(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
2 February 2012 (02.02.2012)

PCT

(10) **International Publication Number**
WO 2012/016172 A2

(51) **International Patent Classification**:
C12P 7/64 (2006.01)

(21) **International Application Number**:
PCT/US2011/045947

(22) **International Filing Date**:
29 July 2011 (29.07.2011)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/369,547 30 July 2010 (30.07.2010) US

(71) **Applicant** *(for all designated States except US)*: **AMYRIS, INC.** [US/US]; 5885 Hollis Street, Suite 100, Emeryville, CA 94608 (US).

(72) **Inventor; and**
(75) **Inventor/Applicant** *(for US only)*: **PLATT, Darren** [AU/US]; 5885 Hollis Street, Suite 100, Emeryville, CA 94608 (US).

(74) **Agent**: **TRAVER, Robert D.**; Sheridan Ross P.C., 1560 Broadway, Suite 1200, Denver, CO 80202 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

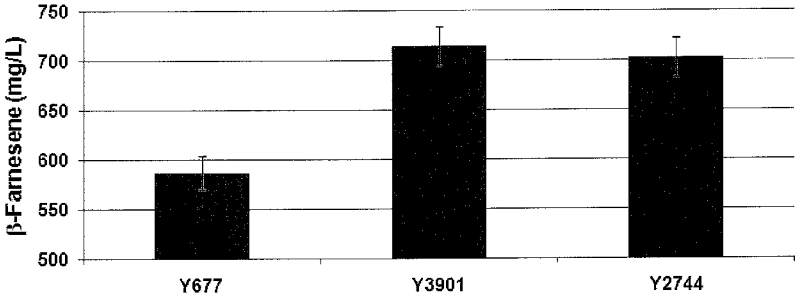
(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: GENETICALLY MODIFIED MICROBES PRODUCING INCREASED LEVELS OF ACETYL-COA DERIVED COMPOUNDS

Figure 6



(57) **Abstract**: Provided herein are genetically modified microorganisms that produce increased amounts of acetyl-CoA derived compounds in industrial fermentation processes, and methods for making and using the same.

WO 2012/016172 A2

GENETICALLY MODIFIED MICROBES PRODUCING INCREASED LEVELS OF ACETYL-COA DERIVED COMPOUNDS

[0001] This application claims benefit of priority of U.S. Provisional Application No. 61/369,547, filed on July 30, 2010, which is hereby incorporated by reference in its entirety.

1. FIELD OF THE INVENTION

[0002] The compositions and methods provided herein generally relate to the industrial use of microorganisms. In particular, provided herein are genetically modified microorganisms that have increased availability of acetyl-CoA for the production of acetyl-CoA derived compounds, and methods for making and using such genetically modified microorganisms.

2. BACKGROUND

[0003] The use of genetically modified microorganisms to produce useful chemicals in more efficient and cost-effective quantities has become a goal of the pharmaceutical, supplement, food and chemical industries. Among such useful chemicals are compounds derived from acetyl-CoA, such as, for example, isoprenoids (used, for example, in pharmaceutical products and as biofuels, food additives, and other specialty chemicals), amino acids (used, for example, in feed additives and pharmaceutical products), fatty acids (used, for example, in solvents, emulsifiers, cleaning compounds, and lubricants), and vitamins (used, for example, as dietary supplements). Given the low yields achieved by extracting many acetyl-CoA derived compounds from natural sources, genetically modified microorganisms present a promising vehicle for their fermentative production. However, there exists a continuing need for microorganisms that are suitable for industrial applications in general, and for the industrial production of acetyl-CoA derived compounds in particular.

3. SUMMARY OF THE INVENTION

[0004] Provided herein are genetically-modified microorganisms (*e.g.,* a genetically modified *Saccharomyces cerevisiae* cell) that produce acetyl-CoA derived compound(s), and methods and materials for generating and using such microorganisms.

[0005] In one aspect, provided herein is a genetically modified microorganism capable of making an acetyl-CoA derived compound, wherein the microorganism comprises heterologous nucleic acid encoding a modified YNG2.

[0006] In another aspect, provided herein is a genetically modified microorganism capable of making an acetyl-CoA derived compound, wherein the microorganism comprises heterologous nucleic acid encoding a modified YNG2, wherein the modified YNG2 is functionally equivalent to a YNG2 that is lacking a PHD-finger domain.

[0007] In another aspect, provided herein is a genetically modified microorganism capable of making an acetyl-CoA derived compound, wherein the microorganism comprises heterologous nucleic acid encoding a modified YNG2 that has at least 15% less nucleosome HAT activity in either the nuA4 or piccolo NuA4 complex.

[0008] In another aspect, provided herein is a genetically modified microorganism capable of making an acetyl-CoA derived compound, wherein the microorganism comprises heterologous nucleic acid encoding a modified YNG2, wherein the modified YNG2 results in at least 15% increased translation of a galactose regulated gene product.

[0009] In some embodiments, the modified YNG2 gene produces a modified YNG2 protein that comprises a modified PHD-finger domain. In some such embodiments, the modified PHD-finger domain comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 227, 235, 243, 252, and 264 of the wild-type YNG2 amino acid sequence.

[0010] In some embodiments, the modified YNG2 lacks a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the omitted segment comprises all or a segment of the PHD-finger domain of YNG2. In some such embodiments, the omitted segment comprises the amino acids at positions selected from the group consisting of positions corresponding to positions 221 to 282, 254 to 282, and 264 to 282 of the wild-type YNG2p amino acid sequence.

[0011] In some embodiments, the modified YNG2 comprises a modification outside of the PHD-finger domain. In some such embodiments, the modified YNG2 comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 40, 54, 60, 61, 91, 103, 129, 133, 134, 147, and 182 of the wild-type YNG2p amino acid sequence.

[0012] In some embodiments, the modified YNG2 has an amino acid sequence selected from the group consisting of SEQ ID NOs: 90, 92, 106, 107, 108, 109, and 110.

[0013] In another aspect, provided herein is a method for enhancing production of an acetyl-CoA derived compound in a microorganism, the method comprising culturing a genetically modified microorganism comprising heterologous nucleic acid encoding a modified YNG2 as described herein in a suitable medium and under conditions that promote production of the acetyl-CoA derived compound, wherein the acetyl-CoA derived compound is produced at a level that is higher than the level of the acetyl-CoA derived compound in a microorganism not comprising the heterologous nucleic acid encoding a modified YNG2.

4. BRIEF DESCRIPTION OF THE FIGURES

[0014] FIG. 1 provides a schematic representation of the mevalonate ("MEV") pathway for the production of isopentenyl diphosphate ("IPP").

[0015] FIG. 2 provides a schematic representation of the conversion of IPP and dimethylallyl pyrophosphate ("DMAPP") to geranyl pyrophosphate ("GPP"), farnesyl pyrophosphate ("FPP"), and geranylgeranyl pyrophosphate ("GGPP").

[0016] FIGs. 3A-E provide maps of the inserts of vectors pAM489, pAM491, pAM493, pAM495, and pAM497.

[0017] FIG. 4 provides structures of chromosomal integration constructs YNG2-C227Y-A and YNG2-C227Y-B, and the YNG2 target locus after integration of the two chromosomal integration constructs by homologous recombination.

[0018] FIG. 5 provides the structure of chromosomal integration construct YNG2-delPHD and the YNG2 target locus after integration of the chromosomal integration construct by homologous recombination.

[0019] FIG. 6 provides β-farnesene production data for strains Y677, Y3901, and Y2744.

[0020] FIG. 7 provides galatose regulated gene expression data for strains Y3900 and Y3901 for endogenous (B, D) and heterologous (A, C) genes under sugar phase sampling conditions (A, B) or ethanol phase sampling conditions (C, D).

[0021] FIG. 8 provides galatose-regulated protein expression data for strains Y3900 and Y3901 for heterologous genes under sugar phase sampling conditions (A) or ethanol phase sampling conditions (B).

5. DETAILED DESCRIPTION OF THE EMBODIMENTS

5.1 Definitions

[0022] As used herein, the term "heterologous" refers to what is not normally found in nature. The term "heterologous nucleotide sequence" refers to a nucleotide sequence not normally found in a given cell in nature. As such, a heterologous nucleotide sequence may be: (a) foreign to its host cell (*i.e.,* is "exogenous" to the cell); (b) naturally found in the host cell (*i.e.,* "endogenous") but present at an unnatural quantity in the cell (*i.e.,* greater or lesser quantity than naturally found in the host cell); or (c) be naturally found in the host cell but positioned outside of its natural locus.

[0023] As used herein, the term "parent" refers to a microorganism that does not comprise all of the genetic modifications of a genetically modified microorganism as described herein, but that serves as the starting point for introducing said genetic modifications leading to the generation of such a genetically modified microorganism.

[0024] As used herein, the term "YNG2p" refers to a polypeptide encoded by a YNG2 gene of *Saccharomyces cerevisiae.* The sequence of the wild-type YNG2 gene of *Saccharomyces cerevisiae* has been previously described. Loewith *et al., Mol. Cell. Biol.* 20 (11):3807 (2000). Representative YNG2 nucleotide sequences of *Saccharomyces cerevisiae* include Genbank accession numbers FN393071 and NM_001179220, and SEQ ID NO: 104 as provided herein. Representative YNG2p sequences of *Saccharomyces cerevisiae* include Genbank accession number CAY80101 and NP_011958, and SEQ ID NO: 105 as provided herein.

[0025] As used herein, the phrase "modified YNG2 gene" refers to a nucleotide sequence of the YNG2 gene of *Saccharomyces cerevisiae* that comprises nucleotide substitutions, deletions, and/or additions as compared to the nucleotide sequence of the wild-type YNG2 gene.

[0026] As used herein, the phrase "modified YNG2 protein" or "modified YNG2p" refers to an amino acid sequence of the YNG2p of *Saccharomyces cerevisiae* that comprises amino acid substitutions, deletions, and/or additions as compared to the amino acid sequence of the wild-type YNG2p.

[0027] In some embodiments, a modified YNGp as described herein is a YNG2 protein whose sequence differs from that of the wild-type protein and that has a decreased activity relative to the wild-type protein (unmodified YNG2p) under the same or equivalent conditions. In some embodiments, the modified YNG2p is encoded by a nucleic acid sequence that is at least 95% identical to the wild-type nucleic acid sequence. In other embodiments, the modified YNG2p is encoded by a nucleic acid sequence that is at least 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, and 35% identical to the nucleic acid sequence encoding the wild-type protein. In still other embodiments, the modified YNG2p has less than 95% identity to SEQ ID NO: 105. In still other embodiments, the modified YNG2p has less than 90%, 85%, 80%, 75% and 70% than

SEQ ID NO: 105. In still further embodiments, the modified YNG2p includes a partial or complete deletion or inactivation of the PHD domain of the YNG2 protein.

[0028] In some embodiments, a modified YNG2p as described herein has at least a 15% decreased activity relative to the wild-type protein. In other embodiments, the modified YNG2p has at least 20%, 25%, 30%, 35%, 40%, 50%, 55%, 60%, 65%, and 70% decreased activity relative to the wild-type protein. In other embodiments, the modified YNG2p has more than 50% decreased activity relative to the wild-type protein. In still other embodiments, the decreased activity is nucleosome HAT activity of the nuA4 or piccolo NuA4 complex.

[0029] In some embodiments, a modified YNG2p as described herein is functionally equivalent to a YNG2p that is lacking a PHD-finger domain. As used herein, the phrase “PHD-finger domain” refers to the amino acids located at positions 220 to 282 of wild-type YNG2p. Thus, a modified YNG2p that is functionally equivalent to a YNG2p that is lacking a PHD-finger domain includes but is not limited to a YNG2p wherein one or more of amino acids 220 to 282 are deleted, substituted or modified so as to render the PHD-finger domain non-functional.

[0030] As used herein, the phrase “isoprenoids” refers to a diverse group of natural compounds that are derived from a single biosynthetic precursor, the five-carbon molecule isopentenyl pyrophosphate (“IPP”). Isoprenoids find commercial application as pharmaceuticals, nutriceuticals, fragrances, flavoring compounds, agricultural pest control agents, and biofuels.

[0031] As used herein, the phrase “galactose regulated gene expression” refers to transcription of coding sequences that are operably linked to a promoter that is induced by galactose. As used herein, the term “galactose regulated protein expression” refers to production of proteins from coding sequences that are operably linked to a promoter that is induced by galactose. Illustrative examples of promoters that are induced by galactose include but are not limited to the promoters of the GAL1, GAL10, GAL7, GAL2, and GCY1 genes of *Saccharomyces cerevisiae*.

[0032] As used herein, the phrase “acetyl-CoA derived compound” refers to a carbon-containing compound that is biosynthesized by a microorganism where one or more of the carbons is derived from an acetyl-CoA molecule. Illustrative examples of acetyl-CoA derived compounds include but are not limited to isoprenoids, amino acids, fatty acids, and vitamins. A “heterologous acetyl-CoA derived compound” refers to either

(i) a compound not normally synthesized by the microorganism or (ii) a compound that is normally synthesized by the microorganism but whose biosynthesis includes at least one heterologous nucleic acid sequence as part of the compound's biosynthetic pathway.

[0033] As used herein, the phrase "biosynthetic enzyme" refers to an enzyme that functions in a biosynthetic pathway leading to the production of an acetyl-CoA derived compound.

5.2 Genetically Modified Microorganisms and Methods for Making the Same

[0034] Provided herein are genetically modified microorganisms (*e.g.,* a genetically modified *Saccharomyces cerevisiae* cell) that produce an acetyl-CoA derived compound, and methods and materials for generating and using such compositions.

[0035] Acetyl-CoA is an intermediate of both primary and secondary metabolism. Because acetyl-CoA does not readily cross membranes, biosynthesis of acetyl-CoA is required in each cellular compartment, where it functions as a precursor for many biomolecules. In the nucleus, acetyl-CoA is the substrate for the acetylation of histones and transcription factors, regulating their function in maintaining or altering chromosome structure and gene transcription. In mitochondria, acetyl-CoA is incorporated into the TCA cycle for use in generating ATP. In the cytosol, acetyl-CoA is metabolized via one of three mechanisms: (1) carboxylation, leading to polyketides and fatty acids, (2) condensation, leading to the biosynthesis of mevalonate-derived isoprenoids, and (3) acetylation, leading to acetylated phenolics, alkaloids, isoprenoids, and sugars. Thus, increased production and availability of acetyl-CoA in the cytosol can increase carbon flux directly into the production of isoprenoid, polyketide and fatty acid compounds, with very limited amounts of the available cytosolic acetyl-CoA diverted into nuclear or mitochondrial metabolism.

[0036] The genetically modified microorganisms provided herein comprise genetic modifications resulting in greater yields of one or more compounds biosynthesized from acetyl-CoA compared to a parent microorganism lacking the genetic modifications described herein.

[0037] In one aspect, provided herein is a genetically modified microorganism comprising a heterologous nucleic acid molecule encoding a modified YNG2p, wherein the microorganism is capable of making a heterologous acetyl-CoA derived compound.

[0038] Without being bound by theory, it is believed that expression of the modified YNG2p as described herein results in an increased production of a heterologous acetyl-CoA derived compound by either increasing acetyl-CoA pools or by increasing the levels of a galactose-regulated gene product. YNG2 is involved in a complex that regulates the acetylation and deacetylation of lysine residues in histones.

[0039] Histone acetylation and deacetylation is mediated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). One such HAT is the Esa1p of *Saccharomyces cerevisiae*, which is the catalytic enzyme of the NuA4 multi-protein complex and the piccolo NuA4 tri-protein complex, the former being implicated in nontargeted, global histone acetylation, and the latter in specific, local histone acetylation. Another component of both complexes is YNG2p, a homeodomain (PHD)-finger protein that has been postulated to play a regulatory role in histone acetylation and transcriptional activation.

[0040] While YNG2p cannot be deleted from a cell without making the cell unsuitable for heterologous compound production, changes to the YNG2 gene or YNG2p, including insertions, deletions and substitutions can be made, which result in changes in histone acetylation and/or gene transcription, leading to an increase in the biosynthesis of acetyl-CoA derived compounds.

[0041] Thus, in one aspect, provided herein is a genetically modified microbial cell capable of making an acetyl-CoA derived compound, wherein the cell comprises a heterologous nucleic acid molecule encoding a modified YNG2p, wherein the cell makes at least 15% more of the acetyl-CoA derived compound relative to a reference cell that does not contain the heterologous nucleic acid molecule encoding the modified YNG2p, but is otherwise genetically identical.

[0042] In another aspect, provided herein is a genetically modified microbial cell capable of making an acetyl-CoA derived compound, wherein the cell comprises a heterologous nucleic acid molecule encoding a modified YNG2p, wherein the modified YNG2p is functionally equivalent to a YNG2p that is lacking a PHD-finger domain.

[0043] In another aspect, provided herein is a genetically modified microbial cell capable of making an acetyl-CoA derived compound, wherein the cell comprises a heterologous nucleotide molecule encoding a modified YNG2p that has at least 15% less nucleosome HAT activity in either the nuA4 or piccolo NuA4 complex relative to an unmodified YNG2p.

[0044] In yet another aspect, provided herein is a genetically modified microbial cell capable of making a heterologous acetyl-CoA derived compound, wherein the cell comprises a heterologous nucleotide molecule encoding a modified YNG2p, wherein expression of the modified YNG2p results in at least 15% increased translation of a galactose regulated gene product, relative to a reference cell that does not contain the heterologous nucleic acid molecule encoding the modified YNG2p, but is otherwise genetically identical.

[0045] In some embodiments, the modified YNG2p comprises a modified PHD-finger domain. In some such embodiments, the modified PHD-finger domain comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 227, 235, 243, 252, and 264 of the wild-type YNG2p amino acid sequence.

[0046] In some embodiments, the modified YNG2p lacks a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the omitted segment comprises all or part of the PHD-finger domain of YNG2p. In some such embodiments, the omitted segment comprises the amino acids at the positions selected from the group consisting of positions corresponding to positions 221 to 282, 254 to 282, and 264 to 282 of the wild-type YNG2p amino acid sequence.

[0047] In some embodiments, the modified YNG2p comprises a modification outside of the PHD-finger domain. In some such embodiments, the modified YNG2p comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 40, 54, 60, 61, 91, 103, 129, 133, 134, 147, and 182 of the wild-type YNG2p amino acid sequence. In some embodiments, the modified YNG2p has an amino acid sequence selected from the group consisting of SEQ ID NOs: 90, 92, 106, 107, 108, 109, and 110.

[0048] In one aspect, provided herein is a genetically modified microorganism comprising (a) a first heterologous nucleic acid molecule encoding a modified YNG2p, and (b) a second heterologous nucleic acid molecule encoding a biosynthetic enzyme, wherein the microorganism produces at least 15% more of an acetyl-CoA derived compound than a parent microorganism not comprising the first heterologous nucleic acid molecule.

[0049] In some embodiments, the second heterologous nucleic acid molecule encodes a biosynthetic enzyme including, but not limited to, mevalonate (“MEV”) pathway enzymes, IPP isomerases, polyprenyl synthases, and enzymes that can modify a polyprenyl to form a hemiterpene, a monoterpene, a sesquiterpene, a diterpene, a triterpene, a tetraterpene, a polyterpene, a steroid compound, a carotenoid, or a modified acetyl-CoA derived compound. In some embodiments, the acetyl-CoA derived compound is an isoprenoid compound.

[0050] In some embodiments, the second heterologous nucleic acid molecule encodes a biosynthetic enzyme including, but not limited to, at least one polyketide synthesis pathway enzyme, and enzymes that can modify acetyl-CoA compound to form a polyketide product such as a macrolide, an antibiotic, an antifungal, a cytostatic compound, an anticholesterolemic compound, an antiparasitic compound, a coccidiostatic compound, an animal growth promoter or an insecticide. In some embodiments, the acetyl-CoA derived compound is a polyene. In some embodiments, the acetyl-CoA derived compound is a cyclic lactone. In some embodiments, the acetyl-CoA derived compound comprises a 14, 15, or 16-membered lactone ring.

[0051] In some embodiments, the second heterologous nucleic acid molecule encodes a biosynthetic enzyme including, but not limited to, at least one fatty acid synthesis pathway enzyme, and enzymes that can modify an acetyl-CoA compound to form a fatty acid product such as a palmitate, palmitoyl CoA, palmitoleic acid, sapienic acid, oleic acid, linoleic acid, α-linolenic acid, arachidonic acid, eicosapentaenoic acid, erucic acid, and docosahexaenoic acid.

[0052] In another aspect, provided herein is a method for producing an acetyl-CoA derived compound comprising: (a) obtaining a plurality of genetically modified microorganisms comprising (i) a first heterologous nucleic acid molecule encoding a modified YNG2p, and (ii) a second heterologous nucleic acid molecule encoding a biosynthetic enzyme; (b) culturing said genetically modified microorganisms in a medium comprising a carbon source under conditions suitable for making the acetyl-CoA derived compound; and (c) recovering the acetyl-CoA derived compound from the medium.

[0053] In some embodiments, the acetyl-CoA derived compound is produced in an amount greater than about 10 grams per liter of fermentation medium. In some such embodiments, the acetyl-CoA derived compound is produced in an amount from about 10 to about 50 grams, more than about 15 grams, more than about 20 grams, more than about 25 grams, or more than about 30 grams per liter of cell culture.

[0054] In some embodiments, the acetyl-CoA derived compound is produced in an amount greater than about 50 milligrams per gram of dry cell weight. In some such embodiments, the acetyl-CoA derived compound is produced in an amount from about 50 to about 1500 milligrams, more than about 100 milligrams, more than about 150 milligrams, more than about 200 milligrams, more than about 250 milligrams, more than about 500 milligrams, more than about 750 milligrams, or more than about 1000 milligrams per gram of dry cell weight.

[0055] In some embodiments, the acetyl-CoA derived compound is produced in an amount that is at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 2.5-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, at least about 30-fold, at least about 40-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 200-fold, at least about 300-fold, at least about 400-fold, at least about 500-fold, or at least about 1,000-fold, or more, higher than the amount of the acetyl-CoA derived compound produced by a microorganism that is not genetically modified as disclosed herein, on a per unit volume of cell culture basis.

[0056] In some embodiments, the acetyl-CoA derived compound is produced in an amount that is at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 2.5-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, at least about 30-fold, at least about 40-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 200-fold, at least about 300-fold, at least about 400-fold, at least about 500-fold, or at least about 1,000-fold, or more, higher than the amount of the acetyl-CoA derived compound produced by a microorganism that is not genetically modified as disclosed herein, on a per unit dry cell weight basis.

[0057] In some embodiments, the acetyl-CoA derived compound is produced in an amount that is at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 2.5-fold, at least about 5-fold, at least about 10-

fold, at least about 20-fold, at least about 30-fold, at least about 40-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 200-fold, at least about 300-fold, at least about 400-fold, at least about 500-fold, or at least about 1,000-fold, or more, higher than the amount of the acetyl-CoA derived compound produced by a microorganism that is not genetically modified as disclosed herein, on a per unit volume of cell culture per unit time basis.

[0058] In some embodiments, the acetyl-CoA derived compound is produced in an amount that is at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 2.5-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, at least about 30-fold, at least about 40-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 200-fold, at least about 300-fold, at least about 400-fold, at least about 500-fold, or at least about 1,000-fold, or more, higher than the amount of the acetyl-CoA derived compound produced by a microorganism that is not genetically modified as disclosed herein, on a per unit dry cell weight per unit time basis.

[0059] In some embodiments, the acetyl-CoA derived compound is an isoprenoid compound. In some such embodiments, the isoprenoid compound is selected from the group consisting of abietadiene, amorphadiene, carene, α-farnesene, β-farnesene, farnesol, geraniol, geranylgeraniol, isoprene, linalool, limonene, myrcene, nerolidol, ocimene, patchoulol, β-pinene, sabinene, γ-terpinene, terpinolene and valencene. Isoprenoid compounds also include, but are not limited to, carotenoids (such as lycopene, α- and β-carotene, α- and β-cryptoxanthin, bixin, zeaxanthin, astaxanthin, and lutein), steroid compounds, and compounds that are composed of isoprenoids modified by other chemical groups, such as mixed terpene-alkaloids, and coenzyme Q-10. In some embodiments, the acetyl-CoA derived compound is a polyketide selected from the group consisting of a polyketide macrolide, antibiotic, antifungal, cytostatic, anticholesterolemic, antiparasitic, a coccidiostatic, animal growth promoter and insecticide. Polyketide compounds also include, but are not limited to, polyenes and cyclic lactones. In some embodiments, the acetyl-CoA derived compound is a fatty acid selected from the group consisting of palmitate, palmitoyl CoA, palmitoleic acid, sapienic acid, oleic acid, linoleic acid, α-linolenic acid, arachidonic acid, eicosapentaenoic acid, erucic acid, and docosahexaenoic acid.

[0060] In another aspect, provided herein is an isolated nucleic acid comprising a modified YNG2 coding sequence encoding a modified YNG2p. In some embodiments, the modified YNG2 coding sequence, when introduced into a parent microorganism provides a genetically modified microorganism that has reduced levels of NuA4 complex and/or piccolo NuA4 complex mediated acetylation of its nucleosomal histones compared to the parent microorganism. In some embodiments, the modified YNG2 coding sequence, when introduced into a parent microorganism, provides a genetically modified microorganism that has similar levels of NuA4 complex and/or piccolo NuA4 complex mediated acetylation of its nucleosomal histones as a genetically modified microorganism comprising a YNG2 coding sequence that encodes a YNG2p lacking a PHD-finger domain.

[0061] In some embodiments, the modified YNG2 coding sequence, when introduced into a parent microorganism, provides a genetically modified microorganism that has increased galactose regulated gene expression compared to the parent microorganism. In some embodiments, the modified YNG2 coding sequence when introduced into a parent microorganism provides a genetically modified microorganism that has similar levels of galactose regulated gene expression as a genetically modified microorganism comprising a YNG2 coding sequence encoding a YNG2p lacking a PHD-finger domain. In some embodiments, the modified YNG2 coding sequence, when introduced into a parent microorganism comprising a heterelogous nucleotide sequence encoding a biosynthetic enzyme, results in a genetically modified microorganism that produces increased amounts of an acetyl-CoA derived compound compared to the parent microorganism. In some embodiments, the modified YNG2 coding sequence, when introduced into a parent microorganism comprising a heterelogous nucleotide sequence encoding a biosynthetic enzyme, results in a genetically modified microorganism that produces levels of an acetyl-CoA derived compound that are similar to the levels of the acetyl-CoA derived compound produced by a genetically modified microorganism comprising a YNG2 coding sequence encoding a YNG2p lacking a PHD-finger domain. In some embodiments, the modified YNG2 coding sequence encodes a YNG2p comprising a modified PHD-finger domain. In some such embodiments, the modified PHD-finger domain comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 227, 235, 243, 252, and

264 of the wild-type YNG2p amino acid sequence. In some embodiments, the modified YNG2 coding sequence encodes a YNG2p lacking a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the omitted segment comprises all or a segment of the PHD-finger domain of YNG2p. In some embodiments, the modified YNG2 coding sequence encodes a modified YNGp having an amino acid sequence selected from the group consisting of SEQ ID NOs: 90, 92, 106, 107, 108, 109, and 110.

5.2.1 Microbe Selection

[0062] Microbes useful in the compositions and methods provided herein include eukaryotic unicellular organisms, in particular, fungi; and more particularly, yeasts.

[0063] Yeast useful in the compositions and methods provided herein include yeast that have been deposited with microorganism depositories (*e.g.* IFO, ATCC, *etc.*) and belong to the genera *Aciculoconidium, Ambrosiozyma, Arthroascus, Arxiozyma, Ashbya, Babjevia, Bensingtonia, Botryoascus, Botryozyma, Brettanomyces, Bullera, Bulleromyces, Candida, Citeromyces, Clavispora, Cryptococcus, Cystofilobasidium, Debaryomyces, Dekkara, Dipodascopsis, Dipodascus, Eeniella, Endomycopsella, Eremascus, Eremothecium, Erythrobasidium, Fellomyces, Filobasidium, Galactomyces, Geotrichum, Guilliermondella, Hanseniaspora, Hansenula, Hasegawaea, Holtermannia, Hormoascus, Hyphopichia, Issatchenkia, Kloeckera, Kloeckeraspora, Kluyveromyces, Kondoa, Kuraishia, Kurtzmanomyces, Leucosporidium, Lipomyces, Lodderomyces, Malassezia, Metschnikowia, Mrakia, Myxozyma, Nadsonia, Nakazawaea, Nematospora, Ogataea, Oosporidium, Pachysolen, Phachytichospora, Phaffia, Pichia, Rhodosporidium, Rhodotorula, Saccharomyces, Saccharomycodes, Saccharomycopsis, Saitoella, Sakaguchia, Saturnospora, Schizoblastosporion, Schizosaccharomyces, Schwanniomyces, Sporidiobolus, Sporobolomyces, Sporopachydermia, Stephanoascus, Sterigmatomyces, Sterigmatosporidium, Symbiotaphrina, Sympodiomyces, Sympodiomycopsis, Torulaspora, Trichosporiella, Trichosporon, Trigonopsis, Tsuchiyaea, Udeniomyces, Waltomyces, Wickerhamia, Wickerhamiella, Williopsis, Yamadazyma, Yarrowia, Zygoascus, Zygosaccharomyces, Zygowilliopsis*, and *Zygozyma*, among others.

[0064] In some embodiments, the microbe is *Saccharomyces cerevisiae, Pichia pastoris, Schizosaccharomyces pombe, Dekkera bruxellensis, Kluyveromyces lactis* (previously called *Saccharomyces lactis*), *Kluveromyces marxianus, Arxula adeninivorans*, or *Hansenula polymorpha (*now known as *Pichia angusta).* In some embodiments, the microbe is a strain of the genus *Candida*, such as *Candida lipolytica, Candida guilliermondii*, *Candida krusei, Candida pseudotropicalis*, or *Candida utilis.*

[0065] In a particular embodiment, the microbe is *Saccharomyces cerevisiae*. In some embodiments, the microbe is a strain of *Saccharomyces cerevisiae* selected from the group consisting of Baker's yeast, CBS 7959, CBS 7960, CBS 7961, CBS 7962, CBS 7963, CBS 7964, IZ-1904, TA, BG-1, CR-1, SA-1, M-26, Y-904, PE-2, PE-5, VR-1, BR-1, BR-2, ME-2, VR-2, MA-3, MA-4, CAT-1, CB-1, NR-1, BT-1, and AL-1. In some embodiments, the microbe is a strain of *Saccharomyces cerevisiae* selected from the group consisting of PE-2, CAT-1, VR-1, BG-1, CR-1, and SA-1. In a particular embodiment, the strain of *Saccharomyces cerevisiae* is PE-2. In another particular embodiment, the strain of *Saccharomyces cerevisiae* is CAT-1. In another particular embodiment, the strain of *Saccharomyces cerevisiae* is BG-1.

[0066] In some embodiments, the microbe is suitable for industrial fermentation, *e.g.,* bioethanol fermentation. In particular embodiments, the microbe is conditioned to subsist under high solvent concentration, high temperature, expanded substrate utilization, nutrient limitation, osmotic stress due, acidity, sulfite and bacterial contamination, or combinations thereof, which are recognized stress conditions of the industrial fermentation environment. The genetically modified microorganism disclosed herein comprises a defective YNG2 gene.

5.2.2 Modified YNG2

[0067] In some embodiments, the microorganism disclosed herein comprises a first heterologous nucleic acid molecule comprising a modified YNG2 gene. The sequence of the wild-type YNG2 gene of *Saccharomyces cerevisiae* has been previously described. Loewith *et al., Mol. Cell. Biol.* 20 (11):3807 (2000). Representative YNG2 nucleotide sequences of *Saccharomyces cerevisiae* include Genbank accession numbers FN393071 and NM_001179220, and SEQ ID NO: 104 as provided herein. Representative YNG2p sequences of *Saccharomyces cerevisiae* include Genbank accession number CAY80101 and NP_011958, and SEQ ID NO: 105 as provided herein. The PHD-finger domain of YNG2p is defined as the segment of YNG2p that consists of the amino acids at positions 220 to 282 of the wild-type YNG2p amino acid sequence.

[0068] In some embodiments, the modified YNG2 gene produces a modified YNG2p that comprises a modified PHD-finger domain. In some such embodiments, the modified PHD-finger domain comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 227, 235, 243, 252, and 264 of the wild-type YNG2p amino acid sequence. In some such embodiments, the modified PHD-finger domain comprises an amino acid substitution from cysteine to tyrosine at a position corresponding to position 227 of the wild-type YNG2p amino acid sequence (C227Y mutation). In other such embodiments, the modified PHD-finger domain comprises an amino acid substitution from methionine to isoleucine at a position corresponding to position 235 of the wild-type YNG2p amino acid sequence (M235I mutation). In other such embodiments, the modified PHD-finger domain comprises an amino acid substitution from cysteine to arginine at a position corresponding to position 243 of the wild-type YNG2p amino acid sequence (C243R mutation). In other such embodiments, the modified PHD-finger domain comprises an amino acid substitution from cysteine to tyrosine at a position corresponding to position 252 of the wild-type YNG2p amino acid sequence (C252Y mutation). In other such embodiments, the modified PHD-finger domain comprises an amino acid substitution from tyrosine to histidine at a position corresponding to position 264 of the wild-type YNG2p amino acid sequence (Y264H mutation). In some embodiments, the modified YNG2p has an amino acid sequence selected from the group consisting of SEQ ID NOs: 90, 107, 108, 109, and 110.

[0069] In some embodiments, the modified YNG2 gene produces a modified YNG2p that lacks a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises all or a segment of the PHD-finger domain of YNG2p. In some such embodiments, the segment comprises the amino acids corresponding to positions 264 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 254 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 221 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 181 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 167 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 111 to 282 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 106 to 282 of the wild-type YNG2p amino acid sequence. In other embodiments, the segment comprises at least the amino acids corresponding to positions 220 to 230, 230 to 240, 240 to 250, 250 to 260, 260 to 270, 270 to 280, 225 to 235, 235 to

245, 245 to 255, 255 to 265, 265 to 275, or 275 to 282 of the wild-type YNG2 amino acid sequence. In some embodiments, the segment comprises at least the N-terminal third, the middle third, or the C-terminal third of the PHD-finger domain of wild-type YNG2p. In some embodiments, the segment comprises at least a fourth, a third, or a half of the PHD-finger domain of wild-type YNG2p. In some embodiments, the modified YNG2p has the amino acid sequence disclosed herein as SEQ ID NO: 92. In some embodiments, the segment comprises amino acids located outside of the PHD-finger domain of YNG2p. In some such embodiments, the segment comprises the amino acids corresponding to positions 181 to 219 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 167 to 219 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 111 to 219 of the wild-type YNG2p amino acid sequence. In some such embodiments, the segment comprises the amino acids corresponding to positions 106 to 219 of the wild-type YNG2p amino acid sequence.

[0070] In some embodiments, the modified YNG2 gene produces a modified YNG2p that comprises a wild-type PHD-finger domain but one or more modifications outside of the PHD-finger domain. In some such embodiments, the modified YNG2p comprises one or more amino acid substitutions at positions selected from the group consisting of positions corresponding to positions 40, 54, 60, 61, 91, 103, 129, 133, 134, 147, and 182 of the wild-type YNG2p amino acid sequence. In some such embodiments, the modified YNG2p comprises an amino acid substitution from lysine to arginine at a position corresponding to position 40 of the wild-type YNG2p amino acid sequence (K40R mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from arginine to glycine at a position corresponding to position 54 of the wild-type YNG2p amino acid sequence (R54G mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from proline to serine at a position corresponding to position 60 of the wild-type YNG2p amino acid sequence (P60S mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from lysine to arginine at a position corresponding to position 61 of the wild-type YNG2p amino acid sequence (K61R mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from asparagine to serine at a position corresponding to position 91 of the wild-type YNG2p amino acid sequence (N91S mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from lysine to arginine at a position corresponding to position 103 of the wild-type YNG2p amino acid sequence (K103R mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from alanine to serine at a position corresponding to position 129 of the wild-type YNG2p amino acid sequence (A129S mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from serine to tyrosine at a position corresponding to position 133 of the wild-type YNG2p amino acid sequence (S133Y mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from arginine to serine at a position corresponding to position 134 of the wild-type YNG2p amino acid sequence (R134S mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from arginine to lysine at a position corresponding to position 147 of the wild-type YNG2p amino acid sequence (R147K mutation). In other such embodiments, the modified YNG2p comprises an amino acid substitution from lysine to arginine at a position corresponding to position 182 of the wild-type YNG2p amino acid sequence (K182R mutation). In some embodiments, the modified YNG2p has the amino acid sequence of SEQ ID NO: 106.

5.2.3 Targeted Modification of YNG2

[0071] The YNG2 gene can be modified in a microorganism by introducing into the microorganism a chromosomal integration construct that can replace an endogenous gene sequence with an altered gene sequence by homologous recombination. In such embodiments, the chromosomal integration construct comprises the altered YNG2 gene sequence flanked by a pair of nucleotide sequences that are homologous to a pair of nucleotide sequences flanking the endogenous gene sequence to be replaced (YNG2 homologous sequences). Upon replacement of the endogenous gene sequence of the YNG2 gene with the altered YNG2 gene sequence, a modified YNG2p is produced from the modified YNG2 gene.

[0072] In some embodiments, the chromosomal integration construct is a linear DNA molecule. In other embodiments, the chromosomal integration construct is a circular DNA molecule. In some such embodiments, the circular chromosomal integration construct comprises a single YNG2 homologous sequence. Upon integration at the target YNG2 locus, such circular chromosomal integration construct comprising a single YNG2 homologous sequence would become linearized, with a portion of the YNG2 homologous sequence positioned at each end and the remaining segments of the chromosomal integration construct inserting into the target YNG2 locus without replacing any of the target locus nucleotide sequence. In particular embodiments, the single YNG2 homologous sequence of a circular chromosomal integration construct is homologous to a sequence located within the coding sequence of the YNG2 gene. A circular integration construct comprising a single YNG2 homologous sequence that is homologus to a sequence located within the coding sequence of the YNG2 gene can be suitable, for example, to introduce a stop codon into the YGN2 gene to result in a modified YNG2 gene that produces a modified YNG2p in which the PHD-finger domain is deleted.

[0073] Parameters of chromosomal integration constructs that may be varied include, but are not limited to, the lengths of the YNG2 homologous sequences; the nucleotide sequence of the YNG2 homologous sequences; the length of the altered YNG2 gene sequence; and the nucleotide sequence of the altered YNG2 gene sequence.

[0074] In some embodiments, an effective range for the length of each YNG2 homologous sequence is 50 to 5,000 base pairs. In particular embodiments, the length of each YNG2 homologous sequence is about 500 base pairs. For a discussion of the length of homology required for gene targeting, *see* Hasty *et al., Mol Cell Biol* 11:5586-91 (1991).

[0075] In some embodiments, the YNG2 homologous sequences comprise coding sequences of the YNG2 gene. In other embodiments, the YNG2 homologous sequences comprise upstream and/or downstream sequences of the YNG2 gene. In some embodiments, one YNG2 homologous sequence comprises a nucleotide sequence that is homologous to a nucleotide sequence located within or 5' of the coding sequence of the YNG2 gene, and the other YNG2 homologous sequence comprises a nucleotide sequence that is homologous to a nucleotide sequence located 3' of the coding sequence of the YNG2 gene, respectively. In some embodiments, one YNG2 homologous sequence comprises a nucleotide sequence that is homologous to a nucleotide sequence located 5' of the coding sequence of the YNG2 gene, and the other YNG2 homologous sequence comprises a nucleotide sequence that is homologous to a nucleotide sequence located within or 3' of the coding sequence of the YNG2 gene, respectively. In some embodiments, both YNG2 homologous sequences comprise nucleotide sequences that are homologous to nucleotide sequences located within the coding sequence of the YNG2 gene.

[0076] In some embodiments, the altered gene sequence is generated via site-directed mutagenesis (*see* Carter, *Bio*C*hem. J.* 237:1-7 (1986); Zoller and Smith, *Methods Enzymol.* 154:329-50 (1987)), cassette mutagenesis, restriction selection mutagenesis

(Wells *et al*., *Gene* 34:315-323 (1985)), oligonucleotide-mediated (site-directed) mutagenesis, PCR mutagenesis, or other techniques known in the art for modifying DNA sequences performed either before or after introduction of the altered gene sequence into the chromosomal integration vector.

[0077] In some embodiments, the length for the altered gene sequence is from 1 to 10,000 base pairs. In some embodiments, the length for the altered gene sequence is from 1 to 8,000 base pairs. In some embodiments, the length for the altered gene sequence is from 1 to 6,000 base pairs. In some embodiments, the length for the altered gene sequence is from 1 to 4,000 base pairs. In some embodiments, the length for the altered gene sequence is from 1 to 2,000 base pairs. In some embodiments, the length for the altered gene sequence is a length approximately equivalent to the distance between the regions of the YNG2 gene that match the YNG2 homologous sequences in the chromosomal integration construct.

[0078] In some embodiments, the altered gene sequence comprises a nucleotide sequence encoding a selectable marker that enables selection of microorganisms comprising the altered gene sequence. In some embodiments, a termination codon is positioned in-frame with, and downstream of, the nucleotide sequence encoding the selectable marker to prevent translational read-through that might yield a fusion protein.

[0079] In some embodiments, the YNG2 gene is modified in a microorganism by introducing into the microorganism more than one chromosomal integration construct wherein only integration of all of the introduced chromosomal integration constructs by homologus recombination results in the desired modification of the YNG2 gene. In some such embodiments, each chromosomal integration construct comprises (a) one or more YNG2 homologous sequences, and (b) a region of homology with one other chromosomal integration construct (construct homolgous sequence), and at least one chromosomal integration construct comprises an altered YNG2 gene sequence flanked by a pair of YNG2 homologous sequences. In some such embodiments, each chromosomal integration construct comprises a nucleotide sequence encoding only a segment of a selectable marker, wherein such segment comprises the construct homologous sequence, such that only integration of all the chromosomal integration constructs by homologous recombination at the YNG2 homologous sequences and at the construct homologous sequence creates functional selectable markers that can be used to identify microorganisms comprising the desired modification of the YNG2 gene.

5.2.4 Random Modification of YNG2

[0080] The YNG2 gene can be modified in a microorganism by random mutagenesis using any of a variety of well-established methods followed by identification and isolation of mutations introduced into the YNG2 coding sequence. Suitable mutagenesis methods include, but are not limited to, chemical mutation methods, radiation-induced mutagenesis, and methods of mutating a nucleic acid during synthesis. In some embodiments, a nucleic acid comprising a nucleotide sequence encoding wild-type YNG2 is exposed to a chemical mutagen, subjected to radiation mutation, or subjected to an error-prone PCR, and the mutagenized nucleic acid is introduced into a genetically modified microorganism. Mutations in the YNG2 coding sequence that increase production of an acetyl-CoA derived compound can be subsequently identified by introducing the mutagenized DNA into a parent microorganism comprising a heterologous nucleotide sequence encoding a biosynthetic enzyme, comparing the amount of acetyl-CoA derived compound produced by the microorganism to the amount produced by the parent cell, and sequencing the introduced YNG2 coding sequence.

5.2.5 Nucleosome HAT Activity

[0081] In some embodiments, the modified YNG2 gene produces a modified YNG2p that when present in a NuA4 complex or a piccolo NuA4 complex causes reduced nucleosome HAT activity by the complex. In some such embodiments, the modified YNG2p comprises a modified PHD-finger domain. In other such embodiments, the modified YNG2p comprises a non-modified PHD-finger domain but one or more modifications outside of the PHD-finger domain. In yet other such embodiments, the modified YNG2p lacks a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the omitted segment comprises all or a part of the PHD-finger domain. In other such embodiments, the omitted segment comprises all or a part of YNG2p outside of the PHD-finger domain.

[0082] The nucleosome HAT activity of YNG2p in the context of a NuA4 complex or piccolo NuA4 complex can be determined by comparing the *in vitro* nucleosome HAT activity of a complex isolated from cells comprising a modified YNG2 gene to that of a complex isolated from cells comprising the wild-type YNG2 gene. Relevant methods for such analyses, including methods for isolation of the NuA4 and piccolo NuA4 complexes, purification of nucleosome histone substrates, and *in vitro* HAT assays, are well known in the art, and are described, for example, by Eberharter et al.

Methods (1998) 15(4):315, Utley at al. *Nature* (1998) 394(6692):498, Lin et al. *Cell* (2009) 136:1073, and Selleck et al. *Mol. Cell. Biol.* (2005) 25(13):5535.

5.2.6 Galactose Mediated Gene Expression

[0083] In some embodiments, the modified YNG2 gene produces a modified YNG2p that when present in a microorganism leads to increased galactose-regulated gene expression. In some such embodiments, the modified YNG2p comprises a modified PHD-finger domain. In other such embodiments, the modified YNG2p comprises a wild-type PHD-finger domain but includes one or more modifications outside of the PHD-finger domain. In yet other such embodiments, the modified YNG2p lacks a segment of the wild-type YNG2p amino acid sequence. In some such embodiments, the omitted segment comprises all or a part of the PHD-finger domain. In other such embodiments, the omitted segment comprises all or a part of YNG2p outside of the PHD-finger domain. In some embodiments, the modified YNG2p effects increased galactose-regulated gene expression only when the microorganism consumes ethanol as a primary carbon source.

[0084] Galactose-regulated gene expression at the mRNA level can, for example, be determined by comparing the *in vitro* or *in vivo* expression of coding sequences operably linked to galactose-inducible promoters using well known techniques such as Northern blotting, reverse transcription quantitative PCR (RT-PCR), expression profiling using microarrays, and serial analysis of gene expression (SAGE). Galactose-regulated gene expression at the protein level can, for example, be determined by comparing the *in vitro* or *in vivo* production of proteins from coding sequences operably linked to galactose-inducible promoters using well known techniques such as Western blotting, fluorescent microscopy, ELISA, and various chromatographic methods.

5.2.7 Selectable Markers

[0085] In some embodiments, the chromosomal integration vector used to genetically modify a microorganism disclosed herein comprises one or more selectable markers useful for the selection of transformed microorganisms.

[0086] In some embodiments, the selectable marker is an antibiotic resistance marker. Illustrative examples of antibiotic resistance markers include, but are not limited to the *BLA, NAT1, PAT, AUR1-C, PDR4, SMR1, CAT*, mouse dhfr, *HPH, DSDA,* KAN^R, and *SH BLE* gene products. The *BLA* gene product from *E. coli* confers resistance to beta-lactam antibiotics (*e.g.,* narrow-spectrum cephalosporins, cephamycins, and carbapenems (ertapenem), cefamandole, and cefoperazone) and to all the anti-gram-negative-bacterium penicillins except temocillin; the *NAT1* gene product from *S. noursei* confers resistance to nourseothricin; the *PAT* gene product from *S. viridochromogenes* Tu94 confers resistance to bialophos; the *AUR1-C* gene product from *Saccharomyces cerevisiae* confers resistance to Auerobasidin A (AbA); the PDR4 gene product confers resistance to cerulenin; the SMR1 gene product confers resistance to sulfometuron methyl; the CAT gene product from Tn9 transposon confers resistance to chloramphenicol; the mouse dhfr gene product confers resistance to methotrexate; the *HPH* gene product of *Klebsiella pneumonia* confers resistance to Hygromycin B; the *DSDA* gene product of *E. coli* allows cells to grow on plates with D-serine as the sole nitrogen source; the KAN^R gene of the Tn903 transposon confers resistance to G418; and the *SH BLE* gene product from *Streptoalloteichus hindustanus* confers resistance to Zeocin (bleomycin). In some embodiments, thes antibiotic resistance marker is deleted after the genetically modified microorganism disclosed herein is isolated.

[0087] In some embodiments, the selectable marker rescues an auxotrophy (*e.g.,* a nutritional auxotrophy) in the genetically modified microorganism. In such embodiments, a parent microorganism comprises a functional disruption in one or more gene products that function in an amino acid or nucleotide biosynthetic pathway, such as, for example, the *HIS3, LEU2, LYS1, LYS2, MET15*, *TRP1, ADE2,* and *URA3* gene products in yeast, which renders the parent microorganism incapable of growing in media without supplementation with one or more nutrients (auxotrophic phenotype). The auxotrophic phenotype can then be rescued by transforming the parent microorganism with an expression vector or chromosomal integration construct encoding a functional copy of the disrupted gene product, and the genetically modified microorganism generated can be selected for based on the loss of the auxotrophic phenotype of the parent microorganism. Utilization of the *URA3, TRP1*, and *LYS2* genes as selectable markers has a marked advantage because both positive and negative selections are possible. Positive selection is carried out by auxotrophic complementation of the *URA3, TRP1*, and *LYS2* mutations, whereas negative selection is based on specific inhibitors, *i.e.,* 5-fluoro-orotic acid (FOA), 5-fluoroanthranilic acid, and a-aminoadipic acid (aAA), respectively, that prevent growth of the prototrophic strains but allows growth of the *URA3, TRP1*, and *LYS2* mutants, respectively.

[0088] In other embodiments, the selectable marker rescues other non-lethal deficiencies or phenotypes that can be identified by a known selection method.

[0089] In some embodiments in which the YNG2 gene is modified in a microorganism by introducing into the microorganism more than one chromosomal integration construct, each chromosomal integration construct can comprise a nucleotide sequence encoding only a segment of a selectable marker, wherein such segment comprises the construct homologous sequence, such that only integration of all the chromosomal integration constructs by homologous recombination at the YNG2 homologous sequences and at the construct homologous sequences creates functional selectable markers that can be used to identify microorganisms comprising the desired modification of the YNG2 gene.

5.2.8 Microorganism Transformations

[0090] Chromosomal integration constructs can be introduced into microorganisms by any method known to one of skill in the art without limitation. *See,* for example, Hinnen *et al.,* Proc. Natl. Acad. Sci. USA 75:1292-3 (1978); Cregg *et al.,* Mol. Cell. Biol. 5:3376-3385 (1985); U.S. Patent NO: 5,272,065; Goeddel et al., eds, 1990, Methods in Enzymology, vol. 185, Academic Press, Inc., CA; Krieger, 1990, Gene Transfer and Expression -- A Laboratory Manual, Stockton Press, NY; Sambrook et al., 1989, Molecular Cloning -- A Laboratory Manual, Cold Spring Harbor Laboratory, NY; and Ausubel et al., eds., Current Edition, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, NY. Exemplary techniques include, but are not limited to, spheroplasting, electroporation, PEG 1000 mediated transformation, and lithium acetate or lithium chloride mediated transformation.

5.2.9 Culturing Genetically Modified Microbes

[0091] Also provided herein are methods for producing an acetyl-CoA derived compound. The methods generally involve culturing the genetically modified microorganisms disclosed herein under suitable conditions in a suitable medium comprising a carbon source.

[0092] Suitable conditions and suitable media for culturing microorganisms are well known in the art. In some embodiments, the suitable medium is supplemented with one or more additional agents, such as, for example, an inducer (*e.g.,* when one or more nucleotide sequences encoding a gene product are under the control of an inducible promoter), a repressor (*e.g.,* when one or more nucleotide sequences encoding a gene product are under the control of a repressible promoter), or a selection agent (*e.g.,* an antibiotic to select for microorganisms comprising the genetic modifications).

[0093] In some embodiments, the carbon source is a monosaccharide (simple sugar), a disaccharide, a polysaccharide, a non-fermentable carbon source, or one or more combinations thereof. Non-limiting examples of suitable monosaccharides include glucose, galactose, mannose, fructose, ribose, and combinations thereof. Non-limiting examples of suitable disaccharides include sucrose, lactose, maltose, trehalose, cellobiose, and combinations thereof. Non-limiting examples of suitable polysaccharides include starch, glycogen, cellulose, chitin, and combinations thereof. Non-limiting examples of suitable non-fermentable carbon sources include acetate and glycerol.

5.2.10 Extracting Acetyl-CoA Derived Compounds from Fermentation

[0094] The acetyl-CoA derived compound produced by the genetically modified microorganisms described herein may be isolated from the fermentation using any suitable separation and purification methods known in the art.

[0095] In some embodiments, an organic phase comprising the acetyl-CoA derived compound is separated from the fermentation by centrifugation. In other embodiments, an organic phase comprising the acetyl-CoA derived compound separates from the fermentation spontaneously. In yet other embodiments, an organic phase comprising the acetyl-CoA derived compound is separated from the fermentation by adding a deemulsifier and/or a nucleating agent into the fermentation reaction. Illustrative examples of deemulsifiers include flocculants and coagulants. Illustrative examples of nucleating agents include droplets of the acetyl-CoA derived compound itself and organic solvents such as dodecane, isopropyl myristrate, and methyl oleate.

[0096] In some embodiments, the acetyl-CoA derived compound is separated from other products that may be present in the organic phase. In some embodiments, separation is achieved using adsorption, distillation, gas-liquid extraction (stripping), liquid-liquid extraction (solvent extraction), ultrafiltration, and standard chromatographic techniques.

[0097] In some embodiments, the acetyl-CoA derived compound is pure, *e.g.,* at least about 40% pure, at least about 50% pure, at least about 60% pure, at least about 70% pure, at least about 80% pure, at least about 90% pure, at least about 95% pure, at least about 98% pure, or more than 98% pure, where "pure" in the context of an acetyl-CoA derived compound refers to an acetyl-CoA derived compound that is free from other acetyl-CoA derived compounds, contaminants, etc.

5.2.11 Heterologous Nucleotide Sequences Encoding Biosynthetic Enzymes for the Production of Isoprenoids

[0098] In some embodiments, the acetyl-CoA derived compound is an isoprenoid. Isoprenoids are derived from IPP, which in yeast is biosynthesized by enzymes of the MEV pathway (FIG. 1). IPP generated via the MEV pathway can be converted to its isomer, DMAPP, condensed, and modified through the action of various additional enzymes to form simple and more complex acetyl-CoA derived isoprenoid compounds (FIG. 2).

[0099] In embodiments in which the acetyl-CoA derived compound is an isoprenoid, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme selected from the group consisting of MEV pathway enzymes, IPP isomerases, polyprenyl synthases, and enzymes that can modify a polyprenyl to form a hemiterpene, a monoterpene, a sesquiterpene, a diterpene, a triterpene, a tetraterpene, a polyterpene, a steroid compound, a carotenoid, or a modified acetyl-CoA derived compound.

[00100] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can condense two molecules of acetyl-coenzyme A to form acetoacetyl-CoA, e.g., an acetyl-CoA thiolase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (NC_000913 REGION: 2324131..2325315; *Escherichia coli*), (D49362; *Paracoccus denitrificans*), and (L20428; *Saccharomyces cerevisiae*).

[00101] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can condense acetoacetyl-CoA with another molecule of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), e.g., a HMG-CoA synthase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (NC_001145. complement 19061..20536; *Saccharomyces cerevisiae*), (X96617; *Saccharomyces cerevisiae*), (X83882; *Arabidopsis thaliana*), (AB037907; *Kitasatospora griseola*), (BT007302; *Homo sapiens*), and (NC_002758, Locus tag SAV2546, GeneID 1122571; *Staphylococcus aureus*).

[00102] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can convert HMG-CoA into mevalonate, e.g., a HMG-CoA reductase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (NM_206548; *Drosophila melanogaster*), (NC_002758, Locus tag SAV2545, GeneID 1122570; *Staphylococcus aureus*), (NM_204485; *Gallus gallus*), (AB015627; *Streptomyces sp.* KO 3988), (AF542543; *Nicotiana attenuata*), (AB037907; *Kitasatospora griseola*), (AX128213, providing the sequence encoding a truncated HMGR; *Saccharomyces cerevisiae*), and (NC_001145: complement (115734..118898; *Saccharomyces cerevisiae*).

[00103] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can convert mevalonate into mevalonate 5-phosphate, e.g., a mevalonate kinase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (L77688; *Arabidopsis thaliana*), and (X55875; *Saccharomyces cerevisiae*).

[00104] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can convert mevalonate 5-phosphate into mevalonate 5-pyrophosphate, e.g., a phosphomevalonate kinase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (AF429385; *Hevea brasiliensis*), (NM_006556; *Homo sapiens*), and (NC_001145. complement 712315..713670; *Saccharomyces cerevisiae*).

[00105] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can convert mevalonate 5-pyrophosphate into IPP, e.g., a mevalonate pyrophosphate decarboxylase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (X97557; *Saccharomyces cerevisiae*), (AF290095; *Enterococcus faecium*), and (U49260; *Homo sapiens*).

[00106] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can convert IPP generated via the MEV pathway into DMAPP, e.g., an IPP isomerase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (NC_000913, 3031087..3031635; *Escherichia coli*), and (AF082326; *Haematococcus pluvialis*).

[00107] In some embodiments, the genetically modified microorganism disclosed herein further comprises a heterologous nucleotide sequence encoding a polyprenyl synthase that can condense IPP and/or DMAPP molecules to form polyprenyl compounds containing more than five carbons.

[00108] In some embodiments, the genetically modified microorganism disclosed herein comprises a heterologous nucleotide sequence encoding an enzyme that can condense one molecule of IPP with one molecule of DMAPP to form one molecule of GPP, e.g., a GPP synthase. Illustrative examples of nucleotide sequences encoding such an enzyme include but are not limited to: (AF513111; *Abies grandis*), (AF513112; *Abies grandis*), (AF513113; *Abies grandis*), (AY534686; *Antirrhinum majus*), (AY534687; *Antirrhinum majus*), (Y17376; *Arabidopsis thaliana*), (AE016877, Locus AP11092; *Bacillus cereus*; ATCC 14579), (AJ243739; *Citrus sinensis*), (AY534745; *Clarkia breweri*), (AY953508; *Ips pini*), (DQ286930; *Lycopersicon esculentum*), (AF182828; *Mentha x piperita*), (AF182827; *Mentha x piperita*), (MPI249453; Mentha x piperita), (PZE431697, Locus CAD24425; *Paracoccus zeaxanthinifaciens*), (AY866498; *Picrorhiza kurrooa*), (AY351862; *Vitis vinifera*), and (AF203881, Locus AAF12843; *Zymomonas mobilis*).

[00109] In some embodiments, the genetically modified microorganism disclosed herein comprises a heterologous nucleotide sequence encoding an enzyme that can condense two molecules of IPP with one molecule of DMAPP, or add a molecule of IPP to a molecule of GPP, to form a molecule of FPP, e.g., a FPP synthase. Illustrative examples of nucleotide sequences that encode such an enzyme include but are not limited to: (ATU80605; *Arabidopsis thaliana*), (ATHFPS2R; *Arabidopsis thaliana*), (AAU36376; *Artemisia annua*), (AF461050; *Bos taurus*), (D00694; *Escherichia coli* K-12), (AE009951, Locus AAL95523; *Fusobacterium nucleatum subsp. nucleatum* ATCC 25586), (GFFPPSGEN; *Gibberella fujikuroi*), (CP000009, Locus AAW60034; *Gluconobacter oxydans* 621H), (AF019892; *Helianthus annuus*), (HUMFAPS; *Homo sapiens*), (KLPFPSQCR; *Kluyveromyces lactis*), (LAU15777; *Lupinus albus*), (LAU20771; *Lupinus albus*), (AF309508; *Mus musculus*), (NCFPPSGEN; *Neurospora crassa*), (PAFPS1; *Parthenium argentatum*), (PAFPS2; *Parthenium argentatum*), (RATFAPS; *Rattus norvegicus*), (YSCFPP; *Saccharomyces cerevisiae*), (D89104; *Schizosaccharomyces pombe*), (CP000003, Locus AAT87386; *Streptococcus pyogenes*), (CP000017, Locus AAZ51849; *Streptococcus pyogenes*), (NC_008022, Locus YP_598856; *Streptococcus pyogenes* MGAS10270), (NC_008023, Locus YP_600845; *Streptococcus pyogenes* MGAS2096), (NC_008024, Locus YP_602832; *Streptococcus pyogenes* MGAS10750), (MZEFPS; *Zea mays*), (AE000657, Locus AAC06913; *Aquifex aeolicus* VF5), (NM_202836; *Arabidopsis thaliana*), (D84432, Locus BAA12575; *Bacillus subtilis*), (U12678, Locus AAC28894; *Bradyrhizobium japonicum* USDA 110), (BACFDPS; *Geobacillus stearothermophilus*), (NC_002940, Locus NP_873754; *Haemophilus ducreyi* 35000HP), (L42023, Locus AAC23087; *Haemophilus influenzae* Rd KW20), (J05262; *Homo sapiens*), (YP_395294; *Lactobacillus sakei subsp. sakei* 23K), (NC_005823, Locus YP_000273; *Leptospira interrogans serovar Copenhageni str. Fiocruz* L1-130), (AB003187; *Micrococcus luteus*), (NC_002946, Locus YP_208768; *Neisseria gonorrhoeae* FA 1090), (U00090, Locus AAB91752; *Rhizobium sp.* NGR234), (J05091; *Saccharomyces cerevisae*), (CP000031, Locus AAV93568; *Silicibacter pomeroyi* DSS-3), (AE008481, Locus AAK99890; *Streptococcus pneumoniae* R6), and (NC_004556, Locus NP 779706; *Xylella fastidiosa* Temecula1).

[00110] In some embodiments, the genetically modified microorganism disclosed herein further comprises a heterologous nucleotide sequence encoding an enzyme that can combine IPP and DMAPP or IPP and FPP to form GGPP. Illustrative examples of nucleotide sequences that encode such an enzyme include but are not limited to: (ATHGERPYRS; *Arabidopsis thaliana*), (BT005328; *Arabidopsis thaliana)*, (NM_119845; *Arabidopsis thaliana*), (NZ_AAJM01000380, Locus ZP_00743052; *Bacillus thuringiensis serovar israelensis,* ATCC 35646 sq1563), (CRGGPPS; *Catharanthus roseus*), (NZ_AABF02000074, Locus ZP_00144509; *Fusobacterium nucleatum subsp. vincentii,* ATCC 49256), (GFGGPPSGN; *Gibberella fujikuroi*), (AY371321; *Ginkgo biloba*), (AB055496; *Hevea brasiliensis*), (AB017971; *Homo sapiens*), (MCI276129; *Mucor circinelloides f. lusitanicus*), (AB016044; *Mus musculus*), (AABX01000298, Locus NCU01427; *Neurospora crassa*), (NCU20940; *Neurospora crassa*), (NZ_AAKL01000008, Locus ZP_00943566; *Ralstonia solanacearum* UW551), (AB118238; *Rattus norvegicus*), (SCU31632; *Saccharomyces cerevisiae*), (AB016095; Synechococcus elongates), (SAGGPS; *Sinapis alba*), (SSOGDS; *Sulfolobus acidocaldarius*), (NC_007759, Locus YP_461832; *Syntrophus aciditrophicus* SB), (NC_006840, Locus YP_204095; *Vibrio fischeri* ES114), (NM_112315; *Arabidopsis thaliana*), (ERWCRTE; *Pantoea agglomerans*), (D90087, Locus BAA14124; *Pantoea ananatis*), (X52291, Locus CAA36538; *Rhodobacter capsulatus*), (AF195122, Locus AAF24294; *Rhodobacter sphaeroides*), and (NC_004350, Locus NP_721015; *Streptococcus mutans* UA159).

[00111] In some embodiments, the genetically modified microorganism disclosed herein further comprises a heterologus nucleotide sequence encoding an enzyme that can modify a polyprenyl to form a hemiterpene, a monoterpene, a sesquiterpene, a diterpene, a triterpene, a tetraterpene, a polyterpene, a steroid compound, a carotenoid, or a modified acetyl-CoA derived compound.

[00112] In some embodiments, the heterologous nucleotide encodes a carene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (AF461460, REGION 43..1926; *Picea abies*) and (AF527416, REGION: 78..1871; *Salvia stenophylla*).

[00113] In some embodiments, the heterologous nucleotide encodes a geraniol synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (AJ457070; *Cinnamomum tenuipilum*), (AY362553; *Ocimum basilicum*), (DQ234300; *Perilla frutescens* strain 1864), (DQ234299; *Perilla citriodora* strain 1861), (DQ234298; *Perilla citriodora* strain 4935), and (DQ088667; *Perilla citriodora*).

[00114] In some embodiments, the heterologous nucleotide encodes a linalool synthase. Illustrative examples of a suitable nucleotide sequence include but are not limited to: (AF497485; *Arabidopsis thaliana*), (AC002294, Locus AAB71482; *Arabidopsis thaliana*), (AY059757; *Arabidopsis thaliana*), (NM_104793; *Arabidopsis thaliana*), (AF154124; *Artemisia annua*), (AF067603; *Clarkia breweri*), (AF067602; *Clarkia concinna*), (AF067601; *Clarkia breweri*), (U58314; *Clarkia breweri*), (AY840091; *Lycopersicon esculentum*), (DQ263741; *Lavandula angustifolia*), (AY083653; *Mentha citrate*), (AY693647; *Ocimum basilicum*), (XM_463918; *Oryza sativa*), (AP004078, Locus BAD07605; *Oryza sativa*), (XM_463918, Locus XP_463918; *Oryza sativa*), (AY917193; *Perilla citriodora*), (AF271259; *Perilla frutescens*), (AY473623; *Picea abies*), (DQ195274; *Picea sitchensis*), and (AF444798; *Perilla frutescens* var. crispa cultivar NO: 79).

[00115] In some embodiments, the heterologous nucleotide encodes a limonene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (+)-limonene synthases (AF514287, REGION: 47..1867; *Citrus limon*) and (AY055214, REGION: 48..1889; *Agastache rugosa*) and (-)-limonene synthases (DQ195275, REGION: 1..1905; *Picea sitchensis*), (AF006193, REGION: 73..1986; *Abies grandis*), and (MHC4SLSP, REGION: 29..1828; *Mentha spicata*).

[00116] In some embodiments, the heterologous nucleotide encodes a myrcene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (U87908; *Abies grandis*), (AY195609; *Antirrhinum majus*), (AY195608;

Antirrhinum majus), (NM_127982; *Arabidopsis thaliana* TPS10), (NM_113485; *Arabidopsis thaliana* ATTPS-CIN), (NM_113483; *Arabidopsis thaliana* ATTPS-CIN), (AF271259; *Perilla frutescens*), (AY473626; *Picea abies*), (AF369919; *Picea abies*), and (AJ304839; *Quercus ilex*).

[00117] In some embodiments, the heterologous nucleotide encodes an ocimene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (AY195607; *Antirrhinum majus*), (AY195609; *Antirrhinum majus*), (AY195608; *Antirrhinum majus*), (AK221024; *Arabidopsis thaliana*), (NM_113485; *Arabidopsis thaliana* ATTPS-CIN), (NM_113483; *Arabidopsis thaliana* ATTPS-CIN), (NM_117775; *Arabidopsis thaliana* ATTPS03), (NM_001036574; *Arabidopsis thaliana* ATTPS03), (NM_127982; *Arabidopsis thaliana* TPS10), (AB110642; *Citrus unshiu* CitMTSL4), and (AY575970; *Lotus corniculatus var. japonicus*).

[00118] In some embodiments, the heterologous nucleotide encodes an α-pinene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (+) α-pinene synthase (AF543530, REGION: 1..1887; *Pinus taeda*), (-)α-pinene synthase (AF543527, REGION: 32..1921; *Pinus taeda*), and (+)/(-)α-pinene synthase (AGU87909, REGION: 6111892; *Abies grandis*).

[00119] In some embodiments, the heterologous nucleotide encodes a β-pinene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (-) β-pinene synthases (AF276072, REGION: 1..1749; *Artemisia annua*) and (AF514288, REGION: 26..1834; *Citrus limon*).

[00120] In some embodiments, the heterologous nucleotide encodes a sabinene synthase. An illustrative example of a suitable nucleotide sequence includes but is not limited to AF051901, REGION: 26..1798 from *Salvia officinalis*.

[00121] In some embodiments, the heterologous nucleotide encodes a γ-terpinene synthase. Illustrative examples of suitable nucleotide sequences include: (AF514286, REGION: 30..1832 from *Citrus limon*) and (AB110640, REGION 1..1803 from *Citrus unshiu*).

[00122] In some embodiments, the heterologous nucleotide encodes a terpinolene synthase. Illustrative examples of a suitable nucleotide sequence include but is not limited to: (AY693650 from *Oscimum basilicum*) and (AY906866, REGION: 10..1887 from *Pseudotsuga menziesii*).

[00123] In some embodiments, the heterologous nucleotide encodes an amorphadiene synthase. An illustrative example of a suitable nucleotide sequence is SEQ ID NO: 37 of U.S. Patent Publication NO: 2004/0005678.

[00124] In some embodiments, the heterologous nucleotide encodes an α-farnesene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to DQ309034 from *Pyrus communis cultivar d'Anjou* (pear; gene name AFS1) and AY182241 from *Malus domestica* (apple; gene AFS1). Pechouus *et al.*, *Planta* **219(1)**:84-94 (2004).

[00125] In some embodiments, the heterologous nucleotide encodes a β-farnesene synthase. Illustrative examples of suitable nucleotide sequences include but is not limited to GenBank accession number AF024615 from *Mentha x piperita* (peppermint; gene Tspa11), and AY835398 from *Artemisia annua*. Picaud et al., *Phytochemistry* **66(9)**: 961-967 (2005).

[00126] In some embodiments, the heterologous nucleotide encodes a farnesol synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to GenBank accession number AF529266 from *Zea mays* and YDR481C from *Saccharomyces cerevisiae* (gene Pho8). Song, L., *Applied Biochemistry and Biotechnology* **128**:149-158 (2006).

[00127] In some embodiments, the heterologous nucleotide encodes a nerolidol synthase. An illustrative example of a suitable nucleotide sequence includes but is not limited to AF529266 from *Zea mays* (maize; gene tps1).

[00128] In some embodiments, the heterologous nucleotide encodes a patchouliol synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to AY508730 REGION: 1..1659 from *Pogostemon cablin*.

[00129] In some embodiments, the heterologous nucleotide encodes a nootkatone synthase. Illustrative examples of a suitable nucleotide sequence includes but is not limited to AF441124 REGION: 1..1647 from *Citrus sinensis* and AY917195 REGION: 1..1653 from *Perilla frutescens*.

[00130] In some embodiments, the heterologous nucleotide encodes an abietadiene synthase. Illustrative examples of suitable nucleotide sequences include but are not limited to: (U50768; *Abies grandis*) and (AY473621; *Picea abies*).

5.2.12 Heterologous Nucleotide Sequences Encoding Biosynthetic Enzymes for the Production of Polyketides

[00131] In some embodiments, the acetyl-CoA derived compound is a polyketide. Polyketides are synthesized by sequential reactions catalysed by a collection of enzyme activities called polyketide synthases (PKSs), which are large multi-enzyme protein complexes that contain a coordinated group of active sites. Polyketide biosynthesis proceeds stepwise starting from simple 2-, 3-, 4-carbon building blocks such as acetyl-CoA, propionyl CoA, butyryl-CoA and their activated derivatives, malonyl-, methylmalonyl- and ethylmalonyl-CoA, primarily through decarboxylative condensation of malonyl-CoA-derived units via Claisen condensation reactions. The PKS genes are usually organized in one operon in bacteria and in gene clusters in eukaryotes. Three types of polyketide synthases have been characterized: Type I polyketide synthases are large, highly modular proteins subdivided into two classes: (1) iterative PKSs, which reuse domains in a cyclic fashion; and (2) modular PKSs, which contain a sequence of separate modules and do not repeat domains. Type II polyketide synthases are aggregates of monofunctional proteins, and Type III polyketide synthases do not use acyl carrier protein domains.

[00132] Unlike fatty acid biosynthesis, in which each successive chain elongation step is followed by a fixed sequence of ketoreduction, dehydration and enoyl, reduction as described below, the individual chain elongation intermediates of polyketide biosynthesis undergo all, some, or no functional group modifications, resulting in a large number of chemically diverse products. Additional degrees of complexity arise from the use of different starter units and chain elongation units as well as the generation of new stereo-isomers.

[00133] The order of complete polyketide-synthesis as directed by a polyketide synthase follows (in the order N-terminus to C-terminus): starting or loading the initial carbon building blocks onto an acyl carrier protein, elongation modules which catalyze the extension of the growing macrolide chain and termination modules that catalyze the release of the synthesized macrolide. Component domains or separate enzyme functionalities active in this biosynthesis include acyl-transferases for the loading of starter, extender and intermediate acyl units; acyl carrier proteins which hold the growing macrolide as a thiol ester; β-keto-acyl synthases which catalyze chain extension; β-keto reductases responsible for the first reduction to an alcohol functionality; dehydratases which eliminate water to give an unsaturated thiolester; enoyl reductases which catalyse the final reduction to full saturation; and thiolesterases which catalyze macrolide release and cyclization.

[00134] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can condense at least one of acetyl-CoA and malonyl-CoA with an acyl carrier protein, e.g. an acyl-transferase.

[00135] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can condense a first reactant selected from the group consisting of acetyl-CoA and malonyl-CoA with a second reactant selected from the group consisting of malonyl-CoA or methylmalonyl-CoA to form a polyketide product, e.g. a β-keto-acyl synthase.

[00136] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can reduce a β-keto chemical group on a polyketide compound to a β-hydroxy group, e.g. a β-keto reductase.

[00137] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can dehydrate an alkane chemical group in a polyketide compound to produce an α-β-unsaturated alkene, e.g. a dehydratase.

[00138] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can reduce an α-β-double-bond in a polyketide compound to a saturated alkane, e.g. an enoyl-reductase.

[00139] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can hydrolyze a polyketide compound from an acyl carrier protein, e.g. a thioesterase.

[00140] In some embodiments, the genetically modified microorganism disclosed herein comprises heterologous nucleotide sequences, for example sequences encoding PKS enzymes and polyketide modification enzymes, capable of producing a polyketide selected from, but not limited to, the following polyketides: Avermectin (*see, e.g.,* U.S. Pat. NO: 5,252,474; U.S. Pat. NO: 4,703,009; EP Pub. NO: 118,367; MacNeil *et al.*, 1993, "Industrial Microorganisms: Basic and Applied Molecular Genetics"; Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, "A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin"; MacNeil *et al.*, 1992, *Gene* 115: 119-125; and Ikeda and Omura, 1997, *Chem. Res*. 97: 2599-2609); Candicidin (FR008) (*see, e.g.,* Hu *et a*l., 1994, *Mol. Microbiol.* 14: 163-172); Carbomycin, Curamycin (*see, e.g.,* Bergh *et al., Biotechnol Appl Biochem.* 1992 Feb;15(1):80-9); Daunorubicin (*see, e.g., J Bacteriol.* 1994 Oct;176(20):6270-80); Epothilone (*see, e.g.,* PCT Pub. NO: 99/66028; and PCT Pub. NO: 00/031247); Erythromycin (*see, e.g.,* PCT Pub. NO: 93/13663; U.S. Pat. NO: 6,004,787; U.S. Pat. NO: 5,824,513; Donadio *et al.*, 1991, *Science* 252:675-9; and Cortes *et al.*, Nov. 8, 1990, *Nature* 348:176-8); FK-506 (*see, e.g.,* Motamedi *et al.*, 1998; *Eur. J Biochem.* 256: 528-534; and Motamedi *et al.*, 1997, *Eur. J Biochem.* 244: 74-80); FK-520 (*see, e.g.,* PCT Pub. NO: 00/020601; and Nielsen *et al.*, 1991, *Biochem.* 30:5789-96); Griseusin (*see, e.g.,* Yu *et al., J Bacteriol.* 1994 May;176(9):2627-34); Lovastatin (*see, e.g.,* U.S. Pat. NO: 5,744,350); Frenolycin (*see, e.g.,* Khosla *et al., Bacteriol.* 1993 Apr;175(8):2197-204; and *Bibb et al., Gene* 1994 May 3;142(1):31-9); Granaticin (*see, e.g.,* Sherman *et al.,* EMBO J. 1989 Sep;8(9):2717-25; and Bechtold *et al., Mol Gen Genet.* 1995 Sep 20;248(5):610-20); Medermycin (*see, e.g.,* Ichinose *et al., Microbiology* 2003 Jul;149(Pt 7):1633-45); Monensin (*see, e.g.,* Arrowsmith *et al., Mol Gen Genet.* 1992 Aug;234(2):254-64); Nonactin (*see, e.g., FEMS Microbiol Lett.* 2000 Feb 1;183(1):171-5); Nanaomycin (*see, e.g.,* Kitao *et al., J Antibiot* (Tokyo). 1980 Jul;33(7):711-6); Nemadectin (*see, e.g.,* MacNeil *et al.*, 1993, supra); Niddamycin (*see, e.g.,* PCT Pub. NO: 98/51695; and Kakavas *et al.*, 1997, *J. Bacteriol.* 179: 7515-7522); Oleandomycin (*see e.g.,* Swan *et al.*, 1994, *Mol. Gen. Genet.* 242: 358-362; PCT Pub. NO: 00/026349; Olano *et al.*, 1998, *Mol. Gen. Genet.* 259(3): 299-308; and PCT Pat. App. Pub. NO: WO 99/05283); Oxytetracycline (*see, e.g.,* Kim *et al., Gene.* 1994 Apr 8;141(1):141-2); Picromycin (*see, e.g.,* PCT Pub. NO: 99/61599; PCT Pub. NO: 00/00620; Xue *et al.*, 1998, *Chemistry & Biology* 5(11): 661-667; Xue *et al.*, October 1998, *Proc. Natl. Acad. Sci.* USA 95: 12111 12116); Platenolide (*see, e.g.,* EP Pub. NO: 791,656; and U.S. Pat. NO: 5,945,320); Rapamycin (*see, e.g.,* Schwecke *et al.*, August 1995, *Proc. Natl. Acad. Sci.* USA 92:7839-7843; and Aparicio *et al.*, 1996, Gene 169: 9-16); Rifamycin (*see, e.g.,* PCT Pub. NO: WO 98/07868; and August *et al.*, Feb. 13, 1998, *Chemistry & Biology*, 5(2): 69-79); Sorangium (*see, e.g.,* U.S. Pat. NO: 6,090,601); Soraphen (*see, e.g.,* U.S. Pat. NO: 5,716,849; Schupp *et a*l., 1995, *J. Bacteriology* 177: 3673-3679); Spinocyn (*see, e.g.,* PCT Pub. NO: 99/46387); Spiramycin (*see, e.g.,* U.S. Pat. NO: 5,098,837); Tetracenomycin (*see, e.g.,* Summers *et al., J Bacteriol.* 1992 Mar;174(6):1810-20; and Shen *et al., J Bacteriol.* 1992 Jun;174(11):3818-21);

Tetracycline (*see, e.g., J Am Chem Soc.* 2009 Dec 9;131(48):17677-89); Tylosin (*see, e.g.,* U.S. Pat. NO: 5,876,991; U.S. Pat. NO: 5,672,497; U.S. Pat. NO: 5,149,638; EP Pub. NO: 791,655; EP Pub. NO: 238,323; Kuhstoss *et al.*, 1996, Gene 183:231-6; and Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355); and 6-methylsalicyclic acid (*see, e.g.,* Richardson *et al.*, *Metab Eng.* 1999 Apr;1(2):180-7; and Shao *et al., Biochem Biophys Res Commun.* 2006 Jun 23;345(1):133-9).

5.2.13 Heterologous Nucleotide Sequences Encoding Biosynthetic Enzymes for the Production of Fatty Acids

[00141] In some embodiments, the acetyl-CoA derived compound is a fatty acid. Fatty acids are synthesized by a series of decarboxylative Claisen condensation reactions from acetyl-CoA and malonyl-CoA catalyzed by fatty acid synthases. Similar to polyketide synthases, fatty acid synthases are not a single enzyme but an enzymatic system composed of 272 kDa multifunctional polypeptide in which substrates are handed from one functional domain to the next. Two principal classes of fatty acid synthases have been characterized: Type I fatty acid synthases are single, multifunctional polypeptides common to mammals and fungi (although the structural arrangement of fungal and mammalian synthases differ) and the CMN group of bacteria (corynebacteria, mycobacteria, and nocardia). Type II synthases, found in archaeabacteria and eubacteria, are a series of discrete, monofunctional enzymes that participate in the synthesis of fatty acids. The mechanisms fatty acid elongation and reduction is the same in the two classes of synthases, as the enzyme domains responsible for these catalytic events are largely homologous amongst the two classes.

[00142] Following each round of elongation of the fatty acid chain in the decarboxylative Claisen condensation reactions, the β-keto group is reduced to a fully saturated carbon chain by the sequential action of a ketoreductase, a dehydratase, and an enol reductase. The growing fatty acid chain moves between these active sites attached to an acyl carrier protein and is ultimately released by the action of a thioesterase upon reaching a carbon chain length of 16 (palmitidic acid).

[00143] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can covalently link at least one of acetyl-CoA and malonyl-CoA with an acyl carrier protein, e.g. an acyl-transferase.

[00144] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can condense acetyl chemical moiety and a malonyl chemical moiety, each bound to an acyl carrier protein (ACP), to form acetoacetyl-ACP, e.g. a β-Ketoacyl-ACP synthase.

[00145] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can reduce the double bond in acetoacetyl-ACP with NADPH to form a hydroxyl group in D-3-hydroxybutyryl hydroxylase-ACP, e.g. a β-Ketoacyl-ACP reductase.

[00146] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can dehydrate D-3-Hydroxybutyryl hydroxylase-ACP to create a double bond between the beta- and gamma-carbons forming crotonyl-ACP, e.g. a β-hydroxyacyl-ACP dehydrase.

[00147] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can reduce crotonyl ACP with NADPH to form butyryl-ACP, e.g. an enoyl ACP reductase.

[00148] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can hydrolyze a C16 acyl compound from an acyl carrier protein to form palmitate, e.g. a thioesterase.

[00149] In some embodiments, the genetically modified microorganism disclosed herein comprises a second heterologous nucleotide sequence encoding an enzyme that can increase acetyl-CoA production, *e.g., pdh*, *panK*, *aceEF* (encoding the Elp dehydrogenase component and the E2p dihydrolipoamide acyltransferase component of the pyruvate and 2-oxoglutarate dehydrogenase complexes), *fabH*, *fabD*, *fabG*, *acpP*, and *fabF*. Illustrative examples of nucleotide sequences encoding such enzymes include, but are not limited to: *pdh* (BAB34380, AAC73227, AAC73226), *panK* (also known as coaA, AAC76952), *aceEF* (AAC73227, AAC73226), *fabH* (AAC74175), fabD (AAC74176), *fabG* (AAC74177), *acpP* (AAC74178), *fabF* (AAC74179).

[00150] In some embodiments, increased fatty acid levels can be effected in the cell by attenuating or knocking out genes encoding proteins involved in fatty acid degradation. For example, the expression levels of *fadE*, *gpsA*, *idhA*, *pflb*, *adhE*, *pta*, *poxB*, *ackA*, and/or *ackB* can be attenuated or knocked-out in an engineered host cell using techniques known in the art. Illustrative examples of nucleotide sequences encoding such proteins include, but are not limited to: *fadE* (AAC73325), *gspA* (AAC76632), *IdhA* (AAC74462), *pflb* (AAC73989), *adhE* (AAC74323), *pta* (AAC75357), *poxB* (AAC73958), *ackA* (AAC75356), and *ackB* (BAB81430). The resulting host cells will have increased acetyl-CoA production levels when grown in an appropriate environment.

[00151] In some embodiments, the fatty acid producing cell comprises a heterologous nucleotide sequence encoding an enzyme that can convert acetyl-CoA into malonyl-CoA, *e.g.,* the multisubunit AccABCD protein. An illustrative example of a suitable nucleotide sequence encoding AccABCD includes but is not limited to accession number AAC73296, EC 6.4.1.2.

[00152] In some embodiments, the fatty acid producing cell comprises a heterologous nucleotide sequence encoding a lipase. Illustrative examples of suitable nucleotide sequences encoding a lipase include, but are not limited to accession numbers CAA89087 and CAA98876.

[00153] In some embodiments, increased fatty acid levels can be effected in the cell by inhibiting PlsB, which can lead to an increase in the levels of long chain acyl-ACP, which will inhibit early steps in the fatty acid biosynthesis pathway (*e.g.*, *accABCD*, *fabH*, and *fabI*). The expression level of PlsB can be attenuated or knocked-out in an engineered host cell using techniques known in the art. An illustrative example of a suitable nucleotide sequence encoding PlsB includes but is not limited to accession number AAC77011. In particular embodiments, the *plsB* D31 IE mutation can be used to increase the amount of available acyl-CoA in the cell.

[00154] In some embodiments, increased production of monounsaturated fatty acids can be effected in the cell by overexpressing an *sfa* gene, which would result in suppression of *fabA*. An illustrative example of a suitable nucleotide sequence encoding *sfa* includes but is not limited to accession number AAN79592.

[00155] In some embodiments, increased fatty acid levels can be effected in the cell by modulating the expression of an enzyme which controls the chain length of a fatty acid substrate, *e.g.,* a thioesterase. In some embodiments, the fatty acid producing cell has been modified to overexpress a *tes* or *fat* gene. Illustrative examples of suitable *tes* nucleotide sequences include but are not limited to accession numbers: (*tesA*: AAC73596, from *E. Coli,* capable of producing $C_{18:1}$ fatty acids) and (*tesB*: AAC73555 from *E. Coli*). Illustrative examples of suitable *fat* nucleotide sequences include but are not limited to: (*fatB*: Q41635 and AAA34215, from *Umbellularia california*, capable of producing $C_{12:0}$

fatty acids), (*fatB2*: Q39513 and AAC49269, from *Cuphea hookeriana*, capable of producing $C_{8:0}$ – $C_{10:0}$ fatty acids), (*fatB3*: AAC49269 and AAC72881, from *Cuphea hookeriana*, capable of producing $C_{14:0}$ – $C_{16:0}$ fatty acids), (*fatB*: Q39473 and AAC49151, from *Cinnamonum camphorum*, capable of producing $C_{14:0}$ fatty acids), (*fatB [M141T]:* CAA85388, from m*Arabidopsis thaliana*, capable of producing $C_{16:1}$ fatty acids), (*fatA*: NP 189147 and NP 193041, from *Arabidopsis thaliana*, capable of producing $C_{18:1}$ fatty acids), (*fatA:* CAC39106, from *Bradvrhiizobium japonicum*, capable of preferentially producing $C_{18:1}$ fatty acids), (*fatA*: AAC72883, from *Cuphea hookeriana,* capable of producing $C_{18:1}$ fatty acids), and (*fatA1*, AAL79361 from *Helianthus annus*).

[00156] In some embodiments, increased levels of C_{10} fatty acids can be effected in the cell by attenuating the expression or activity of thioesterase C_{18} using techniques known in the art. Illustrative examples of suitable nucleotide sequences encoding thioesterase C_{18} include, but are not limited to accession numbers AAC73596 and P0ADA1. In other embodiments, increased levels of C_{10} fatty acids can be effected in the cell by increasing the expression or activity of thioesterase C_{10} using techniques known in the art. An illustrative example of a suitable nucleotide sequence encoding thioesterase C_{10} includes, but is not limited to accession number Q39513.

[00157] In some embodiments, increased levels of C_{14} fatty acids can be effected in the cell by attenuating the expression or activity of endogenous thioesterases that produce non-C_{14} fatty acids, using techniques known in the art. In other embodiments, increased levels of C_{14} fatty acids can be effected in the cell by increasing the expression or activity of thioesterases that use the substrate C_{14}-ACP, using techniques known in the art. An illustrative example of a suitable nucleotide sequence encoding such a thioesterase includes, but is not limited to accession number Q39473.

[00158] In some embodiments, increased levels of C_{12} fatty acids can be effected in the cell by attenuating the expression or activity of endogenous thioesterases that produce non- C_{12} fatty acids, using techniques known in the art. In other embodiments, increased levels of C_{12} fatty acids can be effected in the cell by increasing the expression or activity of thioesterases that use the substrate C_{12}-ACP, using techniques known in the art. An illustrative example of a suitable nucleotide sequence encoding such a thioesterase includes, but is not limited to accession number Q41635.

6. EXAMPLES

Example 1

[00159] This example describes methods for making genetically modified yeast cells that can biosynthesize the acetyl-CoA derived compound β-farnesene.

[00160] Genomic DNA was isolated from *Saccharomyces cerevisiae* strains Y002 and Y003 (CEN.PK2 background MATA or MAT alpha, MATA; ura3-52; trp1-289; leu2-3,112; his3Δ1; MAL2-8C; SUC2; van Dijken et al. (2000) Enzyme Microb. Technol. 26:706-714), Y007 (S288C background MATA trp1Δ63; ATCC number 200873), and EG123 (MATA ura3; trp1; leu2; his4 can1; Michaelis & Herskowitz (1988) Mol. Cell. Biol. 8: 1309-1318). The strains were grown overnight in 10 mL liquid medium containing 1% yeast extract, 2% bacto-peptone, and 2% dextrose (YPD medium). The cultures were centrifuged at 3,100 rpm, cell pellets were washed in 10 mL ultra-pure water and re-centrifuged, and genomic DNA was extracted using the Y-DER yeast DNA extraction kit (Pierce Biotechnologies, Rockford, IL) as per manufacturer's suggested protocol. Extracted genomic DNA was re-suspended in 100 uL 10 mM Tris-Cl, pH 8.5, and OD260/280 readings were taken on a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) to determine genomic DNA concentration and purity.

[00161] DNA amplification by Polymerase Chain Reaction (PCR) was done in an Applied Biosystems 2720 Thermocycler (Applied Biosystems Inc., Foster City, CA) using the Phusion DNA Polymerase system (Finnzymes OY, Espoo, Finland or New England Biolabs, Ipswich, CA) as per manufacturer's suggested protocol. Upon completion of a PCR amplification of a DNA fragment that was to be inserted into the TOPO TA pCR2.1 cloning vector (Invitrogen, Carlsbad, CA), A nucleotide overhangs were created by adding 1 uL of Qiagen Taq Polymerase (Qiagen, Valencia, CA) to the reaction mixture and performing an additional 10 minute, 72°C PCR extension step, followed by cooling to 4°C.

[00162] Agarose gel electrophoresis was performed using a 1% TBE (0.89 M Tris, 0.89 M boric acid, 0.02 M EDTA sodium salt) agarose gel containing 0.5 ug/mL ethidium bromide, at 120 V, 400 mA for 30 minutes. DNA bands were visualized using ultraviolet light. DNA fragments were gel extracted using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, CA) according to manufacturer's suggested protocols. The purified DNA was eluted into 10 uL ultra-pure water, and OD260/280 readings were taken on a ND-1000 spectrophotometer to determine DNA concentration and purity.

[00163] Ligations were performed using High Concentration T4 DNA Ligase (New England Biolabs, Ipswich, MA) as per manufacturer's suggested protocol. For plasmid propagation, ligated constructs were transformed into *Escherichia coli* chemically competent cells (Invitrogen, Carlsbad, CA) as per manufacturer's suggested protocol. Positive transformants were selected on solid medium containing 1.5% bacto agar, 1% tryptone, 0.5% yeast extract, 1% NaCl, and 50 ug/mL of an appropriate antibiotic. Isolated transformants were grown for 16 hours in liquid Luria-Bertoni (LB) medium containing 50 ug/mL of an appropriate antibiotic, and plasmid was isolated and purified using a QIAprep Spin Miniprep kit (Qiagen, Valencia, CA) as per manufacturer's suggested protocol. Constructs were verified by performing diagnostic restriction endonuclease digestions, resolving DNA fragments on an agarose gel, and visualizing the bands using ultraviolet light. Select constructs were also verified by DNA sequencing, which was done by Elim Biopharmaceuticals Inc. (Hayward, CA).

[00164] For yeast cell transformations, 25 ml of YPD medium was inoculated with a single colony of a starting host strain. The culture was grown overnight at 30°C on a rotary shaker at 200 rpm. The OD600 of the culture was measured, and the culture was then used to inoculate 50 ml of YPD medium to an OD600 of 0.15. The newly inoculated culture was grown at 30°C on a rotary shaker at 200 rpm up to an OD600 of 0.7 to 0.9, at which point the cells were transformed with DNA. The cells were allowed to recover in YPD medium for 4 hours before they were plated on agar containing a selective agent to identify the host cell transformants.

[00165] Plasmid pAM489 was generated by inserting the ERG20-P_{GAL}-tHMGR insert of vector pAM471 into vector pAM466. Vector pAM471 was generated by inserting DNA fragment ERG20-P_{GAL}-tHMGR, which comprises the coding sequence of the ERG20 gene of *Saccharomyces cerevisiae* (ERG20 nucleotide positions +1 to +1208; with the A of the ATG codon being nucleotide position +1 and the nucleotide just upstream of the ATG codon being nucleotide position -1) (ERG20), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P_{GAL}), and a truncated coding sequence of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide positions +1586 to +3323) (tHMGR), into the TOPO Zero Blunt II cloning vector (Invitrogen, Carlsbad, CA). Vector pAM466 was generated by inserting DNA fragment TRP1$^{-856\text{ to }+548}$, which comprises a segment of the wild-type TRP1 locus of *Saccharomyces cerevisiae* that extends from nucleotide positions -856 to +548 and harbors a non-native internal *XmaI* restriction site between bases -226 and -225, into the TOPO TA pCR2.1 cloning vector (Invitrogen, Carlsbad, CA). DNA fragments ERG20-P_{GAL}-tHMGR and TRP1$^{-856 \text{ to } +548}$ were generated by PCR amplification as outlined in Table 1. For the construction of pAM489, 400 ng of pAM471 and 100 ng of pAM466 were digested to completion using *XmaI* restriction endonuclease (New England Biolabs, Ipswich, MA), DNA fragments corresponding to the ERG20-P_{GAL}-tHMGR insert and the linearized pAM466 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding pAM489. FIG. 3A shows a map of the ERG20-P_{GAL}-tHMGR insert, and SEQ ID NO: 9 shows the nucleotide sequence of the insert with flanking TRP1 sequences.

Table 1 – PCR amplifications performed to generate pAM489

PCR Round	Template	Primer 1	Primer 2	PCR Product
1	100 ng of Y003 genomic DNA	61-67-CPK001-G (SEQ ID NO: 23)	61-67-CPK002-G (SEQ ID NO: 24)	TRP1$^{-856 \text{ to } -226}$
		61-67-CPK003-G (SEQ ID NO: 25)	61-67-CPK004-G (SEQ ID NO: 26)	TRP1$^{-225\text{-to } +548}$
	100 ng of EG123 genomic DNA	61-67-CPK025-G (SEQ ID NO: 47)	61-67-CPK050-G (SEQ ID NO: 55)	ERG20
	100 ng of Y002 genomic DNA	61-67-CPK051-G (SEQ ID NO: 56)	61-67-CPK052-G (SEQ ID NO: 57)	P_{GAL}
		61-67-CPK053-G (SEQ ID NO: 58)	61-67-CPK031-G (SEQ ID NO: 48)	tHMGR
2	100 ng each of TRP1$^{-856 \text{ to } -226}$ and TRP1$^{-225\text{-to } +548}$ purified PCR products	61-67-CPK001-G (SEQ ID NO: 23)	61-67-CPK004-G (SEQ ID NO: 26)	**TRP1$^{-856 \text{ to } +548}$**
	100 ng each of ERG20 and P_{GAL} purified PCR products	61-67-CPK025-G (SEQ ID NO: 47)	61-67-CPK052-G (SEQ ID NO: 57)	ERG20-P_{GAL}
3	100 ng each of ERG20-P_{GAL} and tHMGR purified PCR products	61-67-CPK025-G (SEQ ID NO: 47)	61-67-CPK031-G (SEQ ID NO: 48)	**ERG20-P_{GAL}-tHMGR**

[00166] Plasmid pAM491 was generated by inserting the ERG13-P_{GAL}-tHMGR insert of vector pAM472 into vector pAM467. Vector pAM472 was generated by inserting DNA fragment ERG13-P_{GAL}-tHMGR, which comprises the coding sequence of the ERG13 gene of *Saccharomyces cerevisiae* (ERG13 nucleotide positions +1 to +1626) (ERG13), the genomic locus containing the divergent GAL1 and GAL10 promoter of

Saccharomyces cerevisiae (GAL1 nucleotide position -1 to -668) (P_{GAL}), and a truncated coding sequence of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide position +1586 to +3323) (tHMGR), into the TOPO Zero Blunt II cloning vector. Vector pAM467 was generated by inserting DNA fragment URA3$^{-723\text{ to }701}$, which comprises a segment of the wild-type URA3 locus of *Saccharomyces cerevisiae* that extends from nucleotide position -723 to position -224 and harbors a non-native internal *XmaI* restriction site between bases -224 and -223, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG13-P_{GAL}-tHMGR and URA3$^{-723\text{ to }701}$ were generated by PCR amplification as outlined in Table 2. For the construction of pAM491, 400 ng of pAM472 and 100 ng of pAM467 were digested to completion using *XmaI* restriction endonuclease, DNA fragments corresponding to the ERG13-P_{GAL}-tHMGR insert and the linearized pAM467 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding pAM491. FIG. 3B shows a map of the ERG13-P_{GAL}-tHMGR insert, and SEQ ID NO: 10 shows the nucleotide sequence of the insert with flanking URA3 sequences.

Table 2 – PCR amplifications performed to generate pAM491

PCR Round	Template	Primer 1	Primer 2	PCR Product
1	100 ng of Y007 genomic DNA	61-67-CPK005-G (SEQ ID NO: 27)	61-67-CPK006-G (SEQ ID NO: 28)	URA3$^{-723\text{ to }-224}$
		61-67-CPK007-G (SEQ ID NO: 29)	61-67-CPK008-G (SEQ ID NO: 30)	URA3$^{-223\text{ to }701}$
	100 ng of Y002 genomic DNA	61-67-CPK032-G (SEQ ID NO: 49)	61-67-CPK054-G (SEQ ID NO: 59)	ERG13
		61-67-CPK052-G (SEQ ID NO: 57)	61-67-CPK055-G (SEQ ID NO: 60)	P_{GAL}
		61-67-CPK031-G (SEQ ID NO: 48)	61-67-CPK053-G (SEQ ID NO: 58)	tHMGR
2	**100 ng each of URA3$^{-723\text{ to }-224}$ and URA3$^{-223\text{ to }701}$ purified PCR products**	**61-67-CPK005-G (SEQ ID NO: 27)**	**61-67-CPK008-G (SEQ ID NO: 30)**	**URA3$^{-723\text{ to }701}$**
	100 ng each of ERG13 and P_{GAL} purified PCR products	61-67-CPK032-G (SEQ ID NO: 49)	61-67-CPK052-G (SEQ ID NO: 57)	ERG13-P_{GAL}
3	100 ng each of ERG13-P_{GAL} and tHMGR purified PCR products	61-67-CPK031-G (SEQ ID NO: 48)	61-67-CPK032-G (SEQ ID NO: 49)	**ERG13-P_{GAL}-tHMGR**

[00167] Plasmid pAM493 was generated by inserting the IDI1-P_{GAL}-tHMGR insert of vector pAM473 into vector pAM468. Vector pAM473 was generated by inserting DNA fragment IDI1-P_{GAL}-tHMGR, which comprises the coding sequence of the IDI1 gene of *Saccharomyces cerevisiae* (IDI1 nucleotide position 1 to 1017) (IDI1), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P_{GAL}), and a truncated coding sequence of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide positions 1586 to 3323) (tHMGR), into the TOPO Zero Blunt II cloning vector. Vector pAM468 was generated by inserting DNA fragment ADE1$^{-825\text{ to }653}$, which comprises a segment of the wild-type ADE1 locus of *Saccharomyces cerevisiae* that extends from nucleotide position -225 to position +653 and harbors a non-native internal *XmaI* restriction site between bases -226 and -225, into the TOPO TA pCR2.1 cloning vector. DNA fragments IDI1-P_{GAL}-tHMGR and ADE1$^{-825\text{ to }653}$ were generated by PCR amplification as outlined in Table 3. For the construction of pAM493, 400 ng of pAM473 and 100 ng of pAM468 were digested to completion using *XmaI* restriction endonuclease, DNA fragments corresponding to the IDI1-P_{GAL}-tHMGR insert and the linearized pAM468 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM493. FIG. 3C shows a map of the IDI1-P_{GAL}-tHMGR insert, and SEQ ID NO: 11 shows the nucleotide sequence of the insert with flanking ADE1 sequences.

Table 3 – PCR amplifications performed to generate pAM493

PCR Round	Template	Primer 1	Primer 2	PCR Product
1	100 ng of Y007 genomic DNA	61-67-CPK009-G (SEQ ID NO: 31)	61-67-CPK010-G (SEQ ID NO: 32)	ADE1$^{-825\text{ to }-226}$
		61-67-CPK011-G (SEQ ID NO: 33)	61-67-CPK012-G (SEQ ID NO: 34)	ADE1$^{-225\text{ to }653}$
	100 ng of Y002 genomic DNA	61-67-CPK047-G (SEQ ID NO: 54)	61-67-CPK064-G (SEQ ID NO: 69)	IDI1
		61-67-CPK052-G (SEQ ID NO: 57)	61-67-CPK065-G (SEQ ID NO: 70)	P_{GAL}
		61-67-CPK031-G (SEQ ID NO: 48)	61-67-CPK053-G (SEQ ID NO: 58)	tHMGR
2	100 ng each of ADE1$^{-825\text{ to }-226}$ and ADE1$^{-225\text{ to }653}$ purified PCR products	61-67-CPK009-G (SEQ ID NO: 31)	61-67-CPK012-G (SEQ ID NO: 34)	**ADE1$^{-825\text{ to }653}$**
	100 ng each of IDI1 and P_{GAL} purified PCR products	61-67-CPK047-G (SEQ ID NO: 54)	61-67-CPK052-G (SEQ ID NO: 57)	IDI1-P_{GAL}
3	100 ng each of IDI1-P_{GAL} and tHMGR purified PCR products	61-67-CPK031-G (SEQ ID NO: 48)	61-67-CPK047-G (SEQ ID NO: 54)	**IDI1-P_{GAL}-tHMGR**

[00168] Plasmid pAM495 was generated by inserting the ERG10-P_{GAL}-ERG12 insert of pAM474 into vector pAM469. Vector pAM474 was generated by inserting DNA fragment ERG10-P_{GAL}-ERG12, which comprises the coding sequence of the ERG10 gene of *Saccharomyces cerevisiae* (ERG10 nucleotide position 1 to 1347) (ERG10), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P_{GAL}), and the coding sequence of the ERG12 gene of *Saccharomyces cerevisiae* (ERG12 nucleotide position 1 to 1482) (ERG12), into the TOPO Zero Blunt II cloning vector. Vector pAM469 was generated by inserting DNA fragment HIS3$^{-32\text{ to }-1000}$-HISMX- HIS3$^{504\text{ to }-1103}$, which comprises two segments of the HIS locus of *Saccharomyces cerevisiae* that extend from nucleotide position -32 to position -1000 and from nucleotide position +504 to position +1103, a HISMX marker, and a non-native *XmaI* restriction site between the HIS3$^{504\text{ to }-1103}$ sequence and the HISMX marker, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG10-P_{GAL}-ERG12 and HIS3$^{-32\text{ to }-1000}$-HISMX- HIS3$^{504\text{ to }-1103}$ were generated by PCR amplification as outlined in Table 4. For construction of pAM495, 400 ng of pAM474 and 100 ng of pAM469 were digested to completion using *XmaI* restriction endonuclease, DNA fragments corresponding to the ERG10-P_{GAL}-ERG12 insert and the linearized pAM469 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM495. FIG. 3D shows a map of the ERG10-P_{GAL}-ERG12 insert, and SEQ ID NO: 12 shows the nucleotide sequence of the insert with flanking HIS3 sequences.

Table 4 – PCR reactions performed to generate pAM495

PCR Round	Template	Primer 1	Primer 2	PCR Product
1	100 ng of Y007 genomic DNA	61-67-CPK013-G (SEQ ID NO: 35)	61-67-CPK014alt-G (SEQ ID NO: 36)	HIS3$^{-32\text{ to }-1000}$
		61-67-CPK017-G (SEQ ID NO: 39)	61-67-CPK018-G (SEQ ID NO: 40)	HIS3$^{504\text{ to }-1103}$
		61-67-CPK035-G (SEQ ID NO: 50)	61-67-CPK056-G (SEQ ID NO: 61)	ERG10
		61-67-CPK057-G (SEQ ID NO: 62)	61-67-CPK058-G (SEQ ID NO: 63)	P_{GAL}
		61-67-CPK040-G (SEQ ID NO: 51)	61-67-CPK059-G (SEQ ID NO: 64)	ERG12
	10 ng of plasmid pAM330 DNA **	61-67-CPK015alt-G (SEQ ID NO: 37)	61-67-CPK016-G (SEQ ID NO: 38)	HISMX
2	100 ng each of HIS3$^{504\text{ to }-1103}$ and HISMX PCR purified products	61-67-CPK015alt-G (SEQ ID NO: 37)	61-67-CPK018-G (SEQ ID NO: 40)	HISMX- HIS3$^{504\text{ to }-1103}$
	100 ng each of ERG10 and P_{GAL} purified PCR products	61-67-CPK035-G (SEQ ID NO: 50)	61-67-CPK058-G (SEQ ID NO: 63)	ERG10-P_{GAL}
3	100ng each of HIS3$^{-32\text{ to }-1000}$ and HISMX-HIS3$^{504\text{ to }-1103}$ purified PCR products	61-67-CPK013-G (SEQ ID NO: 35)	61-67-CPK018-G (SEQ ID NO: 40)	**HIS3$^{-32\text{ to }-1000}$-HISMX- HIS3$^{504\text{ to }-1103}$**
	100 ng each of ERG10-P_{GAL} and ERG12 purified PCR products	61-67-CPK035-G (SEQ ID NO: 50)	61-67-CPK040-G (SEQ ID NO: 51)	**ERG10-P_{GAL}-ERG12**

** The HISMX marker in pAM330 originated from pFA6a-HISMX6-PGAL1 as described by van Dijken *et al.* ((2000) *Enzyme Microb. Technol.* **26**(9-10):706-714).

[00169] Plasmid pAM497 was generated by inserting the ERG8-P_{GAL}-ERG19 insert of pAM475 into vector pAM470. Vector pAM475 was generated by inserting DNA fragment ERG8-P_{GAL}-ERG19, which comprises the coding sequence of the ERG8 gene of *Saccharomyces cerevisiae* (ERG8 nucleotide position +1 to +1512) (ERG8), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P_{GAL}), and the coding sequence of the ERG19 gene of *Saccharomyces cerevisiae* (ERG19 nucleotide position +1 to +1341) (ERG19), into the TOPO Zero Blunt II cloning vector. Vector pAM470 was generated by inserting DNA fragment LEU2$^{-100\text{ to }450}$-HISMX- LEU2$^{1096\text{ to }1770}$, which comprises two segments of the

LEU2 locus of *Saccharomyces cerevisiae* that extend from nucleotide position -100 to position +450 and from nucleotide position +1096 to position +1770, a HISMX marker, and a non-native *XmaI* restriction site between the LEU2$^{1096\text{ to }1770}$ sequence and the HISMX marker, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG8-P$_{GAL}$-ERG19 and LEU2$^{-100\text{ to }450}$-HISMX- LEU2$^{1096\text{ to }1770}$ were generated by PCR amplification as outlined in Table 5. For the construction of pAM497, 400 ng of pAM475 and 100 ng of pAM470 were digested to completion using *XmaI* restriction endonuclease, DNA fragments corresponding to the ERG8-P$_{GAL}$-ERG19 insert and the linearized pAM470 vector were purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM497. FIG. 3E for a map of the ERG8-P$_{GAL}$-ERG19 insert, and SEQ ID NO: 13 shows the nucleotide sequence of the insert with flanking LEU2 sequences.

Table 5 – PCR reactions performed to generate pAM497

PCR Round	Template	Primer 1	Primer 2	PCR Product
1	100 ng of Y007 genomic DNA	61-67-CPK019-G (SEQ ID NO: 41)	61-67-CPK020-G (SEQ ID NO: 42)	LEU2$^{-100 \text{ to } 450}$
		61-67-CPK023-G (SEQ ID NO: 45)	61-67-CPK024-G (SEQ ID NO: 46)	LEU2$^{1096 \text{ to } 1770}$
	10ng of plasmid pAM330 DNA **	61-67-CPK021-G (SEQ ID NO: 43)	61-67-CPK022-G (SEQ ID NO: 44)	HISMX
	100 ng of Y002 genomic DNA	61-67-CPK041-G (SEQ ID NO: 52)	61-67-CPK060-G (SEQ ID NO: 65)	ERG8
		61-67-CPK061-G (SEQ ID NO: 66)	61-67-CPK062-G (SEQ ID NO: 67)	P_{GAL}
		61-67-CPK046-G (SEQ ID NO: 53)	61-67-CPK063-G (SEQ ID NO: 68)	ERG19
2	100 ng each of LEU2$^{1096 \text{ to } 1770}$ and HISMX purified PCR products	61-67-CPK021-G (SEQ ID NO: 43)	61-67-CPK024-G (SEQ ID NO: 46)	HISMX-LEU2$^{1096 \text{ to } 1770}$
	100 ng each of ERG8 and P_{GAL} purified PCR products	61-67-CPK041-G (SEQ ID NO: 52)	61-67-CPK062-G (SEQ ID NO: 67)	ERG8-P_{GAL}
3	100 ng of LEU2$^{-100 \text{ to } 450}$ and HISMX- LEU2$^{1096 \text{ to } 1770}$ purified PCR products	61-67-CPK019-G (SEQ ID NO: 41)	61-67-CPK024-G (SEQ ID NO: 46)	**LEU2$^{-100 \text{ to } 450}$-HISMX-LEU2$^{1096 \text{ to } 1770}$**
	100 ng each of ERG8-P_{GAL} and ERG19 purified PCR products	61-67-CPK041-G (SEQ ID NO: 52)	61-67-CPK046-G (SEQ ID NO: 53)	**ERG8-P_{GAL}-ERG19**

** The HISMX marker in pAM330 originated from pFA6a-HISMX6-PGAL1 as described by van Dijken *et al.* ((2000) *Enzyme Microb. Technol.* **26**(9-10):706-714).

[00170] Expression plasmid pAM353 was generated by inserting a nucleotide sequence encoding a β-farnesene synthase into the pRS425-Gal1 vector (Mumberg et. al. (1994) Nucl. Acids. Res. 22(25): 5767-5768). The nucleotide sequence insert was generated synthetically, using as a template the coding sequence of the β-farnesene synthase gene of *Artemisia annua* (GenBank accession number AY835398) codon-optimized for expression in *Saccharomyces cerevisiae* (SEQ ID NO: 1). The synthetically generated nucleotide sequence was flanked by 5' BamHI and 3' XhoI restriction sites, and could thus be cloned into compatible restriction sites of a cloning vector such as a standard pUC or pACYC origin vector. The synthetically generated nucleotide sequence was isolated by digesting to completion the construct using BamHI and XhoI restriction endonucleases. The reaction mixture was resolved by gel electrophoresis, the approximately 1.7 kb DNA fragment comprising the β-farnesene synthase coding sequence was gel purified, and the isolated DNA fragment was ligated into the BamHI XhoI restriction site of the pRS425-Gal1 vector, yielding expression plasmid pAM353.

[00171] Expression plasmid pAM404 was generated by inserting a nucleotide sequence encoding the β-farnesene synthase of *Artemisia annua* (GenBank accession number AY835398), codon-optimized for expression in *Saccharomyces cerevisiae*, into vector pAM178 (SEQ ID NO: 2). The nucleotide sequence encoding the β-farnesene synthase was PCR amplified from pAM353 using primers 52-84 pAM326 BamHI (SEQ ID NO: 21) and 52-84 pAM326 NheI (SEQ ID NO: 22). The resulting PCR product was digested to completion using BamHI and NheI restriction endonucleases, the reaction mixture was resolved by gel electrophoresis, the approximately 1.7 kb DNA fragment comprising the β-farnesene synthase coding sequence was gel extracted, and the isolated DNA fragment was ligated into the BamHI NheI restriction site of vector pAM178, yielding expression plasmid pAM404.

[00172] Strains Y93 (MAT A) and Y94 (MAT alpha) were generated by replacing the promoter of the ERG9 gene in strains Y002 and Y003, respectively, with the promoter of the MET3 gene of *Saccharomyces cerevisiae.* The KanMX-P_{MET3} region of vector pAM328 (SEQ ID NO: 17), which comprises the P_{MET3} promoter preceded by the kanamycin resistance marker flanked by the promoter and terminator of the *Tef1* gene of *Kluyveromyces lactis*, was PCR amplified using primers 50-56-pw100-G (SEQ ID NO: 19) and 50-56-pw101-G (SEQ ID NO: 20), which include 45 base pairs of homology to the native *ERG9* promoter. 10 ug of the resulting PCR product was transformed into exponentially growing Y002 and Y003 cells using 40% w/w Polyethelene Glycol 3350 (Sigma-Aldrich, St. Louis, MO), 100 mM Lithium Acetate (Sigma-Aldrich, St. Louis, MO), and 10 ug Salmon Sperm DNA (Invitrogen Corp., Carlsbad, CA), and incubating the cells at 30°C for 30 minutes followed by heat shocking them at 42°C for 30 minutes

(Schiestl and Gietz (1989) Curr. Genet. 16:339-346). Positive recombinants were selected for growth on rich medium containing 0.5 ug/mL Geneticin (Invitrogen Corp., Carlsbad, CA), and selected clones were confirmed by diagnostic PCR.

[00173] Strains Y176 (MAT A) and Y177 (MAT alpha) were generated by replacing the coding sequence of the ADE1 gene in strains Y93 and Y94, respectively, with the LEU2 gene of *Candida glabrata* (*CgLEU2*). The 3.5 kb *CgLEU2* genomic locus was PCR amplified from *Candida glabrata* genomic DNA (ATCC, Manassas, VA) using primers 61-67-CPK066-G (SEQ ID NO: 71) and 61-67-CPK067-G (SEQ ID NO: 72), which contain 50 base pairs of flanking homology to the coding sequence of the ADE1 gene. 10 ug of the resulting PCR product were transformed into exponentially growing Y93 and Y94 cells. Positive recombinants were selected for growth on medium lacking leucine, and selected clones were confirmed by diagnostic PCR.

[00174] Strain Y188 was generated by introducing into strain Y176 an additional copy of the coding sequences of the ERG13, truncated HMG1, ERG10, and ERG12 genes of *Saccharomyces cerevisiae* under regulotary control of the galactose inducible promoter of the GAL1 or GAL10 gene of *Saccharomyces cerevisiae.* 2 µg of expression plasmids pAM491 and pAM495 were digested to completion using PmeI restriction endonuclease (New England Biolabs, Beverly, MA), and the purified DNA inserts were introduced into exponentially growing Y176 cells. Positive recombinants were selected for by growth on medium lacking uracil and histidine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

[00175] Strain Y189 was generated by introducing into strain Y188 an additional copy of the coding sequences of the ERG20, truncated HMG1, ERG8, and ERG19 genes of *Saccharomyces cerevisiae* under regulotary control of the galactose inducible promoter of the GAL1 or GAL10 gene of *Saccharomyces cerevisiae.* 2 ug of expression plasmids pAM489 and pAM497 plasmid DNA were digested to completion using PmeI restriction endonuclease, and the purified DNA inserts were introduced into exponentially growing Y177 cells. Positive recombinants were selected for by growth on medium lacking tryptophan and histidine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

[00176] Strain Y238 was generated by mating strains Y188 and Y189, and by introducing an additional copy of the coding sequences of the IDI1 and truncated HMG1 genes of *Saccharomyces cerevisiae* under regulotary control of the galactose inducible promoter of the GAL1 or GAL10 gene of *Saccharomyces cerevisiae*. Approximately 1 x

10^7 cells from strains Y188 and Y189 were mixed on a YPD medium plate for 6 hours at room temperature to allow for mating, and the mixed cell culture was plated on medium lacking histidine, uracil, and tryptophan to select for growth of diploid cells. 2 ug of expression plasmid pAM493 plasmid DNA were digested to completion using PmeI restriction endonuclease, and the purified DNA insert was introduced into the exponentially growing diploid cells. Positive recombinants were selected for by growth on medium lacking adenine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

[00177] To generate haploid strains containing the full complement of introduced coding sequences, strain Y238 was sporulated in 2% potassium acetate and 0.02% raffinose liquid medium. Approximlately 200 genetic tetrads (tetrads are four-spored meiotic products) were isolated using a Singer Instruments MSM300 series micromanipulator (Singer Instrument Co, LTD. Somerset, UK). Independent genetic isolates containing the appropriate complement of introduced genetic material were identified by their ability to grow in the absence of adenine, histidine, uracil, and tryptophan. Integration of all introduced DNA was confirmed by diagnostic PCR. The resultant strains were given the designation Y210 (MAT A) and Y211 (MAT alpha).

[00178] Strain Y221 was generated from strain Y211 by transforming exponentially growing Y211 cells with vector pAM178 (SEQ ID NO: 2). Positive transformants were selected for by growth on complete synthetic medium lacking leucine.

[00179] Strain Y290 was generated from strain Y221 by deleting the coding sequence of the GAL80 gene, and thus rendering the GAL promoters in the strain constitutively active. To this end, exponentially growing Y221 cells were transformed with integration construct Gal80US_hphA_Gal80DS (SEQ ID NO: 100). Host cell transformants were selected based on their resistance to hygromycin B conferred by the hphA gene present in the genomic integration construct, yielding strain Y290.

[00180] Strain Y318 was generated from strain Y290 by removing the pAM178 vector by serial propagation in leucine-rich media.

[00181] Strain Y409 was generated from strain Y318 by introducing a heterologous nucleotide sequence encoding a β-farnesene synthase. To this end, exponentially growing Y318 cells were transformed with expression plasmid pAM404. Host cell transformants were selected on complete synthetic defined medium (CSM) lacking leucine.

[00182] Strain Y419 was generated from strain Y409 by rendering the GAL promoters in the strain constitutively active and able to express higher levels of GAL4p in the presence of glucose (i.e., able to more efficiently drive expression off galactose-inducible promoters in the presence of glucose, as well as assure that there is enough Gal4p transcription factor to drive expression from all the galactose-inducible promoters in the cell). Specifically, the KanMX marker at the ERG9 locus in strain Y409 was replaced by a DNA fragment that comprised the ORF of the GAL4 gene of *Saccharomyces cerevisiae* under the control of an "operative constitutive" version of its native promoter (Griggs & Johnston (1991) *PNAS* 88(19):8597-8601) and the GAL4 terminator (P_{Gal4OC}-GAL4-T_{GAL4}), and the nourseothricin resistance selectable marker gene of *Streptomyces noursei* (NatR) flanked by the promoter and terminator of the *Tef1* gene of *Kluyveromyces lactis*. To this end, exponentially growing Y409 cells were transformed with integration construct ERG9US_PGAL4oc-GAL4-TGAL4_natA_PMet3old-ERG9CDS (SEQ ID NO: 101). Host cell transformants were selected based on their resistance to nourseothricin conferred by the introduction of natA marker in place of the kanMX marker upstream of the chimeric MET3/ERG9 promoter, yielding strain Y419.

[00183] Strain Y677 was generated from strain Y419 by introducing an additional copy of the coding region of the ERG12 gene of *Saccharomyces cerevisiae* under the control of P_{GAL1} at the GAL80 locus. To this end, exponentially growing Y677 cells were transformed with integration construct GAL80US_kanR_PGAL1-ERG12_GAL80DS (SEQ ID NO: 102). Host cell transformants were selected based on their resistance to kanamycin conferred by the presence in the genomic integration construct of the kanR marker, yielding strain Y677.

[00184] Strain Y1213 was generated from strain Y677 by removing the URA3 marker. To this end, strain Y677 was transformed with integration construct URA3US_PGAL10-ERG13_PGAL1-tHMGR_hisG_URA3DS (SEQ ID NO: 103) comprising the HisG gene of *Escherichia coli* (non-functional in yeast) flanked by yeast genomic sequences that are located 5' and 3' of the URA3 marker in strain Y677. Host cell transformants were selected on CSM containing 5-FOA, yielding strain Y1213.

Example 2

[00185] This example describes methods for making genetically modified yeast cells that comprise the C227Y mutation in the PHD-finger domain of YNG2p, and that biosynthesize increased yields of the acetyl-CoA derived compound β-farnesene.

[00186] The genetically modified yeast cells were generated by simultaneous integration via homologous recombination into a strain Y1213 cell's genome of two chromosomal integration constructs, namely chromosomal integration construct YNG2-C227Y-A and chromosomal integration construct YNG2-C227Y-B (FIG. 4).

[00187] Chromosomal integration construct YNG2-C227Y-A comprised in 5' to 3' direction the following two segments:

1. segment <u>YNG2-C227Y</u>: a nucleotide sequence identical to a nucleotide sequence located within and downstream of the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (nucleotide positions +187 to +1104 wherein the A of the ATG codon of the YNG2 gene is position +1) except that it comprised a 30 nucleotide long sequence consisting of 10 codons that code for amino acids 218 to 227 of YNG2p, of which the codon coding for amino acid 227 was changed from TGT to TAT to effect an amino acid change from cysteine to tyrosine (C227Y mutation), and of which the other nine codons were changed to synonymous codons (*i.e.*, different codons coding for the same amino acid as present in the native YNG2p) to a) create a region of heterology between the YNG2-C227Y segment and the concomitant segment in the native YNG2 locus in the *Saccharomyces cerevisiae* genome that increases the frequency of homologous recombination events between the chromosomal integration construct and the native locus involving a cross-over upstream of the codon encoding the C227Y mutation relative to homologous recombination events involving a cross-over downstream of this codon, and b) enable rapid identification of the desired homologous recombinants by colony PCR; and
2. segment <u>URA</u>: a nucleotide sequence comprising the promoter and the N-terminal approximate two thirds of the coding sequence of the URA3 selective marker.

[00188] Chromosomal integration construct YNG2-C227Y-B comprised in 5' to 3' direction the following two segments:

1. segment RA3: a nucleotide sequence comprising the C-terminal approximate two thirds of the coding sequence of the URA3 selective marker (note that this segment overlaps with segment URA described above); and
2. segment YNG2-DS1: a nucleotide sequence identical to a nucleotide sequence located downstream of the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (nucleotide positions +864 to +1355; note that this segment overlaps with segment YNG2-C227Y).

[00189] To generate chromosomal integration constructs YNG2-C227-A and YNG2-C227-B, the DNA fragments listed in Table 6 were amplified by PCR using the primers and templates shown in the table, and then assembled using the terminal primers shown in Table 7. For assembly, 100 fmole of each DNA fragment were combined in a 50 uL PCR reaction without primers, and a first round of PCR amplification was initiated (see legend of Table 7). Samples were placed on ice, 0.5 µM of each terminal primer were added to the reaction mixtures, and a second round of PCR amplification was performed (see legend of Table 7). The reaction mixtures were resolved by gel electrophoresis, and the chromosomal integration constructs were gel purified.

Table 6 - Amplified DNA fragments

DNA Fragment	Primers	Size (bp)	Template
C227Y-1	CR1110 (SEQ ID NO: 5) CR1111 (SEQ ID NO: 6)	493	Y002 genomic DNA
C227Y-2	CR1112 (SEQ ID NO: 7) CR1114 (SEQ ID NO: 8)	477	Y002 genomic DNA
URA	RYSE S002 (SEQ ID NO: 14) CR1207 (SEQ ID NO: 15)	1165	Plasmid comprising the promoter and coding sequence of URA3
RA3	CR1206 (SEQ ID NO: 16) RYSE S005 (SEQ ID NO: 18)	869	Plasmid comprising the promoter and coding sequence of URA3
DS1	CR1116 (SEQ ID NO: 73) CR1115 (SEQ ID NO: 74)	515	Y002 genomic DNA

PCR amplifications were done using the Phusion DNA polymerase (New England Biolabs, Ipswich, MA) as per manufacturer's suggested protocol. The PCR reactions were resolved by 1% agarose gel electrophoresis, and the DNA fragments were purified using a gel extraction kit (Omega Bio-Tek, Norcross, GA).

Table 7 - Terminal primers for assembly of DNA fragments into chromosomal integration constructs

Chromosomal integration construct	DNA Fragments To Be Assembled	Chromosomal integration construct Size (bp) (Sequence)	Terminal Primer 1	Terminal Primer 2
YNG2-C227Y-A	C227Y-1 C227Y-2 URA	2080 (SEQ ID NO: 3)	CR1110 (SEQ ID NO: 5)	CR1207 (SEQ ID NO: 15)
YNG2-C227Y-B	RA3 DS1	1361 (SEQ ID NO: 4)	CR1206 (SEQ ID NO: 16)	CR1116 (SEQ ID NO: 73)

Phusion polymerase reactions contained: 41 uL ddH2O, 20 uL 5x HF Buffer, 5 uL of each terminal primer (1 µM), 2 uL dNTP mix (200 µM), 1.8 uL Phusion DNA Polymerase (New England Biolabs, Ipswich, MA), and 30 µL DNA fragments.

The first round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 1 minute; 5 cycles of denature at 98°C for 20 seconds, anneal at 60°C for 20 seconds, and extend at 72°C for at least 40 seconds per kilobase pair of total assembly length; followed by a hold at 4°C for addition of the two terminal primers. The second round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 1 minutes; 35 cycles of denature at 98°C for 12 seconds, anneal at 60°C for 15 seconds, and extend at 72°C for at least 20 seconds per kilobase pair of total assembly length; 1 cycle of final extend at 72°C for 5 minutes; and a final hold at 4°C.

[00190] Strain Y3901 was generated by transforming strain Y1213 with 300 ng of chromosomal integration construct YNG2-C227Y-A and 200 ng of chromosomal integration construct YNG2-C227-B. Host cell transformants that had integrated the URA3 marker (FIG. 4) were selected on CSM lacking uracil. Isolated clones that had integrated the heterology block and codon encoding the C227Y mutation were identified by colony PCR.

[00191] Strain Y4181 was generated from strain Y3901 by looping out the URA3 marker by homologous recombination between the duplicate nucleotide sequences contained in segments YNG2-C227Y and YNG2-DS1, and comprised a mutant YNG2 locus that differed from the native YNG2 locus only in comprising the heterology block and the codon encoding the C227Y mutation. Cells in which the loop-out had occurred were selected by plating on CSM comprising 5-FOA. Isolated clones that comprised the mutant YNG2 locus were identified by colony PCR.

[00192] Strain Y3901 clones were inoculated into 8 replicate wells of a 1.1 mL/well 96-well plate containing Bird Seed Medium (BSM) with 2% sucrose. After growth for two days, cultures were diluted 1/25 into another 96-well plate containing BSM with 4% sucrose. After another two days of growth, production of β-farnesene was determined. One hundred microliters of each culture was transferred into a 2.2-mL plate, and 400 uL/well of methanol was added. The plate was sealed and shaken for 30-40 minutes. The seal was removed and 600 uL/well of n-heptane containing 0.001% trans-caryophyllene was added. The plate was sealed and shaken for 60 minutes, then centrifuged at 4000 rpm for 10 minutes. Three hundred fifty uL/well of n-heptane containing 0.001% trans-caryophyllene was added into a new 1.1-mL plate. Seventy ul/well of the top heptane layer from the 2.2-mL plate was transferred into each well of the 1.1-mL plate, bringing the total volume in each well to 420 μL. The heptane extracts were analyzed on an Agilent 7890 Gas Chromatography System (Agilent Technologies, Inc., Palo Alto, CA) with flame ionization detection (FID). A 1 μL aliquot of each extract was injected using a 50:1 split ratio, and compounds contained in the sample were separated using a DB1-MS-LTM column (10-m x 0.10-mm ID x 0.10-um film; Agilent Technologies, Inc., Palo Alto, CA), hydrogen carrier gas, and the following temperature program: 100°C hold for 0.15 minute, increasing temperature at 75°C/minute to a temperature of 195°C, then baking the column at 320°C for 0.52 minute. Using this protocol, trans-β-farnesene had a retention time of approximately 1.23 minutes. Farnesene titers were calculated by comparing generated peak areas against a quantitative calibration curve of purified biologically derived trans-β-farnesene (Amyris Biotechnologies, Emeryville, CA) in heptane.

[00193] As shown in FIG. 6, the introduction of the C227Y mutation into the PHD-finger domain of YNG2p resulted in increased production of the acetyl-CoA derived compound β-farnesene.

Example 3

[00194] This example describes methods for making genetically modified yeast cells comprising a deletion of the PHD-finger domain of YNG2p, and that biosynthesize increased yields of the acetyl-CoA derived compound β-farnesene.

[00195] The genetically modified yeast cells were generated by integration via homologous recombination into a strain Y1213 cell's genome of chromosomal integration construct YNG2-delPHD (FIG. 5).

[00196] Chromosomal integration construct YNG2-delPHD comprised in 5' to 3' direction the following three segments:

1. segment YNG2-trunc: a nucleotide sequence identical to a nucleotide sequence located within the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (from nucleotide positions +187 to +663) with the AAA codon coding for lysine at position 221 of the YNG2p amino acid sequence changed to a TAA stop codon;
2. segment URA3: a nucleotide sequence comprising the promoter and the coding sequence of the URA3 selective marker; and
3. segment YNG2-DS2: a nucleotide sequence identical to a nucleotide sequence located downstream of the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (from nucleotide positions +865 to +1259).

[00197] To generate chromosomal integration construct YNG2-delPHD, the DNA fragments listed in Table 8 were PCR amplified using the primers and templates shown in the table, and then assembled using the terminal primers shown in Table 9. For assembly, 200 fmole of each DNA fragment were combined in a 100 uL PCR reaction without primers, and a first round of PCR amplification was initiated (see legend of Table 9). Samples were placed on ice, 0.5 μM of each terminal primer were added to the reaction mixture, and a second round of PCR amplification was performed (see legend of Table 9). The reaction mixture was resolved by gel electrophoresis, and the chromosomal integration construct was gel purified.

Table 8 - Amplified DNA fragments

DNA Fragment	Primers	Size (bp)	Template
trunc	CR476 (SEQ ID NO: 75) CR477 (SEQ ID NO: 76)	501	Y002 genomic DNA
URA3	RYSE S002 (SEQ ID NO: 14) RYSE S005 (SEQ ID NO: 18)	1602	Plasmid comprising the promoter and coding sequence of URA3
DS2	CR478 (SEQ ID NO: 77) CR479 (SEQ ID NO: 78)	418	Y002 genomic DNA

PCR amplifications were done using the Phusion DNA polymerase (New England Biolabs, Ipswich, MA) as per manufacturer's suggested protocol. The PCR reactions were resolved by 1% agarose gel electrophoresis, and the DNA fragments were purified using a gel extraction kit (Omega Bio-Tek, Norcross, GA).

Table 9 - Terminal primers for assembly of DNA fragments into chromosomal integration construct

Chromosomal integration construct	DNA Fragments To Be Assembled	Chromosomal integration construct Size (bp) (Sequence)	Terminal Primer 1	Terminal Primer 2
YNG2-delPHD	trunc URA3 DS2	2474 (SEQ ID NO: 79)	CR476 (SEQ ID NO: 75)	CR479 (SEQ ID NO: 78)

Phusion polymerase reactions contained: 41 uL ddH2O, 20 uL 5x HF Buffer, 5 uL of each terminal primer (1 uM), 2 uL dNTP mix (200 uM), 1.8 uL Phusion DNA Polymerase (New England Biolabs, Ipswich, MA), and 30 uL DNA fragments.

The first round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 1 minute; 5 cycles of denature at 98°C for 20 seconds, anneal at 60°C for 20 seconds, and extend at 72°C for at least 40 seconds per kilobase pair of total assembly length; followed by a hold at 4°C for addition of the two terminal primers. The second round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 1 minutes; 35 cycles of denature at 98°C for 12 seconds, anneal at 60°C for 15 seconds, and extend at 72°C for at least 20 seconds per kilobase pair of total assembly length; 1 cycle of final extend at 72°C for 5 minutes; and a final hold at 4°C.

[00198] Strain Y2744 was generated by transforming strain Y1213 with 500 ng of chromosomal integration construct YNG2-delPHD. Host cell transformants that had incorporated the URA3 marker were selected on CSM lacking uracil. Isolated clones that comprised the truncated YNG2 coding sequence were identified by colony PCR.

[00199] Strain Y2744 clones were cultured and production of β-farnesene was determined as described in Example 2. Experiments were performed using 8 independent clones, and results were averaged.

[00200] As shown in FIG. 6, deletion of the PHD-finger domain of YNG2p resulted in increased production of the acetyl-CoA derived compound β-farnesene.

Example 4

[00201] This example describes methods for making genetically modified yeast cells that comprise modified YNG2p, and that biosynthesize increased yields of the acetyl-CoA derived compound β-farnesene.

[00202] The genetically modified yeast cells were generated by simultaneous integration via homologous recombination into a strain Y1213 cell's genome of two chromosomal integration constructs, namely chromosomal integration construct YNG2-mut-A and chromosomal integration construct YNG2-mut-B.

[00203] Chromosomal integration construct YNG2-mut-A comprised in 5' to 3' direction the following two segments:

1. segment YNG2-US: a nucleotide sequence identical to a nucleotide sequence located upstream of and within the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (nucleotide positions -474 to +1051) except that it comprised a 24 nucleotide long sequence consisting of 8 codons that code for amino acids 2 to 9 of YNG2p and that were changed to synonymous codons (*i.e.*, different codons coding for the same amino acid as present in the native YNG2p) to a) create a region of heterology between the YNG2-US segment and the concomitant segment in the native YNG2 locus in the *Saccharomyces cerevisiae* genome that increases the frequency of homologous recombination events between the chromosomal integration construct and the native locus involving a cross-over upstream of the start codon relative to homologous recombination events involving a cross-over downstream of the start codon, and b) enable rapid identification of the desired homologous recombinants by colony PCR; and
2. segment URA (see Example 2).

[00204] Chromosomal integration construct YNG2-mut-B comprised in 5' to 3' direction the following two segments:

1. segment RA3 (see Example 2); and
2. segment YNG2-DS: a nucleotide sequence identical to a nucleotide sequence located downstream of the YNG2 coding sequence in the *Saccharomyces cerevisiae* genome (nucleotide positions +792 to +1364; note that this segment comprised a nucleotide sequence that was also contained in segment YNG2-US).

[00205] To generate chromosomal integration constructs YNG2-mut-A and YNG2-mut-B, the DNA fragments listed in Table 10 were PCR amplified using the primers and templates shown in the table.

Table 10 - Amplified DNA fragments

DNA Fragment	Primers	Size (bp)	Template
US	TRG-386-59-YNG2-us500Fwd (SEQ ID NO: 98) TRG-386-59-YNG2-PR (SEQ ID NO: 93)	503	Y002 genomic DNA
DS3	TRG-386-59-YNG2-PFWt (SEQ ID NO: 94) TRG-386-59-YNG2-ds200Rev (SEQ ID NO: 95)	1072	Y002 genomic DNA
URA	RYSE S002 (SEQ ID NO: 14) CR1207 (SEQ ID NO: 15)	1165	RABits 555
RA3	CR1206 (SEQ ID NO: 16) RYSE S005 (SEQ ID NO: 18)	869	RABits 555
DS4	TRG-386-59-YNG2-stopFwd (SEQ ID NO: 96) TRG-386-59-YNG2-ds500Rev (SEQ ID NO: 97)	597	Y002 genomic DNA
PCR amplifications were done using the Phusion DNA polymerase (New England Biolabs, Ipswich, MA) as per manufacturer's suggested protocol. The PCR reactions were resolved by gel electrophoresis, the DNA fragments were gel purified, and the purified DNA fragments were treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, MA) as per manufacturer's suggested protocol. The PNK was heat inactivated at 65°C for 20 minutes, and the samples were stored at -20°C.			

[00206] DNA fragment DS3 was subsequently mutagenized using the GeneMorph II Random Mutagenesis Kit (Stratagene, La Jolla, CA) using 100 ng, 10 ng, 1 ng, and 0.1 ng of the DNA fragment as template, and using primers TRG-386-59-YNG2-PFWt (SEQ ID NO: 94) and TRG-386-59-YNG2-ds200Rev (SEQ ID NO: 95), yielding DNA fragments DS3-mut. For each amount of starting material, 4 reactions were set up to obtain a highly diverse mutagenesis profile (total of 16 reactions).

[00207] The DNA fragments were assembled using the terminal primers shown in Table 11 to yield the two chromosomal integration constructs. To this end, 666 fmole of each DNA fragment (except for DNA fragment DS3-mut, of which 1332 fmole were used) were placed together in one tube. The samples were split into three 30 uL reactions; water, buffer, dNTPs, and DNA polymerase were added to each reaction mixture, and a first round of PCR amplification was initiated (see legend of Table 11). Samples were placed on ice, 5 uL of each primer at 10 uM concentration were added to the reaction mixtures, and a second round of PCR amplification was performed (see legend of Table 11). The three PCR reaction mixtures were combined in one tube, the reaction mixtures were resolved by gel electrophoresis, and the chromosomal integration constructs were gel purified.

Table 11 - Terminal primers for assembly of DNA fragments into chromosomal integration constructs

Chromosomal integration construct	DNA Fragments To Be Assembled	Chromosomal integration construct Size (bp) (Sequence)	Terminal Primer 1	Terminal Primer 2
YNG2-mut-A	US DS3-mut * URA	2690	TRG-386-59-YNG2-us500Fwd (SEQ ID NO: 98)	CR1207 (SEQ ID NO: 15)
YNG2-mut-B	RA3 DS4	1442	CR1206 (SEQ ID NO: 16)	TRG-386-59-YNG2-ds500Rev (SEQ ID NO: 97)

PCR reactions contained: 41 uL ddH2O, 20 uL 5x HF Buffer, 5 uL of each terminal primer (1 uM), 2 uL dNTP mix (200 uM), 1.8 uL Phusion DNA Polymerase, and 30 uL DNA fragments (200 fmole of each DNA fragment except DNA fragment YNG2-het-DS200-mut, which was added at 400 fmole).

The first round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 2 minutes; 8 cycles of denature at 98°C for 30 seconds, anneal at 55°C for 30 seconds, and extend at 72°C for 2.5 minutes; followed by a hold at 4°C for addition of the two terminal primers. The second round of PCR amplification was performed as follows: 1 cycle of denature at 98°C for 2 minutes; 30 rounds of denature at 98°C for 12 seconds, anneal at 55°C for 30 seconds, and extend at 72°C for 2.5 minutes; 1 cycle of final extend at 72°C for 7 minutes; and a final hold at 4°C.

* The mutagenized DSc-mut DNA fragments were combined for each starting material concentration so that the entire diversity of mutations generated could be sampled in subsequent transformations.

[00208] Strain Y1213 was transformed with 300 ng of chromosomal integration construct YNG2-mut-A and 100 ng of chromosomal integration construct YNG2-mut-B. Host cell transformants that had integrated the URA3 marker were selected on CSM lacking uracil.

[00209] Individual clones were inoculated into wells of 1.1 mL/well 96-well plates containing 360 μL BSM with 2% sucrose, and the cultures were incubated for 48 hours at 30°C on a rotary shaker at 999 rpm. An aliquot of 14.4 μL was taken from each well and transferred to new 1.1 mL/well 96-well plates containing BSM with 4% sucrose, and cultured for another 48 hours at 30°C on a rotary shaker at 999 rpm. Production of β-farnesene by each clone was determined transferring 98 uL from each well to the wells of a black 96-well plate, and adding 2 μL of a solution of 100 ug Nile Red (Invitrogen, Carlsbad, CA) per mL dimethylsulfoxide to each well. The plates were read for top fluorescence with excitation at 500 nm and emission at 550 nm in a Spectramax M5 fluorescence plate reader (Molecular Devices, Sunnyvale, CA). Clones that produced high levels of β-farnesene were analyzed by colony PCR followed by sequencing to identify the nature of the YNG2 mutation.

[00210] As shown in Table 12, mutations that potentially contributed to an increase in β-farnesene production were located both inside and outside of the PHD-finger domain of YNG2p.

Table 12 – β-farnesene production by strains comprising modified YNG2p

Starting Material	Strain	Nile Red Signal	amino acid substitutions in YNG2p	amino acids deleted in YNG2p	Control
100ng	1	5186.6	none	264 to 282	Y677 avg = 3922.70 STDEV = 2.4%
100ng	2	5235.06	R54G + P60S + N91S	254 to 282	
10ng	3	4857.58	K40R	111 to 282	Y677 avg = 4107.80 STDEV = 1.1%
10ng	4	5168.6	K61R + R134S + C252Y	none	
10ng	5	5203.94	A129S	none	
10ng	6	4751.03	C243R	none	
1ng	7	5591.83	none	264 to 282	Y677 avg = 3912.45 STDEV = 3.6%
1ng	8	5254.38	none	167 to 282	
1ng	9	5166.69	M235I	none	
0.1ng	10	5375.69	S133Y + R147K	181 to 282	Y677 avg = 4100.44 STDEV = 3.3%
0.1ng	11	5316.52	K103R	106 to 282	
0.1ng	12	5284.64	K182R + Y264H	none	

Example 5

[00211] This example describes methods for creating and identifying mutations in YNG2p that effect increased expression of galactose regulated genes.

[00212] Galactose regulated gene expression was compared between strain Y3901 (see above) and strain Y3901. Strain Y3901 comprised the same heterology block in the YNG2 coding sequence as strain Y3901 cells comprised, but did not comprise the codon encoding the C227Y mutation. Strain Y3900 was generated in the same manner as strain Y3901 except that in place of primer CR1112 used to generated DNA fragment C227Y-2 (see Table 6), primer CR1113 (SEQ ID NO: 99) was used, which comprised the same heterology block as primer CR1112 comprised but did not comprise the codon encoding the C227Y mutation.

[00213] Cell banks of strains Y3900 and Y3901 were prepared by growing the cells in Bird Seed medium with 2% sugar at 30°C and 250 rpm until they reached an OD600 of about 0.5 to 1. At that time, the flasks were placed on ice. Three parts culture and two parts ice cold sterile 50% glycerol were combined, and 1 mL aliquots of this mixture were frozen at -80°C in Cryotubes.

[00214] A seed vial of each of strains Y3000 and Y3001 was inoculated into 40 mL BSM with 2% sucrose in a 250 mL unbaffled shake flask , and the cultures were incubated at 30°C and 250 rpm until at least 3 cell doublings were achieved and the OD600 had not exceeded 5. Each culture was next inoculated into 40 mL BSM with 2% sucrose, and cultured for another 24 hours at 30°C and 250 rpm to achieve exponential cell growth. Finally, each culture was diluted to an OD600 of 0.5 into 40 mL BSM with 4% sucrose (for sugar phase sampling) and to an OD600 of 0.1 into 40 mL BSM with 4% sucrose (for ethanol phase sampling). Under sugar phase sampling conditions the cells primarily consumed sugar whereas under ethanol phase sampling conditions the cells primarily consumed the ethanol they had produced. The cultures were grown at 30°C and 250 rpm for 8 hours for sugar phase sampling and 49 hours for ethanol phase sampling before they were analyzed for galactose regulated gene expression.

[00215] For gene expression analysis, RNA was first extracted by resuspending 3 mL OD pellets of each culture in 125 uL Lysing/Binding Buffer from an Ambion RNaqueous Kit (Applied Biosystems/Ambion, Austin, TX). To each tube, ~65 uL Lysing Matrix C beads (MP Biomedicals, Solon, OH) were added, and the mixtures were bead beaten for 10 minutes using a FastPrep 24 (MP Biomedicals, Solon, OH). RNA was quantified using the Qubit™ Quantitation Platform from Invitrogen (Carlsbad, CA) according to manufacturer's suggested protocol. Gene expression was determined using the nCounter™ Gene Expression Assay and nCounter™ Prep Station (Nanostring Technologies, Seattle, WA) using 15-25 ng RNA and hybridizing at 65°C for at least 12 hours. Samples were analyzed using the nCounter™ Digital Analyzer (Nanostring Technologies, Seattle, WA).

[00216] For protein expression analysis, a 10 OD-mL cell pellet was collected for each strain, washed once with 1 mL HPLC grade water, and stored at -80°C. Protein was obtained for analysis by resuspending the cell pellet in 250 uL lysis buffer (6M urea in 500 mM triethylammonium bicarbonate) and transferring the mixture to a 0.5 mL screw cap tube filled with 250 uL Lysing Matrix C beads (MP Biomedicals, Solon, OH). Samples were vortexed on a FastPrep 24 (MP Biomedicals, Solon, OH) at setting 6 M/S for 2x60 seconds with a 1 minute rest on ice between beatings. Samples were centrifuged for 30 minutes at 20,000 RCF and 4°C to obtain clarified cell lysate (CCL). Protein concentration of the CCL was determined using the Quant-iT protein assay kit (Invitrogen, Carlsbad, CA) according to the manufacturer's suggested protocol. A 50 μL portion of each CCL was diluted with an appropriate amount of lysis buffer to obtain samples with a final protein concentration of 1.5 mg/mL. A 15 uL portion of each 1.5 mg/mL CCL was transferred into each of three 1.5 mL Eppendorf tubes on ice (for triplicate analysis of each sample). 10 uL of AQUA peptide mix (containing 0.25 uM of the desired $^{13}C/^{15}N$ labeled peptide internal standards specified below) and 2 uL of 60 mM tris(2-carboxyethyl) phosphine (TCEP) in water was added to each tube. Tubes were incubated at 37°C for 30 minutes to reduce protein disulfide bonds. 5 uL of 111 mM iodoacetamide in water was added, and the tubes incubated at room temperature in the dark for 30 minutes. A 20 ug portion of trypsin (Promega, Madison, WI) was dissolved in 356 uL of HPLC water, and 20 uL of the resulting solution was added to each sample. Samples were incubated at 37°C for 14 hours. A 443 uL portion of 98/2/0.1 water/acetonitrile/formic acid was added to quench each digest. Samples were analyzed by LC-MRM on an AB Sciex 4000 QTrap mass spectrometer running Analyst 1.5 (AB Sciex, Foster City, CA). The transitions monitored to quantify each protein are listed in Table 13. A standard 2.1 mm C_{18} reverse phase column was used for the analysis. Mobile phase A was water/0.1% formic acid, and mobile phase B was acetonitrile/0.1% formic acid. A 33 minute gradient was used to elute the peptides at 2.4%B/minute and 250 uL/min. AB Sciex Multiquant software was used to integrate peaks for calculation of nmol/gDCW values.

Table 13

Transition Label	Parent Ion m/z	Product Ion m/z	Collision Voltage (V)	Internal Standard
.SYNIDTNK.478(2+)-704(y6)	477.73	704.36	26	
.SYNIDTNK-C13.IS	481.73	712.36	26	X
.ALELASAEQR.544(2+)-661(y6)	544.29	661.33	29	
.ALELASAEQR-C13.IS	549.29	671.33	29	X
.DYNEDLVDGR.598(2+)-917(y8)	598.26	917.43	31	
.DYNEDLVDGR-C13.IS	603.26	927.43	31	X
.TLTSAATAPEFER.697(2+)-920(y8)	697.35	920.45	36	
.TLTSAATAPEFER-C13.IS	702.35	930.45	36	X
.GDIVTEDTFK.563(2+)-740(y6)	562.78	740.35	30	
.GDIVTEDTFK-C13.IS	566.78	748.35	30	X
.WVSPNDLK.480(2+)-673(y6)	479.75	673.35	26	
.WVSPNDLK-C13.IS	483.75	681.35	26	X
.SVVAEATIPGDVVR.707(2+)-928(y9)	706.89	927.53	36	
.SVVAEATIPGDVVR-C13.IS	711.89	937.53	36	X
.TPIGSFQGSLSSK.655(2+)-997(y10)	654.84	997.49	34	
.TPIGSFQGSLSSK-C13.IS	658.84	1005.49	34	X
.YEAFVVGLSAR.606(2+)-920(y9)	606.32	919.54	32	
.YEAFVVGLSAR-C13.IS	611.32	929.54	32	X
.SLVFQLFENK.613(2+)-925(y7)	612.83	925.48	32	
.SLVFQLFENK-C13.IS	616.83	933.48	32	X

[00217] As shown in FIGs. 7 and 8, strain Y3901 grown under ethanol phase sampling conditions had increased gene expression of both endogenous galactose regulated genes (GAL1, GAL10, GAL7, GAL2, GCY1) and heterologous galactose regulated coding sequences encoding enzymes of the MEV pathway (ERG10, ERG12, ERG13, tHMG1, ERG8, MVD1=ERG19), IPP isomerase (IDI1), and FPP synthase (ERG20).

[00218] All publications, patents and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be 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.

WHAT IS CLAIMED:

1. A recombinant microorganism comprising heterologous nucleic acid encoding a modified YNG2.
2. The recombinant microorganism of Claim 1, wherein the microorganism has at least 15% less nucleosome HAT activity in at least one of a nuA4 and piccolo NuA4 complex than a wild type microorganism.
3. The recombinant microorganism of Claim 1, wherein the microorganism has at least 15% greater translation of a galactose regulated gene product than a wild type microorganism.
4. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises a deletion of a PHD-finger domain.
5. The recombinant microorganism of Claim 4, wherein the deletion comprises residues 220 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).
6. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises at least one amino acid substitution at a position selected from the group consisting of positions corresponding to positions 227, 235, 243, 252, and 264 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).
7. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises at least one amino acid substitution at positions selected from the group consisting of positions corresponding to positions 221 to 282, 254 to 282, and 264 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).
8. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises at least one amino acid substitution at a position selected from the group consisting of positions corresponding to positions 40, 54, 60, 61, 91, 103, 129, 133, 134, 147, and 182 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).
9. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 90, 92, 106, 107, 108, 109, and 110.
10. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a cysteine to tyrosine substitution (C227Y mutation) at a position corresponding to position 227 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

11. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a methionine to isoleucine substitution (M235I mutation) at a position corresponding to position 235 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

12. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a cysteine to arginine substitution (C243R mutation) at a position corresponding to position 243 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

13. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a cysteine to tyrosine substitution (C252Y mutation) at a position corresponding to position 252 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

14. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a tyrosine to histidine substitution (Y264H mutation) at a position corresponding to position 264 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

15. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a lysine to arginine substitution (K40R mutation) at a position corresponding to position 40 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

16. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising an arginine to glycine substitution (R54G mutation) at a position corresponding to position 54 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

17. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a proline to serine substitution (P60S mutation) at a position corresponding to position 60 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

18. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a lysine to arginine substitution (K61R mutation) at a position corresponding to position 61 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

19. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising an asparagine to serine substitution (N91S mutation)

at a position corresponding to position 91 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

20. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a lysine to arginine substitution (K103R mutation) at a position corresponding to position 103 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

21. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising an alanine to serine substitution (A129S mutation) at a position corresponding to position 129 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

22. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a serine to tyrosine substitution (S133Y mutation) at a corresponding to position 133 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

23. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising an arginine to serine substitution (R134S mutation) at a position corresponding to position 134 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

24. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising an arginine to lysine substitution (R147K mutation) at a position corresponding to position 147 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

25. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a lysine to arginine substitution (K182R mutation) at a position corresponding to position 182 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

26. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 264 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

27. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 254 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

28. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 221 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

29. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 181 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

30. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 167 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

31. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 111 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

32. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 106 to 282 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

33. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 181 to 219 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

34. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 167 to 219 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

35. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 111 to 219 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

36. The recombinant microorganism of Claim 1, wherein the modified YNG2 comprises an amino acid sequence comprising a deletion of residues corresponding to residues 106 to 219 of the wild-type YNG2p amino acid sequence (SEQ ID NO: 105).

37. The recombinant microorganism of Claim 1, further comprising an isoprenoid pathway enzyme to form an isoprenoid product, wherein the isoprenoid pathway enzyme is selected from the group consisting of:

a) an enzyme that condenses two molecules of acetyl-coenzyme A to form acetoacetyl-CoA;

b) an enzyme that condenses acetoacetyl-CoA with another molecule of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA);

c) an enzyme that converts HMG-CoA into mevalonate;

d) an enzyme that converts mevalonate into mevalonate 5-phosphate;

e) an enzyme that converts mevalonate 5-phosphate into mevalonate 5-pyrophosphate;

f) an enzyme that converts mevalonate 5-pyrophosphate into IPP;

g) an enzyme that converts IPP into DMAPP;

h) a polyprenyl synthase that can condense IPP and/or DMAPP molecules to form polyprenyl compounds containing more than five carbons;

i) an enzyme that condenses IPP with DMAPP to form GPP;

j) an enzyme that condenses two molecules of IPP with one molecule of DMAPP;

k) an enzyme that condenses IPP with GPP to form FPP;

l) an enzyme that condenses IPP and DMAPP to form GGPP; and,

m) an enzyme that condenses IPP and FPP to form GGPP.

38. The recombinant microorganism of Claim 1, further comprising a heterologous nucleic acid encoding an enzyme that modifies a polyprenyl to form a product selected from the group consisting of a hemiterpene, a monoterpene, a sesquiterpene, a diterpene, a triterpene, a tetraterpene, a polyterpene, a steroid compound, a carotenoid, and a modified acetyl-CoA derived compound.

39. The recombinant microorganism of Claim 51, wherein the heterologous nucleic acid encoding an enzyme that modifies a polyprenyl is selected from the group consisting of a geraniol synthase, a linalool synthase, a limonene synthase, a myrcene synthase, an ocimene synthase, an α-pinene synthase, β-pinene synthase, a sabinene synthase, a γ-terpinene synthase, a terpinolene synthase, an amorphadiene synthase, an α-farnesene synthase, a β-farnesene synthase, a farnesol synthase, a nerolidol synthase, a patchouliol synthase, a nootkatone synthase, an abietadiene synthase.

40. The recombinant microorganism of Claim 1, further comprising heterologous nucleic acid encoding at least one polyketide synthesis enzyme, wherein the polyketide synthesis enzyme is selected from the group consisting of:

a) an enzyme that condenses at least one of acetyl-CoA and malonyl-CoA with an acyl carrier protein;

b) an enzyme that condenses a first reactant selected from the group consisting of acetyl-CoA and malonyl-CoA with a second reactant selected from the group consisting of malonyl-CoA or methylmalonyl-CoA to form a polyketide product;

c) an enzyme that reduces a β-keto chemical group on a polyketide compound to a β-hydroxy group;

d) an enzyme that dehydrates an alkane chemical group in a polyketide compound to produce an α-β-unsaturated alkene;

e) an enzyme that reduces an α-β-double-bond in a polyketide compound to a saturated alkane; and,

f) an enzyme that hydrolyzes a polyketide compound from an acyl carrier protein.

41. The recombinant microorganism of Claim 40, wherein the polyketide product is a lipid having at least one of antibiotic, antifungal, and antitumor activity.

42. The recombinant microorganism of Claim 40, wherein the polyketide product is selected from the group consisting of a macrolids, an antibiotic, an antifungal, a cytostatic compound, an anticholesterolemic compound, an antiparasitic compound, a coccidiostatic compound, an animal growth promoter and an insecticide.

43. The recombinant microorganism of Claim 40, wherein the polyketide product is a polyene.

44. The recombinant microorganism of Claim 40, wherein the polyketide product is a cyclic lactone.

45. The recombinant microorganism of Claim 44, wherein the cyclic lactone is one of a 14, 15, or 16-membered ring.

46. The recombinant microorganism of Claim 1, further comprising heterologous nucleic acid encoding at least one fatty acid synthesis enzyme, wherein the fatty acid synthesis enzyme is selected from the group consisting of:

a) an enzyme that covalently links at least one of acetyl-CoA and malonyl-CoA to an acyl carrier protein (ACP);

b) an enzyme that condenses acetyl-ACP and malonyl-ACP to form acetoacetyl-ACP;

c) reduce the double bond in acetoacetyl-ACP with NADPH to form a hydroxyl group in D-3-hydroxybutyryl hydroxylase-ACP;

d) an enzyme that dehydrates D-3-Hydroxybutyryl hydroxylase-ACP to create a double bond between the beta- and gamma-carbons forming crotonyl-ACP;

e) an enzyme that reduces crotonyl ACP with NADPH to form butyryl-ACP; and,

f) an enzyme that hydrolyzes a C16 acyl compound from an acyl carrier protein to form palmitate.

47. The recombinant microorganism of Claim 46, wherein the microorganism accumulates a fatty acid product.

48. The recombinant microorganism of Claim 47, wherein the fatty acid product is selected from the group consisting of palmitate, palmitoyl CoA, palmitoleic acid, sapienic acid, oleic acid, linoleic acid, α-linolenic acid, arachidonic acid, eicosapentaenoic acid, erucic acid, and docosahexaenoic acid.

49. The recombinant microorganism of any one of Claims 1-48 wherein the heterologous nucleic acid molecules are comprised on one heterologous nucleic acid molecule.

50. The recombinant microorganism of any one of Claims 1-48 wherein the heterologous nucleic acid molecules are comprised on a plurality of heterologous nucleic acid molecules.

51. The recombinant microorganism of any one of Claims 1-48 wherein the heterologous nucleic acid molecules are under inducible control.

52. The recombinant microorganism of any one of Claims 1-48 wherein the heterologous nucleic acid molecules are integrated into a chromosome.

53. The recombinant microorganism of any one of Claims 1-48 wherein the heterologous nucleic acid molecules are on a plasmid.

54. The recombinant microorganism of any one of Claims 1-48 wherein each heterologous nucleic acid encodes a eukaryotic enzyme.

55. The recombinant microorganism of any one of Claims 1-48 wherein each heterologous nucleic acid encodes a yeast enzyme.

56. The recombinant microorganism of any one of Claims 1-48 wherein each heterologous nucleic acid encodes a eukaryotic enzyme.

57. The recombinant microorganism of any one of Claims 1-48, wherein the microorganism is a prokaryote.

58. The recombinant microorganism of Claim 57, wherein the microorganism is *E. coli*.

59. The recombinant microorganism of any one of Claims 1-48, wherein the microorganism is a eukaryote.

60. The recombinant microorganism of Claim 59, wherein the microorganism is a fungi.

61. The recombinant microorganism of Claim 59, wherein the microorganism is a yeast.

62. The recombinant microorganism of Claim 61, wherein the microorganism is a yeast belonging to the genera selected from the group consisting of *Aciculoconidium, Ambrosiozyma, Arthroascus, Arxiozyma, Ashbya, Babjevia, Bensingtonia, Botryoascus, Botryozyma, Brettanomyces, Bullera, Bulleromyces, Candida, Citeromyces, Clavispora, Cryptococcus, Cystofilobasidium, Debaryomyces, Dekkara, Dipodascopsis, Dipodascus, Eeniella, Endomycopsella, Eremascus, Eremothecium, Erythrobasidium, Fellomyces, Filobasidium, Galactomyces, Geotrichum, Guilliermondella, Hanseniaspora, Hansenula, Hasegawaea, Holtermannia, Hormoascus, Hyphopichia, Issatchenkia, Kloeckera, Kloeckeraspora, Kluyveromyces, Kondoa, Kuraishia, Kurtzmanomyces, Leucosporidium, Lipomyces, Lodderomyces, Malassezia, Metschnikowia, Mrakia, Myxozyma, Nadsonia, Nakazawaea, Nematospora, Ogataea, Oosporidium, Pachysolen, Phachytichospora, Phaffia, Pichia, Rhodosporidium, Rhodotorula, Saccharomyces, Saccharomycodes, Saccharomycopsis, Saitoella, Sakaguchia, Saturnospora, Schizoblastosporion, Schizosaccharomyces, Schwanniomyces, Sporidiobolus, Sporobolomyces, Sporopachydermia, Stephanoascus, Sterigmatomyces, Sterigmatosporidium, Symbiotaphrina, Sympodiomyces, Sympodiomycopsis, Torulaspora, Trichosporiella, Trichosporon, Trigonopsis, Tsuchiyaea, Udeniomyces, Waltomyces, Wickerhamia, Wickerhamiella, Williopsis, Yamadazyma, Yarrowia, Zygoascus, Zygosaccharomyces, Zygowilliopsis*, and *Zygozyma*.

63. The recombinant microorganism of Claim 61, wherein the microorganism is selected from the group consisting of *Saccharomyces cerevisiae, Pichia pastoris, Schizosaccharomyces pombe, Dekkera bruxellensis, Kluyveromyces lactis* (previously called *Saccharomyces lactis*), *Kluveromyces marxianus, Arxula adeninivorans*, and *Hansenula polymorpha (*now known as *Pichia angusta).*

64. The recombinant microorganism of Claim 61, wherein the microorganism is selected from the group consisting of *Candida lipolytica, Candida guilliermondii*, *Candida krusei, Candida pseudotropicalis*, and *Candida utilis.*

65. The recombinant microorganism of Claim 61, wherein the microorganism is *Saccharomyces cerevisiae.*

66. The recombinant microorganism of Claim 65, wherein the microorganism is a strain of *Saccharomyces cerevisiae* selected from the group consisting of Baker's yeast, CBS 7959, CBS 7960, CBS 7961, CBS 7962, CBS 7963, CBS 7964, IZ-1904, TA, BG-1, CR-1, SA-1, M-26, Y-904, PE-2, PE-5, VR-1, BR-1, BR-2, ME-2, VR-2, MA-3, MA-4, CAT-1, CB-1, NR-1, BT-1, and AL-1.

67. The recombinant microorganism of any one of Claims 1-48, wherein the microorganism is a microalga.

68. The recombinant microorganism of Claim 67, wherein the microorganism is *Chlorella* or *Prototheca*.

69. A recombinant nucleic acid molecule comprising a polynucleotide encoding a modified YNG2.

70. A vector comprising the nucleic acid molecule according to claim 69.

71. The vector according to claim 70 which is an expression vector wherein the nucleic acid molecule is operably linked to one or more control sequences allowing the expression of the nucleic acid molecule in host cells.

72. The vector according to claim 71, wherein the control sequences are GAL promoters.

73. The vector according to claim 71, wherein the nucleic acid molecule is selected from the group consisting of a nucleic acid molecule set forth in a sequence selected from the group consisting of SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102 and SEQ ID NO:103.

74. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 95% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

75. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 90% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

76. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 85% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

77. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 75% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

78. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 65% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

79. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is at least 55% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

80. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 that is less than 90% identical to the wild-type YNG2 amino acid sequence (SEQ ID NO: 105).

81. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 enzyme that has at least 15% decreased enzymatic activity relative to the wild-type YNG2 enzyme.

82. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 enzyme that has at least 25% decreased enzymatic activity relative to the wild-type YNG2 enzyme.

83. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 enzyme that has at least 50% decreased enzymatic activity relative to the wild-type YNG2 enzyme.

84. A recombinant microorganism comprising heterologous nucleic acid encoding a YNG2 enzyme that has less than 50% decreased enzymatic activity relative to the wild-type YNG2 enzyme.

85. A method for enhancing production of an acetyl-CoA derived compound in a microorganism, the method comprising culturing the recombinant microorganism of any one of claims 1 to 68 or 74-85 in a suitable medium and under conditions that promote production of the acetyl-CoA derived compound, wherein the acetyl-CoA derived compound is produced at a level that is higher than the level of the acetyl-CoA derived compound in a microorganism not comprising the heterologous nucleic acid encoding a modified YNG2.

Figure 1

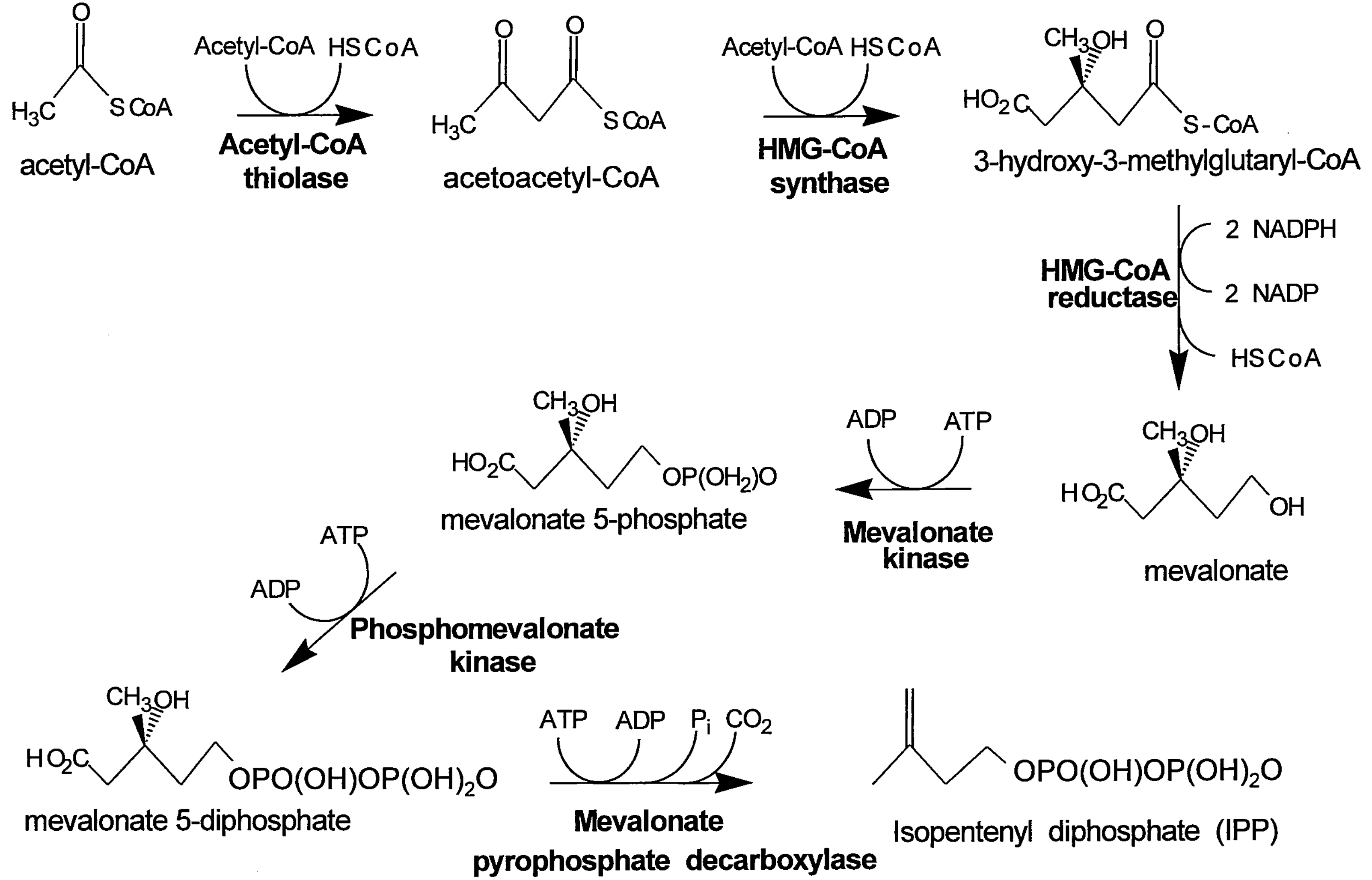

acetyl-CoA
Acetyl-CoA HSCoA
Acetyl-CoA thiolase
acetoacetyl-CoA
Acetyl-CoA HSCoA
HMG-CoA synthase
3-hydroxy-3-methylglutaryl-CoA
2 NADPH
2 NADP
HSCoA
HMG-CoA reductase
mevalonate
ADP
ATP
Mevalonate kinase
mevalonate 5-phosphate
ATP
ADP
Phosphomevalonate kinase
mevalonate 5-diphosphate
ATP
ADP
P_i
CO_2
Mevalonate pyrophosphate decarboxylase
Isopentenyl diphosphate (IPP)

Figure 2

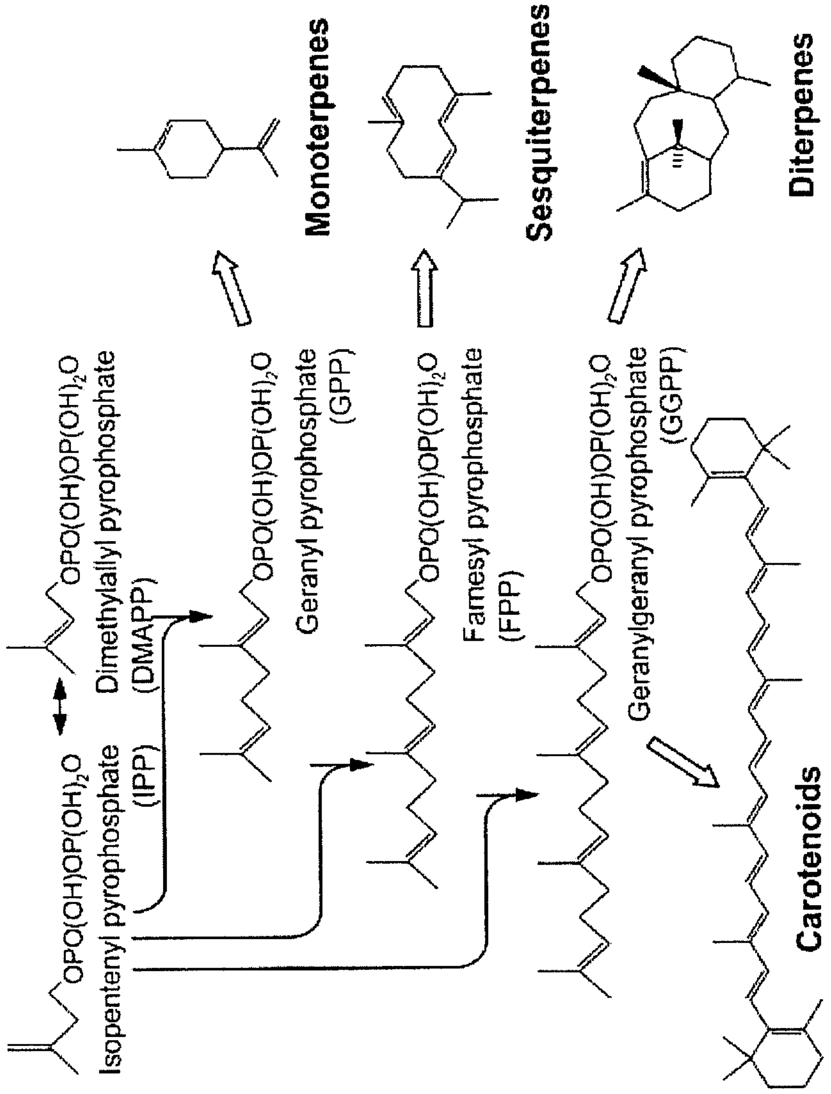

OPO(OH)OP(OH)2O
Isopentenyl pyrophosphate
(IPP)
OPO(OH)OP(OH)2O
Dimethylallyl pyrophosphate
(DMAPP)
OPO(OH)OP(OH)2O
Geranyl pyrophosphate
(GPP)
OPO(OH)OP(OH)2O
Farnesyl pyrophosphate
(FPP)
OPO(OH)OP(OH)2O
Geranylgeranyl pyrophosphate
(GGPP)
Monoterpenes
Sesquiterpenes
Diterpenes
Carotenoids

Figure 3

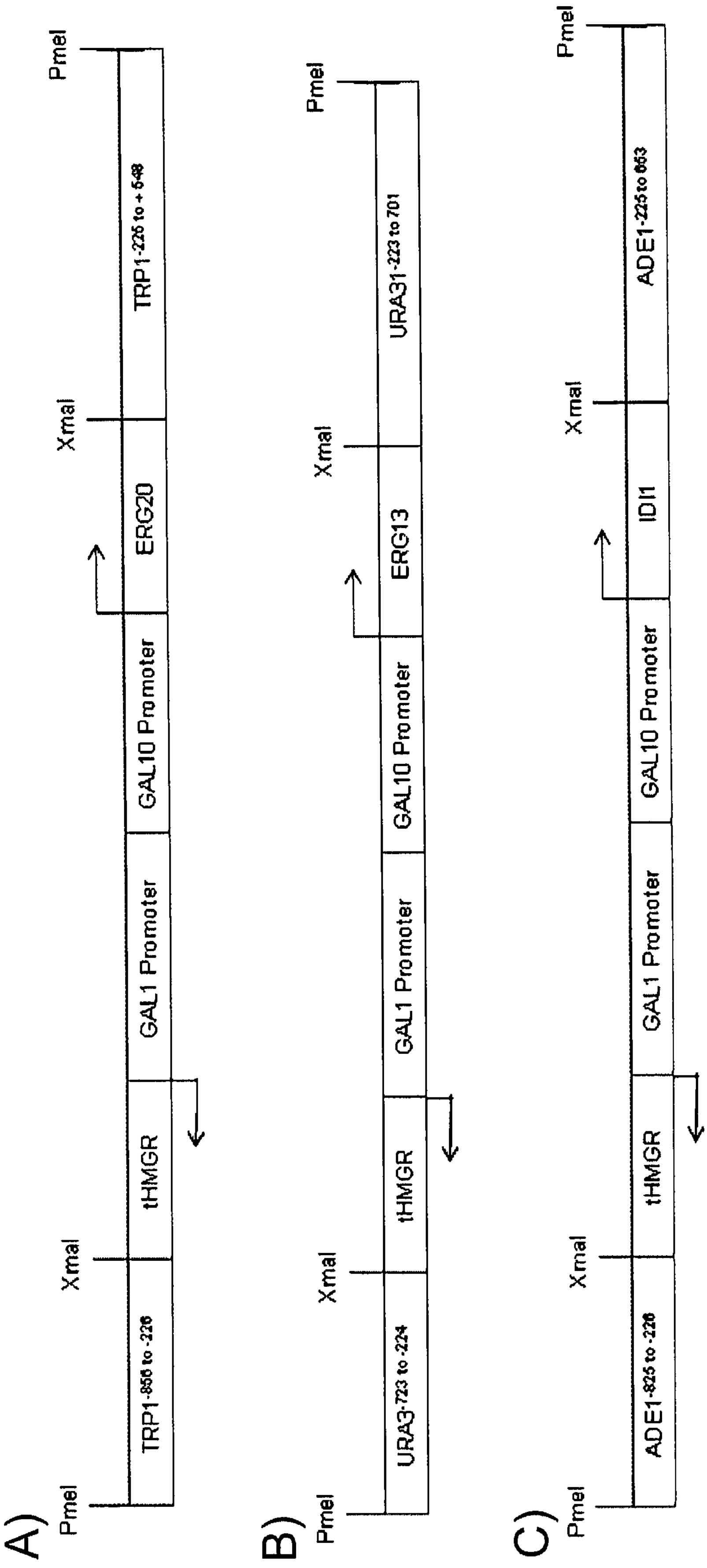

A)
Pmel
TRP1$^{-856 \text{ to } -226}$
XmaI
tHMGR
GAL1 Promoter
GAL10 Promoter
ERG20
XmaI
TRP1$^{-226 \text{ to } +548}$
Pmel
B)
Pmel
URA3$^{-723 \text{ to } -224}$
XmaI
tHMGR
GAL1 Promoter
GAL10 Promoter
ERG13
XmaI
URA31$^{-223 \text{ to } 701}$
Pmel
C)
Pmel
ADE1$^{-825 \text{ to } -226}$
XmaI
tHMGR
GAL1 Promoter
GAL10 Promoter
IDI1
XmaI
ADE1$^{-225 \text{ to } 653}$
Pmel

Figure 3 (continued)

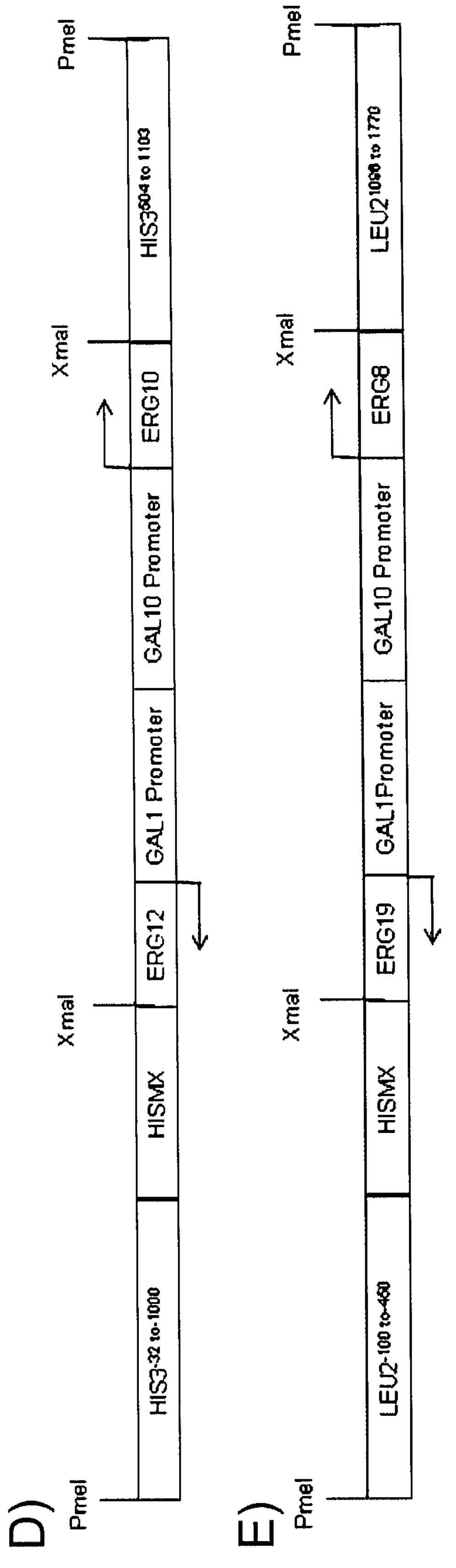


Figure 4

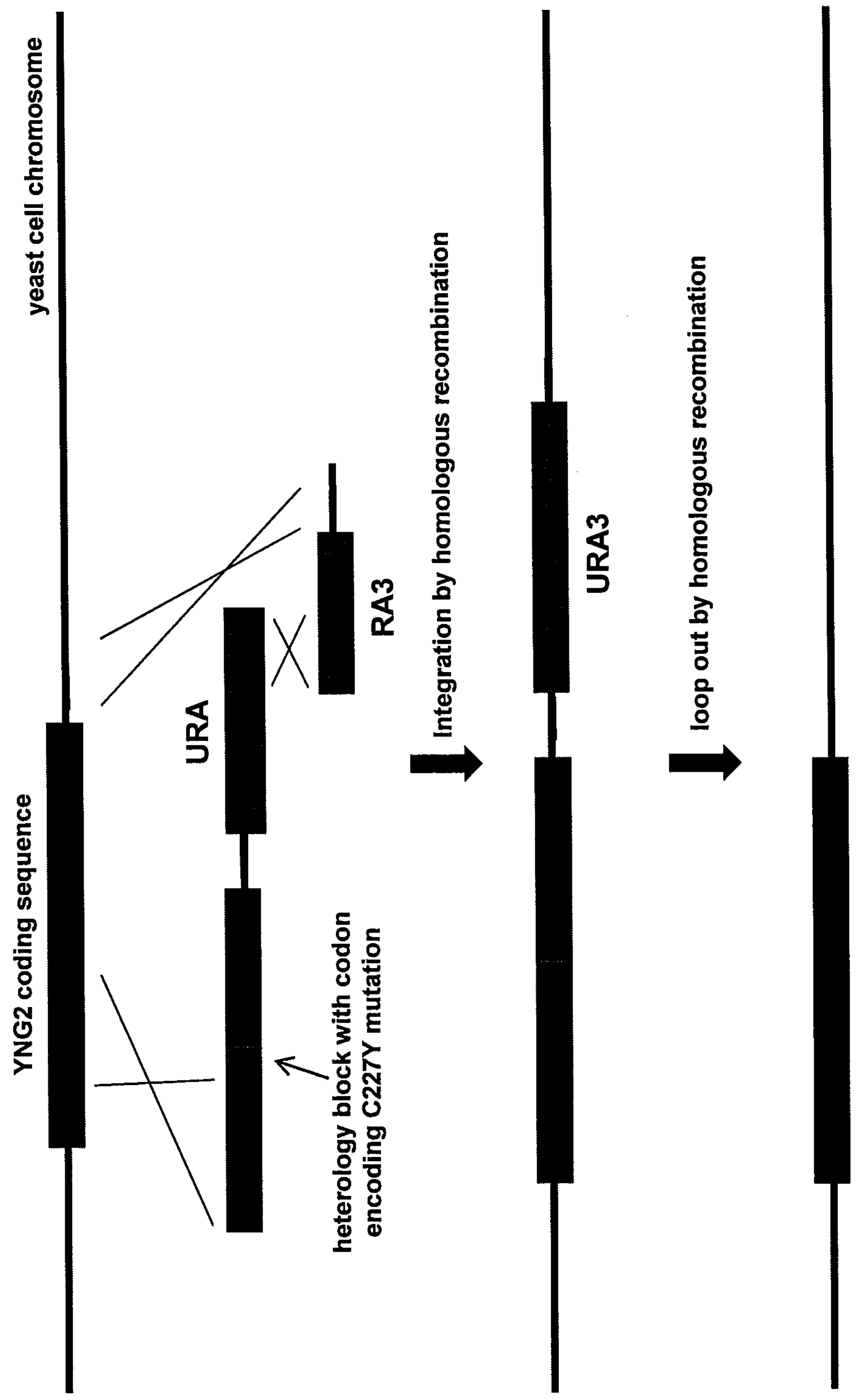

yeast cell chromosome
YNG2 coding sequence
URA
heterology block with codon encoding C227Y mutation
RA3
Integration by homologous recombination
URA3
loop out by homologous recombination

Figure 5

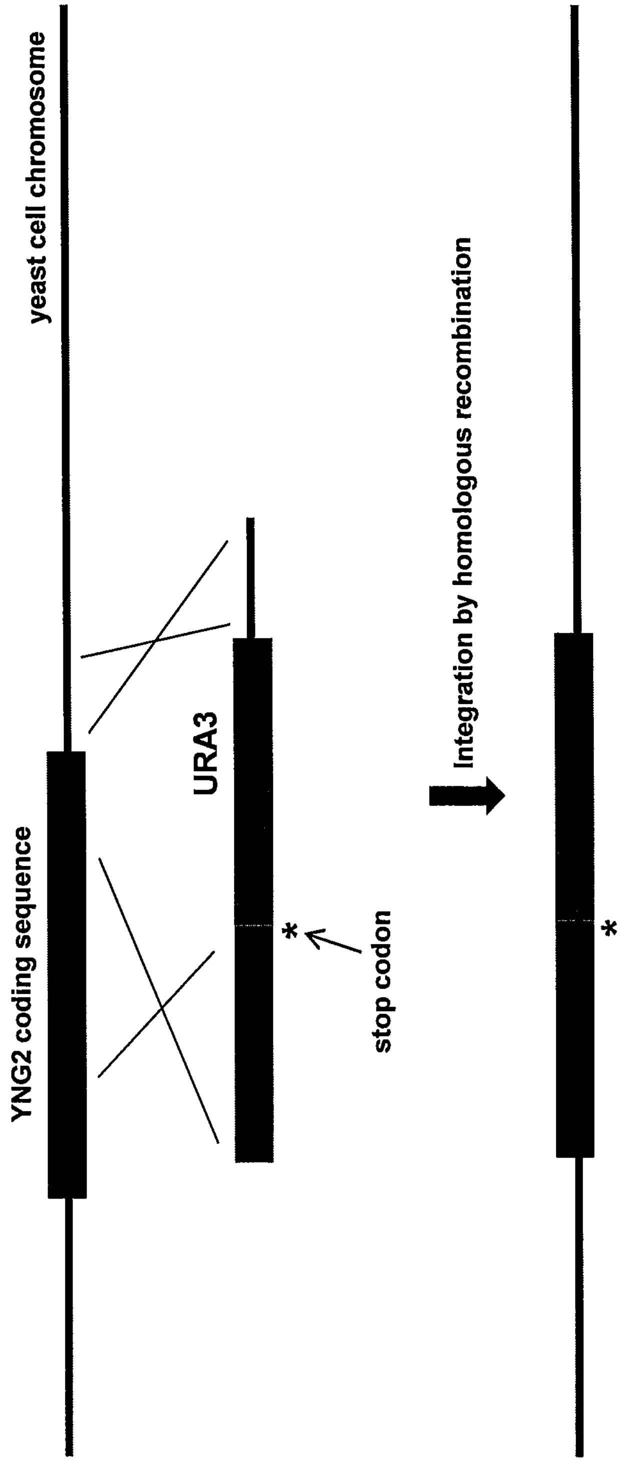

yeast cell chromosome
YNG2 coding sequence
URA3
stop codon
*
Integration by homologous recombination
*

Figure 6

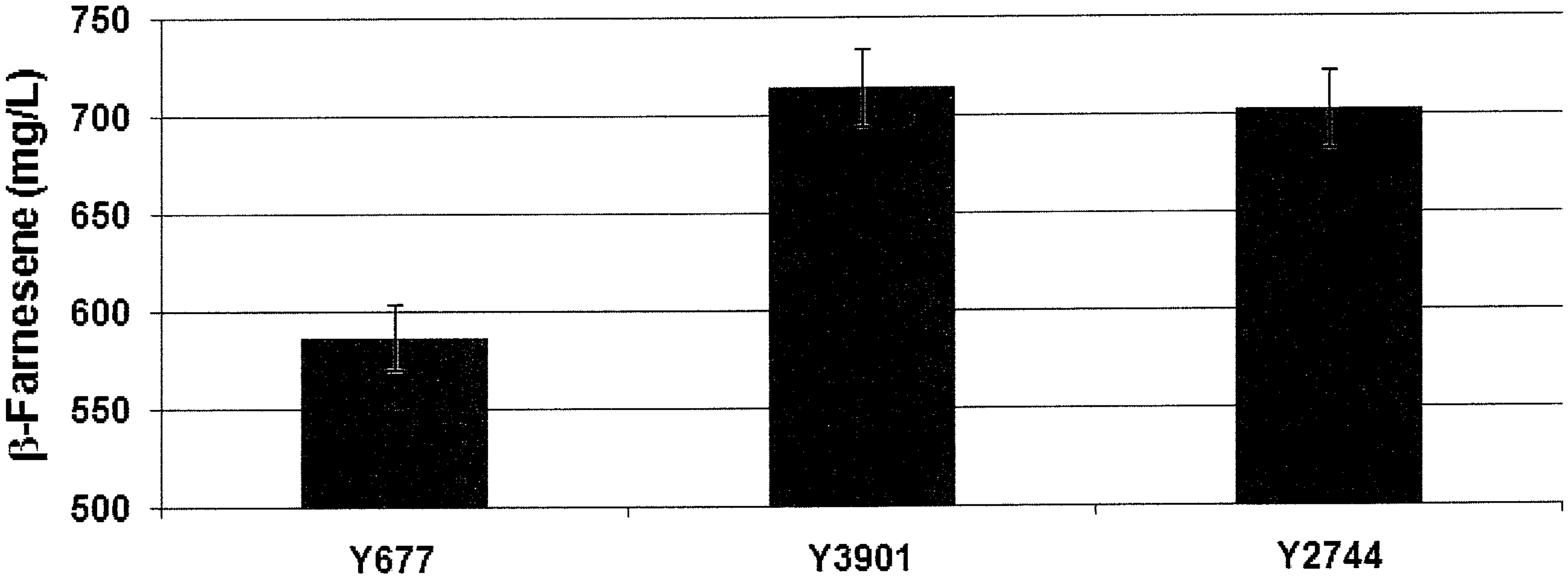


Figure 7

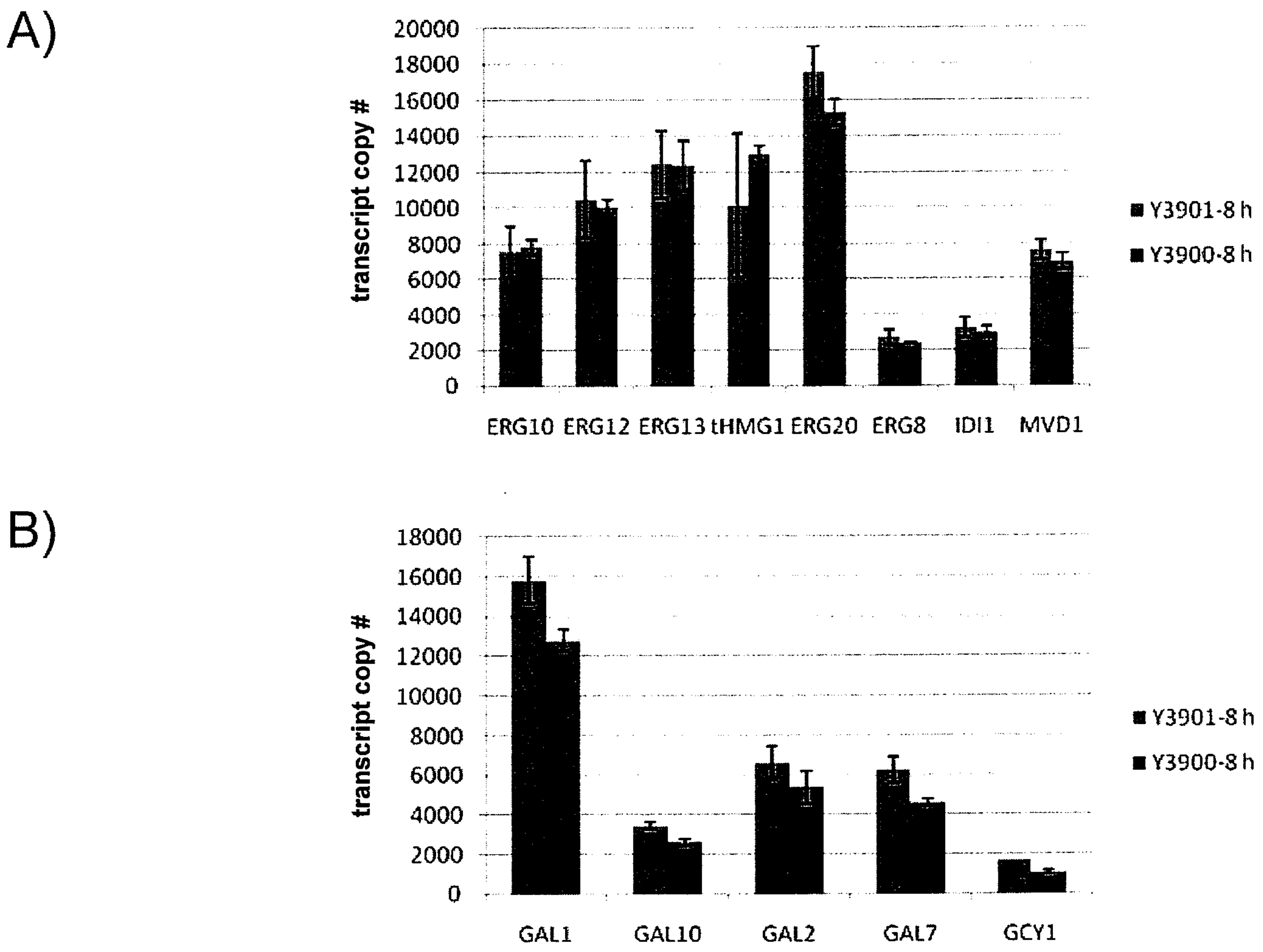

A)
transcript copy #
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
ERG10
ERG12
ERG13
tHMG1
ERG20
ERG8
IDI1
MVD1
■ Y3901-8 h
■ Y3900-8 h
B)
transcript copy #
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
GAL1
GAL10
GAL2
GAL7
GCY1
■ Y3901-8 h
■ Y3900-8 h

Figure 7 (continued)

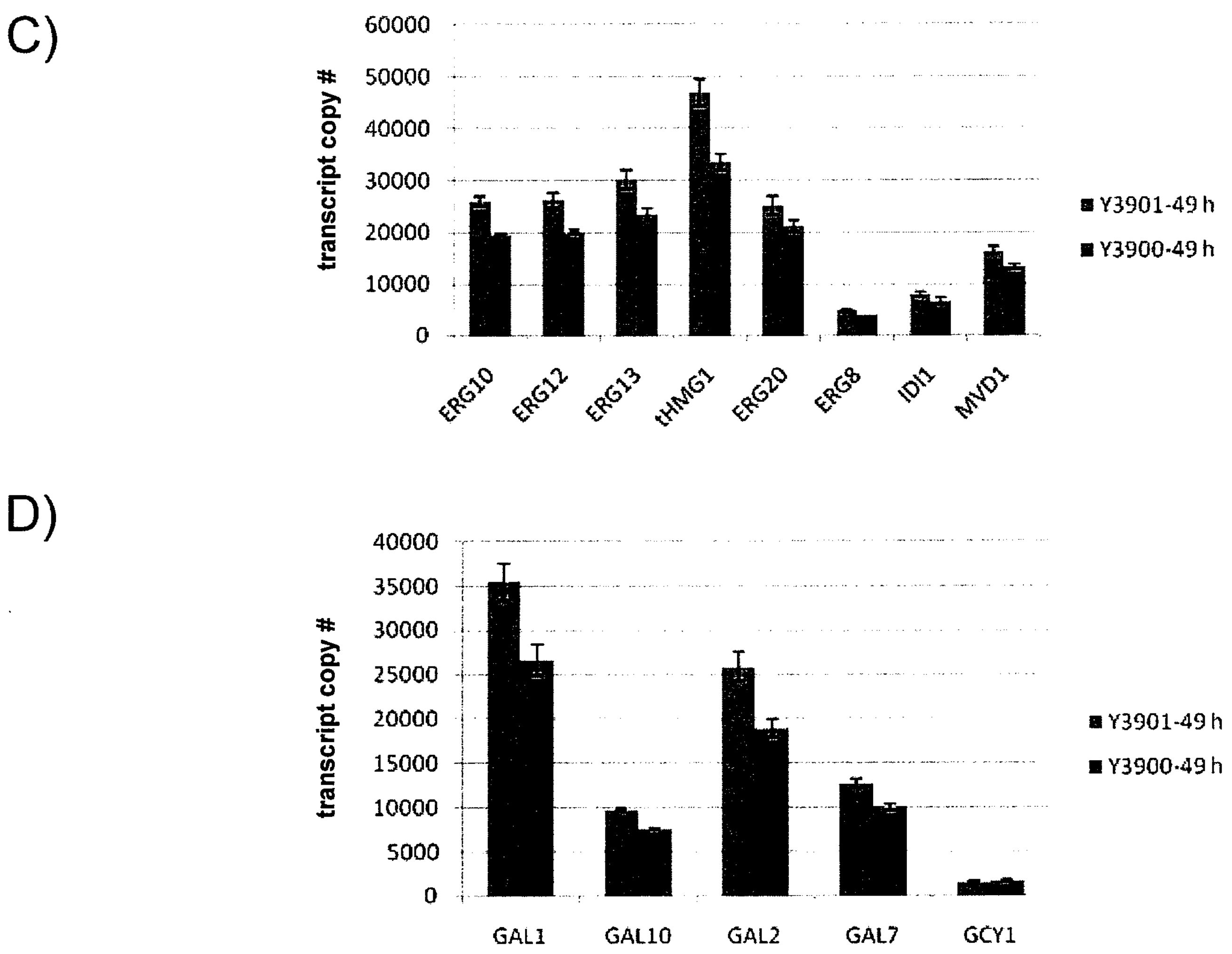
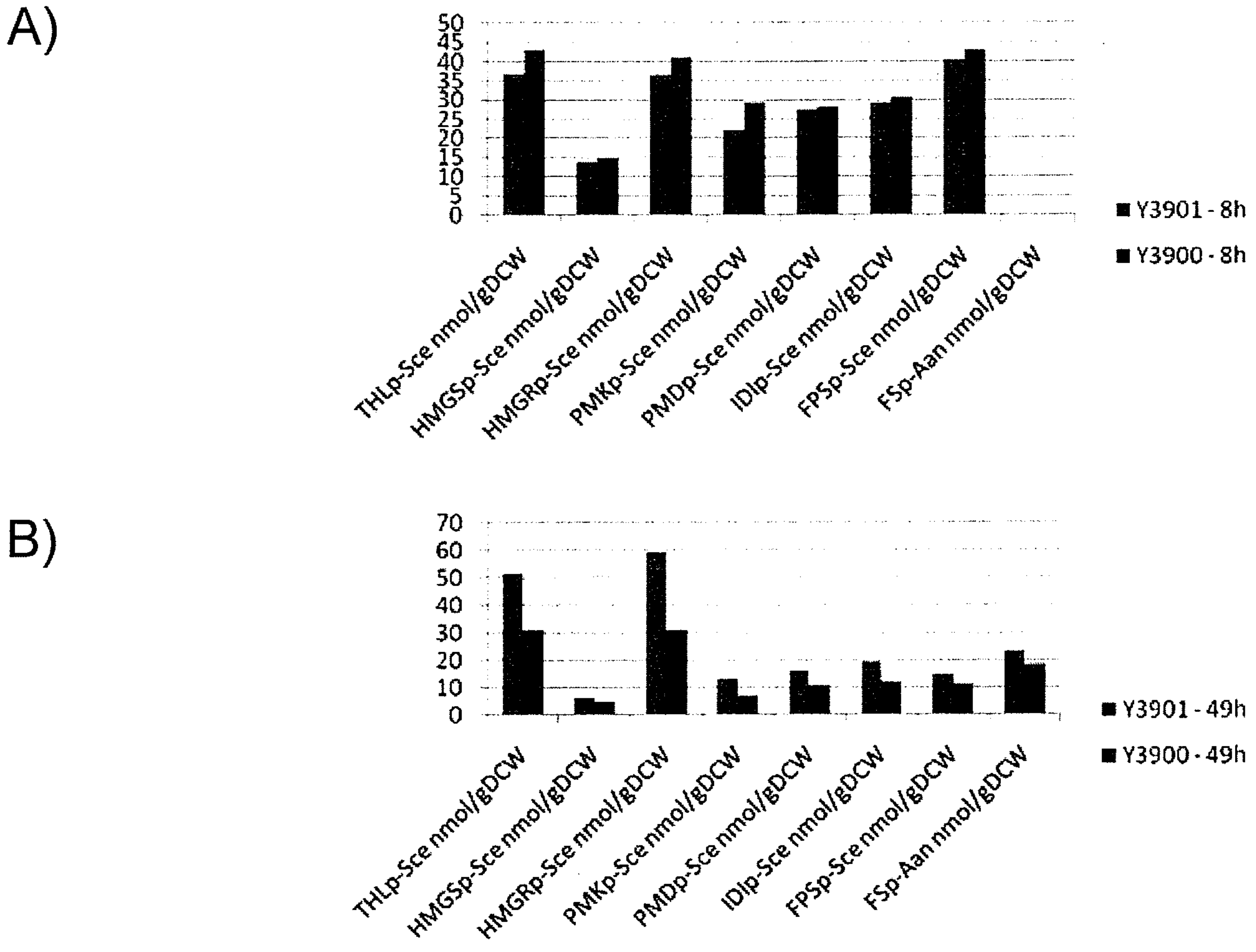


Figure 8


A)
50
45
40
35
30
25
20
15
10
5
0
THLp-Sce nmol/gDCW
HMGSp-Sce nmol/gDCW
HMGRp-Sce nmol/gDCW
PMKp-Sce nmol/gDCW
PMDp-Sce nmol/gDCW
IDIp-Sce nmol/gDCW
FPSp-Sce nmol/gDCW
FSp-Aan nmol/gDCW
■ Y3901 - 8h
■ Y3900 - 8h
B)
70
60
50
40
30
20
10
0
THLp-Sce nmol/gDCW
HMGSp-Sce nmol/gDCW
HMGRp-Sce nmol/gDCW
PMKp-Sce nmol/gDCW
PMDp-Sce nmol/gDCW
IDIp-Sce nmol/gDCW
FPSp-Sce nmol/gDCW
FSp-Aan nmol/gDCW
■ Y3901 - 49h
■ Y3900 - 49h