	O64411
Zea mays	PDS_MAIZE 15-cis-phytoene desaturase	P49086
Zea mays	ZDS_MAIZE Zeta-carotene desaturase	Q9ZTP4

8.1.5 Copper-Dependent Oxidases Sequence Accession for Phylogenetic Analysis

Table S3: A key for the abbreviations of AOC-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. Published in Chavez et al., 2022; Supplemental Dataset; Table S1.

Species Name	Description/Predicted function	Accession Number
Amborella trichopoda	Amt2 uncharacterized protein	XP_006845257.1
Amborella trichopoda	Amt1 primary amine oxidase	XP_006857816.2
Atropa belladonna	AbMPO2 *N*-methylputrescine oxidase	n/a
Brassica juncea	Bj1 diamine oxidase	AAL47166.1
Cicer arietinum	Ca1 copper amine oxidase	NP_001265996.1
Erythroxylum coca	EcAOC1 *N*-methylputrescine oxidase	OP382843
Erythroxylum coca	EcAOC2 *N*-methylputrescine oxidase	OP382844
Erythroxylum coca	EcAOC3 Unknown	n/a
Erythroxylum coca	EcAOC4 Unknown	n/a
Erythroxylum coca	EcAOC5 Unknown	n/a
Erythroxylum coca	EcAOC6 Unknown	n/a
Glycine max	Gm1 copper amino oxidase	NP_001237211.1
Glycine max	Gm9 primary amine oxidase	XP_003546898.1
Glycine max	Gm8 copper methylamine oxidase	XP_003550715.1
Glycine max	Gm7 copper methylamine oxidase	XP_003551224.1
Glycine max	Gm2 primary amine oxidase	XP_003556043.1
Glycine max	Gm3 uncharacterized protein	KRH35530.1
Gossypium hirsutum	Gh polyamine oxidase	AGO02008.1

Hordeum vulgare subsp. vulgare	Hv2 uncharacterized protein	BAJ94038.1
Hordeum vulgare subsp. vulgare	Hv4 uncharacterized protein	BAJ85075.1
Lens culinaris	LSAO Primary amine oxidase	P49252.3
Malus domestica	Md1 amine oxidase 1	AIS23644.1
Malus domestica	Md2 amine oxidase 2	AIS23645.1
Malus domestica	Md3 amine oxidase 3	AIS23646.1
Malus domestica	Md4 amine oxidase 4	AIS23647.1
Malus domestica	Md5 amine oxidase 5	AIS23648.1
Medicago truncatula	Mt1 primary amine oxidase	XP_003592404.1
Medicago truncatula	Mt4 copper methylamine oxidase	XP_003601195.1
Medicago truncatula	Mt3 primary amine oxidase	XP_003601419.1
Medicago truncatula	Mt5 copper methylamine oxidase	XP_003613133.2
Medicago truncatula	Mt2 primary amine oxidase	XP_013455559.1
Nicotiana sylvestris	Ns AOC copper amine oxidase	XP_009778427.1
Nicotiana tabacum	NtMPO *N*-methylputrescine oxidase	BAF49520.1
Nicotiana tabacum	Nt2 copper-containing amine oxidase	AIE54293.1
Oryza sativa subsp. japonica	Os3 amine oxidase	CAD39884.2
Oryza sativa subsp. japonica	Os7 amine oxidase	EAZ31082.1
Oryza sativa subsp. Japonica	Os2 hypothetical protein	EEC80539.1
Oryza sativa subsp. Japonica	Os1 hypothetical protein	EEC82302.1
Pinus sylvestris	Psy1 putative copper-containing diamine oxidase	ADQ37305.1
Pisum sativum	PSAO Primary amine oxidase	Q43077.1
Populus trichocarpa	Pt1 primary amine oxidase	XP_002312527.2
Populus trichocarpa	Pt3 amine oxidase	XP_002314704.2
Populus trichocarpa	Pt5 hypothetical protein	XP_002322194.2
Rc-Ricinus communis	Rc2 Copper Amine oxidase	EEF45396.1
Ricinus communis	Rc5 Primary amine oxidase	XP_002509596.1
Ricinus communis	Rc4 Primary amine oxidase	XP_002509597.1
Ricinus communis	Rc6 copper methylamine oxidase	XP_002511334.1
Ricinus communis	Rc1 Primary amine oxidase	XP_002516777.1
Ricinus communis	Rc3 Primary amine oxidase	XP_002516781.1
Ricinus communis	Rc7 uncharacterized protein	XP_002527922.1
Solanum lycopersicum	Sl4 uncharacterized protein	XP_004239124.1
Solanum lycopersicum	Sl2 uncharacterized protein	XP_004244763.1
Solanum lycopersicum	Sl3 uncharacterized protein	XP_010322539.1
Solanum lycopersicum	Sl1 copper amine oxidase	NP_001296994.1
Sorghum bicolor	Sb2 Primary amine oxidase	XP_002460889.1
Sorghum bicolor	Sb1 Primary amine oxidase	XP_002452575.1
Sorghum bicolor	Sb3 Primary amine oxidase	XP_002446210.1
Sorghum bicolor	Sb4 uncharacterized protein	XP_002448036.1
Triticum urartu	Hv5 primary amine oxidase	EMS63127.1
Vitis vinifera	Vv6 Primary amine oxidase	XP_002278244.1
Vitis vinifera	Vv7 Primary amine oxidase	XP_002278327.1
Vitis vinifera	Vv5 uncharacterized protein	XP_002277961.1

Vitis vinifera	Vv8 unnamed protein product	CBI26238.3
Vitis vinifera	Vv4 unnamed protein product	CBI34761.3
Vitis vinifera	Vv2 uncharacterized protein	XP_002273532.2
Vitis vinifera	Vv3 primary amine oxidase	XP_003635614.2
Zea mays	Zm1 Primary amine oxidase	NP_001145964.1
Zea mays	Zm2 uncharacterized protein	NP_001169559.1

8.1.6 SABATH Methyltransferases Sequence Accession for Phylogenetic Analysis

Table S4: This is the key for the abbreviations of SABATH-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. Published in Chavez et al., 2022; Supplemental Dataset; Table S1.

Species Name	Description/Predicted function	Accession Number
Antirrhinum majus	S-adenosyl-L-methionine: salicylic acid methyltransferase (AmSAMT)	AAN40745.1
Antirrhinum majus	SAM: benzoic acid carboxyl methyltransferase (AmBAMT)	AAF98284.1
Arabidopsis lyrata subsp. lyrata	methyl transferase (AlBSMT1)	AAP57211.1
Arabidopsis thaliana	Paraxanthine methyltransferase 1 (AtPXMT1)	Q6XMI3
Arabidopsis thaliana	Farnesoic acid carboxyl-O-methyltransferase (AtFAMT)	Q9C9M3
Arabidopsis thaliana	Indole-3-acetate O-methyltransferase 1 (AtIAMT1)	Q9FLN8
Arabidopsis thaliana	Paraxanthine methyltransferase 2 (AtPXMT2)	Q9C9M2
Arabidopsis thaliana	Gibberellic acid methyltransferase 1 (AtGAMT1)	F4JUY5
Arabidopsis thaliana	Gibberellic acid methyltransferase 2 (AtGAMT2)	Q5XF78
Arabidopsis thaliana	Paraxanthine methyltransferase 3 (AtPXMT3)	Q9C9M4
Arabidopsis thaliana	S-adenosyl-L-methionine: jasmonic acid carboxyl methyltransferase (AtJAMT)	AAG23343.1
Arabidopsis thaliana	methyl transferase (AtBSMT1)	AAP57210.1
Atropa belladonna	S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase (AbSAMT)	BAB39396.1
Clarkia breweri	Salicylate carboxyl methyltransferase (CbSAMT)	Q9SPV4
Erythroxylum coca	EcSABTH1 MPOB methyltransferase (renamed to EcMPOBMT)	OP382847
Erythroxylum coca	EcSABATH2	n/a
Erythroxylum coca	EcSABATH3	n/a
Oryza sativa subsp. japonica	Indole-3-acetate O-methyltransferase 1 (OsIAMT1)	Q0J998
Zea mays	Anthranilate O-methyltransferase 3 (ZmAAMT3)	D9J100
Zea mays	Anthranilate O-methyltransferase 1 (ZmAAMT1)	D9J0Z7
Zea mays	Anthranilate O-methyltransferase 2 (ZmAAMT2)	B6SU46
Zea mays	Benzoate O-methyltransferase (ZmOMT8)	D9J101
Zea mays	Inactive anthranilate O-methyltransferase 1 (ZmAAMT1I)	D9J0Z8

8.1.7 P450 Cyclases Sequence Accession for Phylogenetic Analysis

Table S5: A key for the abbreviations of Cytochrome P450-related enzymes in the phylogenetic tree and their NCBI or Uniprot accession number. Published in Chavez et al., 2022; Supplemental Dataset; Table S1.

Species Name	Description/Predicted function	Accession Number
Arabidopsis thaliana	CYP81F2	Q9LVD6
Arabidopsis thaliana	CYP81F4	Q9SZU1
Arabidopsis thaliana	CYP81F2	O65790
Arabidopsis thaliana	CYP81F3	Q0WTF4
Arabidopsis thaliana	CYP81D11	Q9LHA1
Arabidopsis thaliana	CYP81D1	Q9FG65
Atropa belladonna	AbCYP82M3	A0A3G4RHY7
Cicer arietinum	CICAR Cytochrome P450	Q9ZRW6
Cicer arietinum	CICAR Cytochrome P450	Q9XFX1
Echinochloa phyllopogon	Cytochrome P450	A0A024FAY6
Echinochloa phyllopogon	Cytochrome P450	A0A024FAY2
Echinochloa phyllopogon	Cytochrome P450	R4WHE6
Erythroxylum coca	EcCYP81AN15 (Methylecgonone cyclase)	OP382847
Erythroxylum coca	CYP79D62	AOW44274.1
Erythroxylum coca	CYP79D63	AOW44272.1
Erythroxylum fischeri	CYP79D60	AOW44273.1
Erythroxylum fischeri	CYP79D61	AOW44271.1
Glycyrrhiza echinata	Isoflavone 2'-hydroxylase	P93147
Helianthus tuberosus	HELTU Cytochrome P450	Q7FPQ4
Jatropha curcas	CYP81Q32	XP_012079324.1
Manihot esculenta	cytochrome P450 81E8-like	XP_021632242.1
Manihot esculenta	isoflavone 2'-hydroxylase-like	XP_021631824.1
Medicago truncatula	Isoflavone 2'-hydroxylase	Q6WNR0
Medicago truncatula	Isoflavone 3'-hydroxylase	Q6WNQ9
Medicago truncatula	CYP81E8	Q6WNQ8
Prunus dulcis	cytochrome P450 81Q32-like	XP_034218022.1
Prunus persica	CYP81E8	XP_007206553.1
Prunus yedoensis var. nudiflora	CYP81E8	PQQ16710.1
Zea mays	CYP81N5	B6SZ21
Zea mays	CYP81N4	B6SSR6
Zea mays	CYP81L6	B6SV20

8.1.8 Hierarchical clustering heatmap showing expression profiles for cyclase

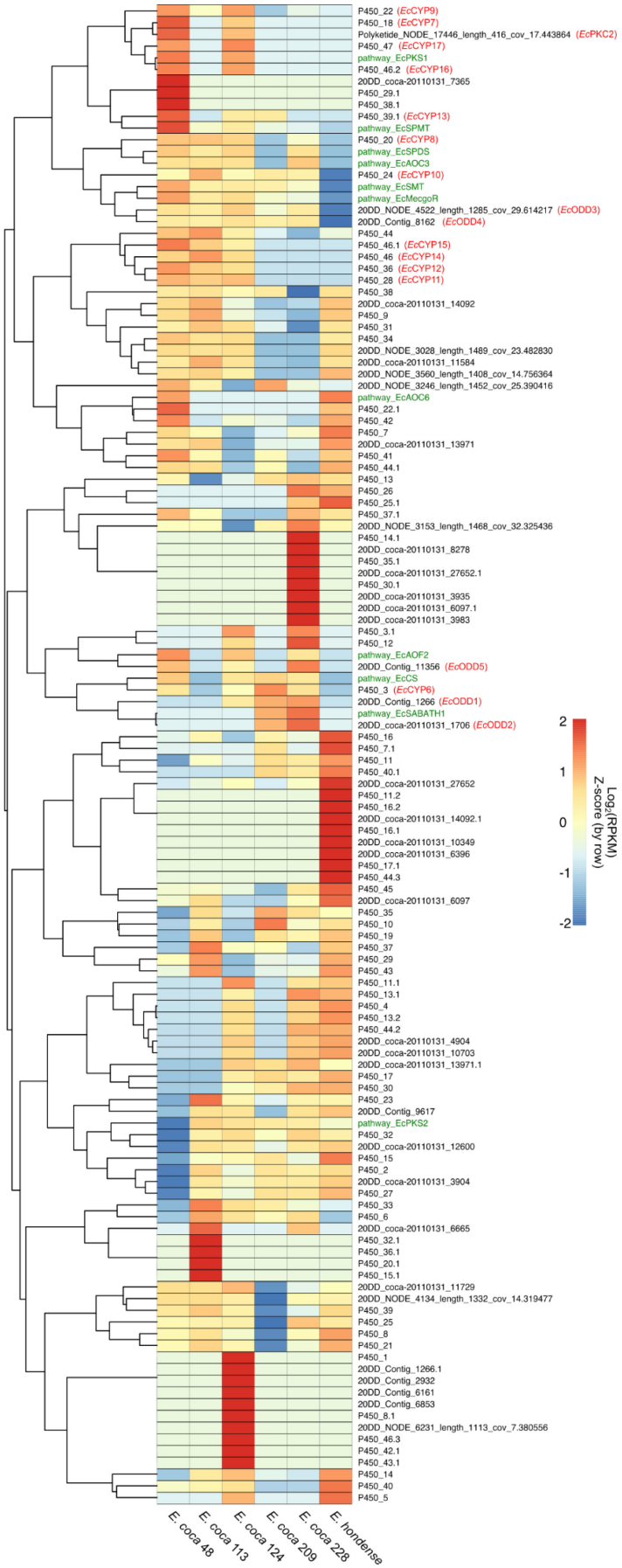

Figure S6: Hierarchical clustering heatmap showing expression profiles for cyclase (methylecgonone synthase) gene candidates (vertical axis) identified from Erythroxylaceae transcriptomes across three *E. coca* plants (48, 113, 124), two *E. novogranatense* plants ("*E. novo*" 209, 228), and a low-cocaine-producing *E. hondense* control (horizontal axis). Transcript expression is scaled by row using a normal distribution (z-score). The dendrogram indicates the hierarchical clustering of candidates by expression profile across *E. coca*, *E. novogranatense*, and *E. hondense*. Color scheme for gene IDs: green, known TA pathway genes; red, putative methylecgonone synthase candidates selected for screening in yeast. Published in Chavez et al., 2022; Page 8; Figure 4B.

8.1.9 Primers for Cloning Synthetic DNA into pHREAC Vector

Primer Name	Sequence 5' to 3'	Purpose of Primer
EcSPDS pHREAC FWD	AGGTAGGGTCTCcaaaaatgggtgaagccgttcaa	Cloning EcSPDS ORF into pHREAC expression vector
EcSPDS pHREAC REV	AGGTAGGGTCTCcagcgttagatggcaatttcacca	
EcSMT pHREAC FWD	CGCAAAGGTCTCgaaaaatggctagagacaacgtc	Cloning EcSMT ORF into pHREAC expression vector
EcSMT pHREAC REV	CGCAAAGGTCTCgagcgttagtacttcaacagcaagc	
EcAOF1 pHREAC FWD	GAGTCAGGTCTCgaaaaatgaagaagttgggcgtt	Cloning EcAOF1 ORF into pHREAC expression vector
EcAOF1 pHREAC REV	GAGTCAGGTCTCgagcgttacaactcttgattgacg	
EcAOC1 pHREAC FWD	CCATAGGGTCTCgaaaaatggctactgctcaagaa	Cloning EcAOC1 ORF into pHREAC expression vector
EcAOC1 pHREAC REV	CCATAGGGTCTCgagcgttacagcttagacaacaaac	
EcAOC2 pHREAC FWD	GTAGTGGGTCTCcaaaaatggcttctgcttctcaa	Cloning EcAOC2 ORF into pHREAC expression vector
EcAOC2 pHREAC REV	GTAGTGGGTCTCcagcgttacagtttggaggccaa	
EcCYP81AN15 pHREAC FWD	TGGTCAGGTCTCgaaaaatggttgatactgtcttgt	Cloning EcCYP81AN15 ORF into pHREAC expression vector
EcCYP81AN15 pHREAC REV	TGGTCAGGTCTCgagcgttagtcgaccatatccaaa	

Table S6: Primer sequences for cloning into pHREAC vector via Golden Gate Cloning. Bases in all capital letters and underlined represent the restriction enzyme recognition site of BsaI. The bases in color represent the start codon (green) and stop codon (red). The four bases between the recognition site and the start/stop codon represent the overhang generated for scarless cloning into the pHREAC vector.

8.1.10 Nanodrop Quantification of RNA Samples from *Erythroxylum coca*

Sample Name	Concentration in ng/µL	A260/A280 ratio	A260/A230 ratio
cocaL1_B1	132.2	2.13	2.26
cocaL1_B2	199.6	2.08	1.59
cocaL1_B3	377.4	2.16	2.32
cocaL2_B1	189.5	2.14	2.20
cocaL2_B2	75.9	2.13	2.08
cocaL2_B3	89.7	2.12	2.02
cocaL3_B1 RA1	414.8	2.16	2.10
cocaL3_B1 RAP	107.6	2.12	2.27
cocaL3_B2	102.5	2.11	1.98
cocaL3_B3	185.6	2.14	2.21

Table S7: Nanodrop readings of RNA quality from *E. coca* leaf tissues. L1 is rolled leaves, L2 is unrolled leaves, and L3 is mature leaves. The RA1 and RAP addition to the sample names are different lysis buffers offered in the NucleoSpin RNA Plant mini kit for different plant tissues.

Sample Name	RNA Concentration in ng/µL	300ng of RNA for RT reaction in µL	Water Added in µL
L1 B1	185	1.62	9.38
L1 B2	243	1.23	9.77
L1 B3	673	0.45	10.55
L2 B1	271	1.11	9.89
L2 B2	98	3.06	7.94
L2 B3	119	2.52	8.48
L3 B1	736	0.41	10.59
L3 B2	128	2.34	8.66
L3 B3	276	1.09	9.91

Table S8: Preparation of RNA samples from *E. coca* for cDNA synthesis via Reverse Transcription using either LunaScript Supermix Reverse Transcriptase or SuperScript IV Reverse Transcriptase protocols. RNA concentrations are based on readings from Agilent Bioanalyzer.

8.1.11 Bioanalyzer Quantification of RNA Samples from *Erythroxylum coca*

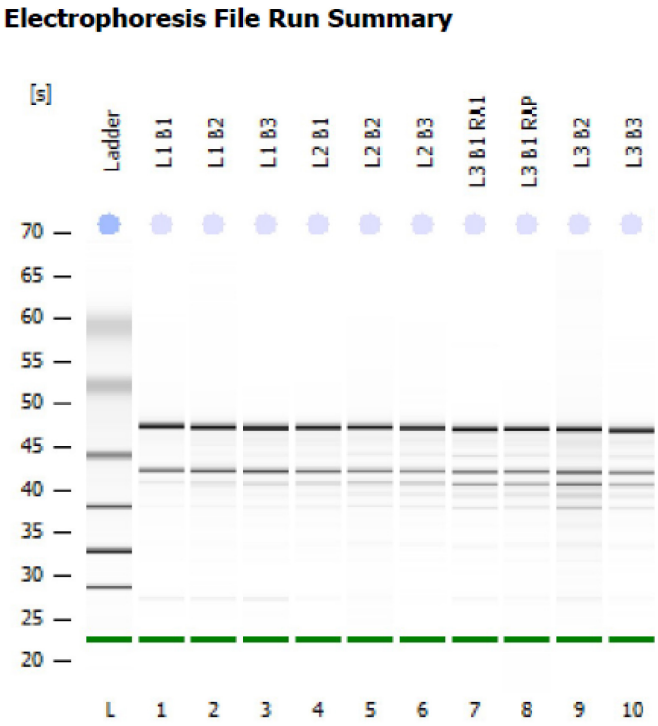

Figure S7: An electrophoresis run of RNA isolated from *Erythroxylum coca* leaf tissues L1, L2, and L3 leaf stages.

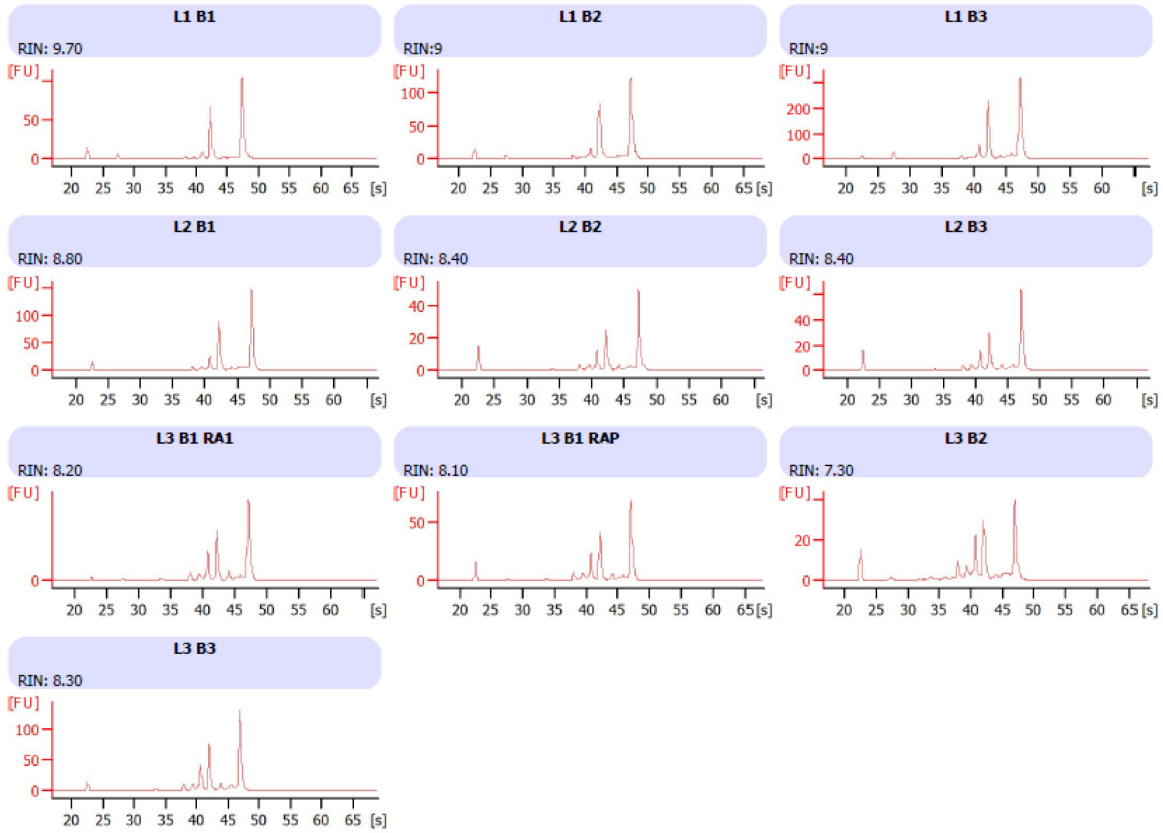

Figure S8: Agilent Bioanalyzer assessment of RNA quality from *E. coca* leaf tissues. L1 is rolled young leaves, L2 is unrolled young leaves, and L3 is mature leaves.

Sample Name	Concentration in ng/µL	RIN value
L1 B1	185	9.70
L1 B2	243	9.0
L1 B3	673	9.0
L2 B1	271	8.80
L2 B2	98	8.40
L2 B3	119	8.40
L3 B1 RA1	736	8.20
L3 B1 RAP	136	8.10
L3 B2	128	7.30
L3 B3	276	8.30

Table S9: Agilent Bioanalyzer assessment of RNA quality from *E. coca* leaf tissues. L1 is rolled young leaves, L2 is unrolled young leaves, and L3 is mature leaves.

8.1.12 qPCR Primer Table

qPCR Primer Name	Primer Sequence in 5' to 3' Direction	Amplicon Size in bp
EcAOF1 qPCR FWD1	TGAGTTCGTTGACCATGGAGAG	101 bp
EcAOF1 qPCR REV1	CCACTGATGAGAGGAACTGCTT	
EcAOC1 qPCR FWD1	GCTTCAATACCACAACCACCAC	79 bp
EcAOC1 qPCR REV1	CTCCAATTCTGAACAGCCCTCT	
EcAOC1 qPCR FWD3	AGAGGGCTGTTCAGAATTGGAG	106 bp
EcAOC1 qPCR REV3	GACTCAACAGGATGGATCAGGG	
EcAOC2 qPCR FWD2	AGTGGAACGCATTGGTTTTACG	101 bp
EcAOC2 qPCR REV2	TGTCCTTGGTATCCAATTCGCA	
EcAOC2 qPCR FWD3	TGCGTCTGTTTGCATGAAGAAG	84 bp
EcAOC2 qPCR REV3	AGATCTGCGTACTTCTGCCAAA	
EcMPOBMT qPCR FWD	TTTTCCCTAGCAACAGTCTGCAT	80 bp
EcMPOBMT qPCR REV	CTCCAGCCCTTCAGGAACCT	
EcCYP81AN15 qPCR FWD1	AAGACCGGTTAGTGGATGAACC	110 bp
EcCYP81AN15 qPCR REV1	CATGTGTGGATCAAGAAGGGGA	
EcMecgoR qPCR FWD	CAACTATACCTCCTGCCGTCAATC	(Jirschitzka et al., 2012)
EcMecgoR qPCR REV	GCTCAATCTTTCGCCCGTCTTC	
Ec6409 qPCR FWD	GAAGAGACAAGTGGTGGGGTGAG	(Docimo et al., 2013)
Ec6409 qPCR REV	AGAAGAGAGCAAAGAGGAAGAGTGG	
Ec10131 qPCR FWD	TGGAAGGGTAGTGGGGTAACAATG	(Docimo et al., 2013)
Ec10131 qPCR REV	GAGCGTAGTCGTCAGAGAAGGC	

Table S10: Primers used to optimize PCR of cDNA targets in *E. coca*.

8.1.13 Mass fragment transitions and parameters for compound identification by LC-MS/MS

Compound	MRM transition (m/z [M+H]+)	Collision energy
Putrescine	89.1 → 72.1	12
Spermidine	146.1 → 72.1	15
***N*-methylspermidine**	160 → 72.1	17
***N*-methylputrescine**	103.1 → 72.1	9
Spermine	203.2 → 112.1	17
***S*-adenosyl-L-methionine**	399.1→ 250.1	18
***S*-adenosyl-L-homocysteine**	385.1→136.1	20
***N*-methylpyrrolinium**	84 → 57	35
D_3 *N*-methylpyrrolinium	87 → 60	20
Hygrine*	142 → 84	13
MPOB	186 → 84	25
MPMOB	200 → 84	17
Methylecgonone	198 →166	21
Methylecgonine	200 → 182	21

Table S11: Multiple Reaction Monitoring (MRM) transition list for LC-MS/MS detection.*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. For experiments in which NMPy was not a key/terminal metabolite, NMPy was analyzed directly.

9 CURRICULUM VITAE

Personal Information

Name: Benjamin Gabriel Chavez

Reference: 0000-0001-5720-1856

Education

IPK-Gatersleben September 2019-to date Ag. Metabolic Diversity	PhD topic "Spermidine *N*-methyltransferases TA biosynthesis in *E. coca*
Texas Tech University September 2018-July 2019	Started PhD Program at Texas Tech University Left Graduate program and started PhD program at IPK-Gatersleben & University of Halle
South Plains College/ Texas Tech University August 2013-May 2018	B.Sc in Cell & Molecular Biology

Peer reviewed publications

Chavez, B. G., Srinivasan, P., Glockzin, K., Kim, N., Montero Estrada, O., Jirschitzka, J., Rowden, G., Shao, J., Meinhardt, L., Smolke, C. D., & D'Auria, J. C. (2022). Elucidation of tropane alkaloid biosynthesis in *Erythroxylum coca* using a microbial pathway discovery platform. Proceedings of the National Academy of Sciences, 119(49), e2215372119. https://doi.org/10.1073/pnas.2215372119.

Restrepo, David A., Ernesto Saenz, Orlando Adolfo Jara-Muñoz, Iván F. Calixto-Botía, Sioly Rodríguez-Suárez, Pablo Zuleta, Benjamin G. Chavez, Juan A. Sanchez, and John C. D'Auria. 2019. "Erythroxylum in Focus: An Interdisciplinary Review of an Overlooked Genus." *Molecules* 24 (20): 3788. https://doi.org/10.3390/molecules24203788.

Irfan, Mohammad, Benjamin Chavez, Paride Rizzo, John C. D'Auria, and Gaurav D. Moghe. 2021. "Evolution-Aided Engineering of Plant Specialized Metabolism." *aBIOTECH* 2 (3): 240–63. https://doi.org/10.1007/s42994-021-00052-3.

Kim, Neill, Benjamin Chavez, Charles Stewart, and John C. D'Auria. 2021. "Structure and Function of Enzymes Involved in the Biosynthesis of Tropane Alkaloids." In *Tropane Alkaloids: Pathways, Potential and Biotechnological Applications*, edited by Vikas Srivastava, Shakti Mehrotra, and Sonal Mishra, 21–50. Singapore: Springer. https://doi.org/10.1007/978-981-33-4535-5_2.

Kim, N., O. Estrada, B. Chavez, C. Stewart, and J. C. D'Auria. 2016. "Tropane and Granatane Alkaloid Biosynthesis: A Systematic Analysis." *Molecules* 21 (11). https://doi.org/10.3390/molecules21111510.

Oral Presentation

Benjamin Chavez, Prashanth Srinivasan, Christina Smolke, and John D'Auria. October 15th, 2023. *N*-methylspermidine is the Start of Tropane Alkaloid Biosynthesis in *Erythroxylum coca*. Postgraduate Workshop of the "Natural Products" section of the German Society for Plant Sciences (DBG) Bad Endbach, Germany.

Benjamin Chavez, Prashanth Srinivasan, Christina Smolke, and John D'Auria. August 30th 2022. Identification of Unknown Steps of Tropane Alkaloid Biosynthesis in *Erythroxylum coca*. Die Botanik-Tagung 2022, International Conference of the German Society for Plant Sciences. Bonn, Germany.

Poster Presentations

Benjamin Chavez, Prashanth Srinivasan, Christina Smolke, and John D'Auria. July 9-July 13, 2022. Identification of Unknown Steps of Tropane Alkaloid Biosynthesis in *Erythroxylum coca* using Synthetic Biology. ASPB 2022, Annual Meeting of the American Society of Plant Biologists (ASPB). Portland, Oregon, USA

Benjamin Chavez, Kayla Anderson, and John D'Auria. December 10th-11th 2019. Moonlighting Activity of Spermidine Synthase in *Erythroxylum coca*. ASPB 2022, Symposium on Small Molecules in Plant Research: Chemistry and Biology Come Together 2019. Valencia Spain

*The underlined author was presenting the work

Benjamin Gabriel Chavez

Gatersleben, den 06.03.2024

10 DECLARATION (ERLÄRUNG)

Eidesstattliche Erklärung / Declaration under Oath

Ich, Benjamin Gabriel Chavez, erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I, Benjamin Gabriel Chavez, declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Datum / Date	Unterschrift des Antragstellers / Signature of the applicant

11 ACKNOWLEDGEMENTS

My gratitude for everyone who participated in my development as a PhD student cannot be understated. First, I must thank my primary research advisor, Dr. John Charles D'Auria, whose passion for biochemistry ignited my passion for science. I have been working with John for almost ten years, and I first joined his lab even before I took an undergraduate biology class. Nevertheless, John took me under his wing and taught me invaluable skills that helped me grow as a young scientist. The most important thing that John instilled within me is the thirst for curiosity and discovery, which was essential in our work to resolve complex missing steps in cocaine biosynthesis in *Erythroxylum coca*. I also thank Dr. Christine Smolke and Dr. Prashanth Srinivasan for their work on testing and validating our candidate genes inside their yeast synthetic biology platform at Stanford University. Their collaboration with the D'Auria lab led to the rapid discovery and reconstruction of the cocaine biosynthetic pathway. I would also like to thank previous members of the D'Auria lab, as their findings were invaluable in continuing to resolve the missing steps in tropane alkaloid biosynthesis. I want to thank Kayla Glockzin, Kyle Glockzin, Neill Kim, Olga Estrada, and Gage Rowden for their efforts. I would also like to thank my friends and colleagues at IPK-Gatersleben for their indispensable advice as I navigated my Ph.D. even after moving from Texas to Germany. I want to thank Paride Rizzo and Isa Mora-Ramirez for their unwavering support as I progressed through my Ph.D. Surround yourself with friends, and you will never work a day in your life. I would also like to thank the technical staff and gardeners at the IPK-Gatersleben campus, as their work was vital to the work presented in this thesis. I would also like to thank Dr. Mohammed-Reza Hajirezaei for assistance with LC-MS/MS measurements.