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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 October 2012 (04.10.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/135591 A2

(51) International Patent Classification:
C12P 9/00 (2006.01)

(21) International Application Number:
PCT/US2012/031395

(22) International Filing Date:
30 March 2012 (30.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/469,767 30 March 2011 (30.03.2011) US

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

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZHAO, Lishan** [CN/US]; 1465 65th Street, Apt. 470, Emeryville, CA 94608 (US). **DIETZEL, Kevin, L.** [US/US]; 5885 Hollis Street, Suite 100, Emeryville, CA 94608 (US). **WICHMANN, Gale, A.** [US/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, QA, RO, RS, RU, RW, 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, RW, 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).

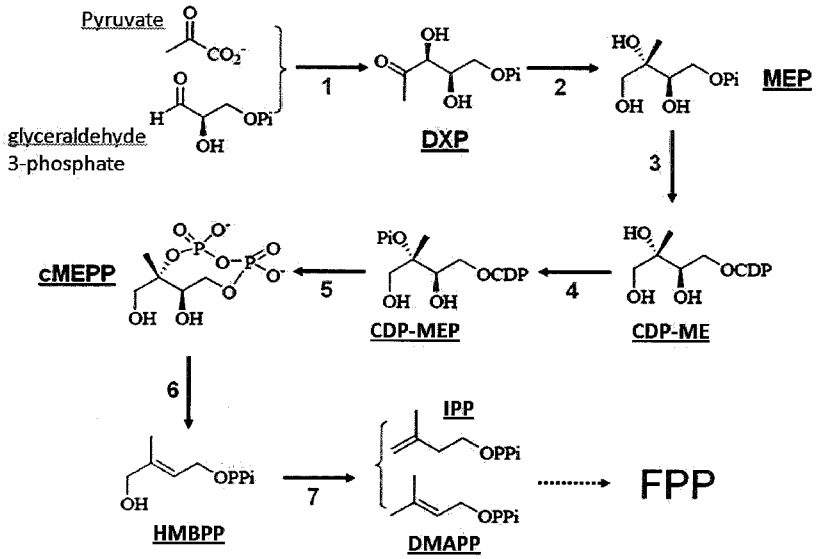
Published:

— *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: MICROBIAL ISOPRENOID PRODUCTION USING A HETEROLOGOUS DXP PATHWAY

Figure 1



(57) Abstract: The present invention provides methods for a robust production of isoprenoids utilizing the DXP biosynthetic pathway. The invention also provides nucleic acids, enzymes, expression vectors, and genetically modified host cells for carrying out the subject methods.

WO 2012/135591 A2

MICROBIAL ISOPRENOID PRODUCTION USING A HETEROLOGOUS DXP PATHWAY

BACKGROUND OF THE INVENTION

Isoprenoids are ubiquitous in nature and are made in all living organisms. They comprise a diverse family of over 40,000 compounds and serve diverse functions including maintaining cellular fluidity, electron transport, and other metabolic functions. A vast number of natural and synthetic isoprenoids are useful as pharmaceuticals, cosmetics, perfumes, pigments and colorants, fungicides, antiseptics, nutraceuticals, and fine chemical intermediates.

An isoprenoid product is typically composed of repeating five carbon isopentenyl diphosphate ("IPP") units, although irregular isoprenoids and polyterpenes have been reported. In nature, isoprenoids are synthesized by consecutive condensations of their precursor IPP and its isomer, dimethylallyl pyrophosphate ("DMAPP"). Two pathways for these precursors are known. Eukaryotes, with the exception of plants, use the mevalonate-dependent ("MEV") pathway to convert acetyl coenzyme A (acetyl-CoA) to IPP, which is subsequently isomerized to DMAPP. Prokaryotes, with some exceptions, typically employ only the mevalonate-independent or deoxyxylulose-5-phosphate ("DXP") pathway to produce IPP and DMAPP. Plants use both the MEV pathway and the DXP pathway.

Traditionally, isoprenoids have been manufactured by extraction from natural sources such as plants, microbes, and animals. However, extraction yield is usually very low due to a number of profound limitations. First, most isoprenoids accumulate in nature in only small amounts. Second, the source organisms are generally not amenable to the large-scale cultivation that is needed to produce commercially-viable quantities of a desired isoprenoid.

Elucidation of the MEV and DXP metabolic pathways, as well as advances in molecular biology have made biosynthetic production of isoprenoids feasible, particularly using the MEV pathway. However, unlike the MEV pathway, the genes encoding the DXP pathway have only recently been isolated and therefore, much less work has been performed in understanding the pathway. In general, to the extent that the DXP pathway has been manipulated, it has been in *E. coli* which includes an endogenous DXP pathway, and thus the host cell necessarily possesses all of the required accessory proteins and cofactors for a functional DXP pathway. In a few instances, the installation of a DXP pathway in a host cell (*e.g.*, *S. cerevisiae*) that does not have an endogenous DXP pathway has been reported. However, only the seven enzymes that convert glyceraldehyde 3 phosphate and pyruvate to IPP and/or DMAPP were expressed and no proof was presented that the IPP or DMAPP made in the host cell was a result of the heterologous *E. coli* DXP enzymes (versus from the endogenously present MEV pathway in *S. cerevisiae*).

SUMMARY OF THE INVENTION

The present invention generally relates to methods and materials useful for isoprenoid compounds using a microbial host cell using the DXP pathway. In particular, provided herein is an iron-sulfur enzyme system (iron-sulfur enzyme plus the two accessory proteins) that can be used for enhancing DXP pathway-mediated isoprenoid production in a host cell with an endogenous DXP pathway, or for engineering a heterologous DXP pathway in a host cell lacking an endogenous DXP pathway, and using the heterologous DXP pathway for isoprenoid production.

In one aspect of the invention, an iron sulfur enzyme system is provided comprising one or more nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) an iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

In another aspect of the invention, an iron sulfur enzyme system is provided comprising one or more nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) an iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting cMEPP to isopentenyl pyrophosphate ("IPP") and dimethylallyl pyrophosphate ("DMAPP").

In another aspect of the invention, a genetically modified microorganism is provided comprising one or more heterologous nucleic acid molecules that encode:

(a) a first reductase that is capable of accepting electrons from an electron donor thereby converting the first reductase from an oxidized state ("oxidized first reductase") to a reduced state ("reduced first reductase");

(b) a first redox protein that is capable of accepting electrons from the reduced first reductase thereby converting the first redox protein from an oxidized state ("oxidized first redox protein") to a reduced state ("reduced first redox protein"); and

(c) a first iron-sulfur enzyme that is capable of accepting electrons from the first reduced redox protein thereby converting the first iron-sulfur enzyme from an oxidized state ("oxidized first iron-sulfur enzyme") to a reduced state ("reduced first iron-sulfur enzyme") and wherein the reduced first iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

In other embodiments, the one or more heterologous nucleic acid molecules further encode:

(d) a second reductase that is capable of accepting electrons from an electron donor thereby converting the second reductase from an oxidized state ("oxidized second reductase") to a reduced state ("reduced second reductase");

(e) a second redox protein that is capable of accepting electrons from the reduced second reductase thereby converting the second redox protein from an oxidized state ("oxidized second redox protein") to a reduced state ("reduced second redox protein"); and

(f) a second iron-sulfur enzyme that is capable of accepting electrons from the second reduced redox protein thereby converting the second iron-sulfur enzyme from an oxidized state ("oxidized second iron-sulfur enzyme") to a reduced state ("reduced second iron-sulfur enzyme") and wherein the reduced second iron-sulfur enzyme is capable of converting HMBPP to isopentenyl pyrophosphate ("IPP") or/and dimethylallyl pyrophosphate ("DMAPP").

In other embodiments, the first reductase and the second reductase are the same reductase. In still other embodiments, the first redox protein and the second redox protein are the same redox protein. In other embodiments, the first reductase and the second reductase are the same reductase and the first redox protein and the second redox protein are the same redox protein.

In another aspect of the invention, a genetically modified microorganism is provided comprising a heterologous DXP pathway that is capable of converting glyceraldehydes 3-phosphate and pyruvate to IPP and DMAPP and wherein the genetically modified microorganism does not include an endogenous DXP pathway.

INCORPORATION BY REFERENCE

All patents, patent applications, and publications mentioned in this specification are herein incorporated by reference to the same extent as though each individual publication or patent document was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic representation of one embodiment of the DXP pathway as well as the additional step of converting IPP and DMAPP to FPP.

Figure 2 shows the internal cellular concentrations of DXP pathway metabolites cMEPP and DXP in yeast strains comprising heterologous DXP pathway enzymes.

Figure 3A shows a schematic map of the pGAL expression vector with an ispG or ispH insert.

Figure 3B shows a Western plot of *Escherichia coli* IspG expressed as soluble protein (lane 2) or targeted to the mitochondria (lane 3) in a yeast host cell (lane 1 = molecular weight ladder).

Figures 4A and 4B show activity profiles of IspG-fld and IspH-fld fusion proteins, respectively.

Figure 5 shows ^{55}Fe incorporation in heterologously-expressed ispG and ispH proteins, normalized to the amount of protein assayed. All ^{55}Fe counts/µg protein were normalized to the level of ^{55}Fe found in LEU1 (defined as 1.0).

Figure 6 shows iron content in representative IspG and IspH proteins purified from *S. cerevisiae*. Leu1, a naturally occurring *S. cerevisiae* protein with four Fe-S clusters, serves as a control. Leu1 was purified using the same methods as the IspG and IspH proteins.

Figure 7 shows the intracellular concentration (µM) of IspG product, HDMAPP, in yeast strains expressing a full DXP pathway (Y6283), IspG and their redox partners only (Y6291 and Y6292), or DXP-TOP pathway only (Y4819).

Figure 8 shows the intracellular levels of DMAPP and FPP in yeast strains expressing a full DXP pathway, which is suppressed in Y6540 and activated in Y6533, when these strains are grown on glucose. Y6533 and Y6540 strains were grown on 2% ^{13}C-glucose synthetic medium without adenine that was supplemented with 2 g/L ^{12}C-mevalonate. Y6540 + ZA and Y6533 + ZA samples represent cultures that also contained 16 µg/mL zaragozic acid.

Figures 9A through 9D show schematic maps of expression vectors used to generate host strains expressing heterologous DXP pathway enzymes.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

Reference is made here to a number of terms that shall be defined to have the following meanings:

The term "optional" or "optionally" means that the subsequently described feature or structure may or may not be present, or that the subsequently described event or circumstance may or may not occur, and that the description includes instances where a particular feature or structure is present and instances where the feature or structure is absent, or instances where the event or circumstance occurs and instances where the event or circumstance does not occur.

The term "deoxyxylulose 5-phosphate pathway" or "DXP pathway" is used herein to refer to a set of enzymes that converts glyceraldehyde-3-phosphate and pyruvate to IPP and/or DMAPP. Figure 1 illustrates one embodiment of the DXP pathway as well as the subsequent interconversion of IPP and DMAPP.

The term "pyrophosphate" is used interchangeably herein with "diphosphate".

The terms "expression vector" or "vector" refer to a nucleic acid that transduces, transforms, or infects a host cell, thereby causing the cell to produce nucleic acids and/or proteins other than those that are native to the cell, or to express nucleic acids and/or proteins in a manner that is not native to the cell.

The term "endogenous" refers to a substance or process that occurs naturally, *e.g.*, in a non-recombinant or unmodified microorganism.

The term "nucleic acid" refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and pyrimidine bases or other natural, chemically, or biochemically modified, non-natural, or derivatized nucleotide bases.

The term "operon" is used to refer to two or more contiguous nucleotide sequences that each encode a gene product such as a RNA or a protein, and the expression of which are coordinately regulated by one or more controlling elements (for example, a promoter).

The term "heterologous nucleic acid" as used herein refers to a nucleic acid wherein at least one of the following is true: (a) the nucleic acid is foreign ("exogenous") to (that is, not naturally found in) a given host cell; (b) the nucleic acid comprises a nucleotide sequence that is naturally found in (that is, is "endogenous to") a given host cell, but the nucleotide sequence is produced in an unnatural (for example, greater than expected or greater than naturally found) amount in the cell; (c) the nucleic acid comprises a nucleotide sequence that differs in sequence from an endogenous nucleotide sequence, but the nucleotide sequence encodes the same protein (having the same or substantially the same amino acid sequence) and is produced in an unnatural

(for example, greater than expected or greater than naturally found) amount in the cell; or (d) the nucleic acid comprises two or more nucleotide sequences that are not found in the same relationship to each other in nature (for example, two or more gene sequences are placed closer together and/or in a different order than naturally found in the host cell).

The term "recombinant host" (also referred to as a "genetically modified host cell" or "genetically modified microorganism") denotes a host cell that comprises a heterologous nucleic acid of the invention.

The term "exogenous nucleic acid" refers to a nucleic acid that is exogenously introduced into a host cell, and hence is not normally or naturally found in and/or produced by a given cell in nature.

The term "regulatory element" refers to transcriptional and translational control sequences, such as promoters, enhancers, polyadenylation signals, terminators, protein degradation signals, and the like, that provide for and/or regulate expression of a coding sequence and/or production of an encoded polypeptide in a host cell.

The term "transformation" refers to a permanent or transient genetic change induced in a cell following introduction of new nucleic acid. Genetic change ("modification") can be accomplished either by incorporation of the new DNA into the genome of the host cell, or by transient or stable maintenance of the new DNA as an episomal element. In eukaryotic cells, a permanent genetic change is generally achieved by introduction of the DNA into the genome of the cell. In prokaryotic cells, a permanent genetic change can be introduced into the chromosome or via extrachromosomal elements such as plasmids and expression vectors, which may contain one or more selectable markers to aid in their maintenance in the recombinant host cell.

The term "operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a nucleotide sequence if the promoter affects the transcription or expression of the nucleotide sequence.

The term "host cell" and "microorganism" are used interchangeably herein to refer to any archae, bacterial, or eukaryotic living cell into which a heterologous nucleic acid can be, or has been inserted. The term also relates to the progeny of the original cell, which may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to natural, accidental, or deliberate mutation.

The term "naturally occurring" as applied to a nucleic acid, an enzyme, a cell, or an organism, refers to a nucleic acid, enzyme, cell, or organism that is found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism that can be isolated from

a source in nature and that has not been intentionally modified by a human in the laboratory is naturally occurring.

The terms "isoprenoid", "isoprenoid compound", "isoprenoid product", "terpene", "terpene compound", "terpenoid", and "terpenoid compound" are used interchangeably herein. They refer to compounds that are capable of being derived from IPP. Exemplary isoprenoids include, but are not limited to, monoterpenes, diterpenes, sesquiterpenes, triterpenes, and polyterpenes.

The singular forms "a," "and," and "the" include plural references unless the context clearly dictates otherwise. For example, reference to "an expression vector" includes a single expression vector as well as a plurality of expression vectors, and reference to "the host cell" includes reference to one or more host cells, and so forth. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

Unless otherwise indicated, this invention is not limited to particular sequences, expression vectors, enzymes, host microorganisms, or processes, as such may vary in accordance with the understanding of those of ordinary skill in the arts to which this invention pertains in view of the teachings herein. Terminology used herein is for purposes of describing particular embodiments only and is not intended to be limiting.

Enzymes of the DXP Pathway

The term "deoxyxylulose 5-phosphate pathway" or "DXP pathway" is used herein to refer to a set of enzymes that converts glyceraldehyde-3-phosphate and pyruvate to IPP or/and DMAPP. Figure 1 illustrates one embodiment of the DXP pathway comprising seven enzymes as well as the subsequent conversion of IPP and DMAPP to farnesyl pyrophosphate ("FPP"). However, the DXP pathway may comprise fewer than seven, or more than seven enzymes, so long as the starting materials are glyceraldehydes 3-phosphate and pyruvate and the products of the pathway are IPP and DMAPP. IPP and DMAPP are referred to as universal isoprenoid intermediates because all isoprenoid compounds are derived therefrom.

The seven enzymatic reactions illustrated in Figure 1 are described in greater detail as follows:

In the first step, pyruvate is condensed with D-glyceraldehyde 3-phosphate to make 1-deoxy-D-xylulose-5-phosphate ("DXP"). An enzyme known to catalyze this step is, for example, 1-deoxy-D-xylulose-5-phosphate synthase. The gene encoding 1-deoxy-D-xylulose-5-phosphate synthase is referred to as dxs. Illustrative examples of nucleotide sequences for dxs encoding 1-deoxy-D-xylulose-5-phosphate synthases include, but are not limited to, (AF035440; *Escherichia coli*), (NC_002947, locus tag PP0527; *Pseudomonas putida* KT2440), (CP000026, locus tag SPA2301;

Salmonella enterica subsp. enterica serovar Paratyphi A str. ATCC 9150), (NC_007493, locus tag RSP_0254; *Rhodobacter sphaeroides* 2.4.1), (NC_005296, locus tag RPA0952; *Rhodopseudomonas palustris* CGA009), (NC_004556, locus tag PD1293; *Xylella fastidiosa Temecula1*), (NC_003076, locus tag AT5G11380; *Arabidopsis thaliana*), (Y18874, *Synechococcus* PCC6301), (AB026631, *Streptomyces sp.* CL190), (AB042821, *Streptomyces griseolosporeus*), (AF111814, *Plasmodium falciparum*), (AF143812, *Lycopersicon esculentum*), (AJ279019, *Narcissus pseudonarcissus*), and (AJ291721, *Nicotiana tabacum*).

In the second step, DXP is converted to 2C-methyl-D-erythritol-4-phosphate ("MEP"). An enzyme known to catalyze this step is, for example, 1-deoxy-D-xylulose-5-phosphate reductoisomerase. The gene encoding this enzyme is referred to as dxr or ispC. Illustrative examples of dxr or ispC nucleotide sequences encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerases include, but are not limited to, (AB013300; *Escherichia coli*), (AF148852; *Arabidopsis thaliana*), (NC_002947, locus tag PP1597; *Pseudomonas putida* KT2440), (AL939124, locus tag SCO5694; *Streptomyces coelicolor* A3(2)), (NC_007493, locus tag RSP_2709; *Rhodobacter sphaeroides* 2.4.1), (NC_007492, locus tag Pfl_1107; *Pseudomonas fluorescens* PfO-1), (AB049187, *Streptomyces griseolosporeus*), (AF111813, *Plasmodium falciparum*), (AF116825, *Mentha x piperita*), (AF182287, *Artemisia annua*), (AF250235, *Catharanthus roseus*), (AF282879, *Pseudomonas aeruginosa*) (AJ242588, *Arabidopsis thaliana*), (AJ250714, *Zymomonas mobilis* strain ZM4), (AJ292312, *Klebsiella pneumoniae*), and (AJ297566, *Zea mays*).

In the third step, MEP is converted to 4-diphosphocytidyl-2C-methyl-D-erythritol ("CDP-ME"). An enzyme known to catalyze this step is, for example, 4-diphosphocytidyl-2C-methyl-D-erythritol synthase. The gene encoding this enzyme is referred to as ispD or ygbP. Illustrative examples of ispD or ygbP nucleotide sequences encoding 4-diphosphocytidyl-2C-methyl-D-erythritol synthases include, but are not limited to, (AF230736; *Escherichia coli*), (NC_007493, locus_tag RSP_2835; *Rhodobacter sphaeroides* 2.4.1), (NC_003071, locus_tag AT2G02500; *Arabidopsis thaliana*), (AB037876, *Arabidopsis thaliana*), (AF109075, *Clostridium difficile*), and (AF230737, Arabidopsis thaliana).

In the fourth step, CDP-ME is converted to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate ("CDP-MEP"). An enzyme known to catalyze this step is, for example, 4-diphosphocytidyl-2C-methyl-D-erythritol kinase. The gene encoding this enzyme is referred to as ispE or ychB. Illustrative examples of ispE or ychB nucleotide sequences encoding 4-diphosphocytidyl-2C-methyl-D-erythritol kinases, include, but are not limited to, (AF216300; *Escherichia coli*), (NC_007493, locus_tag RSP_1779; *Rhodobacter sphaeroides* 2.4.1), (AF263101, *Lycopersicon esculentum*), and (AF288615, *Arabidopsis thaliana*).

In the fifth step, CDP-MEP is converted to 2C-methyl-D-erythritol 2, 4-cyclodiphosphate ("cMEPP"). An enzyme known to catalyze this step is, for example, 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase. The gene encoding this enzyme is referred to as IspF or ygbB. Illustrative examples of IspF or ygbB nucleotide sequences encoding 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthases include, but are not limited to, (AF230738; *Escherichia coli*), (NC_007493, locus_tag RSP_6071; *Rhodobacter sphaeroides* 2.4.1), (NC_002947, locus_tag PP1618; *Pseudomonas putida* KT2440), (AB038256, *Escherichia coli* mecs gene), (AF250236, *Catharanthus roseus* (MECS), (AF279661, *Plasmodium falciparum*), and (AF321531, *Arabidopsis thaliana*).

In the sixth step, cMEPP is converted to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP"). An enzyme known to catalyze this step is, for example, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase. The gene encoding this enzyme is referred to as ispG or gcpE. Illustrative examples of ispG or gcpE nucleotide sequences encoding 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthases include, but are not limited to, (AY033515; *Escherichia coli*), (NC_002947, locus_tag PP0853; *Pseudomonas putida* KT2440), (NC_007493, locus_tag RSP_2982; *Rhodobacter sphaeroides* 2.4.1), (O67496, *Aquifex aeolicus*), (P54482, *Bacillus subtilis*), (Q9pky3, *Chlamydia muridarum*), (Q9Z8H0, *Chlamydophila pneumoniae*), (O84060, *Chlamydia trachomatis*), (P27433, *Escherichia coli*), (P44667, *Haemophilus influenzae*), (Q9ZLL0, *Helicobacter pylori* J99), (O33350, *Mycobacterium tuberculosis*), (S77159, *Synechocystis sp.*), (Q9WZZ3, *Thermotoga maritima*), (O83460, *Treponema pallidum*), (Q9JZ40, *Neisseria meningitidis*), (Q9PPM1, *Campylobacter jejuni*), (Q9RXC9, *Deinococcus radiodurans*), (AAG07190, *Pseudomonas aeruginosa*) and (Q9KTX1, *Vibrio cholerae*).

In the seventh step, HMBPP is converted to isopentyl pyrophosphate ("IPP") and/or its isomer, dimethylallyl diphosphate ("DMAPP"). An enzyme known to catalyze this step is, for example, isopentyl/dimethylallyl diphosphate synthase. The gene encoding this enzyme is referred to as ispH or lytB. Illustrative examples of nucleotide sequences that encode a isopentyl/dimethylallyl diphosphate synthase include, but are not limited to, (AY062212; *Escherichia coli*), (NC_002947, locus_tag PP0606; *Pseudomonas putida* KT2440), (AF027189, *Acinetobacter* sp. BD413), (AF098521, *Burkholderia pseudomallei*), (AF291696, *Streptococcus pneumoniae*), (AF323927, *Plasmodium falciparum*), (M87645, *Bacillus subtillis*), (U38915, *Synechocystis sp.*), and (X89371, *C. jejunisp* O67496)

IPP and DMAPP are enzymatically interconverted. An enzyme known to catalyze this reaction is, for example, IPP isomerase. The gene encoding this enzyme is referred to as idi. Illustrative examples of idi nucleotide sequences include, but are not limited to, (NC_000913,

3031087..3031635; *Escherichia coli*), and (AF082326; *Haematococcus pluvialis*). In certain embodiments, the genetically modified organism encodes a heterologous nucleic acid encoding an enzyme capable of interconverting IPP and DMAPP, which ensures that the conversion of IPP into DMAPP is not a rate-limiting step in the biosynthesis of the desired isoprenoid. IPP and DMAPP are subsequently converted to the desired isoprenoid compound (which occurs in five carbon units) via condensation. For example, condensation of two, three, and four five-carbon units (of either IPP, DMAPP, or combinations thereof) results in geranyl pyrophosphate, farnesyl pyrophosphate and geranylgeranyl pyrophosphate, respectively, which are in turn modified by a terpene synthase to a desired monoterpene, sesquiterpene, and diterpene.

Engineering the DXP Pathway

As described previously, the DXP pathway illustrated in Figure 1 includes seven reactions. Two of these seven reactions are catalyzed by iron-sulfur enzymes. The first iron-sulfur enzyme converts cMEPP to HMBPP and the second iron-sulfur enzyme converts HMBPP to IPP and/or DMAPP (see, for example, reactons 6 and 7 of Figure 1). These enzymes are so-named because they include an iron-sulfur cofactor (also called an iron-sulfur cluster) in the active site of the enzyme that is required for their respective enzymatic reactions (which typically involve a reduction) to occur. Although many heterologous proteins (including the first five enzymes of the DXP pathway illustrated in Figure 1) are expressed successfully in non-native host cells, the expression of iron-sulfur enzymes has been far less successful due to the additional complexity of having multiple proteins involved in the proper assembly of an iron-sulfur cluster and its successful incorporation into an apoprotein. In addition, once properly assembled, the iron-sulfur enzyme still requires at least two partner proteins for it to be functionally competent to carry out its designated reaction. As a result, most heterologous expression of iron-sulfur enzymes have occurred in their endogenous host cells so that the additional complexities are eliminated. In this scenario, the iron-sulfur protein expression essentially becomes simply an over-expression of an endogenous enzyme.

As described previously, the DXP pathway is the endogenous pathway for most bacteria and for plastids, which are organelles found in plants and algae. Generally, successful heterologous expression can be expected within the same classification. For example, expressing a non-native bacterial DXP enzyme in a bacterial host cell with an endogenous DXP pathway, or expressing a non-native plastid DXP enzyme in a plastid with an endogenous DXP pathway. Although there is more complexity in these situations than simple over-expression of an endogenous enzyme, these situations are not too far from the baseline scenario of simple over-expression in a native host.

In contrast to these simpler scenarios of expressing DXP proteins in hosts having endogenous DXP pathways, the successful expression of a heterologous iron-sulfur DXP enzyme cannot be assumed in a host cell lacking an endogenous DXP pathway for the reasons previously described. As an initial matter, a threshold inquiry is whether the host cell without an endogenous DXP pathway has cellular machinery that can properly assemble the iron-sulfur cluster. The present inventor's studies have established that eukaryotic cells without an endogenous DXP pathway do indeed have the proper cellular machinery for assembling an iron-sulfur cluster. Given that the cellular machinery exists, the present inventors have determined that a major consideration in determining whether a DXP iron-sulfur enzyme can be successfully expressed, and can be functionally active in the host cell's cytosol, is its solubility.

For example, if the host cell is yeast, the following unmodified IspG proteins are not soluble in the cytosol and thus are not suitable for heterologous expression in their unmodified form in yeast (and likely not in eukaryotes generally): *Acidithiobacillus ferrooxidans* ATCC 53993 (YP_002220034.1); *Caulobacter crescentus* CB15 (NP_419668.1); *Desulfovibrio vulgaris* RCH1 (ZP_04791624.1); *Escherichia coli* (AAC75568); *Gemmata obscuriglobus* UQM 2246 (ZP_02736855.1); *Micrococcus luteus* NCTC 2665 (YP_002956780.1); *P. falciparum* 3D7 (NC_004314); *Pseudomonas fluorescens* SBW25 (YP_002874560.1); *Planctomyces maris* DSM 8797 (ZP_01855505.1); *Physcomitrella patens subsp. Patens* (XP_001766565); *Pirellula staleyi* DSM 6068 (YP_003370642.1); *Rhodobacter sphaeroides* ATCC 17025 (YP_001168057.1); *Stevia rebaudiana* (ABG75916); *Streptomyces sviceus* ATCC 29083 (ZP_06920121.1); *Streptomyces coelicolor* A3(2) (NP_630839); *Thalassiosira pseudonana* CCMP1335 (XP_002292108.1); *Wolbachia endosymbiont* of *Drosophila melanogaster* (NP_965936.1); and *Xanthomonas campestris pv. vesicatoria* str. 85-10 (YP_363560.1). And, the following unmodified IspH proteins are not soluble in yeast cytosol and thus not suitable for heterologous expression in their unmodified form in yeast (and likely not in eukaryotes generally): *Acidithiobacillus ferrooxidans* ATCC 53993 (YP_002221250.1); *Bacillus coagulans* 36D1 (ZP_04432889.1); *Gemmata obscuriglobus* UQM 2246 (ZP_02735110.1); *Magnetococcus sp.* MC-1 (YP_867320.1); *Micrococcus luteus* NCTC 2665 (YP_002957672.1); *Ostreococcus tauri* (CAL54853.1); *Prevotella buccae* D17 (EFC75366.1); *Plasmodium falciparum* 3D7 (CAD49005); *Pirellula staleyi* DSM 6068 (YP_003369145.1); and *Porphyra yezoensis* (ACI45962.1).

In contrast, the following IspG proteins are soluble in yeast cytosol and thus suitable for heterologous expression: *Algoriphagus* sp. PR1 (ZP_01720138.1); *Arthrospira platensis str. Paraca* (ZP_06384238.1); *Chlamydomonas reinhardtii* (XP_001690937.1); *Brevibacillus brevis* NBRC 100599 (YP_002772921.1); *Catharanthus roseus* (AAO24774.1); *Crocosphaera watsonii* WH 8501

(ZP_00518798.1); *Cyanobium* sp. PCC 7001 (ZP_05045928.1); *Eubacterium dolichum* DSM 3991 (ZP_02077259.1); *Nostoc punctiforme* PCC 73102 (YP_001868335.1); *Ostreococcus tauri* (CAL55258); *Prevotella buccae* D17 (EFC75705); *Prochlorococcus marinus subsp. pastoris str.* CCMP1986 (NP_892794.1); *Porphyra yezoensis* red algae (ACI45961); *Sorghum bicolor* (XP_002454137); *Synechococcus* sp. RS9917 (ZP_01081635.1); *Thermosynechococcus elongatus* BP-1 (NP_681786); and *Vitis vinifera* (XP_002285130.1). The following IspH proteins are soluble in yeast cytosol, and thus suitable for heterologous expression: *Escherichia coli* (AAC73140); *Fibrobacter succinogenes subsp. succinogenes* S85 (YP_003250012.1); Physcomitrella *patens subsp. patens* (XP_001758369.1); *Arabidopsis thaliana* (AY168881); *Sorghum bicolor* (XP_002463933.1); *Stevia rebaudiana* (ABB88836.2); and *Thermosynechococcus elongatus* BP-1 (NP_681832).

In fact, each of the iron-sulfur enzymes described above that were soluble in yeast cytosol were also found to be properly folded and functional as they each were able to catalyze their respective reactions in the presence of artificial electron donors. In addition, if an iron-sulfur protein was found to be insoluble, it was modified or placed in a suitable environment wherein the iron-sulfur protein could be expressed and functional. For example, the yeast mitochondrion is very similar in its environment to a bacterial cytoplasm. As a consequence, when *E. coli* IspG, which is insoluble in yeast cytosol, is targeted to and expressed in the yeast mitochondrion, it is properly folded and functional therein. (Suitable targeting sequences for targeting the DXP proteins to the mitochondrion can be identified using Mitroprot, available from the Institute of Human Genetics, Melmholtz Center, Munich Germany, see Claros, et al., Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* (1996) 241:779-786.) Curiously, all of the previous work relating to the expression of a heterologous DXP pathway, including IspG expression in a host cell without an endogenous DXP pathway, has been done on yeast using *E. coli* IspG.

It is also known that many endogenous DXP proteins in plant plastids have targeting sequences to the chloroplast. Plasmodium also has endogenous DXP pathway proteins that have targeting sequences to Apicoplast. In general, the successful heterologous expression of these proteins in a non-native host can be improved by removing these targeting sequences from these proteins before heterologous expression. The task of removing these targeting sequences from these DXP pathway proteins is well within the skill of one in the arts of molecular biology or bioengineering. Programs that can be used to identify such targeting sequences for identification and removal from these DXP proteins include: TargetP, available from the Center for Biological Sequence Analysis (Technical University of Denmark, DTU) (see, Emanuelsson, et al., Locating proteins in

the cell using TargetP, SignalP, and related tools, *Nature Protocols* (2007) 2:953-971) and PATS (Prediction of Apicoplast Targeted Sequences) made available by the University of Frankfurt, Germany. (*See*, Zuegge et al., Deciphering apicoplast targeting signals - feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene* (2001) 280:19-26.)

In addition to solubility in the desired compartment of the host cell (typically cytoplasm), the reactions mediated by a DXP iron-sulfur enzyme involve electron transfer and at least two accessory proteins that are required for the proper function of an iron-sulfur enzyme. The first accessory protein is a reductase that is capable of accepting one or more electrons from an electron donor. The second accessory protein is a redox protein that is a redox partner to the iron-sulfur enzyme. The redox partner essentially receives the one or more electrons from the reductase (which in turn received the electrons from an electron donor) and transfers the one or more electrons to the iron-sulfur cluster in the iron-sulfur enzyme. Ultimately, the electrons are transferred by the iron-sulfur enzyme to a product of the enzymatic reaction catalyzed by the iron-sulfur enzyme.

Transfer of these electrons amongst the proteins in the iron-sulfur protein and the accessory proteins includes acceptance of one or more electrons by a first receiving protein, which converts it from its oxidized state to its reduced state, and subsequent transfer of the same electrons by the first receiving protein to a second receiving protein, which converts the first receiving protein back to its oxidized state (from its reduced state). Redox balance is achieved within this electron transport system by regenerating the original source of the electrons (*e.g.,* by converting NAD^+ or $NADP^+$ to NADH or NADPH, respectively).

Another complexity in the expression of a heterologous DXP pathway in a host cell without an endogenous DXP pathway is that the two accessory proteins that are involved in each of the two iron-sulfur enzymatic reactions have not been conclusively identified. As a consequence, almost all of the work involving the DXP pathway has been in host cells that have an endogenous DXP pathway. Because host cells with an endogenous DXP pathway necessarily include all of the cofactors and proteins needed for a functional pathway, working in such host cells eliminates the need for also expressing the two accessory proteins for each of the two iron-sulfur enzymes. However, if the DXP pathway is to be a commercially viable option for isoprenoid production, it is almost certain that even in host cells with an endogenous DXP pathway, the respective accessory proteins need to be over-expressed so that neither of the two iron-sulfur enzymes becomes rate limiting to the DXP pathway.

In order to conclusively identify redox partner(s) compatible with IspG or IspH, the present inventors used the representative eukaryotic cell yeast for screening. The ferrodoxin and flavodoxin family of proteins and their cognate reductases were tested for their respective ability to partner with IspG and IspH. This family was selected because they are the most common redox proteins involved in electron transfer in host cells with an endogenous DXP pathway. Unfortunately, if the desired host cell is a eukaryotic cell such as yeast, neither a ferrodoxin nor a flavodoxin is known to be present in the yeast cytoplasm.

Notwithstanding this difficulty, the inventors of the present invention have identified the necessary accessory proteins that are compatible with IspG's and IspH's such that the respective enzyme can carry out its electron transfer reactions. The simplest iron-sulfur clusters have two iron ions bridged by two sulfide ions, and are designated "[2Fe-2S]" clusters. In general, [2Fe-2S]-dependent redox partners were not generally soluble when expressed in yeast. For example the following [2Fe-2S] ferredoxins were not soluble in yeast cytoplasm in their unmodified forms: *Arthrospira platensis* (P00246); *Arabidopsis thaliana* (AAG40057); *Arabidopsis thaliana* (NP_180320); *Bacillus coagulans* (EEN90783); *Chlamydomonas reihardtii* (XP_001692808); *Chlamydomonas reihardtii* (ABC88601); *Escherichia coli* str. K-12 substr. MG1655 (NP_417057.1); *Nostoc* strain MAC (P00252); *Nostoc* strain MAC (0812211B); *Peptococcus aerogenes* (P00193.1); *Plasmodium falciparum* 3D7 (Q8IED5); *Saccharomyces cerevisiae* (NP_015071); *Spinacia oleracea* (P00221.2); *Synechocystis* sp. 6714 (0812212A); and *Thermosynechococcus elongatus* BP-1 (NP_682446.1).

In contrast, the following flavodoxins were found to be both soluble and, when paired with a suitable reductase, capable of working with IspG: *Alteromonas macleodii* ATCC 27126 (ZP_04714638.1); *Bacillus coagulans* 36D1 (ZP_04431144.1); *Bacillus subtilis subsp. subtilis* str. 168 YkuP (NP_389300.2); *Crocosphaera watsonii* WH 8501 (ZP_00515759.1); *Desulfovibrio vulgaris subsp. vulgaris* DP (YP_966026.1); *Escherichia coli* str. K-12 substr. MG1655 (NP_415210); *Nostoc punctiforme* PCC 73102 (YP_001866434.1); *Anabaena* PCC 7119 (AAB20462); *Wolinella succinogenes* DSM 1740 (NP_906993.1); and *Bacillus subtilis subsp. subtilis* str. 168 YkuN (NP_389298.1).

In general, components of a fully functional IspH iron-sulfur enzyme system appeared more robust because they were more readily identified than those for IspG. Certain redox partners were found to be compatible with both IspG and IspH. Illustrative examples of two such redox partners include *Escherichia coli* str. K-12 substr. MG1655 (NP_415210) and *Bacillus subtilis subsp. subtilis str.* 168 YkuN (NP_389298.1). However, each IspG and IspH provided the best

activity if paired with a different redox partner, suggesting that the respective redox partner in the endogenous system is specific and distinct.

Because flavodoxins were identified as compatible redox partners for the two DXP iron-sulfur enzymes, the suitable reductases included ferrodoxin, and flavodoxin NADP+ reductases ("FNRs"). As with the other components of the iron-sulfur enzyme system, the crucial determinant of heterogeneous function was solubility of an FNR in yeast cytoplasm. In general, if an FNR is soluble then it is likely to be compatible with flavodoxins. Illustrative examples of suitable FNRs include, but are not limited to, *Anabaena* PCC 7119 (1QUE_A); *Arabidopsis thaliana* (root) (NP_849734.1); *Escherichia coli* str. K-12 substr. MG1655 (NP_418359); *Nostoc* sp. PCC 7906 (BAG69182); *Spinacia oleracea* (P00455); *Thermosynechococcus elongatus* BP-1 (NP_682001); and *Zea mays* (3LVB_A).

Thus, the present invention provides iron-sulfur enzyme systems comprising an iron-sulfur DXP enzyme and one or more accessory proteins that are compatible with the iron-sulfur enzyme. In one embodiment, the present invention provides iron-sulfur enzyme systems comprising an iron-sulfur DXP enzyme and the two accessory proteins that are compatible with the iron-sulfur enzyme, so that the enzyme is able to catalyze its enzymatic reaction. There is no requirement that the two accessory proteins are those that are the naturally-occurring enzymatic partners for each iron-sulfur enzyme in their endogenous host. The iron-sulfur enzyme system (iron-sulfur enzyme plus the two accessory proteins) of the present invention can be used for enhancing a DXP pathway-mediated isoprenoid production in a host cell with an endogenous DXP pathway, or for engineering a heterologous DXP pathway in a host cell lacking an endogenous DXP pathway, and using the heterologous DXP pathway for isoprenoid production.

Thus, in one aspect of the invention, an iron sulfur enzyme system is provided comprising one or more nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) an iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur

enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

In another aspect of the invention, an iron sulfur enzyme system is provided comprising one or more nucleic acid molecule comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) a iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting cMEPP to isopentenyl pyrophosphate ("IPP") and dimethylallyl pyrophosphate ("DMAPP").

In another aspect of the invention, a genetically modified microorganism is provided comprising one or more heterologous nucleic acid molecules comprising one or more nucleic acid sequences that encode:

(a) a first reductase that is capable of accepting electrons from an electron donor thereby converting the first reductase from an oxidized state ("oxidized first reductase") to a reduced state ("reduced first reductase");

(b) a first redox protein that is capable of accepting electrons from the reduced first reductase thereby converting the first redox protein from an oxidized state ("oxidized first redox protein") to a reduced state ("reduced first redox protein"); and

(c) a first iron-sulfur enzyme that is capable of accepting electrons from the first reduced redox protein thereby converting the first iron-sulfur enzyme from an oxidized state ("oxidized first iron-sulfur enzyme") to a reduced state ("reduced first iron-sulfur enzyme") and wherein the reduced first iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

In other embodiments, the one or more heterologous nucleic acid molecules further encode:

(g) a second reductase that is capable of accepting electrons from an electron donor thereby converting the second reductase from an oxidized state ("oxidized second reductase") to a reduced state ("reduced second reductase");

(h) a second redox protein that is capable of accepting electrons from the reduced second reductase thereby converting the second redox protein from an oxidized state ("oxidized second redox protein") to a reduced state ("reduced second redox protein"); and

(i) a second iron-sulfur enzyme that is capable of accepting electrons from the second reduced redox protein thereby converting the second iron-sulfur enzyme from an oxidized state ("oxidized second iron-sulfur enzyme") to a reduced state ("reduced second iron-sulfur enzyme") and wherein the reduced second iron-sulfur enzyme is capable of converting HMBPP to isopentenyl pyrophosphate ("IPP") or/and dimethylallyl pyrophosphate ("DMAPP").

In another aspect of the invention, a genetically modified microorganism is provided comprising one or more heterologous nucleic acid molecules wherein the nucleic acid molecules encode at least one of the following enzymes:

(a) an enzyme that converts pyruvate and glyceraldehydes 3-phosphate to 1-deoxy-D-xylulose-5-phosphate ("DXP");

(b) an enzyme that converts DXP to 1C-methyl-D-erythritol-4-phosphate ("MEP");

(c) an enzyme that converts MEP to 4-diphosphocytidyl-2C-methyl-D-erythritol ("CDP-ME");

(d) an enzyme that converts CDP-ME to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate ("CDP-MEP"); and

(e) an enzyme that converts CDP-MEP to 2C-methyl-D-erythritol 2,4-cyclodiphosphate ("cMEPP");

and the nucleic acid molecules further encode:

(f) a first reductase that is capable of accepting electrons from an electron donor thereby converting the first reductase from an oxidized state ("oxidized first reductase") to a reduced state ("reduced first reductase");

(g) a first redox protein that is capable of accepting electrons from the reduced first reductase thereby converting the first redox protein from an oxidized state ("oxidized first redox protein") to a reduced state ("reduced first redox protein"); and

(h) a first iron-sulfur enzyme that is capable of accepting electrons from the first reduced redox protein thereby converting the first iron-sulfur cluster enzyme from an oxidized state ("oxidized first iron-sulfur enzyme") to a reduced state ("reduced first iron-sulfur enzyme") and wherein the reduced first iron-sulfur cluster enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP");

(i) a second reductase that is capable of accepting electrons from an electron donor thereby converting the second reductase from an oxidized state ("oxidized second reductase") to a reduced state ("reduced second reductase");

(j) a second redox protein that is capable of accepting electrons from the reduced second reductase thereby converting the second redox protein from an oxidized state ("oxidized second redox protein") to a reduced state ("reduced second redox protein"); and

(k) a second iron-sulfur enzyme that is capable of accepting electrons from the second reduced redox protein thereby converting the second iron-sulfur enzyme from an oxidized state ("oxidized second iron-sulfur enzyme") to a reduced state ("reduced second iron-sulfur enzyme") and wherein the reduced second iron-sulfur enzyme is capable of converting HMBPP to isopentenyl pyrophosphate ("IPP") or/and dimethylallyl pyrophosphate ("DMAPP").

For all of the above aspects, in some embodiments, the first reductase and the second reductase are the same reductase. In still other embodiments, the first redox protein and the second redox protein are the same redox protein. In other embodiments, the first reductase and the second reductase are the same reductase and the first redox protein and the second redox protein are the same redox protein.

In some embodiments, the unmodified microorganism does not comprise an endogenous DXP pathway. In other embodiments, the genetically modified microorganism is a eukaryote. In other embodiments, the genetically modified microorganism is a fungus. In other embodiments, the genetically modified microorganism is yeast. In still other embodiments, the genetically modified microorganism is *S. cerevisiae*.

In another aspect of the invention, the genetically modified microorganism is provided comprising a heterologous DXP pathway that is capable of making IPP or DMAPP therefrom and wherein the unmodified microorganism does not comprise an endogenous DXP pathway.

In some embodiments, first iron-sulfur enzyme is not IspG from *E. coli*. In other embodiments, the first iron-sulfur enzyme is IspG from a *Bacillus*. In other embodiments, the first iron-sulfur enzyme is IspG from *B. coagulans*, *B. thuringiensis*, or *B. subtilis*. In still other embodiments, the first iron-sulfur enzyme is IspG from *Thermus*. In yet other embodiments, the first iron-sulfur enzyme is IspG from *T. thermophilus*.

In some embodiments, the second iron-sulfur enzyme is IspH from *E. coli*. In other embodiments, the second iron-sulfur enzyme is IspH from *F. succinogens*. In other embodiments, the second iron-sulfur enzyme is IspH from *P. patens*. In other embodiments, the second iron-sulfur enzyme is IspH from *S. bicolor*. In other embodiments, the second iron-sulfur enzyme is IspH from *S.*

rebaudiana. In other embodiments, the second iron-sulfur enzyme is IspH from *T. elongatus*. In still other embodiments, the second iron-sulfur enzyme is IspH from *A. thaliana*.

In some cases, the activity of the iron-sulfur enzyme is improved by creating a fusion protein between it and its redox protein. As a result, in some embodiments, the first redox protein and the first iron-sulfur enzyme form a contiguous polypeptide. In other embodiments, the first redox protein and the second iron-sulfur enzyme form a contiguous polypeptide. In still other embodiments, the second redox protein and the second iron-sulfur enzyme form a contiguous polypepetide.

In some embodiments, the electron donor is NADPH. In other embodiments, the reductase is a flavodxin/ferredoxin NADP+ reductase ("FNR"). In other embodiments the reductase is an FNR from *E. coli*. In other embodiments, the reductase is an FNR from *A. thaliana*.

In some embodiments, the first redox protein is a flavodoxin. In still other embodiments the first redox protein is a flavodoxin from *B. subtilis*. In yet still other embodiments, the first redox protein is YkuN from *B. subtilis*.

In some embodiments, the second redox protein is a flavodoxin. In other embodiments, the second redox protein is a flavodoxin from *E. coli*.

In some embodiments, the first reductase is an FNR. In other embodiments, the second reductase is an FNR. In still other embodiments, the first and second reductase is a FNR. In yet other embodiments, the first reductase and the second reductase are different FNRs.

Additional Engineering for Augmenting the DXP Pathway

Flux through the DXP pathway may also be augmented by increasing the efficiency or output of iron-sulfur cluster biogenesis, assembly and/or subsequent insertion of these iron-sulfur clusters into the respective apoprotein. As described previously, complex, cellular machinery is involved in these processes that includes more than 25 proteins that have been identified to date.

The ISC assembly machinery is part of this complex system. It is found in bacteria and in mitochondria and is involved in the biogenesis of all iron-sulfur proteins. Because of its importance in the biogenesis of all iron-sulfur proteins, the over-expression of any subset of its members increases the levels of iron-sulfur proteins generally, including the two DXP iron-sulfur enzymes.

In eukaroyotes, in addition to the ISC assembly machinery, the ISC export machinery and the cytosolic CIA machinery are required. The ISC assembly, found in the mitochondrion in eukaroyotes, comprises approximately 15 proteins. Illustrative examples of proteins that comprise the ISC assembly machinery include, but are not limited to, cysteine desulfurase Nfs1;

iron binding protein frataxin Yfh1; redox proteins Arh1 and Yah1; NADH kinase POS5; scaffold proteins Isu1 and Isu2; and chaperones Ss1-ATP, Jac1 and Grx5.

The mitochondrion also includes the ISC export machinery that exports key cellular components necessary for cytosolic iron-sulfur cluster assembly and subsequent insertion into the apoprotein, out of the mitochondrion. Illustrative examples of proteins that comprise the ISC export machinery include, but are not limited to, ABC transporter Atm1; the sulfhydryl oxidase Erv1, and glutathione. The cytosolic CIA machinery is involved in the maturation of the iron-sulfur apoprotein. Illustrative examples of proteins that comprise the CIA machinery include, but are not limited to, Cfd1-Nbp35; Nar1, Cia1 and Cia2; and Dre2. As with the ISC assembly machinery, the over-expression of any subset of members of either the ISC export machinery, the CIA machinery, or both, increases the levels of iron-sulfur proteins generally including the two DXP iron-sulfur enzymes.

General Metabolic Engineering

Other aspects of the invention include vectors comprising the heterologous nucleic acid sequences as well as methods for using genetically modified host cells of the present invention to make isoprenoid(s).

The heterologous nucleic acid sequences of the present invention can be expressed by a single or by multiple vectors. The nucleic acid sequences can be arranged in any order in a single operon, or in separate operons that are placed in one or multiple vectors. Where desired, two or more expression vectors can be employed, each of which contains one or more heterologous sequences operably linked in a single operon. While the choice of single or multiple vectors and the use of single or multiple promoters may depend on the size of the heterologous sequences and the capacity of the vectors, it will largely depend on the overall yield of a given isoprenoid that the vector is able to provide when expressed in a selected host cell. In some instances, two-operon expression systems provide a higher yield of isoprenoid. The subject vectors can stay replicable episomally, or as an integral part of the host cell genome.

In certain embodiments, the heterologous nucleic acids of the present invention are under the control of a single regulatory element. In some cases, the heterologous nucleic acid sequences are regulated by a single promoter. In other cases, the heterologous nucleic acid sequences are placed within a single operon. In still other cases, the heterologous nucleic acid sequences are placed within a single reading frame.

Where desired, the subject nucleic acid sequences can be modified to reflect the codon preference of a selected host cell to effect a higher expression of such sequences in a host cell. For example, the subject nucleotide sequences will in some embodiments be modified for yeast

codon preference. (See, *e.g.*, Bennetzen and Hall, *J. Biol. Chem.* (1982) 257(6): 3026-3031). As another non-limiting example, the nucleotide sequences will in other embodiments be modified for *E. coli* codon preference. (See, *e.g.*, Gouy and Gautier, Nucleic Acids Res. (1982) 10(22):7055-7074; Eyre-Walker, *Mol. Biol. Evol.* (1996) 13(6):864-872 (see also, Nakamura et al., *Nucleic Acids Res.* (2000) 28(1):292.) Codon usage tables for many organisms are available, which can be used as a reference in designing sequences of the present invention. The use of prevalent codons of a given host microorganism generally increases the likelihood of translation, and hence the expression level of the desired sequences.

Preparation of the subject nucleic acids can be carried out by a variety of routine recombinant techniques and synthetic procedures. Standard recombinant DNA and molecular cloning techniques used in the Examples of this disclosure are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) (Maniatis) and by T. J. Silhavy, M. L. Bennan, and L. W. Enquist, *Experiments with Gene Fusions, Cold Spring Harbor Laboratory*, Cold Spring Harbor, N.Y. (1984) and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987). Briefly, the subject nucleic acids may be genomic DNA fragments, cDNAs, and RNAs, all of which can be synthesized, purchased commercially, extracted directly from a cell or recombinantly produced by various amplification processes including, but not limited to, PCR and rt-PCR.

Direct chemical synthesis of nucleic acids typically involves sequential addition of 3'-blocked and 5'-blocked nucleotide monomers to the terminal 5'-hydroxyl group of a growing nucleotide polymer chain, wherein each addition is effected by nucleophilic attack of the terminal 5'-hydroxyl group of the growing chain on the 3'-position of the added monomer, which is typically a phosphorus derivative, such as a phosphotriester, phosphoramidite, or the like. Such methodology is known to those of ordinary skill in the art and is described in the pertinent texts and literature (see, for example, Matteuci et al. *Tet. Lett.* (1980) 521:719; U.S. Patent No. 4,500,707; 5,436,327; and 5,700,637).

The present invention relates in some embodiments to an increase in the level of transcription of nucleic acids in a host organism. The level of transcription of a nucleic acid in a host microorganism can be increased in a number of ways. For example, this can be achieved by increasing the copy number of the nucleotide sequence encoding the enzyme (*e.g.*, by using a higher copy number expression vector comprising a nucleotide sequence encoding the enzyme, or by introducing additional copies of a nucleotide sequence encoding the enzyme into the genome of the host microorganism, for example, by recA-mediated recombination, use of

"suicide" vectors, recombination using lambda phage recombinase, and/or insertion via a transposon or transposable element). In addition, it can be carried out by changing the order of the coding regions on the polycistronic mRNA of an operon or breaking up an operon into individual genes, each with its own control elements, or increasing the strength of the promoter (transcription initiation or transcription control sequence) to which the enzyme coding region is operably linked (for example, using a consensus arabinose- or lactose-inducible promoter in an *Escherichia coli* host microorganism in place of a modified lactose-inducible promoter, such as the one found in pBluescript and the pBBR1MCS plasmids), or using an inducible promoter and inducing the inducible-promoter by adding a chemical to the growth medium.

The level of translation of a nucleotide sequence in a host microorganism of the invention can be increased in a number of ways, including, but not limited to, increasing the stability of the mRNA, modifying the sequence of the ribosome binding site, modifying the distance or sequence between the ribosome binding site and the start codon of the enzyme coding sequence, modifying the entire intercistronic region located upstream of, or adjacent to the 5' side of the start codon of the enzyme coding region, stabilizing the 3'-end of the mRNA transcript using hairpins and specialized sequences, modifying the codon usage of enzyme, altering expression of rare codon tRNAs used in the biosynthesis of the enzyme, and/or increasing the stability of the enzyme, as, for example, via mutation of its coding sequence. Determination of preferred codons and rare codon tRNAs can be based on a sequence analysis of genes derived from the host microorganism.

The activity of a DXP pathway enzyme or prenyltransferase in a host can be altered in a number of ways, including, but not limited to, expressing a modified form of the enzyme that exhibits increased solubility in the host cell, expressing an altered form of the enzyme that lacks a domain through which the activity of the enzyme is inhibited, expressing a modified form of the enzyme that has a higher Kcat or a lower Km for the substrate, or expressing an altered form of the enzyme that is not affected by feed-back or feed-forward regulation by another molecule in the pathway. Such variant enzymes can also be isolated through random mutagenesis of a broader specificity enzyme, and a nucleotide sequence encoding such variant enzyme can be expressed from an expression vector or from a recombinant gene integrated into the genome of a host microorganism.

The subject vector can be constructed to yield a desired level of copy numbers of the encoded enzyme. In some embodiments, the subject vectors yield at least 10, between 10-20, between 20-50, between 50-100, or even higher than 100 copies of the desired enzymes. Low copy number plasmids generally provide fewer than about 20 plasmid copies per cell; medium copy

number plasmids generally provide from about 20 plasmid copies per cell to about 50 plasmid copies per cell, or from about 20 plasmid copies per cell to about 80 plasmid copies per cell; and high copy number plasmids generally provide from about 80 plasmid copies per cell to about 200 plasmid copies per cell, or more.

Suitable low copy expression vectors for *Escherichia coli* include, but are not limited to, pACYC184, pBeloBac11, pBR332, pBAD33, pBBR1MCS and its derivatives, pSC101, SuperCos (cosmid), and pWE15 (cosmid). Suitable medium copy expression vectors for *Escherichia coli* include, but are not limited to, pTrc99A, pBAD24, and vectors containing a ColE1 origin of replication and its derivatives. Suitable high copy number expression vectors for *Escherichia coli* include, but are not limited to, pUC, pBluescript, pGEM, and pTZ vectors. Suitable low-copy (centromeric) expression vectors for yeast include, but are not limited to, pRS415 and pRS416 (Sikorski & Hieter (1989) Genetics 122:19-27). Suitable high-copy 2 micron expression vectors in yeast include, but are not limited to, pRS425 and pRS426 (Christainson et al. *Gene* (1992) 110:119-122). Alternative 2 micron expression vectors include non-selectable variants of the 2 micron vector (Bruschi & Ludwig Curr. Genet. (1988) 15:83-90) or intact 2 micron plasmids bearing an expression cassette (as exemplified in U.S. Patent No. 7,419,801, issued September 2, 2008) or 2 micron plasmids bearing a defective selection marker such as LEU2d (Erhanrt et al., *J. Bacteriol*. (1983) 156 (2): 625-635) or URA3d (Okkels, *Annals of the New York Academy of Sciences* (1996) 782(1): 202-207).

Regulatory elements include, for example, promoters and operators, which can also be engineered to increase the metabolic flux of the DXP pathways by increasing the expression of one or more genes that play a significant role in determining the overall yield of an isoprenoid produced. A promoter is a sequence of nucleotides that initiates and controls the transcription of a nucleic acid sequence by an RNA polymerase enzyme. An operator is a sequence of nucleotides adjacent to the promoter that functions to control transcription of the desired nucleic acid sequence. The operator contains a protein-binding domain where a specific repressor protein can bind. In the absence of a suitable repressor protein, transcription initiates through the promoter. In the presence of a suitable repressor protein, the repressor protein binds to the operator and inhibits transcription from the promoter.

In some embodiments of the present invention, promoters used in expression vectors are inducible. In other embodiments, the promoters used in expression vectors are constitutive. In some embodiments, one or more nucleic acid sequences are operably linked to an inducible promoter, and one or more other nucleic acid sequences are operably linked to a constitutive promoter.

Non-limiting examples of suitable promoters for use in prokaryotic host cells include a bacteriophage T7 RNA polymerase promoter; a trp promoter; a lac operon promoter; a hybrid promoter (for example, a lac/tac hybrid promoter, a tac/trc hybrid promoter, a trp/lac promoter, a T7/lac promoter; a trc promoter; a tac promoter, and the like); an araBAD promoter; *in vivo* regulated promoters (such as an ssaG promoter or a related promoter; see, for example, U.S. Patent Publication No. 2004/0131637, published July 8, 2004), a pagC promoter (Pulkkinen and Miller, *J. Bacteriol*. (1991) 173(1):86-93; Alpuche-Aranda et al. *PNAS* U.S.A. (1992) 89(21):10079-83), a nirB promoter (Harborne et al. *Mol. Micro*. (1992) 6:2805-2813), and the like (see, for example, Dunstan et al. *Infect. Immun*. (1999) 67:5133-5141; McKelvie et al. *Vaccine* (2004) 22:3243-3255; and Chatfield et al. *Biotechnol*. (1992) 10:888-892); a sigma70 promoter (for example, a consensus sigma70 promoter (see, GenBank Accession Nos. AX798980, AX798961, and AX798183); a stationary phase promoter, for example, a dps promoter, an spv promoter, and the like; a promoter derived from the pathogenicity island SPI-2 (see, PCT publication No. WO 96/17951, published June 13, 1996); an actA promoter (see, Shetron-Rama et al. (2002) *Infect. Immun.* 70:1087-1096); an rpsM promoter (see, Valdivia and Falkow, *Mol. Microbiol*. (1996) 22:367 378); a tet promoter (see, Hillen et al. (1989) In Saenger W. and Heinemann U. (eds) Topics in Molecular and Structural Biology, Protein–Nucleic Acid Interaction. Macmillan, London, UK, Vol. 10, pp. 143–162); an SP6 promoter (see, Melton et al. *Nucl. Acids Res*. (1984) 12:7035-7056); and the like.

In some embodiments of the present invention, the total activity of heterologous enzyme(s) that play a larger role in the overall yield of isoprenoid(s), relative to other enzymes in the respective pathways, may be increased by expressing the enzyme(s) from a strong promoter. Suitable strong promoters for *Escherichia coli* include, but are not limited to, Trc, Tac, T5, T7, and PLambda. In another embodiment of the present invention, the total activity of the one or more MEV pathway enzymes in a host is increased by expressing the enzyme from a strong promoter on a high copy number plasmid. Suitable examples of such high copy number plasmids for Escherichia coli include, but are not limited to, Trc, Tac, T5, T7, and PLambda promoters used with pBAD24, pBAD18, pGEM, pBluescript, pUC, and pTZ vectors.

Non-limiting examples of suitable promoters for use in eukaryotic host cells include, but are not limited to, a CMV immediate early promoter, an HSV thymidine kinase promoter, an early or late SV40 promoter, LTRs from retroviruses, and a mouse metallothionein-I promoter.

Non-limiting examples of suitable constitutive promoters for use in prokaryotic host cells include a sigma70 promoter (for example, a consensus sigma70 promoter). Non-limiting examples of

suitable inducible promoters for use in bacterial host cells include the pL of bacteriophage λ; Plac; Ptrp; Ptac (Ptrp-lac hybrid promoter); an isopropyl-beta-D44 thiogalactopyranoside (IPTG)-inducible promoter (for example, a lacZ promoter); a tetracycline inducible promoter; an arabinose inducible promoter (for example, PBAD; see, Guzman et al. (1995) *J. Bacteriol.* 177:4121-4130); a xylose-inducible promoter (for example, Pxyl; see, Kim et al. (1996) *Gene* 181:71-76); a GAL1 promoter; a tryptophan promoter; a lac promoter; an alcohol-inducible promoter (for example, a methanol-inducible promoter or an ethanol-inducible promoter); a raffinose-inducible promoter; a heat-inducible promoter (for example, heat inducible lambda PL promoter); a promoter controlled by a heat-sensitive repressor (for example, CI857-repressed lambda-based expression vectors; see, Hoffmann et al. (1999) *FEMS Microbiol Lett*. 177(2):327-34); and the like.

Non-limiting examples of suitable constitutive promoters for use in yeast host cells include an ADH1, an ADH2, a PGK, or a LEU2 promoter. Non-limiting examples of suitable inducible promoters for use in yeast host cells include, but are not limited to, a divergent galactose-inducible promoter such as a GAL 1 or a GAL 10 promoter (West at al. (1984) Mol. Cell. Biol. 4(11):2467-2478), or a CUP1 promoter. Where desired, the subject vector may comprise a promoter that is stronger than a native *E. Coli* Lac promoter.

Non-limiting examples of operators for use in bacterial host cells include a lactose promoter operator (the LacI repressor protein changes conformation when contacted with lactose, thereby preventing the LacI repressor protein from binding to the operator), a tryptophan promoter operator (when complexed with tryptophan, the TrpR repressor protein has a conformation that binds the operator; in the absence of tryptophan, the TrpR repressor protein has a conformation that does not bind to the operator), and a tac promoter operator (see, for example, deBoer et al. (1983) *PNAS* U.S.A. 80:21-25). The genes in the expression vector typically also encode a ribosome binding site to direct translation (that is, synthesis) of any encoded mRNA gene product. For suitable ribosome binding sites for use in *Escherichia coli*, see Shine et al. (1975) *Nature* 254:34, and Steitz, in Biological Regulation and Development: Gene Expression (ed. R. F. Goldberger), vol. 1, p. 349, 1979, Plenum Publishing, N.Y.). Insertion of the ribosome binding site encoding nucleotide sequences such as 5'-AAAACA-3' upstream of a coding sequence facilitates efficient translation in a yeast host microorganism (Looman et al. (1993) *Nuc. Ac. Res.* 21:4268-4271; Yun et. al. (1996) *Mol. Microbiol*. 19:1225-1239). In some embodiments of the invention, optimized Shine-Dalgarno sequences are inserted upstream of the coding region of a gene that is placed within a DXP module. The optimized Shine-Dalgarno sequences ensure efficient ribosome binding and translation of the coding sequence.

Other regulatory elements that may be used in an expression vector include transcription enhancer elements and transcription terminators. See, for example, Bitter et al. (1987) *Methods in Enzymology*, 153:516-544.

An expression vector may be suitable for use in particular types of host microorganisms and not others. One of ordinary skill in the art, however, can readily determine through routine experimentation whether a particular expression vector is suited for a given host microorganism. For example, the expression vector can be introduced into the host organism, which is then monitored for viability and expression of any genes contained in the vector.

The expression vector may also contain one or more selectable marker genes that, upon expression, confer one or more phenotypic traits useful for selecting or otherwise identifying host cells that carry the expression vector. Non-limiting examples of suitable selectable markers for eukaryotic cells include dihydrofolate reductase and neomycin resistance. Non-limiting examples of suitable selectable markers for prokaryotic cells include tetracycline, ampicillin, chloramphenicol, carbenicillin, and kanamycin resistance.

For the production of isoprenoid(s) at an industrial scale, it may be impractical or too costly to use a selectable marker that requires the addition of an antibiotic to the fermentation media. Accordingly, some embodiments of the present invention employ host cells that do not require the use of an antibiotic resistance-conferring selectable marker to ensure plasmid (expression vector) maintenance. In these embodiments of the present invention, the expression vector contains a plasmid maintenance system such as the 60-kb IncP (RK2) plasmid, optionally together with the RK2 plasmid replication and/or segregation system, to effect plasmid retention in the absence of antibiotic selection (see, for example, Sia et al. (1995) *J. Bacteriol*. 177:2789-97; or Pansegrau et al. (1994) *J. Mol. Biol*. 239:623-63). A suitable plasmid maintenance system for this purpose is encoded by the parDE operon of RK2, which codes for a stable toxin and an unstable antitoxin. The antitoxin can inhibit the lethal action of the toxin by direct protein-protein interaction. Cells that lose the expression vector that harbors the parDE operon are quickly deprived of the unstable antitoxin, resulting in cell death caused by the stable toxin. The RK2 plasmid replication system is encoded by the trfA gene, which codes for a DNA replication protein. The RK2 plasmid segregation system is encoded by the parCBA operon, which codes for proteins that function to resolve plasmid multimers that may arise from DNA replication.

The subject vectors can be introduced into a host cell stably or transiently by a variety of established techniques. For example, one method involves a calcium chloride treatment wherein the expression vector is introduced via a calcium precipitate. Other salts, for example calcium phosphate, may also be used following a similar procedure. In addition, electroporation (that is,

the application of current to increase the permeability of cells to nucleic acids) may be used. Other transformation methods include microinjection, DEAE dextran mediated transformation, and heat shock in the presence of lithium acetate. Lipid complexes, liposomes, and dendrimers may also be employed to transfect the host microorganism.

Upon transformation, a variety of methods can be practiced to identify the host cells into which the subject vectors have been introduced. One exemplary selection method involves subculturing individual cells to form individual colonies, followed by testing for expression of the desired gene product. Another method entails selecting transformed host cells based upon phenotypic traits conferred through the expression of selectable marker genes contained within the expression vector. Those of ordinary skill can identify genetically-modified host cells using these or other methods available in the art.

In some embodiments, the heterlogous nucleic acid sequences can be inserted into the genome of the host organism. One method useful for the introduction of such sequences is the use of integration cassettes. Integration cassettes are typically linear, double-stranded DNA fragments that can be chromosomally integrated by homologous recombination via the use of two PCR-generated fragments or one PCR-generated fragment. The integration cassette comprises a nucleic acid integration fragment that contains an expressible DNA fragment and a selectable marker bounded by specific recombinase sites responsive to a site-specific recombinase, and homology arms having homology to different portions of the host cell's chromosome. (see, for example, U.S. Patent Publication No. 2004/0219629, published November 4, 2004). Generally, the preferred length of the homology arms is about 10 to about 100 base pairs in length. From 20 to 40 base pairs of homology, the efficiency of homologous recombination increases by four orders of magnitude (Yu et al. *PNAS* (2000) 97:5978-5983). One method of introducing DXP modules into the host genome utilizes the λ-Red recombinase system. The λ-Red system enables the use of homologous recombination as a tool for *in vivo* chromosomal engineering in hosts, such as *E. coli*, normally considered difficult to transform by homologous recombination. The λ-Red system works in other bacteria as well (Poteete, A., *supra*, 2001). Use of the λ-Red recombinase system can be applicable to other hosts generally used for industrial production.

The introduction of various pathway sequences of the invention into a host cell can be confirmed by methods such as PCR, Southern blot or Northern blot hybridization. For example, nucleic acids can be prepared from the resultant host cells, and the specific sequences of interest can be amplified by PCR using primers specific for the sequences of interest. The amplified product is subjected to agarose gel electrophoresis, polyacrylamide gel electrophoresis or capillary electrophoresis, followed by staining with ethidium bromide, SYBR Green solution or the like, or

detection of DNA with UV light. Alternatively, nucleic acid probes specific for the sequences of interest can be employed in a hybridization reaction. The expression of a specific gene sequence can be ascertained by detecting the corresponding mRNA via reverse-transcription coupled PCR, Northern blot hybridization, or by immunoassays using antibodies reactive with the encoded gene product. Exemplary immunoassays include, but are not limited to, ELISA, radioimmunoassays, and sandwich immunoassays.

The enzymatic activity of a given pathway enzyme can be assayed by a variety of methods known in the art. In general, the enzymatic activity can be ascertained by the formation of the product or conversion of a substrate of an enzymatic reaction that is under investigation. The reaction can take place *in vitro* or *in vivo*.

The yield of an isoprenoid via one or more metabolic pathways disclosed herein can be augmented by inhibiting reactions that divert intermediates away from productive metabolic reactions that direct the intermediates toward the formation of the desired isoprenoid product(s). Inhibition of such unproductive reactions can be achieved by reducing the expression and/or activity of enzymes involved in one or more unproductive reactions. Such reactions include side reactions of the TCA cycle that lead to fatty acid biosynthesis, alanine biosynthesis, the aspartate superpathway, gluconeogenesis, heme biosynthesis, and/or glutamate biosynthesis, at a level that affects the overall yield of isoprenoid production. Inhibition can be accomplished by reducing or eliminating the expression of certain genes in the target pathway, or in competing pathways, that may serve as competing sinks for energy or carbon. Where the sequence of the gene to be disrupted is known, one effective method of gene down-regulation is targeted gene disruption, where foreign DNA is inserted into a structural gene so as to disrupt transcription. This can be effected by the creation of genetic cassettes comprising the DNA to be inserted (often a genetic marker) flanked by sequence having a high degree of homology to a portion of the gene to be disrupted. Introduction of the cassette into the host cell results in insertion of the foreign DNA into the structural gene via the native DNA replication mechanisms of the cell or by the λ-Red recombination system. (See, for example, Hamilton et al., *J. Bacteriol.*, (1989) 171:4617-4622); Balbas et al., *Gene* (1993) 136:211-213); Gueldener et al., *Nucleic Acids Res*. (1996) 24:2519-2524); and Smith et al., *Methods Mol. Cell. Biol*. (1996) 5:270-277).

Antisense technology is another method of down-regulating genes where the sequence of the target gene is known. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. This construct is then introduced into the host cell and the antisense strand of RNA is produced. Antisense RNA inhibits gene expression by preventing the accumulation of mRNA that encodes

the protein of interest. A person of skill in the art will know that special considerations are associated with the use of antisense technologies in order to reduce expression of particular genes. For example, the proper level of expression of antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan.

Other less specific methodologies can also be used to down regulate undesired activity. For example, cells may be exposed to UV radiation and then screened for the desired phenotype. Mutagenesis with chemical agents can also be effective for generating mutants and commonly used substances include chemicals that affect non-replicating DNA such as HNO_2 and NH_2OH, as well as agents that affect replicating DNA such as acridine dyes, notable for causing frame-shift mutations. Specific methods for creating mutants using radiation or chemical agents are well documented in the art. See, for example, Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, Mass.; or Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, (1992) 36: 227).

Another non-specific method of gene disruption is the use of transposable elements or "transposons." Transposons are genetic elements that insert randomly into DNA but can be later retrieved on the basis of sequence to determine where the insertion has occurred. Both *in vivo* and *in vitro* transposition methods are known. Both methods involve the use of a transposable element in combination with a transposase enzyme. When the transposable element or transposon is contacted with a nucleic acid fragment in the presence of the transposase, the transposable element will randomly insert into the nucleic acid fragment. The technique is useful for random mutagenesis and for gene isolation, because the disrupted gene may be identified on the basis of the sequence of the transposable element. Kits for *in vitro* transposition are commercially available (see, for example, The Primer Island Transposition Kit, available from Perkin Elmer Applied Biosystems, Branchburg, N.J., based upon the yeast Ty1 element; The Genome Priming System, available from New England Biolabs, Beverly, Mass.; based upon the bacterial transposon Tn7; and the EZ::TN Transposon Insertion System, available from Epicentre Technologies, Madison, Wis., based upon the Tn5 bacterial transposable element). Transposon-mediated random insertion in the chromosome can be used for isolating mutants for any number of applications, including enhanced production of many desired products, including enzymes or other proteins, amino acids, or small organic molecules, including alcohols.

An example where the reduction of a side reaction can increase the levels of a desired isoprenoid compound is the elimination or reduction of squalene synthase activity in host cells with an endogenous mevalonate pathway, such as yeast. In such systems, erg9 mutants that have a reduced ability to convert FPP into squalene have been shown to make more FPP-derived

isoprenoid product (see *e.g.*, Karst and Lacroute, Molec. Gen. Genet., 154, 269-277 (1977); U.S. Patent No. 5,589,372, issued December 31, 1996). Where the erg9 gene is blocked in yeast, such erg9 mutants may need extraneous ergosterol or other sterols added to the medium for the cells to remain viable because yeast strains generally need ergosterol for cell membrane fluidity. The cells normally cannot utilize this additional sterol unless grown under anaerobic conditions. However, erg 9 mutants that take up exogenously-supplied sterols under aerobic conditions have been identified. These include those having genetic modifications in upc (uptake control mutation which allows cells to take up sterols under aerobic conditions); hem1 (the HEM1 gene encodes aminolevulinic acid synthase that is the first committed step to the heme biosynthetic pathway from FPP, and hem1 mutants are capable of taking up ergosterol under aerobic conditions following a disruption in the ergosterol biosynthetic pathway, provided the cultures are supplemented with unsaturated fatty acids); and overexpression of the SUT1 (sterol uptake gene, which can be used to allow for uptake of sterols under aerobic conditions; see, Bourot and Karst, *Gene* (1995) 165:97-102).

Host Cells

A wide variety of host cells can be used in the practice of the present invention. In one embodiment, the host cell is a genetically modified host microorganism in which nucleic acid molecules have been inserted, deleted or modified (*i.e.*, mutated; *e.g.*, by insertion, deletion, substitution, and/or inversion of nucleotides), to produce the desired isoprenoid compound or isoprenoid derivative. Illustrative examples of suitable host cells include any archae, bacterial, or eukaryotic cell. Examples of an archae cell include, but are not limited to, those belonging to the genera: *Aeropyrum, Archaeglobus, Halobacterium, Methanococcus, Methanobacterium, Pyrococcus, Sulfolobus,* and *Thermoplasma*. Illustrative examples of archae strains include, but are not limited to, *Aeropyrum pernix, Archaeoglobus fulgidus, Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Pyrococcus abyssi, Pyrococcus horikoshii, Thermoplasma acidophilum,* and *Thermoplasma volcanium*.

Examples of a bacterial cell include, but are not limited to, those belonging to the genera: *Agrobacterium, Alicyclobacillus, Anabaena, Anacystis, Arthrobacter, Azobacter, Bacillus, Brevibacterium, Chromatium, Clostridium, Corynebacterium, Enterobacter, Erwinia, Escherichia, Lactobacillus, Lactococcus, Mesorhizobium, Methylobacterium, Microbacterium, Phormidium, Pseudomonas, Rhodobacter, Rhodopseudomonas, Rhodospirillum, Rhodococcus, Salmonella, Scenedesmun, Serratia, Shigella, Staphlococcus, Strepromyces, Synnecoccus,* and *Zymomonas*.

Illustrative examples of bacterial strains include, but are not limited to, *Bacillus subtilis, Bacillus amyloliquefacines, Brevibacterium ammoniagenes, Brevibacterium immariophilum, Clostridium beigerinckii, Enterobacter sakazakii, Escherichia coli, Lactococcus lactis, Mesorhizobium loti, Pseudomonas aeruginosa, Pseudomonas mevalonii, Pseudomonas pudica, Rhodobacter capsulatus, Rhodobacter sphaeroides, Rhodospirillum rubrum, Salmonella enterica, Salmonella typhi, Salmonella typhimurium, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Staphylococcus aureus,* and the like.

In general, if a bacterial host cell is used, a non-pathogenic strain is preferred. Illustrative examples of non-pathogenic strains include, but are not limited to, *Bacillus subtilis, Escherichia coli, Lactibacillus acidophilus, Lactobacillus helveticus, Pseudomonas aeruginosa, Pseudomonas mevalonii, Pseudomonas pudita, Rhodobacter sphaeroides, Rodobacter capsulatus, Rhodospirillum rubrum*, and the like.

Illustrative examples of eukaryotic strains include, but are not limited to, *Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Candida albicans, Chrysosporium lucknowense, Fusarium graminearum, Fusarium venenatum, Kluyveromyces lactis, Neurospora crassa, Pichia angusta, Pichia finlandica, Pichia kodamae, Pichia membranaefaciens, Pichia methanolica, Pichia opuntiae, Pichia pastoris, Pichia pijperi, Pichia quercuum, Pichia salictaria, Pichia thermotolerans, Pichia trehalophila, Pichia stipitis, Streptomyces ambofaciens, Streptomyces aureofaciens, Streptomyces aureus, Saccaromyces bayanus, Saccaromyces boulardi, Saccharomyces cerevisiae, Streptomyces fungicidicus, Streptomyces griseochromogenes, Streptomyces griseus, Streptomyces lividans, Streptomyces olivogriseus, Streptomyces rameus, Streptomyces tanashiensis, Streptomyces vinaceus, Trichoderma reesei and Xanthophyllomyces dendrorhous* (formerly *Phaffia rhodozyma*).

In general, if a eukaryotic cell is used, a non-pathogenic strain is preferred. Illustrative examples of non-pathogenic strains include, but are not limited to, *Fusarium graminearum, Fusarium venenatum, Pichia pastoris, Saccaromyces boulardi*, and *Saccaromyces cerevisiae*.

Examples of eukaryotic cells include, but are not limited to, fungal cells. Examples of fungal cell include, but are not limited to, those belonging to the genera: *Aspergillus, Candida, Chrysosporium, Cryotococcus, Fusarium, Kluyveromyces, Neotyphodium, Neurospora, Penicillium, Pichia, Saccharomyces, Trichoderma and Xanthophyllomyces* (formerly *Phaffia*). In addition, certain strains have been designated by the Food and Drug Administration as "GRAS", or Generally Regarded As Safe. These strains include: *Bacillus subtilis, Lactibacillus acidophilus, Lactobacillus helveticus, and Saccharomyces cerevisiae*.

ISOPRENOIDS OF THE PRESENT INVENTION

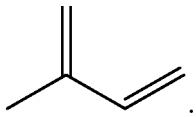
The compositions and methods of the present invention can be employed to produce a wide variety of isoprenoids, including, without limitation, any C5 through C20, or higher, carbon number isoprenoid. The following describes, without limitation, exemplary isoprenoids of the invention.

C5 Compounds

The C5 isoprenoid compounds are also known as hemiterpenes because they are derived from a single isoprene unit (IPP or DMAPP). C5 compounds of the present invention generally are derived from IPP or DMAPP.

Isoprene

Isoprene has the chemical structure:



Isoprene is found in many plants and is made from IPP by isoprene synthase. Illustrative examples of suitable nucleotide sequences encoding isoprene synthases include, but are not limited to, (AB198190; *Populus alba*) and (AJ294819; *Polulus alba x Polulus tremula*).

C10 Compounds

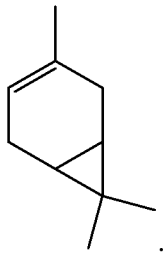
The C10 isoprenoid compounds are also known as monoterpenes because they are derived from two isoprene units. C10 compounds of the invention are generally derived from geranyl pyrophosphate (GPP), which is made by the condensation of IPP with DMAPP. An enzyme known to catalyze this condensation reaction is, for example, geranyl pyrophosphate synthase. In certain embodiments, the host cells of the present invention comprise heterologous nucleic acid sequences that encode an enzyme that converts IPP and DMAPP into GPP.

Illustrative examples of nucleotide sequences for geranyl pyrophosphate synthase 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).

GPP is then subsequently converted to a variety of C10 compounds. Illustrative examples of C10 compounds include but are not limited:

Carene

Carene has the chemical structure:

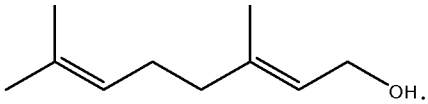


.

Careneis found in the resin of many trees, particularly pine trees. Carene is made from GPP from carene synthase. Illustrative examples of suitable nucleotide sequences encoding carene synthases include, but are not limited to, (AF461460, REGION 43..1926; *Picea abies*) and (AF527416, REGION: 78..1871; *Salvia stenophylla*).

Geraniol

Geraniol (also known as rhodnol) has the chemical structure:

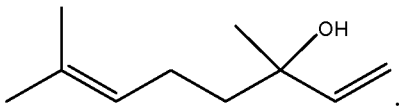


.

Geraniol is the main component of oil-of-rose and palmarosa oil. It also occurs in geranium, lemon, and citronella. Geraniol is made from GPP by geraniol synthase. Illustrative examples of suitable nucleotide sequences encoding geraniol synthases 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*).

Linalool

Linalool has the chemical structure:



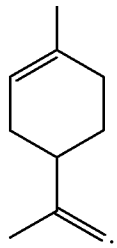
.

Linalool is found in many flowers and spice plants, such as coriander seeds. Linalool is made from GPP by linalool synthase. Illustrative examples of a suitable nucleotide sequence encoding linalool synthases 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).

Limonene

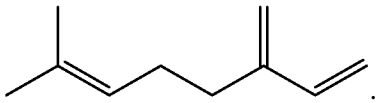
Limonene has the chemical structure:



Limonene is found in the rind of citrus fruits and peppermint. Limonene is made from GPP by limonene synthase. Illustrative examples of suitable nucleotide sequences encoding limonene synthases 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*).

Myrcene

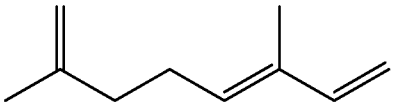
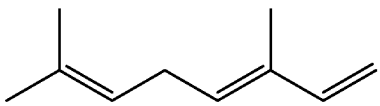
Myrcene has the chemical structure:



Myrcene is found in the essential oil in many plants including bay, verbena, and myrcia from which it gets its name. Myrcene is made from GPP by ocimene synthase. Illustrative examples of suitable nucleotide sequences encoding ocimene synthases 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*).

Ocimene

α- and β-Ocimene have the chemical structures:

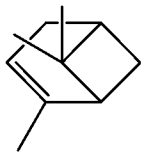
 and , respectively.

α- and β-Ocimene are found in a variety of plants and fruits, including *Ocimum basilicum* and are made from GPP by ocimene synthase. Illustrative examples of suitable nucleotide sequences encoding an ocimene synthase 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*).

α-Pinene

α-Pinene has the chemical structure:

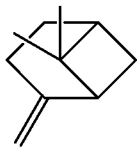


.

α-Pinene is found in pine trees and eucalyptus. α-Pinene is made from GPP by α-pinene synthase. Illustrative examples of suitable nucleotide sequences encoding α-pinene synthases 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*).

β-Pinene

β-Pinene has the chemical structure

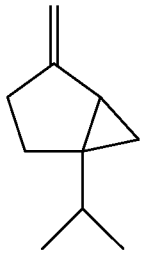


.

β-Pinene is found in pine trees, rosemary, parsley, dill, basil, and rose. β-Pinene is made from GPP by β-pinene synthase. Illustrative examples of suitable nucleotide sequences encoding β-pinene synthase include, but are not limited to, (-) β-pinene synthases (AF276072, REGION: 1..1749; *Artemisia annua*) and (AF514288, REGION: 26..1834; *Citrus limon*).

Sabinene

Sabinene has the chemical structure:

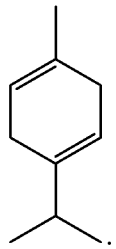


.

Sabinene is found in black pepper, carrot seed, sage, and tea trees. Sabinene is made from GPP by sabinene synthase. An illustrative example of a suitable nucleotide sequence encoding a sabinene synthased includes, but is not limited to, AF051901, REGION: 26..1798 from *Salvia officinalis*.

γ-Terpinene

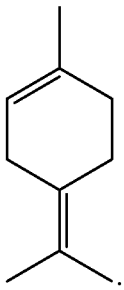
γ-Terpinene has the chemical structure:



γ-Terpinene is a constituent of essential oils from citrus fruits. Biochemically, γ-terpinene is made from GPP by a γ-terpinene synthase. Illustrative examples of suitable nucleotide sequences encoding γ-terpinene synthases include, but are not limited to, (AF514286, REGION: 30..1832 from *Citrus limon*) and (AB110640, REGION 1..1803 from *Citrus unshiu*).

Terpinolene

Terpinolene has the chemical structure:



Terpinolene is found in black currant, cypress, guava, lychee, papaya, pine, and tea. Terpinolene is made from GPP by terpinolene synthase. Illustrative examples of suitable nucleotide sequences encoding terpinolene synthases include, but are not limited to, (AY693650 from *Oscimum basilicum*) and (AY906866, REGION: 10..1887 from *Pseudotsuga menziesii*).

C15 Compounds

The C15 isoprenoid compounds are also known as sesquiterpenes because they are derived from three isoprene units. C15 compounds of the present invention generally derive from farnesyl pyrophosphate (FPP), which is made by the condensation of two molecules of IPP with one molecule of DMAPP. An enzyme known to catalyze this condensation reaction is, for example, farnesyl pyrophosphate synthase. In certain embodiments, the host cells of the present invention comprise a heterologous nucleic acid sequence that encodes an enzyme that converts IPP and DMAPP into FPP. Illustrative examples of nucleotide sequences for farnesyl pyrophosphate synthase and the ispA gene which encodes it are described above.

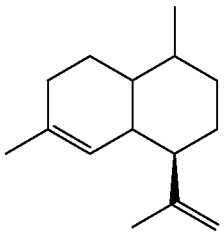
Alternatively, FPP can also be made by adding IPP to GPP. Illustrative examples of nucleotide sequences encoding an enzyme that catalyzes this reaction include, but are not limited to: (AE000657, Locus AAC06913; *Aquifex aeolicus* VF5), (NM_202836; *Arabidopsis thaliana*),

(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; Lept*ospira 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).

FPP is then subsequently converted to a variety of C15 compounds. Illustrative examples of C15 compounds include but are not limited to:

Amorphadiene

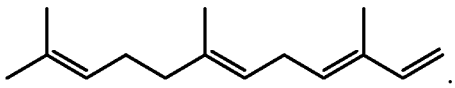
Amorphadiene has the chemical structure:



Amorphadiene is a precursor to artemisinin, which is made by *Artemisia anna*. Amorphadiene is made from FPP by amorphadiene synthase. An illustrative example of a suitable nucleotide sequence encoding amorphadiene synthase is SEQ ID NO. 37 of U.S. Patent No. 7,192,751, issued March 20, 2007.

α-Farnesene

α-Farnesene has the chemical structure:

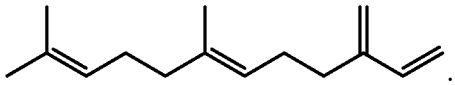


.

α-Farnesene is found in various biological sources including, but not limited to, the Dufour's gland in ants and in the coating of apple and pear peels. α-Farnesene is made from FPP by α-farnesene synthase. Illustrative examples of suitable nucleotide sequences encoding α-Farnesene synthases 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). (see, Pechouus et al., *Planta* (2004) 219(1):84-94).

β-Farnesene

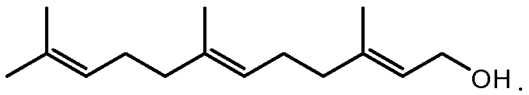
β-Farnesene has the chemical structure:



β-Farnesene is found in various biological sources including, but not limited to, aphids and essential oils, such as peppermint. In some plants, such as wild potato, β-farnesene is synthesized as a natural insect repellent. β-Farnesene is made from FPP by β-farnesene synthase. Illustrative examples of suitable nucleotide sequences encoding β-farnesene synthase include, but are not limited to, GenBank accession number AF024615 from *Mentha x piperita* (peppermint; gene Tspa11), and AY835398 from *Artemisia annua*. (see, Picaud et al., *Phytochemistry* (2005) 66(9):961-967).

Farnesol

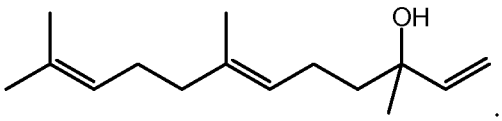
Farnesol has the chemical structure



Farnesol is found in various biological sources including insects and essential oils such as from cintronella, neroli, cyclamen, lemon grass, tuberose, and rose. Farnesol is made from FPP by a hydroxylase such as farnesol synthase. Illustrative examples of suitable nucleotide sequences encoding farnesol synthases include, but are not limited to, GenBank accession number AF529266 from *Zea mays* and YDR481C from *Saccharomyces cerevisiae* (gene Pho8). (see, Song, L., *Applied Biochemistry and Biotechnology*, (2006) 128:149-158).

Nerolidol

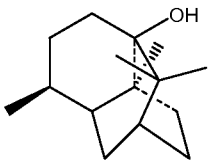
Nerolidol has the chemical structure:



Nerolidol, also known as peruviol, is found in various biological sources as essential oils, such as from neroli, ginger, jasmine, lavender, tea tree, and lemon grass. Nerolidol is made from FPP by a hydroxylase such as nerolidol synthase. An illustrative example of a suitable nucleotide sequence encoding a nerolidol synthase includes, but is not limited to, AF529266 from *Zea mays* (maize; gene tps1).

Patchoulol

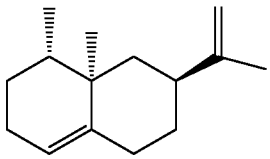
Patchoulol has the chemical structure:



Patchoulol, also known as patchouli alcohol, is a constituent of the essential oil of *Pogostemon patchouli*. Patchoulol is made from FPP by patchoulol synthase. An illustrative example of a suitable nucleotide sequence encoding a patchoulol synthase includes, but is not limited to, AY508730 REGION: 1..1659 from *Pogostemon cablin*.

Valenecene

Valencene has the chemical structure:



.

Valenecene is one of the main chemical components of the smell and flavor of oranges and is found in orange peels. Valencene is made from FPP by nootkatone synthase. Illustrative examples of a suitable nucleotide sequence encoding a nootkatone synthase includes, but are not limited to, AF441124 REGION: 1..1647 from *Citrus sinensis* and AY917195 REGION: 1..1653 from *Perilla frutescens*.

C20 Compounds

The C20 isoprenoid compounds are known as diterpenes because they are derived from four isoprene units. The C20 compounds of the invention are generally derived from geranylgeraniol pyrophosphate (GGPP), which is made by the condensation of three molecules of IPP with one molecule of DMAPP. In certain embodiments, the host cells of the present invention comprise a heterologous nucleic acid sequence that encodes an enzyme that converts IPP and DMAPP into GGPP. An enzyme known to catalyze this step is, for example, geranylgeranyl pyrophosphate synthase.

Illustrative examples of nucleotide sequences encoding geranylgeranyl pyrophosphate synthases 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*

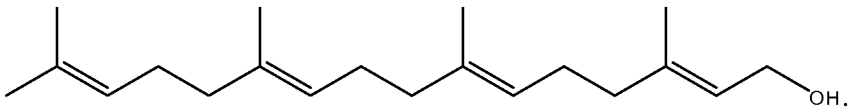
brasiliensis), (AB017971; *Homo sapiens*), (MCI276129; *Mucor circinelloides f. lusitanicus*), (AB016044; *Mus musculus*), (AABX01000298, Locus NCU01427; *Neurospora crassa*), (NCU20940; *Neurospora crassa*), (NZ_AAKL01000008, Locus ZP_00943566; *Ralstonia solanacearum* UW551), (AB118238; *Rattus norvegicus*), (SCU31632; *Saccharomyces cerevisiae*), (AB016095; *Synechococcus elongates*), (SAGGPS; *Sinapis alba*), (SSOGDS; *Sulfolobus acidocaldarius*), (NC_007759, Locus YP_461832; *Syntrophus aciditrophicus* SB), and (NC_006840, Locus YP_204095; *Vibrio fischeri* ES114).

Alternatively, GGPP can also be made by adding IPP to FPP. Illustrative examples of nucleotide sequences encoding an enzyme capable of this reaction include but are not limited to: (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).

GGPP may subsequently be converted to a variety of C20 isoprenoids. Illustrative examples of C20 compounds include but are not limited to:

Geranylgeraniol

Geranylgeraniol has the chemical structure:

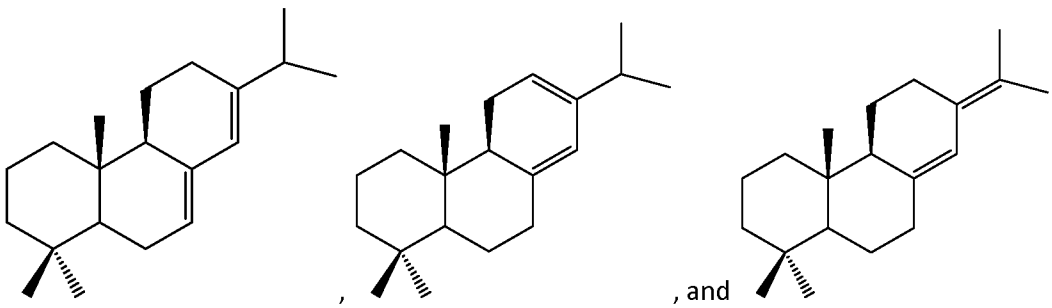


Geranylgeraniol is a constituent of wood oil from *Cedrela toona* and of linseed oil.

Geranylgeraniol can be made by *e.g.*, adding to the expression constructs of the present invention a phosphatase gene, after the gene for GGPP synthase.

Abietadiene

Abietadiene encompasses the following isomers:



and is found in trees such as *Abies grandis*. Abietadiene is made by abietadiene synthase. Illustrative examples of suitable nucleotide sequences encoding an abietadiene synthase include, but are not limited to: (U50768; *Abies grandis*) and (AY473621; *Picea abies*).

C20+ Compounds

C20+ compounds are also within the scope of the present invention. Illustrative examples of such compounds include sesterterpenes (C25 compounds made from five isoprene units), triterpenes (C30 compounds made from six isoprene units), and tetraterpenes (C40 compounds made from eight isoprene units). These compounds are made using methods described herein, while substituting or adding nucleotide sequences for the appropriate synthase(s).

Although the invention has been described in conjunction with specific embodiments thereof, the foregoing description is intended to illustrate and not limit the scope of the invention. Other aspects, advantages, and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

EXAMPLES

Unless otherwise indicated, the practice of the present invention can employ conventional techniques known to those skilled in the arts of the biosynthetic industry and the like. To the extent such techniques are not described fully herein, one can find ample reference to them in the scientific literature.

In the following examples, efforts have been made to ensure accuracy with respect to numbers used (for example, amounts, temperatures, etc.), but variations and deviations can be accommodated, and in the event a clerical error in the numbers reported herein exists, one of ordinary skill in the arts to which this invention pertains can deduce the correct amount in view of the remaining disclosure herein. Unless otherwise indicated, temperature is reported in degrees Celsius, and pressure is at or near atmospheric pressure at sea level. All reagents, unless otherwise indicated, were obtained commercially. The following examples are intended for illustrative purposes only and do not limit in any way the scope of the present invention.

Example 1

This example describes methods for making DNA constructs useful in the generation of host cells comprising heterologous DXP pathway enzymes.

The DNA constructs were generated from PCR-amplified DNA fragments. Table 1 lists the DNA fragments, and the templates and primers used to generate them by PCR amplification. PCR amplifications were done using the PHUSION™ DNA polymerase (New England Biolabs, Ipswich, MA) per the manufacturer's suggested protocol. The PCR reactions were resolved by gel electrophoresis, the DNA fragments were gel-purified using the gel extraction kit from Omega Bio-tek, Inc., (Norcross, GA), and the purified DNA fragments were treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, MA) per the manufacturer's suggested protocol. The PNK was heat-inactivated at 65°C for 20 minutes, and the samples were stored at -20°C.

Table 1 – PCR Amplified DNA Fragments		
DNA Fragment	Primers	Template
ERG13 US	GW392 (SEQ ID NO: 1) GW393 (SEQ ID NO: 2)	Y002 genomic DNA
ERG13 DS	GW394 (SEQ ID NO: 3) GW395 (SEQ ID NO: 4)	Y002 genomic DNA
HO US	2604 0 link-fwd (SEQ ID NO: 5) 2604 2 link-rev (SEQ ID NO: 6)	Y002 genomic DNA
HO DS	2605 0 link-fwd (SEQ ID NO: 7) 2605 2 link-rev (SEQ ID NO: 8)	Y002 genomic DNA
ADE2 US	R3519-fwd (SEQ ID NO: 9) R3519-rev (SEQ ID NO: 10)	Y002 genomic DNA
ADE2 DS	GW522 (SEQ ID NO: 11) GW523 (SEQ ID NO: 12)	Y002 genomic DNA
TRP1 ½	R2862-fwd (SEQ ID NO: 13) R2862-rev (SEQ ID NO: 14)	Y002 genomic DNA
TRP2 2/2-1	GW526 (SEQ ID NO: 15) GW525 (SEQ ID NO: 16)	Y002 genomic DNA
TRP2 2/2-2	GW524 (SEQ ID NO: 60) GW525 (SEQ ID NO: 16)	Y002 genomic DNA
GAL80 US	JWL7 (SEQ ID NO: 17) JWL8 (SEQ ID NO: 18)	Y002 genomic DNA
GAL80 DS	R3209-fwd (SEQ ID NO: 19) R3209-rev (SEQ ID NO: 20)	Y002 genomic DNA
ADH5 US	R4389-fwd (SEQ ID NO: 21) R4389-rev (SEQ ID NO: 22)	Y002 genomic DNA
ADH5 DS	R3202-fwd (SEQ ID NO: 23) R3202-rev (SEQ ID NO: 24)	Y002 genomic DNA
LEU2 US	M187-fwd (SEQ ID NO: 25) M187-rev (SEQ ID NO: 26)	Y002 genomic DNA
LEU2 DS	M3818-F (SEQ ID NO: 27) M3818-R (SEQ ID NO: 28)	Y002 genomic DNA
LEU2	GW396 (SEQ ID NO: 29) GW397 (SEQ ID NO: 30)	Y002 genomic DNA
HIS3	M2850-fwd (SEQ ID NO: 31) M2850-rev (SEQ ID NO: 32)	Y002 genomic DNA
URA3 ½	M2865 5 link-fwd (SEQ ID NO: 33) M2865 9 link-rev (SEQ ID NO: 34)	Y002 genomic DNA
URA3 2/2	2077 4 link-fwd (SEQ ID NO: 35) 2077 9 link-rev (SEQ ID NO: 36)	Y002 genomic DNA
PGAL1/10-1	R2246-fwd (SEQ ID NO: 37) R2246-rev (SEQ ID NO: 38)	Y002 genomic DNA
PTDH3-1	M2820-PTDH3-fwd (SEQ ID NO: 39) M2820-PTDH3-rev (SEQ ID NO: 40)	Y002 genomic DNA
PTDH3-2	M2823-PTDH3-fwd (SEQ ID NO: 41) M2823-PTDH3-rev (SEQ ID NO: 42)	Y002 genomic DNA
PTDH3-3	M3819 Prom-F (SEQ ID NO: 43) M3819 Prom-R (SEQ ID NO: 44)	Y002 genomic DNA
PTEF2-1	M2821-PTEF2-fwd (SEQ ID NO: 45) M2821-PTEF2-rev (SEQ ID NO: 46)	Y002 genomic DNA
PTEF2-2	M2824-PTEF2-fwd (SEQ ID NO: 47) M2824-PTEF2-rev (SEQ ID NO: 48)	Y002 genomic DNA
PTEF2-3	R3419-prom-fwd (SEQ ID NO: 65) R3419-prom-rev (SEQ ID NO: 66)	Y002 genomic DNA

PTEF2-4	GW527 (SEQ ID NO: 67) R5593-prom-rev (SEQ ID NO: 68)	Y002 genomic DNA
PTEF2-5	GW527 (SEQ ID NO: 67) GW546 (SEQ ID NO: 87)	Y002 genomic DNA
PTEF2-6	GW550 (SEQ ID NO: 86) GW546 (SEQ ID NO: 87)	Y002 genomic DNA
PTDH2	M2822-PTDH2-fwd (SEQ ID NO: 49) M2822-PTDH2-rev (SEQ ID NO: 50)	Y002 genomic DNA
Ec_dxs-TSHM2	M2820-dxs-1-fwd (SEQ ID NO: 51) M2820-dxs-2-rev (SEQ ID NO: 52)	SEQ ID NO: 97 [a)]
Ec_dxr-TADO1-1	M2821-dxr-fwd (SEQ ID NO: 83) M2821-dxr-rev (SEQ ID NO: 82)	SEQ ID NO: 98 [a)]
Ec_dxr-TADO1-2	GW577 (SEQ ID NO: 93) GW578 (SEQ ID NO: 94)	SEQ ID NO: 99 [a)]
Ec_ispD-TYSB1	M2822-ispD-fwd (SEQ ID NO: 53) M2822-ispD-rev (SEQ ID NO: 54)	SEQ ID NO: 99 [a)]
hisG	GW563 (SEQ ID NO: 55) GW564 (SEQ ID NO: 56)	SEQ ID NO: 109
Ec_ispE-THOM2	M2823-ispE-fwd (SEQ ID NO: 57) M2823-ispE-rev (SEQ ID NO: 58)	SEQ ID NO: 100 [a)]
Ec_ispF-TADE17	M2824-ispF-fwd (SEQ ID NO: 85) M2824-ispF-rev (SEQ ID NO: 84)	SEQ ID NO: 101 [a)]
Ec_ispF-THOM2	GW565 (SEQ ID NO: 136) GW566 (SEQ ID NO: 139)	SEQ ID NO: 101 [a)]
Ec_xylB-TPUP2	M3819 Gene-F (SEQ ID NO: 61) M3819 Gene-R (SEQ ID NO: 62)	SEQ ID NO: 102 [a)]
PENO2	GW556 (SEQ ID NO: 63) GW574 (SEQ ID NO: 64)	Y002 genomic DNA
BsyN_fld-TSUI3	GW573 (SEQ ID NO: 69) GW544 (SEQ ID NO: 70)	SEQ ID NO: 104 [a)]
At_RFNR2trunc-TOLE1-1	R3419-ORF-fwd (SEQ ID NO: 71) R3419-ORF-rev (SEQ ID NO: 72)	SEQ ID NO: 103 [a)]
At_RFNR2trunc-TOLE1-2	R6121-fwd (SEQ ID NO: 95) R6121-rev (SEQ ID NO: 96)	SEQ ID NO: 103 [a)]
Ec_ispH-TTHS1	R5593-ORF-fwd (SEQ ID NO: 73) GW528 (SEQ ID NO: 74)	SEQ ID NO: 106 [a)]
STREP-Ec_ispH-TTHS1	GW545 (SEQ ID NO: 75) GW528 (SEQ ID NO: 74)	SEQ ID NO: 201 [a)]
STREP-Te_ispH-TOLA1	GW545 (SEQ ID NO: 75) GW529 (SEQ ID NO: 76)	SEQ ID NO: 105
STREP-At_RFNR2trunc-TOLE1	GW545 (SEQ ID NO: 75) GW551 (SEQ ID NO: 77)	SEQ ID NO: 107
STREP-Ec_fldA-Ec_ispH	GW594 (SEQ ID NO: 78) GW595 (SEQ ID NO: 79)	SEQ ID NO: 108
STREP-Bt_ispG-BsyN_fld	GW596 (SEQ ID NO: 80) GW597 (SEQ ID NO: 81)	SEQ ID NO: 59
a) The synthetically generated DNA fragments were provided by Integrated DNA Technologies, Inc., (Coraville, IA).		

The DNA fragments were subsequently assembled into a first set of "stitches" as outlined in Table 2, using methods described in U.S. Patent No. 8,110,360. Additional DNA fragments from Table 1 and stitches from Table 2 were assembled into a second set of stitches as outlined in Table 3. For assembly reactions, 100 fmole of each DNA fragment or stitch were placed together in one tube.

The samples were split into three, 30 µL reactions; water, buffer, dNTPs, and DNA polymerase were added to each reaction mixture, and a first round of PCR-amplification was initiated. 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. The three PCR reaction mixtures were combined in one tube, the reaction mixtures were resolved by gel electrophoresis, and the PCR products were gel purified.

Table 2 - Terminal Primers Used for Assembly of First Set of Stitches

Stitch	DNA Fragments (see Table 1)	Terminal Primer 1	Terminal Primer 2
PTDH3-Ec_dxs-TSHM2	PTDH3-1, Ec_dxs-TSHM2	M2820-PTDH3-fwd (SEQ ID NO: 39)	M2820-dxs-2-rev (SEQ ID NO: 52)
PTEF2-Ec_dxr-TADO1	PTEF2-1, Ec_dxr-TADO1-1	M2821-PTEF2-fwd (SEQ ID NO: 45)	M2821-dxr-rev (SEQ ID NO: 82)
PTDH2-Ec_ispD-TYSB1	PTDH2, Ec_ispD-TYSB1	M2822-PTDH2-fwd (SEQ ID NO: 49)	M2822-ispD-rev (SEQ ID NO: 54)
PTEF2-Ec_ispF-TADE17	PTEF2-2, Ec_ispF-TADE17	M2824-PTEF2-fwd (SEQ ID NO: 47)	M2824-ispF-rev (SEQ ID NO: 84)
PTDH3-Ec_ispE-THOM2	PTDH3-2, Ec_ispE-THOM2	M2823-PTDH3-fwd (SEQ ID NO: 41)	M2823-ispE-rev (SEQ ID NO: 58)
PTDH3-Ec_xylB-TPUP2	PTDH3-3, Ec_xylB-TPUP2	M3819 Prom-F (SEQ ID NO: 43)	M3819 Gene-R (SEQ ID NO: 62)
PENO2-BsyN_fld-TSUI3	PENO2, BsyN_fld-TSUI3	GW556 (SEQ ID NO: 63)	GW544 (SEQ ID NO: 70)
PTEF2-At_RFNR2trunc-TOLE1	PTEF2-3, At_RFNR2trunc-TOLE1-1	R3419-prom-fwd (SEQ ID NO: 65)	R3419-ORF-rev (SEQ ID NO: 72)
PTEF2-Ec_ispH-TTHS1	PTEF2-4, Ec_ispH-TTHS1	GW527 (SEQ ID NO: 67)	GW528 (SEQ ID NO: 74)
PTEF2-STREP-Ec_ispH-TTHS1	PTEF2-5, Strep-Ec_ispH-TTHS1	GW527 (SEQ ID NO: 67)	GW528 (SEQ ID NO: 74)
PTEF2-STREP-Te_ispH-TOLA1	PTEF2-5, Strep-Te_ispH-TOLA1	GW527 (SEQ ID NO: 67)	GW529 (SEQ ID NO: 76)
PTEF2-STREP-At_RFNR2trunc-TOLE1	PTEF2-6, Strep-At_RFNR2trunc-TOLE1	GW550 (SEQ ID NO: 86)	GW551 (SEQ ID NO: 77)

The first round of PCR amplification was performed as follows: one cycle of denature at 98°C for 2 minutes; 5 cycles of denature at 98°C for 30 seconds and anneal/extend at 72°C for 30 seconds per kilobase PCR product. The second round of PCR amplification was performed as follows: one cycle of denature at 98°C for 2 minutes; 35 rounds of denature at 98°C for 12 seconds and anneal/extend at 72°C for 20-25 seconds per kilobase PCR product; one cycle of final extend at 72°C for 7 minutes; and a final hold at 4°C. When the annealing temperature was not 72°C (*i.e.*, when it was either 54°C or 65°C), in the first round of PCR amplification a 1 minute annealing step followed by a 30 seconds per kilobase PCR product extension step at 72°C was used, and for the second round of PCR amplification a 15 second annealing step followed by a 20 seconds per kilobase PCR product extension step at 72°C was used.

Table 3 - Terminal Primers Used for Assembly of Second Set of Stitches

Stitch	DNA Fragments (see Table 1) or Stitches (see Table 2)	Terminal Primer 1	Terminal Primer 2
i3850	Ec_ispD-TYSB1, hisG, Ec_ispF-THOM2	M2822-ispD-fwd (SEQ ID NO: 53)	GW566 (SEQ ID NO: 139)
i3830	ERG13 US, LEU2, ERG13 DS, linearized pAM2041 [a)]	terminal end of linearized pAM2041	terminal end of linearized pAM2041
s2075	HO US, PTDH3-Ec_dxs-TSHM2, PTEF2-Ec_dxr-TADO1, PTDH2-Ec_ispD-TYSB1, URA3 1/2, linearized pAM2041 [a)]		
s2076	HO DS, PTDH3-Ec_ispE-THOM2, PTEF2-Ec_ispF-TADE17, URA3 2/2, linearized pAM2041 [a)]		
i2145	LEU2 US, HIS3, PTDH3-Ec_xylB-TPUP2, PTEF2-Ec_dxr-TADO1, LEU2 DS, linearized pAM2041 [a)]		
s11469	ADE2 US, PENO2-BsyN_fld-TSUI3, PTEF2-At_RFNR2trunc-TOLE1, TRP1 1/2, linearized pAM2041 [a)]		
s11472	ADE2 US, PENO2-BsyN_fld-TSUI3, PTEF2-STREP-At_RFNR2trunc-TOLE1, TRP1 1/2, linearized pAM2041 [a)]		
s11475	ADE2 DS, PTEF2-Ec_ispH-TTHS1, TRP1 2/2, linearized pAM2041 [a)]		
s11478	ADE2 DS, TRP1 2/2-2, linearized pAM2041 [a)]		
s11479	ADE2 DS, PTEF2-STREP-Ec_ispH-TTHS1, TRP1 2/2-1, linearized pAM2041 [a)]		
s11480	ADE2 DS, PTEF2-STREP-Te_ispH-TOLA1, TRP1 2/2-1, linearized pAM2041 [a)]		
s11629	GAL80 US, STREP-Bt_ispG-BsyN_fld, pGAL1/10-1, At_RFNR2trunc-TOLE1-2, TRP1 1/2, linearized pAM2041 [a)]		
s11632	ADH5 US, STREP-Bt_ispG-BsyN_fld, pGAL1/10-1, At_RFNR2trunc-TOLE1-2, TRP1 1/2, linearized pAM2041 [a)]		
s11634	GAL80 DS, STREP-Ec_fldA-Ec_ispH, pGAL1/10-1, Ec_dxr-TADO1-2, TRP1 2/2, linearized pAM2041 [a)]		
s11635	ADH5 DS, STREP-Ec_fldA-Ec_ispH, pGAL1/10-1, Ec_dxr-TADO1-2, TRP1 2/2, linearized pAM2041 [a)]		

The first round of PCR amplification was performed as follows: one cycle of denature at 98°C for 2 minutes; 5 cycles of denature at 98°C for 30 seconds and anneal/extend at 72°C for 30 seconds per kilobase PCR product. The second round of PCR amplification was performed as follows: one cycle of denature at 98°C for 2 minutes; 35 rounds of denature at 98°C for 12 seconds and anneal/extend at 72°C for 20-25 seconds per kilobase PCR product; one cycle of final extend at 72°C for 7 minutes; and a final hold at 4°C. When the annealing temperature was not 72°C (*i.e.*, when it was either 54°C or 65°C), in the first round of PCR amplification a 1 minute annealing step followed by a 30 seconds per kilobase PCR product extension step at 72°C was used, and for the second round of PCR amplification a 15 second annealing step followed by a 20 seconds per kilobase PCR product extension step at 72°C was used.

a) Plasmid pAM2041 (SEQ ID NO: 92) was digested to completion using SchI restriction endonuclease (Fementas Life Sciences, Glen Burnie, Maryland), and the purified, linearized plasmid was included in the PCR reaction.

Example 2

This example describes methods for making yeast strains comprising heterologous DXP enzymes. The yeast strains listed in Table 4 were generated by introducing the indicated genetic modifications into the indicated parental strains. To these ends, exponentially-growing parental strain cells were transformed with the indicated DNA fragment(s), stitch(es), or 2μ plasmid, and host cell transformants were grown on the indicated selective agar. Genomic integrations in selected clones were confirmed by colony PCR using PCR primers homologous to the 5' and 3' junctions of the chromosomal integration sites.

Table 4 – Strains Generated

Strain	Parent Strain	Genetic Modification Introduced Into Parent Strain	Stitch	2μ Plasmid	Selective Agar
Y5821	Y002 [a]	replace ERG13 coding sequence with *S. cerevisiae* LEU2 coding sequence	i3830 [b]	-	CSM-L
Y6080	Y5823 [e]	replace HO coding sequence with *S. cerev.* URA3 coding sequence, E. coli dxs coding sequence with *S. cerev.* TDH3 promoter and SHM2 terminator, E. coli dxr coding sequence with *S. cerev.* TEF2 promoter and ADO1 terminator, E. coli ispD coding sequence with *S. cerev.* TDH2 promoter and YSB1 terminator, E. coli ispF coding sequence with *S. cerev.* TEF2 promoter and ADE17 terminator, and E. coli ispE coding sequence with *S. cerev.* TDH3 promoter and HOM2terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	s2075 [c] [f] s2076 [c] [f]	-	CSM-U
Y6081	Y6080	replace LEU2 coding sequence with *S. cerev.* HIS3 coding sequence, E. coli dxr coding sequence with *S. cerev.* TEF2 promoter and ADO1terminator, E. coli xylB coding sequence with *S. cerev.* TDH3 promoter and PUP2terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	i2145 [b]	-	CSM-H + 2 g/L mevalolactone [d]
Y6533	Y6081	replace GAL80 coding sequence with *S. cerev.* TRP1 coding sequence, StrepII-tagged *Bacillus thuringiensis* ispG fused to *Bacillus subtilis* flavodoxin with *S. cerev.* GAL1 promoter, *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* GAL10 promoter and OLE1 terminator, StrepII-tagged E. coli fldA coding sequence fused to E. coli ispH coding sequence with *S. cerev.* GAL1 promoter, and E. coli dxr coding sequence with *S. cerev.* GAL1 promoter and ADO1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.* Yes, you have the promoters correct	s11629 [c] [f] s11634 [c] [f]	-	CSM-W + 2 g/L mevalolactone [d]
Y6540	Y6081	replace ADH5 coding sequence with *S. cerev.* TRP1 coding sequence, StrepII-tagged *Bacillus thuringiensis* ispG fused to *Bacillus subtilis* flavodoxin with *S. cerev.* GAL1Promoter, *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* GAL10 promoter and OLE1 terminator, and StrepII-tagged E. coli fldA coding sequence fused to E. coli ispH coding sequence with *S. cerev.* GAL1 Promoter, and E. coli dxr coding sequence with *S. cerev.* GAL10 promoter and ADO1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	s11632 [c] [f] s11635 [c] [f]	-	CSM-W + 2 g/L mevalolactone [d]
Y4763	Y192 [f]	express StrepII-tagged Algoriphagus IspG coding sequence under regulatory control of *S. cerev.* GAL1 promoter and *S. cerev.* TILV5 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	-	pAM2133 (SEQ ID NO: 88; FIG. 9A)	CSM-U
Y4161	Y002	replace HO coding sequence with *S. cerev.* URA3 coding sequence, E. coli dxs coding sequence with *S. cerev.* TDH3 promoter and SHM2 terminator, E. coli dxr coding sequence with *S. cerev.* TEF2 promoter and ADO1 terminator, E. coli ispD coding sequence with *S. cerev.* TDH2 promoter and YSB1 terminator, E. coli ispF coding sequence with *S. cerev.* TEF2 promoter and ADE17 terminator, and E. coli ispE coding sequence with *S. cerev.* TDH3 promoter and HOM2terminator,	s2075 [c] [f] s2076 [c] [f]	-	CSM-W + 2 g/L mevalolactone [d]

		all heterologous sequences being codon-optimized for expression in *S. cerev.*			
Y4819	Y4161	replace LEU2 coding sequence with *S. cerev.* HIS3 coding sequence, E. coli dxr coding sequence with *S. cerev.* TEF2 promoter and ADO1terminator, E. coli xylB coding sequence with *S. cerev.* TDH3 promoter and PUP2terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	i2145 [b)]	-	CSM-H + 2 g/L mevalolactone [d)]
Y6237	Y4819	replace URA3 coding sequence at deleted HO locus with hisG coding sequence	i3850	-	CSM + 5 FOA + U
Y6276	Y6237	replace ADE2 coding sequence with *S. cerev.* TRP1 coding sequence, *Bacillus subtilis* flavodoxin coding sequence with *S. cerev.* ENO2 promoter and SUI3 terminator, *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* TEF2 promoter and OLE1 terminator, E. coli ispH coding sequence with *S. cerev.* TEF2 promoter and THS1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	s11469 [c) f)] s11475 [c) f)]	-	CSM-W + 2 g/L mevalolactone [d)]
Y6280	Y6237	replace ADE2 coding sequence with *S. cerev.* TRP1 coding sequence, *Bacillus subtilis* flavodoxin coding sequence with *S. cerev.* ENO2 promoter and SUI3 terminator, StrepII-tagged *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* TEF2 promoter and OLE1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	s11472 [c) f)] s11478 [c) f)]	-	CSM-W + 2 g/L mevalolactone [d)]
Y6292	Y6280	express StrepII-tagged *Thermus thermophilus* ispG with *S. cerev.* GAL1 promoter and ILV5 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	-	PGAL1-STREP-Tt_ispG (SEQ ID NO: 89; FIG. 9B)	CSM-U
Y6291	Y6280	express StrepII-tagged *Bacillus thuringiensis* ispG with *S. cerev.* GAL1 promoter and ILV5 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	-	PGAL1-STREP-Bt_ispG (SEQ ID NO: 90; FIG. 9C)	CSM-U
Y6281	Y6237	express StrepII-tagged *Bacillus thuringiensis* ispG with *S. cerev.* FBA1 promoter and ILV5 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.* replace ADE2 coding sequence with *S. cerev.* TRP1 coding sequence, *Bacillus subtilis* flavodoxin coding sequence with *S. cerev.* ENO2 promoter and SUI3 terminator, StrepII-tagged *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* TEF2 promoter and OLE1 terminator, and StepII-tagged E. coli ispH with *S. cerev.* TEF2 promoter and THS1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	s11472 [c) f)] s11479 [c) f)]	PFBA1-STREP-Bt_ispG (SEQ ID NO: 91; FIG. 9D)	CSM-U
Y6283	Y6276	express StrepII-tagged *Bacillus thuringiensis* ispG with *S. cerev.* GAL1 promoter and ILV5terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	-	PGAL1-STREP-Bt_ispG (SEQ ID NO: 90; FIG. 9C)	CSM-U
Y6279	Y6237	replace ADE2 coding sequence with *S. cerev.* TRP1 coding sequence, *Bacillus subtilis* flavodoxin coding sequence with *S.*	s11472 [c) f)] s11480 [c) f)]	-	CSM-W + 2 g/L

		cerev. ENO2 promoter and SUI3 terminator, StrepII-tagged *Arabidopsis thaliana* FNR truncated coding sequence with *S. cerev.* TEF2 promoter and OLE1 terminator, and StepII-tagged *Thermosynechococcus elongates* ispH with*S. cerev.* TEF2 promoter and OLA1 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*			mevalolactone [d)]
Y6289	Y6279	express StrepII-tagged *Bacillus thuringiensis* ispG with *S. cerev.* GAL1 promoter and ILV5 terminator, all heterologous sequences being codon-optimized for expression in *S. cerev.*	-	PGAL1-STREP-Bt_ispG (SEQ ID NO: 90; FIG. 9C)	CSM-U

CSM = 6.7 g/L Yeast Nitrogen Base (YNB; BD Diagnostics, Franklin Lakes, NJ), 2 g/L Synthetic Complete Mix (CSM; Bufferad, Lake Bluff, IL), 2% glucose, 1% methionine, 1% leucine, 0.2% adenine sulfate, 0.2% uracil, 1% lysine, 1% trypophan, 1% histidine; CSM-L = CSM comprising all amino acids except leucine; CSM-U = CSM comprising all amino acids except uracil; CSM-H = CSM comprising all amino acids except histidine; CSM-W = CSM comprising all amino acids except tryptophan; CSM + 5 FOA + U = CSM comprising all amino acids and 5- fluoroorotic acid.

a) CEN.PK2 background MAT A; ura3-52; trp1-289; leu2-3,112; his3Δ1; MAL2-8C; SUC2 (van Dijken et al. *Enzyme Microb. Technol.* 2000, 26:706-714)
b) The construct was digested to completion using PmeI restriction endonuclease (New England Biolabs, Inc., Ipswich, MA) to liberate the insert from the pAM2041 vector, before the isolated insert was transformed into the parent strain.
c) The constructs were digested to completion using PmeI restriction endonuclease (New England Biolabs, Inc., Ipswich, MA) to liberate the inserts from the pAM2041 vectors, and the purified inserts were then co-transformed into the parent strain where they recombined with each other and with the host cell genome by host cell mediated homologous recombination.
d) Simga-Aldrich, St. Louis, MO.
e) See Example 6.
f) BJ5459, ATCC # 208284

Example 3

This example describes methods for the detection and quantitation of DXP pathway metabolites by Mass Spectrometry.

Engineered yeast cells were rapidly quenched in -80°C 8:1:1 methanol:water:glycerol at a ratio of 1:4 cell broth:quenching solution, and pelleted by centrifugation. After disposing of quench, pellets were extracted by sonication, vortexing with beads, resting on wet ice, and centrifugation in two fractions of 7:3 methanol:20 mM ammonium acetate, pH 10 extraction buffer spiked with farnesyl-thiolodiphosphate (TFPP; Echelon Biosciences, Salt Lake City, UT) as an internal standard. The extracts were combined, diluted in methanol, centrifuged to pellet salts, and transferred to LC vials for injection. Analyte separation was achieved via a gradient from acetonitrile to aqueous 20 mM ammonium acetate, pH 10, on a LUNA™ NH2 column (Phenomenex, Torrance, CA) using high pressure liquid chromatography (Shimadzu, Kyoto, Japan) (see Table 5 for gradient conditions). Mass spectral data was acquired via scheduled multiple reaction monitoring (MRM) on an API 5000 instrument (Applied Biosystems, Carlsbad, CA). Negative ionization mode was used to monitor two MRM transitions each for DXP, cMEPP, HDMAPP, DMAPP/IPP, and FPP, with retention time and fragmentation confirmed with purchased standards (Sigma, St Louis, MO; and

Echelon Biosciences, Salt Lake City, UT) (see Table 6). Experiments involving labeled isotope feeding also included MRM transitions for the aforementioned analytes plus 5 atomic mass units, and plus 15 atomic mass units for FPP. Data was analyzed using ANALYST™ (AB SCIEX, Foster City, CA), and normalized to cell number by measuring the optical density of cultures at 600nm (OD_{600}). Calibration curves using TFPP and standards were used for absolute quantitation.

Table 5 - HPLC Gradient

Minutes	% A	% B
0	30	70
2	35	65
5	45	55
8	55	45
9.99	90	10
10	30	70
12	30	70

A = aqueous 20 mM ammonium acetate, pH 10.
B = acetonitrile.

Table 6 - List of Masses						
Q1 Mass	Q3 Mass	Scheduled Retention Time	Analyte, Transition #	Declustering Potential	Collision Energy	Collision Cell Exit Potential
212.982	78.8	7.5	DXP_1	-55	-42	-21
212.982	96.8	7.5	DXP_2	-55	-20	-13
212.982	138.9	7.5	DXP_3	-55	-18	-19
217.982	78.8	7.5	DXP_1,+5 AMU	-55	-42	-21
217.982	96.8	7.5	DXP_2,+5 AMU	-55	-20	-13
217.982	138.9	7.5	DXP_3,+5 AMU	-55	-18	-19
276.71	79	7.5	cMEPP_1	-100	-38	-9
276.71	158.8	7.5	cMEPP_2	-100	-32	-23
281.71	79	7.5	cMEPP_1,+5 AMU	-100	-38	-9
281.71	158.8	7.5	cMEPP_2,+5 AMU	-100	-32	-23
260.916	78.9	9.3	HDMAPP_1	-75	-52	-9
260.916	62.9	9.3	HDMAPP_2	-75	-130	-13
260.916	158.9	9.3	HDMAPP_3	-75	-22	-21
265.916	78.9	9.3	HDMAPP_1,+5 AMU	-75	-52	-9
265.916	62.9	9.3	HDMAPP_2,+5 AMU	-75	-130	-13
265.916	158.9	9.3	HDMAPP_3,+5 AMU	-75	-22	-21
245.021	78.9	9.3	DMAPP_1	-35	-24	-9
245.021	158.9	9.3	DMAPP_2	-35	-24	-9
250.021	78.9	9.3	DMAPP_1,+5 AMU	-35	-24	-9
250.021	158.9	9.3	DMAPP_2,+5 AMU	-35	-24	-9
381.065	78.9	9	FPP_1	-95	-38	-11
381.065	63	9	FPP_2	-95	-130	-11
386.065	78.9	9	FPP_1,+5 AMU	-95	-38	-11
386.065	63	9	FPP_2,+5 AMU	-95	-130	-11
391.065	78.9	9	FPP_1,+10 AMU	-95	-38	-11
391.065	63	9	FPP_2,+10 AMU	-95	-130	-11
396.065	78.9	9	FPP_1,+15 AMU	-95	-38	-11
396.065	63	9	FPP_2,+15 AMU	-95	-130	-11
397.055	158.9	9.4	TFPP	-65	-26	-21

Example 4

This example describes methods for detection and quantitation of DXP pathway proteins in yeast by Mass Spectrometry.

Engineered yeast cells were grown to yield sufficient biomass (OD_{600} x culture volumes (mL) > 5). Pellets were lysed in the presence of 6 M urea, 500 mM triethylamine bicarbonate pH 8.5, with glass beads on a mini-beadbeater (Biospec, Bartlesville, OK). Total protein measurements were taken using a QUBIT™ Fluorometer (Invitrogen, Carlsbad, CA), and 20-30 µg total protein was reduced by incubating with 6 mM tris(2-carboxyethyl)phosphine at 37°C for 30 min, alkylated by incubating with 11 mM iodoacetamide at room temperature for 30 min in the dark, and digested with tosyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin at a trypsin:total protein ratio of at least 1:20 at 37°C, overnight. The digestion reaction (approximately 50 µL total volume) was then quenched with 450 µL 0.1% formic acid in water, and analyzed on a QTRAP™ 4000 Mass Spectrometer (AB SCIEX, Foster City, CA) (see Table 7 for gradient conditions).

Table 7 – LC Gradient		
Time (min)	% A	% B
0	98	2
13.5	60	40
14.5	15	85
17.5	98	2
A = 0.1% formic acid in water B = 0.1% formic acid in acetonitrile		

For LC-MS/MS acquisition method design, sequence information for proteins of interest were obtained and fed into a FASTA file. The FASTA file was fed into Skyline software (MacCoss Lab, University of Washington, WA), and an in-silico digestion was performed by the software, yielding tryptic fragments amenable to detection by LC-MRM. The resulting MRM transitions were then transferred to an acquisition method in ANALYST™ (AB SCIEX, Foster City, CA). The acquisition method was an MRM-IDA (Multiple Reaction Monitoring – Information Dependent Acquisition), and suitable MRM's triggered MS/MS that were then used to confirm the identity of the desired tryptic fragments belonging to the proteins of interest.

The acquired MS/MS spectra were searched in Mascot against the Swissprot database, and spectra with p-values of less than 0.05 were accepted as positive IDs. The retention time of the positively-identified MS/MS spectra was matched to the retention time of the MRM traces of concurrent y-ions corresponding to the matched tryptic peptide. The identified tryptic peptide signature could then be used to determine the presence of a specific protein, or even to perform relative quantitation of proteins.

Example 5

This example describes the optimization of expression of the first five enzymes in the DXP pathway (DXP-TOP pathway).

Six independent isolates of Y6080 were grown up in CSM comprising 2 g/L mevalolactone to an OD_{600} of between 2 and 3, and analyzed for metabolites of the DXP-TOP pathway as described in Example 3.

As shown in FIG. 2, all Y6080 isolates produced cMEPP at an internal cellular concentration of between 300 to 400 µM, suggesting that the DXP-TOP pathway was active in these cells. However, significant levels of DXP were also detected, suggesting that DXR enzyme activity was limiting.

To increase carbon flux through the first five DXP enzymes to cMEPP, strain Y6081 was generated, which comprises an additional copy of the *Escherichia coli* dxr coding sequence, in addition to a copy of the *Escherichia coli* gene xylB, which encodes an enzyme that can convert

deoxy-xylulose into DXP. Six independent isolates of Y6081 were analyzed for metabolites of the DXP-TOP pathway as described in Example 3. As shown in FIG. 2, all but one Y6081 isolate (#36) showed an increase in cMEPP concentration. Isolate lineage #27 of Y6080 and Y6081 was chosen for all future strain construction work.

Example 6

This example describes the selection of yeast strains that can grow in the absence of an endogenous MEV pathway. Because these strains are mevalonate auxotrophs, they can only survive if provided mevalonate in their growth medium. As a consequence, a heterologous DXP pathway can be engineered in these strains and tested to determine if it can rescue these strains by making IPP and DMAPP through the DXP pathway when placed in medium without mevalonate. In addition, inactivation of YPK1 allows yeast strains to be more efficient in their use of mevalonate.

Three positive transformants of Y5821 were streaked on CSM agar plates lacking leucine and comprising 5 g/L mevalolactone or 0 g/L mevalolactone. Plates were incubated at 30°C for 7 - 10 days. No growth was observed on CSM agar plates containing 0 g/L mevalolactone, confirming that Y5821 is a mevalonate auxotroph. Cells were scraped from the CSM agar plates comprising 5 g/L mevalolactone, washed in sterile water, resuspended to a concentration of 5 x 10^7 cells in 1 mL of sterile water, and plated (5 x 10^6 cells per plate) onto ten CSM agar plates comprising 2 g/L mevalolactone, a mevalolactone concentration that supports colony growth by Y5821 cells only after 14 days. Three single colonies visible after 5 days of growth at 30°C were picked, and were confirmed to grow faster than their Y5821 parent in liquid CSM lacking leucine (see Table 8). The three strains were also plated on CSM agar plates comprising 0 g/L mevalonolactone to confirm that they were still mevalonate auxotrophs.

Y5822, Y5823, and Y5824 were subjected to another round of screening for faster growth on CSM agar plates comprising 0.267 g/L mevalolactone. In this second round of screening, strain Y5825 was derived from Y5824. Y5825 was confirmed to grow even faster than Y5824 in liquid CSM lacking leucine.

Table 8 - Doubling Time (h^{-1}) in Liquid CSM Comprising Mevalonolactone

	5 g/L mevalonolactone	2 g/L mevalonolactone
Y002	2.41	2.39
Y5821	4.63	no growth observed in 10 hours
Y5822	not determined	3.82
Y5823	not determined	2.48
Y5824	not determined	4.74

Additional Y5821 cells were scraped from CSM agar plates comprising 5 g/L mevalolactone, washed in sterile water, resuspended to a concentration of 5 x 10^8 cells in 1 mL of sterile water, and plated onto ten CSM agar plates comprising 0 g/L mevalolactone (5 x 10^7 cells per plate). After three weeks of growth at 30^{0}C, no colonies were visible, indicating the frequency of spontaneous suppression of the erg13 deletion is less than 1 in 5 x 10^8 cells.

Strains Y5822, Y5823, Y5824, and Y5825 were submitted to whole genome sequencing and compared to a reference CENPK parent genome. As shown in Table 9, each mutant strain comprised only a single mutation, except Y5825, which had two mutations. The mutations fell into two genes: YPK1 and PRM10/YJL107C. YPK1 is annotated in the Saccharomyces Genome Database (yeastgenome.org, Feb 21, 2011) as a "serine/threonine protein kinase that phosphorylates and downregulates flippase activator Fpk1p; mutations affect receptor-mediated endocytosis and sphingolipid-mediated and cell integrity signaling pathways." Mutations found in YPK1 in Y5822, Y5824, and Y5823 are presumably inactive or negatively affect YPK1, indicating that loss of function of this protein allows *S. cerevisiae* to grow on a lower concentration of mevalolactone. YJL107C is directly upstream from PRM10 and the coding sequences overlap at the annotated stop (TGA) and start (ATG) codons for the two genes. In the yeast *Zygosaccharomyces rouxii*, these two sequences are annotated as a single gene, and *Saccharomyces cerevisiae* transcript data indicates that PRM10/YJL107C have identical transcriptional profiles (Xu Z. et al. *Nature* 2009, 457:1033-1037). The data therefore support the conclusion that PRM10/YJL107C is a single transcriptional unit. The entire PRM10/YJL107C protein contains a Major Facilitator Superfamily (MFS) domain, which is a domain found in a large family of transporters.

Example 7

This example describes methods for expressing and purifying soluble DXP pathway enzymes in yeast cells.

A pGALl expression vector was generated for use in purifications of heterologous DXP proteins in yeast cells. To this end, the DNA fragments shown in Table 9 were PCR-amplified as described in Example 1.

Table 9 – PCR amplified DNA Fragments		
DNA Fragment	Primers	Template
Ec_ispG-1	GW323 (SEQ ID NO: 110) [a] RYSE7 (SEQ ID NO: 111)	SEQ ID NO: 112
Ec_ispG-2	GW324 (SEQ ID NO: 113) [b] RYSE7 (SEQ ID NO: 111)	Ec_ispG-1
PGAL1	GW327 (SEQ ID NO: 114) GW325 (SEQ ID NO: 115)	Y002 genomic DNA
PRS02	R2870-fwd (SEQ ID NO: 116) R2870-rev (SEQ ID NO: 117)	pAM70 [c]
PRS39	R2871-fwd (SEQ ID NO: 118) R2871-rev (SEQ ID NO: 119)	pAM70 [c]

a) GW323 eliminates the starting ATG from the ispG coding sequence, and adds 30 bp of additional sequence encoding ten additional amino acids that comprise the StrepII tag (WSHPQFEK; Schmidt, TG and Skerra, A. The strep-tag system for one-step purification and high-affinity detection or capturing of proteins. Nature Protocols 2007, 2, 1528-1535).
b) GW324 has 20 bp of overlapping homology to the 5' terminus of GW323, and adds an additional 33 bp of sequence, representing six DNA bases of the optimal Kozack sequence (GTAACC) and encoding nine additional amino acids. The GTAACC sequence, as well as the +2 serine after the +1 methionine at the junction of the nine added amine acids promotes efficient translation initiation and expression in yeast (Macbeth and Bass. Large-scale overexpression and purification of ADARs from *Saccharomyces cerevisiae* for biophysical and biochemical studies. *Methods Enzymol.* 2007, 424, 319-31).
c) Christianson TW, *et al.* Multifunctional yeast high-copy-number shuttle vectors. Gene 1992, 110, 119-22.

The DNA fragments were assembled into stitches as outlined in Table 10 and described in Example 1.

Table 10 - Terminal Primers for Assembly of Stitches			
Stitch	DNA Fragments (see Table 10)	Terminal Primer 1	Terminal Primer 2
PGAL1-StrepII- ispG	PGAL1, Ec_ispG-2	GW327 (SEQ ID NO: 114)	RYSE 7 (SEQ ID NO: 111)
PRS02-PGAL1-StrepII ispG-PRS39-pAM09	PRS02, PGAL1-StrepII- ispG, PRS39, linearized pAM2041 [a]	terminal end of linearized pAM2041	terminal end of linearized pAM2041

[a] Plasmid pAM2041 (SEQ ID NO: 92) was digested to completion using SchI restriction endonuclease, and the isolated linearized plasmid was included in the PCR reaction.

The final stitch, PRS02-PGAL1-StrepII-ispG -PRS39-pAM09, was transformed into chemically competent *Escherchia coli* cells. Positive transformants were selected on Luria Bertani (LB) agar containing carbenicillin. Selected clones were grown in LB containing 100 µg/mL carbenicillin, the plasmid was extracted by standard techniques, and the isolated plasmid was sequenced to verify the DNA was assembled correctly. A sequence-verified plasmid was digested with the restriction enzyme PmeI (New England BioLabs, Inc., Ipswich MA) to release the PRS02-PGAL1-StrepII-ispG-PRS39 insert from the vector backbone. Vector pAM70 was digested with XmaI and SacI-HF restriction endonucleases (New England BioLabs, Inc., Ipswich MA). 100 ng of purified linearized pAM70 was mixed with 150 ng of purified PRS02-PGAL1-StrepII-ispG-PRS39 insert, and the mixture was used to transform a yeast strain in which the final vector pAM2130 (FIG. 3A) was created *in vivo* via host cell mediated gap-repair homologous recombination. The vector was

recovered using the ZYMOPREP™ Yeast Plasmid Miniprep II kit (Zymo Research, Irvine, CA), and sequence verified. pAM2130 was digested using PacII and AatII restriction endonucleases (New England Biolabs, Ipswich, MA) to remove the ispG coding sequence, and the vector backbone was purified using a gel extraction kit (Qiagen, Valencia, CA), yielding the pGAL1 expression vector.

All proteins analyzed by *in vitro* biochemistry in this invention were cloned into the pGAL1 expression vector using gap repair into yeast to create an in-frame fusion protein containing the StrepII tag at the N terminus of the protein. The StrepII tag could then be used to identify the expressed proteins via Western blot, and to purify the expressed proteins using a STREP-TACTIN™ resin. Cloning into the pGAL1 expression vector was accomplished by PCR-amplifying a sequence encoding a protein of interest using a primer comprising, in addition to gene-specific sequences, the 5'-terminal sequence GGGAGCTGAAACCGCGGTCCCAAATTCC, and another primer comprising, in addition to gene-specific sequences, the 5'-terminal sequence AACCCTCACTAAAGGGAACAAAAGCTGG, and by co-transforming yeast host cells with the purified PCR product and the linear pGAL1 expression vector. The terminal sequences of the PCR product served to direct the sequence encoding the protein of interest to be recombined into the co-transformed pGAL1 expression vector by the native recombination machinery in *Saccharomyces cerevisiae* (see, for example, the cloning of fusion proteins in Example 11).

For protein expression, strain Y4743 (Macbeth and Bass. Large-scale overexpression and purification of ADARs from *Saccharomyces cerevisiae* for biophysical and biochemical studies. *Methods Enzymol*. 2007, 424:319-31) was used. This strain is deficient in the expression of several proteases and has an extra copy of the GAL4 gene of *Saccharomyces cerevisiae* under regulatory control of the promoter of the GAL1 gene of *Saccharomyces cerevisiae*, modifications that are expected to increase heterologous protein stability and expression, respectively.

For biomass build and induction, a 50 mL culture of Y4743 cells containing the plasmid of interest were grown overnight in modified CSM (6.7 g/L YNB, 10 mL/L succinic acid, 6 g/L NaOH pellets, 1.92 g/L CSM, 0.002% adenine sulfate) containing all amino acids with the exception of uracil and containing 2% glucose. This culture was diluted to an OD_{600} of 0.2 in 50 mL fresh modified CSM lacking uracil and containing 2% glucose (to maintain selection for the plasmid), and allowed to grow for 24 hours. The complete culture was then harvested by centrifugation, the supernatant discarded, and the culture pellet resuspended to a total volume of 50 mL in modified CSM lacking uracil and containing 2.7% galactose as the only carbon source. This switch to galactose was the induction phase, and lasted a minimum of 16 hours and a maximum of 24 hours. Cell pellets were harvested, washed in 1x phosphate buffered saline (PBS), and stored at -80 °C until needed.

When necessary, cell pellets were taken into the anaerobic chamber to seal the pellet in a layer of gas free of oxygen.

For protein purification, the StrepII tagged proteins were purified using STREP-TACTIN™ sepharose resin or the STREP-TACTIN™ spin columns (IBA GmbH, Goettingen, Germany) according to the manufacturer's suggested protocol with the exception that avidin was added at a final concentration of 40 µg/mL to prevent native biotinylated proteins in the cell free extract from binding the STREP-TACTIN™ resin.

To determine the fraction of soluble *Escherichia coli* ispG protein produced using the pGAL1 expression vector, Y4743 cells comprising pAM2130 were lysed by either the protein denaturing extraction method or the soluble protein extraction method. For the soluble protein extraction method, cell pellets were resuspended in a non-denaturing buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, and 1 complete mini protease inhibitor cocktail tablet (Roche Applied Science, Indianapolis, IN) per 10 mL of buffer), 0.5 mm glass beads were added, and the cells were lysed using a MP beadbeater. Tubes were centrifuged, total protein concentration in the supernatant was determined by Bradford analysis, and 2.5 µg of total protein was separated by SDS PAGE and subjected to Western blot using an anti-ispG primary antibody. For the protein denaturing extraction method, cell pellets were resuspended in a denaturing buffer (20 mM Tris pH 7.4, 8M urea, 2%CHAPS, 1%DTT, and 1 complete mini protease inhibitor cocktail tablet per 10 mL of buffer), 0.5 mm glass beads were added, and the cells were lysed using a MP beadbeater. Tubes were centrifuged, total protein concentration in the supernatant was determined by Bradford analysis, and 3.5 µg of total protein was separated by SDS PAGE and subjected to Western blot using an anti-ispG primary antibody.

As shown in FIG. 3B, the signal obtained with the denaturing protein extraction method (lane 3) was significantly stronger than the signal obtained using the soluble protein extraction method (lane 2), suggesting that the large majority of *Escherichia coli* IspG was insoluble when expressed in the yeast cytosol.

To investigate whether a lack of soluble IspG expression in Y4743 cells comprising pAM2130 was related to an Fe-S incorporation issue, a mitochondrial-targeted *Escherichia coli* ispG expression construct was created. To this end, the DNA fragments shown in Table 11 were PCR-amplified as described in Example 1.

Table 11 – PCR Amplified DNA Fragments		
DNA Fragment	Primers	Template
PGAL1/10-2	GW114 (SEQ ID NO: 120) GW115 (SEQ ID NO: 121)	Y002 genomic DNA
MTS [a)]	GW116 (SEQ ID NO: 123) GW150 (SEQ ID NO: 124)	SEQ ID NO: 122 [b)]
PGAL1-MTS-2	GW400 (SEQ ID NO: 125) GW401 (SEQ ID NO: 126)	PGAL1/10-MTS-1 (see Table 13)
a) MTS = mitochondrial targeting signal of the Su9 gene of *Neurospora crassa* (Westermann and Neupert, Mitochondria-targeted green fluorescent proteins: convenient tools for the study of organelle biogenesis in *S. cerevisiae*. Yeast 2000,16, 1421-1427), codon-optimized for expression in *Saccharomyces cerevisiae.* b) The synthetically-generated DNA fragment was provided by Integrated DNA Technologies, Inc. (Coraville, IA).		

The DNA fragments were assembled into stitches as outlined in Table 13 and as described in Example 1.

Table 13 - Terminal Primers for Assembly of Stitches			
Stitch	DNA Fragments (see Table 12)	Terminal Primer 1	Terminal Primer 2
PGAL1/10-MTS-1	PGAL1/10, MTS	GW114 (SEQ ID NO: 120)	GW150 (SEQ ID NO: 124)

The final PGAL1-MTS-2 PCR fragment had 500 bp of homology to the PGAL1 promoter at its 5'-end and 30 bp of homology to the StrepII tag on its 3'-end. It was mixed with linearized pAM2130, and the mixture was used to transform Y4743 cells, in which the two DNA fragments were ligated via host cell mediated gap-repair homologous recombination to yield a plasmid. The plasmid was recovered using the ZYMOPREP™ Yeast Plasmid Miniprep II kit (Zymo Research, Irvine, CA), and sequence verified. The plasmid was digested using Age1 restriction endonuclease to remove the ispG coding sequence, and the vector backbone was purified using a QUIAGEN™ gel extraction kit, yielding the pGAL1-MTS expression vector. The mitochondrial-targeted IspG protein (Ec_IspG-MTS) was found to be mostly soluble, readily purified, and active.

Example 8

This example describes the *in vitro* identification of active IspG enzymes in *Saccharomyces cerevisiae*.

Using the pGAL1 expression vector, various IspG homologs were expressed in Y4743, extracted from the host cells via the soluble protein extraction method, and analyzed by Western blot as described in Example 7.

As shown in Table 13, 29 of the IspG homologs tested proved to be soluble in the host cell cytosol. The activity of the 29 soluble IspG homologs was then determined using artificial electron donors. For *in vitro* enzyme activity assays with artificial electron donors, the 50 µL reaction contained degassed 100 mM Tris buffer pH8.0, cMEPP at 320 µM, 5 mM dithionite, 1

mM Methyl viologen, and 5 µL of freshly purified IspG protein. The total volume was brought up to 50 µL using degassed, deionized water. Samples were incubated in the anaerobic chamber at ambient temperature. At the desired time points, 10 µL of the reaction was added to 490 µL of methanol with 1 µg/mL of internal standard (TFPP). Samples were then kept at -80°C until they were analyzed by LC-MRM as described in Example 3. The activity of the IspG enzymes was measured by the rate of HDMAPP formation. Relative activity was determined by the relative amount of HDMAPP formation observed with equivalent amounts of IspG protein over the same amount of time.

Table 13 – Solubilities and Activities of IspG Homologs in Yeast Host Cells

IspG Homolog	Species Origin	Protein Accession Number	Solubility in Host Cell Cytosol	Activity with Artificial Electron Donors
Al_G (SEQ ID NO: 153)	*Algoriphagus sp. PR1*	ZP_01720138.1	++++	++
Ap_G (SEQ ID NO: 154)	*Arthrospira platensis str. Paraca*	ZP_06384238.1	++++	++
At_G (SEQ ID NO: 155)	*Arabidopsis thaliana*	AAO15446	++++	–
Bb_G (SEQ ID NO: 156)	*Brevibacillus brevis NBRC 100599*	YP_002772921.1	++++	+
Bc_G (SEQ ID NO: 157)	*Bacillus coagulans 36D1*	ZP_04432900.1	++++	+++
Bs_G (SEQ ID NO: 158)	*Bacillus subtilis subsp. subtilis str. 168*	NP_390386.1	++++	+++
Bt_G (SEQ ID NO: 159)	*Bacillus thuringiensis IBL 200*	ZP_04073932.1	++++	++++
Caro_G (SEQ ID NO: 160)	*Catharanthus roseus*	AAO24774.1	+++	++
Chl_G (SEQ ID NO: 161)	*Chlamydomonas reinhardtii*	XP_001690937.1	+	+
Cw_G (SEQ ID NO: 162)	*Crocosphaera watsonii WH 8501*	ZP_00518798.1	++++	+
Cy_G (SEQ ID NO: 163)	*Cyanobium sp. PCC 7001*	ZP_05045928.1	++++	+
Ec_G-mts (see Table x)	*Escherichia coli*	AAC75568 with MTS	++++	++++
Ed_G (SEQ ID NO: 164)	*Eubacterium dolichum DSM 3991*	ZP_02077259.1	++++	+
Fs_G (SEQ ID NO: 165)	*Fibrobacter succinogenes subsp. succinogenes S85*	YP_003250459	++++	–
Mc_G (SEQ ID NO: 166)	*Magnetococcus sp. MC-1*	YP_867482	++++	–
Mp_G (SEQ ID NO: 167)	*Micromonas pusilla CCMP1545*	XP_003056663.1	++++	–
Np_G (SEQ ID NO: 168)	*Nostoc punctiforme PCC 73102*	YP_001868335.1	++++	+
Ot_G (SEQ ID NO: 169)	*Ostreococcus tauri*	CAL55258	+++	+++

Pb_G (SEQ ID NO: 170)	*Prevotella buccae D17*	EFC75705	++++	+
Pc_G (SEQ ID NO: 171)	*Paulinella chromatophora*	YP_002049460.1	++++	–
Pm_G (SEQ ID NO: 172)	*Prochlorococcus marinus subsp. pastoris str. CCMP1986*	NP_892794.1	++++	+
Py_G (SEQ ID NO: 173)	*Porphyra yezoensis red algae*	ACI45961	++++	+
Sb_G (SEQ ID NO: 174)	*Sorghum bicolor*	XP_002454137	++	++
Sm_G	*Selaginella moellendorffii*	XP_002986683.1	++	–
Syn_G (SEQ ID NO: 175)	*Synechococcus sp. RS9917*	ZP_01081635.1	++++	++
Te_G (SEQ ID NO: 176)	*Thermosynechococcus elongatus BP-1*	NP_681786	++++	++
Tt_G (SEQ ID NO: 177)	*Thermus thermophilus HB8*	YP_143571.1	++++	+++++
Vv_G (SEQ ID NO: 178)	*Vitis vinifera*	XP_002285130.1	++++	++
All IspG homologues tested were codon-optimized for expression in *Saccharomyces cerevisiae*, and devoid of any organelle targeting sequences. – = no activity detected; ++++, +++, ++, + = most soluble or active, highly soluble or active, medium soluble or active, poorly soluble or active, respectively.				

Example 9

This example describes the *in vitro* identification of active IspH enzymes in *Saccharomyces cerevisiae*.

Using the pGAL1 expression vector, various IspHs were expressed in Y4743, extracted from the host cells via the soluble protein extraction method, and analyzed by Western blot as described in Example 7. Table 14 illustrates seven exemplary IspHs that were soluble in the host cell cytosol.

The activity of the exemplary IspH homologs was then determined using artificial electron donors as described for IspG except that in place of cMEPP at 320 µM, HDMAPP at 380 µM was used in the 50 µLreaction, and the activities of the IspH enzymes were measured by the rate of DMAPP/IPP formation.

IspH enzymes were generally much more robust than IspG enzymes. In general, if the enzymes were soluble in the host cell's cytosol, they were more likely to be fully functional, displaying good to excellent enzymatic activity. Illustratives examples of suitable IspH's include but are not limited to IspH enzymes from: *Escherichia coli*, *Fibrobacter succinogens*, *Physcometrella patens*, *Sorghum bicolor*, *Stevia rebaudiana*, *Thermosynecoccus elongatus*, and *Arabidopsis thaliana*.

Table 14 – Solubilities and Activities of IspH Homologs in Yeast Host Cells

IspH Homolog	Species Origin	Protein Accession Number	Solubility in Host Cell Cytosol	Activity with Artificial Electron Donors
At-H (SEQ ID NO: 179)	*Arabidopsis thaliana*	AY168881	++++	++++
Ec-H (SEQ ID NO: 180)	*Escherichia coli*	AAC73140	++	++++
Fs-H (SEQ ID NO: 181)	*Fibrobacter succinogenes subsp. succinogenes S85*	YP_003250012.1	++	++
Pp-H (SEQ ID NO: 182)	*Physcomitrella patens subsp. patens (1)*	XP_001758369.1	++	++++
Sb-H (SEQ ID NO: 183)	*Sorghum bicolor*	XP_002463933.1	++++	++++
Sr-H (SEQ ID NO: 184)	*Stevia rebaudiana*	ABB88836.2	++++	+++
Te-H (SEQ ID NO: 185)	*Thermosynechococcus elongatus BP-1*	NP_681832	++++	++++

All IspH homologs tested were codon-optimized for expression in *Saccharomyces cerevisiae*, and devoid of any organelle targeting sequences.
NA = not available (*e.g.*, not expressed in soluble form in yeast); NT = not tested; ++++, +++, ++, + = most soluble or active, highly soluble or active, medium soluble or active, poorly soluble or active, respectively.

Example 10

This example describes the *in vitro* identification of redox partners for DXP pathway enzymes IspG and IspH in *Saccharomyces cerevisiae*.

Using the pGAL1 expression vector, the 14 flavodoxin and 10 flavodoxin/ferredoxin $NADP^+$-reductases (FNR) homologs listed in Tables 14 and 15, respectively, were expressed in Y4743, extracted from the host cells via the soluble protein extraction method, and analyzed by Western blotting as described in Example 7.

Table 15 shows the results of exemplary flavodoxins that were soluble in the cytosol and their respective degree of compatibility with IspG. As stated previously, IspH's are generally more robust than IspG's and suitable redox partners and reductases that provide a fully functional IspH were more readily identified.

Table 15 – Solubilities and Activities as Redox Partner for IspG Enzymes of Flavodoxin Homologs in Yeast Host Cells

Flavodoxin Homolog	Species Origin	Protein Accession Number	Solubility in Host Cell Cytosol	Redox Partner Activity with IspG
AM_fld (SEQ ID NO: 127)	*Alteromonas macleodii ATCC 27126*	ZP_04714638.1	++++	+
Anabaena_fld (SEQ ID NO: 128)	*Anabaena PCC 7119*	AAB20462	++++	++
Bc_fld (SEQ ID NO: 129)	*Bacillus coagulans 36D1*	ZP_04431144.1	++++	+
BsyN_fld (SEQ ID NO: 130)	*Bacillus subtilis subsp. subtilis str. 168 YkuN*	NP_389298.1	++++	++++
BsyP_fld (SEQ ID NO: 131)	*Bacillus subtilis subsp. subtilis str. 168 YkuP*	NP_389300.2	++++	+
Cw_fld (SEQ ID NO: 132)	*Crocosphaera watsonii WH 8501*	ZP_00515759.1	++++	+
Dv_fld (SEQ ID NO: 133)	*Desulfovibrio vulgaris subsp. vulgaris DP4*	YP_966026.1	++++	+
Ec_fld (SEQ ID NO: 134)	*Escherichia coli str. K-12 substr. MG1655*	NP_415210	++++	+
Np_fld (SEQ ID NO: 135)	*Nostoc punctiforme PCC 73102*	YP_001866434.1	++++	+
Rs_fld (SEQ ID NO: 137)	*Rhodobacter sphaeroides ATCC 17025*	YP_001169375.1	++++	–
Syn_fld (SEQ ID NO: 138)	*Synechococcus sp. CC9605*	YP_382866.2	++++	–
Vh_fld (SEQ ID NO: 140)	*Vibrio harveyi 1DA3*	ZP_06175070.1	++++	–
Ws_fld (SEQ ID NO: 141)	*Wolinella succinogenes DSM 1740*	NP_906993.1	++++	++

All flavodoxin homologs tested were codon-optimized for expression in *Saccharomyces cerevisiae*, and devoid of any organelle targeting sequences.

– = no activity detected; ++++, +++, ++, + = most soluble or active, highly soluble or active, medium soluble or active, poorly soluble or active, respectively.

Table 16 – Solubilities and Activities as Redox Partner for IspG and IspH Enzymes of FNR Homologs in Yeast Host Cells

FNR Homolog	Species Origin	Protein Accession Number	Solubility in Host Cell Cytosol	Redox Partner Activity with IspGs and Flavodoxins	Redox Partner Activity with IspHs and Flavodoxins
Anabaena_FNR (SEQ ID NO: 143)	*Anabaena PCC 7119*	1QUE_A	++++	++++	NT
At_RFNR2 (SEQ ID NO: 145)	*Arabidopsis thaliana (root)*	NP_849734.1	++++	++++	++++
Ec_FNR (SEQ ID NO: 146)	*Escherichia coli str. K-12 substr. MG1655*	NP_418359	++++	++	+++
Nostoc_FNR (SEQ ID NO: 147)	*Nostoc sp. PCC 7906*	BAG69182	++++	++++	++++
So_FNR (SEQ ID NO: 150)	*Spinacia oleracea*	P00455	+++	++	NT
Te_FNR (SEQ ID NO: 151)	*Thermosynechococcus elongatus BP-1*	NP_682001	++++	++++	++++
Zm_FNR (SEQ ID NO: 152)	*Zea mays*	3LVB_A	++++	++++	++++

All FNR homologs tested were codon-optimized for expression in *Saccharomyces cerevisiae*, and devoid of any organelle targeting sequences.
NA = not available (*e.g.*, not expressed in soluble form in yeast); NT = not tested; ++++, +++, ++, + = most soluble or active, highly soluble or active, medium soluble or active, poorly soluble or active, respectively.

Assays with flavodoxin homologs and FNR homologs were performed essentially as described above for artificial electron donors, except the 50 µL reaction contained degassed 100 mM Tris buffer pH8.0, cMEPP at 320 µM (for ispG) or HDMAPP at 380 µM (for ispH), 5 µL of the freshly purified FNR, 5 µL of the freshly purified flavodoxin, and 5 µL of freshly purified IspG or IspH protein.

The four FNR homologs proved to be of equal activity, such that any of them could be used in combination with IspG or IspH and a flavodoxin without a change in overall enzyme activity. Certain redox partners were found to be compatible with both IspG and IspH. Illustrative examples of two such redox partners include *Escherichia coli* str. K-12 substr. MG1655 (NP_415210) and *Bacillus subtilis subsp. subtilis* str. 168 YkuN (NP_389298.1). However, each IspG and IspH provided the best activity if paired with a different redox partner, suggesting that the respective redox partner in the endogenous system is specific and distinct.

The most compatible flavodoxin/IspG pair that was identified from the above screen was the flavodoxin of *Bacillus subtilis* YkuN with IspG from *Bacillus coagulans*, *Bacillus thuringiensis, Bacillus subtilis*, and *Thermus thermophilus*.

The most compatible flavodoxin/IspH pair that was identified was *Escherichia coli* flavodoxin with IspH from *S. bicolor*. The *Bacillus subtilis* flavodoxin YkuN also worked well with IspH but was not nearly as active as the *Escherichia coli* flavodoxin.

Not Furnished Upon Filing

Example 11

This example describes the generation of protein fusions of IspG or IspH with their flavodoxin redox partner proteins.

Fusion proteins of the IspG and IspH enzymes with their preferred flavodoxin redox partner were created to increase the local effective concentration of flavodoxin *in vivo*. Fusion proteins have the additional advantages of ensuring one-to-one stoichiometries of interacting enzymes *in vivo*, and of providing a single template for enzyme engineering to enhance enzyme interactions and catalytic properties.

N- and C-terminal fusions were made between Bc_G, Bt_G, and Tt_G with BsyN_fld, and between Ec_H and Ec_fld. To this end, the DNA fragments shown in Table 17 were PCR-amplified as described in Example 1.

Table 17 – PCR Amplified DNA Fragments

DNA Fragment	Primers	Template
For N-terminal Coding Sequences		
Ec_fld	KLD 119 (SEQ ID NO: 186) KLD126 (SEQ ID NO: 187)	SEQ ID NO: 142
Bc_G	KLD 120 (SEQ ID NO: 188) KLD 128 (SEQ ID NO: 189)	SEQ ID NO: 198
Bt_G	KLD 217 (SEQ ID NO: 190) KLD218 (SEQ ID NO: 191)	SEQ ID NO: 199
Tt_G	KLD 195 (SEQ ID NO: 192) KLD 201 (SEQ ID NO: 193)	SEQ ID NO: 200
For C-terminal Coding Sequences		
Ec_H	KLD 144 (SEQ ID NO: 194) KLD150 (SEQ ID NO: 195)	SEQ ID NO: 201
BsyN_fld	KLD 219 (SEQ ID NO: 196) KLD220 (SEQ ID NO: 197)	SEQ ID NO: 202

The DNA fragments were assembled into stitches as outlined in Table 18 and as described in Example 1.

Table 18 - Terminal Primers for Assembly of Stitches

Stitch	DNA Fragments (see Table 20)	Terminal Primer 1	Terminal Primer 2
Ec_fld-Ec_H	Ec_fld, Ec_H	KLD 173 (SEQ ID NO: 203)	KLD 174 (SEQ ID NO: 204)
Bc_G-BsyN_fld	Bc_G, BsyN_fld		
Bt_G-BsyN_fld	Bt_G, BsyN_fld		
Tt_G-BsyN_fld	Tt_G, BsyN_fld		

Finally, Y4743 cells were co-transformed with the purified stitches and linear pGAL1 expression vector, and the fusion proteins were purified as described.

The activities of the fusion proteins were assayed using the *in vitro* enzyme activity assay with the physiological electron donor as described above except only one source of IspG or IspH and

flavodoxin redox partner was needed. The percent production of HTMAPP or DMAPP/IPP after 16 hours was compared to product production by soluble Bt_G with BsyN_fld (the best IspG flavodoxin pair identified in Example 10) or Ec_H with Ec_fld (the best IspH flavodoxin pair identified in Example 10), respectively.

As shown in FIGs. 4A and 4B, Bc_G-BsyN_fld, Bt_G-BsyN_fld, and Tt_G-BsyN_fld, and Ec_fld-Ec_H fusion proteins were active and converted a significant amount of CMEPP into HDMAPP or HDMAPP to DMAPP/IPP, respectively.

Example 12

This example describes methods for measuring the amount of Fe incorporation in DXP pathway enzymes expressed in yeast.

This method was adapted from a reported procedure (Pierik AG, Netz DJ, Lill R. *Nat Protoc* 2009, 4:753-766). Yeast cells were grown in 50 mL of modified CSM comprising 2% glucose until the OD_{600} was greater than 2. At that point, cells were spun down, resuspended in 50 mL of low iron modified CSM comprising 2.7% galactose, and grown overnight. The next morning, the cell cultures were spun down, cell pellets were resuspended in 87 mL of low iron CSM comprising 2.7% galactose, and the resuspended cell culture was split into four different flasks: three 250 mL disposable flasks received 25 mL of cell culture each. 10 µCi of ^{55}Fe ($^{55}FeCl_3$ (PerkinElmer, Boston MA) added to 250 µL of 0.1 M ascorbate solution per flask) was added to each flask, and one 125 mL flask received 12 mL of culture and unlabelled $FeCl_3$ to a final concentration of 300 µM (25 µL of 150 mM $FeCl_3$ was added to 125 µL of 0.1 M ascorbate solution, and 150 µL of the $FeCl_3$ in ascorbate was added to 12 mL of cell culture). All cultures were then grown for 20-24 hours prior to cell harvesting and processing.

For processing, the cells were pelleted, washed in 10 mL of citrate buffer (50 mM citrate, 1 mM EDTA, pH 7.0), pelleted, washed in 1 mL of HEPES buffer (20 mM HEPES/KOH pH 7.4), transferred to a 2 mL centrifuge tube, and pelleted. The HEPES buffer was carefully removed with a pipette and the wet pellet was weighed. 500 µL of cold Buffer W (100 mM Tirs pH 8.0, 150 mM NaCl) with complete EDTA-free protease inhibitor (Roche Applied Science, Indianapolis, IN) was added per 0.3 g of wet pellet weight. 700 µL of resuspended cells (equivalent to 0.3 g of pellet wet weight) was added to 2.0 mL O-ring tubes containing 500 µL of 0.5 mm diameter glass beads and the cells were lysed by bead-beating in a FASTPREP™ instrument (MP Biomedicals, Solon, OH); cells were beat for 40 sec at 6.0 m/s, iced for 5 min, and beat for another 40 sec at 6.0 m/s then iced for 5 min. All subsequent procedures were done on ice or at 4°C. Tubes were centrifuged at max speed for 5 min, and all supernatant was carefully removed and transferred to a pre-chilled 1.5 mL centrifuge tube. Samples were centrifuged at max speed again for 5 min to remove

cellular debris, and the clarified supernatant was carefully transferred to a fresh 1.5 mL centrifuge tube. 15 µg of avidin was added to each sample, gently mixed, and allowed to sit on ice for 30 min; avidin was added to prevent native biotinylated proteins in the cell-free extract from binding the streptactin resin. Clarified supernatant with avidin was then loaded onto a STREP-TACTIN™ spin column (IBA GmbH) according to the manufacturer's suggested protocol. The concentration of protein in the elution was determined using a QUBIT™ reader (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. All remaining protein elution was added into a glass scintillation vial containing 10 mL of scintillation fluid and the radioactivity was detected on a Beckman Coulter LS 6500 scintillation counter. Total counts of radioactivity were normalized to the amount of protein in the sample.

The unlabelled cell culture was prepped in the method described above, and the eluted protein was run out on a polyacrylamide gel to verify that the protein was expressed, soluble, and pure. The endogenous yeast gene LEU1 is a cytosolic Fe-S containing protein, and this protein was also cloned into the 2 micron PGAL1-StrepII N-terminal tag plasmid as a control. In every experiment Y4763 containing LEU1 was cultured, labeled, and extracted as a positive control and to normalize ^{55}Fe labeling among experiments. Y4763 with the PGAL1-StrepII N-terminal tagged plasmid containing no gene insert was run as a negative control to control for background ^{55}Fe signal. No yeast protein was ever detected on a protein gel using this negative control.

As shown in FIG. 5, heterologously-expressed ispG and ispH proteins incorporated labeled ^{55}Fe, suggesting that the endogenous cytosolic Fe-S machinery in *S. cerevisiae* is capable of incorporating the necessary metallic cofactors needed for catalysis in these proteins. Note that the ispG protein from *Escherichia coli* was shown to be insoluble in *S. cerevisiae*, and therefore no signal was detected for ^{55}Fe as no protein was eluted from the StrepII purification column.

Example 13

This example describes methods for the quantitation of complexed iron in the purified IspG and IspH enzymes.

The amount of iron present in purified IspG and IspH protein samples was quantitated using the method reported previously (Fish, W. *Rapid Colormetric Micromethod for the Quantitation of Complexed Iron in Biological Samples*, Methods in Enzymology, 1988, 158:357-364). Protein concentration was determined using the QUBIT™ Fluorometer (P/N Q32860, Invitrogen, Chicago, IL) and the QUANT-IT™ Protein Assay Kit (P/N Q33212, Invitrogen, Chicago, IL) per the manufacturer's instructions. Glassware used for reagent preparation was de-mineralized by acid treatment, and then rinsed in Hyclone ultra pure water, WFI grade (P/N 16750-108, VWR, Pittsburgh, PA). All reagents were made using Hyclone ultra pure water, WFI grade (P/N 16750-

108, VWR, Pittsburgh, PA). The standard for Fe calibration was prepared as follows: a 0.041% solution of Iron (II) ethylenediammonium sulfate tetrahydrate solution was serially diluted in 0.01 N HCl to a range of 0.1875 µg/mL-6 µg/mL. Purified protein samples were diluted, if necessary, in Protein Diluent (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2.5 mM Biotin). A blank sample to measure reagent background was prepared at the same time using 600 µL of Protein Diluent in place of the purified protein. To release the bound iron, 300 µLof Reagent A (0.6 N HCL, 2.25% (w/v) $KMnO_4$) was added to 600 µL of purified protein or blank solution in an Eppendorf tube. 250 µL of Reagent A was also added to 600 µL of each of the standard solutions. Samples were then incubated for two hours at 60°C. After incubation, 0.06 mL of Reagent B (6.5 mM Ferrozine (P/N 82950, Sigma-Aldrich, Inc., Atlanta, CA), 13.1 mM neocuprione (P/N 72080, Sigma-Aldrich, Inc., Atlanta, GA), 2 M ascorbic acid, 5 M ammonium acetate) was added to each sample. Samples were then incubated at room temperature (22-25°C) for 30 minutes while protected from light. Absorbance was read at 562 nM (A_{562}) on a GENESYS™ 10 spectrophotometer (Thermo Scientific, Asheville, NC) using disposable polystyrene cuvettes. A calibration curve was established and the iron concentrations in the protein samples were calculated from their A_{562} based on the calibration curve.

As shown in FIG. 6, the heterologously-expressed IspG and IspH proteins contained approximately 1 mol Fe per mol of protein. The amount of Fe contained in IspG and IspH proteins was lower than what is expected for a protein with four Fe-S clusters. However, the fact that the native yeast [4Fe-4S] enzyme Leu1, which also contained approximately 1 mol Fe per mol of protein, contained the same amount of Fe as isolated IspG and IspH, suggests that the low level of Fe-S incorporation is due to the limited capacity of the yeast Fe-S cluster assembly machinery.

Example 14

This example describes the detection of IspG metabolites from *S. cerevisiae* strains bearing heterologous DXP pathways.

In the reported *in vitro* experiments, IspG homologues exhibited lower catalytic activity than IspH homologues even though both enzymes contain a [4Fe-4S] cluster. The results point to the IspG-catalyzed conversion of cMEPP to HDMAPP as a likely bottleneck in the DXP pathway in yeast host cells.

To determine whether the *in vivo* activity of the IspG homologues was indeed reduced, strains Y4819, Y6283, Y6291, and Y6292 were induced for IspG expression, and then analyzed for intracellular HDMAPP. To this end, the strains were inoculated into 3 mL of CSM comprising 2% (w/v) glucose and lacking uracil, and grown overnight at 30°C, with shaking. Fresh 50 mL of CSM comprising 2% glucose and lacking uracil was inoculated with 3 mL of the overnight cultures, and

the resultant cultures were allowed to grow at 30°C until they reached an OD_{600} of between 1 and 4. The cultures were spun down at 5,000xg for 10 min to separate cells from the supernatant, the cells were transferred to 50 mL of CSM comprising 4% (w/v) galactose and lacking uracil, and the cultures were incubated for an additional 18-22 hours at 30°C, with shaking. Metabolite extraction and analysis of HDMAPP were carried out as described in Example 3.

As shown in FIG. 7, no HDMAPP accumulation was detected in the control strain, Y4819, which only contains the DXP-TOP pathway. A small amount of HDMAPP was detected in Y6283, which harbors a complete DXP pathway. Much higher levels of HDMAPP (approximately 10 µM) were observed in strains Y6291 and Y6292, which lack an IspH gene but express *B. subtilis* flavodoxin YkuN along with Bt-G and Tt-G, respectively.

To further verify the chemical identity of an HDMAPP signal observed in the strains depicted in FIG. 7, strains Y6291 and Y6289 were grown on fully ^{13}C-labeled sugar as a carbon source, the incorporation of which shifts the HDMAPP metabolite accumulated in these strains by 5 atomic mass units (AMU) from 260.916 to 265.916 in the LC/MS analysis. Strains Y4819, Y6291, and Y6289 were inoculated into 3 mL of CSM comprising 2% (w/v) ^{13}C-glucose and lacking uracil, and grown overnight at 30°C with shaking. A fresh 10 mL aliquot of CSM comprising 2% ^{13}C-glucose and lacking uracil was inoculated with 1 mL of the overnight cultures, and the resultant cultures were allowed to grow at 30°C until they reached an OD_{600} between 1 and 4. A 2 mL aliquot of CSM comprising 4% (w/v) ^{13}C-galactose and lacking uracil was then inoculated with the above cultures to a final OD_{600} of 0.03 and allowed to grow for an additional 48 hours at 30°C, with shaking. Metabolite extraction and analysis of HDMAPP were carried out as previously described. No detectable ^{13}C-labeled HDMAPP signal was observed in Y4819. ^{13}C-labeled HDMAPP signal was also missing in Y6289 and Y6291 when they were cultured in ^{13}C-labeled glucose (IspG expression was not induced). In contrast, ^{13}C-labeled HDMAPP accumulated in Y6289 and Y6291 when ^{13}C-labeled galactose was used as the carbon source to induce IspG and YkuN expression.

Example 15

This example describes the detection of FPP production via a heterologous DXP pathway in *S. cerevisiae*.

To determine whether the heterologous DXP pathway is functional, yeast strains harboring a full DXP pathway and a deleted ERG13 coding sequence were grown on ^{13}C-labeled media supplemented with ^{12}C-mevalonate. Under the cultivation conditions, unlabeled IPP/DMAPP metabolites were only made from the supplemented ^{12}C-mevalonate via the mevalonate pathway. A functional DXP pathway, on the other hand, would produce only ^{13}C-labeled DMAPP/IPP. FPP, which is made by condensing 1 molecule of DMAPP and 2 molecules of IPP, may

therefore exist in either an unlabeled form (mass = 381.065), partially labeled forms (mass + 5 AMU = 386.065, and mass + 10 AMU = 391.065), or a fully labeled form (mass + 15 AMU = 397.055). Detection of any labeled form of FPP (*i.e.* mass + 5 AMU, mass + 10 AMU, or mass + 15 AMU) would demonstrate that the heterologous DXP pathway was capable of providing flux to isoprenoids. The ratio of ^{13}C-labeled FPP to unlabeled FPP was determined by the level of DXP flux in yeast relative to the native mevalonate pathway flux (mevalonate concentrations in the growth medium).

Strains Y6540 and Y6533 were inoculated into 3 mL of CSM comprising 2% (w/v) ^{13}C-glucose and lacking adenine, and grown overnight at 30°C with shaking. Fresh 10 mL of CSM comprising 2% ^{13}C-glucose and lacking adenine was inoculated with 1 mL of the overnight cultures, and the resultant cultures were allowed to grow at 30°C until they reached an OD_{600} of between 1 and 4. The cultures were then equally split into two flasks, and grown in the presence or absence of zaragozic acid (a squalene synthase inhibitor). A final concentration of 16 μg/mL of zaragozic acid was added to one flask while the same volume of water (6 μL) was added to the other. Four of the resultant cultures were then allowed to grow for an additional 18-22 hours at 30°C, with shaking. Metabolite extraction and analysis of DMAPP and FPP were carried out as described in Example 3.

As shown in FIG. 8, the intracellular DMAPP/IPP metabolite concentrations were extremely low in these strains. No ^{13}C-labeled DMAPP/IPP was detected in either strain. In cultures without zaragozic acid, the level of ^{12}C-labeled FPP was very low, and no ^{13}C-labeled FPP was detected. Partially labeled FPP (mass + 5 AMU or 386.065) was detected in Y6533 only when zaragozic acid was added to cause FPP to accumulate by inhibiting the FPP-consuming enzyme squalene synthase. In contrast, Y6540, in which the expression of IspG, IspH, and redox proteins is only induced by the addition of galactose, did not accumulate any ^{13}C-labeled FPP even in the presence of zaragozic acid. It was estimated that the FPP + 5 AMU signal was at approximately 1% of the total FPP, indicating that the DXP flux is about 100-fold less active than the native mevalonate pathway. Since labeled FPP of mass +10 AMU and +15 AMU is made via condensation of 2 or 3 ^{13}C-labeled DMAPP/IPP units, respectively, these labeled metabolites are expected to be below the limit of LC/MS detection, which explains why only the FPP +5 form of the ^{13}C-labeled FPP was detected.

The foregoing examples of the present invention have been presented for purposes of illustration and description. Furthermore, these examples are not intended to limit the invention to the form disclosed herein. Consequently, variations and modifications commensurate with the teachings of the description of the invention, and the skill or knowledge of the relevant art, are

within the scope of the present invention. The specific embodiments described in the examples provided herein are intended to further explain the best mode known for practicing the invention and to enable others skilled in the art to utilize the invention in such, or other, embodiments and with various modifications required by the particular applications or uses of the present invention. It is intended that the appended claims be construed to include alternative embodiments to the extent permitted by the prior art.

CLAIMS

1. An iron sulfur enzyme system comprising one or more nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) an iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

2. An iron sulfur enzyme system comprising one or more nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) a iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting cMEPP to isopentenyl pyrophosphate ("IPP") and dimethylallyl pyrophosphate ("DMAPP").

3. A genetically modified microorganism comprising one or more heterologous nucleic acid molecules comprising one or more nucleic acid sequences encoding:

(a) a first reductase that is capable of accepting electrons from an electron donor thereby converting the first reductase from an oxidized state ("oxidized first reductase") to a reduced state ("reduced first reductase");

(b) a first redox protein that is capable of accepting electrons from the reduced first reductase thereby converting the first redox protein from an oxidized state ("oxidized first redox protein") to a reduced state ("reduced first redox protein"); and

(c) a first iron-sulfur enzyme that is capable of accepting electrons from the first reduced redox protein thereby converting the first iron-sulfur enzyme from an oxidized state ("oxidized first iron-sulfur enzyme") to a reduced state ("reduced first iron-sulfur enzyme") and wherein the reduced first iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

4. The genetically modified microorganism of claim 3 wherein the one or more heterologous nucleic acid molecules further encode:

(d) a second reductase that is capable of accepting electrons from an electron donor thereby converting the second reductase from an oxidized state ("oxidized second reductase") to a reduced state ("reduced second reductase");

(e) a second redox protein that is capable of accepting electrons from the reduced second reductase thereby converting the second redox protein from an oxidized state ("oxidized second redox protein") to a reduced state ("reduced second redox protein"); and

(f) a second iron-sulfur enzyme that is capable of accepting electrons from the second reduced redox protein thereby converting the second iron-sulfur enzyme from an oxidized state ("oxidized second iron-sulfur enzyme") to a reduced state ("reduced second iron-sulfur enzyme") and wherein the reduced second iron-sulfur enzyme is capable of converting HMBPP to isopentenyl pyrophosphate ("IPP") or/and dimethylallyl pyrophosphate ("DMAPP").

5. The genetically modified microorganism of claim 3 wherein the first iron-sulfur enzyme is not IspG from *E. coli*.

6. The genetically modified microorganism of claim 4 wherein the first and second reductases are the same reductase.

7. The genetically modified microorganism of claim 4 wherein the first and second redox proteins are the same redox protein.

8. The genetically modified microorganism of claim 4 wherein the first and second reductases are the same reductase and the first and second redox proteins are the same redox protein.

9. The genetically modified microorganism of claim 3 wherein the first redox protein and the first iron-sulfur enzyme form a contiguous polypeptide.

10. The microorganism of claim 4 wherein the second redox protein and the second iron-sulfur enzyme form a contiguous polypeptide.

11. The genetically modified microorganism of claims 3 or 4 wherein the microorganism does not include an endogenous DXP pathway.

12. The genetically modified microorganism of claims 3 or 4 wherein the microorganism is a eukaroyote.

13. The genetically modified microorganism of claims 3 or 4 wherein the microorganism is fungi.

14. The genetically modified microorganism of claims 3 or 4 wherein the microorganism is yeast.

15. The genetically modified microorganism of claims 3 or 4 wherein the microorganism is *S. cerevisiae*.

16. The genetically modified microorganism of claims 3 or 4 wherein the electron donor is NADPH.

17. The genetically modified microorganism of claims 3 or 4 wherein the first redox protein is a flavodoxin.

18. The genetically modified microorganism of claim 17 wherein the flavodoxin is from *B. subtilis*.

19. The genetically modified microorganism of claim 17 wherein the flavodoxin is YkuN from *B. subtilis*.

20. The genetically modified microorganism of claim 4 wherein the second redox protein is a flavodoxin from *E. coli*.

21. The genetically modified microorganism of claim 4 wherein the second iron-sulfur cluster protein is IspH from *E. coli*.

22. The genetically modified microorganism of claims 3 or 4 wherein the one or more heterologous nucleic acids also encode one or more proteins of the ISC assembly machinery.

23. The genetically modified microorganism of claim 22 wherein the one or more heterologous nucleic acids encode one or more of the ISC assembly machinery proteins selected from Nfs1, Yfh1, Arh1, Yah1, POS5, Isu1, Isu2, Ss1-ATP, Jac1 and Grx5.

24. The genetically modified microorganism of claims 3 or 4 wherein the one or more heterologous nucleic acids also encode one or more proteins of the ISC export machinery.

25. The genetically modified microorganism of claim 24 wherein the one or more heterologous nucleic acids encode one or more of the ISC export machinery proteins selected from Atm1 and Erv1.

26. The genetically modified microorganism of claims 3 or 4 wherein the one or more heterologous nucleic acids also encode one or more proteins of the CIA machinery.

27. The genetically modified microorganism of claim 26 wherein the one or more heterologous nucleic acids also encode one or more proteins of the CIA machinery selected from Cfd1-Nbp35, Nar1, Cia1, Cia2, and Dre2.

28. A genetically modified yeast comprising one or more heterologous nucleic acid molecules that encode the following enzymes:

(a) Dxs;

(b) Dxr;

(c) IspD;

(d) IspE;

(e) IspF;

(f) a flavodoxin/ferrodoxin $NADP^+$-reductase reductase ("FNR");

(g) a first flavodoxin that is compatible with the FNR; and,

(h) IspG that is compatible with the first flavodoxin such the enzymes collectively are capable of converting pyruvate and D-glyceraldehyde-3-phosphate to 1-hydroxy-2-methyl-2-(E)-butenyl- 4-diphosphate in the cytoplasm of said yeast,

wherein the IspG is not IspG from *E. coli*.

29. The genetically modified yeast of claim 28 further comprising:

(i) IspH that is compatible with the first flavodoxin such that the enzymes collectively are capable of converting pyruvate and D-glyceraldehyde-3-phosphate to isopentenyl pyrophosphate ("IPP") and dimethylallyl pyrophosphate ("DMAPP").

30. The genetically modified yeast of claim 28 further comprising:

(i) a second flavodoxin that is compatible with the FNR;

(j) IspH that is compatible with the second flavodoxin such that the enzymes collectively are capable of converting pyruvate and D-glyceraldehyde-3-phosphate to isopentenyl pyrophosphate ("IPP") and dimethylallyl pyrophosphate ("DMAPP").

31. The genetically modified yeast of claims 28, 29, or 30 wherein one or more of Dxs, Dxr, IspD, IspE, and Isp F are *E. coli* enzymes.

32. The genetically modified yeast of claim 31 wherein all of Dxs, Dxr, IspD, IspE, and IspF are *E. coli* enzymes.

33. A microorganism capable of producing an isoprenoid compound comprising:

(1) one or more heterologous nucleic acid molecules comprising one or more nucleic acid sequences encoding an enzyme selected from the group consisting of:

(a) an enzyme that condenses pyruvate with D-glyceraldehyde 3-phosphate to make 1-deoxy-D-xylulose-5-phosphate ("DXP"),

(b) an enzyme that converts DXP to 2C-methyl-D-erythritol-4-phosphate ("MEP"),

(c) an enzyme that converts MEP to 4-diphosphocytidyl-2C-methyl-D-erythritol ("CDP-ME"),

(d) an enzyme that converts CDP-ME is converted to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate ("CDP-MEP")

(e) an enzyme that converts CDP-MEP is converted to 2C-methyl-D-erythritol 2, 4-cyclodiphosphate ("cMEPP")

(f) an enzyme that converts , cMEPP is converted to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP")

(g) an enzyme that converts converted to isopentyl pyrophosphate ("IPP") and/or its isomer, dimethylallyl diphosphate ("DMAPP"), and

(h) an enzyme that converts an enzyme capable of interconverting IPP and DMAPP; AND

(2) one or more heterologous nucleic acid molecules comprising one or more nucleic acid sequences encoding an enzyme selected from the group consisting of:

(a) a reductase that is capable of accepting electrons from an electron donor thereby converting the reductase from an oxidized state ("oxidized reductase") to a reduced state ("reduced reductase");

(b) a redox protein that is capable of accepting electrons from the reduced reductase thereby converting the redox protein from an oxidized state ("oxidized redox protein") to a reduced state ("reduced redox protein"); and

(c) an iron-sulfur enzyme that is capable of accepting electrons from the reduced redox protein thereby converting the iron-sulfur enzyme from an oxidized state ("oxidized iron-sulfur enzyme") to a reduced state ("reduced iron-sulfur enzyme") and wherein the reduced iron-sulfur enzyme is capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP").

34. The microorganism of Claim 33, wherein the microorganism comprises an endogenous DXP metabolic pathway.

35. The microorganism of Claim 33, wherein the microorganism does not contain an endogenous DXP metabolic pathway.

36. The microorganism of Claim 33, wherein at least one heterologous enzyme is encoded by a plurality of nucleic acid molecules.

37. The microorganism of Claim 33, wherein the at least one heterologous nucleic acid molecule is under inducible control.

38. The microorganism of Claim 33, wherein at least one heterologous nucleic acid is integrated into a chromosome.

39. The microorganism of Claim 33, wherein at least one heterologous nucleic acid is on a plasmid.

40. The microorganism of Claim 33, wherein at least one heterologous enzyme is encoded by one recombinant nucleic acid molecule.

41. The microorganism of Claim 33, wherein at least one heterologous enzyme is encoded by a plurality of nucleic acid molecules.

42. The microorganism of Claim 33, wherein at least one heterologous nucleic acid molecule is under inducible control.

43. The microorganism of Claim 33, wherein the microorganism is a yeast.

44. The microorganism of Claim 33, wherein the microorganism is *Saccharomyces cerevisiae*.

45. The microorganism of Claim 33, wherein the microorganism is a fungus.

46. The microorganism of Claim 33, wherein at least one heterologous enzyme is a bacterial enzyme.

47. The microorganism of Claim 33, wherein each of the heterologous enzymes is a eukaryotic enzyme.

48. The microorganism of Claim 33, wherein one of the heterologous enzymes is an IspG protein selected from the group consisting of: *Algoriphagus* sp. PR1 (ZP_01720138.1); *Arthrospira platensis str. Paraca* (ZP_06384238.1); *Chlamydomonas reinhardtii* (XP_001690937.1); *Brevibacillus brevis* NBRC 100599 (YP_002772921.1); *Catharanthus roseus* (AAO24774.1); *Crocosphaera watsonii* WH 8501 (ZP_00518798.1); *Cyanobium* sp. PCC 7001 (ZP_05045928.1); *Eubacterium dolichum* DSM 3991 (ZP_02077259.1); *Nostoc punctiforme* PCC 73102 (YP_001868335.1); *Ostreococcus tauri* (CAL55258); *Prevotella buccae* D17 (EFC75705); *Prochlorococcus marinus subsp. pastoris str*. CCMP1986 (NP_892794.1); *Porphyra yezoensis* red algae (ACI45961); *Sorghum bicolor* (XP_002454137); *Synechococcus* sp. RS9917 (ZP_01081635.1); *Thermosynechococcus elongatus* BP-1 (NP_681786); and *Vitis vinifera* (XP_002285130.1).

49. The microorganism of Claim 33, wherein one of the heterologous enzymes is an IspH protein selected from the group consisting of: *Escherichia coli* (AAC73140); *Fibrobacter succinogenes subsp. succinogenes* S85 (YP_003250012.1); Physcomitrella *patens subsp. patens* (XP_001758369.1); *Arabidopsis thaliana* (AY168881); *Sorghum bicolor* (XP_002463933.1); *Stevia rebaudiana* (ABB88836.2); and *Thermosynechococcus elongatus* BP-1 (NP_681832).

50. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is an Isp G or an IspH protein comprising a targeting sequence that directs the intracellular location of the heterologous enzyme into a yeast mitochondrion.

51. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is a ferrodoxin or a flavodoxin protein.

52. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is a [2Fe-2S] dependent redox protein.

53. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is a flavodoxin selected from the group consisting of: *Alteromonas macleodii* ATCC 27126 (ZP_04714638.1); *Bacillus coagulans* 36D1 (ZP_04431144.1); *Bacillus subtilis subsp. subtilis* str. 168 YkuP (NP_389300.2); *Crocosphaera watsonii* WH 8501 (ZP_00515759.1); *Desulfovibrio vulgaris subsp. vulgaris* DP (YP_966026.1); *Escherichia coli* str. K-12 substr. MG1655 (NP_415210); *Nostoc punctiforme* PCC 73102 (YP_001866434.1); *Anabaena* PCC 7119 (AAB20462); *Wolinella succinogenes* DSM 1740 (NP_906993.1); and *Bacillus subtilis subsp. subtilis* str. 168 YkuN (NP_389298.1).

54. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is a redox protein selected from the group consisting of: *Escherichia coli* str. K-12 substr. MG1655 (NP_415210) and *Bacillus subtilis subsp. subtilis str*. 168 YkuN (NP_389298.1).

55. The microorganism of Claim 33, wherein at least one of the heterologous enzymes is an FNRs selected from the group consisting of: *Anabaena* PCC 7119 (1QUE_A); *Arabidopsis thaliana* (root) (NP_849734.1); *Escherichia coli* str. K-12 substr. MG1655 (NP_418359); *Nostoc* sp. PCC 7906 (BAG69182); *Spinacia oleracea* (P00455); *Thermosynechococcus elongatus* BP-1 (NP_682001); and *Zea mays* (3LVB_A).

56. The microorganism of Claim 33, further comprising at least one heterologous enzyme selected from the group consisting of: ISC assembly machinery protein, an ISC export machinery protein, and a cytosolic CIA machinery protein.

57. The microorganism of Claim 33, further comprising at least one heterologous enzyme selected from the group consisting of: cysteine desulfurase Nfs1; iron binding protein frataxin Yfh1; redox proteins Arh1 and Yah1; NADH kinase POS5; scaffold proteins Isu1 and Isu2; and chaperones Ss1-ATP, Jac1 and Grx5.

58. The microorganism of Claim 33, further comprising at least one heterologous enzyme selected from the group consisting of: ABC transporter Atm1; the sulfhydryl oxidase Erv1, and glutathione.

59. The microorganism of Claim 33, further comprising at least one heterologous enzyme selected from the group consisting of: Cfd1-Nbp35; Nar1, Cia1 and Cia2; and Dre2.

60. The microorganism of Claim 33, further comprising a genetic change that decreases the activity of one or more metabolic pathways that divert intermediates away from the production of isoprenoid compounds.

61. The microorganism of Claim 34, wherein the genetic change results in lower or absent squalene synthase enzymatic activity relative to wild type yeast.

62. The microorganism of Claim 61, wherein the genetic change results in lower or absent enzymatic activity of enzymes that catalyze reactions of the TCA cycle that lead to fatty acid biosynthesis, alanine biosynthesis, the aspartate superpathway, gluconeogenesis, heme biosynthesis, and/or glutamate biosynthesis.

63. The microorganism of Claim 61, wherein the genetic change is a disruption of an endogenous gene.

64. The microorganism of Claim 61, wherein the genetic change is a repression of an endogenous gene by a heterologous repressible promoter.

65. The microorganism of Claim 61, wherein the genetic change is an expression of at least one antisense gene encoding a fragment of a protein having enzymatic activity of enzymes that catalyze reactions of the TCA cycle that lead to fatty acid biosynthesis, alanine biosynthesis, the aspartate superpathway, gluconeogenesis, heme biosynthesis, and/or glutamate biosynthesis.

66. The microorganism of Claim 33, further comprising a recombinant nucleic acid molecule encoding at least one enzyme that increases the carbon flux through geranyl pyrophosphate, farnesyl pyrophosphate, or geranylgeranyl pyrophosphate.

67. The microorganism of Claim 33, further comprising a recombinant nucleic acid molecule encoding at least one enzyme that encodes at least one enzyme in the synthesis of isopentenyl pyrophosphate from an isoprenoid pathway.

68. The microorganism of Claim 33, further comprising a recombinant nucleic acid molecule encoding at least one enzyme that encodes at least one enzyme that increases the supply of a precursor to an isoprenoid biosynthetic pathway, wherein the precursor is selected from the group consisting of pyruvate, and D-glyceraldehyde-3-phosphate.

69. The microorganism of Claim 68, wherein the at least one enzyme that increases the supply of a precursor to an isoprenoid biosynthetic pathway is selected from the group consisting of a pyruvate kinase and a triphosphate isomerase.

70. A fusion protein of an IspG protein and a flavodoxin redox partner protein.

71. A fusion protein of an IspH protein and a flavodoxin redox partner protein.

72. A nucleic acid expression vector comprising at least one nucleic acid sequence encoding a fusion protein selected from the group consisting of: a fusion protein of an IspG protein and a flavodoxin redox partner protein and a fusion protein of an IspH protein and a flavodoxin redox partner protein.

73. A method to produce an isoprenoid comprising:

(a) culturing a microorganism of any one of Claims 33-69 in a fermentation medium to produce the product, wherein the microorganism comprises at least one recombinant nucleic acid molecule encoding at least one heterologous enzyme of the DXP pathway and at least one iron-sulfur enzyme capable of converting 2C-methyl-D-erythritol-2,4-cyclodiphosphate ("cMEPP") to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate ("HMBPP"); and

(b) recovering the product.

Figure 1

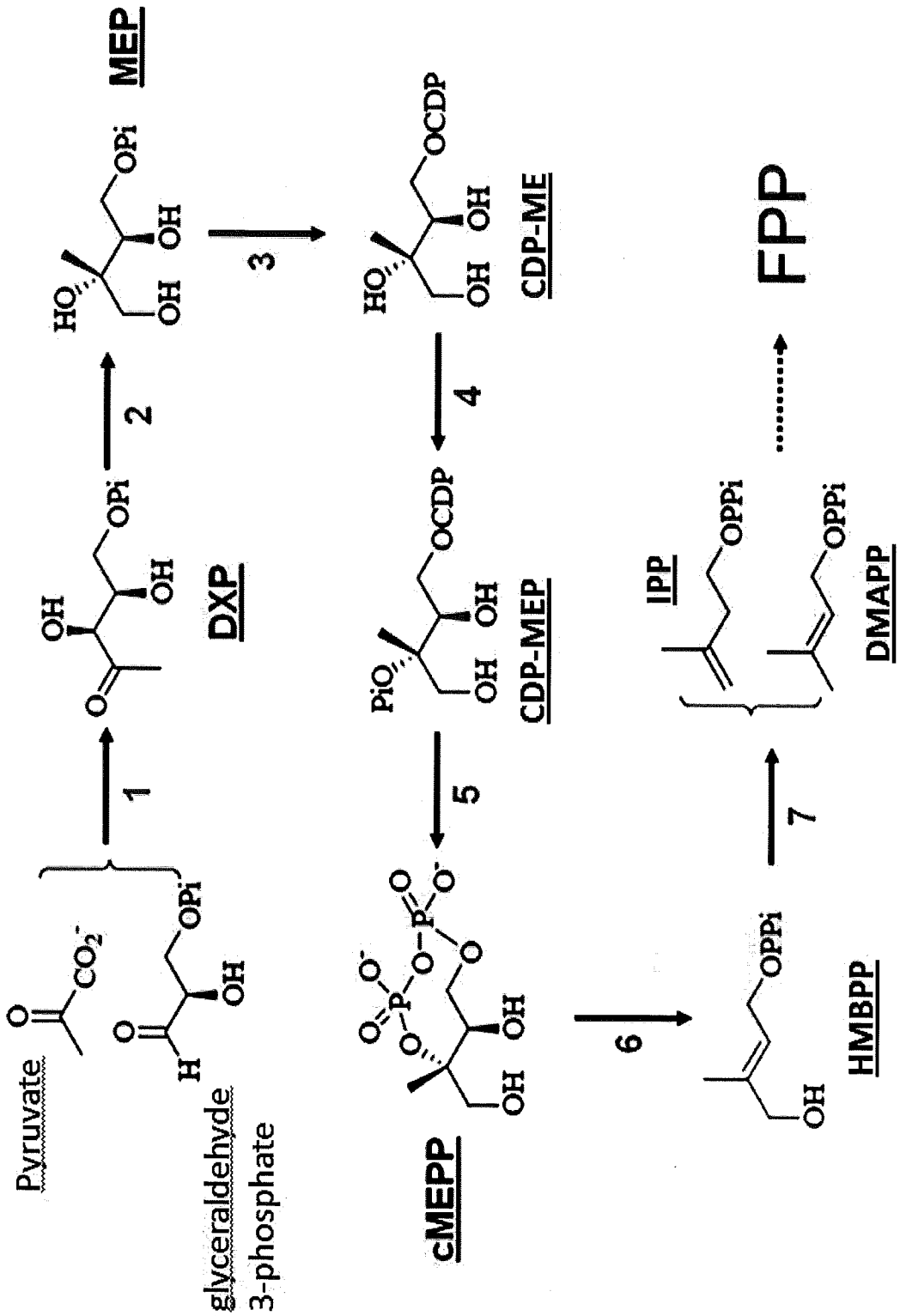

Pyruvate
glyceraldehyde 3-phosphate
DXP
MEP
CDP-ME
CDP-MEP
cMEPP
HMBPP
IPP
DMAPP
FPP
1
2
3
4
5
6
7

Figure 2

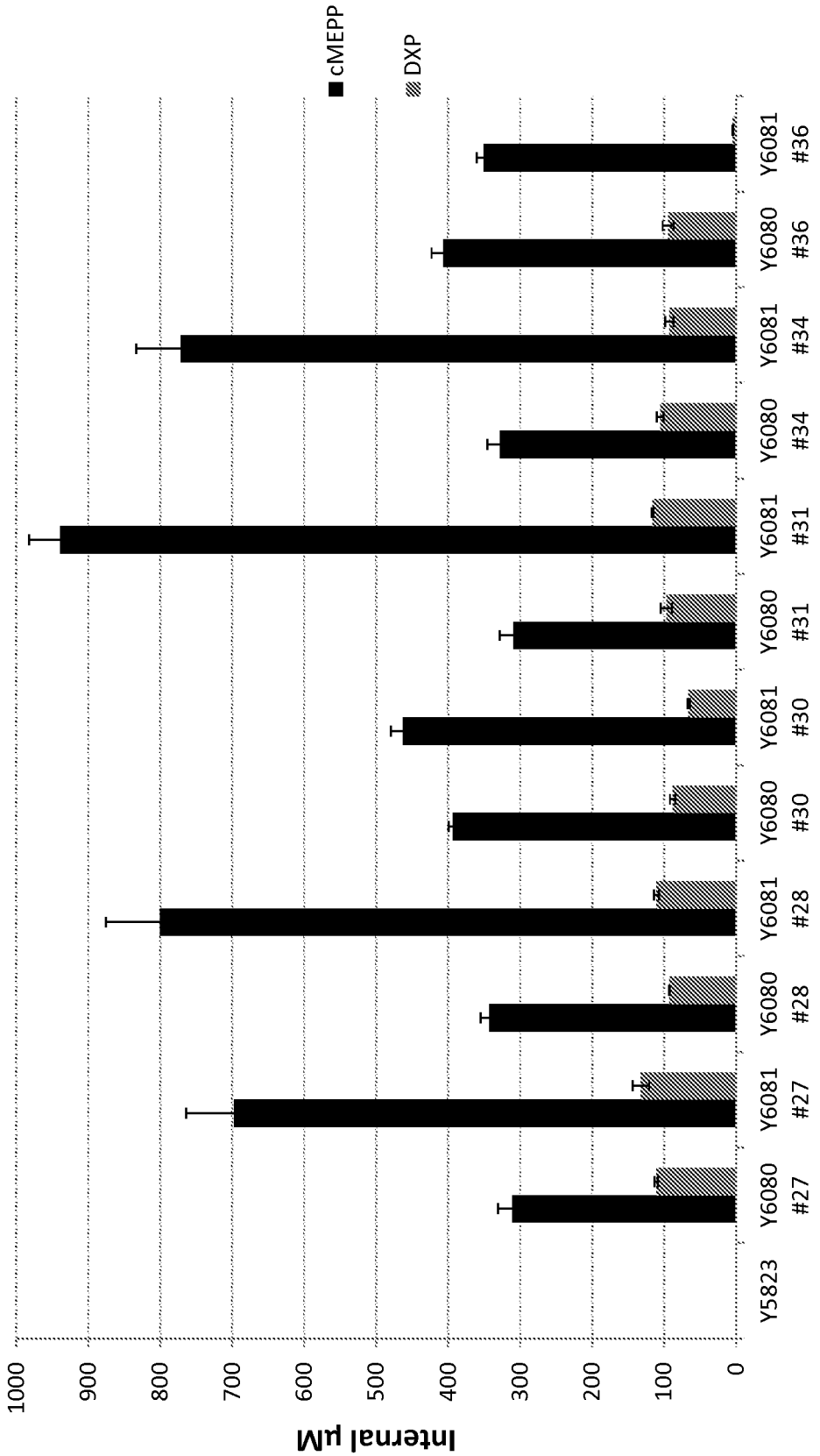

cMEPP
DXP
Internal μM
0
100
200
300
400
500
600
700
800
900
1000
Y5823
Y6080 #27
Y6081 #27
Y6080 #28
Y6081 #28
Y6080 #30
Y6081 #30
Y6080 #31
Y6081 #31
Y6080 #34
Y6081 #34
Y6080 #36
Y6081 #36

Figure 3

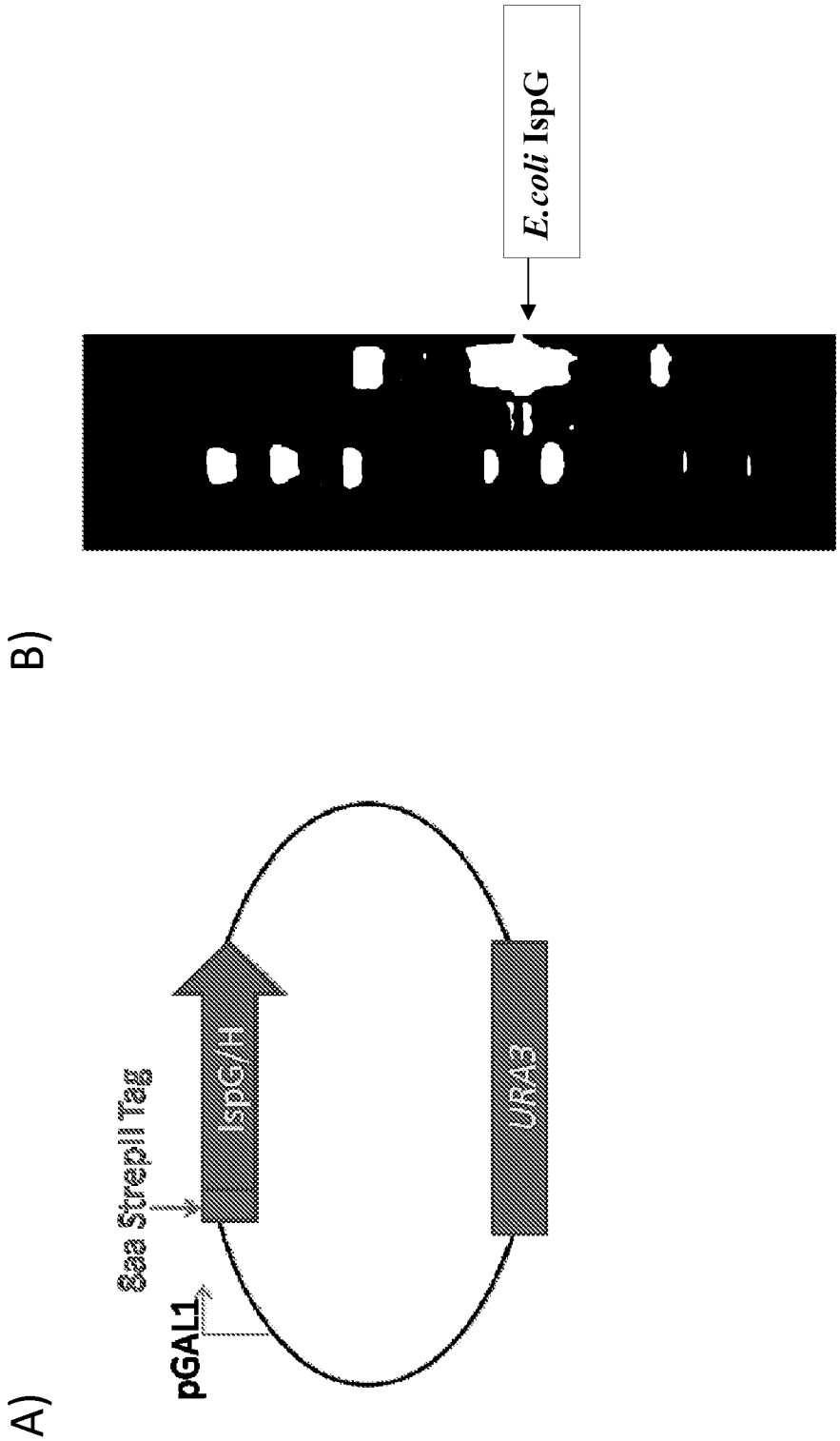


Figure 4

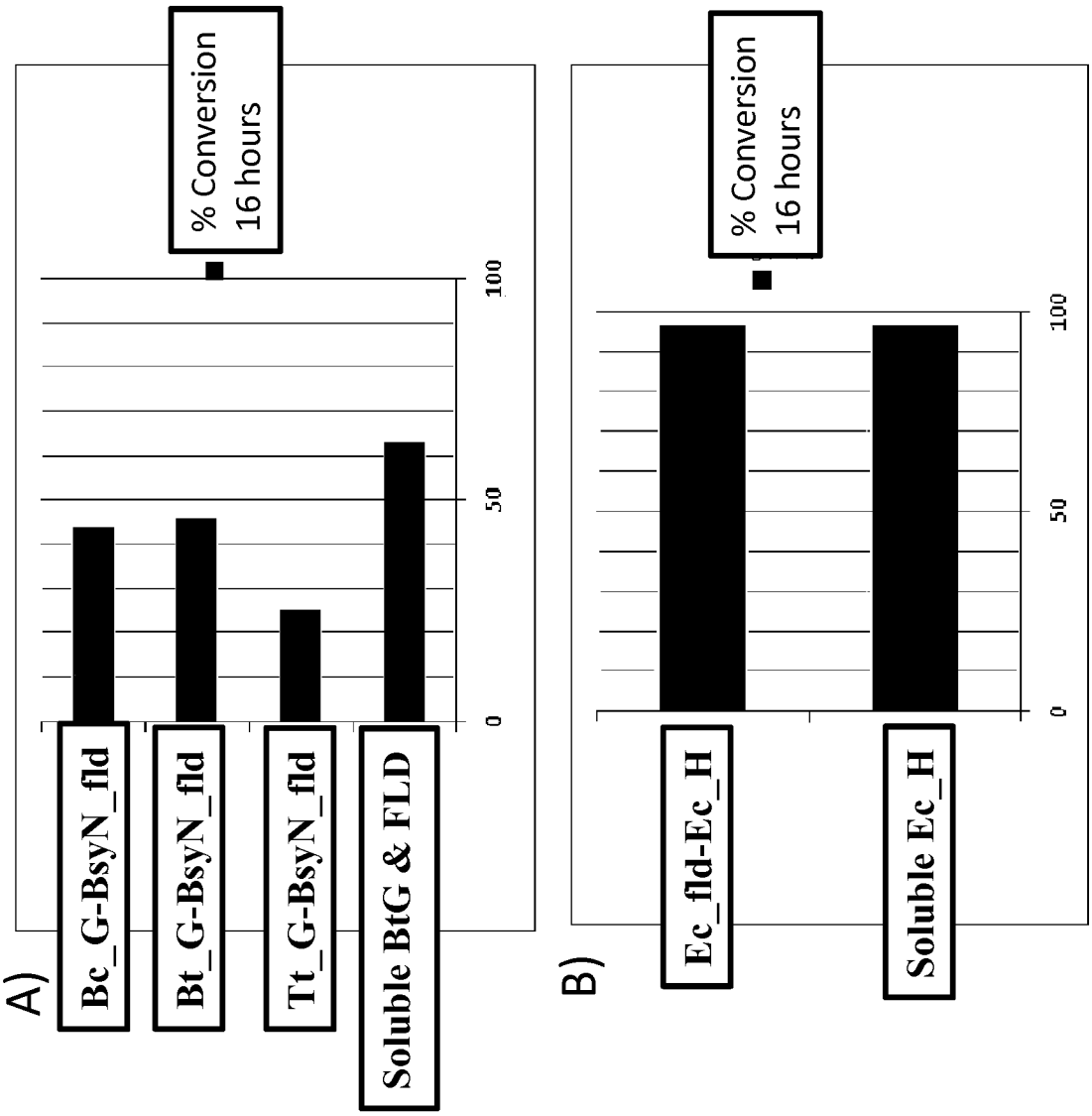

A)
% Conversion 16 hours
Bc_G-BsyN_fld
Bt_G-BsyN_fld
Tt_G-BsyN_fld
Soluble BtG & FLD
0
50
100
B)
% Conversion 16 hours
Ec_fld-Ec_H
Soluble Ec_H
0
50
100

Figure 5

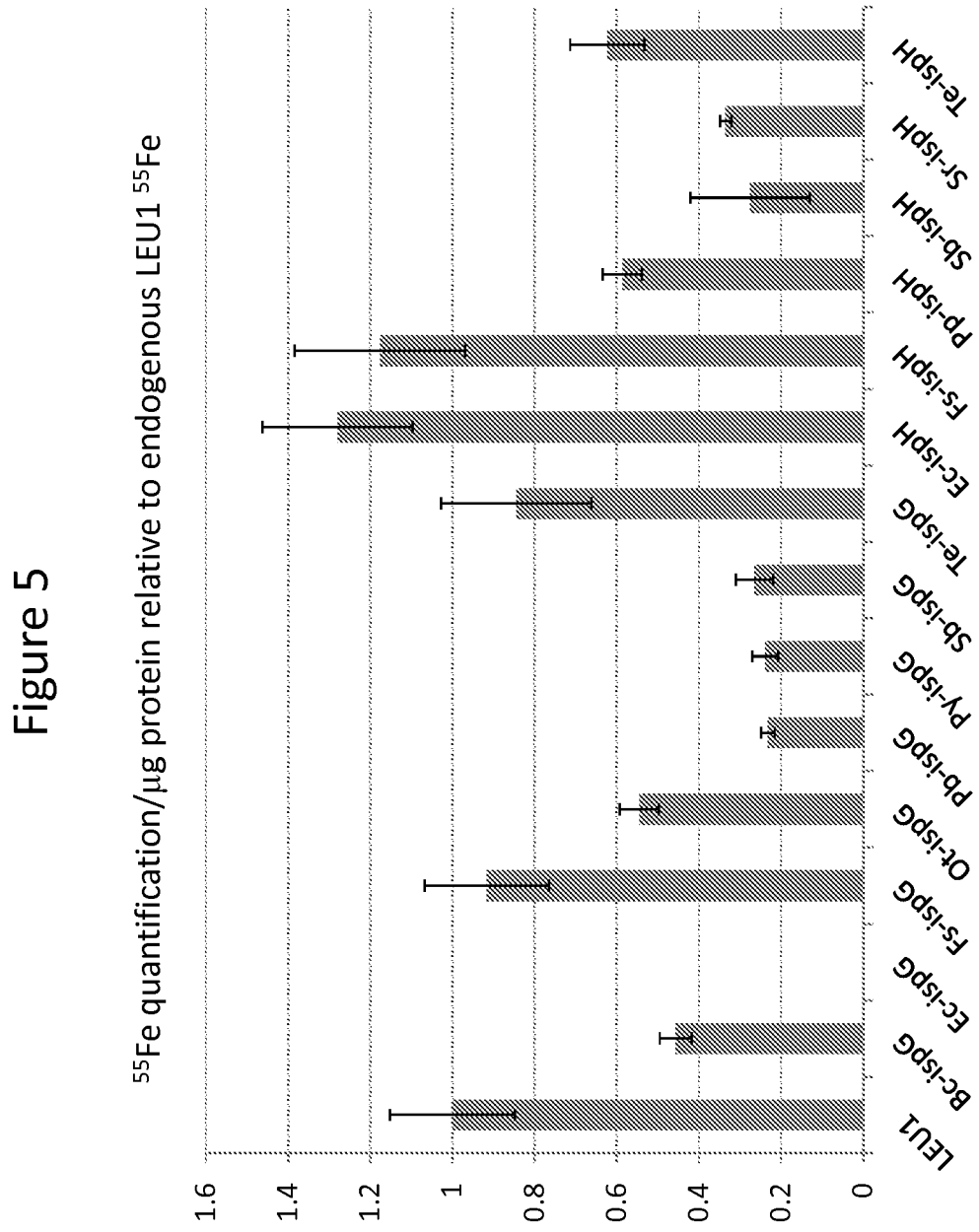

^{55}Fe quantification/µg protein relative to endogenous LEU1 ^{55}Fe
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
LEU1
Bc-ispG
Ec-ispG
Fs-ispG
Ot-ispG
Pb-ispG
Py-ispG
Sb-ispG
Te-ispG
Ec-ispH
Fs-ispH
Pp-ispH
Sb-ispH
Sr-ispH
Te-ispH

Figure 6

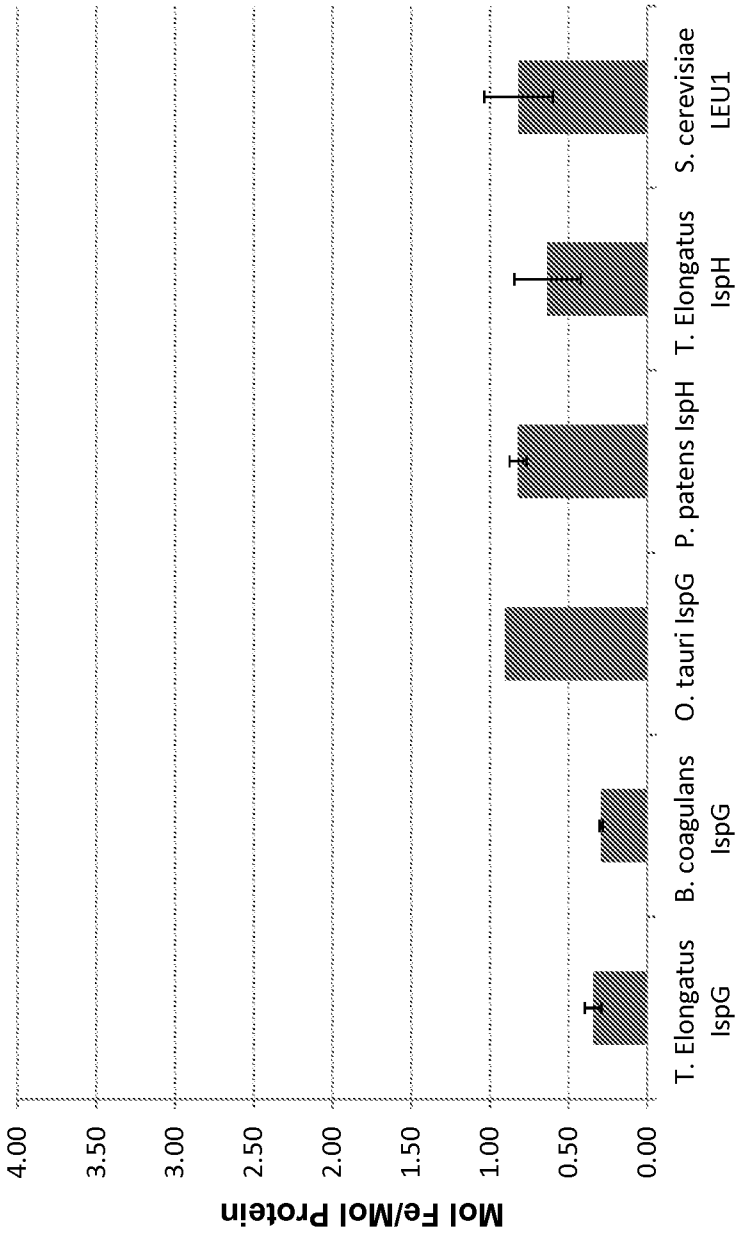

Mol Fe/Mol Protein
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
T. Elongatus IspG
B. coagulans IspG
O. tauri IspG
P. patens IspH
T. Elongatus IspH
S. cerevisiae LEU1

Figure 7

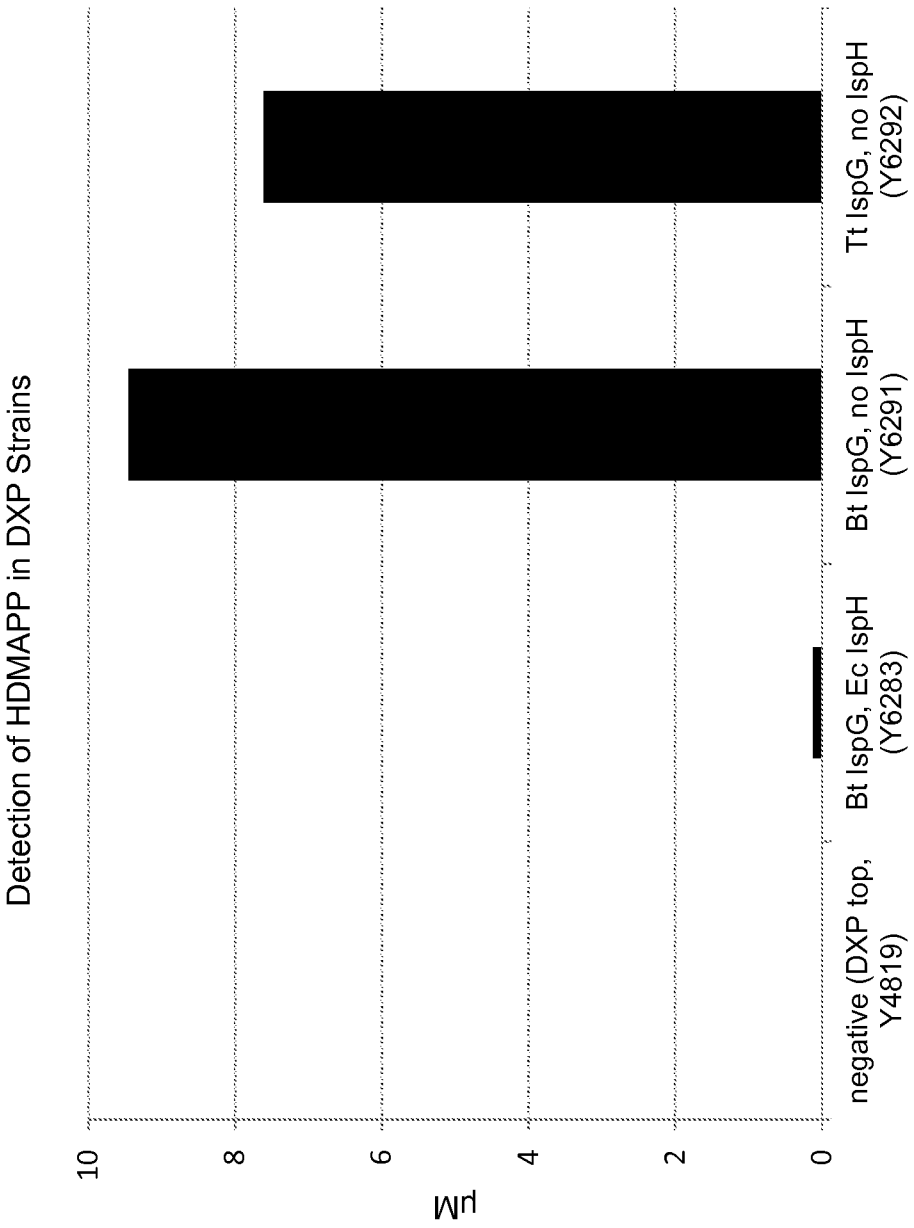

Detection of HDMAPP in DXP Strains
μM
0
2
4
6
8
10
negative (DXP top, Y4819)
Bt IspG, Ec IspH (Y6283)
Bt IspG, no IspH (Y6291)
Tt IspG, no IspH (Y6292)

Figure 8

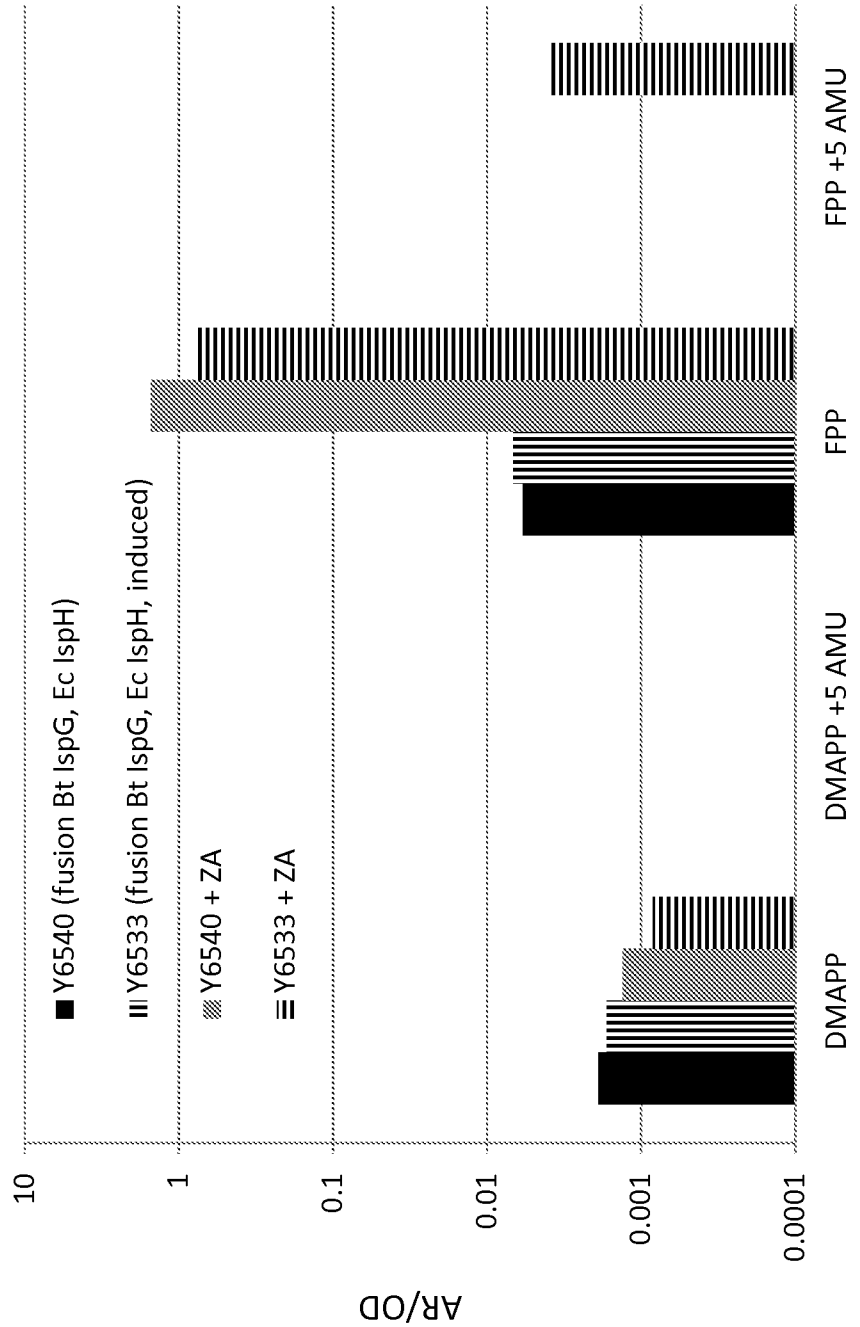
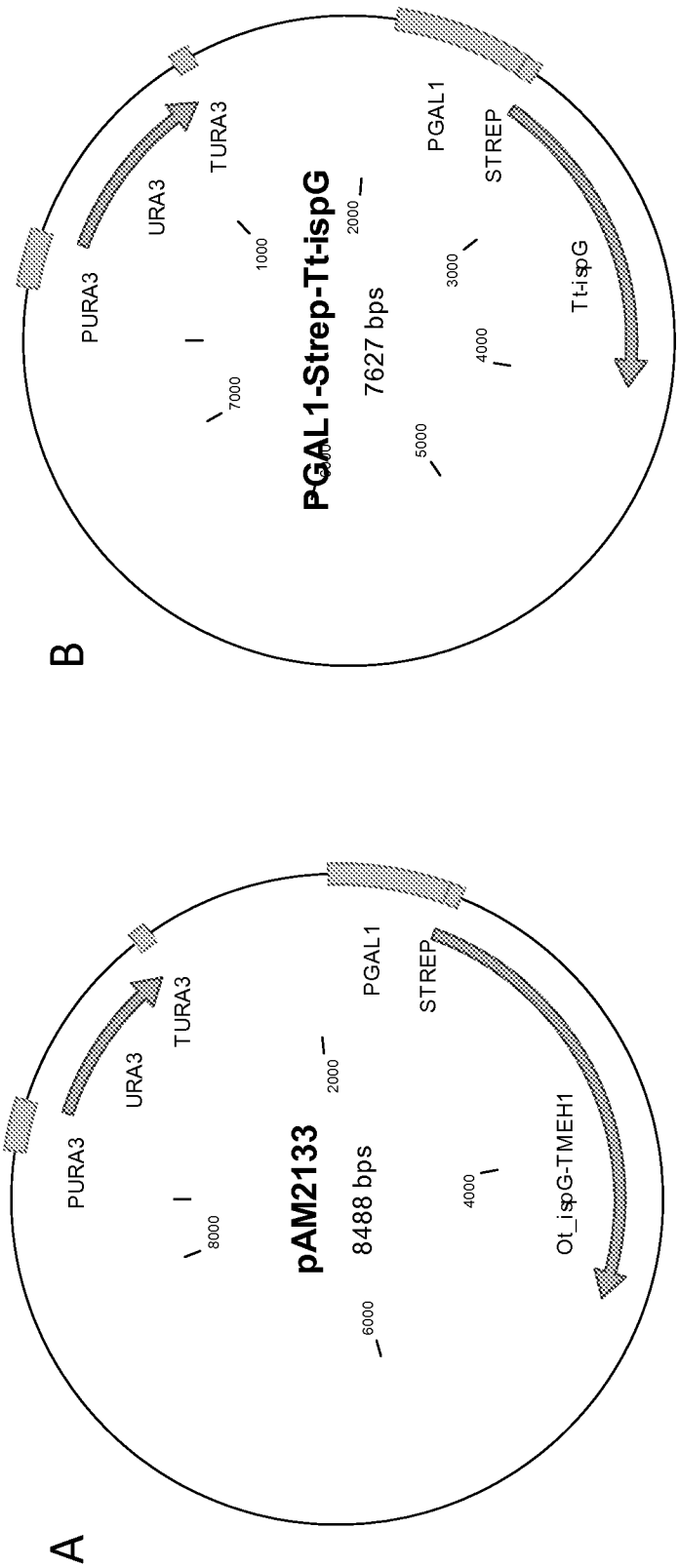

Y6540 (fusion Bt IspG, Ec IspH)
Y6533 (fusion Bt IspG, Ec IspH, induced)
Y6540 + ZA
Y6533 + ZA
AR/OD
10
1
0.1
0.01
0.001
0.0001
DMAPP
DMAPP +5 AMU
FPP
FPP +5 AMU

Figure 9


A
pAM2133
8488 bps
PURA3
URA3
TURA3
PGAL1
STREP
Ot_ispG-TMEH1
2000
4000
6000
8000
B
PGAL1-Strep-Tt-ispG
7627 bps
PURA3
URA3
TURA3
PGAL1
STREP
Tt-ispG
1000
2000
3000
4000
5000
7000

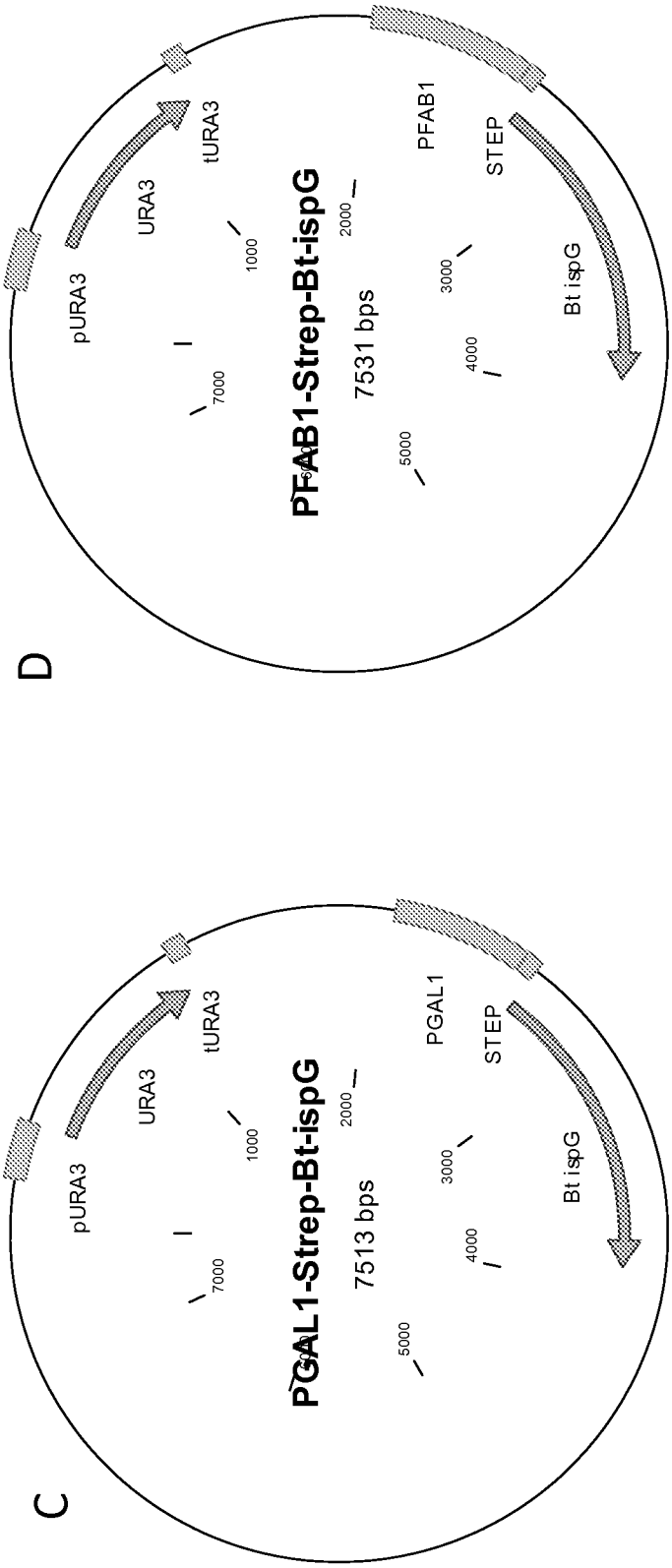



Figure 9 (continued)