



US 20190127770A1

(19) **United States**(12) **Patent Application Publication****Siddiqui et al.**(10) **Pub. No.: US 2019/0127770 A1**(43) **Pub. Date: May 2, 2019**

(54) **BENZYLISOQUINOLINE ALKALOID (BIA)  
PRECURSOR PRODUCING MICROBES, AND  
METHODS OF MAKING AND USING THE  
SAME**

*C12P 7/24* (2006.01)*C12N 15/52* (2006.01)*C12N 9/04* (2006.01)*C12N 9/10* (2006.01)*C12N 9/02* (2006.01)*C12N 9/90* (2006.01)

(71) Applicant: **The Board of Trustees of the Leland  
Stanford Junior University**, Stanford,  
CA (US)

(52) **U.S. Cl.**CPC ..... *C12P 17/12* (2013.01); *C12P 13/225*(2013.01); *C12P 13/001* (2013.01); *C12P 7/24*(2013.01); *C12N 9/90* (2013.01); *C12N**9/0006* (2013.01); *C12N 9/1085* (2013.01);*C12N 9/0071* (2013.01); *C12N 9/0008*(2013.01); *C12N 15/52* (2013.01)

(72) Inventors: **Michael Shareef Siddiqui**, San Mateo,  
CA (US); **Christina D. Smolke**, Menlo  
Park, CA (US)

(21) Appl. No.: **16/224,217**(22) Filed: **Dec. 18, 2018****Related U.S. Application Data**

(63) Continuation of application No. 15/031,618, filed on  
Apr. 22, 2016, filed as application No. PCT/US2014/  
063738 on Nov. 3, 2014.

(60) Provisional application No. 61/899,496, filed on Nov.  
4, 2013.

**Publication Classification**(51) **Int. Cl.***C12P 17/12* (2006.01)*C12P 13/22* (2006.01)*C12P 13/00* (2006.01)

(57)

**ABSTRACT**

Host cells that are engineered to produce benzyloquinoline alkaloid (BIAs) precursors, such as norcodaurine (NC) and norlaudanosoline (NL), are provided. The host cells may have one or more engineered modifications selected from: a feedback inhibition alleviating mutation in a enzyme gene; a transcriptional modulation modification of a biosynthetic enzyme gene; an inactivating mutation in an enzyme; and a heterologous coding sequence. Also provided are methods of producing a BIA of interest or a precursor thereof using the host cells and compositions, e.g., kits, systems etc., that find use in methods of the invention.

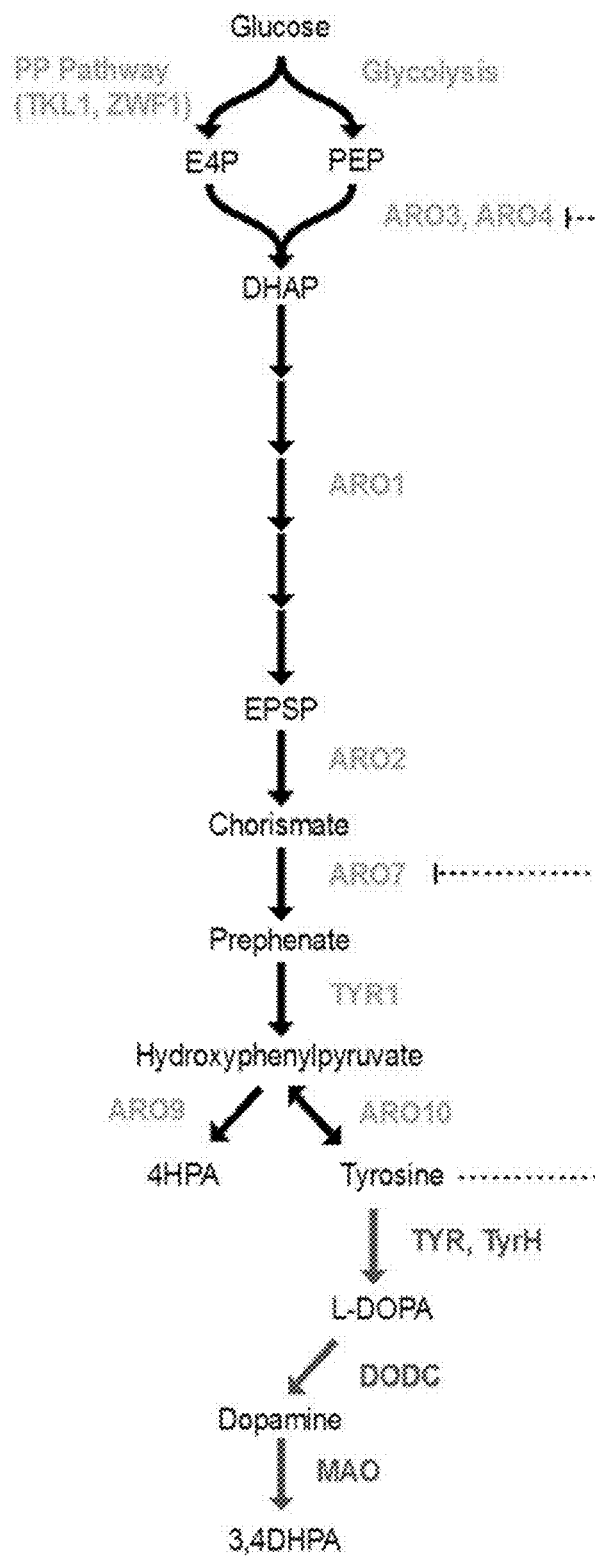
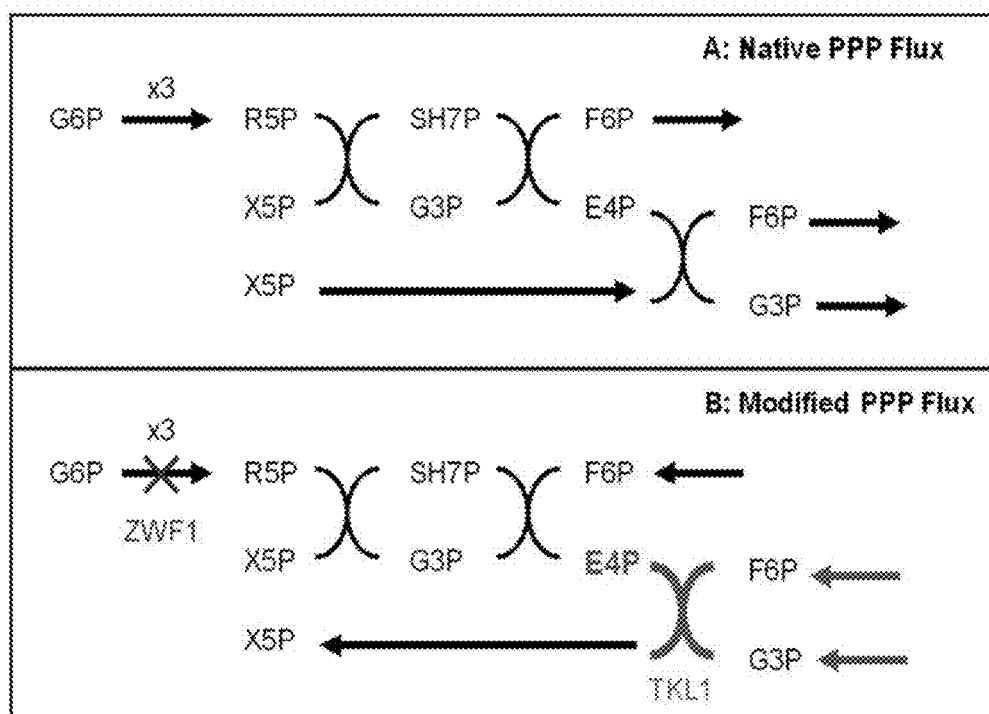


FIG. 1



**FIG. 2**

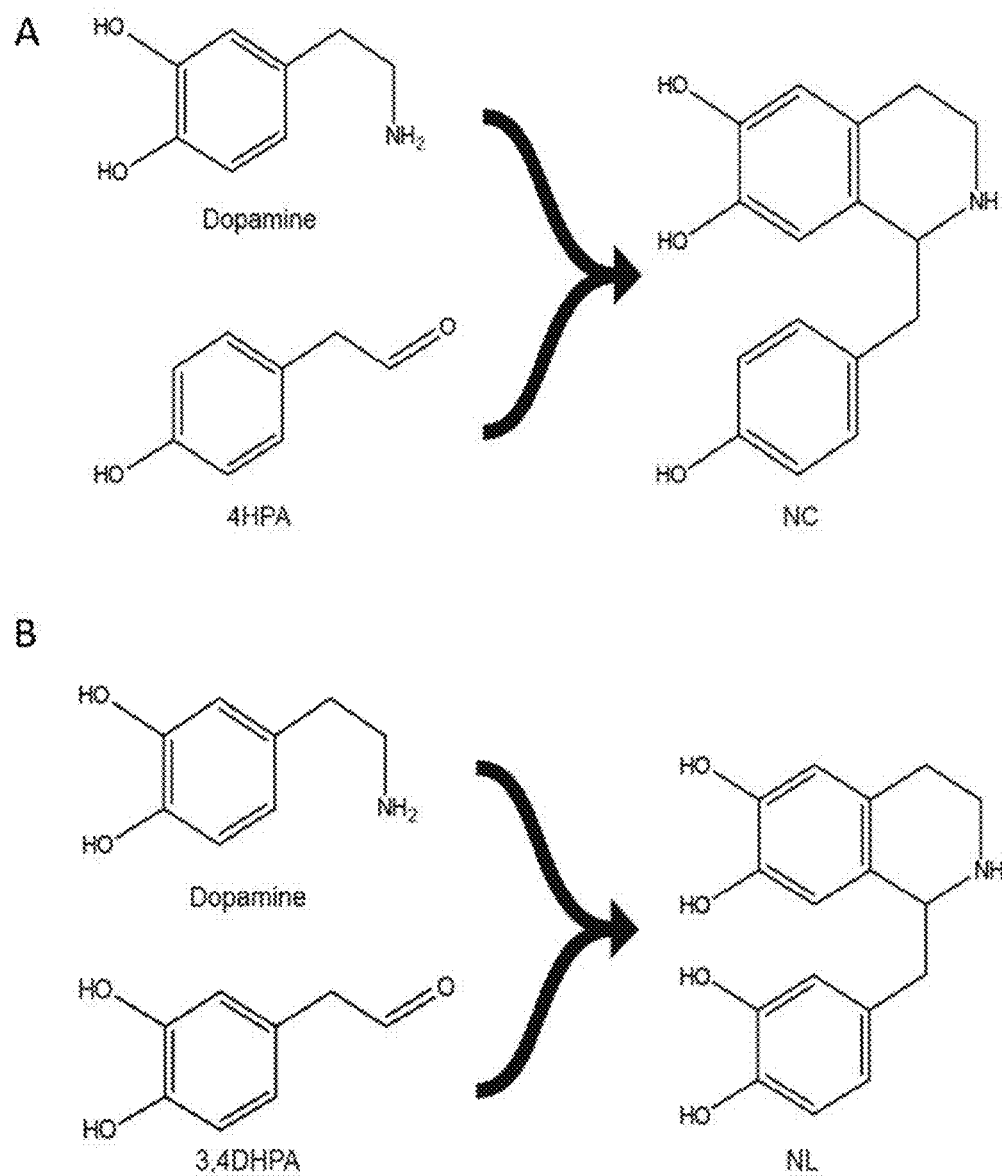


FIG. 3

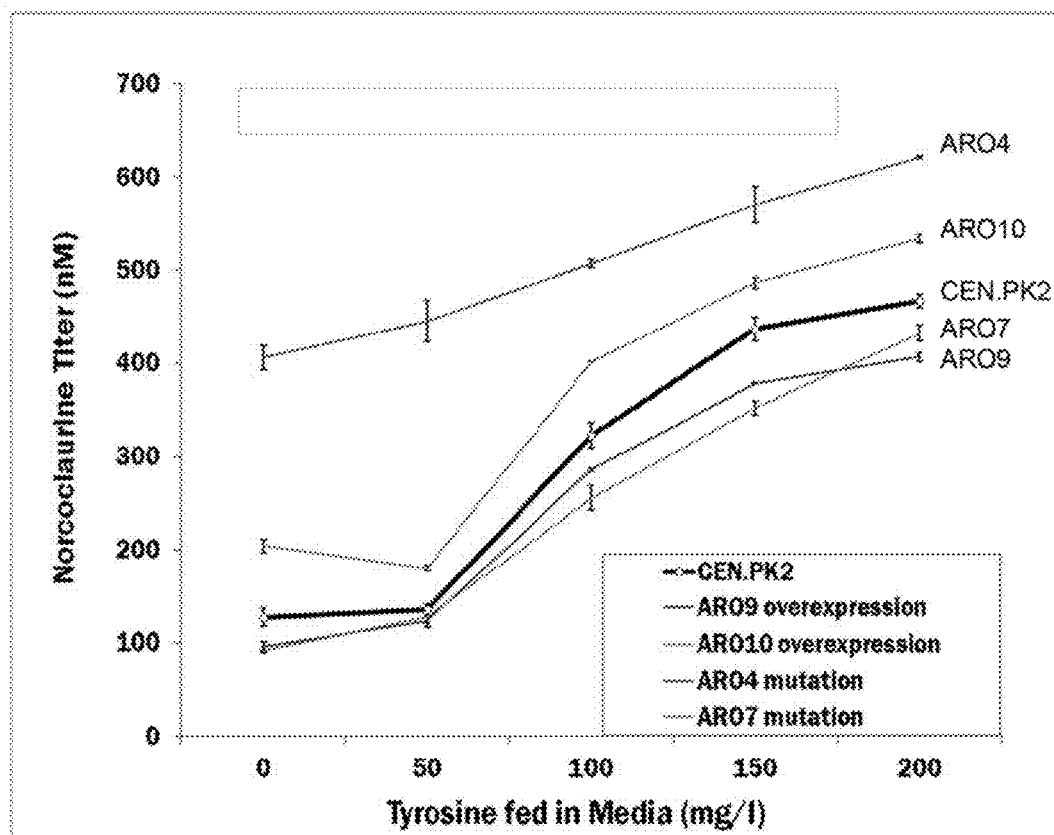


FIG. 4

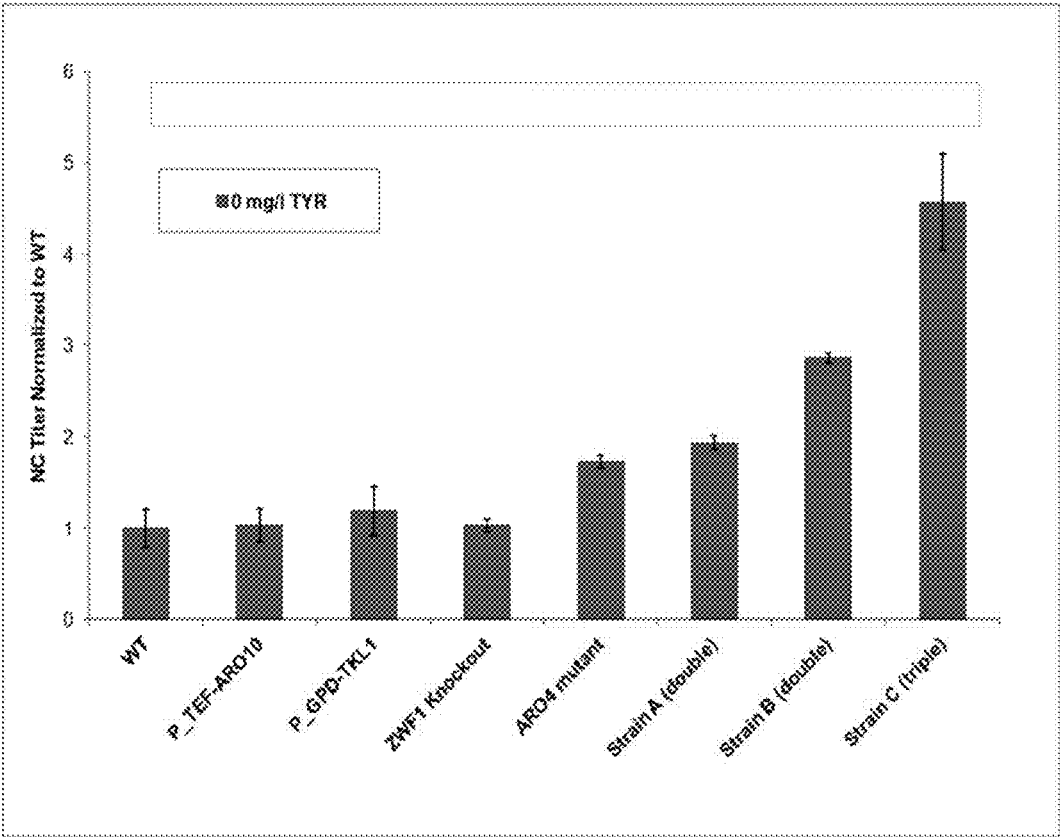


FIG. 5

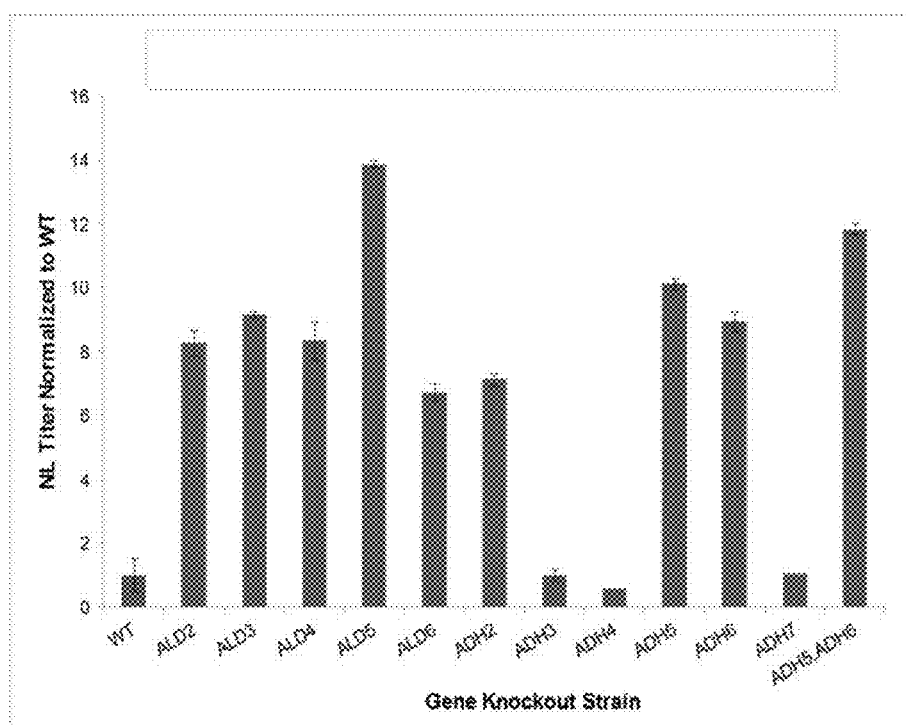
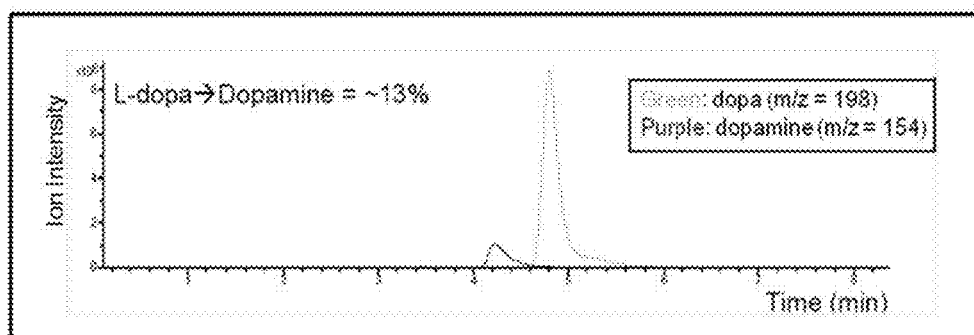


FIG. 6

**FIG. 7**



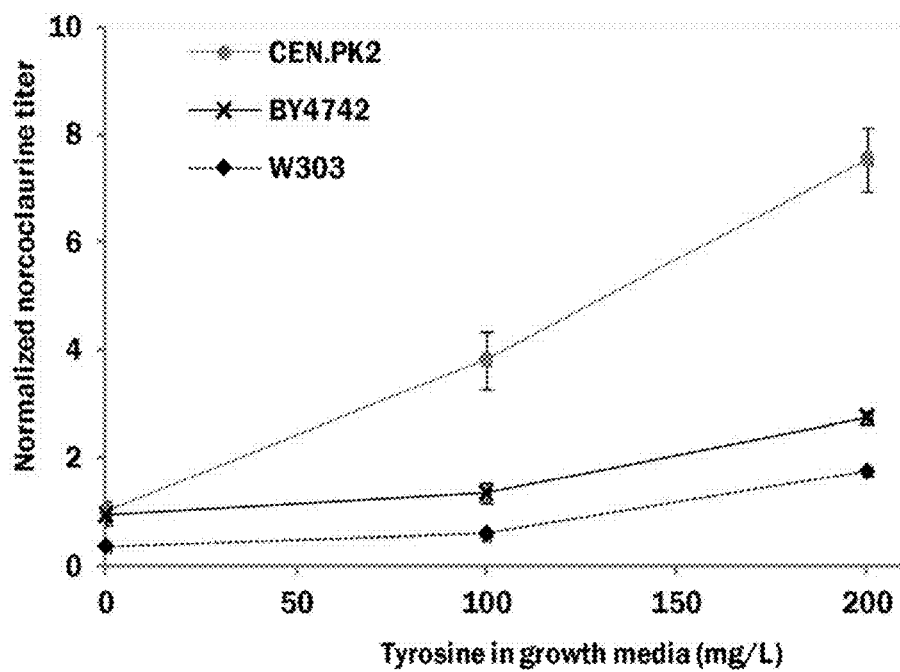
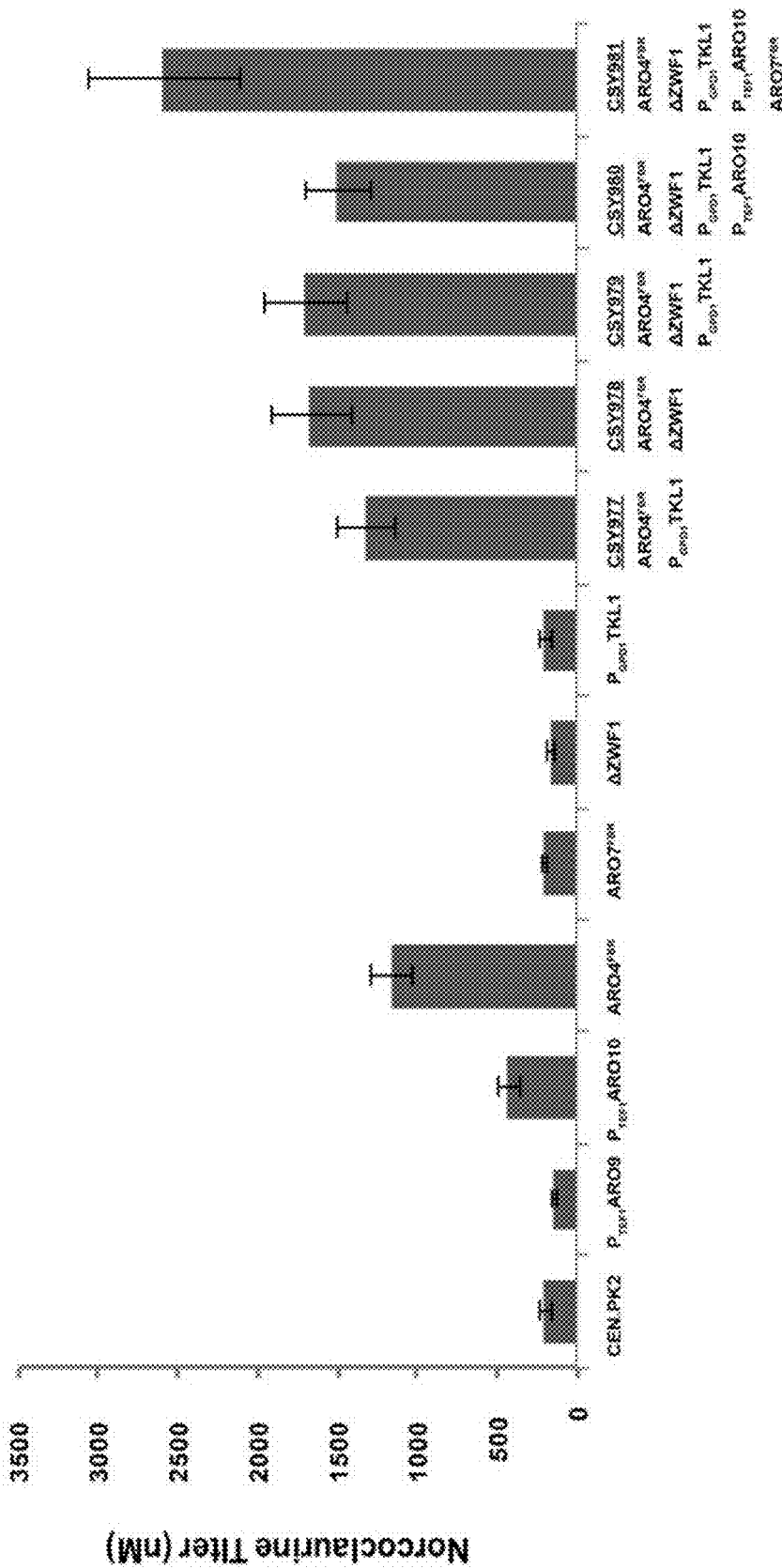


FIG. 8

FIG. 9



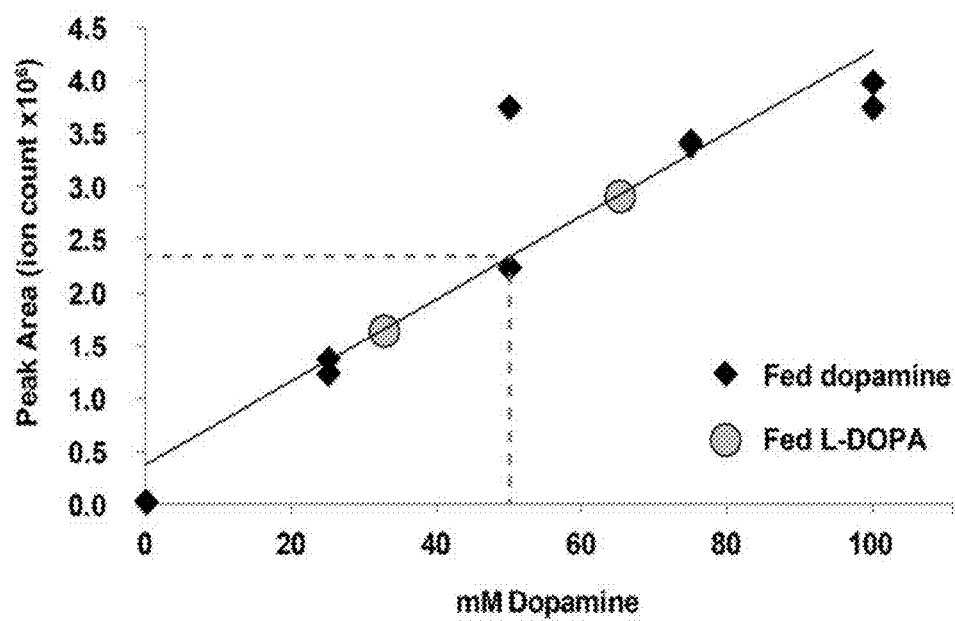


FIG. 10

FIG. 11

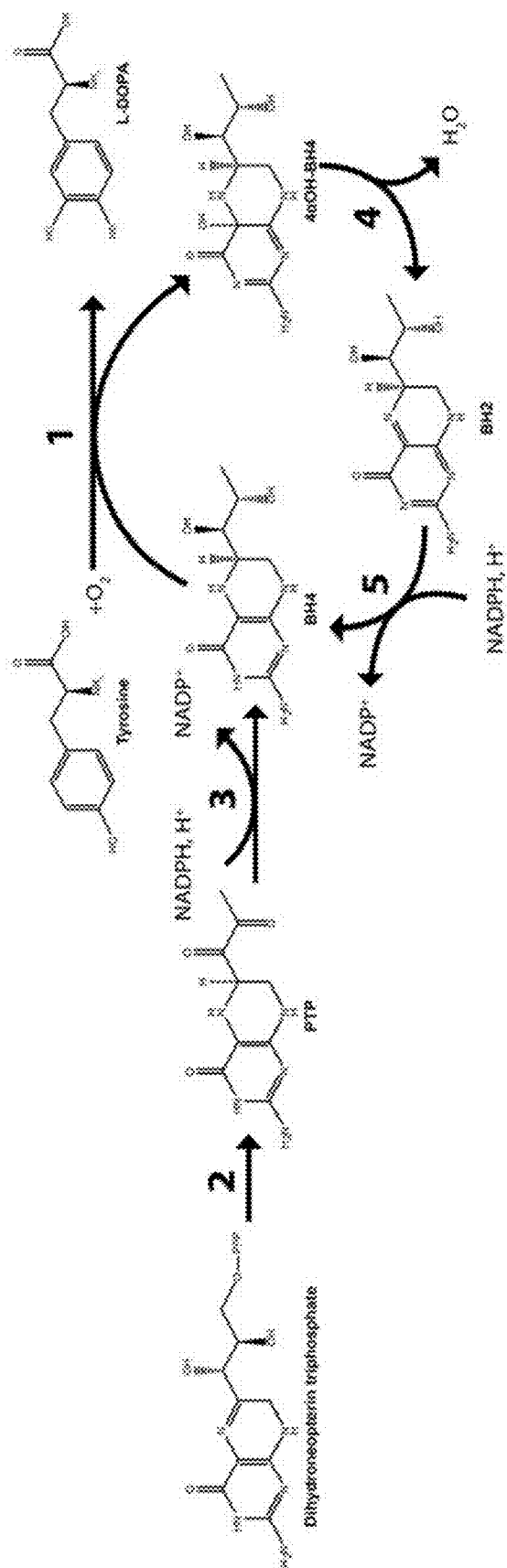


FIG. 12

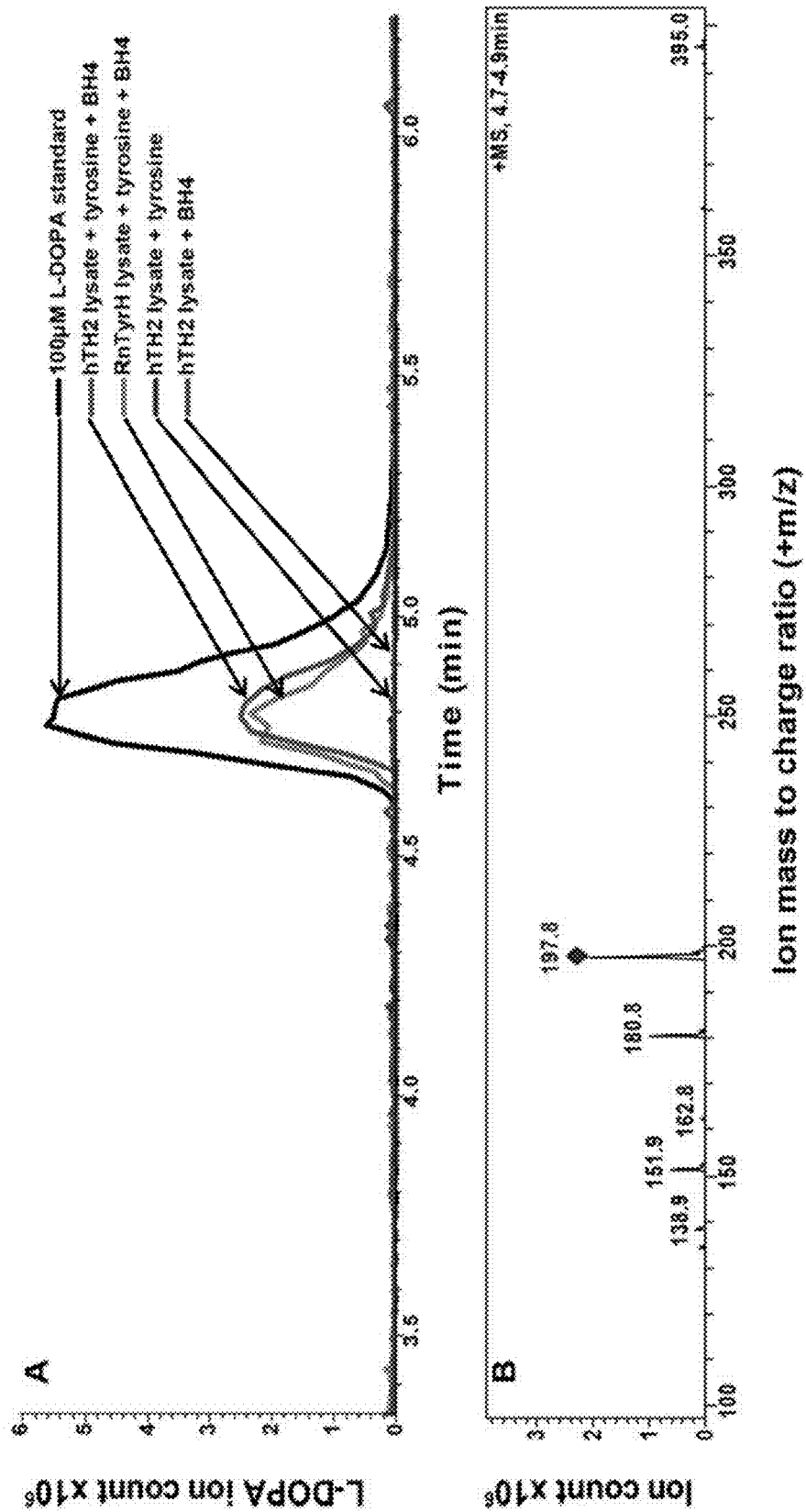
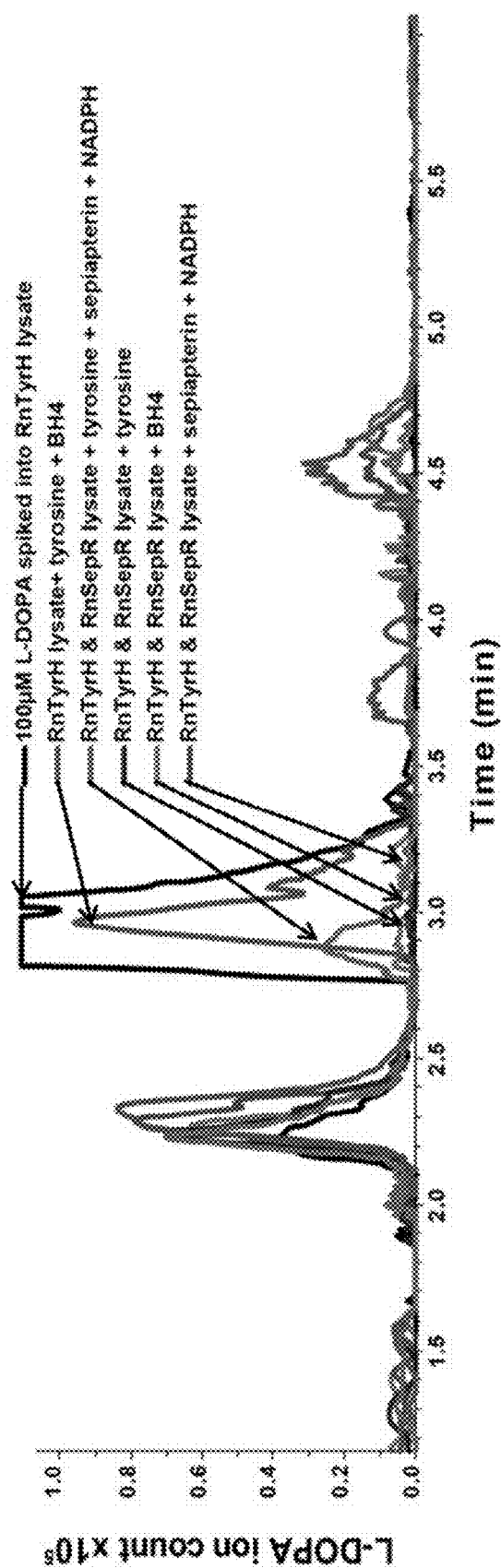


FIG. 13



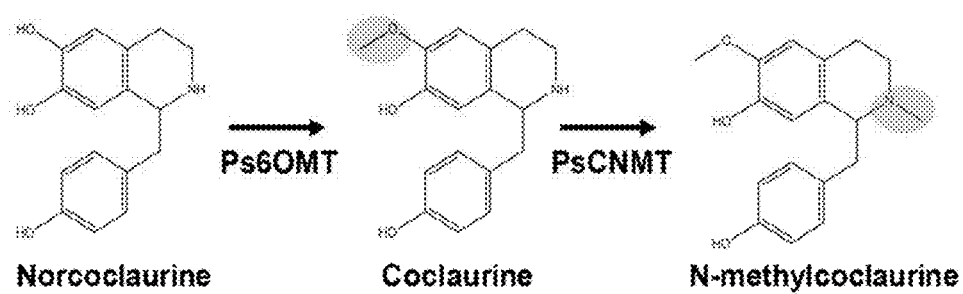


FIG. 14

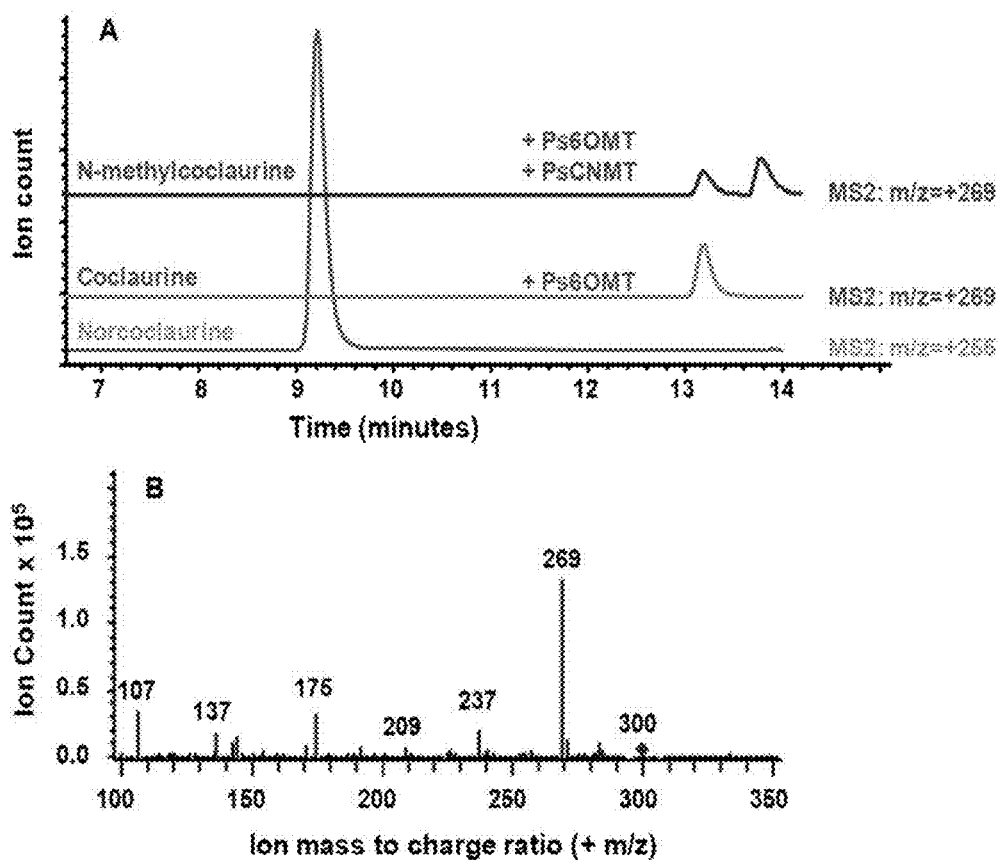


FIG. 15



# BENZYLISOQUINOLINE ALKALOID (BIA) PRECURSOR PRODUCING MICROBES, AND METHODS OF MAKING AND USING THE SAME

## CROSS-REFERENCE TO RELATED APPLICATION

[0001] Pursuant to 35 U.S.C. § 119 (e), this application claims priority to the filing date of U.S. Provisional Patent Application Ser. No. 61/899,496 filed on Nov. 4, 2013; the disclosure of which application is herein incorporated by reference.

## GOVERNMENT RIGHTS

[0002] This invention is made with Government support under grant No. 1066100 awarded by the National Science Foundation. The Government has certain rights in this invention.

## INTRODUCTION

[0003] Benzylisoquinoline alkaloids (BIAs) are a large group of secondary metabolites from plants and other organisms. These molecules have therapeutic functions in the human body, ranging from the established analgesic and antitussive properties of morphine and codeine, to novel activities against cancer and infection observed for molecules such as berberine and sanguinarine. Supply of all these BIA molecules so that they are available to researchers and physicians is of interest. The number of synthetic reactions and requirements for selective stereochemistry means that chemical synthesis of BIAs is low yielding and not a viable means for large-scale production. Instead, for the widely used drugs codeine and morphine, the opium poppy (*Papaver somniferum*) has been bred and developed as a production crop. Intermediates in morphine biosynthesis that find use as drugs and drug precursors do not accumulate because the plant metabolism is evolved to maximize pathway flux to the final opioids. Even for end product metabolites like morphine, accumulation occurs only within specialized cells in the buds and vascular tissue and requires harsh chemical processing of harvested plant material during the extraction process, which may yield less than 2% morphine by dry weight. As such, methods for preparing BIAs are of interest.

## SUMMARY

[0004] Host cells that are engineered to produce benzylisoquinoline alkaloid (BIA) precursors, such as norcoclaurine (NC) and norlaudanosoline (NL), are provided. 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 gene; an inactivating mutation in an enzyme; and a heterologous coding sequence. Also provided are methods of producing a BIA of interest or a precursor thereof using the host cells and compositions, e.g., kits, systems etc., that find use in methods of the invention.

## BRIEF DESCRIPTION OF THE FIGURES

[0005] 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.

[0006] FIG. 1 illustrates the biosynthetic pathway from glucose to tyrosine and other BIA precursor molecules.

[0007] FIG. 2 illustrates the effect of ZWF1 knockout and TKL1 over-expression on the pentose phosphate pathway (PPP). A: native PPP flux, B: Modified PPP flux.

[0008] FIG. 3 illustrates the synthesis of NC (A) and NL (B) from precursor molecules.

[0009] FIG. 4 illustrates the effect of four genetic modifications on NC production with varying fed tyrosine.

[0010] FIG. 5 shows NC production from strains with combinations of genetic modifications.

[0011] FIG. 6 shows the levels of NL production in aldehyde oxidoreductase (ALD) alcohol dehydrogenase (ADH) gene knockout strains FIG. 7 illustrates the activity of a L-DOPA decarboxylase (DODC) enzyme in vivo. Yeast strains transformed with DNA to express *Papaver somniferum* tyrosine/DOPA decarboxylase can convert L-DOPA to dopamine in vivo.

[0012] FIG. 8 shows the production of norcoclaurine (NC) in yeast strains fed 100 mM dopamine and varying concentrations of tyrosine.

[0013] FIG. 9 shows NC production in multiple engineered yeast strains fed 100 mM dopamine and no tyrosine.

[0014] FIG. 10 shows NC production from dopamine or from L-DOPA in an engineered yeast strain (CSY980) with the additional integration of the L-DOPA decarboxylase PpDODC.

[0015] FIG. 11 illustrates a biosynthetic scheme including tyrosine hydroxylation using mammalian tyrosine hydroxylases (TyrHs) with the co-substrate tetrahydrobiopterin (BH4).

[0016] FIG. 12 shows that tyrosine hydroxylases expressed from yeast cells convert tyrosine to L-DOPA: (A) LC-MS chromatogram confirms conversion of tyrosine to L-DOPA in the presence of co-substrate, BH4; and (B) L-DOPA ion fragmentation in lysate samples.

[0017] FIG. 13 shows the co-expression of tyrosine hydroxylase with a BH4 biosynthetic enzyme provides for conversion of tyrosine to L-DOPA.

[0018] FIG. 14 illustrates the synthesis of the BIA precursor molecules coclaurine and N-methylcoclaurine from NC.

[0019] FIG. 15 shows LC-MS analysis (A: ion counts) of the production of NC-derived BIA precursor molecules including N-methylcoclaurine (B: m/z fragmentation pattern) from L-DOPA in the liquid culture of engineered yeast strains.

## DEFINITIONS

[0020] 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.

[0021] 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.

**[0022]** It must be 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 is 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.

**[0023]** As used herein, the terms “determining,” “measuring,” and “assessing,” and “assaying” are used interchangeably and include both quantitative and qualitative determinations.

**[0024]** 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.

**[0025]** 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.

**[0026]** 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.

**[0027]** 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 can 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.

**[0028]** 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 cassettes 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.

**[0029]** 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, 100 or more,  $10^7$  or more,  $10^8$  or more or  $10^9$  or more members.

**[0030]** Numeric ranges are inclusive of the numbers defining the range.

**[0031]** 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.

**[0032]** Other definitions of terms may appear throughout the specification.

#### DETAILED DESCRIPTION

**[0033]** Host cells that are engineered to produce benzyloquinoline alkaloid (BIAs) precursors, such as norco-claurine (NC) and norlaudanosoline (NL), are 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 gene; an inactivating mutation in an enzyme; and a heterologous coding sequence. Also provided are methods of producing a BIA of interest or a precursor thereof using the host cells and compositions, e.g., kits, systems etc., that find use in methods of the invention.

**[0034]** Before the present invention is described in greater detail, it is to be understood that this invention is not limited to particular embodiments described, as such may, of course, 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.

**[0035]** 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.

**[0036]** 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.

**[0037]** 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 can also be used in the practice or testing of the present invention, representative illustrative methods and materials are now described.

**[0038]** 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.

**[0039]** 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. It is further noted that the claims may be 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.

**[0040]** 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.

**[0041]** In further describing the subject invention, benzyloquinoline alkaloid precursors of interest 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.

#### Benzyloquinoline Alkaloid (BIA) Precursors

**[0042]** As summarized above, host cells which produce benzyloquinoline alkaloid precursors (BIA precursors) are provided. The BIA precursor may be any intermediate or precursor compound in a synthetic pathway (e.g., as described herein) that leads to the production of a BIA of interest (e.g., as described herein). In some cases, the BIA precursor has a structure that may be characterized as a BIA or a derivative thereof. In certain cases, the BIA precursor has a structure that may be characterized as a fragment of a BIA. In some cases, the BIA precursor is an early BIA. As used herein, by “early BIA” is meant an early intermediate in the synthesis of a BIA of interest in a cell, where the early BIA is produced by a host cell from a host cell feedstock or

simple starting compound. In some cases, the early BIA is a BIA 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 BIA may refer to a precursor of a BIA end product of interest whether or not the early BIA can be itself be characterized as a benzyloquinoline alkaloid.

**[0043]** In some cases, the BIA precursor is an early BIA, such as a pre-reticuline benzyloquinoline alkaloid. As such, host cells which produce pre-reticuline benzyloquinoline alkaloids (pre-reticuline BIAs) are provided. Reticuline is a major branch point intermediate of interest in the synthesis of downstream BIAs via cell engineering efforts to produce end products such as opioid products. The subject host cells may produce BIA precursors from simple and inexpensive starting materials that may find use in the production of reticuline and downstream BIA end products.

**[0044]** As used herein, the terms “pre-reticuline benzyloquinoline alkaloid”, “pre-reticuline BIA” and “pre-reticuline BIA precursor” are used interchangeably and refer to a biosynthetic precursor of reticuline whether or not the structure of the reticuline precursor itself is characterized as a benzyloquinoline alkaloid. The term pre-reticuline BIA is meant to include biosynthetic precursors, intermediates and metabolites thereof, of any convenient member of a host cell biosynthetic pathway that may lead to reticuline. In some cases, the pre-reticuline BIA includes a benzyloquinoline alkaloid fragment, such as a benzyl fragment, a quinoline fragment or a precursor or derivative thereof. In certain instances, the pre-reticuline BIA has a structure that can be characterized as a benzyloquinoline alkaloid or a derivative thereof.

**[0045]** BIA precursors of interest include, but are not limited to, norcoclaurine (NC) and norlaudanosoline (NL), as well as NC and NL precursors, such as tyrosine, 4-hydroxyphenylacetaldehyde (4-HPA), 4-hydroxyphenylpyruvic acid (4-HPPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA), and dopamine. In some embodiments, the one or more BIA precursors are 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) and dopamine. In certain instances, the one or more BIA precursors are 4-hydroxyphenylacetaldehyde (4-HPA) and dopamine. FIGS. 3A and 3B illustrate the synthesis of NC and NL, respectively from precursor molecules via a Pictet-Spengler condensation reaction, where the reaction may occur spontaneously or may be catalyzed by any convenient enzymes.

**[0046]** Synthetic pathways to a BIA precursor may be generated in the host cells, and may start with any convenient starting compound(s) or materials. FIG. 1 illustrates a synthetic pathway of interest to BIA precursors starting from glucose. 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., tyrosine, 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, dopamine, 4-HPA, 4-HPPA, as well as any of the compounds shown in FIG. 1. 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.

#### Host Cells

**[0047]** As summarized above, one aspect of the invention is a host cell that produces one or more BIA precursors. 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 yeast cells. Any convenient type of host cell may be utilized in producing the subject BIA-producing cells, see, e.g., US2008/0176754, and US2014/0273109 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* and *Salmonella typhimurium* cells, insect cells such as *Drosophila melanogaster* S2 and *Spodoptera frugiperda* Sf9 cells and yeast cells such as *S. cerevisiae* cells, *Schizosaccharomyces pombe* cells and a *Pichia pastoris* 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 BIA precursor of interest. Any of the host cells described in US200810176754, and US2014/0273109 by Smolke et al. may be adapted for use in the subject cells and methods. In certain embodiments, the yeast cells can be of the species *Saccharomyces cerevisiae* (*S. cerevisiae*). In certain embodiments, the yeast cells can be of the species *Schizosaccharomyces pombe*. In certain embodiments, the yeast cells can 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: MATa/ $\alpha$  ura3-52/ura3-52 trp1-28 trp-289 leu2-3<sub>112</sub>/leu2-3<sub>112</sub> his3  $\Delta$ 1/his3  $\Delta$ 1 MAL2-8C/MAL2-8C SUC2/SUC2), S288C, W303, D273-108, X2180, A364A,  $\Sigma$ 1278B, AB972, SK1 and FL100. In certain cases, the yeast strain is any of S288C (MATa; SUC2 mal mel gal2 CUP1 flo8-1 hap1), BY4741 (MATa; his3 $\Delta$ 1; leu2 $\Delta$ 0; met15 $\Delta$ 0; ura3 $\Delta$ 0), BY4742 (MATa; his3 $\Delta$ 1; leu2 $\Delta$ 0; lys2 $\Delta$ 0; ura3 $\Delta$ 0), BY4743 (MATa/MATa; his3 $\Delta$ 1/his3 $\Delta$ 1; leu2 $\Delta$ 0/leu2 $\Delta$ 0; met15 $\Delta$ 0/MET15; LYS2/lys2 $\Delta$ 0; ura3 $\Delta$ 0/ura3 $\Delta$ 0), 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 (MATa; his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1). The identity and genotype of additional yeast strains of interest can be found at EUROSCARF ([web.uni-frankfurt.de/fb15/mikro/euroscarf/col\\_index.html](http://web.uni-frankfurt.de/fb15/mikro/euroscarf/col_index.html)).

**[0048]** 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 BIA precursors 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 gene native to the cell; an inactivating mutation in an enzyme native to the cell; and a heterologous coding sequence that encodes an enzyme. A cell that includes one or more modifications may be referred to as a modified cell.

**[0049]** A modified cell may overproduce one or more BIA precursor molecules. By overproduce is meant that the cell has an improved or increased production of a BIA precursor 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 BIA precursor of interest where the control has no BIA 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 BIA precursor production.

**[0050]** In some cases, the host cell is capable of producing an increased amount of norcodaurine 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 norcodaurine 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.

**[0051]** In some cases, the host cell is capable of producing an increased amount of norlaudanosoline 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 norlaudanosoline 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.

**[0052]** In some embodiments, the host cell is capable of producing a 10% or more yield of norcodaurine from a starting compound such as tyrosine, 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 norcodaurine from a starting compound.

**[0053]** In some embodiments, the host cell is capable of producing a 10% or more yield of norlaundonosoline from a starting compound such as tyrosine, 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 norlaundonosoline from a starting compound.

**[0054]** In some embodiments, the host cell overproduces one or more BIA precursor molecule selected from the group consisting of tyrosine, 4-hydroxyphenylacetaldehyde (4-HPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA), and dopamine.

**[0055]** 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.

**[0056]** In some embodiments of the host cell, when the cell includes one or more heterologous coding sequences that encode one or more enzymes, 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 gene native to the cell; and an inactivating mutation in an enzyme 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 gene native to the cell; an inactivating mutation in an enzyme native to the cell; and a heterologous coding sequence that encode an enzyme. In some embodiments of the host cell, when the cell includes one or more transcriptional modulation modifications of one or more biosynthetic enzyme 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 native to the cell; and a heterologous coding sequence that encodes an enzyme. In certain instances of the host cell, when the cell includes one or more inactivating mutations in one or more enzymes 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 gene native to the cell; and a heterologous coding sequence that encodes an enzyme.

**[0057]** 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 gene 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 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

**[0058]** 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 synthetic pathway of FIG. 1 or FIG. 2. 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.

**[0059]** A variety of feedback inhibition control mechanisms and biosynthetic enzymes native to the host cell that are directed to regulation of levels of BIA 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 a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase and a chorismate mutase. In some embodiments, the one or more biosynthetic enzyme genes encode a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase. In some instances, the one or more biosynthetic enzyme genes encode a chorismate mutase. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene selected from ARO4 and ARO7. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is ARO4. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is ARO7. 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.

**[0060]** 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 2 $\mu$  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.

**[0061]** In certain embodiments, the one or more feedback inhibition alleviating mutations are present in the ARO4 gene. ARO4 mutations of interest include, but are not limited to, substitution of the lysine residue at position 229 with a leucine, a substitution of the glutamine residue at position 166 with a lysine residue, or a mutation as described by Hartmann M, et al. ((2003) *Proc Natl Acad Sci USA* 100(3):862-887) or Fukuda et al. ((1992) *J Ferment Bioeng* 74(2):117-119). In some instances, mutations for conferring feedback inhibition are selected from a mutagenized library of enzyme mutants. Examples of such selections include rescue of growth of o-fluoro-D,L-phenylalanine or growth of aro3 mutant yeast strains in media with excess tyrosine as described by Fukuda et al. ((1990) *Breeding of Brewing Yeast Producing a Large Amount of Beta-Phenylethyl Alcohol and Beta-Phenylethyl Acetate*. *Agr Biol Chem Tokyo* 54(1):269-271).

**[0062]** ARO7 mutations of interest include, but are not limited to, substitution of the threonine residue at position 226 with an isoleucine, as described by Schmidheini et al. ((1989), *J Bacteriol* 171(3):1245-1253) and additional mutations conferring feedback inhibition selected from a mutagenized library of microbial chorismate mutase mutants. Examples of such selections include assays for 5-methyl-tryptophan sensitivity or increased production of melanin pigments in strains expressing heterologous tyrosinase enzymes (1.9) in the absence of externally fed tyrosine.

**[0063]** 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

**[0064]** 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 genes of the cell. In some cases, the one or more biosynthetic enzyme genes are native to the cell. Any convenient biosynthetic enzyme 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 (Ser. No. 14/211,611) by Smolke et al., the disclosure of which is herein incorporated by reference in its entirety.

**[0065]** Any convenient biosynthetic enzyme genes may be transcriptionally modulated, and include but are not limited to, those biosynthetic enzymes described in FIG. 1, such as ARO3, ARO4, ARO1, ARO7, TYR1, TYR, TyrH, DODC, MAO, ARO10, ARO9 and TKL. In some instances, the one or more biosynthetic enzyme genes is selected from ARO10, ARO9 and TKL. In some cases, the one or more biosynthetic enzyme genes is ARO10. In certain instances, the one or more biosynthetic enzyme genes is ARO9. In some embodiments, the one or more biosynthetic enzyme genes is TKL. 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 synthetic pathway of one of FIGS. 1 and 2.

**[0066]** In some embodiments, the transcriptional modulation modification includes substitution of a strong promoter for a native promoter of the one or more biosynthetic enzyme genes. The promoters driving expression of the genes of interest can be constitutive promoters or inducible promoters, provided that the promoters can 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, 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 glucocorticoid 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 can be used to optimize transcription, and hence, enzyme levels to maximize production while minimizing energy resources.

#### Inactivating Mutations

**[0067]** The host cells may include one or more inactivating mutations to an enzyme 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 to increase the levels of a BIA precursor 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 native to the cell FIG. 2 illustrates a native pentose phosphate pathway (PPP) flux and modified PPP flux where that involves inactivation of ZWF1 enzyme. 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 that is inactivated and is part of or connected to the synthetic pathway of a BIA precursor 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, 80% 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 activity of the enzyme.

**[0068]** In some embodiments, the cell includes an inactivating mutation in an enzyme native to the cell. Any convenient enzymes may be targeted for inactivation. Enzymes of interest include, but are not limited to those enzymes, described in FIGS. 1 and 2 whose action in the synthetic pathway of the host cell tends to reduce the levels of a BIA precursor of interest. In some cases, the enzyme has glucose-6-phosphate dehydrogenase activity. In certain embodiments, the enzyme that includes an inactivating mutation is ZWF1 (see e.g., FIG. 2). 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 ALD2, ALD3, ALD4, ALD5 and ALD6. 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 embodiments, the host cell includes one or more inactivating mutations to one or more genes described in Table 1.

#### Heterologous Coding Sequences

**[0069]** 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 BIA precursor(s), 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, that is not normally present in the host organism and can 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 can 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.

**[0070]** In some embodiments, the host cell includes norcoclaurine (NC) synthase activity. Any convenient NC synthase enzymes find use in the subject host cells. NC synthase enzymes of interest include, but are not limited to, enzymes such as EC 4.2.1.78, as described in Table 1. In certain embodiments, the host cell includes a heterologous coding sequence for an NC synthase 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 tyrosine to L-DOPA. In certain cases, the one or more enzymes is selected from bacterial tyrosinases, eukaryotic tyrosinases (e.g., EC 1.14.18.1) and tyrosine hydroxylases (e.g., EC 1.14.16.2.) In some instances, the host cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert L-DOPA to dopamine (e.g., EC 4.1.1.28).

**[0071]** In certain embodiments, the cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert dopamine to 3,4-DHPA. In certain cases, the one or more enzymes is a monoamine oxidase (MAO) (e.g., EC 1.4.3.4). 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.

**[0072]** In some instances, the one or more heterologous coding sequences include a MAO coding sequence integrated at a genomic locus encoding native ARO10. In certain instances, the one or more heterologous coding sequences include a MAO coding sequence operably linked to an inducible promoter. In some embodiments, the inducible promoter is part of an inducible system including a DNA binding protein targeted to a promoter regulating the ARO10 gene. In some embodiments, the host cell includes one or more heterologous coding sequences for one or more enzymes or active fragments thereof described in the genes of Table 1.

**[0073]** 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 code for its equivalent amino acid sequence, provided that the equivalent amino acid sequence produces a functional protein. Such equivalent amino acid sequences can 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 capability indicated herein. Fusions of two or more enzymes are also envisioned to facilitate the transfer of metabolites in the pathway, provided that catalytic activities are maintained.

**[0074]** Operable fragments, mutants or truncated forms may be identified by modeling and/or screening. This is made possible by 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.

**[0075]** 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%, 98%, 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.

**[0076]** 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. Thus, if the activity of the native enzyme were reduced or altogether absent, the produced intermediates would be more readily available for incorporation into the desired product.

**[0077]** In some instances, where ablation of expression of a protein may be of interest, 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 BIA molecules 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 BIA 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. 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 CPR1 to increase activity of heterologous



P450 enzymes. In addition, endogenous enzymes such as ARO8, 9, and 10, which are directly involved in the synthesis of intermediate metabolites, may also be overexpressed.

**[0078]** 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 BIAs and precursors in plants. In some cases, the enzymes for which the heterologous sequences code can be any of the enzymes in the BIA pathway, and can 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, 8 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.

**[0079]** In some cases, peptide 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, 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](http://www.yeastgenome.org).

**[0080]** In some embodiments, the host cell (e.g., a yeast strain) is engineered for selective production of a BIA 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 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 or the periplasm. 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-terminal and/or C-terminal of the enzyme.

**[0081]** 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 route to a variety of yeast organelles: the endoplasmic reticulum (ER), mitochondria (MT), plasma membrane (PM), 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 instances, the tag includes or is derived from, a transmem-

brane 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.

**[0082]** 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 BIA precursor production. Embodiments of the present invention include increased BIA precursor production in a host cell through simultaneous expression of multiple species variants of a single or multiple enzymes. In some cases, additional gene copies of a single or multiple enzymes are included in the host cell. Any convenient methods may be utilized including multiple copies of a heterologous coding sequence for an enzyme in the host cell.

**[0083]** In some embodiments, the host cell includes multiple copies of a heterologous coding sequence for an enzyme, 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 include multiple copies of heterologous coding sequences for one or more enzymes, 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.

**[0084]** In some embodiments of the host cell, the heterologous coding sequence is from a source organism selected from the group consisting of *P. somniferum*, *T. flavum* and *C. japonica*. In some instances, the source organism is *P. somniferum*, *E. caldomicum*, *C. japonica*, *T. flavum*, *Berberis stolonifer*, *T. flavum* subsp. *glaucom*, *Coptis chinensis*, *Thalictrum* spp, *Coptis* spp, *Papaver* spp, *Berberis wilsonae*, *A. mexicana*, or *Berberis* spp. In certain instances, the heterologous coding sequence is from a source organism selected from *P. somniferum*, *T. flavum* and *C. japonica*. In some embodiments, the host cell includes a heterologous coding sequence from one or more of the source organisms described in Table 1.

**[0085]** The engineered host cell medium may be sampled and monitored for the production of BIA precursors of interest. The BIA precursors 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 known 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

**[0086]** As summarized above, aspects of the invention include methods of preparing a benzyloquinoline alkaloid (BIA) 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 BIAs of interest or precursors thereof (e.g., pre-reticuline BIAs). 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 BIAs of interest. In some instances, the method is a method of preparing a benzyloquinoline alkaloid (BIA), include culturing a host cell (e.g., as described herein); adding a starting compound to the cell culture; and recovering the BIA from the cell culture. In some embodiments of the method, the starting compound, BIA product and host cell are described by one of the entries of Table 1.

**[0087]** Any convenient methods of culturing host cells may be employed for producing the BIA precursors and downstream BIAs of interest. The particular protocol that is employed may vary, e.g., depending on host cell, the heterologous coding sequences, the desired BIA precursors, 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 BIA precursors in vivo. In some embodiments, the functional enzymes are extracted from the host for production of BIAs 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.

**[0088]** Any convenient cell culture conditions for a particular cell type may be utilized. In certain embodiments, the host cells that includes one or more modifications is 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 is 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 is grown at any convenient temperature (e.g., 30° C.) 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 can also be scaled up for growth in larger fermentation vessels, for example, as part of an industrial process.

**[0089]** 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.

**[0090]** 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  $\mu$ M 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., 10 $\times$  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).

**[0091]** The subject methods may also include recovering the BIA precursor or downstream BIA 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 BIA of interest or precursor thereof 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) are used to separate the BIA or precursor from other soluble components of the cell culture. In some cases, extraction methods (e.g., liquid extraction, pH based purification, etc.) are used to separate the BIA precursor or BIA from other components of the cell culture.

**[0092]** Also included are methods of engineering host cells for the purpose of producing BIAs 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 BIAs of interest.

**[0093]** 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 biphosphate 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 pro-

motor (ADH), translation-elongation factor-1- $\alpha$  promoter (TEF), cytochrome c-oxidase promoter (CYC1), MRP7 promoter, 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 glucocorticoid 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 a 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 can also use promoter selection to optimize transcript, and hence, enzyme levels to maximize production while minimizing energy resources.

**[0094]** 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 can be broken up into 4 general categories: integrative vectors (YIp), autonomously replicating high copy-number vectors (YE<sub>p</sub> or 2 $\mu$  plasmids), autonomously replicating low copy-number vectors (YC<sub>p</sub> 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

**[0095]** 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 BIAs is of interest.

**[0096]** 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 BIAs is of interest. The host cells described herein produce benzyloquinoline alkaloid precursors (BIA precursors). Reticuline is a major branch point intermediate of interest in the synthesis of BIAs including engineering efforts to produce end products such as opioid products. The subject host cells may be utilized to produce BIA precursors from simple and inexpensive starting materials that may find use in the production of reticuline and BIA end products. As such, the subject host cells find use in the supply of therapeutically active BIAs or precursors thereof.

**[0097]** In some instances, the host cells and methods find use in the production of commercial scale amounts of BIAs 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 BIAs of interest or precursors thereof for therapeutic products. Such applica-

tions may include the industrial-scale production of BIAs of interest from fermentable carbon sources such as cellulose, starch, and free sugars.

**[0098]** 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 BIAs of interest or precursors thereof. In addition, the host cells may be engineered to produce BIAs or precursors thereof that find use in testing for bioactivity of interest in as yet unproven therapeutic functions. 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 BIAs of interest, or precursors thereof. In certain cases, research applications include the production of precursors for therapeutic molecules of interest that can 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 BIA metabolites produced in these strains.

**[0099]** 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 protopine, 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 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

**[0100]** 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.

**[0101]** 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.

**[0102]** Also provided are systems for producing a BIA 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.

**[0103]** In some cases, the system includes components for the large scale fermentation of engineered host cells, and the monitoring and purification of BIA compounds produced by the fermented host cells. In certain embodiments, one or 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 BIA products or precursors thereof. In some instances, the host cells produce a BIA precursor of interest (e.g., as described herein). In certain cases, the BIA products of interest are opioid products, such as codeine, neopine, morphine, neomorphine, hydrocodone, oxycodone, hydromorphone, dihydrocodeine, 14-hydroxycodone, or dihydromorphine.

**[0104]** In some cases, the system includes means for monitoring and or analyzing one or more BIA 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 BIA products or precursors of interest is complete, the fermentation may be halted and purification of the BIA products may be done. As such, in some cases, the subject system includes a purification component suitable for purifying the BIA 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 BIA 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 BIA fermentation products of interest following the input of one or more starting compounds to the system.

**[0105]** 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 aver-

age molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

## EXPERIMENTAL

### Example I

**[0106]** A series of specific genetic modifications provide a biosynthetic process in *Saccharomyces cerevisiae* for the production of BIAs from simple, inexpensive feedstocks or precursor molecules. Methods for constructing novel strains capable of producing the early BIA molecules norcodaurine (NC) and norlaudanosoline (NL) from non-BIA precursors or simple feedstocks are described. NC has never been reported as a product of microbial synthesis and is the natural precursor to all known BIA molecules. Methods for manipulating the regulation of yeast biosynthetic pathways and for optimizing the production of aromatic amino acids and related BIA precursors are also described.

#### *A. tyrosine and Related BIA Precursor Overproducing Yeast Strains*

**[0107]** Strains of *S. cerevisiae* are developed with improved flux through the aromatic amino acid biosynthesis pathway for the purposes of increasing intracellular concentrations of BIA precursor molecules including tyrosine, 4-hydroxyphenylacetaldehyde (4-HPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA), and dopamine. These strains combine genetic modifications to the yeast strain for the purpose of increasing carbon flux from central metabolism towards aromatic amino acid synthesis in general, towards tyrosine in particular, and include the introduction of key heterologous enzymes for the production of BIA precursor molecules not naturally produced by yeast. Genetic modifications are employed including the introduction of feedback inhibition alleviating mutations to genes encoding native biosynthetic enzymes, tuning of transcriptional regulation of native biosynthetic enzymes, deletion of genes encoding enzymes that divert precursor molecules away from the intended pathway, and introduction of heterologous enzymes for the conversion of naturally endogenous molecules into non-native BIA precursor molecules.

#### Specific Description:

**[0108]** 1.1) The biosynthetic pathway in the engineered strain incorporates feedback inhibition alleviating mutations (1.1.1) to the native yeast gene ARO4, which encodes a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase, alone or in combination. This mutation (ARO4<sup>FBR</sup>) is incorporated as a directed mutation to the native gene at the original locus, as an additional copy 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. FBR refers to feedback resistant mutants and mutations. The feedback inhibited copy of the DAHP synthase enzyme is under the native yeast transcriptional regulation or is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0109]** 1.1.1) ARO4<sup>FBR</sup> mutations may include, for example, a substitution of the lysine residue at position 229 with a leucine (see e.g., Hartmann M, et al. (2003) Evolution of feedback-inhibited beta/alpha barrel isoenzymes by gene duplication and a single mutation. *Proc Natl Acad Sci USA*

100(3):862-867), a substitution of the glutamine residue at position 166 with a lysine residue (see e.g., Fukuda K et al. (1992) Feedback-Insensitive Mutation of 3-Deoxy-D-Arabinose-7-Phosphate Synthase Caused by a Single Nucleotide Substitution of Aro4 Structural Gene in *Saccharomyces-Cerevisiae*. *J Ferment Bioeng* 74(2): 117-119), or an additional mutation conferring feedback inhibition selected from a mutagenized library of microbial DHAP synthase mutants. Examples of such selections include rescue of growth on o-fluoro-D,L-phenylalanine (see e.g., Fukuda et al. (1990) Breeding of Brewing Yeast Producing a Large Amount of Beta-Phenylethyl Alcohol and Beta-Phenylethyl Acetate. *Agr Biol Chem Tokyo* 54(1):269-271) or growth of aro3 mutant yeast strains on media with excess tyrosine.

**[0110]** 1.2) The biosynthetic pathway in the engineered strain incorporates a feedback inhibition alleviating mutation (1.2.1) to the native yeast gene ARO7, which encodes the enzyme chorismate mutase. This mutation (ARO7<sup>FBR</sup>) is incorporated as a directed mutation to the native gene at the original locus, as an additional copy 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. The feedback inhibited copy of the chorismate mutase enzyme is under the native yeast transcriptional regulation or is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0111]** 1.2.1) ARO7<sup>FBR</sup> mutant alleles may include, for example, a substitution of the threonine residue at position 226 with an isoleucine (see e.g., Schmidheini et al. (1989) A Single Point Mutation Results in a Constitutively Activated and Feedback-Resistant Chorismate Mutase of *Saccharomyces-Cerevisiae*. *J Bacteriol* 171(3):1245-1253) or an additional mutation conferring feedback inhibition selected from a mutagenized library of microbial chorismate mutase mutants. Examples of such selections include assays for 5-methyltryptophan sensitivity or increased production of melanin pigments in strains expressing heterologous tyrosinase enzymes (1.9) in the absence of externally fed tyrosine.

**[0112]** 1.3) The biosynthetic pathway in the engineered strain incorporates the introduction of a strong promoter element (such as GPD1, TEF1, etc) for the overexpression of the native yeast gene ARO10, which encodes an enzyme with hydroxyphenylpyruvate decarboxylase activity. This genetic modification is incorporated as a directed swapping of the native promoter DNA sequence at the original locus, as an additional copy of the gene under new transcriptional regulation introduced as a genetic integration at a separate locus, or as an additional copy on an episomal vector such as a 2p or centromeric plasmid.

**[0113]** 1.4) The biosynthetic pathway in the engineered strain incorporates the introduction of a strong promoter element (such as GPD1, TEF1, etc) for the overexpression of the native yeast gene ARO9, which encodes an enzyme with hydroxyphenylpyruvate/glutamic acid transaminase activity. This genetic modification is incorporated as a directed swapping of the native promoter DNA sequence at the original locus, as an additional copy of the gene under new transcriptional regulation 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.

**[0114]** 1.5) The biosynthetic pathway in the engineered strain incorporates the introduction of a strong promoter

element (such as GPD1, TEF1, etc) for the overexpression of the native yeast gene TKL, which encodes an enzyme with transketolase activity. This genetic modification is incorporated as a directed swapping of the native promoter DNA sequence at the original locus, as an additional copy of the gene under new transcriptional regulation 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.

**[0115]** 1.6) The biosynthetic pathway in the engineered strain is improved by the incorporation of a deletion or inactivating mutation of the native yeast gene ZWF1, which encodes an enzyme with glucose-6-phosphate dehydrogenase activity.

**[0116]** 1.7) The biosynthetic pathway in the engineered strain is improved by the incorporation of one or more deletion(s) or inactivating mutation(s) of known native alcohol dehydrogenase enzymes, such as the enzymes encoded by the genes ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, or SFA1.

**[0117]** 1.8) The biosynthetic pathway in the engineered strain is improved by the incorporation of one or more deletion(s) or inactivating mutation(s) of known native aldehyde oxidoreductases, such as ALD2, ALD3, ALD4, ALD5, or ALD6.

**[0118]** 1.9) The biosynthetic pathway incorporates a heterologous enzyme for the conversion of tyrosine to L-DOPA. This enzyme may be from one of several classes, including, but not limited to bacterial tyrosinases, eukaryotic tyrosinases, and tyrosine hydroxylases (Table 1). The gene for this enzyme is incorporated as a genetic integration or on an episomal vector such as a 2μ or centromeric plasmid. This L-DOPA producing enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0119]** 1.9.1) In a biosynthetic pathway using a tyrosine hydroxylase enzyme for the conversion of tyrosine to L-DOPA, additional expression of genes encoding enzymes for the synthesis and recycling of the pterin cofactor tetrahydrobiopterin (BH4) and its derivatives are incorporated into the engineered strain in support of the activity of the tyrosine hydroxylase enzyme. These enzymes include GTP cyclohydrolase, 6-pyruvoyl-tetrahydropterin synthase, sepiapterin reductase, 4a-hydroxytetrahydrobiopterin dehydratase, and quinoid dihydropteridine reductase (Table 1). The genes for these enzymes are incorporated as a genetic integration or on an episomal vector such as a 2μ or centromeric plasmid. These BH4 synthesis and recycling enzymes are introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0120]** 1.10) The biosynthetic pathway incorporates a heterologous enzyme for the decarboxylation of L-DOPA to produce dopamine. Enzymes with this activity are encoded by a genes from a variety of organisms including bacteria, plants, and mammals. Examples include *Pseudomonas putida* DOPA decarboxylase (PpDODC), *Rattus norvegicus* DOPA decarboxylase (RnDODC), and *Papaver somniferum* tyrosine/DOPA decarboxylase (PsTYDC) (Table 1). The gene for this enzyme is incorporated as a genetic integration or on an episomal vector such as a 2μ or centromeric plasmid. This dopamine producing enzyme is introduced

with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0121]** 1.11) A biosynthetic pathway for the production of 3,4-DHPA incorporates a heterologous enzyme for the oxidation of dopamine to 3,4-DHPA. Examples of genes encoding this enzyme that may be used in the strain include human monoamine oxidase A (hMAOA). *E. coli* monoamine oxidase (EcMAO), and *Micrococcus luteus* monoamine oxidase (MIMAO) (Table 1). The gene for this enzyme is incorporated as a genetic integration or on an episomal vector such as a 2 $\mu$  or centromeric plasmid. This 3,4-DHPA producing enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

**[0122]** 1.11.1) Strains for the production of NC require dopamine and 4-HPA, while strains for the production of NL require dopamine and 3,4-DHPA, but not 4-HPA. A specific modification for the conversion of an NC producing strain into an NL producing strain is the integration of a MAO gene into the yeast genome at the locus encoding the native yeast gene ARO10. This combines a deletion of the native yeast enzyme responsible for converting a tyrosine biosynthetic precursor to 4-HPA with the introduction of the enzyme capable of converting dopamine to 3,4-DHPA.

**[0123]** 1.11.2) Yeast strains are constructed with a gene that expresses a MAO enzyme (1.11) under the control of an inducible promoter. When the strain is grown in the presence of the inducer it can catalyze the conversion of dopamine to 3,4-DHPA, in the absence of inducer the strain only produces 4-HPA.

**[0124]** 1.11.2.1) Yeast strains are constructed with inducible MAO expression (1.11.2), where the inducible system also contains a DNA binding protein targeted to the promoter regulating the ARO10 gene (1.3). The synthetic promoter controlling ARO10 is therefore repressed when the promoter controlling the MAO gene is activated and ARO10 is only expressed when the MAO gene is not transcriptionally active. This system allows for the construction of a single strain that conditionally only produces the precursors for NC or NL.

## B. NC-Producing Yeast Strains

**[0125]** Methods are developed to produce the BIA molecule NC in yeast and demonstrate a first system for microbial synthesis of NC. With the engineered strains described herein, NC is produced and accumulated for its own value or combined with a biosynthetic pathway of additional heterologous enzymes for the complete synthesis of downstream BIAs.

### Specific Description:

**[0126]** 2.1) Yeast strains are grown in liquid culture to a high cell concentration before back diluting to intermediate concentrations (as measured by optical density or OD) in defined media containing high concentrations of dopamine. The media components only need to satisfy conditions for growth of the strains; various growth feedstocks are used (for example, different sugars, nitrogen sources). The NC produced by these yeast strains is excreted by the yeast cells and is measurable in the spent media. Additional NC retained by cells is recovered via cell lysis and extraction from the lysate.

**[0127]** 2.2) Yeast strains containing various combinations of the modifications as described in (1.1-1.8) substantially improve NC production from that measurable in unmodified strains in fed dopamine assays as described above (2.1). In conditions where no extracellular tyrosine is available in the yeast media, modifications described (1.1-1.8) provide for production of NC from fed dopamine; under these conditions the NC production from unmodified yeast strains is most often undetectable.

**[0128]** 2.3) Yeast strains that produce NC when containing the modification as described in (1.10) and grown as described in (2.1) when the additional BIA precursor added to media is L-DOPA instead of dopamine.

**[0129]** 2.3.1) Yeast strains as described in (2.3) containing various combinations of the modifications as described in (1.1-1.8) substantially improve production of NC.

**[0130]** 2.4) Yeast strains that produce NC when containing both the heterologous enzymes for conversion of tyrosine to dopamine (1.9, 1.10) alongside various combinations of modifications described above (1.1-1.8) and grown in media without supplementation of tyrosine, L-DOPA, or dopamine. This specific example constitutes complete synthesis of NC by the strain from simple carbon and nitrogen sources.

**[0131]** 2.5) Yeast strains are modified and cultured as described above (2.1-2.4) where the biosynthetic pathway includes the incorporation of the heterologous enzyme NCS, or truncated versions of the NCS enzyme, for the stereospecific catalysis of the reaction condensing dopamine and 4-HPA for S-NC production. This enzyme may originate from one of several plants, such as *Papaver somniferum*, *Coptis japonica*, and *Thalictrum flavum* (Table 1). The gene for this enzyme is incorporated as a genetic integration or on an episomal vector such as a 2 $\mu$  or centromeric plasmid. This S-NC producing enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter. The NC ultimately produced will be an enantiomeric mixture with bias towards the S-stereoisomer.

## C. NL-Producing Yeast Strains

**[0132]** Methods are developed to produce the BIA molecule NL from yeast. With the engineered strains described herein, NL is produced and accumulated for its own value or combined with a biosynthetic pathway of further heterologous enzymes for the complete synthesis of downstream BIAs.

### Specific Description:

**[0133]** 3.1) Yeast strains containing modifications as described in (1.11, 1.11.1-1.11.2, 1.11.2.1) are grown in liquid culture as described in (2.1) produce NL.

**[0134]** 3.2) Yeast strains containing various combinations of gene deletions as described in (1.7.1.8) improve NL production from that measurable in unmodified strains in fed dopamine assays as described above (3.1).

**[0135]** 3.3) Yeast strains that produce NL when containing the modifications as described in (1.10, 1.11, 1.11.1-1.11.2, 1.11.2.1) and grown as described in (3.1) when the additional BIA precursor is added to media is L-DOPA instead of dopamine.

**[0136]** 3.3.1) Yeast strains as described in (3.3) containing combinations of gene deletions described in (1.7,1.8) improve production of NL.

**[0137]** 3.4) Yeast strains that produce NL when containing both the heterologous enzymes for conversion of tyrosine to dopamine (1.9, 1.10) and dopamine to 3,4-HPA (1.11) alongside various combinations of modifications described above (1.1-1.8, 1.11.1) are grown in media without supplementation of tyrosine, L-DOPA, or dopamine. This specific example constitutes complete synthesis of NL by the strain from simple carbon and nitrogen sources.

**[0138]** 3.5) Yeast strains modified and cultured as described above (3.1-3.4) where the biosynthetic pathway includes the incorporation of the heterologous enzyme NCS, or truncated versions of the NCS enzyme (Table 1), for the stereospecific catalysis of the reaction condensing dopamine and 3,4-HPA for S-NL production. This enzyme may originate from one of several plants, such as *Papaver somniferum*, *Coptis japonica*, and *Thalictrum flavum* (Table 1). The gene for this enzyme is incorporated as a genetic integration or on an episomal vector such as a 2 $\mu$  or centromeric plasmid. This S-NL producing enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter. The NC ultimately produced is an enantiomeric mixture with bias towards the S-stereoisomer.

FIG. 1: Biosynthesis of Tyrosine and BIA Precursor Molecules

**[0139]** Schematic showing the biosynthetic pathway from glucose to tyrosine and other BIA precursors. Aromatic amino acid intermediates present in native yeast metabolism are written in black. Endogenous yeast enzymes are written in grey (apart from TYR, TyrH, DODC and MAO). Heterologous enzymes and non-natural BIA precursor molecules include TYR, TyrH, DODC and MAO. As described in (1.1, 1.2) wild-type yeast enzymes encoded by ARO4 and ARO7 are allosterically inhibited by tyrosine, indicated here by the dotted grey line. Individual steps in the pentose phosphate pathway and glycolysis are not explicitly detailed in this figure, although the genes TKL1 and ZWF1 (targeted in 1.5, 1.6) are involved in the pentose phosphate pathway, as indicated.

FIG. 2: Effect of ZWF1 Knockout and TKL1 Over-Expression on Pentose Phosphate Pathway (PPP)

**[0140]** Schematic detailing how modifications to TKL1 (1.5) and ZWF1 (1.6) affect the overall carbon flow through the pentose phosphate pathway in yeast when glucose is the primary carbon source. Panel A represents wild-type carbon flow; Panel B represents the relative change in carbon flow in a modified strain.

FIG. 3A: Synthesis of NC from Precursor Molecules

**[0141]** NC is synthesized from one molecule of dopamine and one molecule of 4-HPA via a Pictet-Spengler condensation reaction. This reaction can occur spontaneously to produce a racemic mixture of R- and S-NC. This reaction can alternatively be catalyzed by the plant enzyme NCS, which produces S-NC.

FIG. 3B: Synthesis of NL from Precursor Molecules

**[0142]** NL is synthesized from one molecule of dopamine and one molecule of 3,4-DHPA via a Pictet-Spengler condensation reaction. This reaction can occur spontaneously to

produce a racemic mixture of R- and S-NL. While the natural product of NCS is NC, the enzyme has been shown to catalyze the stereospecific production of S-NL (see e.g., Rueffer et al. (1981) (S)-Norlaudanoline Synthase—the 1st Enzyme in the Benzyloquinoline Biosynthetic-Pathway. *Febs Lett* 129(1):5-9).

**[0143]** Measurement of the BIA molecules is performed by LC-MS analysis, where NC production ( $m/z=+272$ , 19.2 min retention time) and NL production ( $m/z=+288$ , 18.9 min retention time) were observed, with ion MS2 fragmentation agreeing with both standards and published detection methods (see e.g., Schmidt et al. (2007) Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring-13C6]-tyramine feeding. *Phytochemistry* 68(2):189-202).

FIG. 4: Effect of Four Genetic Modifications on NC Production with Varying Fed Tyrosine

**[0144]** NC production was demonstrated at several concentrations of fed tyrosine, including no fed tyrosine, in strains with targeted genetic modifications. Wild-type strain, CEN.PK2, was integrated with constructs conferring one of four genetic changes (as described in 1.1-1.4): overexpression of ARO10 by promoter replacement with  $P_{TEF1}$ , overexpression of ARO9 by promoter replacement with  $P_{TEF1}$ , chromosomal integration of an ARO4<sup>FBR</sup> allele, and chromosomal integration of an ARO7<sup>FBR</sup> allele. When incorporated alone, only the  $P_{TEF1}$ -ARO10 and ARO4<sup>FBR</sup> increase production of NC. While both these modifications increased NC production at all tyrosine concentrations, the ARO4<sup>FBR</sup> integrated strain improved most drastically at zero fed tyrosine.

FIG. 5: NC Production with Combinations of Genetic Modifications

**[0145]** Some genetic modifications as described above (1.5, 1.6) improve NC production only in combination with the integration of the ARO4<sup>FBR</sup> mutant (1.1). This figure shows four strains engineered with single genetic modifications,  $P_{TEF1}$ -ARO10 (1.3),  $P_{GPD}$ -TKL1 (1.5), ZWF1 knockout (1.6), and ARO4<sup>FBR</sup> (1.1), alongside three strains constructed with combinations of genetic modifications; Strain A ( $P_{GPD}$ -TKL1, ARO4<sup>FBR</sup>), Strain B (ZWF1 knockout, ARO4<sup>FBR</sup>), and Strain C ( $P_{GPD}$ -TKL1, ZWF1 knockout, ARO4<sup>FBR</sup>). NC production is shown normalized to the WT strain, with Strain C exhibiting a five-fold increase in NC production.

FIG. 6: NL Production in ALD/ADH Knockout Strains

**[0146]** NL production is improved by the deletion of competing yeast enzymes (1.7,1.8) in a strain expressing human MAOA on a 2 $\mu$  plasmid (1.11) and grown in media containing dopamine. NL production is shown as titer measured in spent media normalized to the WT (with hMAO, but with no deletions). Improvements in production is as much as ten times WT NL production.

FIG. 7: Activity of a DODC Enzyme In Vivo

**[0147]** Yeast strains transformed with DNA to express *Papaver somniferum* tyrosine/DOPA decarboxylase can convert L-DOPA to dopamine in vivo. Strains harboring a 2 $\mu$  plasmid were grown in selective media and then back-diluted into media containing L-DOPA. Spent media was

then measured for concentrations of L-DOPA (retention time 4.8 min,  $m/z=+198$ ) and dopamine (retention time 4.2 min,  $m/z=+154$ ).

FIG. 8

**[0148]** NC is produced as described in (2.1) in multiple wild-type yeast lab strains at varying tyrosine concentrations. Specifically, each yeast strain is inoculated into separate liquid cultures and grown overnight to  $OD_{600}\sim 10$ , then back-diluted in YNB minimal media without tyrosine to an  $OD_{600}\sim 1$  and grown for 3 hours, 100  $\mu$ l of each culture was mixed into 400  $\mu$ l YNB media containing 100 mM dopamine and varying concentrations of tyrosine; each strain was grown in each media condition in triplicate samples. NC titer was measured from culture supernatant on an LC-MS instrument detecting  $m/z+272$  ion count in peaks as described by (see e.g., Schmidt et al. (2007) Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring- $^{13}C_6$ ]-tyramine feeding. *Phytochemistry* 68(2):189-202). The area of each peak was integrated to calculate a relative quantity of the NC in each sample and the results were normalized to the ion count area in CEN.PK2 yeast culture with 0 mg/L tyrosine.

FIG. 9

**[0149]** NC is produced as described in (2.1) in multiple engineered yeast strains at fed 100 mM dopamine and no tyrosine. These data were generated in a separate experiment from those in FIG. 5. The strains CSY977-981 were engineered to contain combinations of the genetic modifications described in (1.1-1.6); the labels underneath each strain name indicate which modifications were incorporated into each strain. Strain CSY981 contains five genetic modifications to yeast native metabolism and exhibits a twelve-fold increase in NC titer above the wild-type yeast strain CEN.PK2.

FIG. 10

**[0150]** NC production as described in (2.1) for black diamonds and as described in (2.3) for gray circles. Here NC was produced in an engineered yeast strain (CSY980) with the additional integration of the L-DOPA decarboxylase PpDODC (1.10). In separate liquid cultures this yeast strain was grown in YNB minimal media containing varying concentrations of dopamine (black diamonds) and YNB minimal media containing L-DOPA (gray circles; cultures not fed dopamine). The solid black line represents a linear regression of the relationship between the measured NC and fed dopamine. The peak area measurements for the L-DOPA fed samples were plotted along the regression line for dopamine fed samples to show an “equivalent fed dopamine” quantity for the cultures fed L-DOPA. L-DOPA media was mixed to achieve a target concentration of 10 mM L-DOPA, however L-DOPA was not fully soluble at that concentration, and the effective concentration of dissolved L-DOPA is estimated to be approximately 6 mM. Based on the average NC titers of the two L-DOPA fed yeast cultures the “equivalent fed dopamine” concentration is approximately 50 mM or  $8\times$  the fed L-DOPA concentration (indicated by the gray dotted lines).

FIG. 11

**[0151]** Mammalian tyrosine hydroxylases (TyrHs) are capable of hydroxylating tyrosine, but are dependent on the

co-substrate tetrahydrobiopterin (BH4) for activity, as described in (1.9.1). During the catalysis of tyrosine to L-DOPA by TyrH, molecular oxygen is split and transferred to tyrosine and BH4, as shown by reaction 1. BH4 is oxidized to BH4-4 $\alpha$ -carbinolamine (4 $\alpha$ OH-8H4). Two heterologous enzymes are expressed in yeast to synthesize BH4 from the folate synthesis pathway intermediate, dihydroneopterin triphosphate. First, 6-pyruvoyltetrahydropterin synthase (PTPS) converts dihydroneopterin to PTP (reaction 2), which is then reduced to 8H4 by sepiapterin reductase (SepR, reaction 3). Two enzymes are responsible for the regeneration of BH4 from its 4 $\alpha$ -carbinolamine form. First, pterin-4 $\alpha$ -carbinolamine dehydratase (PCD) catalyzes a loss of water reaction to form dihydrobiopterin (reaction 4). Dihydrobiopterin is then reduced to tetrahydrobiopterin by quinoid dihydropteridine reductase (QDHPR, reaction 5).

FIG. 12

**[0152]** Tyrosine hydroxylases expressed from yeast cells convert tyrosine to L-DOPA. Yeast strains transformed with plasmids carrying tyrosine hydroxylases from human (hTH2) and rat (RnTyrH) were grown in liquid media and then lysed in buffer containing tyrosine and the co-substrate BH4. After 6-hour incubations at 30° C., L-DOPA was measured in the lysate mixture by LC-MS. (A) LC-MS chromatogram confirms conversion of tyrosine to L-DOPA dependent on the presence of the co-substrate, BH4. (B) Fragmentation of the +198  $m/z$  ion peak further confirms the presence of L-DOPA in lysate samples. (See e.g., Lv et al. (2010) LC-MS-MS Simultaneous Determination of L-Dopa and its prodrug n-Pentyl Hydrochloride in Rat Plasma. *Chromatographia*, 72(3/4), 239-243).

FIG. 13

**[0153]** Co-expression of tyrosine hydroxylase with a 8H4 biosynthetic enzyme enables conversion of tyrosine to L-DOPA in yeast cell lysates. Engineered yeast strains integrated with constructs expressing rat tyrosine hydroxylase (RnTyrH) and rat sepiapterin reductase (RnSepR) were grown in liquid media and then lysed in buffer containing tyrosine, NADPH, and the BH4 biosynthetic precursor, sepiapterin. Co-expression of a TyrH with the BH4 biosynthesis gene provides for activity of the tyrosine hydroxylase in the absence of BH4, but in the presence of the BH4 precursor, sepiapterin.

FIG. 14

**[0154]** Synthesis of the BIA molecules coclaurine and N-methylcoclaurine from NC. Conversion of NC to downstream BIA molecules by plant methyltransferase enzymes extends microbial BIA synthesis. These particular BIA molecules can only be synthesized via a NC-dependent biosynthesis scheme.

FIG. 15

**[0155]** Engineered yeast strains produce NC-derived BIA molecules from L-DOPA in liquid culture. A copy of PpDODC was integrated into the engineered yeast strain, CSY979 (as described in FIG. 9), providing for the production of NC from L-DOPA (panel A, bottom chromatogram). Next a copy of the opium poppy 6-O-methyltransferase (Ps6OMT) gene was integrated into this yeast strain to enable the production of coclaurine from L-DOPA (panel A,



middle chromatogram). Finally, a copy of both Ps60OMT and the opium poppy coclaurine-N-methyltransferase (PsC-NMT) genes were integrated into the CSY979 yeast strain carrying the PpDODC gene to enable the production of N-methylcoclaurine from L-DOPA (panel A, top chromatogram). Both the NC and coclaurine measurements matched chromatograms from chemical standards. The production of

N-methylcoclaurine was further confirmed by matching the fragmentation pattern of the +300 m/z ion peak to patterns in published literature (Panel B). (see e.g., Schmidt et al. (2007) Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring-<sup>13</sup>C<sub>6</sub>]-tyramine feeding. *Phytochemistry* 68(2):189-202).

TABLE 1

Genes of interest as components of the engineered metabolic pathways							
Enzyme	Abbrev.	Catalyzed reactions	Source organisms	Engineered regulaton	Coding sequence changes	Genbank #	Specific description ref.
3-deoxy-d-arabinose-heptulosonate-7-phosphate synthase	ARO4, DHAP synthase	erythrose-4-phosphate + PEP → DHAP (EC 2.5.1.54)	<i>Saccharomyces cerevisiae</i>	native, constitutive, synthetic regulation	Feedback inhibition resistant mutation, K229L, Q166K	CAA85212.1	1.1
Chorismate mutase	ARO7	chorismate → prephenate (EC 5.4.9.5)	<i>Saccharomyces cerevisiae</i>	native, constitutive, synthetic regulation	Feedback inhibition resistant mutation, T226I	NP_015385.1	1.2
Phenylpyruvate decarboxylase	ARO10	hydroxyphenylpyruvate → 4HPA (EC 4.1.1.80)	<i>Saccharomyces cerevisiae</i>	constitutive overexpression, synthetic regulation		NP_010668.3	1.3
Aromatic aminotransferase	ARO9	hydroxyphenylpyruvate + glutamate → tyrosine + alpha-ketoglutarate (EC 2.6.1.57)	<i>Saccharomyces cerevisiae</i>	constitutive overexpression, synthetic regulation		AEC14313.1	1.4
Transketolase	TKL1	fructose-6-phosphate + glyceraldehyde-3-phosphate ↔ xylulose-5-phosphate + erythrose-4-phosphate (EC 2.2.1.1)	<i>Saccharomyces cerevisiae</i>	constitutive overexpression, synthetic regulation		NP_015399.1	1.5
Glucose-6-phosphate dehydrogenase	ZWF1	glucose-6-phosphate → 6-phosphogluconolactone (EC 1.1.1.49)	<i>Saccharomyces cerevisiae</i>		full deletion of coding region	CAA96146.1	1.6
Alcohol dehydrogenase	ADH2-7 SFA1	3HPA → tyrosol (EC 1.1.1.90)	<i>Saccharomyces cerevisiae</i>		full deletion of coding region	NP_014032.1, AAT93007.1, NP_011258.2, NP_009703.3, NP_014051.3, NP_010030.1, NP_01013.1	1.7
Aldehyde oxidase	ALD2-6	4HPA → hydroxyphenylacetic acid (EC 1.2.1.39)	<i>Saccharomyces cerevisiae</i>		full deletion of coding region	NP_013893.1, NP_013892.1, NP_015019.1, NP_010996.2, NP_015264.1	1.8
Tyrosinase	TYR	tyrosine → L-DOPA, L-DOPA → dopaquinone (EC 1.14.18.1)	<i>Ralstonia solanacearum</i> , <i>Agaricus bisporus</i>	constitutive overexpression, synthetic regulation		NP_518458.1, AJ223816,	1.9
Tyrosinase hydroxylase	TyrH	tyrosine → L-DOPA (EC 1.14.16.2)	<i>Homo sapiens</i> , <i>norvegicus</i> , <i>Mus musculus</i>	constitutive overexpression, synthetic regulation		NM_012740, NM_000240,	1.9
GTP cyclohydrolase	FOL2	GTP → dihydroneopterin triphosphate (EC 3.5.4.16)	<i>Saccharomyces cerevisiae</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	native regulation, constitutive overexpression, synthetic regulation		CAA97297.1, NP_001019195.1, NP_032128.1	1.9.1

TABLE 1-continued

Genes of interest as components of the engineered metabolic pathways							
Enzyme	Abbrev.	Catalyzed reactions	Source organisms	Engineered regulation	Coding sequence changes	Genbank #	Specific description ref.
6-pyruvoyl tetrahydrobiopterin (PTP) synthase	PTPS	dihydroneopterin triphosphate → PTP (EC 4.2.3.12)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	constitutive overexpression, synthetic regulation		AAH59140.1, BAA04224.1, AAH29013.1	1.9.1
Sepiapterin reductase	SepR	PTP → BH4 (EC 1.1.1.153)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	constitutive overexpression, synthetic regulation		NP_062054.1, NP_003115.1, NP_035597.2	1.9.1
4a-hydroxytetrahydrobiopterin (pterin-4 $\alpha$ -carbinolamine) dehydratase	PCD	4a-hydroxytetrahydrobiopterin → H <sub>2</sub> O + quinoid dihydropteridine (EC 4.2.1.96)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	constitutive overexpression, synthetic regulation		NP_001007602.1, AAB25581.1, NP_079549.1	1.9.1
Quinoid dihydropteridine reductase	QDHP	quinoid dihydropteridine → BH4 (EC 1.5.1.34)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	constitutive overexpression, synthetic regulation		AAH72536.1, NP_000311.2, AAH02107.1	1.9.1
L-DOPA decarboxylase	DODC	L-DOPA → dopamine (EC 4.1.1.28)	<i>Pseudomonas putida</i> , <i>Rattus norvegicus</i>	constitutive overexpression, synthetic regulation		AE015451.1, NP_001257782.1	1.10
Tyrosine/DOPA decarboxylase	TYDC	L-DOPA → dopamine (EC 4.1.1.28)	<i>Papaver somniferum</i>	constitutive overexpression, synthetic regulation			1.10
Monoamine oxidase	MAO	dopamine → 3,4-DHPA (EC 1.4.3.4)	<i>E. coli</i> , <i>Homo sapiens</i> , <i>Micrococcus luteus</i>	constitutive overexpression, synthetic regulation		J03792, D2367, AB010716.1	1.11, 1.11.1
Norococlaurine synthase	NCS	4HPA + dopamine → S-norococlaurine (EC 4.2.1.78)	<i>Coptis japonica</i> , <i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Thalictrum flavum</i> , <i>Corydalis saxicola</i>	constitutive	N-terminal truncation	BAF45337.1, ACI45396.1, ACO90258.1, ACO90247.1, AEB71889.1	2.5, 3.5
Norococlaurine 6-O-methyltransferase	6OMT	Norococlaurine → coclaurine Norlaudanosoline → 3'-hydroxycoclaurine EC 2.1.1.128	<i>Papaver somniferum</i> , <i>Thalictrum flavum</i> , <i>Coptis japonica</i> , <i>Papaver bracteatum</i> , <i>Eschscholzia californica</i>	constitutive overexpression, synthetic regulation		AY268894 AY610507 D29811 ACO90225.1 BAM37634.1	
Coclaurine-Nmethyltransferase	CNMT	Coclaurine → N-methylcoclaurine 3'-hydroxycoclaurine → 3'-hydroxy-Nmethylcoclaurine EC 2.1.1.140	<i>P. somniferum</i> , <i>T. flavum</i> , <i>C. japonica</i>	constitutive overexpression, synthetic regulation		AY217336 AY610508 AB061863	
4'-O-methyltransferase	4'OMT	3'-hydroxy-N-methylcoclaurine → Reticuline EC 2.1.1.116	<i>P. somniferum</i> , <i>T. flavum</i> , <i>Coptis japonica</i> , <i>Eschscholzia californica</i>	constitutive overexpression, synthetic regulation		AY217333, AY217334 AY610510 D29812 BAM37633.1	

[0156] Notwithstanding the appended clauses, the disclosure set forth herein is also defined by the following clauses:

1. A host cell that produces a benzyloquinoline alkaloid (BIA) precursor, wherein the host cell comprises one or more modifications selected from the group consisting of:

[0157] one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell;

[0158] one or more transcriptional modulation modifications of one or more biosynthetic enzyme gene native to the cell;

[0159] one or more inactivating mutations in one or more enzymes native to the cell; and

[0160] one or more heterologous coding sequences that encode one or more enzymes;

[0161] wherein when the cell comprises one or more heterologous coding sequences that encode one or more enzymes, it comprises 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 gene native to the cell; and an inactivating mutation in an enzyme native to the cell.

2. The host cell according to Clause 1, wherein the BIA precursor is selected from the group consisting of norcoclaurine (NC) and norlaudanosoline (NL).

3. The host cell according to Clause 2, wherein the BIA precursor is norcoclaurine (NC).

4. The host cell according to Clause 2, wherein the BIA precursor is norlaudanosoline (NL).

5. The host cell according to Clause 1, wherein when the cell comprises one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell, it comprises a least one additional modification selected from the group consisting of: a transcriptional modulation modification of a biosynthetic enzyme gene native to the cell; an inactivating mutation in an enzyme native to the cell; and a heterologous coding sequence that encode an enzyme.

6. The host cell according to Clause 1, wherein when the cell comprises one or more transcriptional modulation modifications of one or more biosynthetic enzyme genes native to the cell, it comprises 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 native to the cell; and a heterologous coding sequence that encodes an enzyme.

7. The host cell according to Clause 1, wherein when the cell comprises one or more inactivating mutations in one or more enzymes native to the cell, it comprises 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 gene native to the cell; and a heterologous coding sequence that encodes an enzyme.

8. The host cell according to Clause 1, wherein the cell comprises one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell.

9. The host cell according to Clause 8, wherein the one or more biosynthetic enzyme genes encode one or more

enzymes selected from a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase and a chorismate mutase.

10. The host cell according to Clause 9, wherein the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene selected from ARO4 and ARO7.

11. The host cell according to Clause 10, wherein the one or more feedback inhibition alleviating mutations are present in the ARO4 gene.

12. The host cell according to Clause 9, wherein the mutations are selected from mutation of position 229 and mutation of position 166 of the DAHP.

13. The host cell according to any one of the preceding clauses, wherein the cell overproduces one or more BIA precursor molecules.

14. The host cell according to Clause 13, wherein the one or more BIA precursor molecules are selected from the group consisting of tyrosine, 4-hydroxyphenylacetaldehyde (4-HPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) and dopamine.

15. The host cell according to Clause 14, wherein the one or more BIA precursor molecules are 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) and dopamine.

16. The host cell according to Clause 14, wherein the one or more BIA precursor molecules are 4-hydroxyphenylacetaldehyde (4-HPA) and dopamine.

17. The host cell according to Clause 1, wherein the cell comprises one or more transcriptional modulation modifications of one or more biosynthetic enzyme genes native to the cell.

18. The host cell according to Clause 17, wherein the transcriptional modulation modification is substitution of a strong promoter for a native promoter of the one or more biosynthetic enzyme genes.

19. The host cell according to Clause 18, wherein the one or more biosynthetic enzyme genes is selected from the group consisting of ARO10, ARO9 and TKL.

20. The host cell according to Clause 1, wherein the cell comprises one or more inactivating mutations in one or more enzymes native to the cell.

21. The host cell according to Clause 20, wherein the enzyme is a glucose-6-phosphate dehydrogenase.

22. The host cell according to Clause 21, wherein the enzyme is ZWF1.

23. The host cell according to Clause 20, wherein the cell comprises one or more inactivating mutations in one or more enzymes native to the cell comprising alcohol dehydrogenase activity.

24. The host cell according to Clause 23, wherein the one or more enzymes is selected from the group consisting of ADH2, ADH3, ADH4, ADH5, ADH6, ADH7 and SFA1.

25. The host cell according to Clause 20, wherein the cell comprises one or more inactivating mutations in one or more enzymes native to the cell comprising aldehyde oxidoreductase activity.

26. The host cell according to Clause 25, wherein the one or more enzymes is selected from the group consisting of ALD2, ALD3, ALD4, ALD5, and ALD6.

27. The host cell according to Clause 1, wherein the cell comprises one or more heterologous coding sequences that encode one or more enzymes or active fragments thereof.

28. The host cell according to Clause 27, wherein the host cell comprises an NC synthase activity.

29. The host cell according to Clause 27, wherein the host cell comprises a heterologous coding sequence for an NC synthase or active fragment thereof.

30. The host cell according to Clause 29, wherein the one or more enzymes or active fragments thereof convert tyrosine to L-DOPA.

31. The host cell according to Clause 27, wherein the one or more enzymes or active fragments thereof is selected from the group consisting of bacterial tyrosinases, eukaryotic tyrosinases and tyrosine hydroxylases.

32. The host cell according to Clause 27, wherein the cell comprises one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert L-DOPA to dopamine.

33. The host cell according to Clause 27, wherein the cell comprises one or more heterologous coding sequences for one or more enzymes or active fragments thereof that convert dopamine to 3,4-DHPA.

34. The host cell according to Clause 33, wherein the one or more enzymes is a monoamine oxidase (MAO).

35. The host cell according to Clause 34, wherein the one or more heterologous coding sequences comprises a MAO coding sequence integrated at a genomic locus encoding native ARO10.

36. The host cell according to Clause 34, wherein the one or more heterologous coding sequences comprises a MAO coding sequence operably linked to an inducible promoter.

37. The host cell according to Clause 36, wherein the inducible promoter is part of an inducible system comprising a DNA binding protein targeted to a promoter regulating the ARO10 gene.

38. The host cell according to Clause 27, wherein the heterologous coding sequence is from a source organism selected from the group consisting of *P. somniferum*, *T. flavum* and *C. japonica*.

39. The host cell according to Clause 1, wherein the host cell is a eukaryotic cell.

40. The host cell according to Clause 39, wherein the eukaryotic cell is a yeast cell.

41. The host cell according to Clause 40, wherein the yeast cell is a *S. cerevisiae* cell, a *Schizosaccharomyces pombe* or a *Pichia pastoris* cell.

42. The host cell according to Clause 41, wherein the yeast cell is a *S. cerevisiae* cell.

43. A method of preparing a benzyloquinoline alkaloid (BIA), comprising:

[0162] culturing a host cell according to any one of Clauses 1 to 42;

[0163] adding a growth feedstock to the cell culture; and

[0164] recovering the BIA from the cell culture.

44. The method according to Clause 43, wherein the BIA is a BIA precursor selected from norlaudanoline and norcoclaurine.

45. The method according to Clause 43, further comprising adding a starting compound to the cell culture.

46. The method according to Clause 43, further comprising producing a BIA precursor selected from reticuline, coclaurine, N-methylcoclaurine and norreticuline.

47. The method according to Clause 43, further comprising producing one or more opiate compounds.

48. The method according to Clause 45, wherein the one or more opiate compounds is selected from oripavine, morphine, codeine, hydromorphone, hydrocodone, oxycodone and oxymorphone from thebaine.

49. A kit comprising:

[0165] a host cell according to any one of Clauses 1 to 42; and

[0166] one or more components selected from a starting compound, a growth feedstock, a heterologous coding sequence, cloning vectors, multiple cloning sites (MCS), bi-directional promoters, an internal ribosome entry site (IRES) and a cell culture media.

[0167] 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.

[0168] 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.

What is claimed is:

1. A host cell that produces a benzyloquinoline alkaloid (BIA) precursor, wherein the host cell comprises one or more modifications selected from the group consisting of:

one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell;

one or more transcriptional modulation modifications of one or more biosynthetic enzyme gene native to the cell;

one or more inactivating mutations in one or more enzymes native to the cell; and

one or more heterologous coding sequences that encode one or more enzymes;

wherein when the cell comprises one or more heterologous coding sequences that encode one or more enzymes, it comprises 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

lation modification of a biosynthetic enzyme gene native to the cell; and an inactivating mutation in an enzyme native to the cell.

2. The host cell according to claim 1, wherein the BIA precursor is selected from the group consisting of norcoclaurine (NC) and norlaudanosoline (NL).

3. The host cell according to claim 1, wherein the cell comprises one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell.

4. The host cell according to any one of the preceding claims, wherein the cell overproduces one or more BIA precursor molecules.

5. The host cell according to claim 4, wherein the one or more BIA precursor molecules are selected from the group consisting of tyrosine, 4-hydroxyphenylacetaldehyde (4-HPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) and dopamine.

6. The host cell according to claim 1, wherein the cell comprises one or more transcriptional modulation modifications of one or more biosynthetic enzyme genes native to the cell.

7. The host cell according to claim 6, wherein the transcriptional modulation modification is substitution of a strong promoter for a native promoter of the one or more biosynthetic enzyme genes.

8. The host cell according to claim 1, wherein the cell comprises one or more inactivating mutations in one or more enzymes native to the cell.

9. The host cell according to claim 8, wherein the enzyme is a glucose-6-phosphate dehydrogenase, an alcohol dehydrogenase or an aldehyde oxidoreductase.

10. The host cell according to claim 1, wherein the cell comprises one or more heterologous coding sequences that encode one or more enzymes or active fragments thereof.

11. The host cell according to claim 1, wherein the host cell is a eukaryotic cell.

12. The host cell according to claim 11, wherein the eukaryotic cell is a yeast cell.

13. The host cell according to claim 12, wherein the yeast cell is a *S. cerevisiae* cell, a *Schizosaccharomyces pombe* or a *Pichia pastoris* cell.

14. A method of preparing a benzyloquinoline alkaloid (BIA), comprising:

culturing a host cell according to any one of claims 1 to 13;

adding a growth feedstock to the cell culture; and  
recovering the BIA from the cell culture.

\* \* \* \* \*