TROPANE ALKALOID TRANSPORTERS AND METHODS OF MAKING TROPANE ALKALOIDS USING THE SAME

GOVERNMENT RIGHTS

This invention was made with Government support under contract AT007886 awarded by the National Institutes of Health. The Government has certain rights in the invention.

CROSS-REFERENCE TO RELATED APPLICATIONS

Pursuant to 35 U.S.C. §119(e), this application claims priority to the filing date of the United States Provisional Patent Application Serial No. 63/166,117 filed March 25, 2021; the disclosure of which application is herein incorporated by reference.

INTRODUCTION

Tropane alkaloids (TAs) are a class of alkaloids (nitrogen-containing organic molecules) produced by certain plants of the nightshade family (*Solanaceae*). Owing to their anticholinergic (acetylcholine receptor-modulating) properties, TAs are of interest as neurological medicines. Several TAs, including atropine, hyoscyamine, and scopolamine, are classified as essential medicines by the World Health Organization for the treatment of neurological disorders such as organophosphate and nerve agent poisoning, gastrointestinal spasms, and cardiac arrhythmia, as well as to control salivary and mucous secretions during respiratory disease and to mitigate symptoms of Parkinson's disease. As such, an adequate and consistent supply of these TA molecules so that they are available to researchers and physicians is of interest. Current supply chains for medicinal TAs rely on extraction from unsustainable and geographically restricted plant monocultures, in which TAs accumulate to only 0.2-4% dry weight, and which are susceptible to pests, changes in land use, and climate. No total chemical syntheses for TAs from simple feedstocks have yet proven sufficiently economical for industrial use due to difficulties arising from TA stereochemistry. Moreover, poor economies of scale and long generation times have thus far rendered the engineering of transgenic plants or plant cultures with improved TA production an unviable strategy for sourcing these compounds. As such, methods for preparing TAs are of interest.

Biosynthesis platforms based on genetically engineered non-plant cells, such as microbes, are an effective strategy for synthesizing plant-derived natural products like TAs with reduced economic and environmental cost. Due to their genetic and metabolic

tractability, microbial platforms can expand access to both natural and non-natural TA derivatives with unique bioactivities that may otherwise be produced in trace quantities in native plants, or which may only be accessible through derivatization using hazardous chemicals. However, efforts to engineer non-plant cells for conversion of simple building blocks (sugars, amino acids) into complex plant molecules like TAs are hindered by low enzyme tolerance for non-native substrates or chemical transformations. Thus, new approaches for optimizing flux through engineered metabolic pathways are of interest for the efficient production of TAs and their derivatives. As natural metabolic pathways for plant alkaloids have evolved to take advantage of the unique biochemistries found in different tissues and organelles, specific methods and tools for recapitulating the spatial organization of these pathways in non-plant cells are of particular interest.

SUMMARY

This invention includes non-plant organisms engineered for the production of diverse tropane alkaloids (TAs) from precursors and sugar, and in which the movement of molecules is directed by one or more transporter proteins originating in other organisms. For example, included in this invention are engineered microbial strains containing such heterologous transporters for the production of medicinal TAs, which are hereby defined as naturally occurring TAs with established uses in current medical practice, including hyoscyamine, atropine, anisodamine, and scopolamine, and precursors and derivatives thereof. Also included in this invention are engineered microbial strains containing such heterologous transporters for the production of non-medicinal TAs, which are hereby defined as naturally occurring TAs without established uses in current medical practice but which may possess bioactivities of medicinal interest, including calystegines, cocaine, and precursors and derivatives thereof. This invention further includes engineered microbial strains containing such heterologous transporters for the production of non-natural TAs, which are hereby defined as TAs not produced by unmodified organisms, such as TAs produced via esterification of acyl donor and acyl acceptor compounds which are not esterified in naturally-occurring organisms, including derivatives of medicinal TAs and derivatives of non-medicinal TAs. An example of the schemes included in this invention are detailed in Figures 1-3.

Disclosed herein is an engineered fungal cell that produces a tropane alkaloid product, wherein the engineered fungal cell comprises a plurality of heterologous coding sequences encoding a plurality of enzymes within a pathway for producing the tropane alkaloid product; wherein the cell comprises a plurality of heterologous coding sequences encoding a plurality of transporter proteins selected from the group of purine uptake permease-like transporter, multidrug resistance transporter, pleiotropic drug resistance transporter, ATP-binding cassette transporter, multidrug and toxin extrusion transporter, nitrate transporter 1/peptide

transporter family transporter, and major facilitator superfamily transporter; and wherein the plurality of heterologous coding sequences encoding a plurality of transporter proteins comprises a heterologous coding sequence encoding a purine uptake permease-like transporter.

The invention encompasses methods of producing pseudotropine and alkaloids derived from pseudotropine, for example calystegines, using microorganisms engineered to express at least one heterologous enzyme and at least one heterologous transporter protein as microbial catalysts. This invention further includes methods of producing diverse compounds which can be used as acyl donors for the biosynthesis of TA scaffolds using microorganisms engineered to express at least one heterologous enzyme and at least one heterologous transporter protein as microbial catalysts. This invention also includes methods of esterifying acyl donors and acceptors for the production of TA scaffolds using microorganisms engineered to express at least one heterologous enzyme and at least one heterologous transporter protein as microbial catalysts. The invention further includes methods of modifying and culturing engineered microbial strains for the production of medicinal TAs such as hyoscyamine and scopolamine, non-medicinal TAs such as calystegines, and non-natural TAs such as those derived from esterification of tropine with acyl donor compounds other than 3-phenyllactic acid (PLA). The invention also encompasses methods of altering, enhancing, or modulating the movement of molecules involved in the production of medicinal TAs, non-medicinal TAs, and non-natural TAs in microorganisms by means of transporter proteins. The invention further includes methods of producing derivatives of medicinal TAs, non-medicinal TAs, and non-natural TAs using microorganisms engineered to express at least one heterologous transporter protein and at least one enzyme which chemically affixes or removes one or more functional groups to/from a TA molecule.

Disclosed herein is a method for producing a tropane alkaloid product comprising (a) culturing a cell as defined herein under conditions suitable for protein production; (b) adding a starting compound to the cell culture; and (c) recovering the tropane alkaloid product from the culture.

Host cells that are engineered to produce tropane alkaloids that are of interest, such as hyoscyamine and scopolamine, are provided. Host cells that are engineered to produce derivatives of TAs that are of interest, such as nortropane alkaloids and tropane *N*-oxides, are also provided. TAs of interest may include TA precursors, TAs, and modifications of TAs, including derivatives of TAs. The host cells may have one or more modifications selected from: a feedback inhibition alleviating mutation in an enzyme gene; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene; an inactivating mutation in an enzyme or transporter; a heterologous transporter protein whose

expression influences the movement of one or more TAs, TA precursors, or TA derivatives within or between cells; and any other heterologous coding sequence. Also provided are methods of producing a TA of interest using the host cells and compositions, e.g., kits, systems etc., that find use in methods of the invention.

An aspect of the invention provides a method for forming a product stream having a TA product. The method comprises providing engineered non-plant cells and a feedstock including nutrients and water to a batch reactor, in which engineered non-plant cells have at least one modification selected from the group consisting of: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; an inactivating mutation in an enzyme or transporter native to the cell; a heterologous transporter protein whose expression influences the movement of one or more TAs, TA precursors, or TA derivatives within or between cells; and any other heterologous coding sequence. Additionally, the method comprises, in the batch reactor, subjecting the engineered non-plant cells to fermentation by incubating the engineered non-plant cells for a time period of at least about 5 minutes to produce a solution comprising the TA product and cellular material. The method also comprises using at least one separation unit to separate the TA product from the cellular material to provide said product stream comprising the TA product.

In another aspect, the invention provides a method for forming a product stream having a TA product. The method comprises providing engineered non-plant cells and a feedstock including nutrients and water to a reactor. The method also comprises, in the reactor, subjecting the engineered non-plant cells to fermentation by incubating the engineered yeast cells for a time period of at least about 5 minutes (e.g., 5 minutes or longer) to produce a solution comprising cellular material and the TA product. Additionally, the method comprises using at least one separation unit to separate the TA product from the cellular material to provide the product stream comprising the TA product.

Another aspect of the invention provides an engineered non-plant cell that produces a tropane alkaloid (TA) product, the engineered non-plant cell having at least one modification selected from the group consisting of: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; and an inactivating mutation in an enzyme or transporter native to the cell. The engineered non-plant cell comprises at least one heterologous coding sequence encoding at least one enzyme that is selected from the group of arginine decarboxylase, agmatine ureohydrolase, agmatinase, putrescine N-methyltransferase, N-methylputrescine oxidase, pyrrolidine ketide synthase, tropinone synthase, cytochrome P450 reductase, tropinone reductase, phenylpyruvate reductase, 3-phenyllactic acid UDP-glucosyltransferase 84A27, littorine synthase, littorine mutase,

hyoscyamine dehydrogenase, hyoscyamine 6β-hydroxylase/dioxygenase, cocaine synthase, a metazoan hepatic cytochrome P450, senecionine *N*-oxygenase, a pyrrolizidine *N*-oxygenase, and a TA-detoxifying enzyme. In some examples, the engineered non-plant cell comprises a plurality of heterologous coding sequences encoding an enzyme that is selected from the group of arginine decarboxylase, agmatine ureohydrolase, agmatinase, putrescine N-methyltransferase, N-methylputrescine oxidase, pyrrolidine ketide synthase, tropinone synthase, cytochrome P450 reductase, tropinone reductase, phenylpyruvate reductase, 3-phenyllactic acid UDP-glucosyltransferase 84A27, littorine synthase, littorine mutase, hyoscyamine dehydrogenase, hyoscyamine 6β-hydroxylase/dioxygenase, cocaine synthase, a metazoan hepatic cytochrome P450, senecionine *N*-oxygenase, a pyrrolizidine *N*-oxygenase, and a TA-detoxifying enzyme. The engineered non-plant cell also comprises at least one heterologous coding sequence encoding at least one transporter that is selected from the group of an ATP-binding cassette (ABC) transporter, a multidrug and toxin extrusion (MATE) transporter, a purine uptake permease-like (PUP) transporter, a lactose permease-like (LP) transporter or other nitrate transporter 1/peptide transporter (NPF/NRT), and a major facilitator superfamily (MFS) transporter. In some examples, the engineered non-plant cell comprises a plurality of heterologous coding sequences encoding a transporter that is selected from the group of an ATP-binding cassette (ABC) transporter, a multidrug and toxin extrusion (MATE) transporter, a purine uptake permease-like (PUP) transporter, a lactose permease-like (LP) or other nitrate transporter 1/peptide transporter (NPF/NRT), and a major facilitator superfamily (MFS) transporter. In some examples, the heterologous coding sequences may be operably connected. Heterologous coding sequences that are operably connected may be within the same pathway of producing a particular TA product. In some examples, the engineered non-plant cell comprises one or more modifications to intracellular compartmentalization that is selected from the group including, but not limited to, modified intracellular trafficking of enzymes, modified intracellular localization of enzymes, and modified intracellular transport of metabolites.

In another aspect of the invention, a therapeutic agent is provided. The therapeutic agent comprises a tropane alkaloid product.

BRIEF DESCRIPTION OF THE FIGURES

The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings. It is emphasized that, according to common practice, the various features of the drawings are not to-scale. On the contrary, the dimensions of the various features are arbitrarily expanded or reduced for clarity. Included in the drawings are the following figures.

Figure 1 illustrates an exemplary biosynthetic scheme for converting L-arginine to non-medicinal TAs**.** ADC, arginine decarboxylase; ARG, arginase; AUH, agmatine ureohydrolase; ODC, ornithine decarboxylase; PAO, polyamine oxidase; PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase; spont., spontaneous (non-enzymatic) step; PYKS, pyrrolidine ketide synthase; CYP82M3, tropinone synthase; CPR, cytochrome P450-$NADP^+$ reductase; TR2, tropinone reductase 2; P450, cytochrome P450. Arginine, ornithine, spermine, spermidine, and putrescine are naturally synthesized in yeast. All other metabolites shown are not naturally produced in yeast. The final products, indicated inside the box, are examples of non-medicinal TAs.

Figure 2 illustrates an exemplary biosynthetic pathway by which amino acids can be converted to medicinal TA molecules of interest and precursor molecules thereof. This example shows the conversion of L-arginine and L-phenylalanine to medicinal TAs. ADC, arginine decarboxylase; ARG, arginase; AUH, agmatine ureohydrolase; ODC, ornithine decarboxylase; PAO, polyamine oxidase; PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase; spont., spontaneous (non-enzymatic) step; PYKS, pyrrolidine ketide synthase; CYP82M3, tropinone synthase; CPR, cytochrome P450-NADP+ reductase; TR1, tropinone reductase 1; ArAT, aromatic aminotransferase; PPR, phenylpyruvate reductase; UGT84A27, 3-phenyllactate UDP-glucosyltransferase; LS, littorine synthase; CYP80F1, littorine mutase; HDH, (S)-hyoscyamine dehydrogenase; H6H, (S)-hyoscyamine 6β-hydroxylase/dioxygenase. Arginine, ornithine, spermine, spermidine, putrescine, phenylalanine, 3-phenylpyruvic acid, and trace amounts of 3-phenyllactic acid are naturally synthesized in yeast. All other metabolites shown are not naturally produced in yeast. The final products, indicated inside the box, are examples of medicinal TAs.

Figure 3 illustrates an exemplary biosynthetic pathway by which amino acids can be converted to a non-natural TA and precursor molecules thereof. In this example, L-arginine and L-phenylalanine are converted to non-natural TAs. ADC, arginine decarboxylase; ARG, arginase; AUH, agmatine ureohydrolase; ODC, ornithine decarboxylase; PAO, polyamine oxidase; PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase; spont., spontaneous (non-enzymatic) step; PYKS, pyrrolidine ketide synthase; CYP82M3, tropinone synthase; CPR, cytochrome P450-$NADP^+$ reductase; TR1, tropinone reductase 1; PAL, phenylalanine ammonia-lyase; 4CL, 4-coumarate-CoA ligase; CS, cocaine synthase. Arginine, ornithine, spermine, spermidine, putrescine, and phenylalanine are naturally synthesized in yeast. All other metabolites shown are not naturally produced in yeast. The final product, indicated inside the box, is an example of a non-natural TAs.

Figure 4 illustrates exemplary biosynthetic pathways for the production of putrescine from amino acids and other polyamine molecules. This figure shows how endogenous yeast and heterologous biosynthetic pathways can be used to make putrescine from central metabolites.

Figure 5 shows that yeast strains engineered for overexpression of endogenous biosynthetic enzymes involved in arginine and polyamine metabolism can produce putrescine in liquid culture. Additional copies of native genes were expressed from low-copy plasmids in wild-type yeast (CEN.PK2). Transformed strains were cultured in selective media with 2% dextrose at 30 ºC for 48 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001. Unless otherwise indicated, statistical significance is shown relative to the corresponding control (i.e., CEN.PK2).

Figure 6 shows that yeast strains engineered for heterologous expression of biosynthetic enzymes from organisms other than yeast that are involved in arginine and polyamine metabolism can produce putrescine production in liquid culture. In this example, the yeast strains are engineered to express a heterologous biosynthetic pathway from plants and bacteria. Heterologous enzymes were expressed from low-copy plasmids in wild-type yeast. Transformed strains were cultured in selective media with 2% dextrose at 30 ºC for 48 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001. Unless otherwise indicated, statistical significance is shown relative to the corresponding control (i.e., CEN.PK2).

Figure 7 shows that yeast strains engineered for heterologous expression of biosynthetic enzymes involved in arginine and polyamine metabolism from organisms other than yeast can produce TA precursors and intermediates agmatine, N-carbamoylputrescine, and putrescine in liquid culture. This figure shows the functional validation of agmatine/putrescine biosynthetic pathway genes in yeast. Wild-type yeast strain CEN.PK2 was transformed with three low-copy plasmids to co-express between zero (negative control) and three of the indicated biosynthetic genes. Plasmids expressing blue fluorescent protein (BFP) were used as negative controls for each of the three auxotrophic selection markers *URA3*, *TRP1*, and *LEU2*. Transformed strains were cultured in selective media with 2% dextrose at 30 ºC for 48 h prior to LC-MS/MS analysis of metabolite production. All data show titers as measured by LC-MS/MS peak area relative to the negative control (CEN.PK2). Data represent the mean of at three biological replicates and error bars show

standard deviation.

Figure 8 illustrates the endogenous regulatory pathways that tightly control intracellular putrescine levels during normal yeast growth.

Figure 9 shows a heat map of putrescine production in yeast strains with disruptions to endogenous polyamine biosynthesis regulatory mechanisms. For overexpression of native or heterologous putrescine pathways, indicated genes were expressed from low-copy plasmids in wild-type yeast (WT) or each single disruption strain. Strains were cultured in selective (YNB-DO) media with 2% dextrose at 30 °C for 72 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates. This figure shows that yeast strains that have single disruptions of polyamine metabolism genes and overexpressed endogenous or heterologous putrescine biosynthetic pathways can produce putrescine in liquid culture.

Figure 10 provides a summary of engineering efforts for increasing putrescine production in yeast. '+' symbol indicates expression of at least one gene from the pathway, whereas '–' indicates expression of no genes from the pathway. Strains were cultured in selective media with 2% dextrose at 30 °C for 48 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Unless otherwise indicated, statistical significance is shown relative to the corresponding control (i.e., CEN.PK2).

Figure 11 shows LC-MS/MS chromatograms which illustrate the stepwise conversion of putrescine to the TA intermediate NMPy and the side product 4MAB acid via the intermediates NMP and 4MAB in engineered yeast, in accordance with embodiments of the invention. The proposed mechanism for formation of the 4MAB acid side product via activity of an endogenous yeast enzyme (ALD) is shown. Extracted ion chromatogram MRM traces are shown for each metabolite along the pathway and for authentic standards using the highest precursor ion / product ion transition for each metabolite. Control represents strain CSY1235 (see Example 1.5) expressing *SPE1*, *AsADC*, and *speB* on a low-copy plasmid. Chromatogram traces are representative of three biological replicates. Enzyme symbols: PMT, putrescine N-methyltransferase; MPO, N-methylputrescine oxidase; ALD, aldehyde dehydrogenase.

Figure 12 shows LC-MS/MS chromatograms illustrating relative production of the TA

precursors (A) putrescine, (B) NMP, (C,E) 4MAB, and (D,F) NMPy in liquid cultures of engineered yeast expressing AbPMT1 and an MPO enzyme, in accordance with embodiments of the invention. (A) MRM chromatogram of putrescine (m/z+ 89 → 72) for CSY1235 harboring pCS4239 for putrescine overproduction. (B) MRM chromatogram of NMP (m/z+ 103 → 72) for CSY1235 harboring pCS4239 and expressing AbPMT1 from a low-copy plasmid. (C,D) MRM chromatograms of 4MAB (m/z+ 102 → 71) and NMPy (m/z+ 84 → 57), respectively, for CSY1235 harboring pCS4239 and expressing AbPMT1 and NtMPO1 from low-copy plasmids. (E,F) MRM chromatograms of 4MAB (m/z+ 102 → 71) and NMPy (m/z+ 84 → 57), respectively, for CSY1235 harboring pCS4239 and expressing AbPMT1 and DmMPO1ΔC-PTS1 from low-copy plasmids. Y-axes of traces are raw MRM ion counts. All chromatograms were generated by LC-MS/MS analysis of the extracellular medium after 48 hours of growth at 30°C in selective media with 2% dextrose. Traces are representative of at least three biological replicates.

Figure 13 shows the effect of *MEU1* disruption on SAM-dependent putrescine N-methylation by AbPMT1. Wild-type strain CEN.PK2 or *meu1* disruption strain CSY1229 were co-transformed with low-copy plasmids expressing *SPE1*, *AsADC*, and *speB* and *AbPMT1*. Data indicate mean NMP titer relative to CEN.PK2 control as quantified by LC-MS/MS peak area for three biological replicates after 48 hours of growth at 30 °C in selective media with 2% dextrose. Error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 14 shows an in silico prediction of subcellular localization for NMPy biosynthetic genes in plant and yeast/fungal cells using the SherLoc2 web server. Values and coloring indicate probability scores (0 to 1) for localization to each compartment: CYT, cytosol; NUC, nucleus; VAC, vacuole; CHL, chloroplast; MIT, mitochondria; POX, peroxisome.

Figure 15 illustrates (A) colocalization of N- and C-terminal GFP-tagged NtMPO1 with a PEX3 peroxisomal marker and (B) the effect of N- and C-terminal GFP tagging of NtMPO1 on the production of the TA precursors 4MAB and NMPy in liquid cultures of engineered yeast, in accordance with embodiments of the invention. This figures shows an experimental validation of NtMPO1 subcellular localization. (A) Fluorescence microscopy of NtMPO1 N- and C-terminal GFP fusions co-expressed with peroxisome marker mCherry-PEX3 in wild-type yeast (CEN.PK2). White arrows indicate colocalization of GFP-tagged NtMPO1 with peroxisomes. Scale bar, 10 μm. (B) Effect of forcing cytosolic localization of NtMPO1 on 4MAB or NMPy production. Wild-type yeast (CEN.PK2) was co-transformed

with low-copy plasmids expressing wild-type *NtMPO1* or N- or C-terminal GFP fusions together with low-copy plasmids expressing *SPE1*, *AsADC*, and *speB* and *AbPMT1*. LC-MS/MS analysis was performed after 48 hours of growth at 30 °C in selective media with 2% dextrose. Data represent mean of three biological replicates; error bars show standard deviation. Most probable sub-cellular compartment is indicated based on microscopy data in (a).

Figure 16 provides fluorescence microscopy data depicting the sub-cellular localization of AbPMT1 and NtMPO1 when expressed heterologously in yeast. . Microscopy was performed on wild-type yeast expressing N- or C-terminal GFP-tagged AbPMT1 or NtMPO1 from low-copy plasmids. Scale bar, 10 μm.

Figure 17 illustrates (A) a sequence alignment of *NtMPO1* and the putative MPO enzymes *AbMPO1* and *DmMPO1* identified from plant transcriptome data, (B) a comparison of the production of the TA precursors 4MAB and NMPy in liquid cultures of engineered yeast strains expressing NtMPO1, AbMPO1, or DmMPO1, and (C) a comparison of the predicted three-dimensional structures of NtMPO1, AbMPO1, and DmMPO1 determined from homology modeling, in accordance with embodiments of the invention. (A) Alignment of query NtMPO1 sequence against AbMPO1 and DmMPO1 candidates from 1000 Plants Project database. Blue indicates conservation of amino acid structure; red indicates mismatches. (B) Comparison of relative activities of MPO orthologs. Putrescine overproducing strain CSY1235 (see Example 1.5) was co-transformed with low-copy plasmids expressing *SPE1*, *AsADC*, and *speB*, *AbPMT1*, and one of the three MPO variants. LC-MS/MS analysis was performed after 48 hours of growth in selective media at 30 °C. Data represent mean of three biological replicates; error bars show standard deviation. (C) Homology models of MPO enzymes (pink) constructed based on the crystal structure of Pisum sativum copper-containing amino oxidase (PDB: 1KSI, blue) using the RaptorX web server. Top: NtMPO1; center: AbMPO1; bottom: DmMPO1.

Figure 18 illustrates 4MAB production in liquid culture of engineered yeast strains overproducing putrescine and expressing AbPMT1 and N- and C-terminal truncations of NtMPO1 and DmMPO1. This figure shows the effect of N- and C-terminal truncations to methylputrescine oxidase on 4MAB production in engineered yeast. Wild-type (WT) enzymes and indicated truncations were expressed from low-copy plasmids in putrescine-overproducing strain CSY1235 (see Example 1.5). Strains were cultured in selective media with 2% dextrose at 30 °C for 48 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates and error bars show standard deviation. Student's two-

tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001.

Figure 19 illustrates the production of the TA precursors 4MAB and NMPy and the side product 4MAB acid in liquid cultures of engineered yeast strains harboring single disruptions of one of four native aldehyde dehydrogenase genes. This figures shows the effect of disrupting individual aldehyde dehydrogenases on 4MAB acid accumulation. Putrescine overproducing strain CSY1235 (control) or daughter strains with nonsense mutation disruptions of hfd1, ald4, ald5, or ald6 were transformed with low-copy plasmids expressing *SPE1*, *AsADC*, and *speB*, *AbPMT1*, and *DmMPO1*$^{\Delta C\text{-}PTS1}$. Bars indicate relative 4MAB acid titer as measured by LC-MS/MS peak area normalized to CSY1235 (no ALD disruptions) after 48 hours of growth in selective media at 30 ºC. Data represent mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001.

Figure 20 illustrates production of (A) the 4MAB acid side product as well as (B) the TA precursors 4MAB and NMPy in liquid cultures of engineered yeast strains harboring one or more disruptions to native aldehyde dehydrogenases. This figure shows the effect of aldehyde dehydrogenase gene disruptions on production of (A) the 4MAB acid side product and (B) 4MAB and NMPy in engineered yeast. '+' and '–' symbols indicate presence or absence of functional enzyme, respectively. Strains were cultured in selective (YNB-DO) media with 2% dextrose at 30 ºC for 48 h before LC-MS/MS analysis. All data represent the mean of at least three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001. Unless otherwise indicated, statistical significance is shown relative to the corresponding control (CSY1235).

Figure 21 illustrates a comparison of the production of the TA precursor NMPy in liquid cultures of engineered yeast strains with either low-copy plasmid-based or genomic expression of putrescine overproduction genes, *AbPMT1*, and a *DmMPO1* truncation, in accordance with embodiments of the invention. This figure provides a comparison of 4MAB and NMPy production with plasmid-based (CSY1241) and genomic (CSY1243) expression of NMPy biosynthetic genes. Strain CSY1241 was transformed with low-copy plasmids expressing putrescine overproduction genes (*SPE1*, *AsADC*, *speB*), *AbPMT1*, and *DmMPO1*$^{\Delta C\text{-}PTS1}$. Strain CSY1243 expressed all of the aforementioned genes from genomic integrated copies. NMPy levels were quantified by LC-MS/MS following growth in selective (CSY1241) or non-selective (CSY1243) media at 30 ºC for 48 h. Data represent the mean of at least two biological replicates and error bars indicate standard deviation.

Figure 22 illustrates biosynthetic pathways for the production of the side product hygrine from NMPy and MPOB, in accordance with embodiments of the invention. Putative major and minor side reactions in yeast are indicated by bold and dotted arrows, respectively.

Figure 23 illustrates a comparison of production of the TA precursors tropinone and tropine and the side product hygrine in liquid cultures of engineered yeast strains expressing low-copy plasmid-based *AbPYKS*, *AbCYP82M3*, *DsTR1*, and one of four different CPRs. This figure shows production of tropine and related intermediates with expression of AbPYKS, AbCYP82M3, and DsTR1 in engineered yeast. Indicated genes were expressed from low-copy plasmids in CSY1246; '+' and '–' symbols indicate presence or absence of enzyme. Strains were cultured in selective media with 2% dextrose at 30 °C for 48 h before LC-MS/MS analysis. Data represent the mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 24 illustrates (A) a LC-MS/MS chromatogram illustrating the characteristic triple peak of the TA precursor MPOB produced in liquid cultures of engineered yeast strains, and (B) production of the TA precursors NMPy and MPOB in liquid cultures of yeast strains engineered to express *AbPYKS*, *AbCYP82M3*, and one of four CPRs from plasmids. This figure shows accumulation of NMPy and MPOB in the media of engineered strains expressing AbPYKS. (A) Representative LC-MS/MS multiple reaction monitoring (MRM) chromatogram for detection of MPOB in the extracellular medium of CSY1246 expressing *AbPYKS* only from a low-copy plasmid. The three characteristic MPOB isoform peaks are labelled with (I), (II), and (III). LC-MS/MS analysis was performed after growth in selective media at 30°C for 48 h. (B) Relative abundance of NMPy and MPOB (all 3 peaks) in the extracellular media of CSY1246 expressing *AbPYKS*, *AbCYP82M3*, and one of four CPRs from low-copy plasmids after 48 h of growth at 30 °C in selective media. '+' and '–' symbols indicate presence or absence of gene. Data represent mean of three biological replicates; error bars indicate standard deviation.

Figure 25 illustrates the effect of growth temperature on the production of the TA precursor tropine and the side product hygrine in liquid cultures of engineered yeast. This figure shows the effect of growth temperature on spontaneous hygrine production in the tropine-producing yeast strain (CSY1248). Relative selectivity represents the ratio of relative tropine titer to relative hygrine titer. Strains were cultured in non-selective media with 2% dextrose at 30 °C or 25 °C for 48 h before LC-MS/MS analysis. Data represent the mean of

three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001.

Figure 26 illustrates (A) the effect of *ALD4* and *ALD6* reconstitution on the growth of tropine-producing engineered yeast strains on media with or without acetate supplementation, and (B) the effect of eliminating acetate auxotrophy on the production of the side products 4MAB acid and hygrine in liquid cultures of tropine-producing engineered yeast strains, in accordance with embodiments of the invention. This figure shows the effect of elimination of acetate auxotrophy in engineered tropine-producing yeast strain. (A) Effect of reconstituting functional *ALD4* or *ALD6* genes on the growth of the NMPy-producing yeast strain (CSY1246) with and without acetate supplementation. *ALD4* and *ALD6* were expressed from low-copy plasmids. 'WT' indicates CSY1246 with control (BFP) plasmid. Adjacent columns show ten-fold dilutions. (B) Production of 4MAB acid and hygrine side products with reconstituted acetate metabolism in engineered yeast. '+' and '–' symbols indicate presence or absence of fed metabolite (acetate) or *ALD4* and *ALD6* genes expressed from low-copy plasmids. Strains were cultured in selective (YNB-DO) media with 2% dextrose at 30 °C for 48 h before LC-MS/MS analysis. Data represent the mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001.

Figure 27 illustrates (A) the effect of acetate auxotrophy on the accumulation of the TA precursors between NMPy and tropinone in liquid cultures of yeast strains engineered to produce tropine, and (B) representative LC-MS/MS chromatograms of the TA precursor MPOB produced in liquid cultures of yeast strains engineered to produce tropine with and without acetate auxotrophy, in accordance with embodiments of the invention. This figure shows the effect of reconstituting *ALD6* activity on metabolite flux through NMPy towards tropine in engineered yeast. (A) Production of intermediates between NMPy and tropinone in engineered strains with and without functional Ald6p. Intermediate abundances were measured by LC-MS/MS MRM in the extracellular media of the integrated tropine-producing strain (CSY1248) grown in non-selective media supplemented with 0.1% w/v potassium acetate (grey) or the tropine-producing strain with reconstituted *ALD6* (CSY1249) grown in non-selective media without acetate supplementation (pink) at 25 °C for 48 h. Data represent mean of three biological replicates; error bars indicate standard deviation. Student's two-tailed t-test: * P < 0.05, ** P < 0.01, *** P < 0.001. (B) Representative MRM chromatograms for MPOB production from CSY1248 (grey) and CSY1249 (red) cultured as described in (a).

Figure 28 illustrates the progression of improvements to production of the TA precursor tropine and the side product hygrine in liquid cultures of engineered yeast strains,. This figure provides a summary of strains engineered to increase tropine production in yeast. '–' symbol indicates absence of gene; 'p' and 'i' indicate gene expression from low-copy plasmid or genomic integration, respectively. Strains were cultured in selective or non-selective media with 2% dextrose at 30 ºC or 25 ºC for 48 h before LC-MS/MS analysis. Data represent the mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 29 illustrates the effect of expressing additional copies of the heterologous biosynthetic enzymes PMT, MPO, PYKS, and CYP82M3 on the production of each TA precursor between putrescine and tropine in liquid cultures of engineered yeast, in accordance with embodiments of the invention. This figure identifies of metabolic bottlenecks in optimized tropine-producing strain (CSY1249). Strain CSY1249 was transformed with a control plasmid expressing BFP ("no overexpression") or a low-copy plasmid expressing an additional copy of *AbPMT1* , *DmMPO1*$^{\Delta C\text{-}PTS1}$, *AbPYKS*, or *AbCYP82M3*. Intermediate levels in the extracellular medium were quantified by LC-MS/MS following growth at 25 ºC in selective media for 48 h. Data indicate mean of three biological replicates and error bars show standard deviation.

Figure 30 illustrates the impact of additional copies of bottleneck enzymes PMT and PYKS on tropine production in engineered yeast. This figure shows alleviation of metabolic bottlenecks through genomic integration of additional copies of PMT and PYKS enzymes. Tropine-producing strains CSY1249 and CSY1251 were cultured in non-selective media at 25 ºC for 48 h before LC-MS/MS analysis of growth medium. Data represent mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 31 illustrates the production of the TA precursor acyl donor compound PLA in liquid cultures of engineered yeast strains expressing heterologous lactate dehydrogenase and phenylpyruvate reductase enzymes. This figure shows LC-MS/MS analysis of yeast strains engineered to convert L-phenylalanine to 3-phenyllactic acid. Yeast strains are engineered to have a low-copy CEN/ARS plasmid harboring a LEU2 selection marker, a TDH3 promoter, and a coding sequence for BFP as a negative control; an LDH variant from *B. coagulans* (*BcLLDH*), *L. casei* (*LcLLDH*), *L. plantarum* (*LpLLDH*); or a PPR variant from *A. belladonna* (*AbPPR*), *L. plantarum* (*LpPPR*), *Escherichia coli* (*hcxB*) or *W. fluorescens*

(*WfPPR*). Yeast were grown from freshly transformed colonies in 300 µL selective media (-Leu) in 96-well deep-well microtiter plates. After 72 hours of growth in a shaking incubator at 25 °C and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Data show relative 3-phenyllactic acid titers normalized to trace levels present in negative control based on extracted ion chromatograms (ammonium adduct, EIC m/z^+ = 184). Data represent the mean of three biological replicates and error bars show standard deviation. Student's two-tailed t-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 32 shows LC-MS/MS chromatograms illustrating the production of the TA precursor acyl donor compound cinnamic acid in liquid cultures of engineered yeast strains expressing phenylalanine ammonia-lyase. This figure shows LC-MS/MS analysis of yeast strains engineered to convert L-phenylalanine to cinnamic acid. Yeast strains are engineered to have low-copy CEN/ARS plasmid harboring a *TRP1* selective marker, a *TEF1* promoter, and a coding sequence for (i) BFP or (ii) A. thaliana phenylalanine ammonia-lyase (*AtPAL1*). Yeast were grown from freshly transformed colonies in 300 µL selective media (-Trp) in 96-well deep-well microtiter plates. After 48 hours of growth in a shaking incubator at 30 °C and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Chromatogram traces show cinnamic acid produced by these strains based on the most abundant multiple reaction monitoring (MRM) transition for cinnamic acid (m/z+ 149 → 131). Each trace is representative of three samples.

Figure 33 illustrates the substrate specificity of UDP-glucosyltransferase 84A27 (UGT84A27) orthologs from TA-producing *Solanaceae* expressed in engineered yeast. This figure shows a comparison of the activity of UGT84A27 orthologs on three different phenylpropanoid compounds expressed in engineered yeast. (A) Phenylpropanoids tested as glucose (Glu) acceptors for UGT84A27 in engineered yeast. Top, (D)-3-phenyllactic acid (PLA); middle, *trans*-cinnamic acid (CA); bottom, *trans*-ferulic acid (FA). (B) Heatmap of percent conversion of fed phenylpropanoids to glucosides by yeast engineered for UGT84A27 expression. UGT84A27 orthologs or a BFP negative control were expressed from low-copy plasmids in CSY1251. Transformed cells were cultured in selective media supplemented with 500 µM PLA, CA, or FA for 72 h prior to LC-MS/MS analysis. Data represent the mean of $n = 3$ biologically independent samples ± standard deviation.

Figure 34 illustrates an example of chromatographic and mass spectrometric analysis of UGT84A27 activity. This figure depicts representative LC-MS/MS traces showing conversion of PLA, CA, and FA to cognate glucosides by AbUGT in CSY1251 cultured as in Figure 33B for 120 h to enable more complete glucosylation. For PLA, acid

(top trace in each panel) and glucoside (bottom trace in each panel) were distinguished by different NH_4^+ adduct parent masses as well as different retention times. For CA and FA, rapid fragmentation necessitated detection of the glucosides based on the lower-retention peaks produced by their phenylpropanoid fragments.

Figure 35 illustrates structure-guided active site engineering of AbUGT to alter substrate specificity. This figure shows structural analysis of the AbUGT 3D structure to identify potential mutations which increase activity on PLA. (A) Homology model of AbUGT84A27 constructed based on the crystal structure of *Arabidopsis thaliana* salicylate UDP-glucosyltransferase UGT74F2 with bound UDP (PDB: 5V2K). PLA (orange) is shown in the preferred binding pose with UDP-glucose (pink) based on docking simulations. (B) Zoomed view of AbUGT active site with docked D-PLA and UDP-glucose. Potential mutations identified to improve PLA selectivity (F130Y, L205F, I292Q) are shown; dashed lines indicate putative polar/hydrogen bond interactions.

Figure 36 illustrates the substrate specificity of AbUGT84A27 active site mutants. This figure shows a heatmap of percent conversion of fed phenylpropanoids to glucosides by yeast engineered for expression of AbUGT mutants. AbUGT wild-type, active site mutants, or a BFP negative control were expressed from low-copy plasmids in CSY1251. Transformed cells were cultured in selective media supplemented with 500 μM PLA, CA, or FA for 72 h prior to LC-MS/MS analysis. Data represent the mean of n = 3 biologically independent samples ± standard deviation.

Figure 37 shows LC-MS/MS chromatograms validating the step-wise biosynthesis of PLA glucoside in yeast engineered for tropine production. This figure shows multiple reaction monitoring (MRM) and extracted ion chromatogram (EIC) traces from culture media of yeast strains engineered for step-wise reconstitution of PLA glucoside. Strains were grown in non-selective media for 72 h prior to LC-MS/MS analysis of culture supernatant. Chromatogram traces are representative of three biological replicates.

Figure 38 shows a biosynthetic pathway schematic of the dual metabolic fates of glucose in yeast. This figure illustrates the effect of citrate on glucoside production via inhibition of glycolysis. Abbreviations: HXK, hexokinase; GPI, glucose-6-phosphate isomerase; PFK, phosphofructokinase; PGM, phosphoglucomutase; UGP, UDP-glucose pyrophosphorylase.

Figure 39 illustrates the effect of citrate supplementation on heterologous glucoside

production in engineered yeast. This figure shows the effect of 2% citrate supplementation on conversion of phenylpropanoid acids to glucosides by yeast engineered for AbUGT expression. Strain CSY1288 was cultured in non-selective media with or without 2% citrate and no additional supplementation to evaluate glucosylation of endogenously produced PLA, or with supplementation of 500 µM *trans*-cinnamic acid (CA) or *trans*-ferulic acid (FA). Cultures were grown for 72 h prior to LC-MS/MS analysis. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 40 shows relative PLA glucoside production in yeast strains engineered for overexpression of UDP-glucose biosynthetic enzymes. This figure illustrates the effect of overexpressing native enzymes involved in biosynthesis of the glucoside precursor UDP-glucose on production of PLA glucoside in engineered yeast. Enzymes or negative control (BFP) were expressed from low-copy plasmids in strain CSY1288. Strains were cultured for 72 h in selective media prior to LC-MS/MS analysis of metabolites in culture supernatant. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001. Statistical significance is shown relative to the corresponding control.

Figure 41 shows relative PLA glucoside production in CSY1288 with disruptions to endogenous glucosidases. This figure illustrates the effect of disrupting each of three native glycosidase genes on accumulation of PLA glucoside in engineered yeast. Strains were cultured in non-selective media for 72 h prior to LC-MS/MS analysis of culture supernatant. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001. Statistical significance is shown relative to the corresponding control.

Figure 42 shows LC-MS/MS chromatograms illustrating the production of the medicinal TA precursor hyoscyamine aldehyde from littorine in liquid culture for engineered yeast cells expressing *AbCYP80F1*. This figure shows LC-MS/MS analysis of yeast strains engineered to convert (R)-littorine to hyoscyamine aldehyde. Yeast strains are engineered to have a low-copy CEN/ARS plasmid harboring a *LEU2* selection marker, a *TDH3* promoter, and a coding sequence for littorine mutase CYP80F1 from A. belladonna (*AbCYP80F1*). Strains additionally have a second low-copy plasmid harboring a *TRP1* selection marker, a *TDH3* promoter, and a coding sequence for (i) BFP as a negative control, (ii) S. cerevisiae CPR (*NCP1*), or (iii) A. thaliana CPR (*AtATR1*). Yeast were grown from freshly transformed colonies in 300 µL selective media (-Leu –-Trp) supplemented with

1 mM littorine in 96-well deep-well microtiter plates. After 48 hours of growth in a shaking incubator at 30 °C and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Chromatogram traces show hyoscyamine aldehyde produced by these strains based on the most abundant MRM transition (m/z+ 288 → 124). Arrowheads indicate putative hyoscyamine aldehyde peak. Each trace is representative of three samples.

Figure 43 illustrates the production of the medicinal TA scopolamine from the medicinal TA hyoscyamine in liquid cultures for engineered yeast cells expressing orthologs of hyoscyamine 6β-hydroxylase/dioxygenase (H6H). This figure shows the conversion of (S)-hyoscyamine to (S)-scopolamine by engineered yeast strains expressing H6H orthologs. Yeast strains are engineered to have a low-copy CEN/ARS plasmid harboring a *LEU2* selection marker, a *TDH3* promoter, and a coding sequence for BFP as a negative control or an H6H variant from *D. stramonium* (*DsH6H*), *A. acutangulus* (*AaH6H*), *B. arborea* (*BaH6H*), or *D. metel* (*DmH6H*). Yeast were grown from freshly transformed colonies in 300 µL selective media (-Leu) supplemented with 1 mM hyoscyamine in 96-well deep-well microtiter plates. After 48 hours of growth in a shaking incubator at 30 °C and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Data represent the mean of three biological replicates and are normalized to the quantity of scopolamine contaminant in the fed hyoscyamine. Error bars represent standard deviation. Relative scopolamine titer was quantified based on the peak area of the m/z+ 304 → 138 MRM transition.

Figure 44 illustrates the effect of cofactor availability and supplementation in media on the conversion of hyoscyamine to scopolamine in liquid cultures of engineered yeast cells expressing DsH6H. This figure shows the effect of cofactor supplementation on conversion of (S)-hyoscyamine to (S)-scopolamine in engineered yeast. Yeast strains are engineered to have a low-copy CEN/ARS plasmid harboring a *LEU2* selection marker, a *TDH3* promoter, and a coding sequence for (i) BFP as a negative control or (iii) hyoscyamine 6β-hydroxylase/dioxygenase from D. stramonium (*DsH6H*). Yeast were grown from freshly transformed colonies in 300 µL selective media (-Leu) supplemented with the indicated substrates and/or cofactors in 96-well deep-well microtiter plates. After 48 hours of growth in a shaking incubator at 30 °C and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Relative (S)-scopolamine titers were quantified based on integrated peak area of the m/z+ 304 → 138 MRM transition and normalized to the strain expressing DsH6H and with all supplemented cofactors and substrates. Data represent mean of three biological replicates and error bars indicate standard deviation.

Hyo, (S)-hyoscyamine; 2-OG, 2-oxoglutarate; L-AA, L-ascorbic acid.

Figure 45 shows a hierarchical clustering heatmap of hyoscyamine dehydrogenase gene candidates identified from the *A. belladonna* transcriptome via analysis of tissue coexpression data. This figure shows clustering of tissue-specific expression profiles of transcripts in the *A. belladonna* transcriptome which potentially encode enzymes with hyoscyamine dehydrogenase activity. Transcript expression for each candidate is scaled by row using a normal distribution. Dendrogram indicates hierarchical clustering of candidates by tissue-specific expression profile. Known TA pathway genes are identified by name; putative HDH candidates are indicated with locus ID. Black triangles indicate candidates screened for activity; double black triangle indicates candidate with experimentally verified HDH activity.

Figure 46 illustrates the production of the medicinal TA scopolamine from littorine in liquid cultures of engineered yeast cells expressing hyoscyamine dehydrogenase (HDH) candidates. This figure illustrates the experimental screening for activity of HDH candidates identified from the transcriptome of *A. belladonna* in engineered yeast. Yeast strains are engineered to express *A. belladonna* littorine mutase (AbCYP80F1) and *D. stramonium* hyoscyamine 6β-hydroxylase/dioxygenase (DsH6H) from constitutive promoters within expression cassettes integrated into the genome, as well as one of each of the 13 HDH candidates from a low-copy CEN/ARS plasmid harboring a TRP1 selection marker and a TDH3 promoter. Yeast were grown from freshly transformed colonies in 300 µL selective media (-Trp) supplemented with 1 mM littorine in 96-well deep-well microtiter plates. After 72 hours of growth in a shaking incubator at 30 ºC and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Relative hyoscyamine aldehyde titers were quantified based on integrated peak area of the m/z^+ 288 → 124 MRM transition and normalized to that of the engineered strain expressing BFP instead of an HDH candidate. (S)-scopolamine titers were quantified based on integrated peak area of the m/z^+ 304 → 138 MRM transition and a standard curve of a genuine scopolamine standard. Data represent mean of three biological replicates and error bars indicate standard deviation.

Figure 47 illustrates the three-dimensional structure of hyoscyamine dehydrogenase from *A. belladonna*. This figure shows a cartoon representation of the structure of AbHDH as a homology model constructed based on the crystal structure of *Populus tremuloides* sinapyl alcohol dehydrogenase (PtSAD; PDB: 1YQD) as a template. NADPH and Zn^{2+} are shown in the active site. The inset box shows a zoomed view of the AbHDH active site with NADPH and docked hyoscyamine aldehyde. Dashed lines indicate interactions important

for catalysis.

Figure 48 shows a phylogenetic tree of the three identified HDH orthologs (AbHDH, DiHDH, DsHDH) together with closest protein hits in the UniProt/SwissProt database. This figure shows clustering of the three identified HDH enzyme orthologs with closely related protein sequences based on a BLAST search of the UniProt/SwissProt database. Sequences shown include top 50 BLASTp hits based on E-value, as well as 10 additional hits selected from among the next 100 ranks. Phylogenetic relationships were derived via bootstrap neighbor-joining with n = 1000 trials in ClustalX2 and the resulting tree was visualized with FigTree software. Abbreviations: ADH, alcohol dehydrogenase; CADH, cinnamyl alcohol dehydrogenase; MTDH, mannitol dehydrogenase; DPAS, dehydroprecondylocarpine acetate synthase; 8HGDH, 8-hydroxygeraniol dehydrogenase; GDH, geraniol dehydrogenase; GS, geissoschizine synthase; REDX, unspecified redox protein.

Figure 49 illustrates the production of the medicinal TA scopolamine from littorine in liquid cultures of engineered yeast cells expressing hyoscyamine dehydrogenase orthologs. This figure illustrates a comparison of activities between identified HDH enzyme orthologs expressed in engineered yeast. Yeast strains are engineered to express *A. belladonna* littorine mutase (AbCYP80F1) and *D. stramonium* hyoscyamine 6β-hydroxylase/dioxygenase (DsH6H) from constitutive promoters within expression cassettes integrated into the genome, one of each of the three HDH orthologs (AbHDH, DiHDH, DsHDH) from a low-copy CEN/ARS plasmid harboring a TRP1 selection marker and a TDH3 promoter, and an additional copy of DsH6H from a low-copy CEN/ARS plasmid harboring a LEU2 selection marker and a TDH3 promoter. Yeast were grown from freshly transformed colonies in 300 µL selective media (-Leu -Trp) supplemented with 1 mM littorine in 96-well deep-well microtiter plates. After 72 hours of growth in a shaking incubator at 30 ºC and 460 rpm, the yeast were pelleted and the media supernatant was analyzed by LC-MS/MS. Relative hyoscyamine aldehyde titers were quantified based on integrated peak area of the m/z^+ 288 → 124 MRM transition and normalized to that of the engineered strain expressing AbHDH and BFP instead of DsH6H. (S)-scopolamine titers were quantified based on integrated peak area of the m/z^+ 304 → 138 MRM transition and a standard curve of a genuine scopolamine standard. Data represent mean of three biological replicates and error bars indicate standard deviation.

Figure 50 illustrates experimental validation of conversion of fed littorine to scopolamine by yeast engineered for expression of CYP80F1, HDH, and H6H. This figure

shows multiple reaction monitoring (MRM) LC-MS/MS traces from culture media of yeast strains engineered for conversion of littorine to scopolamine. Strains were cultured for 72 h in non-selective media supplemented with 1 mM littorine prior to LC-MS/MS analysis of metabolites in culture supernatant. Dark trace in bottom-right panel (CSY1294, scopolamine) represents 125 nM (38 µg/L) scopolamine standard. Chromatogram traces are representative of three biological replicates

Figure 51 illustrates the canonical plant ER-to-vacuole trafficking and maturation pathway for SCPL acyltransferases (SCPL-ATs). This figure shows a schematic representation of a typical ER-to-vacuole protein trafficking pathway followed by SCPL-ATs in plants, with *A. belladonna* littorine synthase (AbLS) shown as an example. Circled numbers indicate major steps in SCPL-AT expression and activity, including maturation in the (1) ER lumen and (2) Golgi, (3) trafficking to the vacuole, and vacuolar (4) substrate import and (5) product export.

Figure 52 shows co-localization of wild-type littorine synthase from *A. belladonna* expressed in engineered yeast. This figure shows epifluorescence microscopy of yeast engineered for expression of N-terminal GFP-tagged AbLS (GFP-AbLS) and stained with the vacuolar membrane stain FM4-64. Microscopy was performed on CSY1294 expressing GFP-AbLS from a low-copy plasmid. Scale bar, 5 µm.

Figure 53 illustrates a strategy for forced localization of littorine synthase to different yeast sub-cellular compartments via signal sequence replacement. This figure illustrates a protein engineering approach to modifying the sub-cellular localization of AbLS to address potential restrictions on substrate availability in different compartments. (A) Schematic of yeast sub-cellular compartments targeted for localization of AbLS via signal sequence swapping. Signal sequence source proteins are indicated for each compartment. (B) Termini and residues selected for AbLS signal sequence replacement. Residues comprising each signal sequence domain were selected based on structural annotations in the UniProt/SwissProt database.

Figure 54 shows a Western blot of wild-type AbLS expressed in tobacco and treated with deglycosylases. This figure illustrates the identification of glycosylation modification types for AbLS expressed in plants. C-terminal HA-tagged AbLS was transiently expressed in *N. benthamiana* leaves via agroinfiltration. Crude leaf extracts were either untreated (lane 1: '--'), or treated with peptide N-glycosidase F (PNGase F; lane 2: '*N*') or *O*-glycosidase (lane 3: '*O*') to remove *N*- or *O*-linked glycosylation, respectively. Crude extracts were

separated by electrophoresis on a NuPAGE 4-12% Bis-Tris gel and then transferred to a nitrocellulose membrane for immunodetection using a chimeric rabbit IgGκ anti-HA HRP-conjugated antibody. All electrophoresis and blotting steps were performed under disulfide reducing conditions (see Online Methods). Lane 'L', Bio-Rad Precision Plus Dual Color protein ladder.

Figure 55 shows Western blots of AbLS glycosylation site mutants expressed in yeast and tobacco. This figure shows a comparison of the N-glycosylation patterns present for AbLS expressed in yeast and in tobacco. C-terminal HA-tagged wild-type AbLS, single glycosylation site point mutants (N → Q), or a quadruple mutant were expressed transiently via agroinfiltration in *N. benthamiana* ('*Nb*') (A) or from low-copy plasmids in CSY1294 ('Yeast') (B). Preparation of tobacco and yeast crude extracts was performed under denaturing, disulfide-reducing conditions (see Online Methods). Crude extracts were separated by electrophoresis on a NuPAGE 4-12% Bis-Tris gel and then transferred to a nitrocellulose membrane for immunodetection using a chimeric rabbit IgGκ anti-HA HRP-conjugated antibody. All electrophoresis and blotting steps were performed under disulfide reducing conditions (see Online Methods). For (A) and (B), corresponding yeast- and tobacco-expressed controls are included for comparison. Lane 'L', Bio-Rad Precision Plus Dual Color protein ladder.

Figure 56 shows phylogenetic identification of putative endoproteolytic propeptide removal in littorine synthase. This figure shows a sequence alignment of AbLS with characterized serine carboxypeptidases and SCPL acyltransferases known to possess (AtSCT, AsSCPL1, TaCBP2) or lack (AtSMT, yPRC1) proteolytically-removed internal propeptide linkers (bold, grey). Putative N-terminal signal peptides are indicated in bold (black); disulfide bonds are indicated as connecting lines. AtSCT, *Arabidopsis thaliana* sinapoylglucose: choline sinapoyltransferase; AtSMT, *A. thaliana* sinapoylglucose:malate sinapoyltransferase; AbLS, *Atropa belladonna* littorine synthase; AsSCPL1, *Avena strigosa* avenacin synthase; TaCBP2, *Triticum aestivum* carboxypeptidase 2; yPRC1, yeast carboxypeptidase Y.

Figure 57 shows structural identification of putative endoproteolytic propeptide removal in littorine synthase. This figure shows a comparison of the three-dimensional structures of two SCPL-ATs, one of which is known to contain a proteolytically removed internal propeptide sequence. Left: Crystal structure of TaCBP2 (PDB: 1WHT) in (top) cartoon and (bottom) surface representation showing disulfide bonds and internal propeptide removal sites. Right: Homology model of AbLS based on the crystal structure of

TaCBP2 in (top) cartoon and (bottom) surface representation showing N-terminal signal peptide, disulfide bonds, and putative internal propeptide which appears to block active site access.

Figure 58 shows analysis of proteolytic cleavage patterns for AbLS split controls and putative propeptide-swapped variants in yeast. This figure shows Western blot analysis of protein fragment sizes produced by AbLS split controls and propeptide variants expressed in engineered yeast. C-terminal HA-tagged AbLS variants were expressed from low-copy plasmids in CSY1294 (lanes 1-6); HA-tagged wild-type AbLS expressed in *Nicotiana benthamiana* (Nb) is shown as an additional control (lane 7). Gel electrophoresis and blotting were performed under disulfide-reducing conditions and detection was performed using an anti-HA antibody (see Online Methods). Lane symbols: L, protein molecular weight ladder; WT, wild-type AbLS; SPL, AbLS split at putative propeptide with signal peptides on both fragments; SPL-T, AbLS split at putative propeptide without signal peptides on either fragment; GS, AbLS variant with wild-type propeptide swapped for flexible Gly-Ser linker; SCT, AbLS variant with wild-type propeptide swapped for AtSCT propeptide sequence; CUT, AbLS variant with wild-type propeptide swapped for synthetic poly-arginine site recognized and cleaved by Kex2p protease.

Figure 59 illustrates *de novo* hyoscyamine and scopolamine production in yeast strains engineered for expression of AbLS N-terminal fusions. This figure shows a comparison of *de novo* hyoscyamine and scopolamine production in yeast strains expressing AbLS with different soluble protein domains fused to the N-terminus. Wild-type (control) or AbLS fusions were expressed from low-copy plasmids in CSY1294. Transformed strains were cultured for 96 h in selective media prior to LC-MS/MS analysis of metabolites in culture supernatant. Data represent the mean of $n = 3$ biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 60 shows fluorescence microscopy of tobacco alkaloid transporters expressed in CSY1296 for alleviation of vacuolar TA transport limitations. This figure shows fluorescence microscopy images of engineered yeast expressing tobacco alkaloid transporters fused at their C-termini to GFP, to enable identification of their sub-cellular localization. C-terminal GFP fusions of (A) NtJAT1 and (B) NtMATE2 were expressed from low-copy plasmids in CSY1296. Scale bar, 5 μm.

Figure 61 shows production of tropine, hyoscyamine, and scopolamine in CSY1296

engineered for expression of heterologous alkaloid transporters. This figure illustrates the utility of different plant alkaloid transporters in alleviating intracellular substrate transport limitations in yeast engineered for TA production. *Nicotiana tabacum* jasmonate-inducible alkaloid transporter 1 (NtJAT1), multidrug and toxin extrusion (MATE) transporters 1 or 2, or a negative control (BFP) were expressed from low-copy plasmids in CSY1296. Transformed strains were cultured for 96 h in selective media prior to LC-MS/MS analysis of metabolites in culture supernatant. Data represent the mean of $n = 3$ biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 62 shows LC-MS/MS chromatograms in (A) product ion mode and (B) multiple reaction monitoring mode illustrating the *de novo* production of the non-natural TA cinnamoyltropine in engineered yeast. This figure shows LC-MS/MS analysis of engineered yeast strains producing the non-natural TA cinnamoyltropine. (A) Tandem MS/MS spectra of extracellular medium of (i) tropine-producing strain CSY1251; (ii) CSY1251 expressing phenylalanine ammonia-lyase (AtPAL1), 4-coumarate-CoA ligase 5 (At4CL5), and cocaine synthase (EcCS), denoted CSY1282; or (iii) a genuine cinnamoyltropine standard for a parent mass of m/z+ = 272. Blue diamond indicates parent compound peak. (B) Validation of EcCS acyltransferase activity on cinnamic acid and α-tropine via substrate feeding. Strains were transformed with combinations of plasmids expressing AtPAL1 (low-copy plasmid pCS4252) and/or At4CL5 and EcCS (high-copy plasmid pCS4207), and then cultured in media with different supplemented substrates, as follows: (i) CEN.PK2 + At4CL5 + EcCS + 0.1 mM trans-cinnamic acid; (ii) CEN.PK2 + At4CL5 + EcCS + 0.5 mM α-tropine; (iii) CEN.PK2 + AtPAL1 + At4CL5 + EcCS; (iv) CEN.PK2 + AtPAL1 + At4CL5 + EcCS + 0.5 mM α-tropine; (v) CSY1251 + At4CL5 + EcCS; (vi) CSY1251 + At4CL5 + EcCS + 0.2 mM trans-cinnamic acid; (vii) CSY1251 + AtPAL1 + At4CL5 + EcCS; (viii) 25 nM cinnamoyltropine standard. For (A) and (B), yeast strains were cultured in selective media (YNB-DO + 2% dextrose + 5% glycerol) at 25 °C for 72 h prior to LC-MS/MS analysis.

Figure 63 illustrates the impact of varied carbon sources fed (A) alone or (B) together with dextrose on the production of tropine and related TA precursors in liquid cultures of engineered yeast. This figure shows the optimization of carbon source to improve tropine production in engineered yeast. Overnight cultures of tropine-producing strain CSY1249 (see Example 3.3.4) were grown in non-selective rich media (YPD). Overnight cultures were pelleted and resuspended in non-selective defined medium (YNB-SC) with all amino acids and (A) 2% of each carbon source or (B) 2% dextrose and 2% of each additional carbon source, including dextrose. Cultures were grown at 25 °C for 48 h

prior to analysis of growth medium by LC-MS/MS. Data show relative titer of each metabolite normalized to (A) 2% dextrose or (B) 2% + 2% dextrose. Data represent the mean of three biological replicates and error bars indicate standard deviation.

Figure 64 illustrates metabolic bottleneck analysis of scopolamine-producing strain CSY1296. This figure shows the effect of expressing additional copies of flux-limiting enzymes on production of TAs and TA precursors in engineered yeast. An additional copy of each biosynthetic enzyme between tropine and scopolamine was expressed from the following low-copy plasmids in strain CSY1296: (**A**) WfPPR, pCS4436; (**B**) AbUGT, pCS4440; (**C**) DsRed-AbLS, pCS4526; (**D**) AbCYP80F1, pCS4438; (**E**) DsHDH, pCS4478; (**F**) DsH6H, pCS4439; or a BFP control (pCS4208, pCS4212, or pCS4213) corresponding to the same auxotrophic marker as each biosynthetic gene plasmid. Transformed strains were cultured in appropriate selective media at 25 °C for 96 hours prior to quantification of metabolites in the growth medium by LC-MS/MS. Data indicate the mean of $n = 3$ biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 65 shows the effect of alleviating flux and transport limitations on hyoscyamine and scopolamine production in engineered yeast. This figure shows a comparison of *de novo* hyoscyamine and scopolamine production in yeast strain CSY1296 and CSY1297, where the latter possesses additional genomic copies of flux-limiting enzymes (WfPPR and DsH6H) as well as a tobacco vacuolar alkaloid importer (NtJAT1). Strains were cultured in non-selective media for 96 h prior to LC-MS/MS analysis of metabolites in culture supernatant. Data represent the mean of $n = 3$ biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 66 shows an example of the identification of putative TA transporter candidates via linear regression and hierarchical clustering strategies**.** This figure presents a heatmap and dendrogram that show hierarchical clustering of all TA transporter candidates identified from *A. belladonna* transcriptome via linear regression analysis with slope coefficient $P < 0.01$ relative to bait genes. Known TA-related genes and one of the two novel TA transporter candidates experimentally validated, AbPUP1, are indicated; all other rows represent uncharacterized transporter candidates. The second experimentally validated candidate, AbLP1, did not correlate strongly enough with bait genes to be identified in this analysis. Tissue abbreviations: FL, flower; MS, mature seed; PT, primary taproot; SS, sterile seedling; CA, callus; SR, secondary root; ST, stem; RF, ripe fruit; GF,

green fruit; LF, leaf; FB, flower buds.

Figure 67 depicts the data classification strategy used for Training, optimization, and testing of binary classifier models for prediction of novel TA-related genes from *Atropa belladonna* transcriptome. This figure shows a flowchart used for generation of training output values for *A. belladonna* transcripts following dataset pre-processing. “TA”: transcript encodes gene product with known role in TA biosynthesis; “non-TA”: transcript encodes gene product known not to be involved in TA biosynthesis; “unknown”: transcript encodes gene product with unknown role in TA biosynthesis. Values indicate number of transcripts.

Figure 68 shows binary classifier model performance in training, cross-validation, and testing. This figure depicts a comparison in the TA-related gene predictive performance of three supervised binary classifier models trained and cross-validated on an *A. belladonna* transcriptome. Three binary classifier models were trained using one of four different oversampling methods (none: no oversampling; ROSE: random oversampling examples; SMOTE: synthetic minority oversampling technique; up, random oversampling) and one of two different performance metrics (accuracy: fraction of correctly predicted samples out of total number of samples; ROC: area under the receiver operating characteristic curve, which plots true positive rate versus false positive rate).

Figure 69 also shows binary classifier model performance in training, cross-validation, and testing. This figure depicts a comparison in the computation time required by three supervised binary classifier models for training, cross-validation, and testing on the same dataset as in Figure 68. Three binary classifier models were trained using one of four different oversampling methods (none: no oversampling; ROSE: random oversampling examples; SMOTE: synthetic minority oversampling technique; up, random oversampling) and one of two different performance metrics (accuracy: fraction of correctly predicted samples out of total number of samples; ROC: area under the receiver operating characteristic curve, which plots true positive rate versus false positive rate).

Figure 70 shows binary classifier model performance in testing. This figure depicts confusion matrices showing predictive performance of each of three optimized binary classifier models on testing data (same dataset as in Figs. 66-69). Logistic regression (glm), random forest (ranger), and neural network (nnet) binary classifiers were trained and cross-validated using the oversampling techniques and performance metrics which yielded maximum balanced accuracy and minimum computation time in Figs. 68-69. Note that circular wedges are only shown to scale within each matrix row.

Figure 71 illustrates a simplified schematic of the final optimized neural network for prediction of TA-related genes from *A. belladonna* transcriptome. This figure shows a neural network with 11-5-1 architecture whose predictive performance is shown in Fig. 70 (nnet).

Figure 72 shows a phylogenetic analysis of *A. belladonna* putative TA transporters. This figure depicts a phylogenetic tree showing relationships between *A. belladonna* putative TA transporters (AbPUP1, AbLP1, AbCT1), experimentally characterized plant alkaloid transporters, and selected characterized small molecule transporters. Values at each junction signify number of bootstraps out of *n* = 1000 iterations. Scale bar, number of substitutions per amino acid site. Species: At, *Arabidopsis thaliana*; Nt, *Nicotiana tabacum*; Ps, *Papaver somniferum*; Ec, *Escherichia coli*; Cr, *Catharanthus roseus*; Os, *Oryzia sativa*.

Figure 73 shows cellular metabolite transport limitations in yeast engineered for *de novo* production of TAs. This figure shows a schematic illustrating metabolic modules and key metabolite transport steps in engineered yeast described in this specification. TA biosynthesis is distributed across five metabolic modules (I-V) and five cellular compartments: mitochondrion, peroxisome, endoplasmic reticulum (ER), vacuole, and cytosol. The central esterification reaction assembling the TA scaffold from tropine and phenyllactic acid (PLA) glucoside is catalyzed by littorine synthase (*Ab*LS) with N-terminally fused fluorescent protein DsRed in the vacuole. For this reaction, substrates (tropine and PLA glucoside) must be imported to the vacuole lumen and product (littorine) must be exported to the cytosol. Principal starting materials and products are shown for each module. Charges ('+', positive; no symbol, uncharged) of key intermediates and products relevant to transport are indicated. Protein names: *Nt*JAT1, *Nicotiana tabacum* jasmonate-inducible alkaloid transporter 1; *Nt*MATE2, *N. tabacum* multidrug and toxin extrusion transporter 2; *Ab*PUP1, *Atropa belladonna* purine uptake permease-like transporter 1; *Ab*LP1, *A. belladonna* lactose permease-like transporter 1; PdrXp, yeast pleiotropic drug resistance transporters.

Figure 74 shows genetic modules used for pathway construction via chromosomal integration in yeast. This figure is a schematic depicting the genomic layout of metabolic genes used in yeast strain construction. Block arrows represent gene expression cassettes with promoter (line arrow) and terminator (line 'T'). Genus and species sources for heterologous genes are indicated by two letters preceding gene name; no source symbol indicates native yeast gene (if all uppercase) or *E. coli* gene (*speB*). Superscript annotations on gene names indicate N- or C-terminal modifications; dash symbol (-) indicates fusion

protein.

Figure 75 illustrates an engineered pathway for *de novo* biosynthesis of medicinal tropane alkaloids and derivatives in yeast. This figure shows a schematic of the engineered biochemical pathways which can be used to convert simple precursors into medicinal tropane alkaloids and *N*-substituted derivatives. Enzyme symbols: Arg2p, glutamate N-acetyltransferase; Car1p, arginase; *As*ADC, *Avena sativa* arginine decarboxylase; speB, *Escherichia coli* agmatine ureohydrolase; Spe1p, ornithine decarboxylase; Fms1p, polyamine oxidase; Meu1p, methylthioadenosine phosphorylation; Oaz1p, ornithine decarboxylase antizyme-1; *Ab*PMT1, *Atropa belladonna* putrescine *N*-methyltransferase 1; *Ds*PMT1, *Datura stramonium* putrescine *N*-methyltransferase 1; *Dm*MPO1$^{\Delta C\text{-}PTS1}$, *Datura metel N*-methylputrescine oxidase 1 with peroxisome targeting sequence 1 and truncated C-terminus; Hfd1p and Ald2p-Ald5p, aldehyde dehydrogenases; *Ab*PYKS, *A. belladonna* pyrrolidine ketide synthase; *Ab*CYP82M3, *A. belladonna* tropinone synthase; *At*ATR1, *Arabidopsis thaliana* NADPH:cytochrome P450 reductase; *Ds*TR1, *D. stramonium* tropinone reductase I; Aro8p/Aro9p, aromatic aminotransferases; *Wf*PPR, *Wickerhamia fluorescens* phenylpyruvate reductase; *Ab*UGT, *A. belladonna* UDP-glucosyltransferase 84A27; Egh1p, steryl-β-glucosidase; DsRed-*Ab*LS, DsRed fluorescent protein fused to the N-terminus of *A. belladonna* littorine synthase; *Ab*CYP80F1, *A. belladonna* littorine mutase; *Ds*HDH, *D. stramonium* hyoscyamine dehydrogenase; *Ds*H6H, *D. stramonium* hyoscyamine 6β-hydroxylase/dioxygenase; *Hs*CYP, human liver cytochromes P450; *Zv*PNO, *Zonocerus variegatus* pyrrolizidine *N*-oxygenase.

Figure 76 shows the effect of putative TA transporter expression on TA titers in engineered yeast. This figure illustrates TA production in yeast engineered for expression of putative TA transporters. Transporter candidates or a blue fluorescent protein (BFP) control were expressed from low-copy plasmids in CSY1300 and metabolites in supernatant were quantified by LC-MS/MS after 96 h growth in selective media. Data represent the mean of $n = 3$ biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed *t*-test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

Figure 77 shows sub-cellular localization of TA transporter C-terminal GFP fusions in engineered yeast. This figure depicts fluorescence microscopy images of yeast engineered for expression of TA transporters tagged with fluorescent proteins. AbPUP1-GFP and AbLP1-GFP were expressed from low-copy plasmids in CSY1300 (refer to Example Methods). Scale bar, 2 μm.

Figure 78 shows structural analysis of *A. belladonna* putative TA transporters. This figure shows Kyte-Doolittle hydropathy plots (a, b) and predicted membrane transmembrane topology plots (c, d) for AbPUP1 (a, c) and AbLP1 (b, d). The figure also presents (e) the predicted orientation and cellular function of AbPUP1 (left) and AbLP1 (right) based on *in silico* structural analysis and experimental characterization in yeast. Cylinder representations of transporter homology models are shown.

Figure 79 shows gene ontology (GO) analysis of putative TA transporters. This figure depicts gene ontology prediction maps of *A. belladonna* putative TA transporters. GO analysis of cellular component (a, d), biological process (b, e), and molecular function (c, f) for AbPUP1 (a-c) and AbLP1 (d-f) was performed concurrently with homology modeling (see Fig. 78e). Box shading indicates confidence score (C-score) for corresponding GO term.

Figure 80 illustrates the elucidation of putative TA transporter substrates in yeast. This figure presents growth curves showing the effect of *Ab*PUP1 and *Ab*LP1 on alkaloid tolerance in yeast. Strain AD1-8 expressing blue fluorescent protein (control), AbPUP1, or AbLP1 from low-copy plasmids was cultured in selective media supplemented with 0-20 mM alkaloids; tropine was supplemented in media at only 1 mM as higher tested concentrations severely retarded the growth of AD1-8 such that growth measurements were not feasible. Solid lines and shaded/hatched bars respectively indicate mean and standard deviation of n = 3 biologically independent samples.

Figure 81 also shows the elucidation of putative TA transporter substrates in yeast. This figure shows a heatmap of transporter substrate specificity (θ_{HM}), quantified as the ratio of time required for half of maximum OD_{600} increase (t_{HM}) with maximum alkaloid concentration relative to without alkaloid supplementation ($\theta_{HM} = t_{HM,max_alk} / t_{HM,no_alk}$). Heatmap values were extracted from mean growth curves in Fig. 80. Strain AD1-8 expressing blue fluorescent protein (control), AbPUP1, or AbLP1 from low-copy plasmids was cultured in selective media supplemented with 0-20 mM alkaloids. Note that tropine was supplemented in media at only 1 mM as higher tested concentrations severely retarded the growth of AD1-8 such that growth measurements were not feasible.

Figure 82 shows elucidation of putative TA transporter transport mechanisms in yeast. This figure presents growth curves showing H^+ dependence of TA transport by *Ab*PUP1 and *Ab*LP1 in yeast. Strain AD1-8 expressing blue fluorescent protein (control), AbPUP1, or AbLP1 from low-copy plasmids was cultured in selective media supplemented with 10 mM hyoscyamine and 100 mM potassium phosphate buffer. Solid lines and shaded/hatched bars

respectively indicate mean and standard deviation of *n* = 3 biologically independent samples.

Figure 83 also shows the elucidation of putative TA transporter transport mechanisms in yeast. This figure presents growth curves showing the effect of monovalent cations on TA transport by *Ab*PUP1 and *Ab*LP1 in yeast. Strain AD1-8 expressing blue fluorescent protein (control), AbPUP1, or AbLP1 from low-copy plasmids was cultured in selective media supplemented with 20 mM hyoscyamine and 0-100 mM NaCl or KCl. '--' indicates no salt supplementation in excess of that already present in yeast nitrogen base. Solid lines and shaded/hatched bars respectively indicate mean and standard deviation of *n* = 3 biologically independent samples.

Figure 84 shows the optimization of redox cofactor availability, intracellular transport, prototrophy, and culture conditions improve TA production in yeast. This figure provides a graphical summary of strain optimization to increase *de novo* production of hyoscyamine and scopolamine in engineered yeast. Strains were cultured in appropriate selective or non-selective media for 96 h before quantification of TAs in supernatant by LC-MS/MS. Symbols: --, absence of enzyme and/or disruption of encoding gene; P, expression from low-copy plasmid; +, expression from chromosomal gene copy; ++, expression from native gene copy plus additional chromosomal copy; ub, unbuffered growth media; 5.8, growth media buffered to pH 5.8 with 0.1 M potassium phosphate. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed *t*-test: **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

Figure 85 illustrates the canonical reactions catalyzed by NADPH-regenerating enzyme candidates tested in yeast.

Figure 86 illustrates the effect of NADPH-regenerating enzyme overexpression on TA titers in engineered yeast. This figure shows TA production in engineered TA-producing yeast expressing different NADPH-regenerating enzymes. Enzyme-coding genes or a BFP control were overexpressed from low-copy plasmids in strain CSY1300 and metabolites in supernatant were quantified by LC-MS/MS after 96 h growth in selective media. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation.

Figure 87 shows that uracil prototrophy does not improve TA production in CSY1323. This figure shows TA production in a yeast strain engineered for uracil prototrophy. Uracil prototrophy was conferred by expression of Ura3p from a low-copy plasmid in CSY1323.

Transformed strain was cultured for 96 h in selective media prior to LC-MS/MS analysis of TA titers in supernatant. TA titers produced by CSY1323 (with uracil and tryptophan auxotrophy) as shown in Fig. 84 are indicated for comparison. Data represent the mean of *n* = 3 biologically independent samples (open circles) and error bars show standard deviation.

Figure 88 shows the effect of media pH buffering on TA titers in engineered yeast. This figure shows an optimization curve depicting the effect of extracellular media pH on TA production in engineered yeast. Strain CSY1324 was cultured in selective media buffered between pH 5.4 and 7.0 with 0.1 M potassium phosphate (K_2HPO_4/KH_2PO_4) and metabolites in supernatant were quantified by LC-MS/MS after 96 h growth. Data represent the mean of *n* = 3 biologically independent samples (filled circles).

Figure 89 shows the effect of media pH buffering with 0.1 M potassium phosphate on TA production in CSY1323. This figure shows an optimization curve depicting the effect of extracellular media pH on TA production in engineered yeast. Strain CSY1323 was cultured in non-selective media buffered between pH 5.4 and 7.0 with 0.1 M potassium phosphate (K_2HPO_4/KH_2PO_4) and metabolites in supernatant were quantified by LC-MS/MS after 96 h growth. Data represent the mean of *n* = 3 biologically independent samples (filled circles).

Figure 90 illustrates an example of *in silico* biosynthetic network expansion to predict TA derivatives one step away from hyoscyamine and scopolamine using ATLASx. This figure shows a metabolic map that was generated from a one-generation search of the ATLASx chemATLAS database using hyoscyamine and scopolamine as queries and a default conserved atom ratio (CAR) threshold of 0.34. ATLASx compound ID numbers are shown below each structure. Arrows represent predicted one-step enzymatic reactions; arrow color indicates CAR.

Figure 91 shows an example strategy for the production of nortropane alkaloids in engineered yeast. This figure illustrates a simplified schematic of TA biosynthetic pathway extended using ATLASx for *de novo* production of nortropane alkaloids (norTAs) in yeast. Dashed arrows indicate abbreviated pathway steps. H6H, hyoscyamine 6β-hydroxylase/ dioxygenase; HsCYP, human liver cytochrome P450; CPR, yeast NADPH-CYP450 reductase (Ncp1p); CYB5, human cytochrome b5.

Figure 92 illustrates a schematic of genetic constructs used for heterologous co-expression of yeast NADPH:cytochrome P450 reductase (*NCP1*), human cytochrome B5 (*HsCYB5A*), and human liver CYP450 candidates (*HsCYPXXX*) in yeast. Construct 1,

expressed from a low-copy plasmid with a *TRP1* auxotrophic marker, encoded a fusion protein comprising the complete open reading frame of Ncp1p fused using an Arg-Arg-Ala tripeptide linker to the N-terminus of HsCYB5A truncated after residue 104 to remove its C-terminal ER membrane anchor. Construct 2, expressed from a low-copy plasmid with a *URA3* auxotrophic marker, encoded one of eight HsCYP candidates.

Figure 93 shows example results of the development of a LC-MS/MS method for unambiguous identification of TAs and derivatives. This figure depicts representative chromatograms for 10 μM aqueous standards. Each TA and derivative can be identified using a unique combination of multiple reaction monitoring (MRM) mass transition and retention time; no two compounds with the same parent mass (or with shared mass transitions) elute at the same retention time. Solvent gradient for chromatography and MRM detection parameters are provided in Tables 8-10.

Figure 94 illustrates the effect of human liver CYP expression on nortropane alkaloid production in engineered yeast. This figure shows titers and molar production ratio of (top) norhyoscyamine and (bottom) norscopolamine in engineered yeast. CYPs or a BFP control were expressed from low-copy plasmids in CSY1324 and metabolites in supernatant were quantified by LC-MS/MS after 96 h growth in selective media buffered to pH 5.8 with 0.1 M potassium phosphate (top) or unbuffered (bottom). Data represent the mean of n = 3 biologically independent samples (open circles) and error bars show standard deviation. Student's two-tailed t-test: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

Figure 95 shows analytical validation of nortropane alkaloid production in engineered yeast. This figure depicts representative LC-MS/MS multiple reaction monitoring (MRM) chromatograms for norTAs produced *de novo* by CSY1324 expressing HsCYP2D6 and cultured as in Fig. 94, or authentic chemical standards. Chromatograms are representative of n = 3 biologically independent samples.

Figure 96 shows the results of *in silico* enzyme prediction using BridgIT to enable *de novo* biosynthesis of TA *N*-oxides in engineered yeast. This figure illustrates a comparison of the native reaction (left) and target reactions (center, right) for senecionine *N*-oxygenase (SNO) and pyrrolizidine *N*-oxygenase (PNO), the top-scoring enzymes predicted by BridgIT to *N*-oxygenate TAs. KEGG reaction ID for native SNO/PNO transformation, ATLASx reaction IDs, and BridgIT closest KEGG reaction match for target transformations are shown. Values beside arrows for target reactions indicate BridgIT scores.

Figure 97 shows analytical validation of tropane *N*-oxide production in engineered yeast. This figure depicts LC-MS/MS MRM chromatograms for analysis of *de novo* TA *N*-oxide production by CSY1324 expressing SNO/PNOs, or authentic chemical standards. SNO/PNO candidates from *Tyria jacobaeae* (TjSNO), *Grammia geneura* (GgPNO), and *Zonocerus variegatus* (ZvPNO) were expressed from low-copy plasmids in CSY1324 and metabolites in supernatant were analyzed after 96 h growth in selective media buffered to pH 5.8 with 0.1 M potassium phosphate (for hyoscyamine *N*-oxide) or unbuffered (for scopolamine *N*-oxide). Chromatograms are representative of $n = 3$ biologically independent samples.

DEFINITIONS

Before describing exemplary embodiments in greater detail, the following definitions are set forth to illustrate and define the meaning and scope of the terms used in the description.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., John Wiley and Sons, New York (1994), and Hale & Markham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, N.Y. (1991) provide one of skill with the general meaning of many of the terms used herein. Still, certain terms are defined below for the sake of clarity and ease of reference.

It is noted that as used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the context clearly dictates otherwise. For example, the term “a primer” refers to one or more primers, i.e., a single primer and multiple primers. It is further noted that the claims are drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation.

As used herein, the terms “determining,” “measuring,” “assessing,” and “assaying” are used interchangeably and include both quantitative and qualitative determinations.

As used herein, the term “polypeptide” refers to a polymeric form of amino acids of any length, including peptides that range from 2-50 amino acids in length and polypeptides that are greater than 50 amino acids in length. The terms “polypeptide” and “protein” are used interchangeably herein. The term “polypeptide” includes polymers of coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones in which the conventional backbone has been replaced with non-naturally occurring or synthetic backbones. A polypeptide may be of any convenient length, e.g., 2 or more amino acids, such as 4 or more amino acids, 10 or

more amino acids, 20 or more amino acids, 50 or more amino acids, 100 or more amino acids, 300 or more amino acids, such as up to 500 or 1000 or more amino acids. “Peptides” may be 2 or more amino acids, such as 4 or more amino acids, 10 or more amino acids, 20 or more amino acids, such as up to 50 amino acids. In some embodiments, peptides are between 5 and 30 amino acids in length.

As used herein the term “isolated,” refers to an moiety of interest that is at least 60% free, at least 75% free, at least 90% free, at least 95% free, at least 98% free, and even at least 99% free from other components with which the moiety is associated with prior to purification.

As used herein, the term “encoded by” refers to a nucleic acid sequence which codes for a polypeptide sequence, wherein the polypeptide sequence or a portion thereof contains an amino acid sequence of 3 or more amino acids, such as 5 or more, 8 or more, 10 or more, 15 or more, or 20 or more amino acids from a polypeptide encoded by the nucleic acid sequence. Also encompassed by the term are polypeptide sequences that are immunologically identifiable with a polypeptide encoded by the sequence.

A “vector” is capable of transferring gene sequences to target cells. As used herein, the terms, “vector construct,” “expression vector,” and “gene transfer vector,” are used interchangeably to mean any nucleic acid construct capable of directing the expression of a gene of interest and which may transfer gene sequences to target cells, which is accomplished by genomic integration of all or a portion of the vector, or transient or inheritable maintenance of the vector as an extrachromosomal element. Thus, the term includes cloning, and expression vehicles, as well as integrating vectors.

An “expression cassette” includes any nucleic acid construct capable of directing the expression of a gene/coding sequence of interest, which is operably linked to a promoter of the expression cassette. Such cassette is constructed into a “vector,” “vector construct,” “expression vector,” or “gene transfer vector,” in order to transfer the expression cassette into target cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

A “plurality” contains at least 2 members. In certain cases, a plurality may have 10 or more, such as 100 or more, 1000 or more, 10,000 or more, 100,000 or more, 106 or more, 107 or more, 108 or more, or 109 or more members. In any embodiments, a plurality can have 2-20 members.

The term “tropane alkaloid product” is intended to refer to any molecule whose skeleton contains an 8-azabicyclo[3.2.1]octane core group comprising a cycloheptane ring and a nitrogen bridge connecting carbon atoms 1 and 5, wherein the 8-azabicyclo[3.2.1]octanyl group is covalently bonded to an acyl group by means of an ester linkage at the 3 position, and/or wherein the 8-azabicyclo[3.2.1]octanyl group is

functionalized with a hydroxyl group at the 3 position and one or more hydroxyl groups at the 2, 4, 5, 6, and/or 7 positions. Tropane alkaloid products include, but are not limited to, littorine, hyoscyamine, atropine, anisodamine, scopolamine, cocaine, and any other similar tropine/pseudotropine + acyl group natural or non-natural tropane alkaloids (e.g., calystegines).

The term “precursor of a tropane alkaloid product” is intended to refer to any molecule that can be biosynthesized by an organism from a carbon source and a nitrogen source and which can be converted to a tropane alkaloid product in one or more (e.g., one or two) biosynthetic steps; wherein the carbon source is a carbohydrate, a non-carbohydrate sugar, a sugar alcohol, a lipid, a fatty acid, or a substrate which is converted to one or more of the above carbon sources through a metabolic pathway; and wherein the nitrogen source is ammonia, urea, nitrate, nitrite, any amino acid excluding glutamic acid, arginine, ornithine, and citrulline, a peptide, a protein, or any substrate which is converted to one or more of the above nitrogen sources through a metabolic pathway.

The term “derivative of a tropane alkaloid product” is intended to refer to any molecule not naturally produced by an unmodified organism, wherein the skeleton of the molecule comprises a tropane alkaloid product and which differs from said tropane alkaloid product by the attachment and/or removal of functional groups without modification of the skeleton itself. As used herein, attachment and removal of functional groups includes, but is not limited to, hydroxylation and *N*-oxygenation (also referred to as *N*-oxidization), alkylation and *N*-alkylation, dealkylation and *N*-dealkylation, acetylation and *N*-acetylation, acylation and *N*-acylation, halogenation and *N*-halogenation.

The expression “TA biosynthetic pathway” is intended to refer to any collection of genetic elements, their polypeptide products, and/or small molecules which together constitute a series of chemical steps by which a precursor of a tropane alkaloid product can be converted to a tropane alkaloid product or to a derivative of a tropane alkaloid product. As used herein, the terms “TA biosynthetic pathway”, “TA pathway”, “engineered TA pathway”, “TA biosynthetic complement” and reasonable variations of such terms are used interchangeably. In some embodiments, a TA biosynthetic pathway or TA biosynthetic complement is intended to refer to any plurality of enzymes which enable one or more precursors of a TA product to be converted to one or more TA products or derivatives of TA products, such that a host cell which is described to express a TA biosynthetic pathway or complement can be understood to possess such capability.

A “small molecule” is a molecule with total molecular weight less than 1000 Daltons, and in this specification is generally interpreted to refer to such a molecule found within a biological system and which is not a polypeptide. As used herein, the terms “small molecule” and “metabolite” are interchangeable. Thus, the term includes both molecules

found within biological systems and molecules produced by such biological systems with total molecular weight less than 1000 Daltons and which are not polypeptides.

A “transporter” is a polypeptide capable of translocating a small molecule across a biological membrane. As used herein, the terms “small molecule transporter”, “metabolite transporter”, and “organic molecule transporter” are used interchangeably. The term includes, but is not limited to, polypeptides capable of translocating one or more different small molecules across one or more biological membranes and for which such translocation is powered by a gradient in small molecule concentration, a gradient in pH, the hydrolysis of another small molecule, or the simultaneous transport of another small molecule.

Numeric ranges are inclusive of the numbers defining the range.

The methods described herein include multiple steps. Each step may be performed after a predetermined amount of time has elapsed between steps, as desired. As such, the time between performing each step may be 1 second or more, 10 seconds or more, 30 seconds or more, 60 seconds or more, 5 minutes or more, 10 minutes or more, 60 minutes or more, and including 5 hours or more. In certain embodiments, each subsequent step is performed immediately after completion of the previous step. In other embodiments, a step may be performed after an incubation or waiting time after completion of the previous step, e.g., a few minutes to an overnight waiting time.

Other definitions of terms may appear throughout the specification.

DETAILED DESCRIPTION

Host cells that are engineered to produce TAs that are of interest, such as hyoscyamine and scopolamine, are provided. Host cells that are engineered to produce derivatives of TAs that are of interest, such as nortropane alkaloids and tropane *N*-oxides, are also provided. The host cells may have one or more engineered modifications selected from: a feedback inhibition alleviating mutation in an enzyme gene; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene; an inactivating mutation in an enzyme or transporter; and a heterologous coding sequence. Also provided are methods of producing a TA of interest using the host cells and compositions, e.g., kits, systems etc., that find use in methods of the invention.

Before the present invention is described in greater detail, it is to be understood that this invention is not limited to particular embodiments described, and as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between

the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Certain ranges are presented herein with numerical values being preceded by the term "about." The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein may also be used in the practice or testing of the present invention, representative illustrative methods and materials are now described.

All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present invention. Any recited method is carried out in the order of events recited or in any other order which is logically possible.

In further describing the subject invention, TA precursors of interest, TAs, and modifications of TAs, including derivatives of TAs, are described first in greater detail, followed by host cells for producing the same. Next, methods of interest in which the host cells find use are reviewed. Kits that may be used in practicing methods of the invention

are also described.

TROPANE ALKALOID (TA) PRECURSORS

As summarized above, host cells which produce tropane alkaloid precursors (TA precursors) are provided. The TA precursor may be any intermediate or precursor compound in a synthetic pathway (e.g., as described herein) that leads to the production of a TA of interest (e.g., as described herein). In some cases, the TA precursor has a structure that may be characterized as a TA or a derivative thereof. In certain cases, the TA precursor has a structure that may be characterized as a fragment of a TA. In some cases, the TA precursor is an early TA. As used herein, by "early TA" is meant an early intermediate in the synthesis of a TA of interest in a cell, where the early TA is produced by a host cell from a host cell feedstock or simple starting compound. In some cases, the early TA is a TA intermediate that is produced by the subject host cell solely from a host cell feedstock (e.g., a carbon and nutrient source) without the need for addition of a starting compound to the cells. The term early TA may refer to a precursor of a TA end product of interest whether or not the early TA may itself be characterized as a tropane alkaloid.

In some cases, the TA precursor is an early TA, such as a pre-tropine tropane alkaloid or a pre-littorine tropane alkaloid. As such, host cells which produce pre-tropine tropane alkaloids (pre-tropine TAs) and pre-littorine tropane alkaloids (pre-littorine TAs) are provided. Tropine is a major branch point intermediate of interest in the synthesis of downstream TAs via cell engineering efforts to produce end products such as medicinal TA products derived from littorine (Fig. 2). The subject host cells may produce TA precursors from simple and inexpensive starting materials that may find use in the production of tropine, littorine, and downstream TA end products.

As used herein, the terms "pre-esterification tropane alkaloid", "pre-esterification TA", and "pre- esterification TA precursor" are used interchangeably and refer to a biosynthetic precursor of littorine, cinnamoyltropine, or other product of acyl donor and acyl acceptor esterification, whether or not the structure of the esterification precursor itself is characterized as a tropane alkaloid. The term pre-esterification TA is meant to include biosynthetic precursors, intermediates and metabolites thereof, of any convenient member of a host cell biosynthetic pathway that may lead to esterification products such as littorine. In some cases, the pre-esterification TA includes a tropane alkaloid fragment, such as a tropine fragment, a phenylpropanoid fragment or a precursor or derivative thereof. In certain instances, the pre-esterification TA has a structure that may be characterized as a tropane alkaloid or a derivative thereof.

TA precursors of interest include, but are not limited to, tropine and phenyllactic acid (PLA), as well as tropine and PLA precursors, such as arginine, ornithine, agmatine, N-

carbamoylputrescine (NCP), putrescine, N-methylputrescine (NMP), 4-methylaminobutanal, N-methylpyrrolinium (NMPy), 4-(1-methyl-2-pyrrodinyl)-3-oxobutanoic acid (MPOB), tropinone, phenylalanine, prephenic acid, and phenylpyruvic acid (PPA). In some embodiments, the one or more TA precursors are tropine and PLA. In certain instances, the one or more TA precursors are tropine and a phenylpropanoid carboxylic acid other than PLA, such as cinnamic acid. Figures 1, 2, and 3 illustrate the biosynthesis of non-medicinal, medicinal, and non-natural TAs respectively from various TA and non-TA precursor molecules.

Synthetic pathways to a TA precursor may be generated in the host cells, and may start with any convenient starting compound(s) or materials. Figures 1-4 illustrate a synthetic pathway of interest to TA precursors starting from amino acids. The starting material may be non-naturally occurring or the starting material may be naturally occurring in the host cell. Any convenient compounds and materials may be used as the starting material, based upon the synthetic pathway present in the host cell. The source of the starting material may be from the host cell itself, e.g., arginine or phenylalanine, or the starting material may be added or supplemented to the host cell from an outside source. As such, in some cases, the starting compound refers to a compound in a synthetic pathway of the cell that is added to the host cell from an outside source that is not part of a growth feedstock or cell growth media. Starting compounds of interest include, but are not limited to, N-methylputrescine, 4-methylaminobutanal, tropinone, tropine, PLA, cinnamic acid, as well as any of the compounds shown in Figures 1-4. For example, if the host cells are growing in liquid culture, the cell media may be supplemented with the starting material, which is transported into the cells and converted into the desired products by the cell. Starting materials of interest include, but are not limited to, inexpensive feedstocks and simple precursor molecules. In some cases, the host cell utilizes a feedstock including a simple carbon source as the starting material, which the host cell utilizes to produce compounds of the synthetic pathway of the cell. The host cell growth feedstock may include one or more components, such as a carbon source such as cellulose, starch, free sugars and a nitrogen source, such as ammonium salts or inexpensive amino acids. In some cases, a growth feedstock that finds use as a starting material may be derived from a sustainable source, such as biomass grown on marginal land, including switchgrass and algae, or biomass waste products from other industrial or farming activities.

TROPANE ALKALOIDS (TAS)

As summarized above, host cells which produce tropane alkaloids (TAs) of interest are provided. In some embodiments, the engineered strains of the invention will provide a platform for producing tropane alkaloids of interest and modifications thereof across several

classes including, but not limited to, medicinal TAs such as those derived from tropine and PLA; non-medicinal TAs such as those derived from tropinone, pseudotropine, norpseudotropine, or those derived from the addition and/or removal of functional groups to/from TAs; and non-natural TAs such as those derived from the esterification of TA precursors (e.g., acyl donor and acyl acceptor compounds) other than tropine and PLA, or those derived from the addition and/or removal of functional groups to/from TAs. Each of these classes is meant to include biosynthetic precursors, intermediates, and metabolites thereof, of any convenient member of a host cell biosynthetic pathway that may lead to a member of the class. Non-limiting examples of compounds are given below for each of these classes. In some embodiments, the structure of a given example may or may not be characterized itself as a tropane alkaloid. The present chemical entities are meant to include all possible isomers, including single enantiomers, racemic mixtures, optically pure forms, mixtures of diastereomers and intermediate mixtures.

Medicinal TAs may include, but are not limited to, littorine, hyoscyamine, atropine, anisodamine, scopolamine, and derivatives thereof that are naturally produced by plants.

Non-medicinal TAs may include, but are not limited to, calystegines, cocaine, norlittorine, norhyoscyamine, noranisodamine, norscopolamine, littorine *N*-oxide, hyoscyamine *N*-oxide, anisodamine *N*-oxide, scopolamine *N*-oxide, *N*-methyllittorine, *N*-methylhyoscyamine, *N*-methylanisodamine, *N*-methylscopolamine, and derivatives thereof that are naturally produced by plants or by TA metabolism in non-plant organisms.

Non-natural TAs may include, but are not limited to, cinnamoyltropine, cinnamoyl-3β-tropine, coumaroyltropine, coumaroyl-3β-tropine, benzoyltropine, benzoyl-3β-tropine, caffeoyltropine, caffeoyl-3β-tropine, feruloyltropine, feruloyl-3β-tropine, sinapoyltropine, and sinapoyl-3β-tropine.

MODIFICATIONS OF TAS INCLUDING DERIVATIVES

As summarized above, host cells which produce modified derivatives of TAs of interest are provided. In some embodiments, the engineered strains of the invention will provide a platform for derivatizing TAs of interest, including derivatizing TA precursors, medicinal TAs, non-medicinal TAs, and non-natural TAs which are produced by engineered host cells or which are fed to engineered host cells in the growth media.

As used herein, the terms “derivatization”, “functionalization”, “modification by derivatization”, and “modification by functionalization” refer to the modification of TAs or of TA precursors via the attachment and/or removal of functional groups without modification of the skeleton itself. As used herein, attachment and removal of functional groups includes, but is not limited to, hydroxylation and *N*-oxygenation (also referred to as *N*-oxidization), alkylation and *N*-alkylation, dealkylation and *N*-dealkylation, acetylation and *N*-acetylation,

acylation and *N*-acylation, halogenation and *N*-halogenation.

In some embodiments of the invention, derivatization of TAs of interest may be achieved enzymatically by feeding pre-functionalized TA precursors, for example halogenated or alkylated amino acids, to host cells engineered to uptake and then convert fed TA precursors into TAs of interest. In other embodiments of the invention, derivatization of TAs of interest may be achieved enzymatically by engineering host cells to express enzymes which possess the desired activity in attaching a functional group to a target TA, in addition to the enzymes and cellular modifications required to produce the unmodified TA. In other embodiments of the invention, derivatization of TAs of interest may be achieved enzymatically by treating unmodified TAs produced by engineered host cells with purified enzymes capable of attaching desired functional groups, or with crude lysate of host cells engineered to express enzymes that have the desired derivatizing activity. In other embodiments of the invention, derivatization of TAs of interest may be achieved non-enzymatically by treating unmodified TAs produced by engineered host cells with chemical agents with attach desired functional groups.

Modified derivatives of TAs include, but are not limited to, p-hydroxyatropine, p-hydroxyhyoscyamine, p-fluorohyoscyamine, p-chlorohyoscyamine, p-bromohyoscyamine, p-fluoroscopolamine, p-chloropscopolamine, p-bromoscopolamine, *N*-methyllittorine, *N*-butyllittorine, *N*-methylhyoscyamine, *N*-butylhyoscyamine, *N*-methylanisodamine, *N*-butylanisodamine, *N*-methylscopolamine, *N*-butylscopolamine, *N*-acetyllittorine, N-acetylhyoscyamine, *N*-acetylanisodamine, *N*-acetylscopolamine, *N*-formyllittorine, *N*-formylhyoscyamine, *N*-formylanisodamine, *N*-formylscopolamine, norlittorine, norhyoscyamine, noranisodamine, norscopolamine, littorine *N*-oxide, hyoscyamine *N*-oxide, anisodamine *N*-oxide, and scopolamine *N*-oxide.

HOST CELLS

As summarized above, one aspect of the invention is a host cell that produces one or more TAs of interest. Any convenient cells may be utilized in the subject host cells and methods. In some cases, the host cells are non-plant cells. In some instances, the host cells may be characterized as microbial cells. In certain cases, the host cells are insect cells, mammalian cells, bacterial cells, or fungal cells. Any convenient type of host cell may be utilized in producing the subject TA-producing cells, see, e.g., US2008/0176754 now published as U.S. Patent No. 8,975,063, US2014/0273109 and WO2014/143744); the disclosures of which are incorporated by reference in their entirety. Host cells of interest include, but are not limited to, bacterial cells, such as *Bacillus subtilis*, *Escherichia coli, Streptomyces, Anabaena, Arthrobacter, Acetobacter, Acetobacterium, Bacillus, Bifidobacterium, Brachybacterium, Brevibacterium, Carnobacterium, Clostridium,*

Corynebacterium, Enterobacter, Escherichia, Gluconacetobacter, Gluconobacter, Hafnia, Halomonas, Klebsiella, Kocuria, Lactobacillus, Leucononstoc, Macrococcus, Methylomonas, Methylobacter, Methylocella, Methylococcus, Microbacterium, Micrococcus, Microcystis, Moorella, Oenococcus, Pediococcus, Prochlorococcus, Propionibacterium, Proteus, Pseudoalteromonas, Pseudomonas, Psychrobacter, Rhodobacter, Rhodococcus, Rhodopseudomonas, Serratia, Staphylococcus, Streptococcus, Streptomyces, Synechococcus, Synechocystis, Tetragenococcus, Weissella, Zymomonas, and *Salmonella typhimuium* cells, insect cells such as *Drosophila melanogaster* S2 and *Spodoptera frugiperda* Sf9 cells, and yeast cells such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Yarrowia lipolytica*, *Candida albicans, Aspergillus spp., Rhizopus spp., Penicillium spp., and Trichoderma reesei* cells. In some embodiments, the host cells are yeast cells or *E. coli* cells. In some cases, the host cell is a yeast cell. In some instances the host cell is from a strain of yeast engineered to produce a TA of interest. Any of the host cells described in US2008/0176754 now published as U.S. Patent No. 8,975,063, US2014/0273109 and WO2014/143744, may be adapted for use in the subject cells and methods. In certain embodiments, the yeast cells may be of the species *Saccharomyces cerevisiae* (*S. cerevisiae*). In certain embodiments, the yeast cells may be of the species *Schizosaccharomyces pombe*. In certain embodiments, the yeast cells may be of the species *Pichia pastoris*. Yeast is of interest as a host cell because cytochrome P450 proteins, which are involved in some biosynthetic pathways of interest, are able to fold properly into the endoplasmic reticulum membrane so that their activity is maintained.

Yeast strains of interest that find use in the invention include, but are not limited to, CEN.PK (Genotype: *MAT*a/α *ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112 his3* Δ*1/his3* Δ*1 MAL2-8C/MAL2-8C SUC2/SUC2*), S288C, W303, D273-10B, X2180, A364A, ∑1278B, AB972, SK1, and FL100. In certain cases, the yeast strain is any of S288C (MATα; SUC2 mal mel gal2 CUP1 flo1 flo8-1 hap1), BY4741 (MATα; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0), BY4742 (MATα; his3Δ1; leu2Δ0; lys2Δ0; ura3Δ0), BY4743 (MATa/MATα; his3Δ1/his3Δ1; leu2Δ0/leu2Δ0; met15Δ0/MET15; LYS2/lys2Δ0; ura3Δ0/ura3Δ0), and WAT11 or W(R), derivatives of the W303-B strain (MATa; ade2-1; his3-11, -15; leu2-3,-112; ura3-1; canR; cyr+) which express the *Arabidopsis thaliana* NADPH-P450 reductase ATR1 and the yeast NADPH-P450 reductase CPR1, respectively. In another embodiment, the yeast cell is W303alpha (MATα; his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1). The identity and genotype of additional yeast strains of interest may be found at EUROSCARF (web.uni-frankfurt.de/fb15/mikro/euroscarf/col_index.html).

In some instances, the host cell is a fungal cell. In certain embodiments, the fungal cells may be of the Aspergillus species and strains include *Aspergillus niger* (ATCC 1015,

ATCC 9029, CBS 513.88), *Aspergillus oryzae* (ATCC 56747, RIB40), *Aspergillus terreus* (NIH 2624, ATCC 20542) and *Aspergillus nidulans* (FGSC A4).

In certain embodiments, heterologous coding sequences may be codon optimized for expression in Aspergillus sp. and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from phosphoglycerate kinase promoter (PGK), MbfA promoter, cytochrome c oxidase subunit promoter (CoxA), SrpB promoter, TvdA promoter, malate dehydrogenase promoter (MdhA), beta-mannosidase promoter (ManB). In certain embodiments, a terminator may be selected from glucoamylase terminator (GlaA) or TrpC terminator. In certain embodiments, the expression cassette consisting of a promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome of the host. In certain embodiments, selection of cells maintaining the plasmid or integration cassette may be performed with antibiotic selection such as hygromycin or nitrogen source utilization, such as using acetamide as a sole nitrogen source. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as protoplast transformation, lithium acetate, or electroporation. In certain embodiments, cells may be cultured in liquid ME or solid MEA (3 % malt extract, 0.5 % peptone, and ±1.5 % agar) or in Vogel's minimal medium with or without selection.

In some instances, the host cell is a bacterial cell. The bacterial cell may be selected from any bacterial genus. Examples of genera from which the bacterial cell may come include *Anabaena, Arthrobacter, Acetobacter, Acetobacterium, Bacillus, Bifidobacterium, Brachybacterium, Brevibacterium, Carnobacterium, Clostridium, Corynebacterium, Enterobacter, Escherichia, Gluconacetobacter, Gluconobacter, Hafnia, Halomonas, Klebsiella, Kocuria, Lactobacillus, Leuconostoc, Macrococcus, Methylomonas, Methylobacter, Methylocella, Methylococcus, Microbacterium, Micrococcus, Microcystis, Moorella, Oenococcus, Pediococcus, Prochlorococcus, Propionibacterium, Proteus, Pseudoalteromonas, Pseudomonas, Psychrobacter, Rhodobacter, Rhodococcus, Rhodopseudomonas, Serratia, Staphylococcus, Streptococcus, Streptomyces, Synechococcus, Synechocystis, Tetragenococcus, Weissella,* and *Zymomonas*. Examples of bacterial species which may be used with the methods of this disclosure include *Arthrobacter nicotianae, Acetobacter aceti, Arthrobacter arilaitensis, Bacillus cereus, Bacillus coagulans, Bacillus licheniformis, Bacillus pumilus, Bacillus sphaericus, Bacillus stearothermophilus, Bacillus subtilis, Bifidobacterium adolescentis, Brachybacterium tyrofermentans, Brevibacterium linens, Carnobacterium divergens, Corynebacterium flavescens, Enterococcus faecium, Gluconacetobacter europaeus, Gluconacetobacter johannae, Gluconobacter oxydans, Hafnia alvei, Halomonas elongata, Kocuria rhizophila, Lactobacillus acidifarinae, Lactobacillus jensenii, Lactococcus lactis, Lactobacillus*

yamanashiensis, Leuconostoc citreum, Macrococcus caseolyticus, Microbacterium foliorum, Micrococcus lylae, Oenococcus oeni, Pediococcus acidilactici, Propionibacterium acidipropionici, Proteus vulgaris, Pseudomonas fluorescens, Psychrobacter celer, Staphylococcus condimenti, Streptococcus thermophilus, Streptomyces griseus, Tetragenococcus halophilus, Weissella cibaria, Weissella koreensis, Zymomonas mobilis , Corynebacterium glutamicum, Bifidobacterium bifidum/breve/longum, Streptomyces lividans, Streptomyces coelicolor, Lactobacillus plantarum, Lactobacillus sakei, Lactobacillus casei, Pseudoalteromonas citrea, Pseudomonas putida, Clostridium ljungdahlii/aceticum/acetobutylicum/beijerinckii/butyricum, and *Moorella themocellum/thermoacetica.*

In certain embodiments, the bacterial cells may be of a strain of *Escherichia coli*. In certain embodiments, the strain of *E. coli* may be selected from BL21, DH5α, XL1-Blue, HB101, BL21, and K12. In certain embodiments, heterologous coding sequences may be codon optimized for expression in *E. coli* and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from T7 promoter, tac promoter, trc promoter, tetracycline-inducible promoter (tet), lac operon promoter (lac), lacO1 promoter. In certain embodiments, the expression cassette consisting of a promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome. In certain embodiments, the plasmid is selected from pUC19 or pBAD. In certain embodiments, selection of cells maintaining the plasmid or integration cassette may be performed with antibiotic selection such as kanamycin, chloramphenicol, streptomycin, spectinomycin, gentamycin, erythromycin or ampicillin. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as conjugation, heat shock chemical transformation, or electroporation. In certain embodiments, cells may be cultured in liquid Luria–Bertani (LB) media at about 37°C with or without antibiotics.

In certain embodiments, the bacterial cells may be a strain of *Bacillus subtilis*. In certain embodiments, the strain of *B. subtilis* may be selected from 1779, GP25, RO-NN-1, 168, BSn5, BEST195, 1A382, and 62178. In certain embodiments, heterologous coding sequences may be codon optimized for expression in Bacillus sp. and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from grac promoter, p43 promoter, or trnQ promoter. In certain embodiments, the expression cassette consisting of the promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome. In certain embodiments, the plasmid is selected from pHP13 pE194, pC194, pHT01, or pHT43. In certain embodiments, integrating vectors such as pDG364 or pDG1730 may be used to integrate the expression cassette into the genome. In certain embodiments, selection of cells maintaining the

plasmid or integration cassette may be performed with antibiotic selection such as erythromycin, kanamycin, tetracycline, and spectinomycin. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as natural competence, heat shock, or chemical transformation. In certain embodiments, cells may be cultured in liquid Luria–Bertani (LB) media at 37°C or M9 medium plus glucose and tryptophan.

GENETIC MODIFICATIONS TO HOST CELLS

The host cells may be engineered to include one or more modifications (such as two or more, three or more, four or more, five or more, or even more modifications) that provide for the production of TAs of interest. In some cases, by modification is meant a genetic modification, such as a mutation, addition, or deletion of a gene or fragment thereof, or transcription regulation of a gene or fragment thereof. In some cases, the one or more (such as two or more, three or more, or four or more) modifications is selected from: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; an inactivating mutation in an enzyme or transporter native to the cell; a heterologous coding sequence that encodes an enzyme; and a heterologous coding sequence that encodes a protein which modifies the sub-cellular trafficking and/or localization of an enzyme or a metabolite, such as a transporter protein. A cell that includes one or more modifications may be referred to as a modified cell.

A modified cell may overproduce one or more precursor TA, TA, or modified TA molecules. By overproduce is meant that the cell has an improved or increased production of a TA molecule of interest relative to a control cell (e.g., an unmodified cell). By improved or increased production is meant both the production of some amount of the TA of interest where the control has no TA precursor production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some TA of interest production.

In some cases, the host cell is capable of producing an increased amount of putrescine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of putrescine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of N-

methylpyrrolinium relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of N-methylpyrrolinium is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of tropine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of tropine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of phenylpyruvic acid relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of phenylpyruvic acid is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of phenyllactic acid relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of phenyllactic acid is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of littorine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of littorine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of hyoscyamine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of hyoscyamine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about

100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of scopolamine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of scopolamine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of a nortropane alkaloid, such as norhyoscyamine or norscopolamine, relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of the nortropane alkaloid is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some cases, the host cell is capable of producing an increased amount of a tropane *N*-oxide, such as hyoscyamine *N*-oxide or scopolamine *N*-oxide, relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In certain instances, the increased amount of the tropane *N*-oxide is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, 2-fold or more, 5-fold or more, or even 10-fold or more relative to the control host cell.

In some embodiments, the host cell is capable of producing a 10% or more yield of tropine from a starting compound such as arginine, such as 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of tropine from a starting compound.

In some embodiments, the host cell is capable of producing a 10% or more yield of phenyllactic acid from a starting compound such as phenylalanine, such as 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of phenyllactic acid from a starting compound.

In some embodiments, the host cell is capable of producing a 10% or more yield of hyoscyamine from a starting compound such as arginine or phenylalanine, such as 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of hyoscyamine from a starting compound.

In some embodiments, the host cell is capable of producing a 10% or more yield of scopolamine from a starting compound such as arginine or phenylalanine, such as 20% or

more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of scopolamine from a starting compound.

In some embodiments, the host cell is capable of producing a 10% or more yield of a nortropane alkaloid, such as norhyoscyamine or norscopolamine, from a starting compound such as arginine or phenylalanine, such as 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of nortropane alkaloid(s) from a starting compound.

In some embodiments, the host cell is capable of producing a 10% or more yield of a tropane *N*-oxide, such as hyoscyamine *N*-oxide or scopolamine *N*-oxide, from a starting compound such as arginine or phenylalanine, such as 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, or even 90% or more yield of tropane *N*-oxide from a starting compound.

In some embodiments, the host cell overproduces one or more TA of interest molecules selected from the group consisting of arginine, ornithine, agmatine, putrescine, N-methylputrescine, 4-methylaminobutanal, N-methylpyrrolinium, 4-(1-methyl-2-pyrrodinyl)-3-oxobutanoic acid, tropinone, tropine, nortropine, tropine *N*-oxide, phenylalanine, prephenic acid, phenylpyruvic acid, phenyllactic acid, glucose-1-*O*-phenyllactate, littorine, norlittorine, littorine *N*-oxide, hyoscyamine aldehyde, hyoscyamine, norhyoscyamine, hyoscyamine *N*-oxide, anisodamine, noranisodamine, anisodamine *N*-oxide, scopolamine, norscopolamine, and scopolamine *N*-oxide.

Any convenient combinations of the one or more modifications may be included in the subject host cells. In some cases, two or more (such as two or more, three or more, or four or more) different types of modifications are included. In certain instances, two or more (such as three or more, four or more, five or more, or even more) distinct modifications of the same type of modification are included in the subject cells.

In some embodiments of the host cell, when the cell includes one or more heterologous coding sequences that encode one or more enzymes or transporters, it includes at least one additional modification selected from the group consisting of: a feedback inhibition alleviating mutations in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; and an inactivating mutation in an enzyme or transporter native to the cell. In certain embodiments of the host cell, when the cell includes one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell, it includes a least one additional modification selected from the group consisting of: a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; an inactivating mutation in an enzyme or transporter native to the cell; and a heterologous coding sequence that encode an enzyme or a transporter. In some embodiments of the host

cell, when the cell includes one or more transcriptional modulation modifications of one or more biosynthetic enzyme or transporter genes native to the cell, it includes at least one additional modification selected from the group consisting of: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; an inactivating mutation in an enzyme or transporter native to the cell; a heterologous coding sequence that encodes an enzyme; and a heterologous coding sequence that encodes a protein which modifies the sub-cellular trafficking and/or localization of an enzyme or a metabolite, such as a transporter. In certain instances of the host cell, when the cell includes one or more inactivating mutations in one or more enzymes or transporters native to the cell, it includes at least one additional modification selected from the group consisting of: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme or transporter gene native to the cell; a heterologous coding sequence that encodes an enzyme; and a heterologous coding sequence that encodes a protein which modifies the sub-cellular trafficking and/or localization of an enzyme or a metabolite, such as a transporter.

In certain embodiments of the host cell, the cell includes one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell; and one or more transcriptional modulation modifications of one or more biosynthetic enzyme or transporter genes native to the cell. In certain embodiments of the host cell, the cell includes one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell; and one or more inactivating mutations in an enzyme or transporter native to the cell. In certain embodiments of the host cell, the cell includes one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell; and one or more heterologous coding sequences. In some embodiments, the host cell includes one or more modifications (e.g., as described herein) that include one or more of the genes of interest described in Table 1.

Feedback Inhibition Alleviating Mutations

In some instances, the host cells are cells that include one or more feedback inhibition alleviating mutations (such as two or more, three or more, four or more, five or more, or even more) in one or more biosynthetic enzyme genes of the cell. In some cases, the one or more biosynthetic enzyme genes are native to the cell (e.g., is present in an unmodified cell). As used herein, the term “feedback inhibition alleviating mutation” refers to a mutation that alleviates a feedback inhibition control mechanism of a host cell. Feedback inhibition is a control mechanism of the cell in which an enzyme in the synthetic pathway of a regulated compound is inhibited when that compound has accumulated to a certain level, thereby balancing the amount of the compound in the cell. In some instances, the one or

more feedback inhibition alleviating mutations is in an enzyme described in a biosynthetic pathway of Figures 1-4 or in the schematic of Figure 8. A mutation that alleviates feedback inhibition reduces the inhibition of a regulated enzyme in the cell of interest relative to a control cell and provides for an increased level of the regulated compound or a downstream biosynthetic product thereof. In some cases, by alleviating inhibition of the regulated enzyme is meant that the IC_{50} of inhibition is increased by 2-fold or more, such as by 3-fold or more, 5-fold or more, 10-fold or more, 30-fold or more, 100-fold or more, 300-fold or more, 1000-fold or more, or even more. By increased level is meant a level that is 110% or more of that of the regulated compound in a control cell or a downstream product thereof, such as 120% or more, 130% or more, 140% or more, 150% or more, 160% or more, 170% or more, 180% or more, 190% or more, or 200% or more, such as at least 3-fold or more, at least 5-fold or more, at least 10-fold or more or even more of the regulated compound in the host cell or a downstream product thereof.

A variety of feedback inhibition control mechanisms and biosynthetic enzymes native to the host cell that are directed to regulation of levels of TA precursors may be targeted for alleviation in the host cell. The host cell may include one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell. The mutation may be located in any convenient biosynthetic enzyme genes native to the host cell where the biosynthetic enzyme is subject to regulatory control. In some embodiments, the one or more biosynthetic enzyme genes encode one or more enzymes selected from an ornithine decarboxylase (ODC), an ornithine decarboxylase antizyme, and a putrescine N-methyltransferase. In some embodiments, the one or more biosynthetic enzyme genes encode an ornithine decarboxylase. In some instances, the one or more biosynthetic enzyme genes encode an ornithine decarboxylase antizyme. In some embodiments, the one or more biosynthetic enzyme genes encode a putrescine N-methyltransferase. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene selected from SPE1, OAZ1, and PMT. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is SPE1. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is OAZ1. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is PMT. In some embodiments, the host cell includes one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes such as one of those genes described in Table 1.

Any convenient numbers and types of mutations may be utilized to alleviate a feedback inhibition control mechanism. As used herein, the term "mutation" refers to a deletion, insertion, or substitution of an amino acid(s) residue or nucleotide(s) residue

relative to a reference sequence or motif. The mutation may be incorporated as a directed mutation to the native gene at the original locus. In some cases, the mutation may be incorporated as an additional copy of the gene introduced as a genetic integration at a separate locus, or as an additional copy on an episomal vector such as a 2μ or centromeric plasmid. In certain instances, the feedback inhibited copy of the enzyme is under the native cell transcriptional regulation. In some instances, feedback inhibited copy of the enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

In certain embodiments, the host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more feedback inhibition alleviating mutations, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the host cell.

Transcriptional Modulation Modifications

The host cells may include one or more transcriptional modulation modifications (such as two or more, three or more, four or more, five or more, or even more modifications) of one or more biosynthetic enzyme or transporter genes of the cell. In some cases, the one or more biosynthetic enzyme or transporter genes are native to the cell. Any convenient biosynthetic enzyme or transporter genes of the cell may be targeted for transcription modulation. By transcription modulation is meant that the expression of a gene of interest in a modified cell is modulated, e.g., increased or decreased, enhanced or repressed, relative to a control cell (e.g., an unmodified cell). In some cases, transcriptional modulation of the gene of interest includes increasing or enhancing expression. By increasing or enhancing expression is meant that the expression level of the gene of interest is increased by 2-fold or more, such as by 5-fold or more and sometimes by 25-, 50-, or 100-fold or more and in certain embodiments 300-fold or more or higher, as compared to a control, i.e., expression in the same cell not modified (e.g., by using any convenient gene expression assay). Alternatively, in cases where expression of the gene of interest in a cell is so low that it is undetectable, the expression level of the gene of interest is considered to be increased if expression is increased to a level that is easily detectable. In certain instances, transcriptional modulation of the gene of interest includes decreasing or repressing expression. By decreasing or repressing expression is meant that the expression level of the gene of interest is decreased by 2-fold or more, such as by 5-fold or more and sometimes by 25-, 50-, or 100-fold or more and in certain embodiments 300-fold or more or higher, as compared to a control. In some cases, expression is decreased to a level that is

undetectable. Modifications of host cell processes of interest that may be adapted for use in the subject host cells are described in U.S. Publication No. 20140273109 (14/211,611) by Smolke et al. and in World Publication No. WO2020185626A1 by Smolke et al., the disclosures of which are herein incorporated by reference in their entirety.

Any convenient biosynthetic enzyme genes may be transcriptionally modulated, and include but are not limited to, those biosynthetic enzymes described in Figs. 1-3 and Fig. 75, such as ARG2, CAR1, SPE1, FMS1, PHA2, ARO8, ARO9, and UGP1. In some instances, the one or more biosynthetic enzyme genes is selected from ARG2, CAR1, SPE1, and FMS1. In some cases, the one or more biosynthetic enzyme genes is ARG2. In certain instances, the one or more biosynthetic enzyme genes is CAR1. In some embodiments, the one or more biosynthetic enzyme genes is SPE1. In some embodiments, the one or more biosynthetic enzyme genes is FMS1. Any convenient transporter genes may also be transcriptionally modulated, and include but are not limited to, transporters described in Tables 1, 4, and 6, such as SNQ2, PDR1, PDR3, PDR5, PDR7, PDR8, PDR9, PDR10, PDR11, PDR12, PDR15, PDR16, PDR17, PDR18, YCF1, YNR070W, ADP1, VMR1, NFT1, BPT1, YBT1, YOR1, TPO1, TPO2, TPO3, TPO4, TPO5, FLR1, QDR1, QDR2, QDR3, YHK8, HOL1, DTR1, and AQR1. In some embodiments, the host cell includes one or more transcriptional modulation modifications to one or more genes such as one of those genes described in Table 1. In some embodiments, the host cell includes one or more transcriptional modulation modifications to one or more genes such as one of those genes described in a biosynthetic pathway of one of Figures 1-4 or in the schematic of Figure 8.

In some embodiments, the transcriptional modulation modification includes substitution of a strong promoter for a native promoter of the one or more biosynthetic enzyme or transporter genes or the expression of an additional copy(ies) of the gene or genes under the control of a strong promoter. The promoters driving expression of the genes of interest may be constitutive promoters or inducible promoters, provided that the promoters may be active in the host cells. The genes of interest may be expressed from their native promoters, or non-native promoters may be used. Although not a requirement, such promoters should be medium to high strength in the host in which they are used. Promoters may be regulated or constitutive. In some embodiments, promoters that are not glucose repressed, or repressed only mildly by the presence of glucose in the culture medium, are used. There are numerous suitable promoters, examples of which include promoters of glycolytic genes such as the promoter of the *B. subtilis* tsr gene (encoding fructose biphosphate aldolase) or GAPDH promoter from yeast *S. cerevisiae* (coding for glyceraldehyde-phosphate dehydrogenase) (Bitter G. A., *Meth. Enzymol.* 152:673 684 (1987)). Other strong promoters of interest include, but are not limited to, the ADHI

promoter of baker's yeast (Ruohonen L., et al, *J. Biotechnol.* 39:193 203 (1995)), the phosphate-starvation induced promoters such as the PHO5 promoter of yeast (Hinnen, A., et al, in *Yeast Genetic Engineering,* Barr, P. J., et al. eds, Butterworths (1989), the alkaline phosphatase promoter from *B. licheniformis* (Lee. J. W. K., et al., *J. Gen. Microbiol.* 137:1127 1133 (1991)), GPD1 and TEF1. Yeast promoters of interest include, but are not limited to, inducible promoters such as Gal1-10, Gal1, GalL, GalS, repressible promoter Met25, tetO, and constitutive promoters such as glyceraldehyde 3-phosphate dehydrogenase promoter (GPD), alcohol dehydrogenase promoter (ADH), translation-elongation factor-1-alpha promoter (TEF), cytochrome c-oxidase promoter (CYC1), MRP7 promoter, phosphoglycerate kinase (PGK), triose phosphate isomerase (TPI), etc. In some instances, the strong promoter is GPD1. In certain instances, the strong promoter is TEF1. Autonomously replicating yeast expression vectors containing promoters inducible by hormones such as glucocorticoids, steroids, and thyroid hormones are also known and include, but are not limited to, the glucorticoid responsive element (GRE) and thyroid hormone responsive element (TRE), see e.g., those promoters described in U.S. Pat. No. 7,045,290. Vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of genes of interest. It is understood that any convenient promoters specific to the host cell may be selected, e.g., *E. coli.* In some cases, promoter selection may be used to optimize transcription, and hence, enzyme levels to maximize production while minimizing energy resources.

Inactivating Mutations

The host cells may include one or more inactivating mutations to an enzyme or transporter of the cell (such as two or more, three or more, four or more, five or more, or even more). The inclusion of one or more inactivating mutations may modify the flux of a synthetic pathway of a host cell or the spatial localization of a synthetic pathway of a host cell to increase the levels of a TA of interest or a desirable enzyme or precursor leading to the same. In some cases, the one or more inactivating mutations are to an enzyme or a transporter native to the cell. Figure 8 illustrates the native regulatory mechanisms in yeast which act on polyamine production pathways and Figure 9 shows the effects of disruptions to these native regulatory systems on production of putrescine. As used herein, by “inactivating mutation” is meant one or more mutations to a gene or regulatory DNA sequence of the cell, where the mutation(s) inactivates a biological activity of the protein expressed by that gene of interest. In some cases, the gene is native to the cell. In some instances, the gene encodes an enzyme or a transporter that is inactivated and is part of or

connected to the synthetic pathway of a TA of interest produced by the host cell. In some instances, an inactivating mutation is located in a regulatory DNA sequence that controls a gene of interest. In certain cases, the inactivating mutation is to a promoter of a gene. Any convenient mutations (e.g., as described herein) may be utilized to inactivate a gene or regulatory DNA sequence of interest. By “inactivated” or “inactivates” is meant that a biological activity of the protein expressed by the mutated gene is reduced by 10% or more, such as by 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 90% or more, 95% or more, 97% or more, or 99% or more, relative to a control protein expressed by a non-mutated control gene. In some cases, the protein is an enzyme and the inactivating mutation reduces the catalytic activity of the enzyme. In other cases, the protein is a transporter and the inactivating mutation reduces the transport activity of the transporter.

In some embodiments, the cell includes an inactivating mutation in an enzyme or a transporter native to the cell. Any convenient enzymes or transporter may be targeted for inactivation. Enzymes of interest include, but are not limited to, those enzymes described in Figures 1-4, 8, 11, 22, and 41 whose action in the biosynthetic pathways of the host cell tends to reduce the levels of a TA of interest. In some cases, the enzyme has methylthioadenosine phosphorylase activity. In certain embodiments, the enzyme that includes an inactivating mutation is MEU1 (see e.g., Figures 8, 9, and 13). In some cases, the enzyme has ornithine decarboxylase antizyme activity. In certain embodiments, the enzyme that includes an inactivating mutation is OAZ1. In some cases, the enzyme has spermidine synthase activity. In certain embodiments, the enzyme that includes an inactivating mutation is SPE3. In some cases, the enzyme has spermine synthase activity. In some embodiments, the enzyme that includes an inactivating mutation is SPE4. In some cases, the protein is a membrane transporter with polyamine export activity. In certain embodiments, the enzyme or protein that includes an inactivating mutation is TPO5. In some cases, the protein is a membrane transporter with alkaloid export activity. In certain embodiments, the enzyme or protein that includes an inactivating mutation is a multidrug resistance transporter including but not limited to PDR1, PDR5, PDR10, PDR11, PDR12, PDR15, SNQ2, YOR1, FLR1, QDR1, QDR2, QDR3, YHK8, HOL1, DTR1, and AQR1. In some cases, the enzyme has phenylacrylic acid decarboxylase activity. In certain embodiments, the enzyme that includes an inactivating mutation is PAD1. In some cases, the enzyme has alcohol dehydrogenase activity. In some embodiments, the enzyme that includes an inactivating mutation is selected from ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, and SFA1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH2. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH3. In certain embodiments, the enzyme that includes an inactivating

mutation(s) is ADH4. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH5. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH6. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH7. In some cases, the enzyme has aldehyde oxidoreductase activity. In certain embodiments, the enzyme that includes an inactivating mutation is selected from HFD1, ALD2, ALD3, ALD4, ALD5, and ALD6. In certain embodiments, the enzyme that includes and inactivating mutation(s) is HFD1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD2. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD3. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD4. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD5. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD6. In some cases, the enzyme has glucosidase activity. In certain embodiments, the enzyme that includes an inactivating mutation is selected from EXG1, SPR1, and EGH1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is EXG1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is SPR1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is EGH1. In some embodiments, the host cell includes one or more inactivating mutations to one or more genes described in Table 1.

Methods for performing TA acyl transfer reactions using functional expression of acyltransferases in non-plant hosts

Some methods, processes, and systems provided herein describe the concerted reaction of one or more TA precursors comprising an acyl donor group with one or more TA precursors comprising an acyl acceptor group to produce one or more TAs within a non-plant cell (hereafter referred to as TA acyl transfer reactions). Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the TA acyl transfer reaction is a key step in the conversion of a substrate to a diverse range of alkaloids. In some examples, the TA acyl transfer reaction comprises a condensation reaction.

In some examples, the TA acyl transfer may involve at least one condensation reaction. In some cases, at least one of the condensation reactions is carried out in the presence of an enzyme. In some cases, at least one of the condensation reactions is catalyzed by an enzyme. In some cases, at least one enzyme is useful to catalyze the condensation reaction.

In some methods, processes and systems described herein, a condensation reaction may be performed in the presence of an enzyme. In some examples, the enzyme may be an acyltransferase. The acyltransferase may use a TA with an alcohol or

carboxylate functional group as a substrate. The acyltransferase may use a TA containing a carboxylate group activated via a 1-*O*-β glycosidic linkage to a sugar (hereafter referred to as a glycoside) as a substrate. The acyltransferase may convert the TA alcohol and carboxylate/glycoside functional groups to a corresponding ester derivative. Non-limiting examples of enzymes suitable for condensation of TA precursors in this disclosure include serine carboxypeptidase-like acyltransferases (SCPL-ATs). For example, littorine synthase (EC 2.3.1.-) may condense tropine and other TA precursors containing alcohol functional groups with 1-*O*-β-phenyllactoyl-glucose and other TA glycoside precursors to littorine and other corresponding ester products. In some examples, a protein that comprises an SCPL-AT domain of any one of the preceding examples may perform the condensation. In some examples, the SCPL-AT may catalyze the condensation reaction within a host cell, such as an engineered host cell, as described herein. In yet other examples, the SCPL-AT may catalyze the condensation reaction within a sub-cellular compartment inside a host cell, such as an engineered host cell, as described herein.

In some embodiments of the invention, the amino acid sequence of an acyltransferase enzyme which is used to perform a TA acyl transfer reaction, such as an SCPL-AT enzyme, is subject to one or more modifications which alters the post-translational processing, trafficking, folding, oligomerization, and/or sub-cellular localization of the enzyme. As some acyltransferase enzymes, including SCPL-AT enzymes, have never been demonstrated to exhibit catalytic activity in living, non-plant cells, such modifications may prove useful, or may be necessary, for activity in non-plant host cells. Examples of such modifications include, but are not limited to: addition, removal, or replacement of N-terminal signal peptide sequences; addition, removal, or replacement of internal propeptide sequences; addition or removal of asparagine-linked N-glycosylation sites; addition or removal of serine-linked O-glycosylation sites; and fusion of protein domains to the N- and/or C-terminus of the acyltransferase domain.

In one embodiment of the invention, an SCPL-AT enzyme domain is modified at its N-terminus by fusion of a soluble protein domain. This soluble domain masks any internal signal sequences in the acyltransferase domain, thereby modifying the trafficking and/or sub-cellular localization of the fused SCPL-AT domain. In some examples, the N-terminally fused domain induces trafficking of the SCPL-AT domain to sub-cellular compartments including, but not limited to, the ER membrane, ER lumen, cis-Golgi, trans-Golgi, lysosome, vacuole membrane, and vacuole lumen. The N-terminally fused soluble domain can also modify the oligomerization state of the SCPL-AT domain from its native state (monomer) to any state including, but not limited to, homodimer, heterodimer, homotrimer, heterotrimer, homotetramer, heterotetramer, homohexamer, heterohexamer, homooctamer, heterooctamer, or greater degrees of oligomerization.

In one example, the N-terminally fused soluble protein domain is a fluorescent protein selected from the group including, but not limited to, fluorescent proteins derived from *Aequoria* sp. and fluorescent proteins derived from *Discosoma* sp. In one example, the N-terminally fused soluble protein domain is red fluorescent protein from *Discosoma* sp. (DsRed). In other examples, the N-terminally fused soluble protein domain is another enzyme in the TA biosynthetic pathway, including but not limited to, ornithine decarboxylase, putrescine N-methyltransferase, pyrrolidine ketide synthase, tropinone reductase, phenylpyruvate reductase, phenyllactate UDP-glucosyltransferase 84A27, and hyoscyamine dehydrogenase.

Examples of amino acid sequences of soluble protein domains which can be fused to the N-terminus of a SCPL-AT domain that can then be used to perform a TA acyl transfer reaction within a non-plant cell are provided in Table 3. An amino acid sequence for a SCPL-AT enzyme comprising a fused N-terminal domain and that is utilized in TA acyl transfer reactions in non-plant cells may be 50% or more identical to a given amino acid sequence as listed in Table 3. For example, an amino acid sequence for such an acyltransferase may comprise an amino acid sequence that is at least 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

An engineered non-plant host cell may be provided that produces an acyltransferase that catalyzes a TA acyl transfer reaction, wherein the acyltransferase comprises an amino acid sequence whose N-terminus is fused to the amino acid sequence of a soluble protein domain selected from the group consisting of those sequences in Table 3. The acyltransferase that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. The one or more enzymes that are recovered from the engineered host cell that produces the acyltransferase may be used in a process for carrying out a TA acyl transfer reaction. The process may include contacting the TA precursors possessing an alcohol and/or a carboxylate/glycoside functional group with an acyltransferase in an amount sufficient to convert the alcohol and/or carboxylate/glycoside

group to a corresponding ester group. In examples, the TA precursors possessing an alcohol and/or a carboxylate/glycoside functional group may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said TA precursors are converted to the corresponding ester. In further examples, the TA possessing an alcohol and/or a carboxylate/glycoside functional group may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said TA precursors are converted to the corresponding ester.

The one or more enzymes that may be used to carry out a TA acyl transfer reaction may contact the TA precursors *in vitro*. Additionally, or alternatively, the one or more enzymes that may be used to carry out a TA acyl transfer reaction may contact the TA precursors *in vivo*. Additionally, the one or more enzymes that may be used to carry out a TA acyl transfer reaction may be provided to a cell having the TA precursors within, or may be produced within an engineered non-plant host cell.

In some examples, the methods provide for engineered non-plant host cells that produce an alkaloid product, wherein the TA acyl transfer reaction may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a medicinal TA. In still other embodiments, the alkaloid produced is derived from a medicinal TA, including, for example, non-natural TAs. In still other embodiments, the alkaloid product is selected from the group consisting of medicinal TA, non-medicinal TA, and non-natural TA.

In some examples, the substrates are TA precursors selected from the group consisting of tropine, pseudotropine, ecgonine, methylecgonine, phenyllactic acid, cinnamic acid, ferulic acid, coumaric acid, and glycosides of the listed compounds.

In some examples, the methods provide for engineered non-plant host cells that produce alkaloid products from tropine and 1-*O*-β-phenyllactoylglucose. The condensation of tropine and 1-*O*-β-phenyllactoylglucose to littorine may comprise a key step in the production of diverse alkaloid products from a precursor. In some examples, the precursor is an L-amino acid or a sugar (e.g., glucose). The diverse alkaloid products can include, without limitation, medicinal TAs, non-medicinal TAs, and non-natural TAs.

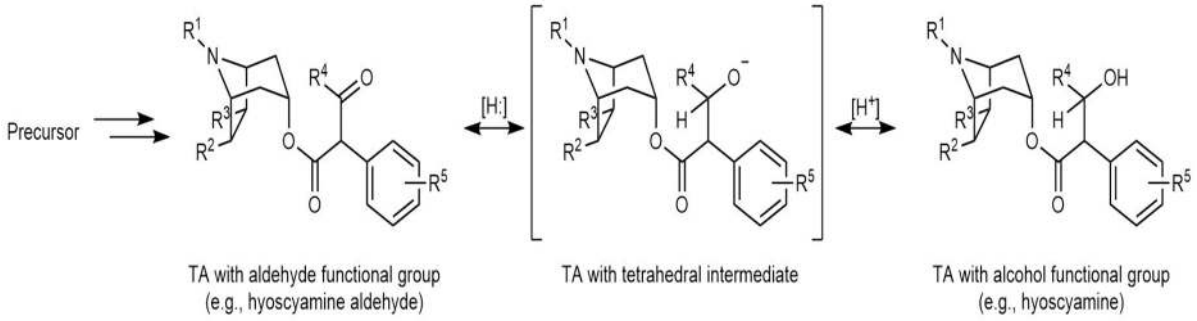
Any suitable carbon source may be used as a precursor toward a TA acyl transfer reaction. Suitable precursors can include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose),

polysaccharides (e.g., starch, cellulose), or a combination thereof. In some examples, unpurified mixtures from renewable feedstocks can be used (e.g., cornsteep liquor, sugar beet molasses, barley malt, biomass hydrolysate). In still other embodiments, the carbon precursor can be a one-carbon compound (e.g., methanol, carbon dioxide) or a two-carbon compound (e.g., ethanol). In yet other embodiments, other carbon-containing compounds can be utilized, for example, methylamine, glucosamine, and amino acids (e.g., L-arginine and L-phenylalanine). In some examples, a TA or a precursor of a TA possessing an alcohol and/or a carboxylate/glycoside functional group may be added directly to an engineered host cell of the invention, including, for example, tropine, pseudotropine, ecgonine, methylecgonine, phenyllactic acid, cinnamic acid, ferulic acid, coumaric acid, and glycosides of the listed compounds.

In some embodiments, the substrate used to carry out the vacuolar TA acyl transfer reaction may comprise one or more alcohol and/or carboxylate/glycoside functional groups, wherein only one of said functional groups is condensed to the corresponding ester.

TA Alcohol-Aldehyde Interconversions

Some methods, processes, and systems provided herein describe the conversion of TAs with aldehyde functional groups to TAs with alcohol (hydroxyl) functional groups, and the conversion of TAs with alcohol functional groups to TAs with aldehyde functional groups (hereafter referred to as TA alcohol-aldehyde interconversions). Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the TA alcohol-aldehyde interconversion is a key step in the conversion of a substrate to a diverse range of alkaloids. In some examples, the conversion of a TA aldehyde group to a TA alcohol group comprises a reduction reaction. In some cases, reduction of a substrate TA aldehyde to an alcohol may be performed by reducing an aldehyde substrate to the corresponding tetrahedral oxyanion intermediate, then protonating this intermediate to a hydroxyl as provided in Figure 2 and as represented generally in Scheme 1. As provided in Scheme 1, R^1 may be H, CH_3, or a higher order alkyl group; R^2 and R^3 may be H, OH, or OCH_3; R^4 may be H; and R^5 may be H, OH, C_1-C_4 alkyl, C_1-C_4 alkoxy, C_1-C_4 acyl, F, Cl, or Br.



Scheme 1

In some examples, the TA alcohol-aldehyde interconversion may involve at least one oxidation reaction or at least one reduction reaction. In some cases, at least one of the oxidation or reduction reactions is carried out in the presence of an enzyme. In some cases, at least one of the oxidation or reduction reactions is catalyzed by an enzyme. In some cases, the oxidation and reduction reactions are both carried out in the presence of at least one enzyme. In some cases, at least one enzyme is useful to catalyze the oxidation and reduction reactions. The oxidation and reduction reactions may be catalyzed by the same enzyme.

In some methods, processes and systems described herein, an oxidation or reduction reaction may be performed in the presence of an enzyme. In some examples, the enzyme may be a dehydrogenase. The dehydrogenase may use a TA with an alcohol or aldehyde functional group as a substrate. The dehydrogenase may convert the TA alcohol or aldehyde functional group to a corresponding aldehyde or alcohol derivative. The dehydrogenase may be referred to as hyoscyamine dehydrogenase (HDH). Non-limiting examples of enzymes suitable for oxidation and/or reduction of TAs in this disclosure include a cytochrome P450 oxidase, a 2-oxoglutarate-dependent oxidase, a flavoprotein oxidase, a short-chain dehydrogenase-reductase (SDR), a medium-chain dehydrogenase-reductase (MDR), a cinnamyl alcohol dehydrogenase (CAD), and an aldo-keto reductase (AKR). For example, tropinone reductase 1 (EC 1.1.1.206) may oxidize tropinone and other TA precursors with ketone functional groups to tropine (3α-tropanol) and other corresponding alcohol products. In some examples, a protein that comprises a dehydrogenase domain of any one of the preceding examples may perform the oxidation or reduction. In some examples, the dehydrogenase may catalyze the oxidation and/or reduction reactions within a host cell, such as an engineered host cell, as described herein.

Examples of amino acid sequences of a dehydrogenase enzyme that may be used to perform a TA alcohol-aldehyde interconversion are provided in Table 2. An amino acid sequence for a dehydrogenase that is utilized in TA alcohol-aldehyde interconversions may be 50% or more identical to a given amino acid sequence as listed in Table 2. For example, an amino acid sequence for such a dehydrogenase may comprise an amino acid sequence that is at least 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

An engineered host cell may be provided that produces a dehydrogenase that catalyzes a TA alcohol-aldehyde interconversion, wherein the dehydrogenase comprises an amino acid sequence selected from the group consisting of those sequences in Table 2. The dehydrogenase that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. The one or more enzymes that are recovered from the engineered host cell that produces the dehydrogenase may be used in a process for carrying out a TA alcohol-aldehyde interconversion. The process may include contacting the TA possessing an alcohol and/or an aldehyde functional group with a dehydrogenase in an amount sufficient to convert the alcohol and/or aldehyde group of the TA to a corresponding aldehyde and/or alcohol group. In examples, the TA possessing an alcohol and/or an aldehyde functional group may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said TA is converted to its corresponding aldehyde and/or alcohol group. In further examples, the TA possessing an alcohol and/or an aldehyde functional group may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said TA is converted to its corresponding aldehyde and/or alcohol group.

The one or more enzymes that may be used to carry out a TA alcohol-aldehyde interconversion may contact the TA *in vitro*. Additionally, or alternatively, the one or more enzymes that may be used to carry out a TA alcohol-aldehyde interconversion may contact the TA *in vivo*. Additionally, the one or more enzymes that may be used to carry out a TA alcohol-aldehyde interconversion may be provided to a cell having the TA within, or may be produced within an engineered host cell.

In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the TA alcohol-aldehyde interconversion may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a medicinal TA. In still other embodiments, the alkaloid produced is derived from a medicinal TA, including, for example, non-natural TAs. In another embodiment, a TA possessing an alcohol and/or an aldehyde functional group is an intermediate toward the product of the engineered host cell. In still other embodiments, the alkaloid product is selected from the

group consisting of medicinal TA, non-medicinal TA, and non-natural TA.

In some examples, the substrate is a TA or a precursor of a TA selected from the group consisting of littorine, hyoscyamine aldehyde, hyoscyamine, anisodamine, and scopolamine.

In some examples, the methods provide for engineered host cells that produce alkaloid products from hyoscyamine aldehyde. The reduction of hyoscyamine aldehyde to hyoscyamine may comprise a key step in the production of diverse alkaloid products from a precursor. In some examples, the precursor is an L-amino acid or a sugar (e.g., glucose). The diverse alkaloid products can include, without limitation, medicinal TAs, non-medicinal TAs, and non-natural TAs.

Any suitable carbon source may be used as a precursor toward a TA alcohol-aldehyde interconversion. Suitable precursors can include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose), polysaccharides (e.g., starch, cellulose), or a combination thereof. In some examples, unpurified mixtures from renewable feedstocks can be used (e.g., cornsteep liquor, sugar beet molasses, barley malt, biomass hydrolysate). In still other embodiments, the carbon precursor can be a one-carbon compound (e.g., methanol, carbon dioxide) or a two-carbon compound (e.g., ethanol). In yet other embodiments, other carbon-containing compounds can be utilized, for example, methylamine, glucosamine, and amino acids (e.g., L-arginine and L-phenylalanine). In some examples, a TA or a precursor of a TA possessing an alcohol and/or an aldehyde functional group may be added directly to an engineered host cell of the invention, including, for example, tropine, pseudotropine, ecgonine, methylecgonine, littorine, hyoscyamine aldehyde, hyoscyamine, anisodamine, and scopolamine.

In some embodiments, the substrate used to carry out the TA alcohol-aldehyde interconversion may comprise one or more alcohol and/or aldehyde functional groups, wherein only one of said functional groups is oxidized or reduced to the corresponding aldehyde or alcohol group.

Methods for modulating cellular metabolite transport

Some methods, processes, and systems provided herein describe the use of proteins (hereafter referred to as ‘transporters’, as defined previously) to translocate metabolites across lipid or lipid-protein membranes (hereafter referred to as ‘transmembrane transport’). Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, transmembrane transport is a key step in the conversion of a substrate to a diverse range of alkaloids.

In certain embodiments, the host cell includes one or more heterologous coding

sequences for one or more transporters or active fragments thereof that localize to a lipid or lipid-protein membrane and translocate a TA or a TA precursor across the same membrane. In some examples, the membrane is the vacuole membrane. In other examples, the membrane is the ER membrane. In some examples, the membrane is the peroxisome membrane. In some examples, the membrane is the mitochondrion membrane. In other examples, the membrane is the cellular plasma membrane.

In some examples, TAs and TA precursors transported in this manner include, but are not limited to, putrescine, *N*-methylputrescine, 4-methylaminobutanal, *N*-methylpyrrolinium, tropinone, tropine, nortropine, tropine *N*-oxide, phenyllactic acid, 1-*O*-β-phenyllactoylglucose, littorine, norlittorine, littorine *N*-oxide, hyoscyamine, norhyoscyamine, hyoscyamine *N*-oxide, anisodamine, noranisodamine, anisodamine *N*-oxide, scopolamine, norscopolamine, and scopolamine *N*-oxide. The accumulation of such TAs or TA precursors in specific sub-cellular compartments can preclude access by operably linked biosynthetic enzymes in different compartments; therefore, the use of transporters which translocate TAs or TA precursors from one compartment to another can mitigate such transport limitations. In certain cases, the expression of heterologous coding sequences for one or more transporters within a host cell can increase production of a TA or a TA precursor.

In some embodiments, the transporter or active fragment thereof is a multidrug and toxin extrusion (MATE) transporter. Any convenient MATE transporters which transport one or more of the aforementioned TAs or TA precursors find use in the subject host cells. Transporter proteins of interest include, but are not limited to, enzymes such as *Nicotiana tabacum* jasmonate-inducible alkaloid transporter 1 (NtJAT1), *N. tabacum* jasmonate-inducible alkaloid transporter 2 (NtJAT2), *N. tabacum* MATE1, *N. tabacum* MATE2, or any others as described in Table 1 and Table 4.

In certain embodiments, the transporter or active fragment thereof is a nitrate transporter 1/peptide transporter family (NPF/NRT) transporter. Any convenient NPF/NRT transporters which transport one or more of the aforementioned TAs or TA precursors find use in the subject host cells. Transporter proteins of interest include, but are not limited to, enzymes such as *Atropa belladonna* lactose permease-like transporter 1 (AbLP1), *Catharanthus roseus* NPF transporter 2.9 (CrNPF2.9), or any others described in Tables 1, 4, and 6.

In some embodiments, the transporter or active fragment thereof is a purine uptake permease-like (PUP) family transporter. Any convenient PUP transporters which transport one or more of the aforementioned TAs or TA precursors find use in the subject host cells. Transporter proteins of interest include, but are not limited to, enzymes such as *N. tabacum* nicotine uptake permease 1 (NtNUP1), *A. belladonna* purine uptake permease-like transporter 1 (AbPUP1), or any others described in Tables 1, 4, and 6.

In other embodiments, the transporter or active fragment thereof is an ATP-binding cassette (ABC) transporter. Any convenient ABC transporters which transport one or more of the aforementioned TAs or TA precursors find use in the subject host cells. In some embodiments, the transporter or active fragment thereof is a pleiotropic drug resistance (PDR) transporter. Any convenient PDR transporters which transport one or more of the aforementioned TAs or TA precursors find use in the subject host cells.

In certain embodiments, the transporter or active fragment thereof is a transporter which translocates tropine or a tropine derivative from the cytosol to the vacuole lumen of a subject host cell, such as NtJAT1, NtJAT2, NtMATE1, or NtMATE2. In some embodiments, the transporter or active fragment thereof is a transporter which translocates an acyl-O-β-glucoside, such as phenyllactyl-O-β-glucoside, from the cytosol to the vacuole lumen of a subject host cell. In certain other embodiments, the transporter or active fragment thereof is a transporter with translocates littorine or a littorine derivative from the vacuole lumen to the cytosol of a subject host cell, such as AbPUP1 or AbLP1. In other embodiments, the transporter or active fragment thereof is a transporter which translocates a TA, a precursor of a TA, or a derivative of a TA between the cytosol and the extracellular space surrounding a subject host cell, such as AbPUP1 or AbLP1.

In certain embodiments, the host cell includes a heterologous coding sequence for a transporter or an active fragment thereof. In some embodiments of the invention, the amino acid sequence of a transporter is subject to one or more modifications which alters the sub-cellular localization, the direction of substrate translocation, and/or the topological orientation of the enzyme. Examples of such modifications include, but are not limited to: addition, removal, or replacement of N-terminal, C-terminal, or internal signal sequences; addition, removal, replacement, or rearrangement of transmembrane helices; and fusion of protein domains to the N- and/or C-terminus of the transporter.

Examples of amino acid sequences of transporters which can be used to mitigate substrate transport limitations and/or to increase accumulation of TAs or TA precursors in specific cellular compartments are provided in Tables 1, 4 and 6. An amino acid sequence for a transporter that is utilized in this manner in non-plant cells may be 50% or more identical to a given amino acid sequence as listed in Tables 1, 4 and 6. For example, an amino acid sequence for such a transporter may comprise an amino acid sequence that is at least 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99%

identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

An engineered non-plant host cell may be provided that produces a transporter which translocates one or more TAs or TA precursors or TA derivatives from one cellular compartment to another, wherein the transporter comprises an amino acid sequence selected from the group consisting of those sequences in Tables 1, 4 and 6. In some examples, the methods provide for engineered non-plant host cells that produce an alkaloid product, wherein TA transmembrane transport may comprise a critical or essential step in the production of an alkaloid product. In some examples, the alkaloid produced is a medicinal TA. In still other embodiments, the alkaloid produced is derived from a medicinal TA, including, for example, non-natural TAs. In still other embodiments, the alkaloid product is selected from the group consisting of medicinal TA, non-medicinal TA, and non-natural TA.

Heterologous Coding Sequences

In some instances, the host cells are cells that harbor one or more heterologous coding sequences (such as two or more, three or more, four or more, five or more, or even more) which encode activity(ies) that enable the host cells to produce desired TAs of interest, e.g., as described herein. As used herein, the term "heterologous coding sequence" is used to indicate any polynucleotide that codes for, or ultimately codes for, a peptide or protein or its equivalent amino acid sequence, e.g., an enzyme or a transporter, that is not normally present in the host organism and may be expressed in the host cell under proper conditions. As such, "heterologous coding sequences" includes multiple copies of coding sequences that are normally present in the host cell, such that the cell is expressing additional copies of a coding sequence that are not normally present in the cells. The heterologous coding sequences may be RNA or any type thereof, e.g., mRNA, DNA or any type thereof, e.g., cDNA, or a hybrid of RNA/DNA. Coding sequences of interest include, but are not limited to, full-length transcription units that include such features as the coding sequence, introns, promoter regions, 3'-UTRs, and enhancer regions.

In examples, the engineered host cell comprises a plurality of heterologous coding sequences each encoding an enzyme or a transporter. In some examples, the plurality of enzymes and transporters encoded by the plurality of heterologous coding sequences may be distinct from each other. In some examples, some of the plurality of enzymes and

transporters encoded by the plurality of heterologous coding sequences may be distinct from each other and some of the plurality of enzymes and transporters encoded by the plurality of heterologous coding sequences may be duplicate copies.

In some examples, the heterologous coding sequences may be operably connected. Heterologous coding sequences that are operably connected may be within the same pathway of producing a particular TA product. In some examples, the operably connected heterologous coding sequences may be directly sequential along the pathway of producing a particular TA product. In some examples, the operably connected heterologous coding sequences may have one or more native enzymes or transport steps between one or more of the enzymes or transporters encoded by the plurality of heterologous coding sequences. In some examples, the heterologous coding sequences may have one or more heterologous enzymes or transporters between one or more of the enzymes or transporters encoded by the plurality of heterologous coding sequences. In some examples, the heterologous coding sequences may have one or more non-native enzymes or transporters between one or more of the enzymes or transporters encoded by the plurality of heterologous coding sequences.

In some embodiments, the host cell includes putrescine N-methyltransferase (PMT) activity. Any convenient PMT enzymes find use in the subject host cells. PMT enzymes of interest include, but are not limited to, enzymes such as EC 2.1.1.53, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for a PMT or an active fragment thereof.

In some instances, the host cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert NMP to 4MAB. In certain cases, the one or more enzymes is selected from plant methylputrescine oxidases (MPOs) eukaryotic MPOs, and eukaryotic polyamine oxidases (e.g., EC 1.4.3.22).

In certain embodiments, the cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert NMPy to MPOB. In certain cases, the one or more enzymes is a type III polyketide synthase (e.g., EC 2.3.1.-). The one or more heterologous coding sequences may be derived from any convenient species (e.g., as described herein). In some cases, the one or more heterologous coding sequences may be derived from a species described in Table 1. In some cases, the one or more heterologous coding sequences are present in a gene or enzyme selected from those described in Table 1.

In certain embodiments, the host cell includes tropinone synthase activity. Any convenient tropinone synthase enzymes (e.g., CYP82M3) find use in the subject host cells. Tropinone synthase enzymes of interest include, but are not limited to, enzymes such as EC 1.14.14.-, as described in Table 1. In certain embodiments, the host cell includes a

heterologous coding sequence for a tropinone synthase or an active fragment thereof.

In certain embodiments, the host cell includes tropinone reductase activity. Any convenient tropinone reductase enzymes find use in the subject host cells. Tropinone reductase enzymes of interest include, but are not limited to, enzymes such as EC 1.1.1.206, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for a tropinone reductase or an active fragment thereof.

In some instances, the host cell includes phenylpyruvate reductase (PPR) activity. Any convenient PPR enzymes find use in the subject host cells. Some PPR enzymes of interest include, but are not limited to, enzymes such as EC 1.1.1.237, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for a PPR or an active fragment thereof.

In certain embodiments, the host cell includes phenyllactate glycosyltransferase activity. Any convenient phenyllactate glycosyltransferase enzymes find use in the subject host cells. Glycosyltransferase enzymes include, but are not limited to, enzymes such as 2.4.1.-, which transfer a glucose moiety from UDP-glucose to phenyllactate by means of a glycosidic ester linkage, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for a phenyllactate glycosyltransferase or an active fragment thereof.

In certain embodiments, the cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert tropine and 1-*O*-β-phenyllactoylglucose to littorine. In some embodiments, the host cell includes littorine synthase activity. Any convenient littorine synthase enzymes or enzymes comprising littorine synthase active fragments find use in the subject host cells. Littorine synthase enzymes of interest include, but are not limited to, enzymes such as EC 2.3.1.-, as described in Table 1, and enzymes comprising littorine synthase enzymes whose N-termini are fused to soluble protein domains described in Table 3. In certain embodiments, the host cell includes a heterologous coding sequence for a littorine synthase or an active fragment thereof.

In certain instances, the host cell includes littorine mutase activity. Any convenient littorine mutase enzymes find use in the subject host cells. Littorine mutase enzymes of interest include, but are not limited to, enzymes such as EC 1.14.19.-, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for a littorine mutase or an active fragment thereof.

In some embodiments, the host cell includes hyoscyamine dehydrogenase (HDH) activity. Any convenient HDH enzymes find use in the subject host cells. Some HDH enzymes of interest include, but are not limited to, those sequences described in Table 2. In certain embodiments, the host cell includes a heterologous coding sequence for an HDH or

an active fragment thereof.

In certain embodiments, the host cell includes hyoscyamine 6β-hydroxylase/dioxygenase (H6H) activity. Any convenient H6H enzymes find use in the subject host cells. Some H6H enzymes of interest include, but are not limited to, enzymes such as EC 1.14.11.11, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for an H6H or an active fragment thereof.

In certain examples, the engineered host cell comprises a plurality of heterologous coding sequences each encoding a transmembrane metabolite transporter. In some examples, the plurality of transporters encoded by the plurality of heterologous coding sequences may be distinct from each other. In some examples, some of the plurality of transporters encoded by the plurality of heterologous coding sequences may be distinct from each other and some of the plurality of transporters encoded by the plurality of heterologous coding sequences may be duplicate copies. In some embodiments, the host cell includes vacuolar tropine import activity. Any convenient transporters capable of transporting tropine from the cytosol to the vacuole lumen find use in the subject host cell. Examples of such transporters include NtJAT1, NtJAT2, NtMATE1, and NtMATE2. In some embodiments, the host cell includes vacuolar 1-*O*-β-phenyllactoylglucose import activity. Any convenient transporters capable of transporting 1-*O*-β-phenyllactoylglucose from the cytosol to the vacuole lumen find use in the subject host cell. In some embodiments, the host cell includes vacuolar littorine export activity. Any convenient transporters capable of transporting littorine from the vacuole lumen to the cytosol find use in the subject host cell. Examples of such transporters include AbPUP1 and AbLP1. In other embodiments, the host cell includes cellular TA import activity. Any convenient transporters capable of transporting littorine, hyoscyamine aldehyde, hyoscyamine, anisodamine, scopolamine or derivatives thereof from the extracellular space into the cytosol find use in the subject host cell. Examples of such transporters include AbPUP1 and AbLP1.

As used herein, the term "heterologous coding sequences" also includes the coding portion of the peptide or enzyme, i.e., the cDNA or mRNA sequence, of the peptide or enzyme, as well as the coding portion of the full-length transcriptional unit, i.e., the gene including introns and exons, as well as "codon optimized" sequences, truncated sequences or other forms of altered sequences that code for the enzyme or transporter or code for its equivalent amino acid sequence, provided that the equivalent amino acid sequence produces a functional protein. Such equivalent amino acid sequences may have a deletion of one or more amino acids, with the deletion being N-terminal, C-terminal, or internal. Truncated forms are envisioned as long as they have the catalytic or transport capability indicated herein. Fusions of two or more enzymes, or one or more enzymes to a transporter, are also envisioned to facilitate the transfer of metabolites in the pathway,

provided that catalytic and transport activities are maintained. Also included are fusions of one or more enzymes or catalytic protein domains with one or more non-catalytic protein domains in a manner by which the non-catalytic protein domain facilitates the solubilization, folding, maturation, and/or activity of the fused catalytic domain.

Operable fragments, mutants or truncated forms may be identified by modeling and/or screening. This is made possible by addition or deletion of, for example, N-terminal, C-terminal, or internal regions of the protein in a step-wise fashion, followed by analysis of the resulting derivative with regard to its activity for the desired reaction compared to the original sequence. If the derivative in question operates in this capacity, it is considered to constitute an equivalent derivative of the enzyme proper.

Aspects of the present invention also relate to heterologous coding sequences that code for amino acid sequences that are equivalent to the native amino acid sequences for the various enzymes. An amino acid sequence that is "equivalent" is defined as an amino acid sequence that is not identical to the specific amino acid sequence, but rather contains at least some amino acid changes (deletions, substitutions, inversions, insertions, etc.) that do not essentially affect the biological activity of the protein as compared to a similar activity of the specific amino acid sequence, when used for a desired purpose. The biological activity refers to, in the example of a decarboxylase, its catalytic activity. Equivalent sequences are also meant to include those which have been engineered and/or evolved to have properties different from the original amino acid sequence. Mutable properties of interest include catalytic activity, substrate specificity, selectivity, stability, solubility, localization, etc. In certain embodiments, an "equivalent" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence, in some cases at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

The host cells may also be modified to possess one or more genetic alterations to accommodate the heterologous coding sequences. Alterations of the native host genome include, but are not limited to, modifying the genome to reduce or ablate expression of a specific protein that may interfere with the desired pathway. The presence of such native proteins may rapidly convert one of the intermediates or final products of the pathway into a metabolite or other compound that is not usable in the desired pathway. Alternately the presence of such native proteins may rapidly transport one of the intermediates or final products of the pathway to a different cellular compartment or outside the cell, such that it is not usable in the desired pathway. Thus, if the activity of the native enzyme or transporter were reduced or altogether absent, the produced intermediates would be more readily

available for incorporation into the desired product and/or in the desired cellular compartment.

In some instances, where ablation of expression of a protein may be of interest, the alteration is in proteins involved in the pleiotropic drug response, including, but not limited to, ATP-binding cassette (ABC) transporters, multidrug resistance (MDR) pumps, and associated transcription factors. These proteins are involved in the export of TA molecules and TA precursors into the culture medium, thus deletion controls the export of the compounds into the media, making them more available for incorporation into the desired product. In some embodiments, host cell gene deletions of interest include genes associated with the unfolded protein response and endoplasmic reticulum (ER) proliferation. Such gene deletions may lead to improved TA production. The expression of cytochrome P450s may induce the unfolded protein response and may cause the ER to proliferate. Deletion of genes associated with these stress responses may control or reduce overall burden on the host cell and improve pathway performance. In other embodiments, host cell gene deletions of interest include genes associated with the lysosomal maturation and vacuolar maturation and proliferation pathways. Such gene deletions may lead to improved TA production. The high-level expression of one or more native or heterologous proteins in the membrane of secretory compartments (such as lysosomes) or of the vacuole may disrupt the normal maturation (proliferation, growth and development, fusion/fission, etc.) of these compartments. Deletion of genes associated with these mechanisms may control the orphology of the compartment(s) in a way that promoters the expression, stability, or activity of enzymes or transporters in the membrane(s) of the compartment(s). Genetic alterations may also include modifying the promoters of endogenous genes to increase expression and/or introducing additional copies of endogenous genes. Examples of this include the construction/use of strains which overexpress the endogenous yeast NADPH-P450 reductase Ncp1p to increase activity of heterologous P450 enzymes. In addition, endogenous enzymes such as Spe1p, Fms1p, Car1p, Arg2p, Aro8p, Aro9p, Pha2p, Ugp1p, Pdc6p, Idp3p, Leu2p, Ura3p, and Trp1p which are directly involved in the synthesis of intermediate metabolites, may also be overexpressed.

Heterologous coding sequences of interest include but are not limited to sequences that encode enzymes, either wild-type or equivalent sequences, that are normally responsible for the production of TAs and precursors in plants. In some cases, the enzymes for which the heterologous sequences code may be any of the enzymes in the TA pathway, and may be from any convenient source. The choice and number of enzymes encoded by the heterologous coding sequences for the particular synthetic pathway may be selected based upon the desired product. In certain embodiments, the host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or

more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more heterologous coding sequences, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 heterologous coding sequences.

In some cases, polypeptide sequences encoded by the heterologous coding sequences are as reported in GENBANK. Enzymes of interest include, but are not limited to, those enzymes described herein and those shown in Table 1. The host cells may include any combination of the listed enzymes or transporters, from any source. Unless otherwise indicated, accession numbers in Table 1 refer to GenBank. Some accession numbers refer to the Saccharomyces genome database (SGD), which is available on the world-wide web at www.yeastgenome.org.

In some embodiments, the host cell (e.g., a yeast strain) is engineered for selective production of a TA of interest by localizing one or more enzymes to a compartment in the cell. In some cases, an enzyme may be located in the host cell such that the compound produced by this enzyme spontaneously rearranges, or is converted by another enzyme to a desirable metabolite before reaching a localized enzyme that may convert the compound into an undesirable metabolite. The spatial distance between two enzymes may be selected to prevent one of the enzymes from acting directly on a compound to make an undesirable metabolite, and restrict production of undesirable end products (e.g., an undesirable opioid by-product). In some other cases, an enzyme may be localized in the host cell such that the sub-cellular compartment in which it is located provides a more optimum pH, cofactor concentration, redox potential, substrate concentration, and/or other biochemical parameter for its activity than the compartment in which the enzyme is naturally found. In certain cases, an enzyme may be localized to a specific compartment within the host cell such that the intracellular trafficking pathway by which the enzyme is transported to said compartment provides the necessary post-translational modifications for the enzyme to exhibit activity. Such post-translational modifications include, but are not limited to, acetylation, acetylglycosylation, amidation, carboxylation, methylation, glutathionylation, hydroxylation, glycosylation, phosphorylation, sulfonation, disulfide bond formation, cleavage of signal sequences, and multi-enzyme complex formation. In certain embodiments, any of the enzymes described herein, either singularly or together with a second enzyme, may be localized to any convenient compartment in the host cell, including but not limited to, an organelle, endoplasmic reticulum, Golgi, vacuole, nucleus, plasma membrane, mitochondrion, peroxisome, periplasm, the lumen of any of the aforementioned organelles, or the membrane enclosing or associated with any of the aforementioned organelles. In cases where one or more enzymes are localized to a membrane associated with any of the aforementioned organelles, the enzyme may be oriented such that the catalytic domain of the enzyme faces the cytosol, the lumen of the organelle, and/or any other intracellular

space. In some embodiments, the host cell includes one or more of the enzymes that include a localization tag. Any convenient tags may be utilized. In some cases, the localization tag is a peptidic sequence that is attached at the N-terminus and/or C-terminus of the enzyme.

Any convenient methods may be utilized for attaching a tag to the enzyme. In some cases, the localization tag is derived from an endogenous yeast protein. Such tags may provide a route to a variety of yeast organelles including, but not limited to, the endoplasmic reticulum (ER), Golgi apparatus (GA), mitochondria (MT), plasma membrane (PM), peroxisome (POX), and vacuole (V). In certain embodiments, the tag is an ER routing tag (e.g., ER1). In certain embodiments, the tag is a vacuole tag (e.g., V1). In certain embodiments, the tag is a plasma membrane tag (e.g., P1). In certain embodiments, the tag is a peroxisome-targeting sequence (e.g., PTS1). In certain instances, the tag includes or is derived from, a transmembrane domain from within the tail-anchored class of proteins. In some embodiments, the localization tag locates the enzyme on the outside of an organelle. In certain embodiments, the localization tag locates the enzyme on the inside of an organelle. In some embodiments, the localization tag locates the enzyme such that one or more portions of the enzyme are found both inside and outside of an organelle.

In some embodiments of the invention, the host cell is modified by expression of one or more coding sequences encoding one or more enzymes comprising a localization tag described above. In certain embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes is expressed in the cytosol. Examples of such enzymes include, but are not limited to, arginine decarboxylases, putrescine N-methyltransferases, pyrrolidine ketide synthases, tropinone reductases, phenylpyruvate reductases, UDP-glucosyltransferases, 2-oxoglutarate-dependent dioxygenases such as hyoscyamine 6β-hydroxylase/dioxygenase, and pyrrolizidine *N*-oxygenases. In certain embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes is expressed in the ER membrane. Examples of such enzymes include, but are not limited to, cytochromes P450 such as tropinone synthase (CYP82M3) littorine mutase (CYP80F1), human hepatic CYP450s including (but not limited to) CYP2D6, CYP2C19, and CYP3A4, and $NADP^+$-cytochrome P450 reductases. In certain embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes is expressed in the mitochondria. Examples of such enzymes include, but are not limited to, N-acetylglutamate synthases. In other embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes is expressed in the peroxisome. Examples of such enzymes include, but are not limited to, amine oxidases such as N-methylputrescine oxidase. In other embodiments, the host cell is

modified by expression of one or more heterologous coding sequences such that one or more enzymes is expressed in the vacuole lumen. Examples of such enzymes include, but are not limited to, serine carboxypeptidase-like acyltransferases such as littorine synthase, and engineered variants thereof. In other embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes or proteins is expressed in the vacuole membrane. Examples of such proteins include, but are not limited to, MATE transporters like NtJAT1 and NtMATE2, NPF transporters like AbLP1, PUP transporters like AbPUP1, and ABC transporters. In other embodiments, the host cell is modified by expression of one or more heterologous coding sequences such that one or more enzymes or proteins is expressed in the plasma membrane. Examples of such proteins include, but are not limited to, ABC transporters, pleiotropic drug resistance transporters, and multidrug resistance transporters.

In some instances, the expression of each type of enzyme is increased through additional gene copies (i.e., multiple copies), which increases intermediate accumulation and/or TA of interest production. Embodiments of the present invention include increased TA of interest production in a host cell through simultaneous expression of multiple species variants of a single or multiple enzymes or transporters. In some cases, additional gene copies of a single or multiple enzymes or transporters are included in the host cell. Any convenient methods may be utilized including multiple copies of a heterologous coding sequence for an enzyme or transporter in the host cell.

In some embodiments, the host cell includes multiple copies of a heterologous coding sequence for an enzyme or transporter, such as 2 or more, 3 or more, 4 or more, 5 or more, or even 10 or more copies. In certain embodiments, the host cell includes multiple copies of heterologous coding sequences for one or more enzymes or transporters, such as multiple copies of two or more, three or more, four or more, etc. In some cases, the multiple copies of the heterologous coding sequence for an enzyme are derived from two or more different source organisms as compared to the host cell. For example, the host cell may include multiple copies of one heterologous coding sequence, where each of the copies is derived from a different source organism. As such, each copy may include some variations in explicit sequences based on inter-species differences of the enzyme of interest that is encoded by the heterologous coding sequence.

In some embodiments of the host cell, the heterologous coding sequence is from a source organism selected from the group consisting of *Escherichia coli, Bacillus coagulans, Lactobacillus casei, Lactobacillus plantarum, Lactobacillus spp, Wickerhamia fluorescens, Aequoria spp, Discosoma spp, Homo sapiens*, *Zonocerus variegatus*, *Arabidopsis thaliana, Avena sativa, Solanum lycopersicum, Solanum tuberosum, Nicotiana tabacum, Nicotiana benthamiana, Atropa belladonna, Hyoscyamus niger, Hyoscyamus muticus, Datura*

stramonium, Datura metel, Datura innoxia, Duboisia myoporoides, Anisodus luridus, Anisodus tanguticus, Anisodus acutangulus, Brugmansia arborea, Brugmansia × candida, Brugmansia sanguinea, Erythroxylum coca, Cochlearia officinalis, Solanum spp, Nicotiana spp, Atropa spp, Hyoscyamus spp, Datura spp, Duboisia spp, Anisodus spp, Brugmansia spp, Erythroxylum spp, or Cochlearia spp. In certain instances, the heterologous coding sequence is from a source organism selected from *A. belladonna, H. niger, and D. stramonium*. In some embodiments, the host cell includes a heterologous coding sequence from one or more of the source organisms described in Table 1.

The engineered host cell medium may be sampled and monitored for the production of TAs of interest. The TAs of interest may be observed and measured using any convenient methods. Methods of interest include, but are not limited to, LC-MS methods (e.g., as described herein) where a sample of interest is analyzed by comparison with a known amount of a standard compound. Identity may be confirmed, e.g., by m/z and MS/MS fragmentation patterns, and quantitation or measurement of the compound may be achieved via LC trace peaks of know retention time and/or EIC MS peak analysis by reference to corresponding LC-MS analysis of a known amount of a standard of the compound.

METHODS

Process Steps

As summarized above, aspects of the invention include methods of preparing a tropane alkaloid (TA) of interest. As such, aspects of the invention include culturing a host cell under conditions in which the one or more host cell modifications (e.g., as described herein) are functionally expressed such that the cell converts starting compounds of interest into product TAs of interest or precursors thereof (e.g., pre-esterification TAs). Also provided are methods that include culturing a host cell under conditions suitable for protein production such that one or more heterologous coding sequences are functionally expressed and convert starting compounds of interest into product TAs of interest. In some instances, the method is a method of preparing a tropane alkaloid (TA), include culturing a host cell (e.g., as described herein); adding a starting compound to the cell culture; and recovering the TA from the cell culture. In some embodiments of the method, the starting compound, TA product and host cell are described by one of the entries of Table 1.

Fermentation media may contain suitable carbon substrates. The source of carbon suitable to perform the methods of this disclosure may encompass a wide variety of carbon containing substrates. Suitable substrates may include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose), polysaccharides (e.g., starch, cellulose), or a combination

thereof. In some cases, unpurified mixtures from renewable feedstocks may be used (e.g., cornsteep liquor, sugar beet molasses, barley malt). In some cases, the carbon substrate may be a one-carbon substrate (e.g., methanol, carbon dioxide) or a two-carbon substrate (e.g., ethanol). In other cases, other carbon containing compounds may be utilized, for example, methylamine, glucosamine, and amino acids.

Any convenient methods of culturing host cells may be employed for producing the TA precursors and downstream TAs of interest. The particular protocol that is employed may vary, e.g., depending on host cell, the heterologous coding sequences, the desired TA precursors and downstream TAs of interest, etc. The cells may be present in any convenient environment, such as an environment in which the cells are capable of expressing one or more functional heterologous enzymes. *In vitro*, as used herein, simply means outside of a living cell, regardless of the location of the cell. As used herein, the term *in vivo* indicates inside a living cell, regardless of the location of the cell. In some embodiments, the cells are cultured under conditions that are conducive to enzyme expression and with appropriate substrates available to allow production of TA precursors and downstream TAs of interest *in vivo*. In some embodiments, the functional enzymes are extracted from the host for production of TAs under *in vitro* conditions. In some instances, the host cells are placed back into a multicellular host organism. The host cells are in any phase of growth, including, but not limited to, stationary phase and log-growth phase, etc. In addition, the cultures themselves may be continuous cultures or they may be batch cultures.

Cells may be grown in an appropriate fermentation medium at a temperature between 20-40ºC. Cells may be grown with shaking at any convenient speed (e.g., 200-500 rpm). Cells may be grown at a suitable pH. Suitable pH ranges for the fermentation may be between pH 4-9. Fermentations may be performed under aerobic, anaerobic, or microaerobic conditions. Any suitable growth medium may be used. Suitable growth media may include, without limitation, common commercially prepared media such as synthetic defined (SD) minimal media or yeast extract peptone dextrose (YEPD) rich media. Any other rich, defined, or synthetic growth media appropriate to the microorganism may be used.

Cells may be cultured in a vessel of essentially any size and shape. Examples of vessels suitable to perform the methods of this disclosure may include, without limitation, multi-well shake plates, test tubes, flasks (baffled and non-baffled), and bioreactors. The volume of the culture may range from 10 microliters to greater than 10,000 liters.

The addition of agents to the growth media that are known to modulate metabolism in a manner desirable for the production of alkaloids may be included. In a non-limiting example, cyclic adenosine 2’3’-monophosphate may be added to the growth media to modulate catabolite repression.

Any convenient cell culture conditions for a particular cell type may be utilized. In certain embodiments, the host cells that include one or more modifications are cultured under standard or readily optimized conditions, with standard cell culture media and supplements. As one example, standard growth media when selective pressure for plasmid maintenance is not required may contain 20 g/L yeast extract, 10 g/L peptone, and 20 g/L dextrose (YPD). Host cells containing plasmids are grown in synthetic complete (SC) media containing 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 20 g/L dextrose supplemented with the appropriate amino acids required for growth and selection. Alternative carbon sources which may be useful for inducible enzyme expression include, but are not limited to, sucrose, raffinose, and galactose. Cells are grown at any convenient temperature (e.g., 30ºC) with shaking at any convenient rate (e.g., 200 rpm) in a vessel, e.g., in test tubes or flasks in volumes ranging from 1-1000 mL, or larger, in the laboratory.

Culture volumes may be scaled up for growth in larger fermentation vessels, for example, as part of an industrial process. The industrial fermentation process may be carried out under closed-batch, fed-batch, or continuous chemostat conditions, or any suitable mode of fermentation. In some cases, the cells may be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for alkaloid production.

A batch fermentation is a closed system, in which the composition of the medium is set at the beginning of the fermentation and not altered during the fermentation process. The desired organism(s) are inoculated into the medium at the beginning of the fermentation. In some instances, the batch fermentation is run with alterations made to the system to control factors such as pH and oxygen concentration (but not carbon). In this type of fermentation system, the biomass and metabolite compositions of the system change continuously over the course of the fermentation. Cells typically proceed through a lag phase, then to a log phase (high growth rate), then to a stationary phase (growth rate reduced or halted), and eventually to a death phase (if left untreated).

A fed-batch fermentation is similar to a batch fermentation, except that the substrate is added in intervals to the system over the course of the fermentation process. Fed-batch systems are used to reduce the impact of catabolite repression on the metabolism of the host cells and under other circumstances where it is desired to have limited amounts of substrate in the growth media.

A continuous fermentation is an open system, in which a defined fermentation medium is added continuously to the bioreactor and an equal amount of fermentation media is continuously removed from the vessel for processing. Continuous fermentation systems are generally operated to maintain steady state growth conditions, such that cell loss due to medium being removed must be balanced by the growth rate in the fermentation. Continuous fermentations are generally operated at conditions where cells are at a constant

high cell density. Continuous fermentations allow for the modulation of one or more factors that affect target product concentration and/or cell growth.

The liquid medium may include, but is not limited to, a rich or synthetic defined medium having an additive component described above. Media components may be dissolved in water and sterilized by heat, pressure, filtration, radiation, chemicals, or any combination thereof. Several media components may be prepared separately and sterilized, and then combined in the fermentation vessel. The culture medium may be buffered to aid in maintaining a constant pH throughout the fermentation.

Process parameters including temperature, dissolved oxygen, pH, stirring, aeration rate, and cell density may be monitored or controlled over the course of the fermentation. For example, temperature of a fermentation process may be monitored by a temperature probe immersed in the culture medium. The culture temperature may be controlled at the set point by regulating the jacket temperature. Water may be cooled in an external chiller and then flowed into the bioreactor control tower and circulated to the jacket at the temperature required to maintain the set point temperature in the vessel.

Additionally, a gas flow parameter may be monitored in a fermentation process. For example, gases may be flowed into the medium through a sparger. Gases suitable for the methods of this disclosure may include compressed air, oxygen, and nitrogen. Gas flow may be at a fixed rate or regulated to maintain a dissolved oxygen set point.

The pH of a culture medium may also be monitored. In examples, the pH may be monitored by a pH probe that is immersed in the culture medium inside the vessel. If pH control is in effect, the pH may be adjusted by acid and base pumps which add each solution to the medium at the required rate. The acid solutions used to control pH may be sulfuric acid or hydrochloric acid. The base solutions used to control pH may be sodium hydroxide, potassium hydroxide, or ammonium hydroxide.

Further, dissolved oxygen may be monitored in a culture medium by a dissolved oxygen probe immersed in the culture medium. If dissolved oxygen regulation is in effect, the oxygen level may be adjusted by increasing or decreasing the stirring speed. The dissolved oxygen level may also be adjusted by increasing or decreasing the gas flow rate. The gas may be compressed air, oxygen, or nitrogen.

Stir speed may also be monitored in a fermentation process. In examples, the stirrer motor may drive an agitator. The stirrer speed may be set at a consistent rpm throughout the fermentation or may be regulated dynamically to maintain a set dissolved oxygen level.

Additionally, turbidity may be monitored in a fermentation process. In examples, cell density may be measured using a turbidity probe. Alternatively, cell density may be measured by taking samples from the bioreactor and analyzing them in a spectrophotometer. Further, samples may be removed from the bioreactor at time intervals

through a sterile sampling apparatus. The samples may be analyzed for alkaloids produced by the host cells. The samples may also be analyzed for other metabolites and sugars, the depletion of culture medium components, or the density of cells.

In another example, a feed stock parameter may be monitored during a fermentation process. In particular, feed stocks including sugars and other carbon sources, nutrients, and cofactors that may be added into the fermentation using an external pump. Other components may also be added during the fermentation including, without limitation, anti-foam, salts, chelating agents, surfactants, and organic liquids.

Any convenient codon optimization techniques for optimizing the expression of heterologous polynucleotides in host cells may be adapted for use in the subject host cells and methods, see e.g., Gustafsson, C. et al. (2004) *Trends Biotechnol*, *22*, 346-353, which is incorporated by reference in its entirety.

The subject method may also include adding a starting compound to the cell culture. Any convenient methods of addition may be adapted for use in the subject methods. The cell culture may be supplemented with a sufficient amount of the starting materials of interest (e.g., as described herein), e.g., a mM to μM amount such as between about 1-5 mM of a starting compound. It is understood that the amount of starting material added, the timing and rate of addition, the form of material added, etc., may vary according to a variety of factors. The starting material may be added neat or pre-dissolved in a suitable solvent (e.g., cell culture media, water, or an organic solvent). The starting material may be added in concentrated form (e.g., 10x over desired concentration) to minimize dilution of the cell culture medium upon addition. The starting material may be added in one or more batches, or by continuous addition over an extended period of time (e.g., hours or days).

Methods for Isolating Products from the Fermentation Medium

The subject methods may also include recovering the TA of interest from the cell culture. Any convenient methods of separation and isolation (e.g., chromatography methods or precipitation methods) may be adapted for use in the subject methods to recover the TA of interest from the cell culture. Filtration methods may be used to separate soluble from insoluble fractions of the cell culture. In some cases, liquid chromatography methods (e.g., reverse phase HPLC, size exclusion, normal phase chromatography) may be used to separate the TA of interest from other soluble components of the cell culture. In some cases, extraction methods (e.g., liquid extraction, pH based purification, etc.) may be used to separate the TA of interest from other components of the cell culture.

The produced alkaloids may be isolated from the fermentation medium using methods known in the art. A number of recovery steps may be performed immediately after (or in some instances, during) the fermentation for initial recovery of the desired product.

Through these steps, the alkaloids (e.g., TAs) may be separated from the cells, cellular debris and waste, and other nutrients, sugars, and organic molecules may remain in the spent culture medium. This process may be used to yield a TA-enriched product.

In an example, a product stream having a tropane alkaloid (TA) product is formed by providing engineered yeast cells and a feedstock including nutrients and water to a batch reactor. The engineered yeast cells may have at least one modification selected from the group consisting of: a feedback inhibition alleviating mutation in a biosynthetic enzyme gene native to the cell; a transcriptional modulation modification of a biosynthetic enzyme gene native to the cell; and an inactivating mutation in an enzyme native to the cell. When the engineered yeast cells are within the batch reactor, the engineered yeast cells may be subjected to fermentation. In particular, the engineered yeast cells may be subjected to fermentation by incubating the engineered yeast cells for a time period of at least about 5 minutes to produce a solution comprising the TA product and cellular material. Once the engineered yeast cells have been subjected to fermentation, at least one separation unit may be used to separate the TA product from the cellular material to provide the product stream comprising the TA product. In particular, the product stream may include the TA product as well as additional components, such as a clarified yeast culture medium. Additionally, a TA product may comprise one or more TAs of interest, such as one or more TA compounds.

Different methods may be used to remove cells from a bioreactor medium that include a TA of interest. In examples, cells may be removed by sedimentation over time. This process of sedimentation may be accelerated by chilling or by the addition of fining agents such as silica. The spent culture medium may then be siphoned from the top of the reactor or the cells may be decanted from the base of the reactor. Alternatively, cells may be removed by filtration through a filter, a membrane, or other porous material. Cells may also be removed by centrifugation, for example, by continuous flow centrifugation or by using a continuous extractor.

If some valuable TAs of interest are present inside the cells, the cells may be permeabilized or lysed and the cell debris may be removed by any of the methods described above. Agents used to permeabilize the cells may include, without limitation, organic solvents (e.g., DMSO) or salts (e.g., lithium acetate). Methods to lyse the cells may include the addition of surfactants such as sodium dodecyl sulfate, or mechanical disruption by bead milling or sonication.

TAs of interest may be extracted from the clarified spent culture medium through liquid-liquid extraction by the addition of an organic liquid that is immiscible with the aqueous culture medium. Examples of suitable organic liquids include, but are not limited to, isopropyl myristate, ethyl acetate, chloroform, butyl acetate, methylisobutyl ketone,

methyl oleate, toluene, oleyl alcohol, ethyl butyrate. The organic liquid may be added to as little as 10% or as much as 100% of the volume of aqueous medium.

In some cases, the organic liquid may be added at the start of the fermentation or at any time during the fermentation. This process of extractive fermentation may increase the yield of TAs of interest from the host cells by continuously removing TA precursors or TAs to the organic phase.

Agitation may cause the organic phase to form an emulsion with the aqueous culture medium. Methods to encourage the separation of the two phases into distinct layers may include, without limitation, the addition of a demulsifier or a nucleating agent, or an adjustment of the pH. The emulsion may also be centrifuged to separate the two phases, for example, by continuous conical plate centrifugation.

Alternatively, the organic phase may be isolated from the aqueous culture medium so that it may be physically removed after extraction. For example, the solvent may be encapsulated in a membrane.

In examples, TAs of interest may be extracted from a fermentation medium using adsorption methods. In particular, TAs of interest may be extracted from clarified spent culture medium by the addition of a resin such as Amberlite® XAD4 or another agent that removes TAs by adsorption. The TAs of interest may then be released from the resin using an organic solvent. Examples of suitable organic solvents include, but are not limited to, methanol, ethanol, ethyl acetate, or acetone.

TAs of interest may also be extracted from a fermentation medium using filtration. At high pH, the TAs of interest may form a crystalline-like precipitate in the bioreactor. This precipitate may be removed directly by filtration through a filter, membrane, or other porous material. The precipitate may also be collected by centrifugation and/or decantation.

The extraction methods described above may be carried out either *in situ* (in the bioreactor) or *ex situ* (e.g., in an external loop through which media flows out of the bioreactor and contacts the extraction agent, then is recirculated back into the vessel). Alternatively, the extraction methods may be performed after the fermentation is terminated using the clarified medium removed from the bioreactor vessel.

Methods for Purifying Products from Alkaloid-Enriched Solutions

Subsequent purification steps may involve treating the post-fermentation TA precursor- or TA-enriched product using methods known in the art to recover individual product species of interest to high purity.

In one example, TA precursors or TAs extracted in an organic phase may be transferred to an aqueous solution. In some cases, the organic solvent may be evaporated by heat and/or vacuum, and the resulting powder may be dissolved in an aqueous solution

of suitable pH. In a further example, the TA precursors or TAs may be extracted from the organic phase by addition of an aqueous solution at a suitable pH that promotes extraction of the TA precursors or TAs into the aqueous phase. The aqueous phase may then be removed by decantation, centrifugation, or another method.

The TA precursor- or TA-containing solution may be further treated to remove metals, for example, by treating with a suitable chelating agent. The TA precursor- or TA-containing solution may be further treated to remove other impurities, such as proteins and DNA, by precipitation. In one example, the TA precursor- or TA-containing solution is treated with an appropriate precipitation agent such as ethanol, methanol, acetone, or isopropanol. In an alternative example, DNA and protein may be removed by dialysis or by other methods of size exclusion that separate the smaller alkaloids from contaminating biological macromolecules.

In further examples, the TA precursor-, TA-, or modified TA-containing solution may be extracted to high purity by continuous cross-flow filtration using methods known in the art.

If the solution contains a mixture of TA precursors or TAs, it may be subjected to acid-base treatment to yield individual TA of interest species using methods known in the art. In this process, the pH of the aqueous solution is adjusted to precipitate individual TA precursors or TAs at their respective pKas.

For high purity, small-scale preparations, the TA precursors or TAs may be purified in a single step by liquid chromatography.

Yeast-Derived Alkaloid APIs Versus Plant-Derived APIs

The clarified yeast culture medium (CYCM) may contain a plurality of impurities. The clarified yeast culture medium may be dehydrated by vacuum and/or heat to yield an alkaloid-rich powder. This product is analogous to the concentrate of nightshade leaves (CNL), which is used by active pharmaceutical ingredient (API) manufacturers for extraction of tropane alkaloids to be subjected to further chemical processing and purification. For the purposes of this invention, CNL is a representative example of any type of purified plant extract from which the desired alkaloids product(s) may ultimately be further purified. Table 11 highlights the impurities in these two products that may be specific to either CYCM or CNL or may be present in both. By analyzing a product of unknown origin for a subset of these impurities, a person of skill in the art could determine whether the product originated from a yeast or plant production host.

API-grade pharmaceutical ingredients are highly purified molecules. As such, impurities that could indicate the plant- or yeast-origin of an API (such as those listed in Tables 2 and 3) may not be present at that API stage of the product. Indeed, many of the

API products derived from yeast strains of the present invention may be largely indistinguishable from the traditional plant-derived APIs. In some cases, however, conventional alkaloid compounds may be subjected to chemical modification using chemical synthesis approaches which may show up as chemical impurities in plant-based products that require such chemical modifications. For example, chemical derivatization may often result in a set of impurities related to the chemical synthesis processes. In certain situations, these modifications may be performed biologically in the yeast production platform, thereby avoiding some of the impurities associated with chemical derivation from being present in the yeast-derived product. In particular, these impurities from the chemical derivation product may be present in an API product that is produced using chemical synthesis processes but may be absent from an API product that is produced using a yeast-derived product. Alternatively, if a yeast-derived product is mixed with a chemically derived product, the resulting impurities may be present but in a lesser amount than would be expected in an API that only or primarily contains chemically derived products. In this example, by analyzing the API product for a subset of these impurities, a person of skill in the art could determine whether the product originated from a yeast production host or the traditional chemical derivatization route.

Non-limiting examples of impurities that may be present in chemically-derivatized tropane alkaloid APIs but not in biosynthesized APIs include hydrogen halides such as hydrogen chloride, hydrogen iodide, and hydrogen bromide formed by chemical N-alkylation, such as N-methylation and N-butylation of hyoscyamine and scopolamine.

However, in the case where the yeast-derived compound and the plant-derived compound are both subjected to chemical modification through chemical synthesis approaches, the same impurities associated with the chemical synthesis process may be expected in the products. In such a situation, the starting material (e.g., CYCM or CNL) may be analyzed as described above.

Methods of Engineering Host Cells

Also included are methods of engineering host cells for the purpose of producing TAs of interest or precursors thereof. Inserting DNA into host cells may be achieved using any convenient methods. The methods are used to insert the heterologous coding sequences into the host cells such that the host cells functionally express the enzymes and convert starting compounds of interest into product TAs of interest.

Any convenient promoters may be utilized in the subject host cells and methods. The promoters driving expression of the heterologous coding sequences may be constitutive promoters or inducible promoters, provided that the promoters are active in the host cells. The heterologous coding sequences may be expressed from their native

promoters, or non-native promoters may be used. Such promoters may be low to high strength in the host in which they are used. Promoters may be regulated or constitutive. In certain embodiments, promoters that are not glucose repressed, or repressed only mildly by the presence of glucose in the culture medium, are used. Promoters of interest include but are not limited to, promoters of glycolytic genes such as the promoter of the *B. subtilis* tsr gene (encoding the promoter region of the fructose bisphosphate aldolase gene) or the promoter from yeast *S. cerevisiae* gene coding for glyceraldehyde 3-phosphate dehydrogenase (GPD, GAPDH, or TDH3), the ADH1 promoter of baker's yeast, the phosphate-starvation induced promoters such as the PHO5 promoter of yeast, the alkaline phosphatase promoter from *B. licheniformis*, yeast inducible promoters such as Gal1-10, Gal1, GalL, GalS, repressible promoter Met25, tetO, and constitutive promoters such as glyceraldehyde 3-phosphate dehydrogenase promoter (GPD), alcohol dehydrogenase promoter (ADH), translation-elongation factor-1-α promoter (TEF), cytochrome c-oxidase promoter (CYC1), MRP7 promoter, phosphoglycerate kinase (PGK), triose phosphate isomerase (TPI), etc. Autonomously replicating yeast expression vectors containing promoters inducible by hormones such as glucocorticoids, steroids, and thyroid hormones may also be used and include, but are not limited to, the glucorticoid responsive element (GRE) and thyroid hormone responsive element (TRE). These and other examples are described U.S. Pat. No. 7,045,290, which is incorporated by reference, including the references cited therein. Additional vectors containing constitutive or inducible promoters such as α factor, alcohol oxidase, and PGH may be used. Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of genes. Any convenient appropriate promoters may be selected for the host cell, e.g., *E. coli*. One may also use promoter selection to optimize transcript, and hence, enzyme levels to maximize production while minimizing energy resources.

Any convenient vectors may be utilized in the subject host cells and methods. Vectors of interest include vectors for use in yeast and other cells. The types of yeast vectors may be broken up into 4 general categories: integrative vectors (YIp), autonomously replicating high copy-number vectors (YEp or 2μ plasmids), autonomously replicating low copy-number vectors (YCp or centromeric plasmids) and vectors for cloning large fragments (YACs). Vector DNA is introduced into prokaryotic or eukaryotic cells via any convenient transformation or transfection techniques.

UTILITY

The host cells and methods of the invention, e.g., as described above, find use in a variety of applications. Applications of interest include, but are not limited to: research

applications and therapeutic applications. Methods of the invention find use in a variety of different applications including any convenient application where the production of TAs is of interest.

The subject host cells and methods find use in a variety of therapeutic applications. Therapeutic applications of interest include those applications in which the preparation of pharmaceutical products that include TAs is of interest. The host cells described herein produce tropane alkaloid precursors (TA precursors) and TAs of interest. Tropinone and tropine are major branch point intermediates of interest in the synthesis of TAs including engineering efforts to produce end products such as medicinal TA products. The subject host cells may be utilized to produce TA precursors from simple and inexpensive starting materials that may find use in the production of TAs of interest, including tropinone, tropine, TA end products, and derivatives of any of these materials. As such, the subject host cells find use in the supply of therapeutically active TAs or precursors thereof.

In some instances, the host cells and methods find use in the production of commercial scale amounts of TAs or precursors thereof where chemical synthesis of these compounds is low yielding and not a viable means for large-scale production. In certain cases, the host cells and methods are utilized in a fermentation facility that would include bioreactors (fermenters) of e.g., 5,000-200,000 liter capacity allowing for rapid production of TAs of interest or precursors thereof for therapeutic products. Such applications may include the industrial-scale production of TAs of interest from fermentable carbon sources such as cellulose, starch, and free sugars.

The subject host cells and methods find use in a variety of research applications. The subject host cells and methods may be used to analyze the effects of a variety of enzymes on the biosynthetic pathways of a variety of TAs of interest or precursors thereof. In addition, the host cells may be engineered to produce TAs or precursors or derivatives thereof that find use in testing for bioactivity of interest in as yet unproven therapeutic functions. For example, TAs in which functional groups have been added or removed to/from the nitrogen heteroatom (*N*-substituted TAs), such as nortropane alkaloids (*N*-demethylation) and tropane *N*-oxides (*N*-oxidization), are of pharmaceutical interest as they differ in solubility, bioavailability, and blood-brain barrier permeability (i.e., nervous system toxicity) relative to unmodified TAs. In some cases, the engineering of host cells to include a variety of heterologous coding sequences that encode for a variety of enzymes elucidates the high yielding biosynthetic pathways towards TAs of interest or precursors thereof. In certain cases, research applications include the production of precursors for therapeutic molecules of interest that may then be further chemically modified or derivatized to desired products or for screening for increased therapeutic activities of interest. In some instances, host cell strains are used to screen for enzyme activities that are of interest in such

pathways, which may lead to enzyme discovery via conversion of TA metabolites produced in these strains.

The subject host cells and methods may be used as a production platform for plant specialized metabolites. The subject host cells and methods may be used as a platform for drug library development as well as plant enzyme discovery. For example, the subject host cells and methods may find use in the development of natural product-based drug libraries by taking yeast strains producing interesting scaffold molecules, such as hyoscyamine and scopolamine, and further functionalizing the compound structure through combinatorial biosynthesis or by chemical means. By producing drug libraries in this way, any potential drug hits are already associated with a production host that is amenable to large-scale culture and production. As another example, these subject host cells and methods may find use in plant enzyme discovery. The subject host cells provide a clean background of defined metabolites to express plant expressed sequence tag (EST) libraries to identify new enzyme activities. The subject host cells and methods provide expression methods and culture conditions for the functional expression and increased activity of plant enzymes in yeast.

KITS AND SYSTEMS

Aspects of the invention further include kits and systems, where the kits and systems may include one or more components employed in methods of the invention, e.g., host cells, starting compounds, heterologous coding sequences, vectors, culture medium, etc., as described herein. In some embodiments, the subject kit includes a host cell (e.g., as described herein), and one or more components selected from the following: starting compounds, a heterologous coding sequence and/or a vector including the same, vectors, growth feedstock, components suitable for use in expression systems (e.g., cells, cloning vectors, multiple cloning sites (MCS), bi-directional promoters, an internal ribosome entry site (IRES), etc.), and a culture medium.

Any of the components described herein may be provided in the kits, e.g., host cells including one or more modifications, starting compounds, culture medium, etc. A variety of components suitable for use in making and using heterologous coding sequences, cloning vectors and expression systems may find use in the subject kits. Kits may also include tubes, buffers, etc., and instructions for use. The various reagent components of the kits may be present in separate containers, or some or all of them may be pre-combined into a reagent mixture in a single container, as desired.

Also provided are systems for producing a TA of interest, where the systems may include engineered host cells including one or more modifications (e.g., as described herein), starting compounds, culture medium, a fermenter and fermentation equipment, e.g.,

an apparatus suitable for maintaining growth conditions for the host cells, sampling and monitoring equipment and components, and the like. A variety of components suitable for use in large scale fermentation of yeast cells may find use in the subject systems.

In some cases, the system includes components for the large scale fermentation of engineered host cells, and the monitoring and purification of TA compounds produced by the fermented host cells. In certain embodiments, one or more starting compounds (e.g., as described herein) are added to the system, under conditions by which the engineered host cells in the fermenter produce one or more desired TA products or precursors thereof. In some instances, the host cells produce a TA of interest (e.g., as described herein). In certain cases, the TA products of interest are medicinal TA products, such as hyoscyamine, N-methylhyoscyamine, anisodamine, scopolamine, N-methylscopolamine, and N-butylscopolamine.

In some cases, the system includes means for monitoring and or analyzing one or more TA compounds or precursors thereof produced by the subject host cells. For example, a LC-MS analysis system as described herein, a chromatography system, or any convenient system where the sample may be analyzed and compared to a standard, e.g., as described herein. The fermentation medium may be monitored at any convenient times before and during fermentation by sampling and analysis. When the conversion of starting compounds to TA products or precursors of interest is complete, the fermentation may be halted and purification of the TA products may be done. As such, in some cases, the subject system includes a purification component suitable for purifying the TA products or precursors of interest from the host cell medium into which it is produced. The purification component may include any convenient means that may be used to purify the TA products or precursors of fermentation, including but not limited to, silica chromatography, reverse-phase chromatography, ion exchange chromatography, HIC chromatography, size exclusion chromatography, liquid extraction, and pH extraction methods. In some cases, the subject system provides for the production and isolation of TA fermentation products of interest following the input of one or more starting compounds to the system.

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.), but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

EXAMPLE METHODS

The following section provides examples of methods and procedures which can be used to construct, culture, and test microbial strains, such as yeast strains, for the production of TA precursors and TAs, as well as to conduct fermentations of such strains to produce TA precursors and TAs. Also included are examples of methods, procedures, and materials which can be used to generate the DNA sequences required for modification of microbial hosts, and to introduce desired DNA sequences into microbial hosts.

Chemical compounds and standards. Chemical standards of TA precursors and TAs for verifying the identity of and quantifying metabolites produced by engineered host cells may be purchased from commercial vendors. For example, putrescine dihydrochloride, N-methylputrescine, hygrine, tropinone, tropine, (S)-hyoscyamine hydrobromide, anisodamine, and (S)-scopolamine hydrobromide may be purchased from Santa Cruz Biotechnology (Dallas, TX). 4-(Methylamino)butyric acid hydrochloride may be purchased from Sigma (St. Louis, MO). γ-Methylaminobutyraldehyde (4MAB) diethyl acetal and (R)-littorine hydrochloride may be purchased from Toronto Research Chemicals (Toronto, ON). A chemical standard for NMPy can be synthesized by deprotecting one volume of the diethyl acetal with five volumes of 2 M HCl at 60 ºC for 30 min as described previously (see Feth, F., Wray, V. & Wagner, K. G. Determination of methylputrescine oxidase by high performance liquid chromatography. Phytochemistry 24, 1653–1655 (1985)), incubating overnight at room temperature, and then washing the resulting concentrate twice with three volumes of diethyl ether to remove residual organic impurities. Norhyoscyamine (noratropine), norscopolamine, and hyoscyamine N-oxide hydrochloride can be purchased from Sigma. Scopolamine N-oxide hydrobromide monohydrate can be purchased from Selleck Chem.

*Plasmid construction***.** Oligonucleotides used for generation of novel DNA sequences by polymerase chain reaction (PCR) and for DNA sequencing can be obtained from a DNA synthesis company, such as IDT DNA, Twist Bioscience, or the Stanford Protein and Nucleic Acid Facility (Stanford, CA). Native yeast genes can be amplified from *S. cerevisiae* genomic DNA via colony PCR (see Kwiatkowski, T. J., Zoghbi, H. Y., Ledbetter, S. A., Ellison, K. A. & Chinault, A. C. Rapid identification of yeast artificial chromosome clones by matrix pooling and crude iysate PCR. Nucleic Acids Res. 18, 7191 (1990)). Gene sequences for heterologous enzymes may be codon-optimized to improve expression in *S. cerevisiae* using suitable codon optimization software, such as the GeneArt GeneOptimizer software (Thermo Fisher Scientific). Heterologous gene sequences can then be synthesized as linear, double-stranded DNA fragments by a commercial DNA synthesis company. Two types of plasmids can be used for gene expression in yeast: direct expression (DE)

plasmids for testing biosynthetic genes of interest and yeast integration (YI) holding plasmids to provide a template for genomic integration of selected promoter-gene-terminator cassettes.

DE plasmids comprise a gene of interest flanked by a constitutive promoter and terminator, a low-copy CEN6/ARS4 yeast origin of replication, and an auxotrophic selection marker. DE plasmids may be constructed by PCR-amplifying genes of interest to append 5' and 3' restriction sites using primer overhangs, digesting PCR products or synthesized gene fragments with appropriate pairs of restriction enzymes (for example, SpeI, BamHI, EcoRI, PstI, or XhoI), and then ligating gene fragments into similarly digested vectors with suitable yeast promoters, terminators, and replication sequences, such as plasmids pAG414GPD-ccdB, pAG415GPD-ccdB, or pAG416GPD-ccdB (see Alberti, S., Gitler, A. D. & Lindquist, S. A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast 24, 913–9 (2007)) using T4 DNA ligase.

YI plasmids comprise a gene of interest flanked by a constitutive promoter and terminator but lack a yeast origin of replication or auxotrophic selection marker. YI plasmids may be constructed by linearizing empty holding vectors with suitable promoters and terminators using 'around-the-horn' PCR with primers designed to bind at the 3' and 5' ends of the promoter and terminator, respectively. Genes of interest can also be PCR-amplified to append 5' and 3' overhangs with 35-40 bp of homology to the termini of the linearized vector backbones. Assembly of genes into YI vectors may then be performed using Gibson assembly. DE plasmids expressing GFP fusions of biosynthetic enzymes may be prepared by first assembling PCR-amplified DNA fragments separately encoding GFP, the target enzyme, and a YI vector backbone using Gibson assembly, and subsequently subcloning the fusion constructs from YI plasmids into DE vectors using restriction enzymes and ligation cloning as described.

PCR amplification may be performed using any high-fidelity recombinant DNA polymerase available from commercial suppliers and linear DNA may be purified using a suitable DNA column purification kit. Assembled plasmids can be propagated in any chemically competent *E. coli* strain using heat-shock transformation and selection in Luria-Bertani (LB) broth or on LB-agar plates with carbenicillin (100 µg/mL), kanamycin (50 µg/mL), or another antibiotic selection. Plasmid DNA can be isolated by alkaline lysis from overnight *E. coli* cultures grown at 37 °C and 250 rpm in selective LB media using plasmid purification columns according to the manufacturer's protocol. Plasmid sequences should be verified by Sanger sequencing.

*Yeast strain construction***.** Any suitable laboratory strain of yeast may be used as a host organism. Yeast strains described in the examples of the Experimental section are derived from the parental strain CEN.PK2-1D (see Entian, K. D. & Kötter, P. 25 Yeast

Genetic Strain and Plasmid Collections. Methods Microbiol. 36, 629–666 (2007)), referred to as CEN.PK2. Strains can be grown non-selectively in yeast-peptone media supplemented with 2% w/v dextrose (YPD media), yeast nitrogen base (YNB) defined media supplemented with synthetic complete amino acid mixture (YNB-SC) and 2% w/v dextrose, or on agar plates of the aforementioned media. Strains transformed with plasmids bearing auxotrophic selection markers (*URA3*, *TRP1*, *HIS3*, and/or *LEU2*) may be grown selectively in YNB media supplemented with 2% w/v dextrose and the appropriate dropout solution (YNB-DO) or on YNB-DO agar plates. Yeast strains which are deficient in acetate metabolism can be grown on the aforementioned media supplemented with 0.1% w/v potassium acetate (i.e., YPAD or YNBA).

Yeast genomic modifications may be performed using the CRISPRm method (see Ryan, O. W. et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife 3, 1–15 (2014)). CRISPRm plasmids express *Streptococcus pyogenes* Cas9 and a single guide RNA (sgRNA) targeting a locus of interest in the yeast genome, and may be constructed by assembly PCR and Gibson assembly of DNA fragments encoding *Sp*Cas9, tRNA promoter and HDV ribozyme, a 20-nt guide RNA sequence, and tracrRNA and terminator. For gene insertions, integration fragments comprising one or more genes of interest flanked by unique promoters and terminators may be constructed using PCR amplification and cloned into holding vectors by Gibson assembly. Integration fragments are PCR amplified using a suitable high-fidelity DNA polymerase with flanking 40 bp microhomology regions to adjacent fragments and/or to the yeast genome at the integration site. For gene disruptions, integration fragments comprise 6-8 stop codons in all three reading frames flanked by 40 bp of microhomology to the disruption site, which is located within the first half of the open reading frame. For complete gene deletions, integration fragments comprise an auxotrophic marker gene flanked by 40 bp of microhomology to the deletion site. Each integration fragment is co-transformed with the CRISPRm plasmid targeting the desired genomic site. Positive integrants may be identified by yeast colony PCR, Sanger sequencing, and/or functional screening by liquid chromatography and tandem mass spectrometry (LC-MS/MS).

*Yeast transformations***.** Yeast strains may be transformed using any suitable method, including heat-shock, electroporation, and chemical transformation. For example, yeast strains described in the examples of the Experimental section were chemically transformed using the Frozen-EZ Yeast Transformation II Kit (Zymo Research). Individual yeast colonies are inoculated into YP(A)D or YNB-SC media and grown overnight at 30 °C and 250-460 rpm. Saturated cultures are back-diluted between 1:10 and 1:50 in fresh media and grown for an additional 5-7 hours to reach exponential phase. Cultures are pelleted by centrifugation at 500×*g* for 4 min and then washed twice by resuspending the pellet in 50

mM Tris-HCl buffer, pH 8.5. Washed pellets are resuspended in 20 µL of EZ2 solution per transformation and then mixed with 100-600 ng of total DNA and 200 µL of EZ3 solution. The yeast suspensions are incubated at 30 °C with gentle rotation for one hour. For plasmid transformations, the transformed yeast are directly plated onto YNB(A)-DO agar plates. For Cas9-mediated chromosomal modifications, yeast suspensions are instead mixed with 1 mL YP(A)D or YNB-SC media, pelleted by centrifugation at 500×*g* for 4 min, and then resuspended in 250 µL of fresh YP(A)D or YNB-SC media. The suspensions are then incubated at 30 °C with gentle rotation for an additional two hours to enable production of G418 resistance proteins and then spread onto YP(A)D or YNB-SC plates containing 400 mg/L G418 (geneticin) sulfate. Plates are then incubated at 30 °C for 48-96 hours to allow colony formation.

*Spot dilution assays***.** Strains are inoculated into YNB(A)-DO media and grown overnight at 30 ºC and 250 rpm. Saturated overnight cultures are pelleted by centrifugation at 500×g for 4 min and resuspended in sterile Tris-HCl buffer, pH 8.0 to a concentration of 10^7 cells/mL based on OD_{600}. Ten-fold serial dilutions of each strain are then prepared in Tris-HCl buffer and 10 µL of each dilution is spotted on pre-warmed YNB(A)-DO plates. Plates are incubated at 30 ºC and imaged after 48 hours.

*Growth conditions for metabolite assays***.** Small-scale metabolite production tests may be conducted in YNB(A)-SC or YNB(A)-DO media. Metabolite production assays can be performed in YNB-SC or YNB-DO media supplemented with 2% dextrose and 5% glycerol (YNB-G) for optimal TA production in at least three replicates (see P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619 (2020)). As TA biosynthesis is significantly enhanced by higher starting cell densities (see preceding reference), multiple yeast colonies may be initially inoculated in triplicate into 1.5 ml selective or non-selective YNB-G and grown to saturation (~20–24 h) at 30 °C and 460 rpm, pelleted by centrifugation at 500g for 4 min and 3,000g for 1 min, resuspended in 1 ml of fresh selective or non-selective YNB-G media supplemented with 15 mg l^{-1} Fe^{2+} from iron (II) sulfate and 50 mM 2-oxoglutarate (5), and then 300 µl transferred into 2 ml deep-well 96-well plates sealed with AeraSeal gas-permeable film (Excel Scientific). Cultures may then be grown for 72–96 h at 25 °C, 460 rpm, and 80% relative humidity in a Lab-Therm LT-X shaker (Adolf Kuhner). For metabolite production assays conducted in pH-buffered media, saturated overnight yeast cultures may be grown and pelleted as described above, then resuspended in 1 ml fresh selective or non-selective YNB-G media supplemented with 15 mg l^{-1} Fe^{2+} from iron (II) sulfate, 50 mM 2-oxoglutarate, and 0.1 M potassium phosphate (K_2HPO_4/KH_2PO_4), pH 5.4 – 7.0.

*Analysis of metabolite production***.** Metabolite profiles and titers may be analyzed using liquid chromatography and tandem mass spectrometry (LC-MS/MS). To separate

cells from media for analysis, fermentation cultures may be pelleted by centrifugation at 3,500×*g* for 5 min at 12 °C and 100-200 µL aliquots of the supernatant can then be removed for direct analysis. Metabolite production may be analyzed by LC-MS/MS using any suitable HPLC device paired with a triple quadrupole mass spectrometer, such as the Agilent 1260 Infinity Binary HPLC and Agilent 6420 Triple Quadrupole mass spectrometer. Chromatography may be performed using a C18 reverse phase column, such as a Zorbax EclipsePlus C18 column (2.1 x 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 is operated with a constant flow rate of 0.4 mL/min at 40 °C and a sample injection volume of 5 µL. Compound separation may be performed using any suitable reverse phase gradient; one such example gradient is as follows: 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 is directed to the MS from 0.01-5 min operating with electrospray ionization (ESI) in positive mode, source gas temperature 350 °C, gas flow rate 11 L/min, and nebulizer pressure 40 psi. Metabolites can be quantified by integrated peak area based on multiple reaction monitoring (MRM) parameters and standard curves.

Fluorescence microscopy. Individual colonies of yeast strains transformed with plasmids encoding biosynthetic enzymes fused to fluorescent protein reporters are inoculated into 1 mL YNB-SC or YNB-DO media and grown overnight (~14-18 h) at 30 °C and 250-460 rpm. Overnight cultures are pelleted by centrifugation at 500×*g* for 4 min and resuspended in 2 mL YNB-DO media with 2% w/v dextrose and then grown at 30 °C and 250 rpm for an additional 4-6 hours to reach exponential phase and allow expressed fluorescent proteins to fold completely. For imaging, approximately 5–10 µl of cell suspension was spotted onto a glass microscope slide and covered with a glass coverslip (Thermo Fisher) and then imaged using a suitable upright epifluorescence microscope, such as a Zeiss AxioImager Epifluorescence/Widefield microscope, with a × 64 oil immersion objective. Fluorescence excitation is performed using any suitable illumination source, such as an EXFO X-Cite 120 illumination source, and suitable filter settings (shown are example settings for Semrock Brightline filters: GFP, 472/30 excitation and 520/35 emission; mCherry/DsRed/Cy3/TexasRed, 562/40 excitation and 624/40 emission). Emitted light is captured with a monochrome camera and appropriate control software, such as a Zeiss Axiocam 503 mono camera and Zen Pro software, and subsequent image analysis may be performed using any suitable image analysis suite, including ImageJ/Fiji (NIH). One example of an image analysis protocol is provided as follows. Images are converted to pseudocolor using the ‘merge channels’ and ‘split channels’ functions. For each sample, linear histogram stretching, if used to improve contrast, must be applied across all images

for a given channel. To reduce the interference of light from other focal planes when imaging sub-cellular organelles, 2D digital deconvolution analysis may be applied. First, a theoretical point-spread function (PSF), which mathematically describes the diffraction of light from a point source in a specific imaging setup, must be computed, for example using the 'Diffraction PSF 3D' plugin in ImageJ, for the green and red channels using the following example parameters: index of refraction of the media, 1.518 (lens oil); numerical aperture, 1.40; wavelength (nm), 520 (green) or 624 (red); longitudinal spherical aberration at max. aperture (nm), 0.00 (default); image pixel spacing (nm), 72; slice spacing (nm), 0; width (pixels), 240; height (pixels), 242; depth (slices), 1; normalization, sum of pixel values = 1. Next, green and red channel images are separately deconvolved against the corresponding PSFs using the 'Parallel Spectral Deconvolution 2D' plugin for ImageJ with default settings and auto regularization.

Yeast toxicity assays and growth curves. Candidate transporters for use in non-plant cells for the purpose of enhancing or modulating metabolite localization may be characterized in any suitable strain of the non-plant organism of interest. However, transporter activity is best characterized in a strain which is deficient in the excretion of xenobiotic (foreign) compounds, as the effects of heterologous transporters may be more easily distinguished. An example protocol for the characterization of TA transporters in yeast is provided as follows. Putative TA transporter candidates were characterized in AD12345678 (AD1-8), a yeast strain in which seven of the major ATP-binding cassette transporters (ABCs) implicated in multidrug and xenobiotic resistance, as well as their principal transcriptional regulator PDR3, have been disrupted (see A. Decottignies, et al., ATPase and multidrug transport activities of the overexpressed yeast ABC protein Yor1p. J. Biol. Chem. 273, 12612–12622 (1998); and see also M. Morita, et al., Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. Proc. Natl. Acad. Sci. U. S. A. 106, 2447–2452 (2009)). Strain AD1-8 is transformed as described earlier with a low-copy plasmid harboring a constitutively expressed transporter candidate or blue fluorescent protein as negative control and a suitable auxotrophic marker, such as URA3, LEU2, TRP1, or HIS2. Transformants are inoculated into 1 mL of selective YNB-G media and grown overnight (~14–18 h) at 30 °C and 460 rpm. Overnight cultures are then diluted in triplicate to a starting concentration of OD_{600} = 0.02 into selective YNB-G media containing 0-20 mM alkaloids (or appropriate substrates to be tested), 0 or 0.1 M potassium phosphate buffer (K_2HPO_4/KH_2PO_4, pH 5.4-6.4) or any suitable pH buffer, and 0-100 mM NaCl or KCl or any other salts of interest in a 96- well plate (200 µl culture volume per well). Culture densities (OD_{600}) are then measured every 30 min for 48 h in a suitable plate reader, such as a SpectraMax i3 plate reader (Molecular Devices), with chamber temperature 30 °C and orbital shaking every 30 min.

*Identification of novel gene variants from transcriptome databases***.** Novel genes and variants thereof may be identified using sequence alignment-based searches of transcriptome and genome databases. For example, orthologs of *N. tabacum* N-methylputrescine oxidase (*NtMPO1*) were identified using a tBLASTn search of the transcriptomes of *D. metel* and *A. belladonna* in the 1000 Plants Project database (see Matasci, N. et al. Data access for the 1,000 Plants (1KP) project. Gigascience 3, 17 (2014)). Coding sequences for putative genes identified using these search strategies can then be optimized for yeast expression and then cloned into expression vectors as described previously.

Transcript co-expression analysis via linear regression and clustering. Linear regression analysis and hierarchical clustering strategies familiar to those with skill in the art may be used for identification of novel gene candidates from plant transcriptome (RNA sequencing) datasets which are delineated in some manner (tissue, elicitor, stimulus, growth conditions, etc.). One example of such a protocol used for identification of TA transporter candidates is described as follows. Tissue- specific abundances (fragments per kilobase of contig per million mapped reads, FPKM) and putative protein structural and functional annotations for each of 43,861 unique transcripts identified from the *A. belladonna* transcriptome may be obtained from the MSU Medicinal Plant Genomics Resource, or from any suitable transcriptomics repository. The transcripts are filtered for those annotated with any of the following transporter protein family (PFAM) IDs: PF00854, PF16974, PF01554, PF00664, PF00005; or any of the following functional annotation keywords: efflux, transporter, ABC. In addition, any transcripts with functional annotations containing the keywords putrescine or tropinone, or with locus ID 4635 (corresponding to hyoscyamine dehydrogenase) may be included in the filter as positive control TA-associated genes to validate clustering with bait genes. Next, mean tissue-specific expression profiles are generated for CYP80F1 and H6H as bait genes. For each of the two, linear regression models are constructed to express the bait gene expression profile as a linear function of each candidate gene profile, and correlation P values are extracted for each candidate (null hypothesis: slope = 0; Student's t-test on estimate of slope coefficient). The candidates identified using each of the two bait genes are pooled and duplicates should be removed. Combined P values for each candidate are computed as the sum of the log10(P values) of the correlations with each of the two bait genes. Candidates are ranked by combined P value; those with $P < 0.01$ may then be further arranged by distance from bait genes via hierarchical clustering of tissue-specific expression profiles.

Phylogenetic analysis of TA transporter candidates. Common phylogenetic analyses familiar to those with skill in the art, such as the construction of phylogenetic trees, may be used to further evaluate gene candidates identified using computational methods described

earlier; an example of such an analysis for three putative TA transporter candidates (AbPUP1, AbCT1, and AbLP1) is provided as follows. Phylogenetic tree construction is based on BLASTp searches using AbPUP1, AbCT1, and AbLP1 as queries against the UniProt/SwissProt database (annotated sequences only). Sequences chosen for tree construction may include any experimentally characterized plant alkaloid transporters and any other selected characterized small molecule transporters. Phylogenetic relationships are derived via bootstrap neighbour-joining with n = 1,000 (or any suitable number sufficient for convergence) iterations in ClustalX2.

Enzyme structural analysis. Heterologous enzymes may be analyzed for structural features that may prove problematic during expression in yeast, such as large unstructured regions, by examining homology models constructed using any suitable homology modeling or *de novo* structure prediction software, such as RaptorX, C-I-TASSER, Rosetta. Prediction of transmembrane domains in a protein or enzyme of interest may be performed using the TMHMM v. 2.0 web server (see A. Krogh et al., Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. 305, 567–580 (2001)) and should be corroborated against hydropathy plots generated using the EMBOSS Pepwindow web server (see F. Madeira, et al., The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641 (2019)). Resultant protein models can be visualized using any three-dimensional molecular viewing software, such as PyMOL (Schrodinger) or UCSF Chimera. Enzyme affinity for specific substrates may be analyzed using any suitable ligand docking simulation software, such as AutoDock, SwissDock, GOLD, or Glide.

Verification of full-length TA transporter cDNA sequences using plant tissue. Sequences encoding proteins, transporters, or enzymes of interest derived from plants may be verified by comparison against authentic plant tissue samples. As an example protocol, the full-length cDNA sequences for AbPUP1 and AbLP1 may be verified using authentic *A. belladonna* tissue samples. *A. belladonna* plants may be purchased from any suitable nursery or vendor and maintained in appropriate potting soil under a 16:8 (or similar) light:dark cycle. Secondary roots are excised with a clean razor blade, washed thoroughly under distilled water, and total RNA is extracted using any suitable kit, such as an RNeasy Plant Mini kit (Qiagen), according to the manufacturer's instructions. The polyA fraction in 1 µg of extracted total RNA is then converted to cDNA using a reverse transcriptase enzyme kit, such as SuperScript III Reverse Transcriptase, and oligo-dT12-18 primers per the manufacturer's instructions. The open reading frames of AbPUP1 and AbLP1 are PCR-amplified from root cDNA and Sanger sequenced using suitable DNA primers.

Analysis of protein expression in yeast by Western blot. For immunoblot analysis of yeast-expressed proteins, a suitable strain is transformed with an expression vector

harboring an epitope-tagged protein of interest. Three days post-transformation, transformed colonies are inoculated into 2 mL YNB-DO media and grown overnight (~16-20 h) to stationary phase at 30 °C and 460 rpm. Cells are pelleted by centrifugation at 3,000 × g for 5 min, resuspended in 200 µL H_2O, mixed with 200 µL of 0.2 M NaOH, and incubated at room temperature for 5 min to allow hydrolysis of cell wall glycoproteins. Cells are re-pelleted at 3,000 × g for 5 min, resuspended in 75 µL H_2O, mixed with 25 µL of 4X NuPAGE LDS sample buffer (Thermo Fisher), and then boiled at 95 °C for 3 min to lyse cells. Suspensions are pelleted by centrifugation at 16,000 × g for 5 min to remove insoluble debris and supernatants are transferred to pre-chilled tubes. For analysis under reducing conditions, protein lysates are mixed with β-mercaptoethanol (final concentration 10%) and incubated at 70 °C for 10 min. Approximately 20-40 µg of total protein is loaded onto NuPAGE Bis-Tris 4-12% acrylamide gels (Thermo Fisher) with Precision Plus Dual Color protein molecular weight marker (BioRad). Electrophoresis is conducted in 1X NuPAGE MOPS SDS running buffer at 150 V for 90 min. Transfer of protein to a nitrocellulose membranes is performed at 15 V for 15 min using a Trans-Blot Semi-Dry apparatus (BioRad) and NuPAGE transfer buffer (Thermo Fisher) per manufacturer instructions. For reducing conditions, NuPAGE antioxidant (Thermo Fisher) is added to a final concentration of 1X to both the running buffer and transfer buffer. Membranes with transferred protein are washed for 5 min in Tris-buffered saline with Tween (TBS-T; 137 mM NaCl, 2.7 mM KCl, 19 mM Tris base, 0.1% Tween20, pH 7.4) and then blocked with 5% skim milk in TBS-T for 1 h at room temperature. Membranes are incubated overnight at 4 °C with a suitable dilution of an HRP-conjugated antibody in TBS-T with 5% milk, washed three times for 5 min each with TBS-T, and then visualized using Western Pico PLUS HRP substrate (Thermo Fisher) and a suitable imager.

EXPERIMENTAL

A series of specific genetic modifications provide a biosynthetic process in *Saccharomyces cerevisiae* for the production of TAs from simple, inexpensive feedstocks or precursor molecules. Methods for constructing novel strains capable of producing the early TA molecules putrescine, N-methylputrescine, 4-methylaminobutanal, N-methylpyrrolinium (NMPy), tropinone, tropine, phenyllactic acid (PLA), and 1-*O*-β-phenyllactoylglucose (PLA glucoside) from non-TA precursors or simple feedstocks are described. NMPy is the natural precursor to all known TA molecules. Methods for manipulating the regulation of yeast biosynthetic pathways and for optimizing the production of amino acid-derived TA precursors are also described. Methods for constructing novel strains capable of producing non-medicinal TAs such as pseudotropine alkaloids and calystegines from simple feedstocks are described. Additionally, methods for constructing novel strains capable of

producing medicinal TAs such as hyoscyamine, anisodamine, and scopolamine from non-TA precursors or simple feedstocks are described. Furthermore, methods for constructing novel strains capable of producing non-natural TAs such as cinnamoyltropine from non-TA precursors or simple feedstocks are described.

Example 1. Engineering a platform yeast strain for high levels of putrescine production

The tropine moiety of TAs is derived from the amino acid arginine via the polyamine molecule putrescine. Strains of *S. cerevisiae* are developed with improved flux through the arginine and polyamine biosynthesis pathways for the purposes of increasing intracellular concentrations of TA precursor molecules including putrescine, NMP, 4MAB, and NMPy. These strains combine genetic modifications for the purpose of increasing carbon and nitrogen flux from central metabolism towards arginine and polyamine biosynthesis in general, and include the introduction of key heterologous enzymes for additional production of the TA precursor putrescine. Genetic modifications are employed including the introduction of feedback inhibition alleviating mutations to genes encoding native biosynthetic enzymes and regulatory proteins, tuning of transcriptional regulation of native biosynthetic enzymes, deletion or disruption of genes encoding enzymes that divert precursor molecules away from the intended pathway, and introduction of heterologous enzymes for the conversion of endogenous molecules into TA precursor molecules.

1.1) The biosynthetic pathway in the engineered strain incorporates overexpression of native yeast genes involved in arginine metabolism and polyamine biosynthesis (Fig. 4).

1.1.1) Examples of overexpressed native genes in yeast include, but are not limited to: glutamate N-acetyltransferase (Arg2p), which catalyzes the first step in arginine biosynthesis from glutamate; arginase (Car1p), which removes the guanidinium group of arginine to produce ornithine in the mitochondrial matrix; a mitochondrial membrane transporter (Ort1p), which exports ornithine from the mitochondrial matrix to the cytosol; ornithine decarboxylase (Spe1p), which decarboxylates cytosolic ornithine to putrescine; and a polyamine oxidase (Fms1p), which dealkylates spermine and spermidine to putrescine.

1.1.2) The impact of overexpression of these native enzymes on putrescine production was examined by co-transforming a yeast strain with different combinations of three low-copy plasmids, each expressing one of *SPE1*, *ORT1*, *CAR1*, *ARG2*, *FMS1*, or blue fluorescent protein (BFP) as a negative control. The titer of putrescine accumulated in the extracellular

medium of co-transformed cells following 48 hours of growth in selective media was quantified by LC-MS/MS (Fig. 5). Overexpression of *SPE1* alone resulted in a 13.4-fold increase in putrescine titer to 23 mg/L. While co-overexpression of *CAR1* or *ARG2* with *SPE1* resulted in 27% and 12% increases in putrescine production relative to *SPE1* alone, overexpression of *ORT1* with *SPE1* caused a 35% decrease in putrescine titer compared to *SPE1*. Overexpression of any three of *SPE1*, *CAR1*, *ARG2*, and *FMS1* collectively increased extracellular putrescine titers to 34-35 mg/L.

1.2) The biosynthetic pathway in the engineered strain incorporates expression of heterologous enzymes from polyamine production pathways found in organisms other than yeast to further increase putrescine production (Fig. 4).

1.2.1) In addition to the ornithine-dependent pathway found in most plants, animals, and fungi, whereby putrescine is synthesized via deguanidination of arginine followed by decarboxylation of ornithine, many bacteria and plants also express an alternate route through which arginine is first decarboxylated by arginine decarboxylase (ADC) to yield agmatine. In plants, the guanidine group of agmatine is then converted to a urea by an iminohydrolase (AIH) to produce N-carbamoylputrescine (NCP), from which the amide group is then removed by an amidase (CPA) to yield putrescine (see Patel, J. et al. Dual functioning of plant arginases provides a third route for putrescine synthesis. Plant Sci. 262, 62–73 (2017)). Some bacteria have evolved an agmatine ureohydrolase (AUH) enzyme that enables direct removal of the guanidine group from agmatine to produce putrescine without an N-carbamoylated intermediate (see Klein, R. D. et al. Reconstitution of a bacterial/plant polyamine biosynthesis pathway in Saccharomyces cerevisiae. Microbiology 145 (Pt 2, 301–7 (1999)).

1.2.2) To reconstruct the heterologous putrescine biosynthetic pathways in yeast, the following enzymes may be used: ADC, AIH, CPA, and AUH. As an example of an engineered strain which possesses these enzymatic activities, an ADC from oat (*Avena sativa*; *AsADC*) with previously demonstrated activity in *S. cerevisiae* (see Klein, R. D. et al. Reconstitution of a bacterial/plant polyamine biosynthesis pathway in Saccharomyces cerevisiae. Microbiology 145 (Pt 2, 301–7 (1999)), an AIH from *Arabidopsis thaliana* (*AtAIH*), two CPA orthologs from tomato (*Solanum lycopersicum*; *SlCPA*) and *A. thaliana* (*AtCPA*), and two AUHs from *E. coli* (*speB*) and *A. thaliana* (*AtARGAH2*) were selected for expression in yeast.

1.2.3) In order to establish the functionality of each heterologous enzyme in yeast, the

three-step (arginine → agmatine → NCP → putrescine) or two-step (arginine → agmatine → putrescine) putrescine pathways were reconstituted in a stepwise fashion by co-transforming the wild-type yeast strain with low-copy plasmids expressing *AsADC*, *AtAIH*, and either *SlCPA* or *AtCPA*; or *AsADC* and either *speB* or *AtARGAH2*. To eliminate effects on cell growth and metabolite production arising from different levels of auxotrophy, all transformations were performed with three low-copy plasmids harboring different auxotrophic markers, using BFP as a negative control in place of a blank or absent plasmid. The relative accumulation of agmatine, NCP, and putrescine in the extracellular medium of transformed cells following 48 hours of growth in selective media were analyzed by LC-MS/MS, which indicated that all enzymes except for SlCPA and AtARGAH2 retained activity in yeast (Fig. 6, 7). Reconstitution of the plant-specific pathway comprising AsADC, AtAIH, and AtCPA enabled putrescine production at titers of 23 mg/L, a 22-fold improvement relative to wild-type titers. The orthologous CPA from tomato (SlCPA) enabled putrescine production at titers of 4.5 mg/L when combined with AsADC and AtAIH, similar to putrescine levels in cells expressing AsADC and AtAIH. Reconstitution of the bacterial shortcut pathway via AsADC and the *E. coli* ureohydrolase (speB) enabled putrescine production at titers of 34 mg/L, 32-fold higher than wild-type.

1.3) The biosynthetic pathway in the engineered strain incorporates overexpression of native yeast genes involved in arginine and polyamine biosynthesis and expression of heterologous biosynthetic enzymes from polyamine production pathways found in organisms other than yeast to further increase putrescine production.

1.3.1) The top-performing triad of overexpressed native genes for putrescine biosynthesis (*SPE1*, *ARG2*, *CAR1; 1.1.2*) was combined with the top-performing heterologous putrescine pathway (*AsADC*, *speB; 1.2.3*) by co-transforming the wild-type yeast strain with a low-copy plasmid encoding *SPE1*, *AsADC*, and *speB* and low-copy plasmids encoding *ARG2* and *CAR1.* Putrescine titers in the culture medium of transformed cells were measured by LC-MS/MS analysis after 48 hours. The resulting strain produced putrescine at titers of 47 mg/L, (Fig. 10).

1.4) Polyamine biosynthesis in yeast is regulated by several mechanisms (Fig. 8). The biosynthetic pathway in the engineered strain incorporates disruptions of one or more of these regulatory mechanisms to reduce feedback inhibition of putrescine production.

1.4.1) Native yeast genes involved in regulation of polyamine biosynthesis, and which may therefore be disrupted to improve intracellular putrescine accumulation, include but are not

limited to the following examples (Fig. 8). Methylthioadenosine phosphorylase (Meu1p) catalyzes the driving step in the recycling pathway for decarboxylated S-adenosylmethionine (dcSAM), which constitutes the alkyl group donor for conversion of putrescine to spermidine and spermine catalyzed by spermidine synthase (Spe3p) and spermine synthase (Spe4p) (see Chattopadhyay, M. K., Tabor, C. W. & Tabor, H. Methylthioadenosine and polyamine biosynthesis in a Saccharomyces cerevisiae meu1Δ mutant. Biochem. Biophys. Res. Commun. 343, 203–207 (2006)). Methylthioadenosine is known to inhibit the activity of spermidine synthase (see Chattopadhyay, M. K., Tabor, C. W. & Tabor, H. Studies on the regulation of ornithine decarboxylase in yeast: Effect of deletion in the MEU1 gene. Proc. Natl. Acad. Sci. 102, 16158–16163 (2005)). Polyamine biosynthesis is regulated by means of an antizyme-mediated negative feedback loop conserved across fungi and metazoans (see Pegg, A. E. Regulation of ornithine decarboxylase. Journal of Biological Chemistry 281, 14529–14532 (2006)). In yeast, the *OAZ1* gene comprises two exons separated by a single nucleotide which collectively encode antizyme-1, a competitive inhibitor of ornithine decarboxylase (Spe1p). A polyamine-induced ribosomal frameshifting mechanism enables translation of the full-length antizyme only at high polyamine levels, thereby imposing feedback inhibition of their biosynthesis. Finally, polyamine uptake from the extracellular environment is mediated by a signaling pathway involving Agp2p, a permease of the plasma membrane with affinity towards carnitine, spermidine, and spermine, and Sky1p, a protein kinase thought to interact with Agp2p.

1.4.2) Yeast single-gene disruption strains for each of *MEU1*, *OAZ1*, *SPE4*, *SKY1*, and *AGP2* were constructed by inserting a series of tandem nonsense mutations within the first third of each open reading frame in wild-type yeast. To characterize the effects of each regulatory disruption in the context of the native and heterologous putrescine production pathways, either yeast ODC (*SPE1*) was overexpressed, or *AsADC* and *speB* were co-expressed from low-copy plasmids in each of the single-gene disruption strains. Putrescine titers in the extracellular medium were measured via LC-MS/MS after 72 hours of growth (Fig. 9). *MEU1* disruption improved putrescine titers by 68% when the native putrescine production pathway via *SPE1* was overexpressed. Similarly, *OAZ1* disruption markedly improved putrescine production by 174% when combined with overexpression of *SPE1*. Disruption of *OAZ1* resulted in a 21-fold increase in putrescine titer in untransformed cells with neither the native nor heterologous putrescine pathways overexpressed. Disruption of *SKY1* and *AGP2* resulted in 29% and 14% respective increases in putrescine titer when overexpressed with *SPE1*. *SKY1* disruption resulted in a 41% decrease in putrescine titer when combined with heterologous expression of *AsADC* and *speB*.

1.5) The biosynthetic pathway in the engineered strain combines the *MEU1* and *OAZ1* regulatory gene knockouts with overexpression of the native and heterologous putrescine biosynthetic genes in order to further increase putrescine production in the engineered strain. Additional copies of the native arginine and polyamine biosynthetic genes *ARG2*, *CAR1*, and *FMS1* were integrated into the genome of a *meu1*/*oaz1* double-disruption strain. This strain was transformed with a low-copy plasmid expressing *SPE1*, *AsADC*, and *speB*. LC-MS/MS analysis of the extracellular medium of this transformed strain indicated that putrescine titers reached 86 mg/L after 48 hours of growth in selective media (Fig. 10).

Example 2. Engineering yeast strains for production of NMPy

Strains of *S. cerevisiae* are developed by modifying the putrescine-overproducing strain developed in Example 1 for the production of the TA precursor NMPy. These strains combine genetic modifications for the purpose of increasing carbon and nitrogen flux from putrescine towards NMPy biosynthesis, and include the introduction of key heterologous enzymes for production of the TA precursors NMP, 4MAB, and NMPy. Genetic modifications are employed including modification of the N- and/or C-terminal domains of enzymes of interest to improve activity in a heterologous host, and deletion or disruption of genes encoding enzymes that diver precursor molecules away from the intended pathway.

2.1) The biosynthetic pathway in the engineered strain enables production of NMPy from endogenous putrescine. Putrescine is first converted to N-methylputrescine (NMP) by a SAM-dependent N-methyltransferase (PMT), which is subsequently oxidized to 4-methylaminobutanal (4MAB) by a copper-dependent diamine oxidase (MPO). 4MAB, like many aldehyde compounds, is unstable in aqueous solution and spontaneously cyclizes via base-catalyzed nucleophilic attack to form NMPy (Fig. 11).

2.1.1) The putrescine overproducing strain of Example 1.5, which harbors a low-copy plasmid expressing *SPE1*, *AsADC*, and *speB* for putrescine overproduction, was co-transformed with additional low-copy plasmids expressing a PMT from *A. belladonna* (*AbPMT1*) and a subsequent MPO enzyme from *Nicotiana tabacum* (*NtMPO1*). The accumulation of intermediates in the extracellular medium of transformed cells expressing each successive enzyme between putrescine and NMPy was compared via LC-MS/MS analysis after 48 hours of growth. The immediate product of NtMPO1 (4MAB) as well as its spontaneous cyclization product (NMPy) were produced with expression of AbPMT1 and NtMPO1 (Fig. 11), as well as their precursors, NMP and putrescine (Fig. 12).

2.1.2) The accumulation of NMP was measured in the growth medium of putrescine-overproducing yeast strains with and without disruption of the *MEU1* gene (described in Example 1.4.2) by LC-MS/MS analysis. This analysis indicated that the prior disruption of *MEU1* in the putrescine-overproducing strain and its concomitant impact on SAM recycling did not inhibit putrescine N-methylation by AbPMT1 (Fig. 13).

2.2) Enzymes may localize to different sub-cellular compartments when heterologously expressed than in their original host organism, resulting in reduced function. The biosynthetic pathway in the engineered strain may incorporate modifications to the polypeptide sequences of native and heterologous enzymes to induce localization of these modified enzymes to sub-cellular compartments other than those to which they localize naturally. For example, prior studies have shown that while NtPMT is expressed in the cytosol of tobacco cells, NtMPO1 localizes to the peroxisome lumen (see Naconsie, M., Kato, K., Shoji, T. & Hashimoto, T. Molecular evolution of n-methylputrescine oxidase in Tobacco. Plant Cell Physiol. 55, 436–444 (2014)).

2.2.1) The sub-cellular localization of NtMPO1 was examined by performing *in silico* prediction of enzyme subcellular localization using the SherLoc2 utility for signal peptide detection (see Briesemeister, S. et al. SherLoc2: A high-accuracy hybrid method for predicting subcellular localization of proteins. J. Proteome Res. 8, 5363–5366 (2009)). This analysis indicated that NtMPO1 harbors a strong yeast consensus peroxisome-targeting sequence (PTS) at its C-terminus (Ala-Lys-Leu, denoted PTS1), which suggests that NtMPO1 may localize to peroxisomes when expressed heterologously in yeast (Fig. 14).

2.2.2) Fluorescence microscopy of wild-type yeast cells expressing either N- or C-terminal GFP-tagged *AbPMT1* and *NtMPO1* from low-copy plasmids indicated that while AbPMT1 is found primarily in the cytosol, localization of NtMPO1 to peroxisomes is contingent on an exposed C-terminal PTS (Fig. 15a, 16).

2.2.3) Cytosolic expression of NtMPO1 achieved by masking the C-terminal PTS with a GFP fusion did not significantly impact extracellular 4MAB or NMPy levels (Fig.15b).

2.3) The biosynthetic pathway in the engineered strain may incorporate orthologs of biosynthetic enzymes other than those listed in Table 1. Different orthologs of an enzyme may exhibit significant differences in activity when expressed in heterologous hosts. Therefore, orthologs of biosynthetic enzymes provided as examples herein and listed in Table 1 may also be used in engineered non-plant cells to perform the same biochemical

conversions.

2.3.1) A tBLASTn search of the transcriptomes of *A. belladonna* and *Datura metel* in the 1000 Plants Project database (see Matasci, N. et al. Data access for the 1,000 Plants (1KP) project. Gigascience 3, 17 (2014)) was performed using the amino acid sequence of NtMPO1 as a query and an E-value threshold of 10^{-150}. Two full-length ortholog sequences denoted *AbMPO1* and *DmMPO1* were identified, which each shared 91% sequence identity with *NtMPO1* (Fig. 17a).

2.3.2) Yeast codon-optimized sequences for *AbMPO1* and *DmMPO1* were obtained and cloned into low-copy expression plasmids. To evaluate their activity, each of the three MPO variants was co-expressed with *AbPMT1* from low-copy plasmids in the putrescine-overproducing strain of Example 1.5, and 4MAB and NMPy accumulation were measured in the extracellular medium by LC-MS/MS following 48 hours of growth in selective media. DmMPO1 showed comparable levels of 4MAB and NMPy production to the original NtMPO1 variant (Fig. 17b).

2.3.3) Differences in activity between orthologous enzymes can often be at least partially attributed to structural differences in their active sites. Template-based homology models of NtMPO1, AbMPO1, and DmMPO1 were constructed based on the crystal structure of a *Pisum sativum* copper-containing amino oxidase (PDB: 1KSI) using the RaptorX web server (see Källberg, M. et al. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511–22 (2012)). The homology models indicated that the orthologs possess long, unstructured N- and C-terminal tail regions (Fig. 17c).

2.3.4) Truncations of the two active orthologs, NtMPO1 and DmMPO1, were tested for activity in engineered yeast. N-terminal truncations removed the first 84 and 81 residues of the two orthologs, respectively. C-terminal truncations removed the last 21 residues. C-terminal truncations were also constructed wherein the unstructured tail was removed but the PTS was retained (denoted $^{\Delta C\text{-}PTS1}$). Each of the MPO truncations was coexpressed with *AbPMT1* from low-copy plasmids in the putrescine-overproducing strain of Example 1.5, and 4MAB and NMPy accumulation in the media after 48 hours of growth were quantified by LC-MS/MS. No significant differences in activity were observed between the NtMPO1 truncations (Fig. 18). Removal of the C-terminal unstructured region from DmMPO1 while retaining the C-terminal PTS tripeptide resulted in a 31% increase in extracellular 4MAB levels relative to the wild-type DmMPO1 enzyme.

2.4) The biosynthetic pathway in the engineered strain incorporates one or more genetic modifications to reduce or eliminate the metabolic flux of undesirable side reactions. Biosynthetic enzymes expressed in heterologous hosts may participate in undesirable side reactions that draw metabolite flux away from the biosynthesis of desired compounds. For example, yeast aldehyde dehydrogenases may oxidize heterologous aldehyde molecules, such as 4MAB, to their cognate carboxylic acids. Based on LC-MS/MS analysis, accumulation of 4MAB acid was observed in the growth media of the putrescine-overproducing strain of Example 1.5 when *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$ were co-expressed from low-copy plasmids, but not in the absence of the MPO enzyme (Fig. 11).

2.4.1) Six yeast genes (*ALD2*-*ALD6* and *HFD1*) have been demonstrated in the literature to encode enzymes with aldehyde dehydrogenase activity (see Datta, S., Annapure, U. S. & Timson, D. J. Different specificities of two aldehyde dehydrogenases from Saccharomyces cerevisiae var. boulardii. Biosci. Rep. 37, BSR20160529 (2017); and also Nakahara, K. et al. The Sjögren-Larsson Syndrome Gene Encodes a Hexadecenal Dehydrogenase of the Sphingosine 1-Phosphate Degradation Pathway. Mol. Cell 46, 461–471 (2012)). The *ALD2* and *ALD3* genes encode a pair of nearly identical cytosolic dehydrogenases which catalyze the oxidation of 3-aminopropanal to β-alanine in the biosynthesis of pantothenic acid (see White, W. H., Skatrud, P. L., Xue, Z. & Toyn, J. H. Specialization of Function Among Aldehyde Dehydrogenases : Genetics 163, 69–77 (2003)). The *ALD4*, *ALD5*, and *ALD6* genes respectively encode two mitochondrial and one cytosolic acetaldehyde dehydrogenase which, in addition to oxidizing acetaldehyde to acetate during fermentative growth on glucose and ethanol (see Saint-Prix, F., Bönquist, L. & Dequin, S. Functional analysis of the ALD gene family of Saccharomyces cerevisiae during anaerobic growth on glucose: The NADP+-dependent Ald6p and Ald5p isoforms play a major role in acetate formation. Microbiology 150, 2209–2220 (2004)), have been shown to oxidize an array of diverse aliphatic and aromatic aldehydes to carboxylic acids (see Datta, S., Annapure, U. S. & Timson, D. J. Different specificities of two aldehyde dehydrogenases from Saccharomyces cerevisiae var. boulardii. Biosci. Rep. 37, BSR20160529 (2017)). Individual knockouts strains for these four target genes were constructed by inserting a series of tandem nonsense mutations within the first third of their open reading frames in the putrescine-overproducing strain of Example 1.5. The contribution of each of the four dehydrogenases toward 4MAB oxidation was evaluated by co-expressing *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$ from low-copy plasmids in each single disruption strain and measuring 4MAB acid accumulation in the media by LC-MS/MS after 48 hours of growth. Marginal decreases in 4MAB acid levels were observed with the individual *HFD1* and *ALD4-6* disruptions (Fig. 19).

2.4.2) Although *ALD4-6* are considered essential genes due to their role in acetate and acetyl-CoA production, prior studies have demonstrated that the three genes are at least partially redundant and that the lethal phenotype of double and triple knockouts can be rescued by supplementing media with acetate (see Saint-Prix, F., Bönquist, L. & Dequin, S. Functional analysis of the ALD gene family of Saccharomyces cerevisiae during anaerobic growth on glucose: The NADP+-dependent Ald6p and Ald5p isoforms play a major role in acetate formation. Microbiology 150, 2209–2220 (2004); and also Luo, Z., Walkey, C. J., Madilao, L. L., Measday, V. & Van Vuuren, H. J. J. Functional improvement of Saccharomyces cerevisiae to reduce volatile acidity in wine. FEMS Yeast Res. 13, 485–494 (2013)). A quadruple knockout yeast strain was constructed with disruptions to the open reading frames of *HFD1* and *ALD4-6*, and which expressed both *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$ from low-copy plasmids. This strain showed a 45% reduction in 4MAB acid levels (Fig. 20a) and a concomitant 46% increase in NMPy production compared to the strain with no disruptions (Fig. 20b).

2.4.3) An *ALD*-null strain was constructed by deleting the tandem *ALD2-ALD3* genes from the genome of the quadruple knockout strain of example 2.4.2 and co-expressing *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$ from low-copy plasmids. Following 48 hours of growth, LC-MS/MS analysis indicated that deletion of *ALD2* and *ALD3* completely eliminated the 4MAB acid side product and increased 4MAB and NMPy production by 83% and 75%, respectively, compared to the strain with all six *ALD* genes intact (Fig. 20a, b).

2.4.4) An NMPy-producing yeast strain was constructed by integrating a previously plasmid-borne putrescine-overproduction gene cassette (*SPE1*, *AsADC*, *speB*) into the genome of the *ALD*-null strain of Example 2.4.3, and additionally integrating *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$. LC-MS/MS analysis confirmed that NMPy production in this strain after 48 hours of growth in non-selective media was comparable to that of the ALD-null strain of example 2.4.3 expressing the requisite putrescine production genes, *AbPMT1* and *DmMPO1*$^{\Delta C\text{-}PTS1}$, from low-copy plasmids and cultured in selective media (Fig. 21).

Example 3. Engineered yeast strains for production of tropine from simple sugars and nutrients

A type III polyketide synthase (PKS) and a cytochrome P450 enable conversion of NMPy to tropinone by way of the TA precursor MPOB. Tropinone can be reduced by a stereospecific reductase, denoted tropinone reductase 1 (TR1), to produce tropine (see Kim, N., Estrada, O., Chavez, B., Stewart, C. & D'Auria, J. C. Tropane and Granatane

Alkaloid Biosynthesis: A Systematic Analysis. Molecules 21, (2016)) (Fig. 22).

3.1) The biosynthetic pathway in the engineered strain incorporates a pyrrolidine ketide synthase, a tropinone synthase CYP82M3, one or more cytochrome P450 reductases, and a tropinone reductase 1 to convert NMPy to tropine.

3.1.1) Yeast codon-optimized DNA sequences encoding *A. belladonna* pyrrolidine ketide synthase (*AbPYKS*), tropinone synthase (*AbCYP82M3*), and *Datura stramonium* tropinone reductase 1 (*DsTR1*) were obtained. Yeast codon-optimized sequences for a panel of four different CPRs, including three plant CPRs from *A. thaliana*, *Eschscholzia californica* (California poppy), and *Papaver somniferum* (opium poppy), and the native yeast CPR (*NCP1*), were also obtained for expression in yeast, since P450 enzymes require $NADP^+$-cytochrome P450 reductase (CPR) partners for continued electron exchange. A yeast strain was constructed by integrating *DsTR1* into the genome of the NMPy-producing strain of Example 2.4.4, and expressing *AbPYKS*, *AbCYP82M3*, and each of the four CPRs from low-copy plasmids. To validate enzyme activity and identify potential bottlenecks, the accumulation of NMPy, MPOB, tropinone, and tropine were monitored by LC-MS/MS in the media of the transformed strains after 48 hours of growth (Fig. 23). Comparable levels of *de novo* tropine production (175-210 µg/L) were observed with all four CPR partners under the assay conditions.

3.2) The presence of metabolic bottlenecks, which are defined as biosynthetic enzymes or spontaneous steps whose low activity limits flux through a portion of a biosynthetic pathway, can result in sub-optimal production of desired TAs and precursors.

3.2.1) For example, analysis of the accumulation of TA intermediates in the media of the engineered strains of Example 3.1.1 indicated that although accumulation of tropinone, the product of AbCYP82M3, was minimal, a substantial portion of MPOB produced by AbPYKS remained unconsumed by AbCYP82M3 (Fig. 24).

3.2.2) Integration of the tropine biosynthesis genes into the yeast genome can improve tropine production by enabling more stable *AbCYP82M3* expression. A tropine-producing platform strain was constructed by integrating *AtATR1* with *AbPYKS* and *AbCYP82M3* into the genome of the NMPy-producing strain of Example 3.1.1. Tropine and hygrine accumulation for the integrated strain was compared to plasmid-based expression of the same genes via LC-MS/MS analysis after 48 hours (Fig. 28). Genomic expression of *AbPYKS*, *AbCYP82M3*, and *AtATR1* increased tropine titers by nearly three-fold (565 µg/L)

relative to plasmid-based expression (189 µg/L). The engineered strain also showed a 2.6-fold increase in hygrine accumulation.

3.3) The accumulation of side products in the biosynthetic pathway of the engineered strain can result in sub-optimal production of desired TAs and precursors.

3.3.1) For example, analysis of the accumulation of TA intermediates in the media of the engineered strains of Example 3.1.1 indicated substantial accumulation of hygrine, a derivative of NMPy, to titers almost four-fold greater than tropine (775-900 µg/L). In the relevant literature, hygrine has been observed to accumulate via spontaneous decarboxylation of MPOB (see Bedewitz, M. A., Jones, A. D., D'Auria, J. C. & Barry, C. S. Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. Nat. Commun. 9, 5281 (2018)) (Fig. 22). As another example, LC-MS/MS analysis of the growth media of the engineered strains of Example 3.1.1 indicated that hygrine also accumulated in the negative control strain lacking AbPYKS and AbCYP82M3 due to decarboxylative condensation with NMPy (Fig. 22).

3.3.2) Modulation of growth temperature may be used to reduce the accumulation of side products in the biosynthetic pathway of the engineered strain to increase flux towards desired TAs and precursors. In one example, the impact of temperature on spontaneous hygrine production was evaluated by leveraging a kinetic principle that the rates of enzymatic and spontaneous reactions are decreased at lower temperatures. Since *A. belladonna* and other TA-producing *Solanaceae* are adapted for optimal growth at cooler climates, growth of yeast strains expressing *Solanaceae* genes at 25 ºC may improve enzyme folding and/or activity, enabling comparable production of enzymatically-generated tropine to growth at 30 ºC while reducing the rate of spontaneous hygrine production. Cultures of the tropine-producing strain of Example 3.2.2 were grown in non-selective defined media at 30 ºC and 25 ºC and the accumulation of tropine and hygrine was compared via LC-MS/MS analysis of the growth medium after 48 hours. Tropine titers were minimally impacted by the decrease in temperature. Hygrine accumulation was decreased by 42% at 25 ºC compared to at 30 ºC, resulting in a 60% increase in the ratio of tropine to hygrine produced (Fig. 25).

3.3.3) Reduction or elimination of undesirable side reactions can be used to improve metabolite flux towards desirable TAs and TA precursors in the biosynthetic pathway of the engineered strain. In one example, flux towards the TA precursor tropine may be improved by reducing hygrine production resulting from spontaneous decarboxylative condensation

with acetate. The impact of removing fed acetate from the media of the NMPy-producing strain of Example 2.4.4 on hygrine and tropine production was evaluated. The effect of abolishing acetate auxotrophy in the engineered strain of Example 2.4.4 was evaluated by expressing functional copies of *ALD4* and *ALD6* on low-copy plasmids and then monitoring the accumulation of hygrine and 4MAB acid via LC-MS/MS analysis after 48 hours of growth. While reconstitution of *ALD4* or *ALD6* enabled growth on selective media in the absence of fed acetate (Fig. 26a), addition of *ALD4* caused a five-fold increase in the accumulation of 4MAB acid while *ALD6* did not produce a significant increase (Fig. 26b). Moreover, the elimination of acetate feeding with either *ALD4* or *ALD6* resulted in 38% and 59% decreases in hygrine accumulation, respectively (Fig. 26b).

3.3.4) A functional copy of the *ALD6* gene was re-integrated into the tropine-producing strain of Example 3.2.2 at the previously disrupted *ald6* locus. The impact of this integration on the accumulation of all metabolites between NMPy and tropine was measured via LC-MS/MS analysis after 48 hours of growth in non-selective media. Restoration of acetate metabolism via Ald6p resulted in a 2.7-fold increase in tropine titers, as well as a 1.6-fold increase in hygrine accumulation (Fig. 28). Moreover, *ALD6* integration resulted in substantial increases in the production of NMPy and tropinone as well as increased consumption of MPOB (Fig. 27).

3.3.5) An additional copy of each biosynthetic enzyme gene between putrescine and tropine (i.e., *AbPMT1*, *DmMPO1*$^{\Delta C\text{-}PTS1}$, *AbPYKS*, and *AbCYP82M3*) was expressed from a low-copy plasmid in the engineered strain of Example 3.3.4 and production of TA intermediates was compared to that of the same strain expressing BFP by LC-MS/MS after 48 hours of growth in selective media. Expression of an additional copy of *AbPYKS* resulted in a 4.3-fold increase in NMP accumulation and a 1.3-fold increase in tropine production (Fig. 29). Expression of an additional copy of *AbPMT1* resulted in significant improvements in the production of all TA precursors between NMP and tropinone, as well as a 2.4-fold increase in tropine production (Fig. 29). Accordingly, additional copies of PMT (*AbPMT1* and *DsPMT1*) and PYKS (*AbPYKS*) were integrated into the genome of the tropine-producing strain of Example 3.3.4 (CSY1249) at the *PAD1* locus. The resulting engineered strain (CSY1251) was grown at 25°C in non-selective media for 48 hours, resulting in tropine production at titers of 3.4 mg/L, 2.2-fold greater than the tropine-producing strain (CSY1249) in Example 3.3.4 (Fig. 30).

Example 4: Yeast engineered for the production of pseudotropine alkaloids from L-arginine.

Yeast strains can be engineered for the production of non-medicinal TAs from early amino acid precursors such as L-arginine. As an example, the platform yeast strains described in Example 3 can be further engineered to produce pseudotropine alkaloids from L-arginine (Fig. 1).

The platform yeast strain producing tropinone from L-arginine (see descriptions in Example 3) can be further engineered to incorporate a stereospecific reductase, for example tropinone reductase 2 (TR2; EC 1.1.1.236), to convert the biosynthesized tropinone to pseudotropine. An expression cassette harboring a strong constitutive promoter such as *TDH3* and a coding sequence for a TR2 variant, for example TR2 from *Datura stramonium* (*DsTR2*), can be integrated into the genome of the tropinone-producing platform yeast strain. The resulting strain can be further engineered to produce hydroxylated derivatives of pseudotropine, for example calystegines, by integrating one or more expression cassettes harboring a strong constitutive promoter such as *PGK1* and a hydroxylating enzyme such as a cytochrome P450 that acts on the pseudotropine scaffold. By incorporating multiple P450 enzymes, each acting on a different position of the pseudotropine skeleton, a variety of calystegines and derivatives thereof can be biosynthesized. The engineered strains can then be cultured in nonselective synthetic complete media at 30 ºC or 25 ºC for 48 to 96 hours, after which the accumulation of pseudotropine alkaloids in the culture media can be analyzed by LC-MS/MS.

Example 5: Yeast engineered for overproduction of phenylpyruvate and associated TA precursors.

Yeast strains can be engineered for the overproduction of phenylpyruvate, which represents the precursor of acyl donor molecules required for production of medicinal TAs (Fig. 2), for the purpose of increasing carbon and nitrogen flux from central metabolism towards desired TAs and TA precursors. Yeast strains can be engineered for overproduction of phenylpyruvate by incorporating genetic modifications, including but not limited to the tuning of transcriptional regulation of native biosynthetic enzymes, deletion or disruption of genes encoding enzymes that divert precursor molecules away from the intended pathway, and introduction of heterologous enzymes for the conversion of endogenous molecules into TA precursor molecules.

In one example, a yeast strain can be engineered for increased phenylpyruvate production by incorporating additional copies of native genes which encode biosynthetic enzymes that produce phenylpyruvate from amino acids or other central metabolites. These additional copies can be controlled by strong constitutive promoters, such as GPD, TEF1, or PGK1. Examples of native gene targets include, but are not limited to, the aromatic acid aminotransferases *ARO8* and *ARO9*, and the dehydratase *PHA2*. In one instance, one or

more additional copies of *ARO8* can be incorporated into the engineered strain under the control of a strong constitutive promoter. In one instance, one or more additional copies of *ARO9* can be incorporated into the engineered strain under the control of a strong constitutive promoter. In another instance, one or more additional copies of *PHA2* can be incorporated into the engineered strain under the control of a strong constitutive promoter. In one embodiment of the invention, one or more additional copies of one or more genes selected from the group including *ARO8*, *ARO9*, and *PHA2* can be incorporated into the engineered strain under the control of unique, strong constitutive promoters.

Example 6: Yeast engineered for the production of acyl donors from L-phenylalanine or L-tyrosine for biosynthesis of TA scaffolds.

Yeast strains can be engineered for the production of diverse phenylpropanoid acyl donor compounds from L-phenylalanine and L-tyrosine, including PLA, cinnamic acid, coumaric acid, ferulic acid, benzoic acid, and coenzyme A thioester and glycoside derivatives of these compounds, which can undergo esterification with tropine, pseudotropine, or derivatives thereof to biosynthesize medicinal TAs, non-medicinal TAs, and non-natural TAs (Fig. 1-3).

6.1) As wild-type yeast produce only trace levels of PLA, production of this TA precursor must be increased to permit sufficient accumulation of downstream TAs. To improve PLA production, heterologous phenylpyruvate reductases (PPRs) may be expressed in the engineered host cells. PPR orthologs from *E. coli*, *Lactobacillus*, *A. belladonna*, and *Wickerhamia fluorescens*, as well as lactate dehydrogenases (LDHs) from *Bacillus* and *Lactobacillus* with reported activity on 3-phenylpyruvate (Table 1) were screened for activity in yeast by expressing each enzyme from a low-copy plasmid in CSY1251 and measuring PLA production by LC-MS/MS after 72 h of growth in selective media. All LDH candidates as well as the PPRs from *L. plantarum*, *E. coli*, and *A. belladonna* yielded modest (1.3- to 3.5-fold) improvements in PLA production relative to control, whereas expression of the PPR from *W. fluorescens* resulted in a nearly 80-fold increase in PLA production to ~250 mg/L (Fig. 31). As such, *WfPPR* was selected for integration into CSY1251 to make strain CSY1287.

6.2) As another example, yeast strains can be engineered for the production of cinnamic acid and coumaric acid, which are phenylpropanoids that can be used as acyl donor compounds for esterification with tropine or pseudotropine to form non-natural TAs, from L-phenylalanine and L-tyrosine, respectively. Yeast can be engineered for production of

cinnamic acid from L-phenylalanine by incorporating an ammonia-lyase such as a phenylalanine ammonia-lyase (PAL; EC 4.3.1.24). Similarly, yeast can be engineered for production of coumaric acid from L-tyrosine by incorporating an ammonia-lyase such as a tyrosine ammonia-lyase (TAL; EC 4.3.1.23). A yeast strain was engineered to produce cinnamic acid from L-phenylalanine by transforming it with a low-copy CEN/ARS plasmid with a *TRP1* selective marker, *TEF1* promoter, and a coding sequence for a PAL variant from *Arabidopsis thaliana* (*AtPAL1*). The resulting strain harboring the low-copy plasmid was grown in synthetic complete media with the appropriate amino acid dropout solution (-Ura) at 30 °C. After 48 hours of growth, the media was analyzed for cinnamic acid content by LC-MS/MS analysis (Fig. 32).

6.3) In *A. belladonna*, PLA is activated for acyl transfer to tropine via glucosylation by UDP-glucosyltransferase 84A27 (AbUGT) (see Qiu, F. et al., Functional genomics analysis reveals two novel genes required for littorine biosynthesis. New Phytol., nph.16317 (2019)). As plant UGTs participate in the biosynthesis of diverse phenylpropanoids and often exhibit broad substrate scope (see Ross, J., Li, Y., Lim, E.-K., D. J. Bowles, Higher plant glycosyltransferases. Genome Biol. 2, 3004.1–3004.6 (2001)), it is necessary to select a UGT with sufficiently high activity on a desired acyl donor.

6.3.1) As an example, the activity of AbUGT on different phenylpropanoid acyl donors, including the canonical substrate, PLA, was evaluated by expressing AbUGT from a low-copy plasmid in CSY1251 and measuring conversion of three phenylpropanoid acyl donors (PLA, cinnamic acid, ferulic acid) to their respective glucosides. While AbUGT glucosylated ~60% and 90% of cinnamic acid and ferulic acid, respectively, glucosylation of PLA was the lowest of the tested substrates at <3% conversion (Fig. 33).

6.3.2) Orthologs of AbUGT from other TA-producing *Solanaceae* may be evaluated for activity on PLA and other phenylpropanoids. In this example, transcripts encoding UGT84A27 from the transcriptomes of *Brugmansia sanguinea* (BsUGT) and *D. metel* (DmUGT) in the 1000Plants Database using a tBLASTn search. Yeast codon-optimized sequences encoding these orthologous UGTs were screened for activity by expressing AbUGT, BsUGT, DmUGT, or a BFP negative control from low-copy plasmids in CSY1251. Glucoside production was measured in cultures of the transformed strains via LC-MS/MS after 72 h of growth in selective media supplemented with 500 µM PLA, cinnamic acid (CA), or ferulic acid (FA) as glucose acceptors. All three UGT orthologs exhibited substantial glucosylation of CA (34-65% conversion) and FA (85-90% conversion) and only trace activity on PLA (<3% conversion), with AbUGT showing the greatest conversion of PLA

(2.7%) (Fig. 33, 34).

6.3.3) Given the disproportionate variation in activity of AbUGT on the structurally similar substrates cinnamate, ferulate, and PLA, a structure-guided rational mutagenesis approach may be implemented to engineer the active site of AbUGT for improved activity on PLA. In this example, a homology model of AbUGT bound to UDP-glucose was first constructed based on the crystal structure of *Arabidopsis thaliana* salicylate UDP-glucosyltransferase UGT74F2 (PDB: 5V2K) using the RaptorX web server (Fig. 35). Next, the docking of D-PLA in the active site was simulated using the Maestro/GlideXP software suite. Based on the energy-minimizing binding mode, the aryl ring of D-PLA is likely stabilized by pi-stacking interactions with F130, while its α-hydroxyl and carboxylate groups are respectively stabilized by hydrogen bonds with Q151 and H24, such that the nucleophilic carboxylate oxygen is within 4 Å of the electrophilic C1 carbon of UDP-glucose (Fig. 35). D-PLA is additionally adjacent to the residues L205 and I292, neither of which appear to interact with either substrate. This suggests that mutation of (i) F130 to tyrosine might preserve pi-stacking with the aryl ring of D-PLA while providing an additional hydrogen bond to stabilize the α-hydroxyl oxygen of D-PLA, which is absent from cinnamate and ferulate; (ii) L205 to phenylalanine might increase pi-stacking stabilization of D-PLA with F130Y; and (iii) I292 to glutamine would generate two additional stabilizing hydrogen bonds with D-PLA and UDP-glucose (Fig. 35). The F130Y, L205F, and I292Q point mutants of AbUGT were screened for activity by expressing each mutant, wild-type AbUGT, or a BFP control from a low-copy plasmid in CSY1251 and measured glucoside production by LC-MS/MS after 72 h of growth in selective media supplemented with 500 μm PLA, CA, or FA. All three mutants exhibited comparably low (and statistically indistinguishable) activity on PLA relative to wild-type AbUGT (<3% conversion), although the F130Y and I292Q mutations significantly decreased UGT activity on CA (Fig. 36).

6.3.4) Based on the results described in sections 6.1 and 6.3, strain CSY1288 was constructed by integrating yeast codon-optimized *WfPPR* and *AbUGT* into the genome of CSY1251, validated by verification of PLA production (66 mg/L) and minimal PLA glucoside accumulation (Fig. 37).

6.4) As poor activity of AbUGT on PLA is likely to limit flux of TA precursors towards downstream TAs, flux of phenylalanine to PLA glucoside may be increased by incorporating genetic modifications which promote UDP-glucose accumulation and decrease glycoside degradation.

6.4.1) UDP-glucose is critical for the formation of storage polysaccharides, cell wall glucans, and glycoproteins, and thus its biosynthesis is tightly regulated (see Nishizawa, M., Tanabe, M., Yabuki, N., Kitada, K., Toh-e, A. Pho85 kinase, a yeast cyclin-dependent kinase, regulates the expression of UGP1 encoding UDP- glucose pyrophosphorylase. Yeast. 18, 239–249 (2001)). During growth on glucose, yeast direct glucose-6-phosphate along two major metabolic routes, glycolysis and starch biosynthesis. As citrate is an allosteric inhibitor of the glycolytic rate-limiting enzyme phosphofructokinase (see Li, Y. et al., Production of Rebaudioside A from Stevioside Catalyzed by the Engineered Saccharomyces cerevisiae. Appl. Biochem. Biotechnol. 178, 1586–1598 (2016)), partial suppression of glycolysis via citrate supplementation might increase UDP-glucose availability and glucoside production (Fig. 38). Strain CSY1288, which encodes genomic WfPPR and AbUGT for endogenous PLA glucoside production, was cultured in media supplemented with 2% citrate and 500 μM CA or FA, and glucoside production was compared by LC-MS/MS after 72 h of growth. Citrate supplementation decreased glucosylation of PLA, CA, and FA by 83%, 56%, and 78%, respectively (Fig. 39).

6.4.2) Overexpression of *PGM2* and *UGP1*, whose gene products respectively catalyze the isomerization of glucose-6-phosphate to glucose-1-phosphate and conversion of glucose-1-phosphate to UDP-glucose, can be used to increase UDP-glucose supply.

6.4.2.1) Extra copies of *PGM2* and *UGP1* were expressed from low-copy plasmids in CSY1288 and PLA glucoside production was measured following 72 h of growth in selective media. While *PGM2* overexpression yielded no improvement relative to control, overexpression of *UGP1* resulted in a ~1.8-fold increase in PLA glucoside production (Fig. 40), supporting that increasing the UDP-glucose pool improves PLA utilization by AbUGT.

6.4.2.2) Native glucosidases may act on PLA and other TA precursor glucosides to reduce accumulation, as other heterologous glucosides have been shown to be hydrolyzed in this manner in yeast (see Schmidt, S., Rainieri, S., Witte, S., Matern, U., Martens, S., Identification of a Saccharomyces cerevisiae glucosidase that hydrolyzes flavonoid glucosides. Appl. Environ. Microbiol. 77, 1751–1757 (2011); see also Wang, H. et al., Engineering Saccharomyces cerevisiae with the deletion of endogenous glucosidases for the production of flavonoid glucosides. Microb. Cell Fact. 15, 1–12 (2016)). In this example, three native glucosidase genes—*EXG1*, *SPR1*, and *EGH1*—were disrupted in CSY1288 and PLA glucoside production was measured following 72 h of growth of disruption mutants in non-selective media. The disruption of *EGH1* more than doubled PLA glucoside production (Fig. 41), indicating that hydrolysis by Egh1p constitutes a substantial loss of TA

precursor flux.

Example 7: Yeast engineered for conversion of littorine to hyoscyamine aldehyde

Yeast strains can be engineered for the conversion of littorine to hyoscyamine aldehyde (Fig. 2). For example, the tropine- and PLA glucoside-producing yeast strain described in Example 6 can be further engineered to express a cytochrome P450 CYP80F1 (EC 1.14.19.-) to catalyze the rearrangement of littorine to hyoscyamine aldehyde, and a cytochrome P450 reductase (CPR; EC 1.6.2.4) to support the activity of the P450 enzyme. A yeast strain was engineered to convert fed littorine to hyoscyamine aldehyde by transforming it with a low-copy CEN/ARS plasmid with a *LEU2* selection marker, *TDH3* promoter, and coding sequence for a CYP80F1 variant from *A. belladonna* (*AbCYP80F1*); and with a low-copy CEN/ARS plasmid with a *TRP1* selection marker, *TEF1* promoter, and a coding sequence for a cytochrome P450 reductase (CPR) from *S. cerevisiae* (*NCP1*) or from *A. thaliana* (*AtATR1*). The resulting strain harboring the low-copy plasmids was grown in synthetic complete media with the appropriate amino acid dropout solution (-Leu -Trp) supplemented with 1 mM littorine at 30 ºC. After 48 hours of growth, the media was analyzed for hyoscyamine aldehyde content by LC-MS/MS analysis (Fig. 42).

Example 8: Yeast engineered for conversion of hyoscyamine to scopolamine.

Yeast strains can be engineered for conversion of hyoscyamine to scopolamine (Fig. 2). For example, the yeast strain described in Example 7 can be further engineered to incorporate enzymes which possess hydroxylase activity at the 6β position of hyoscyamine to form anisodamine and enzymes which possess dioxygenase activity at the 6β-hydroxyl position of anisodamine to form scopolamine, or enzymes which possess both of these activities (EC 1.14.11.11). Yeast strains were engineered to convert fed hyoscyamine to scopolamine by transforming them with a low-copy CEN/ARS plasmid with a *LEU2* selection marker, *TDH3* promoter, and coding sequence for a hyoscyamine 6β-hydroxylase/dioxygenase (H6H) from *D. stramonium* (*DsH6H*), *Anisodus acutangulus* (*AaH6H*), *Brugmansia arborea* (*BaH6H*), or *Datura metel* (*DmH6H*). The resulting strains harboring the low-copy plasmids were grown in synthetic complete media with the appropriate amino acid dropout solution (-Leu) and supplemented with 1 mM hyoscyamine at 30 ºC. After 72 hours of growth, the media was analyzed for scopolamine content by LC-MS/MS analysis (Fig. 43). The strain expressing the H6H variant from *D. stramonium* exhibited the greatest conversion of fed hyoscyamine to scopolamine, although all tested variants showed H6H activity in vivo. Further optimization of cofactor requirements was performed by supplementing different cofactors in the culture media of this engineered yeast strain and analyzing the media by LC-MS/MS after 72 hours of growth. This analysis

identified that ferrous iron supplementation increases conversion of hyoscyamine to scopolamine (Fig. 44).

Example 9. Identification of hyoscyamine dehydrogenase enzyme candidates and reduction of hyoscyamine aldehyde to hyoscyamine in engineered non-plant cells

To identify a dehydrogenase enzyme suitable for performing the TA alcohol-aldehyde interconversions of the methods disclosed herein, and in particular to reduce hyoscyamine aldehyde to hyoscyamine, a hyoscyamine dehydrogenase (HDH) open reading frame was identified from publically available plant RNA sequencing data.

9.1) Tissue-specific abundances (fragments per kilobase of contig per million mapped reads, FPKM) and putative protein structural and functional annotations for each of 43,861 unique transcripts identified from the *A. belladonna* transcriptome were obtained from the Michigan State University Medicinal Plant Genomics Resource (see http://medicinalplantgenomics.msu.edu/). Transcripts encoding hyoscyamine dehydrogenase candidates were identified based on clustering of tissue-specific expression profiles with those of the bait genes CYP80F1 (littorine mutase) and H6H (hyoscyamine 6β-hydroxylase/dioxygenase), which respectively precede and follow the dehydrogenase step in the TA biosynthetic pathway, using the following computational filtering algorithm.

First, the complete list of 43,861 transcripts was filtered for those annotated with any of the following protein family (PFAM) IDs: PF00106, PF13561, PF08659, PF08240, PF00107, PF00248, PF00465, PF13685, PF13823, PF13602, PF16884, PF00248; or any of the following functional annotation keywords: alcohol dehydrogenase, aldehyde reductase, short chain, aldo/keto. Additionally, any transcripts with functional annotations containing the keywords putrescine, tropinone, and tropine were included in the filter as positive control TA-associated genes to validate clustering with bait genes. Next, mean tissue-specific expression profiles were generated for the CYP80F1 and H6H bait genes. For each of the two bait genes, linear regression models were constructed to express the bait gene expression profile (in FPKM) as a linear function of each candidate gene profile and correlation p-values were computed for each candidate. The candidates identified using each of the two bait genes were pooled and duplicates were removed. Combined p-values for each candidate were computed as the sum of the log10 p-values of the correlations with each of the two bait genes. Transcripts matching known dehydrogenases in the TA biosynthetic pathway (i.e., tropinone reductases I and II) were removed, and the remaining candidates were ranked by combined p-value and by distance from bait genes via hierarchical clustering of tissue-specific expression profiles (Fig. 45).

9.2) Nearly all candidates identified in Example 9.1 exhibited the same secondary root-specific expression pattern observed for known TA biosynthetic genes. A BLASTp search of the resulting ~30 candidates against the UniPROT/SwissPROT database revealed that many transcripts were missing terminal or internal sequence regions. To address this, *de novo* transcriptome assembly was repeated from deposited raw RNAseq reads (see http://medicinalplantgenomics.msu.edu/) using the Trinity software package (see Haas, B. J. et al., De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–512 (2013)), and all missing sequence fragments for twelve of the HDH candidates were reconstituted by performing BLAST alignments of incomplete sequence regions against the newly assembled transcriptome (Table 2).

9.3) The missing HDH activity was identified by screening the candidates generated in Examples 9.1 and 9.2 in yeast.

9.3.1) Lack of an authentic commercial standard for hyoscyamine aldehyde and insufficient yield from chemical syntheses necessitated co-expression of HDH candidates with the upstream biosynthetic enzyme—the cytochrome P450 littorine mutase (CYP80F1)—for activity screening *in vivo* via fed littorine (see Example 7). As littorine exhibits similar chromatographic and mass spectrometric properties as the HDH product hyoscyamine, an HDH screening strain (CSY1292) was constructed by integrating yeast codon-optimized *AbCYP80F1* and *DsH6H* (see Example 8) into the genome of CSY1251, enabling screening of HDH candidates via detection of scopolamine (m/z^+ 304) produced from fed littorine (m/z^+ 290) via a three-step biosynthetic pathway (Fig. 2).

9.3.2) Yeast codon-optimized sequences encoding each of the twelve HDH candidates were expressed from a low-copy plasmid in strain CSY1292, and scopolamine production was measured following 72 h of growth in media supplemented with 1 mM littorine. One of the twelve candidates, HDH2 (referred to as AbHDH), exhibited a 35% decrease in hyoscyamine aldehyde levels and measurable accumulation of scopolamine (7.2 µg/L), indicating that it encoded the missing HDH activity (Fig. 46).

9.4) Structural and phylogenetic analyses provided further insight into the catalytic mechanism and evolutionary history of HDH.

9.4.1) A homology model of AbHDH was constructed based on the crystal structure of *Populus tremuloides* sinapyl alcohol dehydrogenase (PtSAD; PDB: 1YQD) (Fig. 47).

AbHDH is a member of the zinc-dependent alcohol dehydrogenase (ZADH) family within the medium-chain dehydrogenase/reductase (MDR) superfamily. Typical of this family, AbHDH exhibits a bi-lobular structure with a well-conserved nucleotide-binding domain and a more variable substrate-binding domain. Alignment of residues S216, T217, S218, and K221 within the AbHDH nucleotide-binding domain with the phosphate-stabilizing residues S214, T215, S216, and K219 in PtSAD suggests that AbHDH is an NADPH-dependent oxidoreductase. Also typical of ZADHs, AbHDH appears to bind a structural Zn^{2+} using a tetrad of cysteine residues near the protein surface (C105, C108, C111, and C119) and a catalytic Zn^{2+} within the active site.

9.4.2) The catalytic mechanism of AbHDH was elucidated via molecular docking of the substrate, hyoscyamine aldehyde, into the active site using the Maestro/Glide software package (Fig. 47). The most favorable binding mode positions the aldehyde group of the substrate within ~5 Å of both the catalytic Zn^{2+} and NADPH hydride donor. The docking result and the general mechanism of ZADHs (see Bomati, E.K., Noel, J.P., Structural and kinetic basis for substrate selectivity in Populus tremuloides sinapyl alcohol dehydrogenase. Plant Cell. 17, 1598–1611 (2005)) suggests the following catalytic mechanism for AbHDH. In the absence of substrate, the catalytic Zn^{2+} within the active site is stabilized by C52, H74, C168, and a water molecule, which is positioned via polar interaction with S54 and displaced upon binding of hyoscyamine aldehyde. Nucleophilic attack of the aldehyde carbonyl by a dihydronicotinamide hydride forms an oxyanion intermediate stabilized by interaction with the catalytic Zn^{2+}, and which is likely protonated via a proton shuttle between the ribose group of $NADP^+$ and S54.

9.5) To confirm whether orthologous oxidoreductases catalyze hyoscyamine biosynthesis in other TA-producing *Solanaceae*, variants of the *AbHDH* coding sequence were identified from transcriptomes of *Datura innoxia* and *Datura stramonium* (see https://medplantrnaseq.org) using a tBLASTx search (Fig. 48). HDH activity of the two identified orthologs (*DiHDH*, *DsHDH*) was validated by co-expressing yeast codon-optimized sequences with an additional copy of the flux-limiting *DsH6H* from low-copy plasmids in CSY1292 and measuring scopolamine production in media supplemented with 1 mM littorine. DsHDH showed the highest substrate depletion and product accumulation of the variants tested (Fig. 49).

9.6) The medicinal TA biosynthetic branch comprising optimal enzyme variants and overexpression of a flux-limiting enzyme was integrated into a platform yeast strain. Strain CSY1294 was constructed by integrating yeast codon-optimized *WfPPR* and *AbUGT*,

DsHDH, and a second copy of *DsH6H* into CSY1292. Scopolamine production from fed littorine was verified in CSY1294 (Fig. 50).

Example 10: Yeast engineered for the esterification of acyl donors and acceptors for production of TA scaffolds.

Yeast strains can be engineered to express enzymes which catalyze the esterification of activated acyl donor compounds and acyl acceptor compounds to produce diverse TA scaffolds (Fig. 2, 3). Activation of the acyl donor group can be achieved by engineering an acyl donor-producing yeast strain to incorporate an enzyme which appends a chemical moiety with high group-transfer potential, such as coenzyme A (CoA) or glucose (glucoside), to the carboxyl group of the acyl donor, as described in Example 6. Examples of acyl donor-activating enzymes which can be utilized in this capacity include CoA ligases and UDP-glycosyltransferases. Examples of esterifying enzymes which can be used to catalyze the esterification of activated acyl donor compounds and acyl acceptors such as tropine and pseudotropine are acyltransferases, including serine carboxypeptidase-like acyltransferases (SCPL-ATs) and BAHD-type acyltransferases. The coding sequence of such acyltransferases may be modified to improve their activity when expressed in a heterologous host such as yeast.

10.1) In plants, where SCPL-ATs are typically found to occur naturally, the coding sequence of SCPL-ATs include N-terminal signal peptides which direct the nascent polypeptide to the endoplasmic reticulum (ER). Once localized to the ER, the SCPL-AT polypeptide is transported by way of the secretory trafficking pathway through the Golgi to the vacuole lumen, where they are found to exhibit activity. During this ER-to-vacuole trafficking process, they undergo several post-translational modifications (Fig. 51), including but not limited to signal peptide cleavage, N-glycosylation, removal of internal propeptide sequences, and disulfide bond formation (see Stehle, F., Stubbs, M.T., Strack, D., & Milkowski, C. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism, FEBS Journal, 275, (2008)). However, as intracellular trafficking pathways and patterns of post-translational modifications differ between organisms, expression of SCPL-ATs in heterologous hosts may result in incorrect sub-cellular localization and/or incorrect post-translational modification for activity. As an example, the coding sequence of a SCPL-AT such as littorine synthase (LS) (Table 1) may be modified to improve activity when expressed in yeast.

10.1.1) Signal peptide sequences can impact the processing and localization of SCPL-ATs in yeast.

10.1.1.1) The presence of a putative N-terminal signal peptide in AbLS suggests that it follows the expected SCPL ER-to-vacuole trafficking pathway *in planta*. AbLS localization in yeast was examined by expressing N- and C-terminal GFP fusions of AbLS from low-copy plasmids in CSY1294. Fluorescence microscopy revealed that the N-terminal fusion (GFP-AbLS) co-localized with the vacuolar membrane stain FM4-64 (Fig. 52). No fluorescence was detected for the C-terminal fusion (AbLS-GFP), consistent with reports that a native C-terminus is crucial for stability of SCPL acyltransferases (see Stehle, F., Stubbs, M.T., Strack, D., & Milkowski, C. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism, FEBS Journal, 275, (2008)).

10.1.1.2) Vacuolar sequestration of SCPL-ATs in yeast might preclude access to cytosolic substrate pools, as yeast likely lack the requisite tonoplastic transporters present in plants for exchange of secondary metabolites with the cytosol. To determine whether forced localization of AbLS to other yeast compartments—presumably, with improved access to cytosolic metabolites—would enable activity, the wild-type N-terminal SP sequence was replaced with a panel of N-terminal signal sequences taken from yeast proteins targeted to the vacuole lumen (Prc1p and Pep4p), vacuole membrane facing the lumen (Dap2p), *trans*-Golgi network (Och1p), ER membrane facing the lumen (Mns1p), and mitochondrial matrix (Cit1p) (Fig. 53). The wild-type SP was also removed entirely, and for another variant a canonical peroxisome-targeting sequence (PTS1) was appended to the C-terminus. These chimeric AbLS variants were expressed from high-copy plasmids in CSY1294 and transformants were screened for activity by LC-MS/MS after 96 h of growth in selective media. No production of littorine or downstream intermediates was observed with any of the variants (Fig. 53).

10.1.2) Incorrect post-translational processing of SCPL-ATs in yeast might prevent expression of active enzyme.

10.1.2.1) Protein N-glycosylation patterns differ between yeast and plants, and previous reports have suggested that correct N-glycosylation of diverse plant enzymes is important for their folding, stability, and/or activity (see Kar, B., Verma, P., den Haan, R., Sharma, A.K., Effect of N-linked glycosylation on the activity and stability of a β-glucosidase from Putranjiva roxburghii. Int. J. Biol. Macromol. 112, 490–498 (2018); see also Podzimek, T. et al., N-glycosylation of tomato nuclease TBN1 produced in N. benthamiana and its effect on the enzyme activity. Plant Sci. 276, 152–161 (2018); see also Strasser, R., Plant protein glycosylation. Glycobiology. 26, 926–939 (2016)). *In silico* analysis of the AbLS polypeptide

predicted four N-glycosylation sites (N152, N320, N376, N416) and no O-glycosylation of this protein was detected in *N. benthamiana* (Fig. 54). C-terminally HA-tagged wild-type AbLS, each of the four N→Q mutants (where mutation of N to Q abolishes N-glycosylation (*23*)), or a quadruple N→Q mutant were expressed in CSY1294 and in *N. benthamiana*, and glycosylation profiles were compared by Western blot. Whereas wild-type AbLS, N→Q single mutants, and the quadruple mutant all appeared as single bands in *N. benthamiana*, indicating a single glycosylation state, only the quadruple N→Q mutant produced a single band in yeast; all other variants appeared as either double or triple bands, indicating a combination of multiple glycosylation states (Fig. 55). However, as the denser of the two wild-type AbLS bands in yeast showed partial overlap with that of wild-type AbLS in tobacco, at least some fraction of yeast-expressed AbLS must be in a correct glycosylation state, and mis-glycosylation is unlikely to account for the complete lack of AbLS activity in yeast. (Fig. 54-55).

10.1.2.2) A subset of SCPL acyltransferases, including sinapoylglucose:choline sinapoyltransferase from *Arabidopsis thaliana* (AtSCT) and an avenacin synthase from *Avena strigosa* (AsSCPL1), have been shown to contain an internal propeptide linker which is proteolytically removed to produce an active heterodimer joined by disulfide bonds (see Shirley, A.M., Chapple, C., Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J. Biol. Chem. 278, 19870–19877 (2003); see also Mugford, S.T. et al., A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. Plant Cell. 21, 2473–2484 (2009)). Comparison of the AbLS amino acid sequence with those of previously characterized plant serine carboxypeptidases and SCPL acyltransferases revealed the presence of an internal 25- to 30-residue sequence which aligns with the highly variable propeptide of AtSCT, AsSCPL1, and wheat carboxypeptidase 2 (TaCBP2), suggesting that AbLS too undergoes endoproteolytic cleavage to form a heterodimer (Fig. 56). Additionally, a homology model of AbLS suggested that the predicted internal propeptide blocks the active site, thereby necessitating removal for activity (Fig. 57). However, wild-type AbLS expressed in *N. benthamiana* does not appear to undergo proteolytic cleavage, as no expected ~20-25 kDa C-terminal fragment was detected by Western blot under disulfide-reducing conditions (Fig. 54, 55). As the putative propeptide does not appear to be cleaved or removed *in planta*, AbLS might adopt a native conformation in plants which shifts the propeptide away from the active site, but differences in the biochemical environment of the yeast secretory pathway and/or vacuole prevents this shift, blocking activity.

To address this failure mode, split AbLS controls were constructed in which the N- and C-terminal domains flanking the putative propeptide linker were expressed independently, with or without separate signal peptides. Additionally, AbLS variants in which the putative propeptide was replaced with either a flexible $(GGGGS)_n$ linker, the internal propeptide from AtSCT previously demonstrated to be cleaved in yeast (see Shirley, A.M., Chapple, C., Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J. Biol. Chem. 278, 19870–19877 (2003)), or a synthetic linker containing a poly-arginine site cleaved by the *trans*-Golgi protease Kex2p (see Chen, X., Zaro, J.L., Shen, W.C., Fusion protein linkers: Property, design and functionality. Adv. Drug Deliv. Rev. 65, 1357–1369 (2013); see also Redding, K., Seeger, M., Payne, G.S., Fuller, R.S., The effects of clathrin inactivation on localization of Kex2 protease are independent of the TGN localization signal in the cytosolic tail of Kex2p. Mol. Biol. Cell. 7, 1667–1677 (1996)) were constructed (Fig. 58). Each of the split AbLS controls and propeptide/linker variants were expressed from low-copy plasmids in CSY1294 and transformants were screened for LS activity by LC-MS/MS after 96 h of growth in selective media. No production of littorine or downstream TAs was observed with any of these variants.

To troubleshoot protein expression, each of the above C-terminal HA-tagged AbLS variants was expressed from low-copy plasmids in CSY1294 and apparent protein sizes were compared to split-AbLS controls by Western blot (Fig. 58). Neither the AtSCT nor the poly-arginine linkers produced the 20-25 kDa C-terminal fragment expected from proteolytic cleavage. In the latter case, failure of the poly-arginine AbLS variant to be cleaved suggests that the protein becomes stalled in the secretion pathway upstream of the *trans*-Golgi network (TGN; also referred to as the late Golgi in yeast), which may account for the severe growth defect observed in CSY1294 expressing wild-type or Golgi-targeted (Och1p SP-fused) AbLS.

10.1.3) Functional expression of SCPL-ATs in yeast can be achieved by engineering N-terminal fusions that alter sorting from the TGN. Transport of soluble yeast proteins from the TGN to the vacuole requires recognition of a typically N-terminal signal sequence by vacuole protein sorting (Vps) cargo transport proteins, whereas integral membrane proteins which reach the yeast TGN appear to be sorted to the vacuole by default (see Stack, J.H., Receptor-Mediated Protein Sorting to the Vacuole in Yeast: Roles for Protein Kinase, Lipid Kinase and GTP-Binding Proteins. Annu. Rev. Cell Dev. Biol. 11, 1–33 (1995); see also Roberts, C.J., Nothwehr, S.F., Stevens, T.H., Membrane protein sorting in the yeast secretory pathway: Evidence that the vacuole may be the default compartment. J. Cell Biol.

119, 69–83 (1992)). Conversion of SCPL-ATs into transmembrane proteins by masking the SP with an N-terminally fused soluble domain can therefore resolve the obstruction in TGN sorting.

10.1.3.1) In one example, AbLS variants were constructed with a panel of N-terminally fused soluble domains, including fluorescent proteins from the *Aequoria* (GFP, BFP, mVenus) and *Discosoma* (mCherry, DsRed) families; small ubiquitin-related modifier (Smt3p) with a mutated protease cleavage site (SUMO*); and the upstream enzyme in the TA pathway, AbUGT. These variants and wild-type AbLS were expressed from low-copy plasmids in CSY1294 and screened for littorine synthase activity following 96 h of growth in selective media. All N-terminally fused AbLS variants exhibited measurable accumulation of hyoscyamine and scopolamine. Fusion of *Aequoria* GFP-derived fluorescent proteins to AbLS resulted in hyoscyamine and scopolamine production of ~1 µg/L and ~0.1 µg/L, respectively; whereas fusion of *Discosoma*-derived fluorescent proteins led to considerably higher TA production, with the greatest titers achieved via DsRed fusion (10.3 µg/L hyoscyamine, 0.87 µg/L scopolamine) (Fig. 59). Enhancement of AbLS activity appeared to be correlated with the oligomerization state of the N-terminal domain, with scopolamine production increasing in order from monomeric (GFP, BFP, mVenus, mCherry, SUMO*) to homodimeric (AbUGT) to homotetrameric (DsRed) domains.

10.2) To generate a strain capable of complete TA biosynthesis, a yeast codon-optimized *DsRed-AbLS* and a second copy of *UGP1* were integrated into the genome of CSY1294 at the disrupted *EGH1* site to generate CSY1296. CSY1296 exhibited *de novo* hyoscyamine and scopolamine production at titers of 10.2 µg/L and 1.0 µg/L, respectively.

Example 11. Alleviation of intracellular substrate transport limitations using heterologous transporters

As the enzymes which carry out TA biosynthesis are distributed across multiple sub-cellular compartments (cytosol, ER membrane, peroxisome, vacuole, mitochondria), and yeast are unlikely to possess the transporters found in plants which enable mobilization of TA biosynthetic intermediates between different compartments, intracellular metabolite transport is likely to restrict TA production.

11.1) Inter-compartment transport limitations may be addressed by functional expression of plant transporters in non-plant host cells. Vacuolar compartmentalization of DsRed-AbLS (Fig. 60) necessitates the import of cytosolic tropine and PLA glucoside to the vacuole lumen and export of vacuolar littorine to the cytosol. Several multidrug and toxin extrusion

(MATE) transporters responsible for vacuolar alkaloid and glycoside sequestration have been identified in *Solanaceae*, including three with observed or predicted activity on TAs (see Morita, M. et al., Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. Proc. Natl. Acad. Sci. U. S. A. 106, 2447–2452 (2009); see also Shoji, T. et al., Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. Plant Physiol. 149, 708–718 (2009)). In one example, *N. tabacum* jasmonate-inducible alkaloid transporter 1 (NtJAT1) and two MATEs (NtMATE1, NtMATE2) were expressed from low-copy plasmids in CSY1296 and accumulation of TAs was measured following 96 h of growth in selective media. Expression of NtJAT1 and NtMATE2 improved TA production, with the former resulting in 74% and 18% increases in hyoscyamine and scopolamine titers, respectively (Fig. 61).

11.2) To evaluate the subcellular localization of these transporters, and determine likely mechanisms of action, fluorescence microscopy of CSY1296 expressing C-terminal GFP fusions of NtJAT1 or NtMATE2 from low-copy plasmids was performed. The analysis supports that NtJAT1 localizes almost exclusively to the vacuolar membrane (co-localizing with DsRed-AbLS), whereas NtMATE2 is partitioned between the vacuolar and plasma membranes (Fig. 60), suggesting that both transporters might function to dissipate vacuolar substrate transport limitations while the latter might also improve cellular TA export.

Example 12: Yeast engineered for the production of non-natural TAs from L-phenylalanine or L-tyrosine and L-arginine

In addition to being engineered for the production of medicinal and non-medicinal TAs which occur naturally in organisms, yeast can also be engineered for the production of non-natural TAs (Fig. 3). For example, yeast can be engineered to express biosynthetic pathways for the production of acyl donor compounds not naturally incorporated into TAs by plants.

12.1) In one example, the platform tropine-producing yeast strain described in Example 3 can be further engineered to produce the acyl donor compound cinnamic acid (as described in Example 6) and to express cinnamate-activating enzymes and esterifying enzymes to produce non-natural TAs such as cinnamoyltropine.

12.1.1) Cinnamate can be produced from phenylalanine via a phenylalanine ammonia-lyase, for example PAL1 from *A. thaliana* (AtPAL1). Since EcCS requires a coenzyme A (CoA)-activated acyl donor, a 4-coumarate-CoA ligase with established activity on

cinnamate, such as 4CL5 from *A. thaliana* (At4CL5) (see Eudes, A. et al. Exploiting members of the BAHD acyltransferase family to synthesize multiple hydroxycinnamate and benzoate conjugates in yeast. Microbial Cell Factories, 15, (2016)), can be expressed to enable cinnamoyl-CoA biosynthesis in yeast. The platform tropine-producing yeast strain described in Example 3 was transformed with a low-copy plasmid enabling production of cinnamic acid as described in Example 6.

12.1.2) The engineered strain of Example 12.1.1 was further modified to produce cinnamoyltropine by transforming it with a high-copy 2μ plasmid with a *URA3* selective marker, *HXT7* and *PMA1* promoters, and coding sequences for a 4-coumarate-CoA ligase variant from *A. thaliana* (*At4CL5*) and a cocaine synthase from *Erythroxylum coca* (*EcCS*). The resulting strain harboring the low- and high-copy plasmids was grown in synthetic complete media with the appropriate amino acid dropout solution (-Ura -Trp) at 25 ºC. After 72 hours of growth, the culture medium was analyzed for cinnamoyltropine by LC-MS/MS analysis (Fig. 62). Tandem MS/MS and fragmentation analysis were used to detect and verify the identity of cinnamoyltropine. Comparison of MS/MS spectra corresponding to the parent mass of cinnamoyltropine ($m/z^+ = 272$) revealed a novel peak at a retention time of 3.684 min, and which produced fragments whose masses appeared to match transitions for a genuine cinnamoyltropine standard (Fig. 62a). The most abundant mass transition, m/z^+ $272 \rightarrow 124$, is consistent with the primary $m/z^+ = 124$ tropine fragment produced during fragmentation of hyoscyamine (see Bedewitz, M.A., et al. A Root-Expressed L-Phenylalanine:4-Hydroxyphenylpyruvate Aminotransferase Is Required for Tropane Alkaloid Biosynthesis in Atropa belladonna. The Plant Cell, 26, (2014)).

12.1.3) Based on the $272 \rightarrow 124$ LC-MS/MS transition for cinnamoyltropine described in Example 12.1.2, a multiple reaction monitoring (MRM) LC-MS/MS method was developed to measure *de novo* cinnamoyltropine production. Cinnamoyltropine accumulated to substantial levels in the extracellular medium of the engineered strain of example 12.1.2, but not in the absence of *AtPAL1*, *At4CL5*, and *EcCS* (Fig. 62b). The titer of cinnamoyltropine produced *de novo* was estimated to be 6.0 μg/L based on a standard curve.

Example 13: Modification of growth media to improve production of TAs and TA precursors

Production titers of TA precursors and TAs can be improved by modifying the culture media composition. For example, the media types can vary in the media base (e.g., yeast peptone, yeast nitrogen base), carbon source (e.g., glucose, maltodextrin), and nitrogen

source (e.g., amino acids, ammonium sulfate, urea). Media types can also vary in the relative proportions of each component, such as the concentration of carbon source and the concentration of nitrogen source, or the concentration of each individual amino acid.

13.1) Tropine-producing yeast strains (as described in Example 3) were initially grown in defined media (i.e., YNB with ammonium sulfate and all amino acids) with varying carbon sources and tropine production assayed after 48 hours of growth at 25 °C. The highest production of tropine was observed with 2% galactose (Fig. 63a). However, some engineered strains either failed to grow or suffered severely reduced growth with certain carbon sources, such as glycerol, arabinose, and sorbitol, likely due to an inability to assimilate these carbon sources into central metabolism.

13.2) Tropine-producing yeast strains (as described in Example 3) were cultured in defined media with 2% dextrose for growth and supplemented with 2% of an additional carbon source, and tropine production was assayed after 48 hours of growth at 25 °C. The highest production of tropine was observed with 2% dextrose and 2% glycerol (Fig. 63b). Glycerol is a non-sugar carbon source which may contribute to higher production of TA precursors and TAs through several mechanisms, including stabilization of cellular lipid membranes, improved folding and stability of heterologous proteins, and regeneration of the NADPH cofactor required for the activity of cytochrome P450s and some short-chain dehydrogenase/reductase enzymes (see Li, Y. *et al*. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. Proc. Natl. Acad. Sci. U.S.A. 2018, 115(17) E3922-E3931).

13.3) Improvements in *de novo* medicinal TA biosynthesis in engineered yeast can be achieved via alleviation of flux bottlenecks and transport limitations.

13.3.1) Improvements in TA production were achieved via overexpression of bottleneck enzymes and media optimization. As production of tropine in CSY1296 (~mg/L) is unlikely to limit flux to scopolamine (~µg/L), metabolic bottlenecks limiting scopolamine production were identified by expressing an additional copy of each heterologous enzyme between phenylpyruvate and scopolamine (Fig. 2) from low-copy plasmids in CSY1296 and measuring production of TAs and intermediates. Additional copies of *WfPPR* and *DsH6H* resulted in 64% and 89% increases in hyoscyamine and scopolamine titers, respectively, indicating that these enzymes were primary limiters of pathway flux (Fig. 64).

13.3.2) An improved scopolamine-producing strain was constructed by integrating *NtJAT1*

and a second copy of *WfPPR* and *DsH6H* into CSY1296. The resulting strain CSY1297 showed 2.4- and 7.1-fold respective increases in hyoscyamine and scopolamine accumulation relative to CSY1296 (Fig. 65).

Example 14: Prediction of putative TA transporters using supervised learning models.

Although MATE transporters described in previous examples (NtJAT1, NtMATE2) may be used to facilitate import of tropine and tropine-like TA intermediates into the vacuole for esterification by littorine synthase, export of assembled tropane esters from the vacuole to the cytosol is a critical limitation to metabolic pathway flux. Thus, identification of new TA transporters with potential activity on assembled TA esters like littorine may enhance TA production in engineered non-plant eukaryotic cells such as yeast.

14.1) TA transporter candidates were identified from a publicly available *A. belladonna* transcriptome (see Michigan State University Medicinal Plant Genomics Resource) via conventional co-expression analysis strategies. This dataset comprised abundances for each of over 40,000 transcripts across 11 different tissues. After filtering the dataset to retain only transcripts annotated by BLAST and protein family (PFAM) searches for roles in small molecule transport, linear regression models were generated to compare the tissue-specific expression profile of each transcript with that of the bait genes littorine mutase (CYP80F1) and hyoscyamine 6β-hydroxylase/dioxygenase (H6H), and then candidates were further arranged by expression profile similarity via hierarchical clustering (Fig. 66). Nearly 100 transporter candidates were found to co-express ($P < 0.01$) and cluster with known TA biosynthetic genes.

14.2) As the conventional regression/clustering analysis yielded a candidate list too large for efficient experimental validation, supervised classification methods were evaluated for use in providing a greater reduction in search space. Several supervised learning models were trained and evaluated for predictive performance in identifying genes involved in TA biosynthesis from the same transcriptome dataset as in Example 14.1.

14.2.1) Using BLAST functional annotations, each of the ~27,000 transcripts which passed an initial filtering step to remove incomplete or zero-abundance transcripts was assigned one of three *a priori* output values, depending on their involvement in TA biosynthesis (Fig. 67); the 15,351 transcripts either known to be (denoted ‘TA’) or not to be involved (denoted ‘nonTA’) in the pathway were used for model training, cross-validation, and testing. Three binary classifier models (logistic regression, LR; random forest, RF; feed-forward neural

network, NN) were trained, cross-validated, and tested for prediction of TA-related genes from tissue abundance profiles using two different training objectives, four different data re-sampling techniques intended for highly class-imbalanced datasets, and two performance metrics (balanced accuracy and computation time). NN models showed the greatest average predictive accuracy on holdout testing data across training parameters (Fig. 68), whereas LR models required the least average computation time—although artificial resampling techniques enabled RF and NN models to be trained and tested in similar durations (Fig. 69).

14.2.2) As all three classifier types with optimal training parameters in Example 14.2.1 showed comparable balanced accuracy (>0.99) and computation times (<3 min) on the testing data, the false positive and false negative rates were next evaluated for each of the three optimized classifiers (Fig. 70). All three models correctly predicted 100% of 'TA' genes (i.e., zero false negatives), whereas the NN model yielded far fewer false positives than either the LR or RF classifiers; 50% of transcripts predicted by the NN classifier to be involved in TA biosynthesis corresponded to known TA-related genes, in contrast to 11% and 14% for the LR and RF models, respectively. The architecture of the optimized NN model with one hidden layer of five nodes is shown in Fig. 71.

14.3) The trained NN classifier was used to identify genes with uncharacterized roles in TA biosynthesis, including TA transport.

14.3.1) The 11,409 transcripts initially identified as encoding genes with 'unknown' involvement in the TA pathway (Fig. 67) and which were therefore not used for model training or testing were presented to the NN classifier, which predicted 33 transcripts to encode novel TA-related gene products (Table 5). Consistent with localization of *Solanaceae* TA biosynthesis in lateral roots, most of the 33 genes show strong expression in secondary roots, moderate expression in primary taproots and calluses, and low expression in stems, flowers, fruits, and leaves. Three transcripts—denoted *A. belladonna* lactose permease-like 1 (*AbLP1*), carbohydrate transporter-like 1 (*AbCT1*), and purine uptake permease-like 1 (*AbPUP1*)—contained amino acid sequences annotated by homology and protein family (PFAM) searches to encode organic small-molecule transporters (Table 5). *Ab*PUP1 shows a highly similar root-specific expression profile to that of known TA enzyme-coding genes, whereas *Ab*LP1 and *Ab*CT1 are more abundant in roots but also expressed in aerial tissues.

14.3.2) As sequence alignments revealed missing regions in the open reading frames of

some of these transcripts, complete protein sequences were generated for the putative TA transporter candidates from Example 14.3.1 via BLAST searches against a more complete transcriptome assembly (see P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619 (2020)) and verified against cDNA synthesized from *A. belladonna* lateral root tissue (Table 6).

14.3.3) Phylogenetic analyses indicated that each of the three putative TA transporters belong to different small-molecule transporter families (Fig. 72). *Ab*PUP1 resembles members of the purine uptake permease (PUP) family, including two transporters previously demonstrated to perform cellular import of nicotine (*Nt*NUP1) in tobacco (see S. B. Hildreth, et al., Tobacco nicotine uptake permease (NUP1) affects alkaloid metabolism. Proc. Natl. Acad. Sci. U. S. A. 108, 18179–18184 (2011)) and benzylisoquinoline alkaloids (BIAs) (*Ps*BUP1) in opium poppy (see M. Dastmalchi, et al., Purine Permease-Type Benzylisoquinoline Alkaloid Transporters in Opium Poppy. Plant Physiol. 181, 916–933 (2019)). In contrast, *Ab*LP1 belongs to the NPF family, which includes a tonoplast-localized strictosidine exporter (*Cr*NPF2.9) implicated in vinca MIA biosynthesis (see R. M. E. Payne, et al., An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. Nat. Plants 3, 1–9 (2017)); and *Ab*CT1, although initially annotated as a carbohydrate transporter, appears to be more closely related to vacuolar metal cation transporters of the zinc-induced facilitator (ZIF) family in *Arabidopsis*.

Example 15: Characterization of putative TA transporters in yeast.

The three TA transporter candidates identified in Example 14 were initially screened for activity in the context of an engineered biosynthetic pathway in which TA production is limited by intracellular transport. As described in the preceding examples, a yeast strain (CSY1297) was engineered for *de novo* biosynthesis of hyoscyamine and scopolamine by expressing 21 different enzymes and disrupting eight endogenous proteins across five metabolic modules and five sub-cellular compartments (see P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619 (2020)) (Figs. 73-75): Module I, designed to increase accumulation of the TA precursor putrescine; Module II, designed to enable conversion of putrescine to the TA acyl acceptor tropine; Module III, designed to convert phenylalanine to the TA acyl donor phenyllactic acid (PLA) glucoside; Module IV, designed to convert the TA esterification product littorine to the downstream TAs hyoscyamine and scopolamine; and Module V, which comprises the central TA esterification reaction in which *A. belladonna* littorine synthase, previously observed to be functional in yeast via expression in the vacuolar membrane as an N-

terminal fusion with the fluorescent protein DsRed (DsRed-AbLS), condenses tropine and PLA glucoside to produce littorine in the vacuole lumen. As uncharged PLA glucoside is relatively membrane-permeable, vacuolar import and export of protonated tropine and littorine, respectively, are primary limitations to flux through this Module (Fig. 73). The tobacco vacuolar nicotine importers *Nt*JAT1 and *Nt*MATE2 were observed in Example 11 to improve vacuolar tropine import in this pathway when expressed individually in the TA-producing strain (CSY1297). However, no transporters have yet been identified to facilitate vacuolar littorine export.

15.1) The capacity of the three newly identified putative TA transporters (see Example 14) to influence TA accumulation in this yeast platform was evaluated. *AbPUP1*, *AbCT1*, and *AbLP1* were each expressed from low-copy plasmids in strain CSY1300, which was constructed by integrating expression cassettes for *NtMATE2* and the selection marker *LEU2* (shown to increase scopolamine accumulation; see P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619 (2020)) into the genome of the TA production strain CSY1297, and transformants were cultured in selective media for 96 h. Liquid chromatography and tandem mass spectrometry (LC-MS/MS) analysis of culture supernatants indicated that *Ab*PUP1 and *Ab*LP1, but not *Ab*CT1, increase TA production (Fig. 76). *Ab*PUP1 expression increased accumulation of hyoscyamine and scopolamine by 2.4- and 1.5-fold (41 µg/l and 102 µg/l), respectively, relative to a blue fluorescent protein (BFP) control, while *Ab*LP1 expression increased accumulation of these products by 2.0- and 1.3-fold (35 µg/l and 88 µg/l). The lack of any substantial increase in accumulation of tropine, as was previously observed with expression of *Nt*JAT1 and *Nt*MATE2 in Example 11, suggested that *Ab*PUP1 and *Ab*LP1 may instead transport substrates downstream of vacuolar tropine esterification.

15.2) The subcellular localization of the putative TA transporters was evaluated using fluorescence microscopy of CSY1300 expressing *Ab*PUP1 and *Ab*LP1 C-terminal GFP fusions from low-copy plasmids (Fig. 77). Both transporters co-localized strongly with *Ds*Red-fused littorine synthase in the yeast vacuole membrane, suggesting a role in vacuolar substrate transport.

15.3) To elucidate the structure and function of these transporters, we constructed plots of protein hydropathy and putative membrane topology using TMHMM, and generated homology models and gene ontology (GO) maps using C-I-TASSER (Figs. 78, 79). *Ab*PUP1 and *Ab*LP1 comprise 10 and 12 transmembrane α-helices, respectively (Fig. 78a-d), and both transporters are predicted to be oriented in the vacuole membrane with

cytosol-facing N- and C-termini (Fig. 78e). For eukaryotic purine permeases, this topology is consistent with transport of substrates from the vacuole lumen (or extracellular space) into the cytosol (see Y. Alguel, et al., Structure of eukaryotic purine/H+ symporter UapA suggests a role for homodimerization in transport activity. Nat. Commun. 7, 1–9 (2016)). GO prediction corroborated that *Ab*PUP1 is likely an intracellular transporter with specificity for nitrogen-containing organic compounds like purines or alkaloids (Fig. 79a-c), whereas *Ab*LP1 showed features consistent with symport of organic substrates and cations (Fig. 79d-f).

15.4) To corroborate their predicted role as vacuolar alkaloid transporters, the substrate specificity and transport mechanism of *Ab*PUP1 and *Ab*LP1 was characterized via growth and toxicity assays in yeast. AD12345678 (abbreviated AD1-8), a yeast strain engineered for increased sensitivity to xenobiotics due to disruptions to seven major multidrug resistance ABCs and their transcriptional regulator, was used for these assays. Although *Ab*PUP1 and *Ab*LP1 are primarily found in the yeast vacuole membrane, a fraction of cellular *Ab*PUP1, and to a lesser extent *Ab*LP1, also appears to localize to the plasma membrane (Fig. 77). Based on the topological equivalence of the vacuole lumen and extracellular space, transporters exporting substrates from the vacuole to the cytosol should also be capable of importing the same extracellular substrates when situated in the plasma membrane, causing increased sensitivity to high substrate concentrations.

15.4.1) The growth of strain AD1-8 expressing *AbPUP1*, *AbLP1*, or a BFP control from low-copy plasmids was monitored in liquid selective media supplemented with various alkaloids at concentrations between 0 and 20 mM (Fig. 80). Expression of either *Ab*PUP1 or *Ab*LP1 conferred acute sensitivity to hyoscyamine (θ_{HM}, ratio of time for half-maximal OD_{600} at 20 mM relative to at 0 mM: control, θ_{HM} ~ 1.1; *Ab*PUP1, θ_{HM} > 3; *Ab*LP1, θ_{HM} ~ 1.5) and modest sensitivity to littorine (control, θ_{HM} ~ 1.0; *Ab*PUP1, θ_{HM} ~ 1.3; *Ab*LP1, θ_{HM} ~ 1.1), but not to tropine or scopolamine (Fig. 81). *Ab*LP1 additionally increased sensitivity to high concentrations of noscapine (θ_{HM} ~ 1.5 to 1.9), a non-TA control. These data suggest that *Ab*PUP1 and *Ab*LP1 export vacuolar littorine and hyoscyamine to the yeast cytosol, and also enable yeast to reuptake these metabolites from the extracellular space.

15.4.2) Many plant alkaloid transporters not driven by ATP hydrolysis are instead dependent on proton gradients for substrate translocation (see N. Shitan, K. Kato, T. Shoji, Alkaloid transporters in plants. Plant Biotechnol. 31, 453–463 (2014)). The proton dependence of *Ab*PUP1- and *Ab*LP1-mediated alkaloid sensitivity in yeast was investigated by monitoring the growth of AD1-8 expressing transporters or a BFP control in pH-buffered

media supplemented with 10 mM hyoscyamine (Fig. 82). Hyoscyamine sensitivity conferred by *Ab*PUP1 was abrogated by increasing buffer pH, indicating that cellular hyoscyamine import (and by equivalence, vacuolar export) by *Ab*PUP1 occurs via a proton-dependent symport mechanism. In contrast, hyoscyamine transport by *Ab*LP1 did not appear to be pH-dependent.

15.4.3) To determine whether transport by *Ab*PUP1 or *Ab*LP1 can be powered by gradients in other common symported cations, the growth of the transporter-expressing yeast strains was monitored in media supplemented with hyoscyamine and either sodium or potassium chloride (Fig. 83). Transport by *Ab*PUP1 was impaired by either Na^+ or K^+, as indicated by reduced alkaloid sensitivity at higher cation concentrations; whereas *Ab*LP1-mediated transport appeared to be slightly enhanced at high [Na^+] but not at high [K^+].

Example 16. Yeast strain engineering and media optimization for increased TA production.

Production of both canonical TAs and novel TA derivatives in the engineered yeast platform described in the prior Examples is hindered by low pathway flux to hyoscyamine and scopolamine. Thus, combining several metabolic engineering strategies, including transport engineering, cofactor regeneration, and media optimization, enables the development of an improved TA production platform suitable for *de novo* derivatization studies.

16.1) The previously identified vacuolar tropine importers *Nt*JAT1 and *Nt*MATE2 were incorporated into a single yeast strain as a starting point for TA titer optimization. Co-expression of *NtMATE2* from a low-copy plasmid with chromosomal *NtJAT1* in strain CSY1299 increased scopolamine production by 29% (24 µg/l to 31 µg/l) and chromosomal expression of both transporters and the *LEU2* selection marker in CSY1300 afforded a further increase of 26% (38 µg/l) (Fig. 84).

16.2) The engineered TA pathway described in the preceding Examples incorporates five enzymes (*Ab*CYP82M3, *Ds*TR1, *Wf*PPR, *Ab*CYP80F1, and *Ds*HDH) that use NADPH for electron transfer. As NADPH depletion by overexpressed enzymes may limit pathway flux, expanding cellular NADPH availability may increase TA accumulation.

16.2.1) Following identification of native yeast enzymes driving NADPH regeneration in central metabolism (Fig. 85) via substrate oxidation (Ald6p, Idp1p-3p, Pdc6p, Zwf1p) or NAD(H) phosphorylation (Pos5p, Yef1p, Utr1p), each corresponding gene was

overexpressed from a low-copy plasmid in CSY1300, and TA accumulation in media was measured following 96 h of growth of transformed strains (Fig. 86). Overexpression of peroxisomal isocitrate dehydrogenase (Idp3p) and pyruvate decarboxylase (Pdc6p) respectively increased production of hyoscyamine 2.5-fold (29 µg/l) and 3.5-fold (41 µg/l) relative to control, although no changes in scopolamine production were noted.

16.2.2) To further increase TA production, the NADPH regeneration mechanisms described in Example 16.2.1 were incorporated together with vacuolar export optimization in the construction of strain CSY1323. Expression cassettes for *AbPUP1*, *AbLP1*, *IDP3*, and *PDC6* were integrated simultaneously with disruption of *TPO5*, which encodes a transporter responsible for Golgi-mediated exocytosis of the TA precursor putrescine (see K. Tachihara et al., Excretion of putrescine and spermidine by the protein encoded by YKL174c (TPO5) in Saccharomyces cerevisiae. J. Biol. Chem. 280, 12637–12642 (2005)), into the genome of CSY1300. CSY1323 produced hyoscyamine and scopolamine at titers 26-fold (56 µg/l) and 1.5-fold (59 µg/l) greater than those of CSY1300 following 96 h of growth in selective media (Fig. 84).

16.3) The effects of amino acid prototrophy on TA production in CSY1323 were examined. It has been previously observed that elimination of leucine auxotrophy and enhanced NADH regeneration via *LEU2* overexpression promotes scopolamine accumulation (see P. Srinivasan, C. D. Smolke, Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619 (2020)). As the TA-producing strains described in the preceding Examples are derived from a histidine prototroph, strains prototrophic for uracil and tryptophan (the remaining two auxotrophies) were constructed by overexpressing the *URA3* and *TRP1* genes from low-copy plasmids in CSY1323. LC-MS/MS analysis of media supernatant from 96 h cultures of wild-type and prototrophic CSY1323 strains indicated that whereas uracil prototrophy does not improve TA production (Fig. 87), tryptophan prototrophy (denoted CSY1324) promotes conversion of hyoscyamine (36% decrease; 56 µg/l to 36 µg/l) to scopolamine (2.9-fold increase; 59 µg/l to 172 µg/l) (Fig. 84).

16.4) Methods for modulating the relative accumulation of different TA products may facilitate the discovery and production of specific derivatives. As *Ab*PUP1 transports hyoscyamine but not scopolamine in a proton-dependent and potassium-inhibited manner, media pH buffering and its concomitant effect on cellular TA transport may be used to influence the hyoscyamine:scopolamine production ratio. TA production was measured in cultures of CSY1324 following 96 h of growth in selective media buffered between 5.4 and 7.0 with 0.1 M potassium phosphate (K_2HPO_4/KH_2PO_4). In this buffer system, [H^+] and [K^+]

respectively decrease and increase with higher pH. At all tested buffer conditions, hyoscyamine accumulation was substantially increased (≥ 5.3-fold) at the expense of decreased scopolamine production relative to CSY1324 grown without buffering (Fig. 88, 84). Maximum hyoscyamine production was obtained at pH 5.8, whereas scopolamine production decreased monotonically with increasing pH (Fig. 88). Similar trends in TA accumulation were observed in buffered cultures of CSY1323, although production of both TAs decreased with higher pH (Fig. 89). This result revealed a simple method for tuning the TA production ratio using pH buffers: hyoscyamine accumulation is maximized during buffered growth at pH 5.8, whereas scopolamine production is maximized in unbuffered media. Under optimal conditions for each TA, hyoscyamine and scopolamine production by CSY1324 reached titers of 480 µg/l and 172 µg/l, representing improvements of 114- and 7.2-fold relative to the starting strain (CSY1298; Fig. 84).

Example 17. *De novo* biosynthesis of computationally predicted medicinal TA derivatives in yeast.

Heterologous production of PNP derivatives has been challenged by limited information on the non-native substrate tolerance and activity of most biosynthetic enzymes. Thus, computational tools for *in silico* biosynthetic pathway expansion can be used to extend an engineered TA pathway for *de novo* production of useful derivatives in non-plant cells such as yeast.

17.1) Using the ATLASx pathway prediction utility (see H. Mohammadi-peyhani et al. ATLASx : a computational map for the exploration of biochemical space. bioRxiv, 1–34 (2021)), a map of all possible derivatives that might be produced in one biosynthetic step from hyoscyamine and scopolamine by known enzyme activities was generated (Fig. 90). The majority of commercial TA-derived drugs are *N*-functionalized derivatives of tropane esters (e.g., trospium, ipratropium, tiotropium, oxitropium, *N*-butylscopolamine), since quaternary amines exhibit reduced blood-brain barrier permeability and central nervous system (CNS) toxicity (see K. L. Kohnen-Johannsen, O. Kayser, Tropane alkaloids: Chemistry, pharmacology, biosynthesis and production. Molecules 24, 1–23 (2019)). As microbial production of secondary amine TAs might facilitate screening, discovery, and *de novo* biosynthesis of pharmaceutical derivatives without *N*-methyl groups (e.g., trospium), the ATLASx-generated map was searched for both secondary and quaternary amine TA derivatives. Two derivative classes of interest were predicted to be accessible via one-step transformations: *N*-demethylated nortropane alkaloids (norTAs) and tropane *N*-oxides (Fig. 90).

17.2) *N*-demethylated derivatives are byproducts of hepatic catabolism of diverse alkaloid drugs, including the TA cocaine (see B. W. LeDuc, et al., Norcocaine and N-hydroxynorcocaine formation in human liver microsomes: Role of cytochrome P-450 3A4. Pharmacology 46, 294–300 (1993)). Thus, one or more human liver CYP450 enzymes (*Hs*CYPs) may be capable of hyoscyamine and scopolamine *N*-demethylation in yeast (Fig. 91). As prior attempts to express *Hs*CYPs in yeast observed increased heterologous activity when paired with yeast NADPH:CYP450 reductase (Ncp1p) fused to human cytochrome b5 (*Hs*CYB5A) (see H. Inui et al., Molecular characterization of specifically active recombinant fused enzymes consisting of CYP3A4, NADPH-cytochrome P450 oxidoreductase, and cytochrome b5. Biochemistry 46, 10213–10221 (2007)), each of the eight *Hs*CYPs collectively responsible for >80% of hepatic drug metabolism (see U. M. Zanger, M. Schwab, Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol. Ther. 138, 103–141 (2013)) was coexpressed with a Ncp1p-HsCYB5A fusion protein from low-copy plasmids (Fig. 92) in the optimized TA production strain described in Example 16 (CSY1324). Using an extended LC-MS/MS method developed for unambiguous identification of TAs and their derivatives (Fig. 93), the culture media of each *Hs*CYP-expressing strain was analyzed for production of norhyoscyamine and norscopolamine (Fig. 94) following 96 h of growth in pH-buffered or non-buffered media, respectively. Low (≤5 µg/l) accumulation of both derivatives was observed in the negative control, suggesting an existing pathway for norTA synthesis in this engineered yeast platform. Expression of *Hs*CYP2D6 improved production of norhyoscyamine and norscopolamine by over threefold (5 µg/l to 18 µg/l) and nearly fourfold (3 µg/L to 11 µg/l) in buffered (pH 5.8) and unbuffered media respectively, while expression of *Hs*CYP2C19 only improved production of norscopolamine in unbuffered media (3 µg/l to 9 µg/l), reflecting an increase in the norTA:TA production ratio from ≤ 3% to nearly 10% (Figs. 94, 95).

17.3) Due to their increased water solubility and lower membrane permeability compared to free bases, drugs derived from alkaloid *N*-oxides may pose lower risk of CNS disruption. Although *N*-oxides are also thought to be trace byproducts of alkaloid catabolism in mammals (see J. D. Phillipson et al., Metabolic N-oxidation of atropine, hyoscine and the corresponding nor-alkaloids by guinea-pig liver microsomal preparations. J. Pharm. Pharmacol. 28, 687–691 (1976), no TA *N*-oxygenases have yet been identified, and no *N*-oxide accumulation was observed in TA-producing yeast strains expressing *Hs*CYPs (see Example 17.2). BridgIT (see N. Hadadi et al., Enzyme annotation for orphan and novel reactions using knowledge of substrate reactive sites. Proc. Natl. Acad. Sci. U. S. A. 116,

7298–7307 (2019)), a computational enzyme prediction tool integrated into the ATLASx webserver, was used to identify putative enzyme candidates for conversion of hyoscyamine and scopolamine to cognate *N*-oxides (Table 7). The highest-scoring enzyme prediction for both TAs was senecionine *N*-oxygenase (SNO; KEGG R07373, EC 1.14.13.101), an NADPH- and flavin-dependent monooxygenase responsible for detoxification and weaponization of plant pyrrolizidine alkaloids in some herbivorous insects (see T. Hartmann, D. Ober, Defense by pyrrolizidine alkaloids: Developed by plants and recruited by insects. Induc. Plant Resist. To Herbiv. 213-231, (2008)). The pyrrolizidine moiety of senecionine, the canonical substrate of SNO and related pyrrolizidine *N*-oxygenases (PNOs), closely resembles the functional groups and atom connectivity found in the tropane ring, suggesting that these enzymes may also be capable of *N*-oxygenating TAs (Fig. 96). As the substrate scope of PNOs appears to have co-evolved with the dietary alkaloid diversity of their host insects (see M. Macel, Attract and deter: A dual role for pyrrolizidine alkaloids in plant-insect interactions. Phytochem. Rev. 10, 75–82 (2011)), PNO orthologs from three different insect species were selected for testing in our yeast TA production platform: *Tj*SNO from *Tyria jacobaeae* (cinnabar moth), a specialist feeder; *Gg*PNO from *Grammia geneura* (Nevada tiger moth), a generalist herbivore of wildflowers; and *Zv*PNO from *Zonocerus variegatus* (harlequin locust), an omniherbivorous agricultural pest (see C. Naumann et al., Evolutionary recruitment of a flavin-dependent monooxygenase for the detoxification of host plant-acquired pyrrolizidine alkaloids in the alkaloid-defended arctiid moth Tyria jacobaeae. Proc. Natl. Acad. Sci. U. S. A. 99, 6085– 6090 (2002); see also S. Sehlmeyer, et al., Flavin-dependent monooxygenases as a detoxification mechanism in insects: New insights from the arctiids (Lepidoptera). PLoS One 5 (2010); see also L. Wang et al., Independent recruitment of a flavin-dependent monooxygenase for safe accumulation of sequestered pyrrolizidine alkaloids in grasshoppers and moths. PLoS One 7, (2012)). Each of the three candidates was expressed from low-copy plasmids in CSY1324 and the culture media was analyzed by LC-MS/MS after 96 h of growth of transformed strains in pH-buffered or non-buffered media. *De novo* production of both hyoscyamine *N*-oxide (71 µg/l in buffered media, pH 5.8) and scopolamine *N*-oxide (14 µg/l in unbuffered media) by the strain expressing *Zv*PNO was observed (Fig. 97), corroborating ATLASx/BridgIT predictions that a SNO/PNO enzyme can *N*-oxygenate TAs.

Table 1: Genes of interest as components of the engineered metabolic pathways

Enzyme	**Abbrev.**	**Catalyzed Reactions**	**Source organisms**	**GenBank #**
N-acetylglutamate synthase	GAT, ARG2	Acetyl-CoA + L-glutamate → CoA + N-acetyl-L-glutamate (EC 2.3.1.1)	*Saccharomyces cerevisiae*	NP_012464.1
Arginine decarboxylase	ADC	L-arginine → agmatine + CO_2 (EC 4.1.1.19)	*Arabidopsis thaliana* *Avena sativa* *Escherichia coli* *Erythroxylum coca* *Nicotiana tabacum*	NP_179243.1, NP_195197.1, CAA40137.1 NP_417413.1 AEQ02349.1 BAA21617.1
Arginase	CAR1	H_2O + L-arginine → L-ornithine + urea (EC 3.5.3.1)	*Saccharomyces cerevisiae*	NP_015214.1
Agmatine ureohydrolase	AUH, speB	Agmatine + H_2O → putrescine + urea (EC 3.5.3.11)	*Escherichia coli* *Bacillus subtilis* *Homo sapiens*	NP_417412.1 NP_391629.1 NP_079034.3
Ornithine decarboxylase	ODC, SPE1	L-ornithine → CO_2 + putrescine (EC 4.1.1.17)	*Saccharomyces cerevisiae* *Homo sapiens* *Erythroxylum coca* *Datura stramonium* *Atropa belladonna*	NP_012737.1 NP_001274118.1 AEQ02350.1 CAA61121.1 AIC34713.1

Polyamine oxidase	PAO, FMS1	$H_2O + O_2$ + spermine → 3-aminopropanal + H_2O_2 + spermidine $H_2O + O_2$ + spermidine → 3-aminopropanal + H_2O_2 + putrescine (EC 1.5.3.17)	*Saccharomyces cerevisiae*	NP_013733.1
Aldehyde dehydrogenase	HFD1, ALD2-6	4-methylaminobutanal → 4-methylaminobutyric acid (EC 1.2.1.3)	*Saccharomyces cerevisiae*	NP_013828.1 NP_013893.1 NP_013892.1 NP_015019.1 NP_010996.2 NP_015264.1
Putrescine N-methyltransferase	PMT	Putrescine + S-adenosyl-L-methionine → N-methylputrescine + S-adenosyl-L-homocysteine (EC 2.1.1.53)	*Nicotiana tabacum* *Atropa belladonna* *Hyoscyamus niger* *Calystegia sepium* *Anisodus acutangulus* *Datura stramonium* *Datura innoxia*	NP001312037.1 BAA82261.1, BAA82262.1 BAA82263.1 CAJ46252.1 ACF21005.1 CAE47481.1 CAJ46254.1
N-methylputrescine oxidase	MPO	N-methylputrescine → 4-methylaminobutanal (EC 1.4.3.22)	*Nicotiana tabacum* *Datura metel*	NP_001312728, NP_001311739 JNVS_scaffold_2009311 (from 1000Plants database)
Pyrrolidine ketide synthase	PYKS	N-methylpyrrolinium + 2 malonyl-CoA → 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid (EC 2.3.1.-)	*Atropa belladonna* *Datura stramonium*	AYU65302.1 n/a

Tropinone synthase	CYP82M3	4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid → tropinone (EC 1.14.14.-)	*Atropa belladonna*	AYU65303.1
Cytochrome P450-$NADP^+$ reductase	CPR	NADPH + H+ + n oxidized hemoprotein = NADP+ + a reduced hemoprotein (EC 1.6.2.4)	*Eschscholzia californica* *Papaver somniferum* *Homo sapiens* *Saccharomyces cerevisiae* *Arabidopsis thaliana*	NM118585, many others (Ref PMID 19931102)
Tropinone reductase 1	TR1	Tropinone + NADPH + H^+ → tropine + $NADP^+$ (EC 1.1.1.206)	*Datura stramonium* *Atropa belladonna* *Hyoscyamus niger* *Datura innoxia* *Brugmansia arborea* *Datura metel* *Anisodus luridus*	AAA33281.1 AFP55030.1 BAA13547.1 AIN39992.1 AIN39993.1 AKY01854.1 AGL76989.1
Tropinone reductase 2	TR2	Tropinone + NADPH + H^+ → pseudotropine + $NADP^+$ (EC 1.1.1.236)	*Datura stramonium* *Hyoscyamus niger* *Atropa belladonna* *Anisodus luridus*	AAA33282.1 AAB09776.1 AGH24753.1 AGL76990.1
Prephenate dehydratase	PHA2	Prephenate + H^+ → 3-phenylpyruvate + CO_2 + H_2O (EC 4.2.1.51)	*Saccharomyces cerevisiae*	NP_014083.2

Aromatic aminotransferase	ARO8 ARO9	2-oxoglutarate + aromatic L-amino acid (phenylalanine, tyrosine) → aromatic oxoacid (phenylpyruvate, hydroxyphenylpyruvate) + L-glutamate (EC 2.6.1.57)	*Saccharomyces cerevisiae*	NP_011313.1 NP_012005.1
Phenylpyruvate reductase	PPR, hcxB	3-phenylpyruvate + NADH + H^+ → 3-phenyllactate + NAD^+ (EC 1.1.1.237)	*Escherichia coli* *Lactobacillus sp.* *Wickerhamia fluorescens* *Lactobacillus plantarum* *Atropa belladonna*	NP_415322.1 ALB78224.1 BAK09193.1 AWO81908.1 AZL88830.1
Lactate dehydrogenase	LDH, L-LDH, D-LDH	3-phenylpyruvate + NADH + H^+ → 3-phenyllactate + NAD^+ (EC 1.1.1.27)	*Bacillus coagulans* *Lactobacillus casei* *Lactobacillus plantarum*	ADN38376.1 WP_003567646.1 WP_003642078.1
3-phenyllactic acid UDP-glucosyltransferase 84A27	UGT84A27	3-phenyllactate + UDP-glucose → 1-O-β-phenyllactoyl-glucose + UDP (EC 2.4.1.-)	*Atropa belladonna*	ATG80135.1
UDP-glucose pyrophosphorylase	UGP	Glucose-1-phosphate + UTP → pyrophosphate + UDP-glucose (EC 2.7.7.9)	*Saccharomyces cerevisiae*	P32861
Littorine synthase	LS	1-O-β-phenyllactoyl-glucose + tropine → (*R*)-littorine (EC 2.3.1.-)	*Atropa belladonna*	ATG80136.1

Littorine mutase	CYP80F1	(*R*)-Littorine → hyoscyamine aldehyde (EC 1.14.19.-)	*Atropa belladonna* *Duboisia myoporoides* *Hyoscyamus niger* *Anisodus luridus*	AHZ34577.1 AQU12715.1 ABD39696.1 AGL76933.1
Hyoscyamine 6β-hydroxylase/dioxygenase	H6H	2-oxoglutarate + (*S*)-hyoscyamine + O_2 → (*S*)-anisodamine + CO_2 + succinate (*S*)-anisodamine → (*S*)-scopolamine (EC 1.14.11.11)	*Datura stramonium* *Atropa belladonna* *Hyoscyamus niger* *Brugmansia arborea* *Anisodus luridus* *Anisodus acutangulus* *Datura metel*	ALD59774.1 AEN79443.1 AAA33387.1 ALD59773.1 AGL76991.1 ABM74185.1 AAQ04302.1
Multidrug and toxin extrusion transporter / jasmonate-inducible transporter	MATE / JAT	TA precursor (compartment A) → TA precursor (compartment B) TA (compartment A) → TA (compartment B)	*Nicotiana tabacum*	AM991692 A3KDM4 A3KDM5
Phenylalanine ammonia-lyase	PAL	L-phenylalanine → NH_4^+ + *trans*-cinnamate (EC 4.3.1.24)	*Arabidopsis thaliana* *Zea mays*	NP_181241.1 AAL40137.1
Tyrosine ammonia-lyase	TAL	L-tyrosine → NH_4^+ + *trans*-4-coumarate (EC 4.3.1.25)	*Rhodosporidium toruloides*	CAA31209.1

4-coumarate-CoA ligase	4CL	4-coumarate + ATP + CoA → 4-coumaroyl-CoA + AMP + diphosphate Cinnamate + ATP + CoA → cinnamoyl-CoA + AMP + diphosphate (promiscuous activity on a variety of aromatic acids: 4CL1-5)	*Arabidopsis thaliana* *Oryza sativa* *Salvia miltiorrhiza* *Solanum tuberosum*	NP_175579.1, NP_188761.1, NP_176686.1, NP_188760.3 XP_015650830.1 AGW27193.1 P31685 Many others
Cocaine synthase	CS	Promiscuous acyltransferase: Aromatic acyl-CoA + tropine or pseudotropine → tropane ester E.g., cinnamoyl-CoA + tropine → cinnamoyltropine	*Erythroxylum coca*	AGT56097.1
3-isopropylmalate dehydrogenase	LEU2	3-isopropylmalate + NAD+ → 4-methyl-2-oxopentanoate + CO2 + NADH	*Saccharomyces cerevisiae*	P04173
Magnesium-activated aldehyde dehydrogenase, cytosolic	ALD6	aldehyde + H2O + NADP+ → carboxylate + 2 H+ + NADPH	*Saccharomyces cerevisiae*	P54115
Isocitrate dehydrogenase [NADP], mitochondrial	IDP1	D-threo-isocitrate + NADP+ → 2-oxoglutarate + CO2 + NADPH	*Saccharomyces cerevisiae*	P21954
Isocitrate dehydrogenase [NADP], cytosolic	IDP2	D-threo-isocitrate + NADP+ → 2-oxoglutarate + CO2 + NADPH	*Saccharomyces cerevisiae*	P41939

Isocitrate dehydrogenase [NADP], peroxisomal	IDP3	D-threo-isocitrate + NADP+ → 2-oxoglutarate + CO2 + NADPH	*Saccharomyces cerevisiae*	P53982
Pyruvate decarboxylase isozyme 3	PDC6	H+ + pyruvate → acetaldehyde + CO2	*Saccharomyces cerevisiae*	P26263
NADH kinase, mitochondrial	POS5	ATP + NADH → ADP + H+ + NADPH	*Saccharomyces cerevisiae*	Q06892
NAD(+) kinase	UTR1	ATP + NAD+ → ADP + H+ + NADP+	*Saccharomyces cerevisiae*	P21373
ATP-NADH kinase	YEF1	ATP + NADH → ADP + H+ + NADPH	*Saccharomyces cerevisiae*	P32622
Glucose-6-phosphate 1-dehydrogenase	ZWF1	D-glucose 6-phosphate + NADP+ → 6-phospho-D-glucono-1,5-lactone + H+ + NADPH	*Saccharomyces cerevisiae*	P11412
Polyamine transporter	TPO5	Putrescine (cytosol) → putrescine (Golgi) → putrescine (extracellular)	*Saccharomyces cerevisiae*	P36029
Orotidine 5'-phosphate decarboxylase	URA3	H+ + orotidine 5'-phosphate → CO2 + UMP	*Saccharomyces cerevisiae*	P03962

N-(5'-phosphoribosyl) anthranilate isomerase	TRP1	N-(5-phospho-β-D-ribosyl)anthranilate → 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate	*Saccharomyces cerevisiae*	P00912
NADPH--cytochrome P450 reductase	NCP1	NADPH + 2 oxidized [cytochrome P450] → H+ + NADP+ + 2 reduced [cytochrome P450]	*Saccharomyces cerevisiae*	P16603
Cytochrome P450 1A2	HsCYP1A2	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P05177
Cytochrome P450 2A6	HsCYP2A6	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P11509
Cytochrome P450 2A13	HsCYP2A13	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	Q16696
Cytochrome P450 2C9	HsCYP2C9	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P11712
Cytochrome P450 2C19	HsCYP2C19	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P33261

Cytochrome P450 2D6	HsCYP2D6	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P10635
Cytochrome P450 2E1	HsCYP2E1	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P05181
Cytochrome P450 3A4	HsCYP3A4	an organic molecule + O2 + reduced [NADPH—hemoprotein reductase] → an alcohol + H+ + H2O + oxidized [NADPH—hemoprotein reductase]	*Homo sapiens*	P08684
Cytochrome b5	HsCYB5A	Electron transfer between cytochromes P450 and NADPH-cytochrome P450 reductases	*Homo sapiens*	P00167
Senecionine *N*-oxygenase	TjSNO	NADPH + O2 + senecionine → H2O + NADP+ + senecionine N-oxide	*Tyria jacobaeae*	Q8MP06
Pyrrolizidine alkaloid *N*-oxygenase	GgPNO	NADPH + O2 + a pyrrolizidine → H2O + NADP+ + a pyrrolizidine N-oxide	*Grammia geneura*	CBI83748
Pyrrolizidine alkaloid *N*-oxygenase	ZvPNO	NADPH + O2 + a pyrrolizidine → H2O + NADP+ + a pyrrolizidine N-oxide	*Zonocerus variegatus*	CBX26645
Purine permease 1	AtPUP1	Purine (extracellular) → purine (cytosol)	*Arabidopsis thaliana*	Q9FZ96

Purine permease 2	AtPUP2	Purine (extracellular) → purine (cytosol)	*Arabidopsis thaliana*	Q94GB1
Purine permease 3	AtPUP3	Purine (extracellular) → purine (cytosol)	*Arabidopsis thaliana*	Q9FZ95
Purine permease 14	AtPUP14	Purine (extracellular) → purine (cytosol)	*Arabidopsis thaliana*	Q9FXH5
Nicotine uptake permease 1	NtNUP1	Nicotine (extracellular) → nicotine (cytosol)	*Nicotiana tabacum*	ADP30798
Benzylisoquinoline alkaloid uptake permease 1	PsBUP1	A benzylisoquinoline alkaloid (extracellular) → a benzylisoquinoline alkaloid (cytosol)	*Papaver somniferum*	MH830312
Purine-uracil permease	AtNCS1	Purine (extracellular) → purine (cytosol)	*Arabidopsis thaliana*	Q9LZD0
Purine-cytosine permease	ScFCY2	Purine (extracellular) → purine (cytosol)	*Saccharomyces cerevisiae*	P17064
Nucleoside permease	EcnupG	Nucleoside (extracellular) → nucleoside (cytosol)	*Escherichia coli*	P0AFF4
Jasmonate-inducible alkaloid transporter 1	NtJAT1	Alkaloid (cytosol) → alkaloid (vacuole lumen) Alkaloid (cytosol) → alkaloid (extracellular)	*Nicotiana tabacum*	AM991692
Multidrug and toxic compound extrusion 2	NtMATE2	Alkaloid (cytosol) → alkaloid (vacuole lumen) Alkaloid (cytosol) → alkaloid (extracellular)	*Nicotiana tabacum*	A3KDM5

Nitrate/peptide family 2.9; vacuolar strictosidine exporter	CrNPF2.9	Strictosidine (vacuole lumen) → strictosidine (cytosol)	*Catharanthus roseus*	KX372303
Protein NRT1/ PTR family 6.3 transporter	AtNPF6.3	Nitrate (extracellular) → nitrate (cytosol)	*Arabidopsis thaliana*	Q05085
Protein zinc induced facilitator-like transporter 1	AtZIFL1	Auxin, hormone transport	*Arabidopsis thaliana*	Q94BZ1
High-affinity nitrate transporter-activating protein 2.1	OsNAR2.1	Nitrate transport	*Oryza sativa*	Q6ZI50
Protein zinc induced facilitator transporter 1	AtZIF1	Zinc (cytosol) → zinc (vacuole lumen)	*Arabidopsis thaliana*	Q8RWN2

Table 2: Example full-length amino acid sequences of hyoscyamine dehydrogenase (HDH) candidates and experimentally validated enzymes.

Sequence	Description	SEQ. ID NO.
Experimentally validated, no observable HDH activity		
MLETVKYLLGSAGPSGYGSKSTAEKVTEQSIHLRSITAIITGATSGIGAETARV LAKRGAKLILPARSLKAAEETKSRILSESPDADIIVMSLDLSSLSSVRKFVAQF EYLNFRLNILINNAGKFAHQHAISEDGIEMTFATNHLGHFLLTKLLLKNMIETA NKTGVQGRIVNVSSSIHGWFSGDAIQYLRLITKDKSQYDATRAYALSKLANVLH TKELAQILKKMVANVTVNCVHPGIVRTRLTREREGLVTDLVFFLTSKLLKTIPQ AAATTCYVATHPRLADVSGKYFADCNEISSSKLGSNLTEAARIWSASEIMVAKN SNAN*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_3722_iso_1_len_1302_ver_2 >HDH1	SEQ. ID NO. 1
MSKTTPNHTQAVSGWAALDSSGKITPYIFNRRENGVNDVTIKILYCGICHTDLH YAKNDWGVTIYPVVPGHEITGIVVEVGSNVTNFKTGDKVGVGCMSASCLQCESC KNSEENYCDKVQFTYNGVFWDGSITYGGYSKMLVADYRFVVAVPENLPMDRAAP LLCAGVTVFVPMKDNNLIGSPRKNIGVIGLGGLGHLAIKFAKAFGHRVTVISTS LSKEKDAKTKLGADDFIVSSNAQQMQSRQKTLDFILDTVSADHSLGPYLELLKI KGTFVIVGAPDKPMGLPAFPLIFGKRTVKGSMIGSIKETQEMLDICGKYNIMCD IEIVTPDRINEAYERIEKNDIKYRFVIDIDGQSSKL*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_5175_iso_1_len_1282_ver_2 >HDH3	SEQ. ID NO. 2
MAMEGTKVARIKLGSDGLEVSAQGLGCMGMSAFYGPPKPEPDMIQLIHHAINSG VTFLDTSDIYGPHTNEILLGKALKGGIRERVELATKFGISFADGKREVRGDPAY VRATCVASLKRLDVDCIDLYYQHRIDTRVPIEVTVGELKKLVEEGKVKYIGLSE ASASTIRRAHAVHPITAVELEWSLWSRDVEEELVPTCRELGIGIVAYSPLGRGF LSSGSKLLEDMSNEDYRKHLPRFQSENLEHNKKLYERICQTAARMGCTPSQLAL AWVHHQGNDVCPIPGTTKIENLNQNIEALSIKLTSEDMTELESIASANAVQGDR YGSGASTYKDSETPPLSAWKVT*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_5694_iso_2_len_1279_ver_2 >HDH4	SEQ. ID NO. 3

MEVKNKYVAIKSNINGAPQESHFEIKVENLSLIVEPDSKEVIIKNLFVSIDPYQ LNRMKSESSSQAAISYASAITPGKAIDTYGVGRVLVSDRPEFKKDDLVAGLLTW GEYTVVKEGSLLNKLDPLGFPLSNHVGVLGFSGLAAYGGFFEVCKPKPGEKVFV SAASGSVGNLVGQYAKLLGCHVVGSAGSQEKVKLLKETLGFDDAFNYKEETDLK SALKRCFPQGIDVCFDNVGGKMLEAAVANMNLFGRVAICGVISEYTNASTRAAP EMLDIVYKRITIQGFLAADFMKVYADFLSETVEYLQDGKLKAVEDVSEGVESIP SAFIGLFNGDNIGKKIVKVADE*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_6801_iso_1_len_1156_ver_2 >HDH5	SEQ. ID NO. 4
MLRIRSRIISISRSLILRQTSSNKFSTHSERKLEGKVAVITGAASGIGKETAAK FISHGAKVIIADIQKQLGQETASELGPNATFVSCDVTKESDISDVVDFAVSKHG QLDIMYNNAGIACRTTFSIVDLDLAQFDRIMAINVRGVVAGIKHAARVMIPQGS GCILCTGSITGVMGGLAQPTYSTTKSCVIGIVKSTTGELCKHGIRINCISPFAI PTAFSLDEMKEYFPGVEPEGLVKILQNASELKGAYCEPIDVANAAIFLASEDAK FISGENLMVDGGFTSFKKLNLSHLVQ*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_8950_iso_1_len_1109_ver_2 >HDH6	SEQ. ID NO. 5
MASNGISHVNGTLAKVITCRAAVAYGPGQPLVVEQVQVDPPQKMEVRIKILFTS ICHTDLSAWKGENEAQRVYPRILGHEASGVVESVGEGVTDMKTGDHVVPIFNGE CGECVYCNSSKKTNLCGKFRVNPFKSVMANDGKCRFRNKDGNPIYHFLNTSTFS EYTVVDSACLVNIDPHAPLDKMTLLSCGVSTGLGAAWNTADVQTGETVAVFGLG AVGLAVVEGARTRGASRIIGVDINSEKRIKGQAIGITDFINPKEIDVPVHEKIR EMTGGGVHYSFECAGNLEVLREAFSSTHDGWGMTIVLGIHPTPRLLPLHPMELF DGRRIVASVFGDFKGKSQLPFFAKQCMAGVVKLDEFITHELPFEKINEGFQLLV DGKSLRCLLHL*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_11748_iso_1_len_557_ver_2 >HDH7	SEQ. ID NO. 6
MAEKITSLESTRYAVVTGGNKGIGYETCRQLVSKGVVVVLTARDEKRGIEATER LKEESSFTDDQIMFHQLDVVDPDSISSLVDFINTKFGRLDILVNNAGVGGLMVE GDVVILKDLIEGDFVSVSTENEEEGDTEKSIEGIVTNYELTKQCVETNFYGAKR MSEAFIPLLQLSNSPTIVNVASFLGKLKLLCNEWAIKVLSNANNLTEDRVDEVV NEFLKDFTEKSIEAKGWPTYFAAYKVSKAAMIAYTRVLATKYPNFRINSVCPGY CKTDLTANTGSLTAEEGAESLVKLALLPNDGPSGLFFYRKDVAAL*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_12989_iso_1_len_1050_ver_2 >HDH8	SEQ. ID NO. 7
MASVSFLSTIGKRLEGKVAMVTGGASGIGEAIAKLFYEHGAKVAIADVQDELGN SVSNALGGSSNSIYIHCDVTNEDDVQEAVDKTISTFGKLDIMICNAGISDETKP	*A. belladonna* plant source; full-length amino acid sequence	SEQ. ID NO. 8

RIIDNTKADFERVLSINVTGVFLTMKHAARVMVPARIGCIISTSSVSSRVGAAA SHAYCSSKHAVLGLTKNLAVELGQFGIRVNCLSPYAMVTPLAEKVIGLENEELE KALDMVGNLKGVTLRVDDVAKAALFLASDDSKYISGHNLFIDGGFTVYNPGLGM FKYPES*	>aba_locus_13944_iso_1_len_941_ver_2 >HDH9	
MLRIASRGGITSRSLQLLQTFNKEFSTHIERKLEGKVALITGAASGIGKETAAK FINNGAKVIIADVQKQLGQETASQLGPNATFVLCDVTKESDVSNAVDFAVSNHG QLDIMYNNAGIICRTPRNIADLDLDAFDRVMAINVRGMMAGIKHAARVMIPRKA GSILCTASITGTMGGLAQPTYSTTKSCVIGMMRSVTAELCQNGIRINCISPFAI PTPFYIDEMKSYYPGVEPEVLVKMLYRASELNGAYCEPVDVANAAVFLASDDAK YVSGQNLVIDGGFTSYKSLNFPMSDQE*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_16663_iso_1_len_466_ver_2 >HDH10	SEQ. ID NO. 9
MGIPSSVTPIVRRLEGKVAVITGGASGIGEAATRLFVKHGAKVVVADVRDDLGR ALCKELGSNDTISFAHCSVTDENDVQNAIDGAVSRYGMLDIMFNNAGITGNMKD PSILATDYKNFKNVFDVNVYGAFLGARIAAKAMIPTKQGSILFTASIASVIGGI ASPITYASSKHAVVGLTNHLAVELGQYGIRVNCISPYTVATPLVREILGKMDKE KAEEVIMETANLKGKILEPEDIAEAAVYLGSDESKYVSGINLVIDGGYSKTNPL ASMVMQNYI*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_114040_iso_1_len_645_ver_2 >HDH11	SEQ. ID NO. 10
MESKSGEGKIVCVTGASGFIASWLVKLLLHRGYTVNATVRNLKDTSKVAHLLGL DGANERLHLFKAELLEEQSFDAAVDGCEGVFHTASPVSLTAKSKEELVDPAVKG TLNVLRSCAKSPSVLRVVITSSTASVICNKNMSTPGAVADETWYSDPEFCEERE EWYQLSKTLAEQAAWKFAKENEMDLVTLHPGLVIGPLLQPTLNFSCEAIVNFIK EGKEAWSGGVYRFVDVRDVANAHILAFEVPSANGRYCLVGVNGYSSLVLKIVQK LYPSITLPENFEDGLPLTPHFQVSSERAKGLGVKFTPLELSVKDTVESLMEKNF LHI*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_125882_iso_1_len_348_ver_2A >HDH13	SEQ. ID NO. 11
MESKSGEGKIVCVTGASGFIASWLVKLLLHRGYTVNATVRNLKDTSKVAHLLGL DGANERLHLFKAELLEEQSFDAAVDGCEGVFHTASPVSLTAKSKEELVDPAVKG TLNVLRSCAKSPSVLRVVITSSTASVICNKNMSTPGAVADETWYSDPEFCEERK EWYQLSKTLAEKAARRFAKENGIDLVTLHPGLVIGPLLQPTLNFSCEAIVNFIK EGKEAWSGGVYRFVDVRDVANAHILAFEVPSANGRYCLVGVNGYSSLVLKIVQK	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_125882_iso_1_len_348_ver_2B >HDH14	SEQ. ID NO. 12

LYPSITLPENFEDGLPLTPHFQVSSERAKGLGVKFTPLELSVKDTVESLMEKNF LHI*		
Experimentally validated, observable HDH activity		
MASEKSLEEKQAENTFGWAAMDSSGVLSPFTFSRRATGEEDVRLKVLYCGICHS DLGCIKNEWGWCSYPLVPGHEIVGIATEVGSKVTKFKVGDRVGVGCMVGSCGTC QNCTQNQESYCPEVIMTCASAYPDGTPTYGGFSNQMVANEKFVIRIPNSLPLDA AAPLLCAGSTVYSAMKFYGLCSQGLHLGVVGLGGLGHVAVKFAKAFGMKVTVIS TSLGKKEEAINQLGADSFLINTDTEQMQGAMEVMDGIIDTVSALHPIEPLLGLL KSHQGKLIIVGLPNKQPELPVFSLINGRKMIGGSAVGGVKETQEMIDFAAEHNI TADIEIVPMDYVNTAMERLEKGDVKFRFVIDVENTLVAAQT*	*A. belladonna* plant source; full-length amino acid sequence >aba_locus_4635_iso_1_len_1351_ver_2 >AbHDH (HDH2)	SEQ. ID NO. 13
MAAEKLSEEEAVKTFGWAAMDSSGVLSPFEFSRRATGAEDVRLKVLYCGICHSD LGCVKNEWGWCSYPLVPGHEIVGIATEVGSRVTKFKVGDRVGVGCMVGSCGSCQ NCSQNLESYCPEVIMTCASAYPDGTPTYGGFSNQMVANEKFVIQIPEKLPLDAA APLLCAGSTVYSPMKFYGLCSPGLHLGVVGLGGLGHVAVKFAKAFGMKVTVIST SIGKKEEAINQLGADSFLTSTDTEQMQGAMETMDGIIDTVSALHPIEPLVGLLK SHQGKLIIVGLPNKQPELPVFSLINGRKMIGGSAVGGVKETQEMIDFAAKHNIT ADIEIVRMDYVNTAMERLEKGDVKFRFVIDVENTLVPAQT*	*D. innoxia* plant source; full-length amino acid sequence >DiHDH	SEQ. ID NO. 14
MAAEKLEERKRWETFGWAAMDSSGVLSPFEFSRRATGEEDVRLKVLYCGICHSD LGCIKNEWGWCSYPLVPGHEIVGIATEVGSRVTKFKVGDRVGVGCMVGSCGSCQ NCSQNLESYCPEVIMTCASAYPDGTPTYGGFSNQMVANEKFVIQIPEKLPLDAA APLLCAGSTVYSPMKFYGLCSPGLHLGVVGLGGLGHVAVKFAKAFGMKVTVIST SIGKKEEAINQLGADSFLISTDTEQMQGAMETMDGIIDTVSALHPIEPLVGLLK SHRGKLIIVGLPNKQPELPVFSLINGRKMIGGSAVGGVKETQEMIDFAAKHNIT ADIEIVGMDYVNTAMERLEKGDVKFRFVIDVENTLVPAQT*	*D. stramonium* plant source; full-length amino acid sequence >DsHDH	SEQ. ID NO. 15

Table 3: Example full-length amino acid sequences of soluble protein domains that can be fused to the N-terminus of serine carboxypeptidase-like acyltransferases to enable functional acyltransferase expression in non-plant cells.

Sequence	Description	SEQ. ID NO.
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLP VPWPTLVTTFGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRD HMVLLEFVTAAGIIHGMDELYK	Green fluorescent protein; full-length amino acid sequence >GFP	SEQ. ID NO. 16
MSELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTMRIKVVEGGPLPFA FDILATSFLYGSKTFINHTQGIPDFFKQSFPEGFTWERVTTYEDGGVLTATQDT SLQDGCLIYNVKIRGVNFTSNGPVMQKKTLGWEAFTETLYPADGGLEGRNDMAL KLVGGSHLIANIKTTYRSKKPAKNLKMPGVYYVDYRLERIKEANNETYVEQHEV AVARYCDLPSKLGHKLN	Blue fluorescent protein; full-length amino acid sequence >BFP	SEQ. ID NO. 17
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKLICTTGKLP VPWPTLVTTLGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIK ANFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRD HMVLLEFVTAAGITHGMDELYK	Yellow fluorescent protein, variant mVenus; full-length amino acid sequence >mVenus	SEQ. ID NO. 18
MSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAK RQGKEMDSLTFLYDGIEIQADQTPEDLDMEDNDIIEAHREQIGG	Small ubiquitin-related modified with mutated protease cleavage site; full-length amino acid sequence >SUMO*	SEQ. ID NO. 19

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT KGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGG VVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGAL KGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIV EQYERAEGRHSTGGMDELYK	Red-fluorescent protein, variant mCherry; full-length amino acid sequence >mCherry	SEQ. ID NO. 20
MGSQGTNIDSIIHVFLISFPGQGHVNPLLRLGKRLASKGVLVSFCAPECVGKDM RAANNNIISDEPTPYGDGFIRFEFFDGWEYTQPKENRQLEIELANLEVVGRAVL PAMLKENEAKGRPVSCLINNPFIPWVCDVADSLGIPCAVLWVQSCASFSAYYHY HFNLAPFPNESNPNIDVHLPNMPILKWDELPSFLLPSNPYPALANAILRQFNYL SKPIRIFIESFDELEKDIVDYMSDFLPIKTVGPLLVEDPKIEQVVRADLVKADS SITQWLNSKPPSSVVYISFGSIVVPSQEQVDEIAYGILNSGLNFLWIMKPPRKN SSFPTVVLPQGYLDKIGDKGKVVEWCLQEQVLAHPSLACFVTHCGWNSSMEVIA NGVPIVAFPQWGDQVTDAKYLVDEFKIGVRLSRGVTENRVIPRDEVERSLHDVT SGPKVAEMKENALKWKMKATEAVAEGGSSDLNLKSFVDELRTLQNSNKNLAKLA PLSN	*Atropa belladonna* UDP-glucosyltransferase 84A27; full-length amino acid sequence >AbUGT84A27	SEQ. ID NO. 21
MASSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPL PFAWDILSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVT QDSSLQDGRFIYKVKFIGVNFPSDGPVMQKKTMGWEPSTERLYPRDGVLKGEIH KALKLKDGGHYLVEFKSIYMAKKPVQLPGYYYVDSKLDITSHNEDYTIVEQYER TEGRHHLFL	Red-fluorescent protein, variant DsRed; full-length amino acid sequence >DsRed	SEQ. ID NO. 22

Table 4: Example full-length amino acid sequences of heterologous transporter which can be expressed in engineered non-plant cells to translocate TAs, TA derivatives, and/or TA precursors across cellular lipid membranes.

Sequence	Description	SEQ. ID NO.
MVEELPQSLKEKKWQINWDAVSQELKKTSRFMAPMVAVTVFQYLLQVVSVMMVG HLGELALSSVAIATSLTNVTGFSLLTGLVGGMETLCGQAYGAQQYHKLSTYTYT AIISLFLVCIPICVLWCFMDKLLILTGQDHSISVEARKYSLWVIPAIFGGAISK PLSRYSQAQSLILPMLLSSFAVLCFHLPISWALIFKLELGNIGAAIAFSISSWL YVLFLASYVKLSSSCEKTRAPFSMEAFLCIRQFFRLAVPSAVMVCLKWWSFEVL ALVSGLLPNPKLETSVMSICITISQLHFSIPYGFGAAASTRVSNELGAGNPQKA RMAVQVVMFLTVVETLVFNTSLFGSRHVLGKAFSNEKQVVDYIAAMTPFLCLSI VTDSLQIVITGIARGSGWQHIGAYINLVVFYVIAIPLAVVLGFVLHLKAKGLWI GIVVGCAIQSIVLSIVTGFTDWEKQAKKARERVHEGRS	*Nicotiana tabacum* jasmonate-inducible alkaloid transporter 1; full-length amino acid sequence >NtJAT1	SEQ. ID NO. 23
MGKSMKSEVEQPLLIAAHGGSSELEEVLSDTQLPYFRRLRYASWIEFQLLYRLA APSVAVYMINNAMSMSTRIFSGQLGNLQLAAASLGNQGIQLFAYGLMLGMGSAV ETLCGQAYGAHRYEMLGVYLQRATVVLSVTGIPLTVVYLFSKNILLALGESKLV ASAAAVFVYGLIPQIFAYAVNFPIQKFLQAQSIVAPSAFISLGTLFVHILLSWV VVYKIGLGLLGASLVLSFSWWIIVVAQFIYIIKSERCKATWAGFRWEAFSGLCQ FVKLSAGSAVMLCLETWYMQILVLLSGLLKNPEIALASISVCLAVNGLMFMVAV GFNAAASVRVSNELGAAHSKSAAFSVFMVTFISFLIAVVEAIIVLSLRNVISYA FTEGEIVAKEVSELCPFLAVTLILNGIQPVLSGVAVGCGWQAFVAYVNVGCYYG VGIPLGCLLGFKFDLGAKGIWTGMIGGTVMQTVILLWVTFRTDWNKKVECAKKR LDKWENLKGPLNKE	*Nicotiana tabacum* multidrug and toxin extrusion transporter 1; full-length amino acid sequence >NtMATE1	SEQ. ID NO. 24
MGKSMKSEVEQPLLAAAHGGSSELEEVLSDSQLPYFRRLRYASWIEFQLLYRLA APSVAVYMINNAMSMSTRIFSGQLGNLQLAAASLGNQGIQLFAYGLMLGMGSAV ETLCGQAYGAHRYEMLGVYLQRATVVLSLTGIPLAVVYLFSKNILLALGESKLV ASAAAVFVYGLIPQIFAYAVNFPIQKFLQSQSIVAPSAFISLGTLFVHILLSWV VVYKIGLGLLGASLVLSFSWWIIVVAQFIYILKSERCKATWAGFRWEAFSGLWQ FVKLSAGSAVMLCLETWYFQILVLLSGLLKNPEIALASISVCLAVNGLMFMVAV	*Nicotiana tabacum* multidrug and toxin extrusion transporter 2; full-length amino acid sequence >NtMATE2	SEQ. ID NO. 25

GFNAAASVRVSNELGAAHPKSAAFSVFMVTFISFLIAVVEAIIVLSLRNVISYA FTEGEVVAKEVSSLCPYLAVTLILNGIQPVLSGVAVGCGWQAFVAYVNVGCYYG VGIPLGCLLGFKFDFGAKGIWTGMIGGTVMQTIILLWVTFSTDWNKEVESARKR LDKWENLKGPLNKE		

Table 5. Example of uncharacterized *A. belladonna* genes with putative roles in TA biosynthesis identified via a neural network binary classifier. The following genes were originally classified as 'unknown' with respect to involvement in TA biosynthesis, and were later predicted to be involved in TA biosynthesis by the trained neural network classifier. Putative TA metabolite transporters based on unified functional annotation are shaded. Values indicate transcript abundance in each tissue, shown as base-2 logarithm of fragments per kilobase of transcript per million mapped reads, $\log_2$(FPKM). Tissue abbreviations: FL, flower; MS, mature seed; PT, primary taproot; SS, sterile seedling; CA, callus; SR, secondary root; ST, stem; RF, ripe fruit; GF, green fruit; LF, leaf; FB, flower buds.

Locus ID	Unified functional annotation	FL	MS	PT	SS	CA	SR	ST	RF	GF	LF	FB
aba_locus_10339_iso_2_len_1347_ver_2	Flavonol synthase/flavanone 3-hydroxylase	0	0	6.23	5.91	0.94	9.17	0	0	0	-0.97	0
aba_locus_10736_iso_1_len_1970_ver_2	AMP dependent ligase	2.07	0	4.59	5.24	2.93	8.59	0	0	0	-0.15	3.38
aba_locus_11227_iso_4_len_1737_ver_2	Lactose permease (**AbLP1**)	0.27	1.07	6.56	6.31	5.25	8.44	2.43	1.40	2.10	0.38	1.86
aba_locus_11280_iso_2_len_1135_ver_2	Asparagine synthetase	3.93	5.01	3.86	6.43	2.11	7.32	0	0	0	0.91	2.77
aba_locus_11352_iso_2_len_1615_ver_2	Carbohydrate transporter (**AbCT1**)	0.63	0.42	6.32	5.34	4.39	8.18	1.73	1.21	2.66	1.34	2.95
aba_locus_11585_iso_3_len_1163_ver_2	Peroxidase 2	1.52	-0.25	4.12	7.62	5.79	7.38	0	0	0	0.78	0.35
aba_locus_116117_iso_1_len_854_ver_2	Pathogen-inducible alpha-dioxygenase	0	0	1.87	3.77	0	6.25	0	0	0	0	0
aba_locus_117832_iso_1_len_388_ver_2	Asparagine synthetase [glutamine-hydrolyzing]	3.91	4.34	3.60	5.88	2.13	6.89	0	0	0	0	2.08
aba_locus_121045_iso_1_len_502_ver_2	High affinity sulfate transporter	0	0	1.81	5.03	1.53	4.93	0	0	0	0	0
aba_locus_122682_iso_1_len_290_ver_2	Methyltransferase family protein	0	0	1.85	4.13	0	7.11	0	0	0	0	1.79
aba_locus_123681_iso_1_len_824_ver_2	Endo-1,4-beta-glucanase	-0.01	0	6.66	4.14	3.71	3.58	1.44	0	0	1.84	5.05
aba_locus_125349_iso_1_len_663_ver_2	Sulfate transporter	0	0	1.17	4.73	0.59	5.06	0.20	0	0	0	0
aba_locus_126696_iso_1_len_540_ver_2	Tyrosine aminotransferase	1.22	0	3.12	4.93	0	7.03	0.52	0	0.80	0	1.82
aba_locus_126900_iso_1_len_415_ver_2	Aminotransferase family protein	0	0	2.70	4.08	0	6.01	0	0	0	0	0.97
aba_locus_127277_iso_1_len_328_ver_2	2-oxoglutarate-dependent dioxygenase	0	0	3.43	2.48	0	8.22	0	0	0	0	0
aba_locus_12794_iso_2_len_956_ver_2	Glutathione S-transferase parC	2.4	5.45	7.96	7.26	4.52	9.43	3.26	0	0	0	3.31
aba_locus_128470_iso_1_len_287_ver_2	Sulphate transporter	0	0	1.54	5.37	2.78	6.03	0	0	2.25	0	0
aba_locus_128994_iso_1_len_358_ver_2	Brassinosteroid insensitive 1-associated receptor kinase 1	0	0	0	6.30	0	3.60	0	0	0	0	0
aba_locus_14132_iso_1_len_1526_ver_2	Hydrolase	0.76	5.63	4.14	3.28	4.66	5.58	-0.91	-0.34	-0.64	0	2.44
aba_locus_15658_iso_1_len_1012_ver_2	Peroxidase 10	0	0	2.02	2.39	0	7.14	0	0	0	0	0
aba_locus_19023_iso_1_len_932_ver_2	1-aminocyclopropane-1-carboxylate oxidase	0.07	0	1.22	5.34	2.21	5.82	0	0	0	0	1.71
aba_locus_19485_iso_1_len_540_ver_2	Limonoid UDP-glucosyltransferase	0	0	5.95	5.96	4.86	8.27	2.18	0	2.31	0	1.98
aba_locus_21462_iso_1_len_1108_ver_2	Purine transporter (**AbPUP1**)	-0.02	0	3.72	5.28	3.27	7.8	0.50	0	0	0	-0.05
aba_locus_22387_iso_2_len_498_ver_2	2-oxoglutarate-dependent dioxygenase	0	0	3.63	3.48	1.22	9.03	0	0	0	1.10	0
aba_locus_4041_iso_2_len_1474_ver_2	Branched-chain amino acid aminotransferase	3.26	-1.13	6.66	4.28	1.66	7.04	1.03	0	0	0.16	3.92
aba_locus_41289_iso_1_len_300_ver_2	Caffeic acid 3-O-methyltransferase	0	0	2.74	1.69	0	6.57	0	0	0	0	0
aba_locus_5126_iso_1_len_553_ver_2	Hydrolase, alpha/beta fold family protein	0	0	3.79	3.77	0	6.54	0	0	0	0	0.26
aba_locus_55852_iso_1_len_329_ver_2	Cembratrienol synthase 2a	1.98	0	3.81	5.20	1.77	5.63	0	0	0	0	0
aba_locus_5597_iso_1_len_1730_ver_2	Beta-glucosidase	0.33	6.81	4.25	5.47	0.87	7.61	0	1.22	0.36	-0.17	2.86
aba_locus_5879_iso_2_len_1038_ver_2	Polyneuridine-aldehyde esterase	0	-0.08	4.24	1.96	-0.08	1.69	1.88	0	0	0.26	2.40
aba_locus_5921_iso_1_len_775_ver_2	2-oxoglutarate-dependent dioxygenase	1.93	0	4.22	2.83	4.96	9.69	0.09	0	2.72	0	0
aba_locus_7330_iso_1_len_1201_ver_2	Orcinol O-methyltransferase	0	0	6.74	6.91	6.34	9.22	-0.33	0.59	2.41	-0.80	1.30
aba_locus_77118_iso_1_len_595_ver_2	Aminotransferase family protein	0	0	3.61	4.94	0.27	7.39	0	0.80	0.95	0	0.35

Table 6. Example cDNA and amino acid sequences of experimentally validated putative TA transporters from *A. belladonna*.

Gene	Type	Sequence (5' ⇒ 3' or N ⇒ C)	SEQ. ID NO.
AbPUP1	cDNA	atggacaatcaaggcagcagcaaatggagaaaacttatacttgtactcaactgtatcact ctaggcataggctattgcggtggtcctttaatttcacgtctctatttcctccacggcggc cgaagaatttggctttctagttggttagcatgcattggatggccaattttcctcattgtt cttatcatcgcctatttttccgtcgtaaaacccagggctctgatatcaagtttctcatc gtgactcgtcaggagttcacaacatccgctgcgctaggcatacttgtcggcctaggcggt tatctttatgcatgggggcccgcgaaactccctgtttctacttcaactgttgtttatgcg gcacaactcggattcaccgtagtatttgcttttctcatagtgaagcagaaattaacggcg tattcggcgaatgcggttgttttgttgatcgttggagcggggactttggcgcttcgagcg gatagtgatcgaccatcaggagagtcaactaaacagtatgttttggggtttgtaatgact cttttggctgcggcgtcaacggggcttgcgttgcctttactggagttgatttatatgaag gctcaaaaaactgttacttatactatggtgttggagattcagatgattctcaatatattc ggttctgctttttgcactgttggaatgattgttaacaaggattttcaggcgatttcgagg gaggcaagccaatttgagcttggagaagcaaaatattatttggtgctagtgtggagtgcc attatttggcaattttcaacactaggaatcgttggagtagtccaatacggctcctcattg ctgtgtggaattttaattgctttcttacttccacttacagaactattaggtgtactttta tatcatgaagtatttcaagctacaaaggggcttgctattttcctctctctttggggattt gtttcttacatctatggtgaagtgaaacaaagccagaagaagaaagaggagaacctaatt ccccaaacagaaatggttcatatcttgcctgtctga	SEQ. ID NO. 26
	Protein	MDNQGSSKWRKLILVLNCITLGIGYCGGPLISRLYFLHGGRRIWLSSWLACIGWPIFLIV LIIAYFFRRKTQGSDIKFLIVTRQEFTTSAALGILVGLGGYLYAWGPAKLPVSTSTVVYA AQLGFTVVFAFLIVKQKLTAYSANAVVLLIVGAGTLALRADSDRPSGESTKQYVLGFVMT LLAAASTGLALPLLELIYMKAQKTVTYTMVLEIQMILNIFGSAFCTVGMIVNKDFQAISR EASQFELGEAKYYLVLVWSAIIWQFSTLGIVGVVQYGSSLLCGILIAFLLPLTELLGVLL YHEVFQATKGLAIFLSLWGFVSYIYGEVKQSQKKKEENLIPQTEMVHILPV*	SEQ. ID NO. 27
AbLP1	cDNA	atgacggacaatgatgaaccaaagccccttggaaggtgttctgtctttacttacggcgtg gggcacatgttgaatgatataacatctgcctgttggtatacatatctcttgctctacttg acagatattggattgtccccaagcaatgctgcaacagtcacacttactggccagtttgct	SEQ. ID NO. 28

gatgcatttacgacaataatagctggtgaactgatagacaggtttggtcatttcaaaata
tggcatggcacgggatgcctgttagttgctctatcttttacatcactcttcggtggttgc
ctcccgtgcaaaattcttggatctgattcacctgttgttcaaactgttggttactgcata
tctgcagtgatcttctgcagcggatggtcttgcactcagatatcacacatgtcaatggtg
aatggcgtcacactaaacccaacaagcagagtagcatgtgttagctgtcgcaatgcttgt
acaatgattgcaagcctaatcttatatgcaatcgcatttttcgtcttcaataagaacact
tctagctcacaggttgacattaagaatcagtaccaattgatggcttatatttcgattttg
attggagcagtcttcgtgattatatttcagcttggaactaaagagccaagtctaaaacaa
gaatctgacggaaaacgtgctacagcttcttggacatactggttgaagaaagtcctatat
tatcaagtacttagcatctatgttttaactagagttgtgactaatgtttcacagaccttc
ttagctgtttacgtaataaatgatctacggatgagcccatcttcgaaagccttgattcct
gccaccatctacttatccagcttcatcgtgtctgtcctcctccaggaactcatatggagt
ggtgaacggttgaaggccttttttttccgctgggggtctcctttggctattttgtggtgct
acagtaattttcctcccaataaacatgaatgcttttatGtaCgtcttgtccgtacttatt
ggcatagcaaacgccttgatgatggtaacatcaatagggatggaaagcgagctagttgat
aaagatcttgatggatctgctttcgtttatggatcattgggctttgttgacaaagttcta
tgtggagtgtttttgtactttctcgagtcatatgagagtgcagatcctgttacctgtgat
ccagcatatccttgtttctccgttacacgattttctctaggtttcattccagggatttct
gccataattggagtcgtagtcacattcttcataagattccacacatcccatcctaagcca
ttaacagaacccctcttggcatag

Protein MTDNDEPKPLGRCSVFTYGVGHMLNDITSACWYTYLLLYLTDIGLSPSNAATVTLTGQFA SEQ. ID NO. 29
DAFTTIIAGELIDRFGHFKIWHGTGCLLVALSFTSLFGGCLPCKILGSDSPVVQTVGYCI
SAVIFCSGWSCTQISHMSMVNGVTLNPTSRVACVSCRNACTMIASLILYAIAFFVFNKNT
SSSQVDIKNQYQLMAYISILIGAVFVIIFQLGTKEPSLKQESDGKRATASWTYWLKKVLY
YQVLSIYVLTRVVTNVSQTFLAVYVINDLRMSPSSKALIPATIYLSSFIVSVLLQELIWS
GERLKAFFSAGGLLWLFCGATVIFLPINMNAFMYVLSVLIGIANALMMVTSIGMESELVD
KDLDGSAFVYGSLGFVDKVLCGVFLYFLESYESADPVTCDPAYPCFSVTRFSLGFIPGIS
AIIGVVVTFFIRFHTSHPKPLTEPLLA*

Table 7. Example of BridgIT scores and closest database matches for top ten enzyme candidates predicted to catalyze *N*-oxidation of hyoscyamine and/or scopolamine.

Rank	Score	Predicted most similar KEGG reaction	Predicted most similar enzyme (EC number)	BridgIT reaction rule
Hyoscyamine -> Hyoscyamine *N*-oxide (ATLASx reaction ID: 1737015052)				
1	0.6367	R07373	1.14.13.101	1.14.13.-
2	0.5783	R00448	1.14.13.59	1.14.13.-
3	0.5740	R10728	1.14.13.194	1.14.13.-
4	0.5695	R11358	1.14.13.221	1.14.13.-
5	0.5695	R09859	1.14.13.141	1.14.13.-
Scopolamine -> Scopolamine *N*-oxide (ATLASx reaction ID: 1737911337)				
1	0.5761	R07373	1.14.13.101	1.14.13.-
2	0.4890	R10662	1.14.13.187	1.14.13.-
3	0.4712	R10728	1.14.13.194	1.14.13.-
4	0.4704	R00448	1.14.13.59	1.14.13.-
5	0.4691	R08925	1.14.13.105	1.14.13.-

Table 8. Example solvent gradient used in rapid LC-MS/MS (method 1) for identification and quantification of tropine, hyoscyamine and scopolamine. Solvent A is 0.1% v/v formic acid in water; solvent B is 0.1% v/v formic acid in acetonitrile.

Time (min)	Volume % solvent B
0.00–0.75	1
0.75–1.33	1–25
1.33–2.70	25–40
2.70–3.70	40–60
3.70–3.71	60–95
3.71–4.33	95
4.33–4.34	95–1
4.34–5.00	1

Table 9. Example solvent gradient used in extended LC-MS/MS (method 2) for identification and quantification of TA derivatives. Solvent A is 0.1% v/v formic acid in water; solvent B is 0.1% v/v formic acid in acetonitrile.

Time (min)	Volume % solvent B
0.00–0.10	3
0.10–5.50	3–28
5.50–6.00	28–98
6.00–7.00	98
7.00–7.50	98–3
7.50–8.50	3

Table 10. Example LC-MS/MS multiple reaction monitoring (MRM) parameters for identification and quantification of TAs and derivatives.

Compound	MRM transition (m/z $[M+H]^+$)	Fragmentor	Collision energy
Tropine	142 → 98	50	21
Hyoscyamine / littorine	290 → 124	150	26
Anisodamine	306 → 140	150	26
Scopolamine	304 → 138	122	22
Norhyoscyamine	276 → 93	150	26
Norscopolamine	290 → 103	94	42
Hyoscyamine N-oxide	306 → 124	150	34
Scopolamine N-oxide	320 → 154	122	22

Table 11: Comparison of impurities that may be present in concentrate of nightshade leaves and clarified yeast culture medium. Notwithstanding the appended clauses, the disclosure may be defined by the following clauses:

Impurities:		**Concentrate of Nightshade Leaves**	**Clarified Yeast Culture Medium**
Inorganic	Sodium	✓	✓
	Magnesium	✓	✓
	Silicon	✓	× (not in culture medium)
	Phosphorus	✓	✓
	Sulfur	✓	✓
	Chloride	✓	✓
	Potassium	✓	✓
	Calcium	✓	✓
	Copper	✓	✓
	Zinc	✓	✓
	Molybdenum	✓	✓ (sodium molybdate in medium)
	Iron	✓	✓
	Manganese	✓	✓
	Ammonium	✓	✓
	Boron	✓	✓
Organic	Polysaccharides (starch, cellulose, xylan)	✓	× (yeast fed simple sugars)
	Lignin (p-cournaryl, coniferyl, sinapyl alcohols)	✓	×
	Pigments (chlorophyll, anthocyanins, carotenoids)	✓	×
	Flavonoids	✓	×
	Phenanthreoids	✓	×
	Latex, gum, and wax	✓	×
	Rubisco	✓	×
	Cuscohygrine	✓	×
Other	Pesticides, Fungicides, Herbicides	✓	×
	Pollen	✓	×

Clause 1. An engineered non-plant cell that produces a tropane alkaloid product, a precursor of a tropane alkaloid product, or a derivative of a tropane alkaloid product.

Clause 2. The cell of clause 1, wherein the cell is a microbial cell.

Clause 3. The cell of clauses 1 or 2, wherein the engineered cell comprises a plurality of heterologous coding sequences for encoding a plurality of enzymes, wherein at least one of the enzymes is selected from the group consisting of arginine decarboxylase, agmatine ureohydrolase, agmatinase, putrescine N-methyltransferase, N-methylputrescine oxidase, pyrrolidine ketide synthase, tropinone synthase, cytochrome P450 reductase, tropinone reductase, phenylpyruvate reductase, 3-phenyllactic acid UDP-glucosyltransferase 84A27, littorine synthase, littorine mutase, hyoscyamine dehydrogenase, hyoscyamine 6β-hydroxylase/dioxygenase, cocaine synthase, a metazoan hepatic cytochrome P450, senecionine *N*-oxygenase, pyrrolizidine *N*-oxygenase, and a tropane alkaloid-detoxifying enzyme.

Clause 4. The cell of any of clauses 1-3, wherein the engineered cell comprises one or more heterologous coding sequences encoding one or more transporters selected from the group of multidrug resistance transporter, pleiotropic drug resistance transporter, ATP-binding cassette transporter, multidrug and toxin extrusion transporter, nitrate transporter 1/peptide transporter family transporter, purine uptake permease-like transporter, and major facilitator superfamily transporter.

Clause 5. The cell of any of clauses 1-4, wherein endogenous arginine metabolism is modified in the cell.

Clause 6. The cell of any of clauses 1-5, wherein endogenous phenylalanine and phenylpropanoid metabolism is modified clauses the cell.

Clause 7. The cell of any of claims 1-6, wherein endogenous polyamine regulatory mechanisms are disrupted in the cell.

Clause 8. The cell of any of the clauses 1-7, wherein endogenous acetate metabolism is modified in the cell.

Clause 9. The cell of any of the clauses 1-8, wherein endogenous glycoside metabolism is modified in the cell.

Clause 10. The cell of any of clauses 1-9, wherein the cell produces a tropane alkaloid product, a precursor of a tropane alkaloid product, or a derivative of a tropane alkaloid product selected from the group consisting of a littorine, hyoscyamine, atropine, anisodamine, scopolamine, calystegine, cocaine, or a non-natural tropane alkaloid.

Clause 11. The cell of any of clauses 1-10, wherein the cell produces a derivative of a tropane alkaloid product selected from the group consisting of a nortropane alkaloid and a tropane *N*-oxide.

Clause 12. The cell of any of the clauses 1-11, wherein the engineered cell comprises a plurality of heterologous coding sequences encoding for a plurality of enzymes which comprise one or more soluble protein domains fused to the N-terminus of a serine carboxypeptidase-like acyltransferase domain.

Clause 13. The cell of any of the clauses 1-12, wherein the transport of TAs, TA precursors, and/or TA derivatives across intracellular membranes or across the plasma membrane is modified in the cell.

Clause 14. A method for producing a tropane alkaloid, a precursor of a tropane alkaloid product, or a derivative of a tropane alkaloid product comprising

(a) culturing a cell of any of clauses 1-13 under conditions suitable for protein production;

(b) adding a starting compound to the cell culture; and

(c) recovering the tropane alkaloid or the precursor of a tropane alkaloid product from the culture.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Accordingly, the preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents

and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and described herein. Rather, the scope and spirit of present invention is embodied by the appended claims.

SEQUENCE LISTING

<110> The Board of Trustees of the Leland Stanford Junior University
Smolke, Christina D.
Srinivasan, Prashanth

<120> Tropane Alkaloid Transporters and Methods of Making Tropane Alkaloids Using the Same

<130> STAN-1840WO

<140> PCT/US2022/021547
<141> 2022-03-23

<150> US 63/166,117
<151> 2021-03-25

<160> 38

<170> PatentIn version 3.5

<210> 1
<211> 328
<212> PRT
<213> Atropa belladonna

<400> 1

Met Leu Glu Thr Val Lys Tyr Leu Leu Gly Ser Ala Gly Pro Ser Gly
1 5 10 15

Tyr Gly Ser Lys Ser Thr Ala Glu Lys Val Thr Glu Gln Ser Ile His
20 25 30

Leu Arg Ser Ile Thr Ala Ile Ile Thr Gly Ala Thr Ser Gly Ile Gly
35 40 45

Ala Glu Thr Ala Arg Val Leu Ala Lys Arg Gly Ala Lys Leu Ile Leu
50 55 60

Pro Ala Arg Ser Leu Lys Ala Ala Glu Glu Thr Lys Ser Arg Ile Leu
65 70 75 80

```
Ser Glu Ser Pro Asp Ala Asp Ile Ile Val Met Ser Leu Asp Leu Ser
                85                  90                  95

Ser Leu Ser Ser Val Arg Lys Phe Val Ala Gln Phe Glu Tyr Leu Asn
            100                 105                 110

Phe Arg Leu Asn Ile Leu Ile Asn Asn Ala Gly Lys Phe Ala His Gln
        115                 120                 125

His Ala Ile Ser Glu Asp Gly Ile Glu Met Thr Phe Ala Thr Asn His
    130                 135                 140

Leu Gly His Phe Leu Leu Thr Lys Leu Leu Leu Lys Asn Met Ile Glu
145                 150                 155                 160

Thr Ala Asn Lys Thr Gly Val Gln Gly Arg Ile Val Asn Val Ser Ser
                165                 170                 175

Ser Ile His Gly Trp Phe Ser Gly Asp Ala Ile Gln Tyr Leu Arg Leu
            180                 185                 190

Ile Thr Lys Asp Lys Ser Gln Tyr Asp Ala Thr Arg Ala Tyr Ala Leu
        195                 200                 205

Ser Lys Leu Ala Asn Val Leu His Thr Lys Glu Leu Ala Gln Ile Leu
    210                 215                 220

Lys Lys Met Val Ala Asn Val Thr Val Asn Cys Val His Pro Gly Ile
225                 230                 235                 240

Val Arg Thr Arg Leu Thr Arg Glu Arg Glu Gly Leu Val Thr Asp Leu
                245                 250                 255

Val Phe Phe Leu Thr Ser Lys Leu Leu Lys Thr Ile Pro Gln Ala Ala
            260                 265                 270

Ala Thr Thr Cys Tyr Val Ala Thr His Pro Arg Leu Ala Asp Val Ser
```

275 280 285

Gly Lys Tyr Phe Ala Asp Cys Asn Glu Ile Ser Ser Ser Lys Leu Gly
290 295 300

Ser Asn Leu Thr Glu Ala Ala Arg Ile Trp Ser Ala Ser Glu Ile Met
305 310 315 320

Val Ala Lys Asn Ser Asn Ala Asn
325

<210> 2
<211> 360
<212> PRT
<213> Atropa belladonna

<400> 2

Met Ser Lys Thr Thr Pro Asn His Thr Gln Ala Val Ser Gly Trp Ala
1 5 10 15

Ala Leu Asp Ser Ser Gly Lys Ile Thr Pro Tyr Ile Phe Asn Arg Arg
20 25 30

Glu Asn Gly Val Asn Asp Val Thr Ile Lys Ile Leu Tyr Cys Gly Ile
35 40 45

Cys His Thr Asp Leu His Tyr Ala Lys Asn Asp Trp Gly Val Thr Ile
50 55 60

Tyr Pro Val Val Pro Gly His Glu Ile Thr Gly Ile Val Val Glu Val
65 70 75 80

Gly Ser Asn Val Thr Asn Phe Lys Thr Gly Asp Lys Val Gly Val Gly
85 90 95

Cys Met Ser Ala Ser Cys Leu Gln Cys Glu Ser Cys Lys Asn Ser Glu
100 105 110

```
Glu Asn Tyr Cys Asp Lys Val Gln Phe Thr Tyr Asn Gly Val Phe Trp
        115                 120                 125

Asp Gly Ser Ile Thr Tyr Gly Gly Tyr Ser Lys Met Leu Val Ala Asp
    130                 135                 140

Tyr Arg Phe Val Val Ala Val Pro Glu Asn Leu Pro Met Asp Arg Ala
145                 150                 155                 160

Ala Pro Leu Leu Cys Ala Gly Val Thr Val Phe Val Pro Met Lys Asp
                165                 170                 175

Asn Asn Leu Ile Gly Ser Pro Arg Lys Asn Ile Gly Val Ile Gly Leu
            180                 185                 190

Gly Gly Leu Gly His Leu Ala Ile Lys Phe Ala Lys Ala Phe Gly His
        195                 200                 205

Arg Val Thr Val Ile Ser Thr Ser Leu Ser Lys Glu Lys Asp Ala Lys
    210                 215                 220

Thr Lys Leu Gly Ala Asp Asp Phe Ile Val Ser Ser Asn Ala Gln Gln
225                 230                 235                 240

Met Gln Ser Arg Gln Lys Thr Leu Asp Phe Ile Leu Asp Thr Val Ser
                245                 250                 255

Ala Asp His Ser Leu Gly Pro Tyr Leu Glu Leu Leu Lys Ile Lys Gly
            260                 265                 270

Thr Phe Val Ile Val Gly Ala Pro Asp Lys Pro Met Gly Leu Pro Ala
        275                 280                 285

Phe Pro Leu Ile Phe Gly Lys Arg Thr Val Lys Gly Ser Met Ile Gly
    290                 295                 300
```

Ser Ile Lys Glu Thr Gln Glu Met Leu Asp Ile Cys Gly Lys Tyr Asn
305 310 315 320

Ile Met Cys Asp Ile Glu Ile Val Thr Pro Asp Arg Ile Asn Glu Ala
325 330 335

Tyr Glu Arg Ile Glu Lys Asn Asp Ile Lys Tyr Arg Phe Val Ile Asp
340 345 350

Ile Asp Gly Gln Ser Ser Lys Leu
355 360

<210> 3
<211> 346
<212> PRT
<213> Atropa belladonna

<400> 3

Met Ala Met Glu Gly Thr Lys Val Ala Arg Ile Lys Leu Gly Ser Asp
1 5 10 15

Gly Leu Glu Val Ser Ala Gln Gly Leu Gly Cys Met Gly Met Ser Ala
20 25 30

Phe Tyr Gly Pro Pro Lys Pro Glu Pro Asp Met Ile Gln Leu Ile His
35 40 45

His Ala Ile Asn Ser Gly Val Thr Phe Leu Asp Thr Ser Asp Ile Tyr
50 55 60

Gly Pro His Thr Asn Glu Ile Leu Leu Gly Lys Ala Leu Lys Gly Gly
65 70 75 80

Ile Arg Glu Arg Val Glu Leu Ala Thr Lys Phe Gly Ile Ser Phe Ala
85 90 95

Asp Gly Lys Arg Glu Val Arg Gly Asp Pro Ala Tyr Val Arg Ala Thr
100 105 110

```
Cys Val Ala Ser Leu Lys Arg Leu Asp Val Asp Cys Ile Asp Leu Tyr
        115                 120                 125

Tyr Gln His Arg Ile Asp Thr Arg Val Pro Ile Glu Val Thr Val Gly
    130                 135                 140

Glu Leu Lys Lys Leu Val Glu Glu Gly Lys Val Lys Tyr Ile Gly Leu
145                 150                 155                 160

Ser Glu Ala Ser Ala Ser Thr Ile Arg Arg Ala His Ala Val His Pro
                165                 170                 175

Ile Thr Ala Val Glu Leu Glu Trp Ser Leu Trp Ser Arg Asp Val Glu
            180                 185                 190

Glu Glu Leu Val Pro Thr Cys Arg Glu Leu Gly Ile Gly Ile Val Ala
        195                 200                 205

Tyr Ser Pro Leu Gly Arg Gly Phe Leu Ser Ser Gly Ser Lys Leu Leu
    210                 215                 220

Glu Asp Met Ser Asn Glu Asp Tyr Arg Lys His Leu Pro Arg Phe Gln
225                 230                 235                 240

Ser Glu Asn Leu Glu His Asn Lys Lys Leu Tyr Glu Arg Ile Cys Gln
                245                 250                 255

Thr Ala Ala Arg Met Gly Cys Thr Pro Ser Gln Leu Ala Leu Ala Trp
            260                 265                 270

Val His His Gln Gly Asn Asp Val Cys Pro Ile Pro Gly Thr Thr Lys
        275                 280                 285

Ile Glu Asn Leu Asn Gln Asn Ile Glu Ala Leu Ser Ile Lys Leu Thr
    290                 295                 300
```

```
Ser Glu Asp Met Thr Glu Leu Glu Ser Ile Ala Ser Ala Asn Ala Val
305                 310                 315                 320


Gln Gly Asp Arg Tyr Gly Ser Gly Ala Ser Thr Tyr Lys Asp Ser Glu
                325                 330                 335


Thr Pro Pro Leu Ser Ala Trp Lys Val Thr
            340                 345


<210>  4
<211>  346
<212>  PRT
<213>  Atropa belladonna

<400>  4

Met Glu Val Lys Asn Lys Tyr Val Ala Ile Lys Ser Asn Ile Asn Gly
1               5                   10                  15


Ala Pro Gln Glu Ser His Phe Glu Ile Lys Val Glu Asn Leu Ser Leu
            20                  25                  30


Ile Val Glu Pro Asp Ser Lys Glu Val Ile Ile Lys Asn Leu Phe Val
        35                  40                  45


Ser Ile Asp Pro Tyr Gln Leu Asn Arg Met Lys Ser Glu Ser Ser Ser
    50                  55                  60


Gln Ala Ala Ile Ser Tyr Ala Ser Ala Ile Thr Pro Gly Lys Ala Ile
65                  70                  75                  80


Asp Thr Tyr Gly Val Gly Arg Val Leu Val Ser Asp Arg Pro Glu Phe
                85                  90                  95


Lys Lys Asp Asp Leu Val Ala Gly Leu Leu Thr Trp Gly Glu Tyr Thr
            100                 105                 110


Val Val Lys Glu Gly Ser Leu Leu Asn Lys Leu Asp Pro Leu Gly Phe
```

115 120 125

Pro Leu Ser Asn His Val Gly Val Leu Gly Phe Ser Gly Leu Ala Ala
130 135 140

Tyr Gly Gly Phe Phe Glu Val Cys Lys Pro Lys Pro Gly Glu Lys Val
145 150 155 160

Phe Val Ser Ala Ala Ser Gly Ser Val Gly Asn Leu Val Gly Gln Tyr
165 170 175

Ala Lys Leu Leu Gly Cys His Val Val Gly Ser Ala Gly Ser Gln Glu
180 185 190

Lys Val Lys Leu Leu Lys Glu Thr Leu Gly Phe Asp Asp Ala Phe Asn
195 200 205

Tyr Lys Glu Glu Thr Asp Leu Lys Ser Ala Leu Lys Arg Cys Phe Pro
210 215 220

Gln Gly Ile Asp Val Cys Phe Asp Asn Val Gly Gly Lys Met Leu Glu
225 230 235 240

Ala Ala Val Ala Asn Met Asn Leu Phe Gly Arg Val Ala Ile Cys Gly
245 250 255

Val Ile Ser Glu Tyr Thr Asn Ala Ser Thr Arg Ala Ala Pro Glu Met
260 265 270

Leu Asp Ile Val Tyr Lys Arg Ile Thr Ile Gln Gly Phe Leu Ala Ala
275 280 285

Asp Phe Met Lys Val Tyr Ala Asp Phe Leu Ser Glu Thr Val Glu Tyr
290 295 300

Leu Gln Asp Gly Lys Leu Lys Ala Val Glu Asp Val Ser Glu Gly Val
305 310 315 320

```
Glu Ser Ile Pro Ser Ala Phe Ile Gly Leu Phe Asn Gly Asp Asn Ile
                325                 330                 335


Gly Lys Lys Ile Val Lys Val Ala Asp Glu
            340                 345


<210>  5
<211>  296
<212>  PRT
<213>  Atropa belladonna

<400>  5

Met Leu Arg Ile Arg Ser Arg Ile Ile Ser Ile Ser Arg Ser Leu Ile
1               5                   10                  15


Leu Arg Gln Thr Ser Ser Asn Lys Phe Ser Thr His Ser Glu Arg Lys
            20                  25                  30


Leu Glu Gly Lys Val Ala Val Ile Thr Gly Ala Ala Ser Gly Ile Gly
        35                  40                  45


Lys Glu Thr Ala Ala Lys Phe Ile Ser His Gly Ala Lys Val Ile Ile
    50                  55                  60


Ala Asp Ile Gln Lys Gln Leu Gly Gln Glu Thr Ala Ser Glu Leu Gly
65                  70                  75                  80


Pro Asn Ala Thr Phe Val Ser Cys Asp Val Thr Lys Glu Ser Asp Ile
                85                  90                  95


Ser Asp Val Val Asp Phe Ala Val Ser Lys His Gly Gln Leu Asp Ile
            100                 105                 110


Met Tyr Asn Asn Ala Gly Ile Ala Cys Arg Thr Thr Phe Ser Ile Val
        115                 120                 125
```

```
                                                10
Asp Leu Asp Leu Ala Gln Phe Asp Arg Ile Met Ala Ile Asn Val Arg
    130                 135                 140


Gly Val Val Ala Gly Ile Lys His Ala Ala Arg Val Met Ile Pro Gln
145                 150                 155                 160


Gly Ser Gly Cys Ile Leu Cys Thr Gly Ser Ile Thr Gly Val Met Gly
                165                 170                 175


Gly Leu Ala Gln Pro Thr Tyr Ser Thr Thr Lys Ser Cys Val Ile Gly
            180                 185                 190


Ile Val Lys Ser Thr Thr Gly Glu Leu Cys Lys His Gly Ile Arg Ile
        195                 200                 205


Asn Cys Ile Ser Pro Phe Ala Ile Pro Thr Ala Phe Ser Leu Asp Glu
    210                 215                 220


Met Lys Glu Tyr Phe Pro Gly Val Glu Pro Glu Gly Leu Val Lys Ile
225                 230                 235                 240


Leu Gln Asn Ala Ser Glu Leu Lys Gly Ala Tyr Cys Glu Pro Ile Asp
                245                 250                 255


Val Ala Asn Ala Ala Ile Phe Leu Ala Ser Glu Asp Ala Lys Phe Ile
            260                 265                 270


Ser Gly Glu Asn Leu Met Val Asp Gly Gly Phe Thr Ser Phe Lys Lys
        275                 280                 285


Leu Asn Leu Ser His Leu Val Gln
    290                 295


<210>  6
<211>  389
<212>  PRT
<213>  Atropa belladonna
```

<400> 6

```
Met Ala Ser Asn Gly Ile Ser His Val Asn Gly Thr Leu Ala Lys Val
1               5                  10                  15

Ile Thr Cys Arg Ala Ala Val Ala Tyr Gly Pro Gly Gln Pro Leu Val
            20                  25                  30

Val Glu Gln Val Gln Val Asp Pro Pro Gln Lys Met Glu Val Arg Ile
        35                  40                  45

Lys Ile Leu Phe Thr Ser Ile Cys His Thr Asp Leu Ser Ala Trp Lys
    50                  55                  60

Gly Glu Asn Glu Ala Gln Arg Val Tyr Pro Arg Ile Leu Gly His Glu
65                  70                  75                  80

Ala Ser Gly Val Val Glu Ser Val Gly Glu Gly Val Thr Asp Met Lys
                85                  90                  95

Thr Gly Asp His Val Val Pro Ile Phe Asn Gly Glu Cys Gly Glu Cys
            100                 105                 110

Val Tyr Cys Asn Ser Ser Lys Lys Thr Asn Leu Cys Gly Lys Phe Arg
        115                 120                 125

Val Asn Pro Phe Lys Ser Val Met Ala Asn Asp Gly Lys Cys Arg Phe
    130                 135                 140

Arg Asn Lys Asp Gly Asn Pro Ile Tyr His Phe Leu Asn Thr Ser Thr
145                 150                 155                 160

Phe Ser Glu Tyr Thr Val Val Asp Ser Ala Cys Leu Val Asn Ile Asp
                165                 170                 175

Pro His Ala Pro Leu Asp Lys Met Thr Leu Leu Ser Cys Gly Val Ser
            180                 185                 190
```

```
Thr Gly Leu Gly Ala Ala Trp Asn Thr Ala Asp Val Gln Thr Gly Glu
        195                 200                 205

Thr Val Ala Val Phe Gly Leu Gly Ala Val Gly Leu Ala Val Val Glu
    210                 215                 220

Gly Ala Arg Thr Arg Gly Ala Ser Arg Ile Ile Gly Val Asp Ile Asn
225                 230                 235                 240

Ser Glu Lys Arg Ile Lys Gly Gln Ala Ile Gly Ile Thr Asp Phe Ile
                245                 250                 255

Asn Pro Lys Glu Ile Asp Val Pro Val His Glu Lys Ile Arg Glu Met
            260                 265                 270

Thr Gly Gly Gly Val His Tyr Ser Phe Glu Cys Ala Gly Asn Leu Glu
        275                 280                 285

Val Leu Arg Glu Ala Phe Ser Ser Thr His Asp Gly Trp Gly Met Thr
    290                 295                 300

Ile Val Leu Gly Ile His Pro Thr Pro Arg Leu Leu Pro Leu His Pro
305                 310                 315                 320

Met Glu Leu Phe Asp Gly Arg Arg Ile Val Ala Ser Val Phe Gly Asp
                325                 330                 335

Phe Lys Gly Lys Ser Gln Leu Pro Phe Phe Ala Lys Gln Cys Met Ala
            340                 345                 350

Gly Val Val Lys Leu Asp Glu Phe Ile Thr His Glu Leu Pro Phe Glu
        355                 360                 365

Lys Ile Asn Glu Gly Phe Gln Leu Leu Val Asp Gly Lys Ser Leu Arg
    370                 375                 380
```

Cys Leu Leu His Leu
385

<210> 7
<211> 315
<212> PRT
<213> Atropa belladonna

<400> 7

Met Ala Glu Lys Ile Thr Ser Leu Glu Ser Thr Arg Tyr Ala Val Val
1 5 10 15

Thr Gly Gly Asn Lys Gly Ile Gly Tyr Glu Thr Cys Arg Gln Leu Val
20 25 30

Ser Lys Gly Val Val Val Val Leu Thr Ala Arg Asp Glu Lys Arg Gly
35 40 45

Ile Glu Ala Thr Glu Arg Leu Lys Glu Glu Ser Ser Phe Thr Asp Asp
50 55 60

Gln Ile Met Phe His Gln Leu Asp Val Val Asp Pro Asp Ser Ile Ser
65 70 75 80

Ser Leu Val Asp Phe Ile Asn Thr Lys Phe Gly Arg Leu Asp Ile Leu
85 90 95

Val Asn Asn Ala Gly Val Gly Gly Leu Met Val Glu Gly Asp Val Val
100 105 110

Ile Leu Lys Asp Leu Ile Glu Gly Asp Phe Val Ser Val Ser Thr Glu
115 120 125

Asn Glu Glu Glu Gly Asp Thr Glu Lys Ser Ile Glu Gly Ile Val Thr
130 135 140

Asn Tyr Glu Leu Thr Lys Gln Cys Val Glu Thr Asn Phe Tyr Gly Ala
145 150 155 160

```
Lys Arg Met Ser Glu Ala Phe Ile Pro Leu Leu Gln Leu Ser Asn Ser
                165                 170                 175

Pro Thr Ile Val Asn Val Ala Ser Phe Leu Gly Lys Leu Lys Leu Leu
            180                 185                 190

Cys Asn Glu Trp Ala Ile Lys Val Leu Ser Asn Ala Asn Asn Leu Thr
        195                 200                 205

Glu Asp Arg Val Asp Glu Val Val Asn Glu Phe Leu Lys Asp Phe Thr
    210                 215                 220

Glu Lys Ser Ile Glu Ala Lys Gly Trp Pro Thr Tyr Phe Ala Ala Tyr
225                 230                 235                 240

Lys Val Ser Lys Ala Ala Met Ile Ala Tyr Thr Arg Val Leu Ala Thr
                245                 250                 255

Lys Tyr Pro Asn Phe Arg Ile Asn Ser Val Cys Pro Gly Tyr Cys Lys
            260                 265                 270

Thr Asp Leu Thr Ala Asn Thr Gly Ser Leu Thr Ala Glu Glu Gly Ala
        275                 280                 285

Glu Ser Leu Val Lys Leu Ala Leu Leu Pro Asn Asp Gly Pro Ser Gly
    290                 295                 300

Leu Phe Phe Tyr Arg Lys Asp Val Ala Ala Leu
305                 310                 315
```

<210> 8
<211> 276
<212> PRT
<213> Atropa belladonna

<400> 8

```
                                        15
Met Ala Ser Val Ser Phe Leu Ser Thr Ile Gly Lys Arg Leu Glu Gly
1               5                   10                  15


Lys Val Ala Met Val Thr Gly Gly Ala Ser Gly Ile Gly Glu Ala Ile
            20                  25                  30


Ala Lys Leu Phe Tyr Glu His Gly Ala Lys Val Ala Ile Ala Asp Val
        35                  40                  45


Gln Asp Glu Leu Gly Asn Ser Val Ser Asn Ala Leu Gly Gly Ser Ser
    50                  55                  60


Asn Ser Ile Tyr Ile His Cys Asp Val Thr Asn Glu Asp Asp Val Gln
65                  70                  75                  80


Glu Ala Val Asp Lys Thr Ile Ser Thr Phe Gly Lys Leu Asp Ile Met
                85                  90                  95


Ile Cys Asn Ala Gly Ile Ser Asp Glu Thr Lys Pro Arg Ile Ile Asp
            100                 105                 110


Asn Thr Lys Ala Asp Phe Glu Arg Val Leu Ser Ile Asn Val Thr Gly
        115                 120                 125


Val Phe Leu Thr Met Lys His Ala Ala Arg Val Met Val Pro Ala Arg
    130                 135                 140


Ile Gly Cys Ile Ile Ser Thr Ser Ser Val Ser Ser Arg Val Gly Ala
145                 150                 155                 160


Ala Ala Ser His Ala Tyr Cys Ser Ser Lys His Ala Val Leu Gly Leu
                165                 170                 175


Thr Lys Asn Leu Ala Val Glu Leu Gly Gln Phe Gly Ile Arg Val Asn
            180                 185                 190


Cys Leu Ser Pro Tyr Ala Met Val Thr Pro Leu Ala Glu Lys Val Ile
```

195 200 205

Gly Leu Glu Asn Glu Glu Leu Glu Lys Ala Leu Asp Met Val Gly Asn
210 215 220

Leu Lys Gly Val Thr Leu Arg Val Asp Asp Val Ala Lys Ala Ala Leu
225 230 235 240

Phe Leu Ala Ser Asp Asp Ser Lys Tyr Ile Ser Gly His Asn Leu Phe
245 250 255

Ile Asp Gly Gly Phe Thr Val Tyr Asn Pro Gly Leu Gly Met Phe Lys
260 265 270

Tyr Pro Glu Ser
275

<210> 9
<211> 297
<212> PRT
<213> Atropa belladonna

<400> 9

Met Leu Arg Ile Ala Ser Arg Gly Gly Ile Thr Ser Arg Ser Leu Gln
1 5 10 15

Leu Leu Gln Thr Phe Asn Lys Glu Phe Ser Thr His Ile Glu Arg Lys
20 25 30

Leu Glu Gly Lys Val Ala Leu Ile Thr Gly Ala Ala Ser Gly Ile Gly
35 40 45

Lys Glu Thr Ala Ala Lys Phe Ile Asn Asn Gly Ala Lys Val Ile Ile
50 55 60

Ala Asp Val Gln Lys Gln Leu Gly Gln Glu Thr Ala Ser Gln Leu Gly
65 70 75 80

```
Pro Asn Ala Thr Phe Val Leu Cys Asp Val Thr Lys Glu Ser Asp Val
                85                  90                  95

Ser Asn Ala Val Asp Phe Ala Val Ser Asn His Gly Gln Leu Asp Ile
            100                 105                 110

Met Tyr Asn Asn Ala Gly Ile Ile Cys Arg Thr Pro Arg Asn Ile Ala
        115                 120                 125

Asp Leu Asp Leu Asp Ala Phe Asp Arg Val Met Ala Ile Asn Val Arg
    130                 135                 140

Gly Met Met Ala Gly Ile Lys His Ala Ala Arg Val Met Ile Pro Arg
145                 150                 155                 160

Lys Ala Gly Ser Ile Leu Cys Thr Ala Ser Ile Thr Gly Thr Met Gly
                165                 170                 175

Gly Leu Ala Gln Pro Thr Tyr Ser Thr Thr Lys Ser Cys Val Ile Gly
            180                 185                 190

Met Met Arg Ser Val Thr Ala Glu Leu Cys Gln Asn Gly Ile Arg Ile
        195                 200                 205

Asn Cys Ile Ser Pro Phe Ala Ile Pro Thr Pro Phe Tyr Ile Asp Glu
    210                 215                 220

Met Lys Ser Tyr Tyr Pro Gly Val Glu Pro Glu Val Leu Val Lys Met
225                 230                 235                 240

Leu Tyr Arg Ala Ser Glu Leu Asn Gly Ala Tyr Cys Glu Pro Val Asp
                245                 250                 255

Val Ala Asn Ala Ala Val Phe Leu Ala Ser Asp Asp Ala Lys Tyr Val
            260                 265                 270
```

Ser Gly Gln Asn Leu Val Ile Asp Gly Gly Phe Thr Ser Tyr Lys Ser
275 280 285

Leu Asn Phe Pro Met Ser Asp Gln Glu
290 295

<210> 10
<211> 279
<212> PRT
<213> Atropa belladonna

<400> 10

Met Gly Ile Pro Ser Ser Val Thr Pro Ile Val Arg Arg Leu Glu Gly
1 5 10 15

Lys Val Ala Val Ile Thr Gly Gly Ala Ser Gly Ile Gly Glu Ala Ala
20 25 30

Thr Arg Leu Phe Val Lys His Gly Ala Lys Val Val Val Ala Asp Val
35 40 45

Arg Asp Asp Leu Gly Arg Ala Leu Cys Lys Glu Leu Gly Ser Asn Asp
50 55 60

Thr Ile Ser Phe Ala His Cys Ser Val Thr Asp Glu Asn Asp Val Gln
65 70 75 80

Asn Ala Ile Asp Gly Ala Val Ser Arg Tyr Gly Met Leu Asp Ile Met
85 90 95

Phe Asn Asn Ala Gly Ile Thr Gly Asn Met Lys Asp Pro Ser Ile Leu
100 105 110

Ala Thr Asp Tyr Lys Asn Phe Lys Asn Val Phe Asp Val Asn Val Tyr
115 120 125

Gly Ala Phe Leu Gly Ala Arg Ile Ala Ala Lys Ala Met Ile Pro Thr
130 135 140

```
Lys Gln Gly Ser Ile Leu Phe Thr Ala Ser Ile Ala Ser Val Ile Gly
145                 150                 155                 160

Gly Ile Ala Ser Pro Ile Thr Tyr Ala Ser Ser Lys His Ala Val Val
                165                 170                 175

Gly Leu Thr Asn His Leu Ala Val Glu Leu Gly Gln Tyr Gly Ile Arg
            180                 185                 190

Val Asn Cys Ile Ser Pro Tyr Thr Val Ala Thr Pro Leu Val Arg Glu
        195                 200                 205

Ile Leu Gly Lys Met Asp Lys Glu Lys Ala Glu Glu Val Ile Met Glu
    210                 215                 220

Thr Ala Asn Leu Lys Gly Lys Ile Leu Glu Pro Glu Asp Ile Ala Glu
225                 230                 235                 240

Ala Ala Val Tyr Leu Gly Ser Asp Glu Ser Lys Tyr Val Ser Gly Ile
                245                 250                 255

Asn Leu Val Ile Asp Gly Gly Tyr Ser Lys Thr Asn Pro Leu Ala Ser
            260                 265                 270

Met Val Met Gln Asn Tyr Ile
        275
```

<210> 11
<211> 327
<212> PRT
<213> Atropa belladonna

<400> 11

```
Met Glu Ser Lys Ser Gly Glu Gly Lys Ile Val Cys Val Thr Gly Ala
1               5                   10                  15
```

```
                                        20
Ser Gly Phe Ile Ala Ser Trp Leu Val Lys Leu Leu Leu His Arg Gly
            20                  25                  30

Tyr Thr Val Asn Ala Thr Val Arg Asn Leu Lys Asp Thr Ser Lys Val
        35                  40                  45

Ala His Leu Leu Gly Leu Asp Gly Ala Asn Glu Arg Leu His Leu Phe
    50                  55                  60

Lys Ala Glu Leu Leu Glu Glu Gln Ser Phe Asp Ala Ala Val Asp Gly
65                  70                  75                  80

Cys Glu Gly Val Phe His Thr Ala Ser Pro Val Ser Leu Thr Ala Lys
                85                  90                  95

Ser Lys Glu Glu Leu Val Asp Pro Ala Val Lys Gly Thr Leu Asn Val
            100                 105                 110

Leu Arg Ser Cys Ala Lys Ser Pro Ser Val Leu Arg Val Val Ile Thr
        115                 120                 125

Ser Ser Thr Ala Ser Val Ile Cys Asn Lys Asn Met Ser Thr Pro Gly
    130                 135                 140

Ala Val Ala Asp Glu Thr Trp Tyr Ser Asp Pro Glu Phe Cys Glu Glu
145                 150                 155                 160

Arg Glu Glu Trp Tyr Gln Leu Ser Lys Thr Leu Ala Glu Gln Ala Ala
                165                 170                 175

Trp Lys Phe Ala Lys Glu Asn Glu Met Asp Leu Val Thr Leu His Pro
            180                 185                 190

Gly Leu Val Ile Gly Pro Leu Leu Gln Pro Thr Leu Asn Phe Ser Cys
        195                 200                 205

Glu Ala Ile Val Asn Phe Ile Lys Glu Gly Lys Glu Ala Trp Ser Gly
```

210 215 220

Gly Val Tyr Arg Phe Val Asp Val Arg Asp Val Ala Asn Ala His Ile
225 230 235 240

Leu Ala Phe Glu Val Pro Ser Ala Asn Gly Arg Tyr Cys Leu Val Gly
245 250 255

Val Asn Gly Tyr Ser Ser Leu Val Leu Lys Ile Val Gln Lys Leu Tyr
260 265 270

Pro Ser Ile Thr Leu Pro Glu Asn Phe Glu Asp Gly Leu Pro Leu Thr
275 280 285

Pro His Phe Gln Val Ser Ser Glu Arg Ala Lys Gly Leu Gly Val Lys
290 295 300

Phe Thr Pro Leu Glu Leu Ser Val Lys Asp Thr Val Glu Ser Leu Met
305 310 315 320

Glu Lys Asn Phe Leu His Ile
325

<210> 12
<211> 327
<212> PRT
<213> Atropa belladonna

<400> 12

Met Glu Ser Lys Ser Gly Glu Gly Lys Ile Val Cys Val Thr Gly Ala
1 5 10 15

Ser Gly Phe Ile Ala Ser Trp Leu Val Lys Leu Leu Leu His Arg Gly
20 25 30

Tyr Thr Val Asn Ala Thr Val Arg Asn Leu Lys Asp Thr Ser Lys Val
35 40 45

```
Ala His Leu Leu Gly Leu Asp Gly Ala Asn Glu Arg Leu His Leu Phe
    50                  55                  60

Lys Ala Glu Leu Leu Glu Glu Gln Ser Phe Asp Ala Ala Val Asp Gly
65                  70                  75                  80

Cys Glu Gly Val Phe His Thr Ala Ser Pro Val Ser Leu Thr Ala Lys
                85                  90                  95

Ser Lys Glu Glu Leu Val Asp Pro Ala Val Lys Gly Thr Leu Asn Val
            100                 105                 110

Leu Arg Ser Cys Ala Lys Ser Pro Ser Val Leu Arg Val Val Ile Thr
        115                 120                 125

Ser Ser Thr Ala Ser Val Ile Cys Asn Lys Asn Met Ser Thr Pro Gly
    130                 135                 140

Ala Val Ala Asp Glu Thr Trp Tyr Ser Asp Pro Glu Phe Cys Glu Glu
145                 150                 155                 160

Arg Lys Glu Trp Tyr Gln Leu Ser Lys Thr Leu Ala Glu Lys Ala Ala
                165                 170                 175

Arg Arg Phe Ala Lys Glu Asn Gly Ile Asp Leu Val Thr Leu His Pro
            180                 185                 190

Gly Leu Val Ile Gly Pro Leu Leu Gln Pro Thr Leu Asn Phe Ser Cys
        195                 200                 205

Glu Ala Ile Val Asn Phe Ile Lys Glu Gly Lys Glu Ala Trp Ser Gly
    210                 215                 220

Gly Val Tyr Arg Phe Val Asp Val Arg Asp Val Ala Asn Ala His Ile
225                 230                 235                 240
```

Leu Ala Phe Glu Val Pro Ser Ala Asn Gly Arg Tyr Cys Leu Val Gly
245 250 255

Val Asn Gly Tyr Ser Ser Leu Val Leu Lys Ile Val Gln Lys Leu Tyr
260 265 270

Pro Ser Ile Thr Leu Pro Glu Asn Phe Glu Asp Gly Leu Pro Leu Thr
275 280 285

Pro His Phe Gln Val Ser Ser Glu Arg Ala Lys Gly Leu Gly Val Lys
290 295 300

Phe Thr Pro Leu Glu Leu Ser Val Lys Asp Thr Val Glu Ser Leu Met
305 310 315 320

Glu Lys Asn Phe Leu His Ile
325

<210> 13
<211> 365
<212> PRT
<213> Atropa belladonna

<400> 13

Met Ala Ser Glu Lys Ser Leu Glu Glu Lys Gln Ala Glu Asn Thr Phe
1 5 10 15

Gly Trp Ala Ala Met Asp Ser Ser Gly Val Leu Ser Pro Phe Thr Phe
20 25 30

Ser Arg Arg Ala Thr Gly Glu Glu Asp Val Arg Leu Lys Val Leu Tyr
35 40 45

Cys Gly Ile Cys His Ser Asp Leu Gly Cys Ile Lys Asn Glu Trp Gly
50 55 60

Trp Cys Ser Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Ile Ala
65 70 75 80

```
Thr Glu Val Gly Ser Lys Val Thr Lys Phe Lys Val Gly Asp Arg Val
                85                  90                  95

Gly Val Gly Cys Met Val Gly Ser Cys Gly Thr Cys Gln Asn Cys Thr
            100                 105                 110

Gln Asn Gln Glu Ser Tyr Cys Pro Glu Val Ile Met Thr Cys Ala Ser
        115                 120                 125

Ala Tyr Pro Asp Gly Thr Pro Thr Tyr Gly Gly Phe Ser Asn Gln Met
    130                 135                 140

Val Ala Asn Glu Lys Phe Val Ile Arg Ile Pro Asn Ser Leu Pro Leu
145                 150                 155                 160

Asp Ala Ala Ala Pro Leu Leu Cys Ala Gly Ser Thr Val Tyr Ser Ala
                165                 170                 175

Met Lys Phe Tyr Gly Leu Cys Ser Gln Gly Leu His Leu Gly Val Val
            180                 185                 190

Gly Leu Gly Gly Leu Gly His Val Ala Val Lys Phe Ala Lys Ala Phe
        195                 200                 205

Gly Met Lys Val Thr Val Ile Ser Thr Ser Leu Gly Lys Lys Glu Glu
    210                 215                 220

Ala Ile Asn Gln Leu Gly Ala Asp Ser Phe Leu Ile Asn Thr Asp Thr
225                 230                 235                 240

Glu Gln Met Gln Gly Ala Met Glu Val Met Asp Gly Ile Ile Asp Thr
                245                 250                 255

Val Ser Ala Leu His Pro Ile Glu Pro Leu Leu Gly Leu Leu Lys Ser
            260                 265                 270
```

```
His Gln Gly Lys Leu Ile Ile Val Gly Leu Pro Asn Lys Gln Pro Glu
        275                 280                 285


Leu Pro Val Phe Ser Leu Ile Asn Gly Arg Lys Met Ile Gly Gly Ser
    290                 295                 300


Ala Val Gly Gly Val Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala
305                 310                 315                 320


Glu His Asn Ile Thr Ala Asp Ile Glu Ile Val Pro Met Asp Tyr Val
                325                 330                 335


Asn Thr Ala Met Glu Arg Leu Glu Lys Gly Asp Val Lys Phe Arg Phe
            340                 345                 350


Val Ile Asp Val Glu Asn Thr Leu Val Ala Ala Gln Thr
        355                 360                 365
```

<210> 14
<211> 364
<212> PRT
<213> Datura innoxia

<400> 14

```
Met Ala Ala Glu Lys Leu Ser Glu Glu Glu Ala Val Lys Thr Phe Gly
1               5                   10                  15


Trp Ala Ala Met Asp Ser Ser Gly Val Leu Ser Pro Phe Glu Phe Ser
            20                  25                  30


Arg Arg Ala Thr Gly Ala Glu Asp Val Arg Leu Lys Val Leu Tyr Cys
        35                  40                  45


Gly Ile Cys His Ser Asp Leu Gly Cys Val Lys Asn Glu Trp Gly Trp
    50                  55                  60


Cys Ser Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Ile Ala Thr
```

65 70 75 80

Glu Val Gly Ser Arg Val Thr Lys Phe Lys Val Gly Asp Arg Val Gly
85 90 95

Val Gly Cys Met Val Gly Ser Cys Gly Ser Cys Gln Asn Cys Ser Gln
100 105 110

Asn Leu Glu Ser Tyr Cys Pro Glu Val Ile Met Thr Cys Ala Ser Ala
115 120 125

Tyr Pro Asp Gly Thr Pro Thr Tyr Gly Gly Phe Ser Asn Gln Met Val
130 135 140

Ala Asn Glu Lys Phe Val Ile Gln Ile Pro Glu Lys Leu Pro Leu Asp
145 150 155 160

Ala Ala Ala Pro Leu Leu Cys Ala Gly Ser Thr Val Tyr Ser Pro Met
165 170 175

Lys Phe Tyr Gly Leu Cys Ser Pro Gly Leu His Leu Gly Val Val Gly
180 185 190

Leu Gly Gly Leu Gly His Val Ala Val Lys Phe Ala Lys Ala Phe Gly
195 200 205

Met Lys Val Thr Val Ile Ser Thr Ser Ile Gly Lys Lys Glu Glu Ala
210 215 220

Ile Asn Gln Leu Gly Ala Asp Ser Phe Leu Thr Ser Thr Asp Thr Glu
225 230 235 240

Gln Met Gln Gly Ala Met Glu Thr Met Asp Gly Ile Ile Asp Thr Val
245 250 255

Ser Ala Leu His Pro Ile Glu Pro Leu Val Gly Leu Leu Lys Ser His
260 265 270

```
Gln Gly Lys Leu Ile Ile Val Gly Leu Pro Asn Lys Gln Pro Glu Leu
        275                 280                 285

Pro Val Phe Ser Leu Ile Asn Gly Arg Lys Met Ile Gly Gly Ser Ala
    290                 295                 300

Val Gly Gly Val Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys
305                 310                 315                 320

His Asn Ile Thr Ala Asp Ile Glu Ile Val Arg Met Asp Tyr Val Asn
                325                 330                 335

Thr Ala Met Glu Arg Leu Glu Lys Gly Asp Val Lys Phe Arg Phe Val
            340                 345                 350

Ile Asp Val Glu Asn Thr Leu Val Pro Ala Gln Thr
        355                 360
```

<210> 15
<211> 364
<212> PRT
<213> Datura stramonium

<400> 15

```
Met Ala Ala Glu Lys Leu Glu Glu Arg Lys Arg Trp Glu Thr Phe Gly
1               5                   10                  15

Trp Ala Ala Met Asp Ser Ser Gly Val Leu Ser Pro Phe Glu Phe Ser
            20                  25                  30

Arg Arg Ala Thr Gly Glu Glu Asp Val Arg Leu Lys Val Leu Tyr Cys
        35                  40                  45

Gly Ile Cys His Ser Asp Leu Gly Cys Ile Lys Asn Glu Trp Gly Trp
    50                  55                  60
```

```
Cys Ser Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Ile Ala Thr
65                  70                  75                  80

Glu Val Gly Ser Arg Val Thr Lys Phe Lys Val Gly Asp Arg Val Gly
                85                  90                  95

Val Gly Cys Met Val Gly Ser Cys Gly Ser Cys Gln Asn Cys Ser Gln
            100                 105                 110

Asn Leu Glu Ser Tyr Cys Pro Glu Val Ile Met Thr Cys Ala Ser Ala
        115                 120                 125

Tyr Pro Asp Gly Thr Pro Thr Tyr Gly Gly Phe Ser Asn Gln Met Val
    130                 135                 140

Ala Asn Glu Lys Phe Val Ile Gln Ile Pro Glu Lys Leu Pro Leu Asp
145                 150                 155                 160

Ala Ala Ala Pro Leu Leu Cys Ala Gly Ser Thr Val Tyr Ser Pro Met
                165                 170                 175

Lys Phe Tyr Gly Leu Cys Ser Pro Gly Leu His Leu Gly Val Val Gly
            180                 185                 190

Leu Gly Gly Leu Gly His Val Ala Val Lys Phe Ala Lys Ala Phe Gly
        195                 200                 205

Met Lys Val Thr Val Ile Ser Thr Ser Ile Gly Lys Lys Glu Glu Ala
    210                 215                 220

Ile Asn Gln Leu Gly Ala Asp Ser Phe Leu Ile Ser Thr Asp Thr Glu
225                 230                 235                 240

Gln Met Gln Gly Ala Met Glu Thr Met Asp Gly Ile Ile Asp Thr Val
                245                 250                 255

Ser Ala Leu His Pro Ile Glu Pro Leu Val Gly Leu Leu Lys Ser His
```

260 265 270

Arg Gly Lys Leu Ile Ile Val Gly Leu Pro Asn Lys Gln Pro Glu Leu
275 280 285

Pro Val Phe Ser Leu Ile Asn Gly Arg Lys Met Ile Gly Gly Ser Ala
290 295 300

Val Gly Gly Val Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys
305 310 315 320

His Asn Ile Thr Ala Asp Ile Glu Ile Val Gly Met Asp Tyr Val Asn
325 330 335

Thr Ala Met Glu Arg Leu Glu Lys Gly Asp Val Lys Phe Arg Phe Val
340 345 350

Ile Asp Val Glu Asn Thr Leu Val Pro Ala Gln Thr
355 360

<210> 16
<211> 238
<212> PRT
<213> Artificial sequence

<220>
<223> Green Fluorescent Protein

<400> 16

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35 40 45

30
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
50 55 60

Gly Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Gln
65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85 90 95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
210 215 220

Thr Ala Ala Gly Ile Ile His Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 17

<211> 233
<212> PRT
<213> Artificial sequence

<220>
<223> Blue Fluorescent Protein

<400> 17

Met Ser Glu Leu Ile Lys Glu Asn Met His Met Lys Leu Tyr Met Glu
1 5 10 15

Gly Thr Val Asp Asn His His Phe Lys Cys Thr Ser Glu Gly Glu Gly
20 25 30

Lys Pro Tyr Glu Gly Thr Gln Thr Met Arg Ile Lys Val Val Glu Gly
35 40 45

Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Thr Ser Phe Leu Tyr
50 55 60

Gly Ser Lys Thr Phe Ile Asn His Thr Gln Gly Ile Pro Asp Phe Phe
65 70 75 80

Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Val Thr Thr Tyr
85 90 95

Glu Asp Gly Gly Val Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp
100 105 110

Gly Cys Leu Ile Tyr Asn Val Lys Ile Arg Gly Val Asn Phe Thr Ser
115 120 125

Asn Gly Pro Val Met Gln Lys Lys Thr Leu Gly Trp Glu Ala Phe Thr
130 135 140

Glu Thr Leu Tyr Pro Ala Asp Gly Gly Leu Glu Gly Arg Asn Asp Met
145 150 155 160

Ala Leu Lys Leu Val Gly Gly Ser His Leu Ile Ala Asn Ile Lys Thr
165 170 175

Thr Tyr Arg Ser Lys Lys Pro Ala Lys Asn Leu Lys Met Pro Gly Val
180 185 190

Tyr Tyr Val Asp Tyr Arg Leu Glu Arg Ile Lys Glu Ala Asn Asn Glu
195 200 205

Thr Tyr Val Glu Gln His Glu Val Ala Val Ala Arg Tyr Cys Asp Leu
210 215 220

Pro Ser Lys Leu Gly His Lys Leu Asn
225 230

<210> 18
<211> 238
<212> PRT
<213> Artificial sequence

<220>
<223> Yellow Fluorescent Protein

<400> 18

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Leu Ile Cys
35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50 55 60

Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Gln
65 70 75 80

```
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                85                  90                  95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
            100                 105                 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
    130                 135                 140

Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160

Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Gly Val
                165                 170                 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Lys Leu Ser
        195                 200                 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
    210                 215                 220

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 19
<211> 98
<212> PRT
<213> Artificial sequence

<220>
<223> Small Ubiquitin-like Modifier

<400> 19

Met Ser Asp Ser Glu Val Asn Gln Glu Ala Lys Pro Glu Val Lys Pro
1 5 10 15

Glu Val Lys Pro Glu Thr His Ile Asn Leu Lys Val Ser Asp Gly Ser
20 25 30

Ser Glu Ile Phe Phe Lys Ile Lys Lys Thr Thr Pro Leu Arg Arg Leu
35 40 45

Met Glu Ala Phe Ala Lys Arg Gln Gly Lys Glu Met Asp Ser Leu Thr
50 55 60

Phe Leu Tyr Asp Gly Ile Glu Ile Gln Ala Asp Gln Thr Pro Glu Asp
65 70 75 80

Leu Asp Met Glu Asp Asn Asp Ile Ile Glu Ala His Arg Glu Gln Ile
85 90 95

Gly Gly

<210> 20
<211> 236
<212> PRT
<213> Artificial sequence

<220>
<223> Red Fluorescent Protein, mCherry

<400> 20

Met Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe
1 5 10 15

Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe
20 25 30

Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr

```
                                        35
        35                  40                  45

Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
    50                  55                  60

Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His
65                  70                  75                  80

Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe
                85                  90                  95

Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val
            100                 105                 110

Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys
        115                 120                 125

Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys
    130                 135                 140

Thr Met Gly Trp Glu Ala Ser Ser Glu Arg Met Tyr Pro Glu Asp Gly
145                 150                 155                 160

Ala Leu Lys Gly Glu Ile Lys Gln Arg Leu Lys Leu Lys Asp Gly Gly
                165                 170                 175

His Tyr Asp Ala Glu Val Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val
            180                 185                 190

Gln Leu Pro Gly Ala Tyr Asn Val Asn Ile Lys Leu Asp Ile Thr Ser
        195                 200                 205

His Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly
    210                 215                 220

Arg His Ser Thr Gly Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 21
<211> 490
<212> PRT
<213> Atropa belladonna

<400> 21

```
Met Gly Ser Gln Gly Thr Asn Ile Asp Ser Ile Ile His Val Phe Leu
1               5                   10                  15


Ile Ser Phe Pro Gly Gln Gly His Val Asn Pro Leu Leu Arg Leu Gly
            20                  25                  30


Lys Arg Leu Ala Ser Lys Gly Val Leu Val Ser Phe Cys Ala Pro Glu
        35                  40                  45


Cys Val Gly Lys Asp Met Arg Ala Ala Asn Asn Asn Ile Ile Ser Asp
    50                  55                  60


Glu Pro Thr Pro Tyr Gly Asp Gly Phe Ile Arg Phe Glu Phe Phe Asp
65                  70                  75                  80


Gly Trp Glu Tyr Thr Gln Pro Lys Glu Asn Arg Gln Leu Glu Ile Glu
                85                  90                  95


Leu Ala Asn Leu Glu Val Val Gly Arg Ala Val Leu Pro Ala Met Leu
            100                 105                 110


Lys Glu Asn Glu Ala Lys Gly Arg Pro Val Ser Cys Leu Ile Asn Asn
        115                 120                 125


Pro Phe Ile Pro Trp Val Cys Asp Val Ala Asp Ser Leu Gly Ile Pro
    130                 135                 140


Cys Ala Val Leu Trp Val Gln Ser Cys Ala Ser Phe Ser Ala Tyr Tyr
145                 150                 155                 160
```

```
His Tyr His Phe Asn Leu Ala Pro Phe Pro Asn Glu Ser Asn Pro Asn
                165                 170                 175

Ile Asp Val His Leu Pro Asn Met Pro Ile Leu Lys Trp Asp Glu Leu
            180                 185                 190

Pro Ser Phe Leu Leu Pro Ser Asn Pro Tyr Pro Ala Leu Ala Asn Ala
        195                 200                 205

Ile Leu Arg Gln Phe Asn Tyr Leu Ser Lys Pro Ile Arg Ile Phe Ile
    210                 215                 220

Glu Ser Phe Asp Glu Leu Glu Lys Asp Ile Val Asp Tyr Met Ser Asp
225                 230                 235                 240

Phe Leu Pro Ile Lys Thr Val Gly Pro Leu Leu Val Glu Asp Pro Lys
                245                 250                 255

Ile Glu Gln Val Val Arg Ala Asp Leu Val Lys Ala Asp Ser Ser Ile
            260                 265                 270

Thr Gln Trp Leu Asn Ser Lys Pro Pro Ser Ser Val Val Tyr Ile Ser
        275                 280                 285

Phe Gly Ser Ile Val Val Pro Ser Gln Glu Gln Val Asp Glu Ile Ala
    290                 295                 300

Tyr Gly Ile Leu Asn Ser Gly Leu Asn Phe Leu Trp Ile Met Lys Pro
305                 310                 315                 320

Pro Arg Lys Asn Ser Ser Phe Pro Thr Val Val Leu Pro Gln Gly Tyr
                325                 330                 335

Leu Asp Lys Ile Gly Asp Lys Gly Lys Val Val Glu Trp Cys Leu Gln
            340                 345                 350

Glu Gln Val Leu Ala His Pro Ser Leu Ala Cys Phe Val Thr His Cys
```

355 360 365

Gly Trp Asn Ser Ser Met Glu Val Ile Ala Asn Gly Val Pro Ile Val
370 375 380

Ala Phe Pro Gln Trp Gly Asp Gln Val Thr Asp Ala Lys Tyr Leu Val
385 390 395 400

Asp Glu Phe Lys Ile Gly Val Arg Leu Ser Arg Gly Val Thr Glu Asn
405 410 415

Arg Val Ile Pro Arg Asp Glu Val Glu Arg Ser Leu His Asp Val Thr
420 425 430

Ser Gly Pro Lys Val Ala Glu Met Lys Glu Asn Ala Leu Lys Trp Lys
435 440 445

Met Lys Ala Thr Glu Ala Val Ala Glu Gly Gly Ser Ser Asp Leu Asn
450 455 460

Leu Lys Ser Phe Val Asp Glu Leu Arg Thr Leu Gln Asn Ser Asn Lys
465 470 475 480

Asn Leu Ala Lys Leu Ala Pro Leu Ser Asn
485 490

<210> 22
<211> 225
<212> PRT
<213> Artificial sequence

<220>
<223> Red Fluorescent Protein, DsRed

<400> 22

Met Ala Ser Ser Glu Asp Val Ile Lys Glu Phe Met Arg Phe Lys Val
1 5 10 15

```
Arg Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
            20                  25                  30

Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys Val
        35                  40                  45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
    50                  55                  60

Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65                  70                  75                  80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
                85                  90                  95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
            100                 105                 110

Leu Gln Asp Gly Arg Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
        115                 120                 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
    130                 135                 140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145                 150                 155                 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
                165                 170                 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
            180                 185                 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
        195                 200                 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
```

40
210 215 220

Leu
225

<210> 23
<211> 470
<212> PRT
<213> Nicotiana tabacum

<400> 23

Met Val Glu Glu Leu Pro Gln Ser Leu Lys Glu Lys Lys Trp Gln Ile
1 5 10 15

Asn Trp Asp Ala Val Ser Gln Glu Leu Lys Lys Thr Ser Arg Phe Met
20 25 30

Ala Pro Met Val Ala Val Thr Val Phe Gln Tyr Leu Leu Gln Val Val
35 40 45

Ser Val Met Met Val Gly His Leu Gly Glu Leu Ala Leu Ser Ser Val
50 55 60

Ala Ile Ala Thr Ser Leu Thr Asn Val Thr Gly Phe Ser Leu Leu Thr
65 70 75 80

Gly Leu Val Gly Gly Met Glu Thr Leu Cys Gly Gln Ala Tyr Gly Ala
85 90 95

Gln Gln Tyr His Lys Leu Ser Thr Tyr Thr Tyr Thr Ala Ile Ile Ser
100 105 110

Leu Phe Leu Val Cys Ile Pro Ile Cys Val Leu Trp Cys Phe Met Asp
115 120 125

Lys Leu Leu Ile Leu Thr Gly Gln Asp His Ser Ile Ser Val Glu Ala
130 135 140

```
Arg Lys Tyr Ser Leu Trp Val Ile Pro Ala Ile Phe Gly Gly Ala Ile
145                 150                 155                 160

Ser Lys Pro Leu Ser Arg Tyr Ser Gln Ala Gln Ser Leu Ile Leu Pro
                165                 170                 175

Met Leu Leu Ser Ser Phe Ala Val Leu Cys Phe His Leu Pro Ile Ser
            180                 185                 190

Trp Ala Leu Ile Phe Lys Leu Glu Leu Gly Asn Ile Gly Ala Ala Ile
        195                 200                 205

Ala Phe Ser Ile Ser Ser Trp Leu Tyr Val Leu Phe Leu Ala Ser Tyr
    210                 215                 220

Val Lys Leu Ser Ser Ser Cys Glu Lys Thr Arg Ala Pro Phe Ser Met
225                 230                 235                 240

Glu Ala Phe Leu Cys Ile Arg Gln Phe Phe Arg Leu Ala Val Pro Ser
                245                 250                 255

Ala Val Met Val Cys Leu Lys Trp Trp Ser Phe Glu Val Leu Ala Leu
            260                 265                 270

Val Ser Gly Leu Leu Pro Asn Pro Lys Leu Glu Thr Ser Val Met Ser
        275                 280                 285

Ile Cys Ile Thr Ile Ser Gln Leu His Phe Ser Ile Pro Tyr Gly Phe
    290                 295                 300

Gly Ala Ala Ala Ser Thr Arg Val Ser Asn Glu Leu Gly Ala Gly Asn
305                 310                 315                 320

Pro Gln Lys Ala Arg Met Ala Val Gln Val Val Met Phe Leu Thr Val
                325                 330                 335
```

```
Val Glu Thr Leu Val Phe Asn Thr Ser Leu Phe Gly Ser Arg His Val
            340                 345                 350


Leu Gly Lys Ala Phe Ser Asn Glu Lys Gln Val Val Asp Tyr Ile Ala
        355                 360                 365


Ala Met Thr Pro Phe Leu Cys Leu Ser Ile Val Thr Asp Ser Leu Gln
    370                 375                 380


Ile Val Ile Thr Gly Ile Ala Arg Gly Ser Gly Trp Gln His Ile Gly
385                 390                 395                 400


Ala Tyr Ile Asn Leu Val Val Phe Tyr Val Ile Ala Ile Pro Leu Ala
                405                 410                 415


Val Val Leu Gly Phe Val Leu His Leu Lys Ala Lys Gly Leu Trp Ile
            420                 425                 430


Gly Ile Val Val Gly Cys Ala Ile Gln Ser Ile Val Leu Ser Ile Val
        435                 440                 445


Thr Gly Phe Thr Asp Trp Glu Lys Gln Ala Lys Lys Ala Arg Glu Arg
    450                 455                 460


Val His Glu Gly Arg Ser
465                 470


<210>  24
<211>  500
<212>  PRT
<213>  Nicotiana tabacum

<400>  24

Met Gly Lys Ser Met Lys Ser Glu Val Glu Gln Pro Leu Leu Ile Ala
1               5                   10                  15


Ala His Gly Gly Ser Ser Glu Leu Glu Glu Val Leu Ser Asp Thr Gln
            20                  25                  30
```

```
Leu Pro Tyr Phe Arg Arg Leu Arg Tyr Ala Ser Trp Ile Glu Phe Gln
        35                  40                  45

Leu Leu Tyr Arg Leu Ala Ala Pro Ser Val Ala Val Tyr Met Ile Asn
    50                  55                  60

Asn Ala Met Ser Met Ser Thr Arg Ile Phe Ser Gly Gln Leu Gly Asn
65                  70                  75                  80

Leu Gln Leu Ala Ala Ala Ser Leu Gly Asn Gln Gly Ile Gln Leu Phe
                85                  90                  95

Ala Tyr Gly Leu Met Leu Gly Met Gly Ser Ala Val Glu Thr Leu Cys
            100                 105                 110

Gly Gln Ala Tyr Gly Ala His Arg Tyr Glu Met Leu Gly Val Tyr Leu
        115                 120                 125

Gln Arg Ala Thr Val Val Leu Ser Val Thr Gly Ile Pro Leu Thr Val
    130                 135                 140

Val Tyr Leu Phe Ser Lys Asn Ile Leu Leu Ala Leu Gly Glu Ser Lys
145                 150                 155                 160

Leu Val Ala Ser Ala Ala Ala Val Phe Val Tyr Gly Leu Ile Pro Gln
                165                 170                 175

Ile Phe Ala Tyr Ala Val Asn Phe Pro Ile Gln Lys Phe Leu Gln Ala
            180                 185                 190

Gln Ser Ile Val Ala Pro Ser Ala Phe Ile Ser Leu Gly Thr Leu Phe
        195                 200                 205

Val His Ile Leu Leu Ser Trp Val Val Val Tyr Lys Ile Gly Leu Gly
    210                 215                 220
```

```
Leu Leu Gly Ala Ser Leu Val Leu Ser Phe Ser Trp Trp Ile Ile Val
225                 230                 235                 240

Val Ala Gln Phe Ile Tyr Ile Ile Lys Ser Glu Arg Cys Lys Ala Thr
                245                 250                 255

Trp Ala Gly Phe Arg Trp Glu Ala Phe Ser Gly Leu Cys Gln Phe Val
            260                 265                 270

Lys Leu Ser Ala Gly Ser Ala Val Met Leu Cys Leu Glu Thr Trp Tyr
        275                 280                 285

Met Gln Ile Leu Val Leu Leu Ser Gly Leu Leu Lys Asn Pro Glu Ile
    290                 295                 300

Ala Leu Ala Ser Ile Ser Val Cys Leu Ala Val Asn Gly Leu Met Phe
305                 310                 315                 320

Met Val Ala Val Gly Phe Asn Ala Ala Ala Ser Val Arg Val Ser Asn
                325                 330                 335

Glu Leu Gly Ala Ala His Ser Lys Ser Ala Ala Phe Ser Val Phe Met
            340                 345                 350

Val Thr Phe Ile Ser Phe Leu Ile Ala Val Val Glu Ala Ile Ile Val
        355                 360                 365

Leu Ser Leu Arg Asn Val Ile Ser Tyr Ala Phe Thr Glu Gly Glu Ile
    370                 375                 380

Val Ala Lys Glu Val Ser Glu Leu Cys Pro Phe Leu Ala Val Thr Leu
385                 390                 395                 400

Ile Leu Asn Gly Ile Gln Pro Val Leu Ser Gly Val Ala Val Gly Cys
                405                 410                 415
```

Gly Trp Gln Ala Phe Val Ala Tyr Val Asn Val Gly Cys Tyr Tyr Gly
420 425 430

Val Gly Ile Pro Leu Gly Cys Leu Leu Gly Phe Lys Phe Asp Leu Gly
435 440 445

Ala Lys Gly Ile Trp Thr Gly Met Ile Gly Gly Thr Val Met Gln Thr
450 455 460

Val Ile Leu Leu Trp Val Thr Phe Arg Thr Asp Trp Asn Lys Lys Val
465 470 475 480

Glu Cys Ala Lys Lys Arg Leu Asp Lys Trp Glu Asn Leu Lys Gly Pro
485 490 495

Leu Asn Lys Glu
500

<210> 25
<211> 500
<212> PRT
<213> Nicotiana tabacum

<400> 25

Met Gly Lys Ser Met Lys Ser Glu Val Glu Gln Pro Leu Leu Ala Ala
1 5 10 15

Ala His Gly Gly Ser Ser Glu Leu Glu Glu Val Leu Ser Asp Ser Gln
20 25 30

Leu Pro Tyr Phe Arg Arg Leu Arg Tyr Ala Ser Trp Ile Glu Phe Gln
35 40 45

Leu Leu Tyr Arg Leu Ala Ala Pro Ser Val Ala Val Tyr Met Ile Asn
50 55 60

Asn Ala Met Ser Met Ser Thr Arg Ile Phe Ser Gly Gln Leu Gly Asn
65 70 75 80

```
Leu Gln Leu Ala Ala Ala Ser Leu Gly Asn Gln Gly Ile Gln Leu Phe
                85                  90                  95

Ala Tyr Gly Leu Met Leu Gly Met Gly Ser Ala Val Glu Thr Leu Cys
            100                 105                 110

Gly Gln Ala Tyr Gly Ala His Arg Tyr Glu Met Leu Gly Val Tyr Leu
        115                 120                 125

Gln Arg Ala Thr Val Val Leu Ser Leu Thr Gly Ile Pro Leu Ala Val
    130                 135                 140

Val Tyr Leu Phe Ser Lys Asn Ile Leu Leu Ala Leu Gly Glu Ser Lys
145                 150                 155                 160

Leu Val Ala Ser Ala Ala Ala Val Phe Val Tyr Gly Leu Ile Pro Gln
                165                 170                 175

Ile Phe Ala Tyr Ala Val Asn Phe Pro Ile Gln Lys Phe Leu Gln Ser
            180                 185                 190

Gln Ser Ile Val Ala Pro Ser Ala Phe Ile Ser Leu Gly Thr Leu Phe
        195                 200                 205

Val His Ile Leu Leu Ser Trp Val Val Val Tyr Lys Ile Gly Leu Gly
    210                 215                 220

Leu Leu Gly Ala Ser Leu Val Leu Ser Phe Ser Trp Trp Ile Ile Val
225                 230                 235                 240

Val Ala Gln Phe Ile Tyr Ile Leu Lys Ser Glu Arg Cys Lys Ala Thr
                245                 250                 255

Trp Ala Gly Phe Arg Trp Glu Ala Phe Ser Gly Leu Trp Gln Phe Val
            260                 265                 270
```

```
Lys Leu Ser Ala Gly Ser Ala Val Met Leu Cys Leu Glu Thr Trp Tyr
        275                 280                 285

Phe Gln Ile Leu Val Leu Leu Ser Gly Leu Leu Lys Asn Pro Glu Ile
    290                 295                 300

Ala Leu Ala Ser Ile Ser Val Cys Leu Ala Val Asn Gly Leu Met Phe
305                 310                 315                 320

Met Val Ala Val Gly Phe Asn Ala Ala Ala Ser Val Arg Val Ser Asn
                325                 330                 335

Glu Leu Gly Ala Ala His Pro Lys Ser Ala Ala Phe Ser Val Phe Met
            340                 345                 350

Val Thr Phe Ile Ser Phe Leu Ile Ala Val Val Glu Ala Ile Ile Val
        355                 360                 365

Leu Ser Leu Arg Asn Val Ile Ser Tyr Ala Phe Thr Glu Gly Glu Val
    370                 375                 380

Val Ala Lys Glu Val Ser Ser Leu Cys Pro Tyr Leu Ala Val Thr Leu
385                 390                 395                 400

Ile Leu Asn Gly Ile Gln Pro Val Leu Ser Gly Val Ala Val Gly Cys
                405                 410                 415

Gly Trp Gln Ala Phe Val Ala Tyr Val Asn Val Gly Cys Tyr Tyr Gly
            420                 425                 430

Val Gly Ile Pro Leu Gly Cys Leu Leu Gly Phe Lys Phe Asp Phe Gly
        435                 440                 445

Ala Lys Gly Ile Trp Thr Gly Met Ile Gly Gly Thr Val Met Gln Thr
    450                 455                 460
```

Ile Ile Leu Leu Trp Val Thr Phe Ser Thr Asp Trp Asn Lys Glu Val
465 470 475 480

Glu Ser Ala Arg Lys Arg Leu Asp Lys Trp Glu Asn Leu Lys Gly Pro
485 490 495

Leu Asn Lys Glu
500

<210> 26
<211> 1056
<212> DNA
<213> Atropa belladonna

<400> 26
atggacaatc aaggcagcag caaatggaga aaacttatac ttgtactcaa ctgtatcact 60

ctaggcatag gctattgcgg tggtccttta atttcacgtc tctatttcct ccacggcggc 120

cgaagaattt ggctttctag ttggttagca tgcattggat ggccaatttt cctcattgtt 180

cttatcatcg cctatttttt ccgtcgtaaa acccagggct ctgatatcaa gtttctcatc 240

gtgactcgtc aggagttcac aacatccgct gcgctaggca tacttgtcgg cctaggcggt 300

tatctttatg catgggggcc cgcgaaactc cctgtttcta cttcaactgt tgtttatgcg 360

gcacaactcg gattcaccgt agtatttgct tttctcatag tgaagcagaa attaacggcg 420

tattcggcga atgcggttgt tttgttgatc gttggagcgg ggactttggc gcttcgagcg 480

gatagtgatc gaccatcagg agagtcaact aaacagtatg ttttggggtt tgtaatgact 540

cttttggctg cggcgtcaac ggggcttgcg ttgcctttac tggagttgat ttatatgaag 600

gctcaaaaaa ctgttactta tactatggtg ttggagattc agatgattct caatatattc 660

ggttctgctt tttgcactgt tggaatgatt gttaacaagg attttcaggc gatttcgagg 720

gaggcaagcc aatttgagct tggagaagca aaatattatt tggtgctagt gtggagtgcc 780

attatttggc aattttcaac actaggaatc gttggagtag tccaatacgg ctcctcattg 840

ctgtgtggaa ttttaattgc tttcttactt ccacttacag aactattagg tgtactttta 900

tatcatgaag tatttcaagc tacaaagggg cttgctattt tcctctctct ttggggattt 960

gtttcttaca tctatggtga agtgaaacaa agccagaaga agaaagagga gaacctaatt 1020

ccccaaacag aaatggttca tatcttgcct gtctga 1056

<210> 27
<211> 351
<212> PRT
<213> Atropa belladonna

<400> 27

Met Asp Asn Gln Gly Ser Ser Lys Trp Arg Lys Leu Ile Leu Val Leu
1 5 10 15

Asn Cys Ile Thr Leu Gly Ile Gly Tyr Cys Gly Gly Pro Leu Ile Ser
20 25 30

Arg Leu Tyr Phe Leu His Gly Gly Arg Arg Ile Trp Leu Ser Ser Trp
35 40 45

Leu Ala Cys Ile Gly Trp Pro Ile Phe Leu Ile Val Leu Ile Ile Ala
50 55 60

Tyr Phe Phe Arg Arg Lys Thr Gln Gly Ser Asp Ile Lys Phe Leu Ile
65 70 75 80

Val Thr Arg Gln Glu Phe Thr Thr Ser Ala Ala Leu Gly Ile Leu Val
85 90 95

Gly Leu Gly Gly Tyr Leu Tyr Ala Trp Gly Pro Ala Lys Leu Pro Val
100 105 110

Ser Thr Ser Thr Val Val Tyr Ala Ala Gln Leu Gly Phe Thr Val Val
115 120 125

Phe Ala Phe Leu Ile Val Lys Gln Lys Leu Thr Ala Tyr Ser Ala Asn
130 135 140

Ala Val Val Leu Leu Ile Val Gly Ala Gly Thr Leu Ala Leu Arg Ala
145 150 155 160

Asp Ser Asp Arg Pro Ser Gly Glu Ser Thr Lys Gln Tyr Val Leu Gly
165 170 175

Phe Val Met Thr Leu Leu Ala Ala Ala Ser Thr Gly Leu Ala Leu Pro
180 185 190

Leu Leu Glu Leu Ile Tyr Met Lys Ala Gln Lys Thr Val Thr Tyr Thr
195 200 205

Met Val Leu Glu Ile Gln Met Ile Leu Asn Ile Phe Gly Ser Ala Phe
210 215 220

Cys Thr Val Gly Met Ile Val Asn Lys Asp Phe Gln Ala Ile Ser Arg
225 230 235 240

Glu Ala Ser Gln Phe Glu Leu Gly Glu Ala Lys Tyr Tyr Leu Val Leu
245 250 255

Val Trp Ser Ala Ile Ile Trp Gln Phe Ser Thr Leu Gly Ile Val Gly
260 265 270

Val Val Gln Tyr Gly Ser Ser Leu Leu Cys Gly Ile Leu Ile Ala Phe
275 280 285

Leu Leu Pro Leu Thr Glu Leu Leu Gly Val Leu Leu Tyr His Glu Val
290 295 300

Phe Gln Ala Thr Lys Gly Leu Ala Ile Phe Leu Ser Leu Trp Gly Phe
305 310 315 320

Val Ser Tyr Ile Tyr Gly Glu Val Lys Gln Ser Gln Lys Lys Lys Glu
325 330 335

Glu Asn Leu Ile Pro Gln Thr Glu Met Val His Ile Leu Pro Val

340 345 350

<210> 28
<211> 1344
<212> DNA
<213> Atropa belladonna

<400> 28
atgacggaca atgatgaacc aaagcccctt ggaaggtgtt ctgtctttac ttacggcgtg 60

gggcacatgt tgaatgatat aacatctgcc tgttggtata catatctctt gctctacttg 120

acagatattg gattgtcccc aagcaatgct gcaacagtca cacttactgg ccagtttgct 180

gatgcattta cgacaataat agctggtgaa ctgatagaca ggtttggtca tttcaaaata 240

tggcatggca cgggatgcct gttagttgct ctatctttta catcactctt cggtggttgc 300

ctcccgtgca aaattcttgg atctgattca cctgttgttc aaactgttgg ttactgcata 360

tctgcagtga tcttctgcag cggatggtct tgcactcaga tatcacacat gtcaatggtg 420

aatggcgtca cactaaaccc aacaagcaga gtagcatgtg ttagctgtcg caatgcttgt 480

acaatgattg caagcctaat cttatatgca atcgcatttt tcgtcttcaa taagaacact 540

tctagctcac aggttgacat taagaatcag taccaattga tggcttatat ttcgattttg 600

attggagcag tcttcgtgat tatatttcag cttggaacta aagagccaag tctaaaacaa 660

gaatctgacg gaaaacgtgc tacagcttct tggacatact ggttgaagaa agtcctatat 720

tatcaagtac ttagcatcta tgttttaact agagttgtga ctaatgtttc acagaccttc 780

ttagctgttt acgtaataaa tgatctacgg atgagcccat cttcgaaagc cttgattcct 840

gccaccatct acttatccag cttcatcgtg tctgtcctcc tccaggaact catatggagt 900

ggtgaacggt tgaaggcctt tttttccgct gggggtctcc tttggctatt ttgtggtgct 960

acagtaattt tcctcccaat aaacatgaat gcttttatgt acgtcttgtc cgtacttatt 1020

ggcatagcaa acgccttgat gatggtaaca tcaataggga tggaaagcga gctagttgat 1080

aaagatcttg atggatctgc tttcgtttat ggatcattgg gctttgttga caaagttcta 1140

tgtggagtgt ttttgtactt tctcgagtca tatgagagtg cagatcctgt tacctgtgat 1200

ccagcatatc cttgtttctc cgttacacga ttttctctag gtttcattcc agggatttct 1260

gccataattg gagtcgtagt cacattcttc ataagattcc acacatccca tcctaagcca 1320

ttaacagaac ccctcttggc atag 1344

<210> 29
<211> 447
<212> PRT
<213> Atropa belladonna

<400> 29

Met Thr Asp Asn Asp Glu Pro Lys Pro Leu Gly Arg Cys Ser Val Phe
1 5 10 15

Thr Tyr Gly Val Gly His Met Leu Asn Asp Ile Thr Ser Ala Cys Trp
20 25 30

Tyr Thr Tyr Leu Leu Leu Tyr Leu Thr Asp Ile Gly Leu Ser Pro Ser
35 40 45

Asn Ala Ala Thr Val Thr Leu Thr Gly Gln Phe Ala Asp Ala Phe Thr
50 55 60

Thr Ile Ile Ala Gly Glu Leu Ile Asp Arg Phe Gly His Phe Lys Ile
65 70 75 80

Trp His Gly Thr Gly Cys Leu Leu Val Ala Leu Ser Phe Thr Ser Leu
85 90 95

Phe Gly Gly Cys Leu Pro Cys Lys Ile Leu Gly Ser Asp Ser Pro Val
100 105 110

Val Gln Thr Val Gly Tyr Cys Ile Ser Ala Val Ile Phe Cys Ser Gly
115 120 125

Trp Ser Cys Thr Gln Ile Ser His Met Ser Met Val Asn Gly Val Thr
130 135 140

```
Leu Asn Pro Thr Ser Arg Val Ala Cys Val Ser Cys Arg Asn Ala Cys
145                 150                 155                 160

Thr Met Ile Ala Ser Leu Ile Leu Tyr Ala Ile Ala Phe Phe Val Phe
                165                 170                 175

Asn Lys Asn Thr Ser Ser Ser Gln Val Asp Ile Lys Asn Gln Tyr Gln
            180                 185                 190

Leu Met Ala Tyr Ile Ser Ile Leu Ile Gly Ala Val Phe Val Ile Ile
        195                 200                 205

Phe Gln Leu Gly Thr Lys Glu Pro Ser Leu Lys Gln Glu Ser Asp Gly
    210                 215                 220

Lys Arg Ala Thr Ala Ser Trp Thr Tyr Trp Leu Lys Lys Val Leu Tyr
225                 230                 235                 240

Tyr Gln Val Leu Ser Ile Tyr Val Leu Thr Arg Val Val Thr Asn Val
                245                 250                 255

Ser Gln Thr Phe Leu Ala Val Tyr Val Ile Asn Asp Leu Arg Met Ser
            260                 265                 270

Pro Ser Ser Lys Ala Leu Ile Pro Ala Thr Ile Tyr Leu Ser Ser Phe
        275                 280                 285

Ile Val Ser Val Leu Leu Gln Glu Leu Ile Trp Ser Gly Glu Arg Leu
    290                 295                 300

Lys Ala Phe Phe Ser Ala Gly Gly Leu Leu Trp Leu Phe Cys Gly Ala
305                 310                 315                 320

Thr Val Ile Phe Leu Pro Ile Asn Met Asn Ala Phe Met Tyr Val Leu
                325                 330                 335
```

```
Ser Val Leu Ile Gly Ile Ala Asn Ala Leu Met Met Val Thr Ser Ile
            340                 345                 350

Gly Met Glu Ser Glu Leu Val Asp Lys Asp Leu Asp Gly Ser Ala Phe
        355                 360                 365

Val Tyr Gly Ser Leu Gly Phe Val Asp Lys Val Leu Cys Gly Val Phe
    370                 375                 380

Leu Tyr Phe Leu Glu Ser Tyr Glu Ser Ala Asp Pro Val Thr Cys Asp
385                 390                 395                 400

Pro Ala Tyr Pro Cys Phe Ser Val Thr Arg Phe Ser Leu Gly Phe Ile
                405                 410                 415

Pro Gly Ile Ser Ala Ile Ile Gly Val Val Val Thr Phe Phe Ile Arg
            420                 425                 430

Phe His Thr Ser His Pro Lys Pro Leu Thr Glu Pro Leu Leu Ala
        435                 440                 445
```

<210> 30
<211> 790
<212> PRT
<213> Artificial sequence

<220>
<223> NtMPO1 from plant transcriptome data

<400> 30

```
Met Ala Thr Thr Lys Gln Lys Val Thr Ala Pro Ser Pro Ser Pro Ser
1               5                   10                  15

Ser Ser Thr Ala Ser Cys Cys Pro Ser Thr Ser Ile Leu Arg Arg Glu
            20                  25                  30

Ala Thr Ala Ala Ile Ala Val Val Gly Asp Gly Leu Gln Asn Trp Thr
        35                  40                  45
```

```
Asn Ile Pro Ser Val Asp Glu Lys Gln Lys Lys Thr Ala Ser Ser Ala
    50                  55                  60

Leu Ala Ser Leu Pro Thr Thr Glu Pro Leu Ser Thr Asn Thr Ser Thr
65                  70                  75                  80

Lys Gly Ile Gln Ile Met Thr Arg Ala Gln Thr Cys His Pro Leu Asp
                85                  90                  95

Pro Leu Ser Ala Ala Glu Ile Ser Val Ala Val Ala Thr Val Arg Ala
            100                 105                 110

Ala Gly Glu Phe Pro Glu Val Arg Asp Gly Met Arg Phe Ile Glu Val
        115                 120                 125

Val Leu Val Glu Pro Asp Lys Ser Val Val Ala Leu Ala Asp Ala Tyr
    130                 135                 140

Phe Phe Pro Pro Phe Gln Ser Ser Leu Met Pro Arg Thr Lys Gly Gly
145                 150                 155                 160

Ser Gln Ile Pro Thr Lys Leu Pro Pro Arg Arg Ala Arg Leu Ile Val
                165                 170                 175

Tyr Asn Lys Lys Thr Asn Glu Thr Ser Ile Trp Ile Val Glu Leu Asn
            180                 185                 190

Glu Val His Ala Ala Ala Arg Gly Gly His His Arg Gly Lys Val Ile
        195                 200                 205

Ala Ser Asn Val Val Pro Asp Val Gln Pro Pro Ile Asp Ala Gln Glu
    210                 215                 220

Tyr Ala Glu Cys Glu Ala Val Val Lys Ser Tyr Pro Pro Phe Arg Asp
225                 230                 235                 240
```

```
Ala Met Arg Arg Arg Gly Ile Asp Asp Leu Asp Leu Val Met Val Asp
               245                 250                 255

Pro Trp Cys Val Gly Tyr His Ser Glu Ala Asp Ala Pro Ser Arg Arg
            260                 265                 270

Leu Ala Lys Pro Leu Val Phe Cys Arg Thr Glu Ser Asp Cys Pro Met
        275                 280                 285

Glu Asn Gly Tyr Ala Arg Pro Val Glu Gly Ile Tyr Val Leu Val Asp
    290                 295                 300

Val Gln Asn Met Lys Ile Ile Glu Phe Glu Asp Arg Lys Leu Val Pro
305                 310                 315                 320

Leu Pro Pro Val Asp Pro Leu Arg Asn Tyr Thr Ala Gly Glu Thr Arg
               325                 330                 335

Gly Gly Val Asp Arg Ser Asp Val Lys Pro Leu His Ile Ile Gln Pro
            340                 345                 350

Glu Gly Pro Ser Phe Arg Ile Ser Gly Asn Tyr Val Glu Trp Gln Lys
        355                 360                 365

Trp Asn Phe Arg Ile Gly Phe Thr Pro Arg Glu Gly Leu Val Ile His
    370                 375                 380

Ser Val Ala Tyr Leu Asp Gly Ser Arg Gly Arg Arg Pro Ile Ala His
385                 390                 395                 400

Arg Leu Ser Phe Val Glu Met Val Val Pro Tyr Gly Asp Pro Asn Asp
               405                 410                 415

Pro His Tyr Arg Lys Asn Ala Phe Asp Ala Gly Glu Asp Gly Leu Gly
            420                 425                 430

Lys Asn Ala His Ser Leu Lys Arg Gly Cys Asp Cys Leu Gly Tyr Ile
```

435 440 445

Lys Tyr Phe Asp Ala His Phe Thr Asn Phe Thr Gly Gly Val Glu Thr
450 455 460

Thr Glu Asn Cys Val Cys Leu His Glu Glu Asp His Gly Met Leu Trp
465 470 475 480

Lys His Gln Asp Trp Arg Thr Gly Leu Ala Glu Val Arg Arg Ser Arg
485 490 495

Arg Leu Thr Val Ser Phe Val Cys Thr Val Ala Asn Tyr Glu Tyr Ala
500 505 510

Phe Tyr Trp His Phe Tyr Gln Asp Gly Lys Ile Glu Ala Glu Val Lys
515 520 525

Leu Thr Gly Ile Leu Ser Leu Gly Ala Leu Gln Pro Gly Glu Tyr Arg
530 535 540

Lys Tyr Gly Thr Thr Ile Leu Pro Gly Leu Tyr Ala Pro Val His Gln
545 550 555 560

His Phe Phe Val Ala Arg Met Asn Met Ala Val Asp Cys Lys Pro Gly
565 570 575

Glu Ala His Asn Gln Val Val Glu Val Asn Val Lys Val Glu Glu Pro
580 585 590

Gly Lys Glu Asn Val His Asn Asn Ala Phe Tyr Ala Glu Glu Thr Leu
595 600 605

Leu Arg Ser Glu Leu Gln Ala Met Arg Asp Cys Asp Pro Phe Ser Ala
610 615 620

Arg His Trp Ile Val Arg Asn Thr Arg Thr Val Asn Arg Thr Gly Gln
625 630 635 640

```
Leu Thr Gly Tyr Lys Leu Val Pro Gly Pro Asn Cys Leu Pro Leu Ala
                645                 650                 655

Gly Pro Glu Ala Lys Phe Leu Arg Arg Ala Ala Phe Leu Lys His Asn
            660                 665                 670

Leu Trp Val Thr Gln Tyr Ala Pro Gly Glu Asp Phe Pro Gly Gly Glu
        675                 680                 685

Phe Pro Asn Gln Asn Pro Arg Val Gly Glu Gly Leu Ala Ser Trp Val
    690                 695                 700

Lys Gln Asp Arg Pro Leu Glu Glu Ser Asp Ile Val Leu Trp Tyr Ile
705                 710                 715                 720

Phe Gly Ile Thr His Val Pro Arg Leu Glu Asp Trp Pro Val Met Pro
                725                 730                 735

Val Glu His Ile Gly Phe Val Leu Gln Pro His Gly Tyr Phe Asn Cys
            740                 745                 750

Ser Pro Ala Val Asp Val Pro Pro Pro Phe Ala Cys Asp Ser Glu Ser
        755                 760                 765

Arg Asp Ser Asp Val Thr Glu Thr Ser Val Ala Lys Ser Thr Ala Thr
    770                 775                 780

Ser Leu Leu Ala Lys Leu
785                 790
```

<210> 31
<211> 646
<212> PRT
<213> Artificial sequence

<220>
<223> AbMPO1 from plant transcriptome data

<400> 31

```
Met Phe Pro Pro Phe Gln Ser Ser Leu Met Pro Arg Thr Lys Gly Gly
1               5                  10                  15

Leu Leu Val Pro Ser Arg Leu Pro Pro Arg Arg Ala Arg Leu Ile Val
            20                  25                  30

Tyr Asn Lys Lys Thr Asn Glu Thr Ser Val Trp Ile Val Gln Leu Thr
        35                  40                  45

Glu Val His Ala Ile Ala Arg Gly Gly Gln His Arg Gly Lys Val Ile
    50                  55                  60

Ser Ser Lys Val Val Pro Asp Val Gln Pro Pro Ile Asp Ala Gln Glu
65                  70                  75                  80

Tyr Ala Asp Cys Glu Ser Val Val Lys Asn Tyr Pro Ser Phe Ile Glu
                85                  90                  95

Ala Met Lys Arg Arg Gly Ile Asp Asp Met Asp Leu Val Met Val Asp
            100                 105                 110

Pro Trp Cys Val Gly Tyr His Ser Glu Ala Asp Ala Pro Ser Arg Arg
        115                 120                 125

Leu Ala Lys Pro Leu Val Phe Cys Arg Thr Glu Ser Asp Cys Pro Met
    130                 135                 140

Glu Asn Gly Tyr Ala Arg Pro Val Glu Gly Ile His Val Leu Val Asp
145                 150                 155                 160

Val Gln Asn Met Gln Val Ile Glu Phe Glu Asp Arg Lys Val Val Pro
                165                 170                 175

Leu Pro Pro Ala Asp Pro Leu Arg Asn Tyr Thr Ala Gly Glu Thr Arg
            180                 185                 190
```

```
Gly Gly Val Asp Arg Ser Asp Val Lys Pro Leu Gln Ile Ile Gln Pro
        195                 200                 205

Asp Gly Pro Ser Phe Arg Val Asp Gly Asn Tyr Val Gln Trp Gln Lys
    210                 215                 220

Trp Asn Phe Arg Val Gly Phe Thr Pro Arg Glu Gly Leu Val Ile His
225                 230                 235                 240

Ser Val Ala Tyr Leu Asp Gly Ser Arg Gly Arg Arg Ser Ile Ala His
                245                 250                 255

Arg Leu Ser Phe Val Glu Met Val Val Pro Tyr Gly Asp Pro Asn Asp
            260                 265                 270

Pro His Tyr Arg Lys Asn Ala Phe Asp Ala Gly Glu Asp Gly Leu Gly
        275                 280                 285

Lys Asn Ala His Ser Leu Lys Arg Gly Cys Asp Cys Leu Gly Tyr Ile
    290                 295                 300

Lys Tyr Phe Asp Ala Asn Phe Thr Asn Phe Thr Gly Gly Val Glu Thr
305                 310                 315                 320

Ile Glu Asn Cys Val Cys Leu His Glu Glu Asp His Gly Met Leu Trp
                325                 330                 335

Lys His Gln Asp Trp Arg Thr Gly Leu Ala Glu Val Arg Arg Ser Arg
            340                 345                 350

Arg Leu Thr Val Ser Phe Ile Cys Thr Val Ala Asn Tyr Glu Tyr Gly
        355                 360                 365

Phe Tyr Trp His Phe Tyr Gln Asp Gly Lys Ile Glu Ala Glu Val Lys
    370                 375                 380
```

```
Leu Thr Gly Ile Leu Ser Leu Gly Ala Leu Gln Pro Gly Glu Ser Arg
385                 390                 395                 400

Lys Tyr Gly Thr Thr Ile Thr Pro Gly Leu Tyr Ala Pro Val His Gln
                405                 410                 415

His Phe Phe Val Ala Arg Met Asn Met Ala Val Asp Cys Lys Pro Gly
            420                 425                 430

Glu Ala His Asn Gln Ile Val Glu Val Asn Val Arg Val Glu Glu Pro
        435                 440                 445

Gly Lys Glu Asn Val His Asn Asn Ala Phe Tyr Ala Glu Glu Thr Leu
    450                 455                 460

Leu Arg Ser Glu Leu Gln Ala Met Arg Asp Cys Asp Pro Leu Ser Ala
465                 470                 475                 480

Arg His Trp Ile Val Arg Asn Thr Arg Thr Ala Asn Arg Thr Gly Lys
                485                 490                 495

Leu Thr Gly Tyr Lys Leu Val Pro Gly Pro Asn Cys Leu Pro Leu Ala
            500                 505                 510

Gly Pro Glu Ala Lys Phe Leu Arg Arg Ala Ala Phe Leu Lys His Asn
        515                 520                 525

Leu Trp Val Thr Gln Tyr Ala Pro Gly Glu Asp Phe Pro Gly Gly Glu
    530                 535                 540

Phe Pro Asn Gln Asn Pro Arg Val Gly Glu Gly Leu Ala Ser Trp Val
545                 550                 555                 560

Lys Gln Asp Arg Ser Leu Glu Glu Ser Asp Ile Val Leu Trp Tyr Val
                565                 570                 575
```

Phe Gly Ile Thr His Val Pro Arg Leu Glu Asp Trp Pro Val Met Pro
580 585 590

Val Glu His Ile Gly Phe Met Leu Gln Pro His Gly Phe Phe Asn Cys
595 600 605

Ser Pro Ala Val Asp Val Pro Pro Pro Arg Gly Cys Asp Leu Glu Ser
610 615 620

Lys Asp Ser Asp Val Ser Glu Asn Gly Val Gly Lys His Thr Ala Thr
625 630 635 640

Gly Leu Met Ala Lys Leu
645

<210> 32
<211> 786
<212> PRT
<213> Artificial sequence

<220>
<223> DmMPO1 from plant transcriptome data

<400> 32

Met Ser Gln Lys Val Thr Lys Pro Ser Ala Ala Ala Ala Val Cys Cys
1 5 10 15

Pro Ser Thr Pro Val Leu Arg Arg Glu Ala Thr Ala Ala Ala Ala Ala
20 25 30

Ile Met Ser Asp Gly Gly Leu Gln Asn Trp Thr Asn Val Arg Ser Val
35 40 45

Asp His Asn Gln Lys Lys Lys Thr Ser Ser Ala Ile Ala Ser Leu Glu
50 55 60

Ser Asn Thr Glu Thr Arg Ser Ser Asn Thr Ser Thr Lys Gly Ile Gln
65 70 75 80

```
Ile Met Pro Arg Ala Gln Thr Cys His Pro Leu Asp Pro Leu Cys Ala
                85                  90                  95

Ala Glu Ile Ser Val Ala Val Ala Thr Val Arg Ala Ala Cys Glu Phe
            100                 105                 110

Pro Glu Val Arg Asp Gly Met Arg Phe Ile Glu Val Val Leu Leu Glu
        115                 120                 125

Pro Asp Lys Asn Val Val Ala Leu Ala Asp Ala Tyr Phe Phe Pro Pro
    130                 135                 140

Phe Gln Ser Ser Leu Ile Arg Arg Thr Lys Gly Gly Ser Leu Ile Pro
145                 150                 155                 160

Thr Lys Leu Pro Pro Arg Arg Ala Arg Leu Ile Val Tyr Asn Lys Lys
                165                 170                 175

Thr Asn Glu Thr Ser Ile Trp Ile Val Glu Leu Thr Glu Val His Ala
            180                 185                 190

Ala Ala Arg Gly Gly His His Arg Gly Lys Val Ile Ser Ser Asn Gly
        195                 200                 205

Val Pro Asp Val Gln Pro Pro Ile Asp Ala Gln Glu Tyr Ala Glu Cys
    210                 215                 220

Glu Ala Val Val Lys Asn Tyr Pro Pro Phe Arg Glu Ala Met Lys Arg
225                 230                 235                 240

Arg Gly Ile Asp Asp Met Asp Leu Val Met Val Asp Pro Trp Cys Val
                245                 250                 255

Gly Tyr His Ser Glu Ala Asp Ala Pro Ser His Arg Leu Ala Lys Pro
            260                 265                 270
```

```
Leu Val Phe Cys Arg Thr Glu Ser Glu Cys Pro Met Glu Asn Gly Tyr
        275                 280                 285

Ala Arg Pro Val Glu Gly Ile Tyr Val Leu Val Asp Val Gln Asn Met
    290                 295                 300

Gln Ile Ile Glu Phe Glu Asp Arg Lys Leu Val Pro Leu Pro Pro Ile
305                 310                 315                 320

Asp Pro Leu Arg Asn Tyr Thr Ala Gly Glu Thr Arg Gly Gly Val Asp
                325                 330                 335

Arg Ser Asp Val Lys Pro Leu His Ile Ile Gln Pro Glu Cys Pro Ser
            340                 345                 350

Phe Arg Ile Ser Gly Asn Tyr Val Glu Trp Gln Lys Trp Asn Phe Arg
        355                 360                 365

Ile Gly Phe Thr Pro Arg Glu Gly Leu Val Ile His Ser Val Ala Tyr
    370                 375                 380

Leu Asp Gly Ser Arg Gly Arg Arg Pro Ile Ala His Arg Leu Ser Phe
385                 390                 395                 400

Val Glu Met Val Val Pro Tyr Gly Asp Pro Asn Asp Pro His Tyr Arg
                405                 410                 415

Lys Asn Ala Phe Asp Ala Gly Glu Asp Gly Leu Gly Lys Asn Ala His
            420                 425                 430

Ser Leu Lys Arg Gly Cys Asp Cys Leu Gly Tyr Ile Lys Tyr Phe Asp
        435                 440                 445

Ala His Phe Thr Asn Phe Thr Gly Gly Val Glu Thr Thr Glu Asn Cys
    450                 455                 460

Val Cys Leu His Glu Glu Asp His Gly Met Leu Trp Lys His Gln Asp
```

```
465                 470                 475                 480

Trp Arg Thr Gly Leu Ala Glu Val Arg Arg Ser Arg Arg Leu Thr Val
                485                 490                 495

Ser Phe Val Cys Thr Val Ala Asn Tyr Glu Tyr Ala Phe Tyr Trp His
            500                 505                 510

Phe Tyr Gln Asp Gly Lys Ile Glu Ala Glu Val Lys Leu Thr Gly Ile
        515                 520                 525

Leu Ser Leu Gly Ala Leu Gln Pro Gly Glu Ser Arg Lys Tyr Gly Thr
    530                 535                 540

Thr Ile Ala Pro Gly Leu Tyr Ala Pro Val His Gln His Phe Phe Val
545                 550                 555                 560

Ala Arg Met Asn Met Ala Val Asp Cys Lys Pro Gly Glu Ala His Asn
                565                 570                 575

Gln Val Val Glu Val Asn Val Lys Val Glu Glu Pro Gly Lys Glu Asn
            580                 585                 590

Val His Asn Asn Ala Phe Tyr Ala Glu Glu Thr Leu Leu Arg Ser Glu
        595                 600                 605

Ser Glu Ala Met Arg Asp Cys Asp Pro Phe Ser Ala Arg His Trp Ile
    610                 615                 620

Val Arg Asn Thr Arg Thr Val Asn Arg Thr Gly Gln Leu Thr Gly Tyr
625                 630                 635                 640

Lys Leu Val Pro Gly Pro Asn Cys Leu Pro Leu Ala Gly His Glu Ala
                645                 650                 655

Lys Phe Leu Arg Arg Ala Ala Phe Leu Lys His Asn Leu Trp Val Thr
            660                 665                 670
```

```
Gln Tyr Ala Pro Gly Glu Asp Phe Pro Gly Gly Glu Phe Pro Asn Asp
        675                 680                 685

Asn Pro Arg Ala Gly Glu Gly Leu Ala Ser Trp Val Lys Gln Asp Arg
    690                 695                 700

Pro Leu Glu Glu Ser Asp Ile Val Leu Trp Tyr Ile Phe Gly Ile Thr
705                 710                 715                 720

His Val Pro Arg Leu Glu Asp Trp Pro Val Met Pro Val Glu His Ile
                725                 730                 735

Gly Phe Met Leu Gln Pro His Gly Phe Phe Asn Cys Ser Pro Ala Val
            740                 745                 750

Asp Val Pro Pro Pro Ser Gly Cys Asp Ser Glu Ser Arg Asp Ser Asp
        755                 760                 765

Val Thr Glu Asn Ser Val Ala Lys His Thr Pro Thr Gly Leu Met Ala
    770                 775                 780

Lys Leu
785
```

<210> 33
<211> 465
<212> PRT
<213> Arabidopsis thaliana

<400> 33

```
Met Arg Asn Leu Ser Phe Ile Val Leu Phe Leu Leu Thr Leu Phe Phe
1               5                   10                  15

Ile His His Leu Val Asp Ala Ser Leu Leu Val Lys Ser Leu Pro Gly
            20                  25                  30
```

```
Phe Glu Gly Pro Leu Pro Phe Glu Leu Glu Thr Gly Tyr Val Ser Ile
        35                  40                  45

Gly Glu Ser Gly Asp Val Glu Leu Phe Tyr Tyr Phe Val Lys Ser Glu
    50                  55                  60

Arg Asn Pro Glu Asn Asp Pro Leu Met Ile Trp Leu Thr Gly Gly Pro
65                  70                  75                  80

Gly Cys Ser Ser Ile Cys Gly Leu Leu Phe Ala Asn Gly Pro Leu Ala
                85                  90                  95

Phe Lys Gly Asp Glu Tyr Asn Gly Thr Val Pro Pro Leu Glu Leu Thr
            100                 105                 110

Ser Phe Ser Trp Thr Lys Val Ala Asn Ile Leu Tyr Leu Glu Ala Pro
        115                 120                 125

Ala Gly Ser Gly Tyr Ser Tyr Ala Lys Thr Arg Arg Ala Phe Glu Ser
    130                 135                 140

Ser Asp Thr Lys Gln Met His Gln Ile Asp Gln Phe Leu Arg Ser Trp
145                 150                 155                 160

Phe Val Lys His Pro Glu Phe Ile Ser Asn Pro Phe Tyr Val Gly Gly
                165                 170                 175

Asp Ser Tyr Ser Gly Lys Ile Val Pro Gly Ala Val Gln Gln Ile Ser
            180                 185                 190

Leu Gly Asn Glu Lys Gly Leu Thr Pro Leu Ile Asn Ile Gln Gly Tyr
        195                 200                 205

Val Leu Gly Asn Pro Val Thr Asp Lys Asn Ile Glu Thr Asn Tyr Arg
    210                 215                 220

Val Pro Phe Ala His Gly Met Gly Leu Ile Ser Asp Glu Leu Phe Glu
```

225 230 235 240

Ser Leu Glu Arg Ser Cys Gly Gly Lys Phe Phe Asn Val Asp Pro Ser
245 250 255

Asn Ala Arg Cys Ser Asn Asn Leu Gln Ala Tyr Asp His Cys Met Ser
260 265 270

Glu Ile Tyr Ser Glu His Ile Leu Leu Arg Asn Cys Lys Val Asp Tyr
275 280 285

Val Leu Ala Asp Thr Pro Asn Ile Arg Thr Asp Arg Arg Arg Val Met
290 295 300

Lys Glu Phe Ser Val Asn Asp Ser Ser Ser Leu Pro Pro Pro Ser Cys
305 310 315 320

Phe Thr Tyr Arg Tyr Phe Leu Ser Ala Phe Trp Ala Asn Asp Glu Asn
325 330 335

Val Arg Arg Ala Leu Gly Val Lys Lys Glu Val Gly Lys Trp Asn Arg
340 345 350

Cys Asn Ser Gln Asn Ile Pro Tyr Thr Phe Glu Ile Phe Asn Ala Val
355 360 365

Pro Tyr His Val Asn Asn Ser Leu Lys Gly Phe Arg Ser Leu Ile Tyr
370 375 380

Ser Gly Asp His Asp Ser Thr Trp Pro Phe Ser Ser Thr Gln Met Ile
385 390 395 400

Ile Arg Ala Leu Asn Tyr Ser Ile Val Asp Asp Trp Arg Pro Trp Met
405 410 415

Met Ser Ser Asn Gln Val Ala Gly Tyr Thr Arg Thr Tyr Ala Asn Lys
420 425 430

```
Met Thr Phe Ala Thr Ile Lys Gly Gly Gly His Thr Ala Glu Tyr Thr
        435                 440                 445


Pro Asp Gln Cys Ser Leu Met Phe Arg Arg Trp Ile Asp Gly Glu Pro
    450                 455                 460


Leu
465


<210>  34
<211>  433
<212>  PRT
<213>  Arabidopsis thaliana

<400>  34

Met Ser Leu Lys Ile Lys Phe Leu Leu Leu Leu Val Leu Tyr His His
1               5                   10                  15


Val Asp Ser Ala Ser Ile Val Lys Phe Leu Pro Gly Phe Glu Gly Pro
            20                  25                  30


Leu Pro Phe Glu Leu Glu Thr Gly Tyr Ile Gly Ile Gly Glu Asp Glu
        35                  40                  45


Asn Val Gln Phe Phe Tyr Tyr Phe Ile Lys Ser Glu Asn Asn Pro Lys
    50                  55                  60


Glu Asp Pro Leu Leu Ile Trp Leu Asn Gly Gly Pro Gly Cys Ser Cys
65                  70                  75                  80


Leu Gly Gly Ile Ile Phe Glu Asn Gly Pro Val Gly Leu Lys Phe Glu
                85                  90                  95


Val Phe Asn Gly Ser Ala Pro Ser Leu Phe Ser Thr Thr Tyr Ser Trp
            100                 105                 110
```

```
Thr Lys Met Ala Asn Ile Ile Phe Leu Asp Gln Pro Val Gly Ser Gly
        115                 120                 125

Phe Ser Tyr Ser Lys Thr Pro Ile Asp Lys Thr Gly Asp Ile Ser Glu
    130                 135                 140

Val Lys Arg Thr His Glu Phe Leu Gln Lys Trp Leu Ser Arg His Pro
145                 150                 155                 160

Gln Tyr Phe Ser Asn Pro Leu Tyr Val Val Gly Asp Ser Tyr Ser Gly
                165                 170                 175

Met Ile Val Pro Ala Leu Val Gln Glu Ile Ser Gln Gly Asn Tyr Ile
            180                 185                 190

Cys Cys Glu Pro Pro Ile Asn Leu Gln Gly Tyr Met Leu Gly Asn Pro
        195                 200                 205

Val Thr Tyr Met Asp Phe Glu Gln Asn Phe Arg Ile Pro Tyr Ala Tyr
    210                 215                 220

Gly Met Gly Leu Ile Ser Asp Glu Ile Tyr Glu Pro Met Lys Arg Ile
225                 230                 235                 240

Cys Asn Gly Asn Tyr Tyr Asn Val Asp Pro Ser Asn Thr Asp Cys Leu
                245                 250                 255

Lys Leu Thr Glu Glu Tyr His Lys Cys Thr Ala Lys Ile Asn Ile His
            260                 265                 270

His Ile Leu Thr Pro Asp Cys Asp Val Thr Asn Val Thr Ser Pro Asp
        275                 280                 285

Cys Tyr Tyr Tyr Pro Tyr His Leu Ile Glu Cys Trp Ala Asn Asp Glu
    290                 295                 300

Ser Val Arg Glu Ala Leu His Ile Glu Lys Gly Ser Lys Gly Lys Trp
```

305 310 315 320

Ala Arg Cys Asn Arg Thr Ile Pro Tyr Asn His Asp Ile Val Ser Ser
325 330 335

Ile Pro Tyr His Met Asn Asn Ser Ile Ser Gly Tyr Arg Ser Leu Ile
340 345 350

Tyr Ser Gly Asp His Asp Ile Ala Val Pro Phe Leu Ala Thr Gln Ala
355 360 365

Trp Ile Arg Ser Leu Asn Tyr Ser Pro Ile His Asn Trp Arg Pro Trp
370 375 380

Met Ile Asn Asn Gln Ile Ala Gly Tyr Thr Arg Ala Tyr Ser Asn Lys
385 390 395 400

Met Thr Phe Ala Thr Ile Lys Gly Gly Gly His Thr Ala Glu Tyr Arg
405 410 415

Pro Asn Glu Thr Phe Ile Met Phe Gln Arg Trp Ile Ser Gly Gln Pro
420 425 430

Leu

<210> 35
<211> 476
<212> PRT
<213> Atropa belladonna

<400> 35

Met Lys Lys Thr Ile Val Val Pro Ile Leu Lys Tyr His Lys Ile Leu
1 5 10 15

Leu Leu Phe Val Leu Ile Leu Val Leu Leu Leu Cys Pro Leu Ile Glu
20 25 30

```
Ala Ala Gly Thr Pro Val Lys Tyr Leu Pro Gly Phe Gly Pro Leu Pro
        35                  40                  45

Phe Glu Phe Glu Thr Gly Tyr Val Gly Leu Gly Glu Ser Glu Glu Val
    50                  55                  60

Gln Leu Phe Tyr Tyr Phe Phe Lys Ser Glu Ser Asn Pro Glu Val Asp
65                  70                  75                  80

Pro Leu Ile Leu Trp Ile Thr Gly Gly Pro Gly Cys Gly Ala Leu Asn
                85                  90                  95

Ala Ile Thr Thr Glu Leu Gly Pro Val Leu Leu Asp Ala Lys Glu Tyr
            100                 105                 110

Asp Gly Ser Leu Pro Thr Leu Ser Leu Asn Pro Tyr Ser Tyr Thr Lys
        115                 120                 125

Val Ala Asn Ile Ile Phe Leu Asp Ser Pro Val Thr Gly Gly Phe Ser
    130                 135                 140

Tyr Ala Thr Thr Lys Glu Ala Asn His Ser Asp Asn Val Gln Met Ala
145                 150                 155                 160

Leu His Thr His Gln Phe Val Gln Lys Trp Leu Val Asp His Ser Glu
                165                 170                 175

Tyr Leu Ser Asn Asp Phe Tyr Val Ala Gly Asp Ser Tyr Ser Gly Ile
            180                 185                 190

Ser Val Pro Ile Ile Thr Gln Val Ile Ser Asp Gly Asn Glu Ala Gly
        195                 200                 205

Asn Lys Pro Trp Ile Asn Leu Lys Gly Tyr Ile Leu Gly Asn Ala Val
    210                 215                 220
```

```
Thr Phe Arg Pro Asp Glu Gln Asn Tyr Arg Ile Pro His Ala His Gly
225                 230                 235                 240

Leu Ala Leu Ile Ser Asp Glu Leu Tyr Lys Ser Leu Glu Ser Ser Cys
                245                 250                 255

Gly Gly Glu Tyr Gln Tyr Ile Asp Gln Thr Asn Thr His Cys Leu Gln
            260                 265                 270

His Val Gln Thr Phe Asn Arg Leu Val Ser Gly Ile Tyr Phe Glu His
        275                 280                 285

Ile Leu Glu Pro Ile Cys Asn Pro Val Ser Thr Lys Ala Arg His Leu
    290                 295                 300

Ser Pro Gln Arg Arg Tyr Leu Asn Gln Lys Leu Gly Gln Leu Lys Asn
305                 310                 315                 320

Pro Thr Met Leu Pro Gly Val Lys Cys Arg Asp Glu Trp His Leu Leu
                325                 330                 335

Ser Glu Ile Trp Val Asn Asp Glu Thr Val Gln Glu Ala Leu His Val
            340                 345                 350

Arg Lys Gly Thr His Gly Ile Trp Lys Gln Cys Pro Asn Tyr Glu Lys
        355                 360                 365

Met Pro Phe Thr Arg Thr Ile Asn Asn Thr Ile Pro Phe His Ala Ser
    370                 375                 380

Leu Ser Lys Lys Gly Tyr Arg Ser Leu Ile Tyr Ser Gly Asp Tyr Asp
385                 390                 395                 400

Leu Tyr Val Pro Phe Leu Ser Thr Gln Ala Trp Ile Arg Ser Leu Asn
                405                 410                 415

Tyr Ser Ile Asp Thr Glu Trp Arg Arg Trp Phe Val Asp Gly Gln Val
```

420 425 430

Ala Gly Tyr Val Thr Thr Tyr Ser Asn Gln Met Thr Phe Thr Thr Ile
435 440 445

Lys Gly Ala Gly His Thr Ala Pro Glu Tyr Lys Pro Ala Glu Cys Leu
450 455 460

Ala Asn Leu Lys Arg Trp Ile Tyr Tyr Gln Pro Leu
465 470 475

<210> 36
<211> 493
<212> PRT
<213> Avena strigosa

<400> 36

Met Glu Lys Leu Leu Val Val Val Leu Leu Leu Val Thr Ile Leu Ala
1 5 10 15

Leu Gly Ala Ala Ala Glu Arg Thr Arg Val Thr His Leu Lys Gly Phe
20 25 30

Asp Gly Pro Leu Pro Phe Ser Leu Glu Thr Gly Tyr Val Glu Val Asp
35 40 45

Glu Thr His Gly Val Glu Leu Phe Tyr Tyr Phe Ile Glu Ser Glu Arg
50 55 60

Lys Pro Ala Glu Asp Pro Val Ile Leu Trp Val Ser Gly Gly Pro Gly
65 70 75 80

Cys Ser Gly Leu Asn Ala Leu Phe Phe Glu Ile Gly Pro Leu Lys Leu
85 90 95

Asp Met Ala Ser Tyr Ala Ala Thr Gly Gly Lys Gly Phe Pro Gly Leu
100 105 110

```
Leu Tyr Phe Glu Asp Ala Trp Thr Lys Ala Ser Asn Met Ile Phe Leu
        115                 120                 125

Asp Ala Pro Val Gly Ala Gly Phe Ser Tyr Ala Arg Gln Thr Glu Gly
    130                 135                 140

Leu Asn Ser Thr Val Thr Gly Leu Gly Arg His Val Arg Val Phe Leu
145                 150                 155                 160

Gln Lys Trp Met Ala Gln His Pro Glu Leu Ala Ser Asn Pro Leu Tyr
                165                 170                 175

Ile Gly Gly Asp Ser Phe Ser Gly Tyr Thr Val Thr Val Ser Ala Leu
            180                 185                 190

Glu Val Ala Asn His Pro Ala Ala Ser Ser Glu Leu Asn Leu Lys Gly
        195                 200                 205

Tyr Met Val Gly Asn Ala Arg Gly Glu Val Asn Asn Asp Asn Ala Cys
    210                 215                 220

Arg Ile Pro Tyr Leu His Gly Met Gly Leu Ile Ser Asp Glu Leu Tyr
225                 230                 235                 240

Glu Ala Ala Leu Ser Ser Cys Val Val Gly Thr Asp Ser Lys Asn Lys
                245                 250                 255

Gln Gln Gln Ser Ala Ala Arg Cys Ser Glu Ala Gln Gln Ala Ile Ser
            260                 265                 270

Glu Ala Thr Thr Asp Leu Asn Pro Ala His Ile Leu Glu Pro Ala Cys
        275                 280                 285

Gly Ala Asp Phe Ser Pro Arg Ala Pro Tyr Leu Ser Leu Thr Thr Pro
    290                 295                 300
```

Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
305 310 315 320

Ser Tyr Tyr Tyr Leu Ser Leu Ser Ser Val Arg Ser Arg Thr Pro Thr
325 330 335

Lys Glu Met Leu Leu Glu Cys Arg Val Tyr Gly Tyr Glu Leu Ser Tyr
340 345 350

Met Trp Ala Asn Asp Ala Glu Val Arg Glu Asn Leu Gly Val Arg Glu
355 360 365

Gly Thr Ile Gly Asp Gly Asn Trp Ala Leu Cys Pro Glu Val Pro Lys
370 375 380

Leu His Leu Thr Asn Asp Val Pro Thr Thr Val Pro Tyr His Arg Arg
385 390 395 400

Leu Thr Gln Arg Gly Tyr Arg Ala Leu Val Tyr Asn Gly Asp His Asp
405 410 415

Leu Leu Met Thr His Ile Gly Thr His Ala Trp Ile Arg Ser Leu Gly
420 425 430

Tyr Pro Val Val Ala Pro Trp Arg Ala Trp Tyr Ser Asn Asn Glu Val
435 440 445

Ala Gly Phe Thr Val Glu Tyr Ser Asn Asn Leu Thr Phe Ala Thr Val
450 455 460

Lys Gly Ala Gly His Met Ala Pro Glu Ser Arg Pro Lys Gln Cys Leu
465 470 475 480

Asp Met Val Arg Arg Trp Ile Ser Pro Ala Gly Lys Leu
485 490

<210> 37

<211> 444
<212> PRT
<213> Triticum aestivum

<400> 37

```
Val Glu Pro Ser Gly His Ala Ala Asp Arg Ile Ala Arg Leu Pro Gly
1               5                   10                  15


Gln Pro Ala Val Asp Phe Asp Met Tyr Ser Gly Tyr Ile Thr Val Asp
            20                  25                  30


Glu Gly Ala Gly Arg Ser Leu Phe Tyr Leu Leu Gln Glu Ala Pro Glu
        35                  40                  45


Asp Ala Gln Pro Ala Pro Leu Val Leu Trp Leu Asn Gly Gly Pro Gly
    50                  55                  60


Cys Ser Ser Val Ala Tyr Gly Ala Ser Glu Glu Leu Gly Ala Phe Arg
65                  70                  75                  80


Val Lys Pro Arg Gly Ala Gly Leu Val Leu Asn Glu Tyr Arg Trp Asn
                85                  90                  95


Lys Val Ala Asn Val Leu Phe Leu Asp Ser Pro Ala Gly Val Gly Phe
            100                 105                 110


Ser Tyr Thr Asn Thr Ser Ser Asp Ile Tyr Thr Ser Gly Asp Asn Arg
        115                 120                 125


Thr Ala His Asp Ser Tyr Ala Phe Leu Ala Lys Trp Phe Glu Arg Phe
    130                 135                 140


Pro His Tyr Lys Tyr Arg Asp Phe Tyr Ile Ala Gly Glu Ser Tyr Ala
145                 150                 155                 160


Gly His Tyr Val Pro Glu Leu Ser Gln Leu Val His Arg Ser Lys Asn
                165                 170                 175
```

```
Pro Val Ile Asn Leu Lys Gly Phe Met Val Gly Asn Gly Leu Ile Asp
            180                 185                 190

Asp Tyr His Asp Tyr Val Gly Thr Phe Glu Phe Trp Trp Asn His Gly
        195                 200                 205

Ile Val Ser Asp Asp Thr Tyr Arg Arg Leu Lys Glu Ala Cys Leu His
    210                 215                 220

Asp Ser Phe Ile His Pro Ser Pro Ala Cys Asp Ala Ala Thr Asp Val
225                 230                 235                 240

Ala Thr Ala Glu Gln Gly Asn Ile Asp Met Tyr Ser Leu Tyr Thr Pro
                245                 250                 255

Val Cys Asn Ile Thr Ser Ser Ser Ser Ser Ser Ser Ser Ser Leu Ser
            260                 265                 270

Gln Gln Arg Arg Ser Arg Gly Arg Tyr Pro Trp Leu Thr Gly Ser Tyr
        275                 280                 285

Asp Pro Cys Thr Glu Arg Tyr Ser Thr Ala Tyr Tyr Asn Arg Arg Asp
    290                 295                 300

Val Gln Met Ala Leu His Ala Asn Val Thr Gly Ala Met Asn Tyr Thr
305                 310                 315                 320

Trp Ala Thr Cys Ser Asp Thr Ile Asn Thr His Trp His Asp Ala Pro
                325                 330                 335

Arg Ser Met Leu Pro Ile Tyr Arg Glu Leu Ile Ala Ala Gly Leu Arg
            340                 345                 350

Ile Trp Val Phe Ser Gly Asp Thr Asp Ala Val Val Pro Leu Thr Ala
        355                 360                 365
```

Thr Arg Tyr Ser Ile Gly Ala Leu Gly Leu Pro Thr Thr Thr Ser Trp
370 375 380

Tyr Pro Trp Tyr Asp Asp Gln Glu Val Gly Gly Trp Ser Gln Val Tyr
385 390 395 400

Lys Gly Leu Thr Leu Val Ser Val Arg Gly Ala Gly His Glu Val Pro
405 410 415

Leu His Arg Pro Arg Gln Ala Leu Val Leu Phe Gln Tyr Phe Leu Gln
420 425 430

Gly Lys Pro Met Pro Gly Gln Thr Lys Asn Ala Thr
435 440

<210> 38
<211> 532
<212> PRT
<213> Saccharomyces cerevisiae

<400> 38

Met Lys Ala Phe Thr Ser Leu Leu Cys Gly Leu Gly Leu Ser Thr Thr
1 5 10 15

Leu Ala Lys Ala Ile Ser Leu Gln Arg Pro Leu Gly Leu Asp Lys Asp
20 25 30

Val Leu Leu Gln Ala Ala Glu Lys Phe Gly Leu Asp Leu Asp Leu Asp
35 40 45

His Leu Leu Lys Glu Leu Asp Ser Asn Val Leu Asp Ala Trp Ala Gln
50 55 60

Ile Glu Arg Leu Tyr Pro Asn Gln Val Met Ser Leu Glu Thr Ser Thr
65 70 75 80

Lys Pro Lys Phe Pro Glu Ala Ile Lys Thr Lys Lys Asp Trp Asp Phe
85 90 95

```
Val Val Lys Asn Asp Ala Ile Glu Asn Tyr Gln Leu Arg Val Asn Lys
            100                 105                 110

Ile Lys Asp Pro Lys Ile Leu Gly Ile Asp Pro Asn Val Thr Gln Tyr
        115                 120                 125

Thr Gly Tyr Leu Asp Val Glu Asp Glu Asp Lys His Phe Phe Phe Trp
    130                 135                 140

Thr Phe Glu Ser Arg Asn Asp Pro Ala Lys Asp Pro Val Ile Leu Trp
145                 150                 155                 160

Leu Asn Gly Gly Pro Gly Cys Ser Ser Leu Thr Gly Leu Phe Phe Glu
                165                 170                 175

Leu Gly Pro Ser Ser Ile Gly Pro Asp Leu Lys Pro Ile Gly Asn Pro
            180                 185                 190

Tyr Ser Trp Asn Ser Asn Ala Thr Val Ile Phe Leu Asp Gln Pro Val
        195                 200                 205

Asn Val Gly Phe Ser Tyr Ser Gly Ser Ser Gly Val Ser Asn Thr Val
    210                 215                 220

Ala Ala Gly Lys Asp Val Tyr Asn Phe Leu Glu Leu Phe Phe Asp Gln
225                 230                 235                 240

Phe Pro Glu Tyr Val Asn Lys Gly Gln Asp Phe His Ile Ala Gly Glu
                245                 250                 255

Ser Tyr Ala Gly His Tyr Ile Pro Val Phe Ala Ser Glu Ile Leu Ser
            260                 265                 270

His Lys Asp Arg Asn Phe Asn Leu Thr Ser Val Leu Ile Gly Asn Gly
        275                 280                 285
```

```
Leu Thr Asp Pro Leu Thr Gln Tyr Asn Tyr Tyr Glu Pro Met Ala Cys
    290                 295                 300

Gly Glu Gly Gly Glu Pro Ser Val Leu Pro Ser Glu Glu Cys Ser Ala
305                 310                 315                 320

Met Glu Asp Ser Leu Glu Arg Cys Leu Gly Leu Ile Glu Ser Cys Tyr
                325                 330                 335

Asp Ser Gln Ser Val Trp Ser Cys Val Pro Ala Thr Ile Tyr Cys Asn
            340                 345                 350

Asn Ala Gln Leu Ala Pro Tyr Gln Arg Thr Gly Arg Asn Val Tyr Asp
        355                 360                 365

Ile Arg Lys Asp Cys Glu Gly Gly Asn Leu Cys Tyr Pro Thr Leu Gln
    370                 375                 380

Asp Ile Asp Asp Tyr Leu Asn Gln Asp Tyr Val Lys Glu Ala Val Gly
385                 390                 395                 400

Ala Glu Val Asp His Tyr Glu Ser Cys Asn Phe Asp Ile Asn Arg Asn
                405                 410                 415

Phe Leu Phe Ala Gly Asp Trp Met Lys Pro Tyr His Thr Ala Val Thr
            420                 425                 430

Asp Leu Leu Asn Gln Asp Leu Pro Ile Leu Val Tyr Ala Gly Asp Lys
        435                 440                 445

Asp Phe Ile Cys Asn Trp Leu Gly Asn Lys Ala Trp Thr Asp Val Leu
    450                 455                 460

Pro Trp Lys Tyr Asp Glu Glu Phe Ala Ser Gln Lys Val Arg Asn Trp
465                 470                 475                 480
```

Thr Ala Ser Ile Thr Asp Glu Val Ala Gly Glu Val Lys Ser Tyr Lys
485 490 495

His Phe Thr Tyr Leu Arg Val Phe Asn Gly Gly His Met Val Pro Phe
500 505 510

Asp Val Pro Glu Asn Ala Leu Ser Met Val Asn Glu Trp Ile His Gly
515 520 525

Gly Phe Ser Leu
530

What is claimed is:

1. An engineered fungal cell that produces a tropane alkaloid product, wherein the engineered fungal cell comprises a plurality of heterologous coding sequences encoding a plurality of enzymes within a pathway for producing the tropane alkaloid product;

wherein the cell comprises a plurality of heterologous coding sequences encoding a plurality of transporter proteins selected from the group of purine uptake permease-like transporter, multidrug resistance transporter, pleiotropic drug resistance transporter, ATP-binding cassette transporter, multidrug and toxin extrusion transporter, nitrate transporter 1/peptide transporter family transporter, and major facilitator superfamily transporter; and

wherein the plurality of heterologous coding sequences encoding a plurality of transporter proteins comprises a heterologous coding sequence encoding a purine uptake permease-like transporter.

2. The cell of claim 1, wherein the cell is a yeast cell.

3. The cell of claim 1 or claim 2, wherein the engineered cell comprises one or more heterologous coding sequences for one or more enzymes, wherein at least one of the enzymes is selected from the group consisting of arginine decarboxylase, agmatine ureohydrolase, agmatinase, putrescine N-methyltransferase, N-methylputrescine oxidase, pyrrolidine ketide synthase, tropinone synthase, cytochrome P450 reductase, tropinone reductase, phenylalanine ammonia-lyase, tyrosine ammonia-lyase, phenylpyruvate reductase, 4-coumarate-CoA ligase, 3-phenyllactic acid UDP-glucosyltransferase 84A27, littorine synthase, littorine mutase, hyoscyamine dehydrogenase, hyoscyamine 6β-hydroxylase/dioxygenase, cocaine synthase, a metazoan hepatic cytochrome P450, senecionine *N*-oxygenase, pyrrolizidine *N*-oxygenase, and a tropane alkaloid-detoxifying enzyme.

4. The cell of any of claims 1-3, wherein the engineered cell comprises one or more heterologous coding sequences encoding one or more multidrug resistance (MDR) transporters or pleiotropic drug resistance (PDR) transporters.

5. The cell of any of claims 1-4, wherein the engineered cell comprises one

or more heterologous coding sequences encoding one or more ATP-binding cassette (ABC) transporters.

6. The cell of any of claims 1-5, wherein:

(a) the engineered cell comprises one or more heterologous coding sequences encoding one or more multidrug and toxin extrusion (MATE) transporters;

(b) the engineered cell comprises one or more heterologous coding sequences encoding one or more nitrate transporter 1/peptide transporter family (NPF/NRT) transporters;

and/or

(c)the engineered cell comprises one or more heterologous coding sequences encoding one or more major facilitator superfamily (MFS) transporters.

7. The cell of any of claims 1-6, wherein the cell produces a tropane alkaloid product selected from the group consisting of a littorine, hyoscyamine, atropine, anisodamine, scopolamine, calystegine, cocaine, or a non-natural tropane alkaloid.

8. A method for producing a tropane alkaloid product comprising

(a) culturing a cell of any of claims 1-7 under conditions suitable for protein production;

(b) adding a starting compound to the cell culture; and

(c) recovering the tropane alkaloid product from the culture.

9. The method of claim 8, wherein:

(a) the cells are cultured in a fed-batch or batch fermentation;

(b) the starting compound added to the cell culture is a sugar or a substrate which contains one or more sugars, or which is converted to one or more sugars during fermentation;

(c) the starting compound added to the cell culture is an amino acid or a mixture comprising one or more amino acids, or a substrate which is converted to one or more amino acids during fermentation;

(d) the starting compound added to the cell culture is a precursor of a tropane alkaloid product; or

(e) the tropane alkaloid product is recovered via a process comprising a liquid-

liquid extraction, chromatography separation, distillation, or recrystallization.