



(51) International Patent Classification:
C12N 15/09 (2006.01) C12N 15/63 (2006.01)
C12N 15/113 (2010.01)

(21) International Application Number:
PCT/US2014/071536

(22) International Filing Date:
19 December 2014 (19.12.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/920,382 23 December 2013 (23.12.2013) US

(71) Applicant: DANISCO US INC. [US/US]; 925 Page Mill
Road, Palo Alto, CA 94304-1013 (US).

(72) Inventors: EULER, Lori, J.; C/o Danisco US Inc., 925
Page Mill Road, Palo Alto, CA 94304-1013 (US). MUIR,
Rachel, E.; C/o Danisco US Inc., 925 Page Mill Road,
Palo Alto, CA 94304-1013 (US). POLLAK, Dana, M.,
W.; C/o Danisco US Inc., 925 Page Mill Road, Palo Alto,
CA 94304-1013 (US). VAN DYK, Tina, K.; C/o Danisco
US Inc., 925 Page Mill Road, Palo Alto, CA 94304-1013
(US).

(74) Agents: TESSMAN, John, W. et al.; Morrison & Foerster
LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018
(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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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.

[Continued on next page]

(54) Title: COMPOSITION AND METHODS FOR CONTROL OF INDUSTRIAL SCALE PRODUCTION OF BIO-PRODUCTS

FIG. 1A

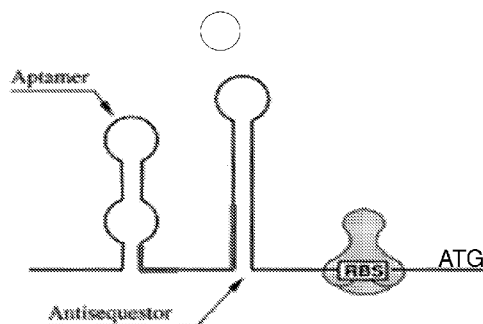
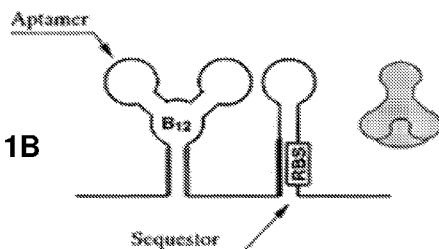


FIG. 1B



(57) Abstract: The invention provides compositions and
methods for microorganisms that have been engineered to
produce and/or to improve efficiency of production of indus-
trial bio-products using regulatable gene expression control
constructs to control the expression of heterologously ex-
pressed nucleic acids.



(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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

COMPOSITIONS AND METHODS FOR CONTROL OF INDUSTRIAL SCALE PRODUCTION OF BIO-PRODUCTS

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Patent Application No. 61/920,382, filed December 23, 2013, which is incorporated by reference herein in its entirety.

SUBMISSION OF SEQUENCE LISTING ON ASCII TEXT FILE

[0002] The content of the following submission on ASCII text file is incorporated herein by reference in its entirety: a computer readable form (CRF) of the Sequence Listing (file name: 246252007940SeqList.txt, date recorded: December 16, 2014, size: 9 KB).

FIELD OF THE INVENTION

[0003] The invention provides compositions and methods for the control of industrial scale production of products produced by genetically engineered microorganisms, and methods to produce and/or to improve efficiency of production of industrial products.

BACKGROUND OF THE INVENTION

[0004] Industrial scale production of bio-products can be improved using microorganisms genetically engineered to express one or more heterologous nucleic acids. Production can be further improved if heterologous gene expression is regulated so that bio-product production is “off” when microorganisms are cultured as a seed stock under conditions suitable for initially establishing a population for large scale production which can then be turned “on” once the seed stock has been placed into a suitable bioreactor to begin large-scale fermentation. Currently, many organisms are engineered for the production of bio-products using heterologous genes under the control of promoters activated by chemical compounds such as isopropyl thiogalactoside (IPTG), which can be added at the bioreactor stage to start the production process. Unfortunately, compounds such as IPTG can be cost prohibitive.

Therefore, alternative means for controlling gene regulation in engineered microorganisms during the industrial production of bio-products are desirable.

[0005] Throughout this specification, various patents, patent applications and other types of publications (*e.g.*, journal articles) are referenced. The disclosure of all patents, patent applications, and publications cited herein are hereby incorporated by reference in their entirety for all purposes.

BRIEF SUMMARY OF THE INVENTION

[0006] The invention provides, *inter alia*, regulatable gene expression constructs for use in controlling gene expression in engineered microorganisms during the industrial production of one or more bio-products as well as engineered microorganisms for use in methods for producing one or more bio-products.

[0007] Accordingly, in some aspects, provided herein is a regulatable gene expression construct comprising: (i) a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest. In some embodiments, the riboswitch comprises an aptamer domain and an expression platform domain. In some embodiments of any of the embodiments disclosed herein, said riboswitch is responsive to a small molecule selected from the group consisting of cobalamin (coenzyme of vitamin B12), flavin mononucleotide (FMN), glutamine, glycine, lysine, purine, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), tetrahydrofolate, or thiamin pyrophosphate (TPP). In some embodiments of any of the embodiments disclosed herein, said riboswitch is responsive to cobalamin (coenzyme of vitamin B12). In some embodiments of any of the embodiments disclosed herein, said sigma factor is selected from the group consisting of sigma 70 (RpoD), sigma 19 (Fecl), sigma 24 (RpoE), sigma 28 (RpoF), sigma 32 (RpoH), sigma 38 (RpoS) or sigma 54 (RpoN). In some embodiments of any of the embodiments disclosed herein, said sigma factor is sigma 28 (RpoF). In some embodiments of any of the embodiments disclosed herein, said sigma 28 is encoded by the *fliA* gene. In some embodiments of any of the embodiments disclosed herein, said sigma factor-dependent promoter is activated by a sigma factor selected from the group

consisting of sigma 70 (RpoD), sigma 19 (Fecl), sigma 24 (RpoE), sigma 28 (RpoF), sigma 32 (RpoH), sigma 38 (RpoS) or sigma 54 (RpoN). In some embodiments of any of the embodiments disclosed herein, said sigma factor-dependent promoter is activated by sigma 28 (RpoF). In some embodiments of any of the embodiments disclosed herein, said sigma factor-dependent promoter is the *pfliC* promoter.

[0008] In other aspects, provided herein is a recombinant bacterial cell comprising the regulatable gene construct of any one of embodiments disclosed herein. In some embodiments, when said bacterial cell is cultured in the presence of said small molecule, the expression of said sigma factor is repressed, thereby inhibiting the expression of said gene of interest operably linked to said sigma factor-dependent promoter. In some embodiments, when said bacterial cell is cultured in the absence of said small molecule, the expression of said sigma factor is activated, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter. In some embodiments of any of the embodiments disclosed herein, said cell comprises: (i) a first extrachromosomal plasmid comprising a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a second extrachromosomal plasmid comprising a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest. In some embodiments of any of the embodiments disclosed herein, said cell comprises a single extrachromosomal plasmid comprising: (i) a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest. In some embodiments of any of the embodiments disclosed herein, said bacterial cell comprises: (i) a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest, wherein at least one of (i) or (ii) is integrated into the chromosome of said bacterial cell. In some embodiments of any of the embodiments disclosed herein, said bacterial cell further comprises a genetic modification to increase intracellular transport of a small molecule selected from the group consisting of cobalamin (coenzyme of vitamin B12), flavin mononucleotide (FMN), glutamine, glycine, lysine, purine, S-

adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), tetrahydrofolate, or thiamin pyrophosphate (TPP). In some embodiments of any of the embodiments disclosed herein, said bacterial cell further comprises a genetic modification to increase intracellular transport of cobalamin (coenzyme of vitamin B12).

[0009] In yet other aspects, provided herein are methods for regulating the expression of a gene of interest comprising: (i) initially culturing the recombinant bacterial cell of any of the embodiments disclosed herein in a first culture condition comprising said small molecule in an amount sufficient to repress expression of said sigma factor, thereby inhibiting the expression of said gene of interest operably linked to said sigma factor-dependent promoter; and (ii) subsequently culturing said recombinant bacterial cells under a second culture condition comprising an amount of said small molecule that is not sufficient to repress expression of said sigma factor, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in the second culture condition. In some embodiments, said initial culture comprising the recombinant bacterial cell of any of the embodiments disclosed herein is added into said second culture, wherein the addition of the initial culture to the second culture results in the dilution of said small molecule such that the amount of said small molecule in said second culture is not sufficient to repress said sigma factor expression, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in the second culture condition.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] FIG. 1A and 1B shows a representation of the cobalamin dependent riboswitch RNA comprising an aptamer domain and an expression platform domain. Figure 1A shows the riboswitch in the on configuration, with the ribosome binding site accessible to begin translating at the start codon. Figure 1B shows the configuration of the riboswitch upon binding of cobalamin (B₁₂), wherein the ribosome binding site is sequestered, resulting in the shutting off of translation.

[0011] FIG. 2 shows a schematic representation of integration of *E. coli* gapA promoters, btuB leader sequence, btuB ribosome binding site, and BtuB-LuxC fusion protein into plasmid pDEW201.

[0012] FIG. 3A and 3B shows the expression of LuxC in RLU as a function of cobalamin concentration diluted from 0.01 mg/L to 0.001 mg/L in *E. coli* transfected with pDMWP156 plasmid.

[0013] FIG. 4 shows plasmid pDMWP165 comprising glucose isomerase promoter p1.20, BtuB2 comprising the cobalamin riboswitch operably linked to the *fliA* gene encoding the sigma factor 28.

[0014] FIG. 5 shows plasmid pDMWP174 comprising the *PfliC* promoter, *mvaE* and *mvaS* genes.

[0015] FIG. 6 shows plasmid pDMWP180 comprising glucose isomerase promoter p1.20, BtuB2 comprising the cobalamin riboswitch operably linked to the *fliA* gene encoding the sigma factor 28.

[0016] FIG. 7 shows plasmid pDMWP181 comprising glucose isomerase promoter p1.5, BtuB2 comprising the cobalamin riboswitch operably linked to the *fliA* gene encoding the sigma factor 28.

[0017] FIG. 8 shows plasmid pDMWP175 comprising the *PfliC* promoter, *mvaE* and *mvaS* genes.

[0018] FIG. 9 shows the microbial fuels that can be produced from a carbon source via cellular pathways.

[0019] FIG. 10 shows the classical and modified MVA pathways. 1, acetyl-CoA acetyltransferase (AACT); 2, HMG-CoA synthase (HMGS); 3, HMG-CoA reductase (HMGR); 4, mevalonate kinase (MVK); 5, phosphomevalonate kinase (PMK); 6, diphosphomevalonate decarboxylase (MVD or DPMDC); 7, isopentenyl diphosphate isomerase (IDI); 8, phosphomevalonate decarboxylase (PMDC); 9, isopentenyl phosphate kinase (IPK). The classical MVA pathway proceeds from reaction 1 through reaction 7 via reactions 5 and 6, while a modified MVA pathway goes through reactions 8 and 9. P and PP in the structural formula are phosphate and pyrophosphate, respectively. This figure was taken from Koga and Morii, *Microbiology and Mol. Biology Reviews*, 71:97-120, 2007, which is incorporated by reference in its entirety, particularly with respect to nucleic acids and

polypeptides of the modified MVA pathway. The modified MVA pathway is present, for example, in some archaeal organisms, such as *Methanosarcina mazei*.

[0020] FIG. 11 shows a schematic representation of an obligate anaerobe expressing (a) a heterologous IspS polypeptide, (b) a heterologous DXS polypeptide, and (c) a heterologous IDI polypeptide to increase DXP pathway flux and isoprene production.

[0021] FIG. 12 shows a schematic representation of an obligate anaerobe engineered with *mvaE* and *mvaS* to express the upper MVA pathway.

[0022] FIG. 13 shows a schematic representation of expressing the lower MVA pathway in an obligate anaerobe including expressing (a) a heterologous MVK polypeptide, (b) a heterologous PMK polypeptide, and (c) a heterologous MVD polypeptide in the cells expressing heterologous IDI polypeptide and heterologous IspS polypeptide for the purpose of increasing isoprene production.

[0023] FIG. 14 shows a schematic representation of expressing the entire MVA pathway in an obligate anaerobe by introducing *mvaE* and *mvaS* in the cells expressing (a) a heterologous MVK polypeptide, (b) a heterologous PMK polypeptide, (c) a heterologous MVD polypeptide, (d) a heterologous IDI polypeptide, and (e) a heterologous IspS polypeptide for the purpose of increasing isoprene production.

[0024] FIG. 15 shows a schematic representation of redirecting carbon flux away from acetate by reducing expression of *ack* and *adhE* to reduce loss of carbon to side products. The arrows next to *Ack* or *AdhE* used in the production of acetate and ethanol, respectively, indicate a reduction of activity or enzyme expression for pathways leading to fermentation products such as acetate, ethanol, or any other alcohol, or carbon containing end product. The purpose is to maximize carbon channeling to isoprene via genetic manipulation.

[0025] FIG. 16 shows exemplary pathways for production of butadiene from acetyl-CoA, glutaconyl-CoA, glutaryl-CoA, 3-aminobutyryl-CoA or 4-hydroxybutyryl-CoA via crotyl alcohol. Enzymes for transformation of the identified substrates to products include: A. acetyl-CoA:acetyl-CoA acyltransferase, B. acetoacetyl-CoA reductase, C. 3-hydroxybutyryl-CoA dehydratase, D. crotonyl-CoA reductase (aldehyde forming), E. crotonaldehyde reductase (alcohol forming), F. crotyl alcohol kinase, G. 2-butenyl-4-phosphate kinase, H. butadiene

synthase, I. crotonyl-CoA hydrolase, synthetase, transferase, J. crotonate reductase, K. crotonyl-CoA reductase (alcohol forming), L. glutaconyl-CoA decarboxylase, M., glutaryl-CoA dehydrogenase, N. 3-aminobutyryl-CoA deaminase, O. 4-hydroxybutyryl-CoA dehydratase, P. crotyl alcohol diphosphokinase.

[0026] FIG. 17 shows exemplary pathways for production of butadiene from erythrose-4-phosphate. Enzymes for transformation of the identified substrates to products include: A. Erythrose-4-phosphate reductase, B. Erythritol-4-phosphate cytidyltransferase, C. 4-(cytidine 5'-diphospho)-erythritol kinase, D. Erythritol 2,4-cyclodiphosphate synthase, E. 1-Hydroxy-2-butenyl 4-diphosphate synthase, F. 1-Hydroxy-2-butenyl 4-diphosphate reductase, G. Butenyl 4-diphosphate isomerase, H. Butadiene synthase I. Erythrose-4-phosphate kinase, J. Erythrose reductase, K. Erythritol kinase.

[0027] FIG. 18 shows an exemplary pathway for production of butadiene from malonyl-CoA plus acetyl-CoA. Enzymes for transformation of the identified substrates to products include: A. malonyl-CoA:acetyl-CoA acyltransferase, B. 3-oxoglutaryl-CoA reductase (ketone-reducing), C. 3-hydroxyglutaryl-CoA reductase (aldehyde forming), D. 3-hydroxy-5-oxopentanoate reductase, E. 3,5-dihydroxypentanoate kinase, F. 3H5PP kinase, G. 3H5PDP decarboxylase, H. butenyl 4-diphosphate isomerase, I. butadiene synthase, J. 3-hydroxyglutaryl-CoA reductase (alcohol forming), K. 3-oxoglutaryl-CoA reductase (aldehyde forming), L. 3, 5-dioxopentanoate reductase (ketone reducing), M. 3,5-dioxopentanoate reductase (aldehyde reducing), N. 5-hydroxy-3-oxopentanoate reductase, O. 3-oxo-glutaryl-CoA reductase (CoA reducing and alcohol forming). Compound abbreviations include: 3H5PP = 3-Hydroxy-5-phosphonooxypentanoate and 3H5PDP = 3-Hydroxy-5-[hydroxy(phosphonooxy)phosphoryl]oxy pentanoate.

DETAILED DESCRIPTION OF THE INVENTION

[0028] The invention provides, *inter alia*, regulatable gene expression constructs for use in controlling gene expression in engineered microorganisms during the industrial production of one or more bio-products as well as engineered microorganisms comprising the regulatable gene expression constructs for use in methods for producing one or more bio-products.

[0029] Use of this expression control system is useful in the engineering of microorganisms for the biological production of various industrial products (*e.g.*, bio-products).

I. General Techniques

[0030] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as Biotechnology: A Textbook of Industrial Microbiology (Brock, Sinauer Associates, Inc., Second Edition, 1989), Molecular Cloning: A Laboratory Manual (Sambrook *et al.*, 1989, 2nd ed.); Oligonucleotide Synthesis (M.J. Gait, ed., 1984); Methods in Enzymology (Academic Press, Inc.); Current Protocols in Molecular Biology (F.M. Ausubel *et al.*, eds., 1987, and periodic updates); PCR: The Polymerase Chain Reaction (Mullis *et al.*, eds., 1994), Dictionary of Microbiology and Molecular Biology (Singleton *et al.*, 2nd ed., J. Wiley and Sons, New York, NY, 1994); and Advanced Organic Chemistry Reactions, Mechanisms and Structure (March, 4th ed., John Wiley and Sons, New York, NY, 1992), which provide one skilled in the art with a general guide to many of the terms and methods used in the present disclosure. Additional methods used in the Examples are described in manuals including Advanced Bacterial Genetics (Davis, Roth and Botstein, Cold Spring Harbor Laboratory, 1980), Experiments with Gene Fusions (Silhavy, Berman and Enquist, Cold Spring Harbor Laboratory, 1984), Experiments in Molecular Genetics (Miller, Cold Spring Harbor Laboratory, 1972) Experimental Techniques in Bacterial Genetics (Maloy, in Jones and Bartlett, 1990), and A Short Course in Bacterial Genetics (Miller, Cold Spring Harbor Laboratory 1992).

[0031] Unless defined otherwise, 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.

II. Definitions

[0032] “Isoprene” refers to 2-methyl-1,3-butadiene (CAS# 78-79-5). It can refer to the direct and final volatile C5 hydrocarbon product from the elimination of pyrophosphate from

3,3-dimethylallyl pyrophosphate (DMAPP). Isoprene is not limited by the method of its manufacture.

[0033] “Industrial bio-products” can include, but are not limited to, isoprene, isoprenoids, isoprenoid precursors, butadiene and ethanol. Industrial products can also include, but are not limited to, bio-products derived directly or indirectly from 2-keto acids, malonyl-CoA, and acetoacetyl-CoA. Industrial bio-products can also include, but are not limited to, monoterpenes, diterpenes, triterpenes, tetraterpenes, sesquiterpene, polyterpene, abietadiene, amorphadiene, carene, α -farnesene, β -farnesene, farnesol, geraniol, geranylgeraniol, linalool, limonene, myrcene, nerolidol, ocimene, patchoulol, β -pinene, sabinene, γ -terpinene, terpineol, valencene. Industrial bio-products can further include, but are not limited to, non-fermentative alcohols (*e.g.*, 1-propanol, 1-butanol, isobutanol, 2-methyl-1-butanol, 3-methyl-1-butanol, 3-methyl-1-pentanol, 4-methyl-1-pentanol and 1-hexanol), fatty acid-derived hydrocarbons (fatty alcohols, fatty esters, olefins, and alkanes), and fermentative alcohols (*e.g.*, butanol). Industrial bio-products can also include, but are not limited to, enzyme products such as amylases, cellulases, glucyltransferases (“gtf”), lipases, xylanases, proteases, phytases, *etc.* or protein products such as aquaporins.

[0034] A “nucleic acid” or “polynucleotide” refers to two or more deoxyribonucleotides and/or ribonucleotides in either single or double-stranded form.

[0035] A “nucleic acid of interest” refers to a polynucleotide encoding a polypeptide that is a part of the synthetic pathway for any industrial product. Alternatively, a “nucleic acid of interest” can refer to a polynucleotide encoding a polypeptide that is the desired product of a bio-process (*e.g.*, an industrial bio-product).

[0036] An “endogenous nucleic acid” is a nucleic acid whose nucleic acid sequence is naturally found in the host cell. In some aspects, an endogenous nucleic acid is identical to a wild-type nucleic acid that is found in the host cell in nature. In some aspects, one or more copies of endogenous nucleic acids are introduced into a host cell.

[0037] A “heterologous nucleic acid” can be a nucleic acid whose nucleic acid sequence is from another species than the host cell or another strain of the same species of the host cell. In some aspects, the sequence is not identical to that of another nucleic acid naturally found in the

same host cell. In some aspects, a heterologous nucleic acid is not identical to a wild-type nucleic acid that is found in the same host cell in nature. In various embodiments of the invention, a heterologous nucleic acid encodes for one or more industrial bio-products.

[0038] “Polypeptides” includes polypeptides, proteins, peptides, fragments of polypeptides, fusion polypeptides and variants.

[0039] An “endogenous polypeptide” is a polypeptide whose amino acid sequence is naturally found in the host cell. In some aspects, an endogenous polypeptide is identical to a wild-type polypeptide that is found in the host cell in nature.

[0040] A “heterologous polypeptide” is a polypeptide encoded by a heterologous nucleic acid. In some aspects, the sequence is not identical to that of another polypeptide encoded by a nucleic acid naturally found in the same host cell.

[0041] As used herein, the terms “minimal medium” or “minimal media” refer to growth medium containing the minimum nutrients possible for cell growth, generally without the presence of amino acids. Minimal medium typically contains: (1) a carbon source for bacterial growth; (2) various salts, which can vary among bacterial species and growing conditions; and (3) water. The carbon source can vary significantly, from simple sugars like glucose to more complex hydrolysates of other biomass, such as yeast extract, as discussed in more detail below. The salts generally provide essential elements such as magnesium, nitrogen, phosphorus, and sulfur to allow the cells to synthesize proteins and nucleic acids. Minimal medium can also be supplemented with selective agents, such as antibiotics, to select for the maintenance of certain plasmids and the like. For example, if a microorganism is resistant to a certain antibiotic, such as ampicillin or tetracycline, then that antibiotic can be added to the medium in order to prevent cells lacking the resistance from growing. Medium can be supplemented with other compounds as necessary to select for desired physiological or biochemical characteristics, such as particular amino acids and the like.

[0042] As used herein, the term “isoprenoid” refers to a large and diverse class of naturally-occurring class of organic compounds composed of two or more units of hydrocarbons, with each unit consisting of five carbon atoms arranged in a specific pattern. As used herein, “isoprene” is expressly excluded from the definition of “isoprenoid.”

[0043] As used herein, the term “terpenoid” refers to a large and diverse class of organic molecules derived from five-carbon isoprenoid units assembled and modified in a variety of ways and classified in groups based on the number of isoprenoid units used in group members. Hemiterpenoids have one isoprenoid unit. Monoterpenoids have two isoprenoid units. Sesquiterpenoids have three isoprenoid units. Diterpenoids have four isoprene units. Sesterterpenoids have five isoprenoid units. Triterpenoids have six isoprenoid units. Tetraterpenoids have eight isoprenoid units. Polyterpenoids have more than eight isoprenoid units.

[0044] As used herein, “isoprenoid precursor” refers to any molecule that is used by organisms in the biosynthesis of terpenoids or isoprenoids. Non-limiting examples of isoprenoid precursor molecules include, *e.g.*, isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP).

[0045] As used herein, the term “mass yield” refers to the mass of the product produced by the bacterial cells divided by the mass of the glucose consumed by the bacterial cells multiplied by 100.

[0046] By “specific productivity,” it is meant the mass of the product produced by the bacterial cell divided by the product of the time for production, the cell density, and the volume of the culture.

[0047] By “titer,” it is meant the mass of the product produced by the bacterial cells divided by the volume of the culture.

[0048] As used herein, the term “cell productivity index (CPI)” refers to the mass of the product produced by the bacterial cells divided by the mass of the bacterial cells produced in the culture.

[0049] Unless defined otherwise herein, 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 pertains.

[0050] As used herein, the singular terms “a,” “an,” and “the” include the plural reference unless the context clearly indicates otherwise.

[0051] It is intended that every maximum numerical limitation given throughout this specification includes every lower numerical limitation, as if such lower numerical limitations were expressly written herein. Every minimum numerical limitation given throughout this specification will include every higher numerical limitation, as if such higher numerical limitations were expressly written herein. Every numerical range given throughout this specification will include every narrower numerical range that falls within such broader numerical range, as if such narrower numerical ranges were all expressly written herein.

III. Small molecule binding riboswitches for control of gene expression

[0052] A "riboswitch" as described herein means an expression control element, which comprises an RNA molecule that comprises an aptamer region, also known as a ligand binding region, and an expression platform domain, also known as a regulatory domain. As described herein, in bacteria, the riboswitch is upstream of an mRNA coding for a sigma factor, wherein the expression platform domain regulates expression of the sigma factor. The aptamer binds to a specific small molecule, depending on the identity of the riboswitch. The expression platform domain comprises a ribosome binding site, such that when the aptamer binds the specific small molecule, the configuration of the ribosome binding site is transformed and included in a secondary structure of the RNA such that the ribosome can no longer bind. Thus, the conformation of the aptamer and expression platform domain can be altered depending on the concentration of the aptamer binding small molecule. The small molecule can be used at higher concentrations to turn off the expression of the sigma factor, or can be diluted out to turn on expression of the sigma factor. This regulation of gene expression by a riboswitch is shown schematically in Figure 1.

[0053] In some systems, such as eukaryotic microorganisms, the riboswitch interacts directly with mRNA to regulate the expression of the protein of interest. In such instances, the RNA molecule that comprises the riboswitch can be similarly used to control the expression of genes of interest as described herein, where the small molecule interaction of the riboswitch is used to directly control the expression of the gene of interest. See for example Suess et al., *Methods in Molecular Biology*, 2012, 824:381-391; Kubodera et al., *FEBS Lett*, 555(3):516-520; Wittmann et al., *FEBS Lett*, 2012, 586:2076-2083; Sudarsan et al., *RNA*, 9(6): 644-647; and Wachter A., *RNA Biology*, 2010 7(1):67-76, the disclosures of which are hereby

incorporated by reference in their entirety, and particularly as they relate to riboswitch systems.

[0054] In bacteria, any riboswitch can be used to regulate the expression of any gene downstream of the riboswitch. As described herein, the riboswitch is used to control the expression of a sigma factor, which interacts with a sigma factor-dependent promoter. A sigma factor-dependent promoter can be operably linked to one or more genes of interest in order to control the expression of these genes of interest. Since the sigma factor-dependent promoter is controlled by the expression of sigma factor, the riboswitch mechanism can be used to regulate the gene expression of the gene of interest operably linked to the sigma factor-dependent promoter. The riboswitch system described herein is ideal for controlling gene expression in engineered microorganisms during the industrial production of one or more bio-products. The bio-product production is readily turned “off” when microorganisms are cultured as a seed stock with a sufficient concentration of the aptamer binding small molecule so that the riboswitch is “off”, *i.e.* the expression of the sigma factor is repressed. The seed stock can be maintained under conditions suitable for initially establishing a population for large scale production which can then be turned “on” once the seed stock has been placed into a suitable bioreactor to begin large-scale fermentation. By dilution of the seed stock to the large-scale bioreactor volume (*e.g.* 10-fold dilution), the small molecule concentration is sufficiently diluted so that binding to aptamer is reduced such that the riboswitch is turned “on” and the expression of the sigma factor is activated, resulting in beginning the bio-production process. Such a system is considerably more cost effective than current methods, such as adding IPTG to turn on an inducible promoter.

[0055] In some embodiments, a riboswitch is any riboswitch known in the art, for example as described in PCT publication WO 2012153142, the disclosure of which is hereby incorporated by reference as it relates to riboswitches. In some embodiments, the riboswitch is selected from the group consisting of cobalamin riboswitch, cyclic di-GMP riboswitches, FMN riboswitch, *glmS* riboswitch, glutamine riboswitches, glycine riboswitch, lysine riboswitch, PreQ1 riboswitches, purine riboswitches, SAH riboswitches, SAM riboswitches, SAM-SAH riboswitches, Tetrahydrofolate riboswitches, TPP riboswitches, Moco riboswitch, and adenine sensing add-A riboswitch. In some embodiments, the specific small molecule that binds to the riboswitch is selected from the group consisting of cobalamin (coenzyme of vitamin B12),

cyclic di-GMP, flavin mononucleotide (FMN), glucosamine-6-phosphate, glutamine, glycine, lysine, pre-queosine, purines, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), both SAH and SAM, and thiamine pyrophosphate (TPP). In some instances, the riboswitch is a cobalamin riboswitch, wherein binding of a sufficient concentration of cobalamin to the riboswitch represses the expression of the gene operably linked to the riboswitch, such as repressing the expression of a sigma factor.

[0056] The riboswitch system as described herein in bacteria is operably linked to a sigma factor, such that the sigma factor expression is repressed or activated depending on the concentration of the small molecule that binds to the riboswitch aptamer domain. Any suitable sigma factor can be used in the systems and methods as described herein, in combination with a suitable sigma factor-dependent promoter that is operably linked to one or more genes of interest (see, *inter alia*, Hiu Yin Yu et al., *J. Bacteriol.*, 2006, 188(15):5524-5531; Campbell et al., *Curr Opin Microbiol.*, 2008, 11(2):121-127; Watson et al., *Nat Struct Mol Biol.*, 2011, 18(3):359-363, the disclosures of which are hereby incorporated by reference in their entirety, and particularly as they relate to sigma factors and sigma factor-dependent promoters). Thus the regulatable gene expression constructs can be engineered to operably link a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch to one or more nucleic acids encoding a sigma factor, and to operably link a suitable sigma factor-dependent promoter to one or more nucleic acids encoding a gene of interest. The sigma factor can be an *E. coli* sigma factor selected from the group consisting of sigma 70 (RpoD), sigma 19 (Fecl), sigma 24 (RpoE), sigma 28 (RpoF), sigma 32 (RpoH), sigma 38 (RpoS) or sigma 54 (RpoN). In some instances, the sigma factor is sigma 28 encoded by the *fliA* gene. In some instances, the sigma factor is sigma 28 encoded by the *fliA* gene and the sigma factor-dependent promoter is the *pfliC* promoter. In some instances, the sigma factor is sigma 28 and the sigma factor-dependent promoter is the *pfliC* promoter as described in the Examples herein.

[0057] The riboswitch system as described herein in bacteria may further comprise nucleic acid encoding at least a portion of the native protein that is controlled by the riboswitch. As such the nucleic acid encoding the RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor, further comprises the nucleic acid encoding at least a portion of the native protein controlled by the riboswitch. In

one instance, the nucleic acid encodes an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a fusion protein comprising at least a portion of the native protein that is controlled by the riboswitch and a sigma factor. In some instances, the nucleic acid comprises the BtuB2 leader sequence, at least a portion of BtuB2 gene and the fliA gene. In some instances, the fusion protein comprises 5 to 50 amino acids of the BtuB2 gene product, 10 to 40 amino acids of the BtuB2 gene product, 15 to 35 amino acids of the BtuB2 gene product, or 33 amino acids of the BtuB2 gene product.

[0058] The recombinant cells comprising a regulatable gene construct as described herein, *i.e.* bacterial cells comprising (i) a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest, are useful in the industrial-scale production of bio-products. The recombinant cells described herein include cells comprising a single extrachromosomal plasmid comprising (i) the nucleic acid molecule encoding an RNA comprising the small molecule-responsive riboswitch operably linked to the one or more nucleic acids encoding the sigma factor; and (ii) the sigma factor-dependent promoter operably linked to the one or more nucleic acids encoding a gene of interest; or cells comprising (i) a first extrachromosomal plasmid comprising a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a second extrachromosomal plasmid comprising a sigma factor-dependent promoter operably linked one or more nucleic acids encoding a gene of interest; or cells comprising (i) a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest, wherein at least one of (i) or (ii) is integrated into the chromosome of said bacterial cell.

[0059] The recombinant bacterial cell as described herein include cells further comprising a genetic modification to increase intracellular transport of a small molecule that binds the riboswitch aptamer domain, such as a small molecule selected from the group consisting of cobalamin (coenzyme of vitamin B12), cyclic di-GMP, flavin mononucleotide (FMN), glucosamine-6-phosphate, glutamine, glycine, lysine, pre-queosine, purines, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), and thiamine pyrophosphate

(TPP). In any embodiments described herein, the bacterial cell further comprises a genetic modification to increase intracellular transport of cobalamin (coenzyme of vitamin B12).

[0060] The recombinant bacterial cell as described herein for use in the industrial-scale production of a bio-product, cultured in the presence of the small molecule that binds the riboswitch aptamer at a sufficient concentration to repress the expression of the sigma factor, the expression of the sigma factor is repressed, thereby inhibiting the expression of the one or more genes operably linked to the sigma factor-dependent promoter; and when said bacterial cell is cultured in the absence of the small molecule, or in the presence of the small molecule at sufficiently low concentrations that do not repress the expression of the sigma factor, the expression of the sigma factor is activated, thereby activating the expression of the one or more genes operably linked to the sigma factor-dependent promoter.

[0061] The bacterial cells can be used in a method for regulating the expression of a gene of interest comprising (i) initially culturing the recombinant bacterial cell in a first culture condition comprising said small molecule in an amount sufficient to repress expression of said sigma factor, thereby inhibiting the expression of said gene of interest operably linked to said sigma factor-dependent promoter; and (ii) subsequently culturing said recombinant bacterial cells under a second culture condition comprising an amount of said small molecule that is not sufficient to repress expression of said sigma factor, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in second culture conditions. In some instances, the initial culturing of the recombinant cells under conditions to repress expression of the sigma factor does not completely inhibit the expression of the gene operably linked to the sigma factor-dependent promoter, as there is some level of sigma factor produced, resulting in some level of the desired bio-product being produced in the initial culture. In some instances, the level of desired bio-product produced in the second culture condition is at least about 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, or 10-fold greater than the level of desired bio-product produced in the first culture condition.

[0062] In some instances, the method comprises said initial culture comprising the recombinant bacterial cell added into said second culture, wherein the addition of the initial culture to the second culture results in the dilution of said small molecule such that the amount

of said small molecule in said second culture is not sufficient to repress said sigma factor expression, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in second culture conditions. In some instances, the level of desired bio-product produced upon dilution into the second culture condition is at least about 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, or 10-fold greater than the level of desired bio-product produced in the first culture condition.

[0063] It is to be understood that compositions and/or systems, methods of making and using these aspects and/or embodiments are encompassed within the scope of the invention.

A. Exemplary Nucleic Acids and Polypeptides

[0064] The inventors have created (and herein describe) polynucleotides, polypeptides, plasmids, vectors, expression systems, host cells, etc. based on the components of this gene expression control system, as well as methods of making and using these components to facilitate the genetic manipulation of microorganisms to produce one or more industrial bio-products such as (but not limited to) isoprene, butadiene, isoprenoids, and ethanol.

[0065] Various nucleic acids as components of the gene expression control system, or encoding components of the gene expression control system described herein, including an RNA comprising small molecule responsive riboswitch, sigma factors, sigma factor-dependent promoters, one or more genes of interest and other polypeptides and nucleic acids can be used (either individually or in any combination) in the compositions and methods as described herein.

[0066] In some embodiments, a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch is operably linked to one or more nucleic acids encoding a sigma factor. In some embodiments, a sigma factor-dependent promoter is operably linked to one or more nucleic acids encoding a gene of interest. "Operably linked" refers to one or more genes that have been placed under the regulatory control of a promoter, which then controls the transcription and optionally the translation of those genes. In the construction of heterologous promoter/structural gene combinations, it is generally preferred to position the genetic sequence or promoter at a distance from the gene transcription start site that is approximately the same as the distance between that genetic sequence or promoter and the

gene it controls in its natural setting; *i.e.* the gene from which the genetic sequence or promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting; *i.e.*, the gene from which it is derived.

[0067] In some embodiments, a nucleic acid as described herein has one or more mutations compared to the sequence of a wild-type (*i.e.*, a sequence occurring in nature) nucleic acid, including, *inter alia*, the nucleic acid encoding the RNA comprising a small molecule responsive riboswitch, the nucleic acid encoding the sigma factors, the sigma factor-dependent promoter nucleic acid, or nucleic acid encoding the one or more genes of interest. In some embodiments, the nucleic acid has one or more mutations (*e.g.*, a silent mutation) that increase the transcription or translation of the nucleic acid. In some embodiments, the nucleic acid is a degenerate variant of any nucleic acid as described herein.

[0068] As will be understood by those skilled in the art, the polynucleotide sequences of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

[0069] Polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

[0070] Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence) or may comprise a variant, or a biological functional equivalent of such a sequence.

Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions. In some embodiments, the enzymatic activity of the encoded polypeptide is not substantially diminished relative to the unmodified polypeptide. In some embodiments, the enzymatic activity of the encoded polypeptide is improved (*e.g.*, optimized) relative to the unmodified polypeptide. In other embodiments, the enzymatic activity of the encoded

polypeptide is substantially diminished relative to the unmodified polypeptide. The effect on the enzymatic activity of the encoded polypeptide may generally be assessed as described herein.

[0071] As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence. Such nucleotides are typically referred to as “codon-optimized.” Any of the nucleotide sequences described herein may be utilized in such a “codon-optimized” form. Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, expression and/or activity of the gene product.

[0072] Polynucleotides may comprise a “heterologous nucleic acid,” whose sequence is from another species than the host cell or another strain of the same species of host cell. In some embodiments, the sequence is not identical to that of another nucleic acid naturally found in the same host cell. In some embodiments, a heterologous nucleic acid is not identical to a wild-type nucleic acid that is found in the same host cell in nature.

[0073] The polynucleotides of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a polynucleotide fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

[0074] Polynucleotides and fusions thereof may be prepared, manipulated and/or expressed using any of a variety of well-established techniques known and available in the art. For example, polynucleotide sequences which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to

direct expression of a selected enzyme in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

[0075] In some embodiments, the polypeptide is an isolated polypeptide. As used herein, an "isolated polypeptide" is not part of a library of polypeptides, such as a library of 2, 5, 10, 20, 50 or more different polypeptides and is separated from at least one component with which it occurs in nature. An isolated polypeptide can be obtained, for example, by expression of a recombinant nucleic acid encoding the polypeptide.

[0076] In some embodiments, the polypeptide is a heterologous polypeptide. By "heterologous polypeptide," it is meant a polypeptide whose amino acid sequence is not identical to that of another polypeptide naturally expressed in the same host cell. In particular, a heterologous polypeptide is not identical to a wild-type polypeptide that is found in the same host cell in nature.

[0077] In order to express a desired polypeptide, a nucleotide sequence encoding the polypeptide, or a functional equivalent, may be inserted into appropriate expression vector, *i.e.*, a vector that contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described in Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (1989), and Ausubel *et al.*, Current Protocols in Molecular Biology (1989).

[0078] "Polypeptide," "polypeptide fragment," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues are synthetic non-naturally occurring amino acids, such as a chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally-occurring amino acid polymers. Polypeptides include the polypeptides involved in the regulatable gene expression constructs as described herein (*e.g.* sigma factor), including genes

of interests, such as enzymatic polypeptides, or “enzymes,” which typically catalyze (*i.e.*, increase the rate of) various chemical reactions, such as the enzymes of metabolic pathways as described herein.

[0079] “Sequence identity,” as used herein, refers to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, I) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

[0080] Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include “reference sequence”, “comparison window”, “sequence identity”, “percentage of sequence identity” and “substantial identity”. A “reference sequence” is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (*i.e.*, only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a “comparison window” to identify and compare local regions of sequence similarity. A “comparison window” refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software

Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected.

[0081] Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, *Nucl. Acids Res.* 25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc., 1994-1998, Chapter 15.

[0082] Additionally, standard sequence alignment and/or structure prediction programs can be used to identify additional sigma factor or genes of interest polypeptides and nucleic acids based on the similarity of their primary and/or predicted polypeptide secondary structure with that of known sigma factor or genes of interest polypeptides and nucleic acids. Standard databases such as the SwissProt-Trembl database (world-wide web at "expasy.org", Swiss Institute of Bioinformatics Swiss-Prot group CMU-1 rue Michel Servet CH-1211 Geneva 4, Switzerland) can also be used to identify sigma factor or genes of interest polypeptides and nucleic acids. The secondary and/or tertiary structure of a sigma factor or genes of interest polypeptide can be predicted using the default settings of standard structure prediction programs, such as PredictProtein. Alternatively, the actual secondary and/or tertiary structure of a sigma factor or genes of interest polypeptide can be determined using standard methods.

B. Exemplary Vectors, Promoters and Other Elements

1. Vectors

[0083] Any of the RNA comprising a small molecule responsive riboswitch, sigma factor, one or more gene of interest and other polypeptide, encoded by nucleic acids described herein (alone or in any combination), as well as any sigma factor-dependent promoter, can be included in one or more vectors. Accordingly, provided herein are vectors with one or more sigma factor-dependent promoter or one or more nucleic acids encoding any of RNA comprising a small molecule responsive riboswitch, sigma factors, one or more genes of interest and other polypeptides that are described herein. As used herein, a "vector" means a construct that is capable of delivering, and desirably expressing, one or more nucleic acids of interest in a host cell. Examples of vectors include, but are not limited to, plasmids, viral

vectors, DNA or RNA expression vectors, cosmids, and phage vectors. In some embodiments, the vector contains a nucleic acid under the control of an expression control sequence.

[0084] As used herein, an "expression control sequence" means a nucleic acid sequence that directs transcription of a nucleic acid of interest. An expression control sequence can be a promoter, such as a constitutive or an inducible promoter, or an enhancer. An "inducible promoter" is a promoter that is active under environmental or developmental regulation, such as an arabinose-inducible promoter. The expression control sequence is operably linked to the nucleic acid segment to be transcribed.

[0085] In some embodiments, the vector contains a selective marker. The term "selective marker" refers to a nucleic acid capable of expression in a host cell that allows for ease of selection of those host cells containing an introduced nucleic acid or vector. Examples of selectable markers include, but are not limited to, antibiotic resistance nucleic acids (*e.g.*, erythromycin, chloramphenicol, thiamphenicol, kanamycin, ampicillin, carbenicillin, gentamicin, hygromycin, streptomycin, phleomycin, bleomycin, or neomycin,) and/or nucleic acids that confer a metabolic advantage, such as a nutritional advantage on the host cell.

[0086] Suitable vectors are those which are compatible with the host cell employed. Suitable vectors can be derived, for example, from a bacterium, a virus (such as bacteriophage T7 or an M-13 derived phage), a cosmid, a yeast, or a plant. In some instances, the suitable vector is a plasmid, such as a plasmid as described herein, such as those described in the Examples. Protocols for obtaining and using such vectors are known to those in the art (see, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to the use of vectors).

2. Promoters

[0087] Suitable promoters are used to express any of the heterologous nucleic acids described herein. Suitable promoters may be used to drive the production in host cells of RNA comprising a small molecule responsive riboswitch, sigma factors, one or more genes of interest and other nucleic acids and polypeptides as described herein, or to reduce degradation of RNA comprising a small molecule responsive riboswitch, sigma factors, one or more genes

of interest and nucleic acids and polypeptides as described herein. In some embodiments, a suitable promoter is used to express nucleic acid encoding an RNA comprising a small molecule responsive riboswitch operably linked to a sigma factor. Any of the promoters described herein may be used, including *E. coli* gapA promoter (See Example 1), *Streptomyces lividans* glucose isomerase p1.20 promoter (See Example 3), *Streptomyces lividans* glucose isomerase p1.5 promoter (See Example 5) and *E. coli* fliC promoter (See Example 4).

[0088] One of skill in the art can readily select and/or use suitable promoters to optimize the expression of RNA comprising a small molecule responsive riboswitch, or a sigma factor in a host cell. Similarly, one of skill in the art can readily select and/or use suitable sigma factor-dependent promoters to optimize the expression, or the control of the expression, of genes of interest as described herein.

[0089] In some instances, suitable sigma factor-dependent promoters may be used to control the expression of genes of interest, such that the sigma factor-dependent promoters are regulated by a sigma factor, which itself is regulated by a riboswitch as described herein. Any sigma factor-dependent promoter described herein may be used, including any FliA-dependent promoter such as *E. coli* fliC promoter (See Example 4).

[0090] Suitable promoters may be used to optimize the expression of RNA comprising a small molecule responsive riboswitch, or a sigma factor in a host cell. Any of the nucleic acids described herein (*e.g.*, a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor) may be operably linked to a promoter. Any of the promoters described herein may be used, including *E. coli* gapA promoter (See Example 1), *Streptomyces lividans* glucose isomerase p1.20 promoter (See Example 3) and *Streptomyces lividans* glucose isomerase p1.5 promoter (See Example 5).

[0091] High expression levels in certain bacterial cells may cause degradation of engineered polypeptide(s) including sigma factors and genes of interest. To improve production of sigma factors or genes of interest, an inducible expression system that allows both the timing and magnitude of expression of engineered polypeptide(s) to be controlled may be used. The tighter control may facilitate the expression of engineered polypeptide(s) at a concentration and period during the growth of the cells that is toxic to the cells, and results in the production of higher amounts of the desired polypeptide.

[0092] Any one of the promoters characterized or used in the Examples of the present disclosure may be used.

[0093] Promoters are well known in the art, and any promoter that functions in the host cell can be used for expression of a sigma factor in the host cell. Initiation control regions or promoters, which are useful to drive expression of polypeptides in various host cells are numerous and familiar to those skilled in the art (see, for example, WO 2004/033646 and references cited therein, which are each hereby incorporated by reference in their entireties, particularly with respect to vectors for the expression of nucleic acids of interest). Virtually any promoter capable of driving these nucleic acids is suitable for the present invention including, but not limited to, lac, trp, T7, tac, and trc, (useful for expression in *E. coli*).

3. Plasmids

[0094] In various embodiments, a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor, or a gene of interest, or any other nucleic acid or polypeptide as described herein, is contained in a low copy plasmid (*e.g.*, a plasmid that is maintained at about 1 to about 4 copies per cell), medium copy plasmid (*e.g.*, a plasmid that is maintained at about 10 to about 15 copies per cell), or high copy plasmid (*e.g.*, a plasmid that is maintained at about 50 or more copies per cell). In some embodiments, the nucleic acid encoding RNA comprising a small molecule responsive riboswitch and a sigma factor is operably linked to a *Streptomyces lividans* glucose isomerase p1.20 promoter or *Streptomyces lividans* glucose isomerase p1.5 promoter, such as is contained in pDMWP165, pDMWP180, or pDMWP181, as described in the Examples. In some embodiments, the nucleic acid encoding RNA comprising a small molecule responsive riboswitch and a sigma factor operably linked to a promoter is contained in a medium or high copy plasmid.

[0095] In some embodiments, the vector is a replicating plasmid that does not integrate into a chromosome in the cells. In some embodiments, part or all of the vector integrates into a chromosome in the cells. Additional examples of suitable expression and/or integration vectors are provided in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor, 1989, and Current Protocols in Molecular Biology (F. M. Ausubel et al. (eds) 1987, Supplement 30, section 7.7.18) which are both hereby incorporated by reference in

their entirety, particularly with respect to vectors. Particularly useful vectors include pFB6, pBR322, pUC18, pUC100, pBBR1, pCL1920, pCC2Fos, and pENTR/D.

4. Other Elements

[0096] Other molecular biology elements may also be used, such as termination sequence, origins of replication, and the like.

[0097] In some embodiments, the expression vector also includes a termination sequence. Termination control regions may also be derived from various genes native to the host cell. In some embodiments, the termination sequence and the promoter sequence are derived from the same source. In another embodiment, the termination sequence is endogenous to the host cell. Optionally, a termination site may be included. For effective expression of the polypeptides, DNA encoding the polypeptide are linked operably through initiation codons to selected expression control regions such that expression results in the formation of the appropriate messenger RNA.

[0098] A sigma factor-dependent promoter or a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor, or a gene of interest, or any other nucleic acid as described herein, can be incorporated into a vector, such as an expression vector, using standard techniques (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982, which is hereby incorporated by reference in its entirety, particularly with respect to the screening of appropriate DNA sequences and the construction of vectors). Methods used to ligate the DNA construct comprising a nucleic acid of interest (such as a sigma factor-dependent promoter or a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor, or a gene of interest), a promoter, a terminator, and other sequences and to insert them into a suitable vector are well known in the art. For example, restriction enzymes can be used to cleave the sigma factor-dependent promoter or a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor, or a gene of interest and the vector. Then, the compatible ends of the cleaved sigma factor-dependent promoter or a nucleic acid encoding RNA comprising a small molecule responsive riboswitch, or a sigma factor, or a gene of interest and the cleaved vector can be ligated. Linking is generally accomplished by ligation at convenient restriction sites. If

such sites do not exist, the synthetic oligonucleotide linkers are used in accordance with conventional practice (see, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor, 1989, and Bennett and Lasure, More Gene Manipulations in Fungi, Academic Press, San Diego, pp. 70-76, 1991, which are both hereby incorporated by reference in their entirety, particularly with respect to oligonucleotide linkers). Additionally, vectors can be constructed using known recombination techniques (e.g., Invitrogen Life Technologies, Gateway Technology).

[0099] Different types of origins of replication can be used. One, two or more origins of replication can be used. The origins of replication can be from different organisms and/or gram positive or gram negative organisms. Exemplary uses of origins of replication to practice the invention are further described in the Examples.

IV. Host Cells for Production of Industrial Bio-Products

[0100] One of skill in the art will recognize that expression vectors are designed to contain certain components which optimize gene expression for certain host strains. Such optimization components include, but are not limited to origin of replication, promoters, and enhancers. The vectors and components referenced herein are described for exemplary purposes and are not meant to narrow the scope of the invention.

[0101] Any microorganism or progeny thereof that can be used to heterologously express nucleic acids can be used with the methods and compositions for regulating gene expression in engineered cells described herein. Exemplary host cells include, for example, yeasts, such as species of *Saccharomyces* (e.g., *S. cerevisiae*), bacteria, such as species of *Escherichia* (e.g., *E. coli*), archaea, such as species of *Methanosarcina* (e.g., *Methanosarcina mazei*), plants, such as kudzu or poplar (e.g., *Populus alba* or *Populus alba x tremula* CAC35696) or aspen (e.g., *Populus tremuloides*).

[0102] Bacteria cells, including gram positive or gram negative bacteria can be used to express any of the nucleic acids or polypeptides described above. In some embodiments, the host cell is a gram-positive bacterium. Non-limiting examples include strains of *Streptomyces* (e.g., *S. lividans*, *S. coelicolor*, *S. rubiginosus*, or *S. griseus*), *Streptococcus*, *Bacillus* (e.g., *B. licheniformis* or *B. subtilis* or *B. coagulans*), *Listeria* (e.g., *L. monocytogenes*), *Corynebacteria*,

or *Lactobacillus* (e.g., *L. spp*). In some embodiments, the host organism is a gram-negative bacterium. Non-limiting examples include strains of *Escherichia* (e.g., *E. coli*), *Pseudomonas* (e.g., *P. alcaligenes*), *Pantoea* (e.g., *P. citrea*), *Enterobacter*, or *Helicobacter* (*H. pylori*). Other host cells for use in the methods and gene expression-control systems disclosed above include in any one of *P. citrea*, *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, *B. thuringiensis*, *S. albus*, *S. lividans*, *S. coelicolor*, *S. griseus*, *Pseudomonas sp.*, and *P. alcaligenes* cells.

[0103] There are numerous types of anaerobic cells that can be used as host cells in the compositions and methods of the present invention. In one aspect of the invention, the cells described in any of the compositions or methods described herein are obligate anaerobic cells and progeny thereof. Obligate anaerobes typically do not grow well, if at all, in conditions where oxygen is present. It is to be understood that a small amount of oxygen may be present, that is, there is some tolerance level that obligate anaerobes have for a low level of oxygen. In one aspect, obligate anaerobes engineered to produce isoprene can serve as host cells for any of the methods and/or compositions described herein and are grown under substantially oxygen-free conditions, wherein the amount of oxygen present is not harmful to the growth, maintenance, and/or fermentation of the anaerobes.

[0104] In another aspect of the invention, the host cells described and/or used in any of the compositions or methods described herein are facultative anaerobic cells and progeny thereof. Facultative anaerobes can generate cellular ATP by aerobic respiration (e.g., utilization of the TCA cycle) if oxygen is present. However, facultative anaerobes can also grow in the absence of oxygen. This is in contrast to obligate anaerobes which die or grow poorly in the presence of greater amounts of oxygen. In one aspect, therefore, facultative anaerobes can serve as host cells for any of the compositions and/or methods provided herein and can be engineered to produce isoprene. Facultative anaerobic host cells can be grown under substantially oxygen-free conditions, wherein the amount of oxygen present is not harmful to the growth, maintenance, and/or fermentation of the anaerobes, or can be alternatively grown in the presence of greater amounts of oxygen.

[0105] The host cell can additionally be a filamentous fungal cell and progeny thereof. (See, e.g., Berka & Barnett, *Biotechnology Advances*, 7(2):127-154 (1989)). In some aspects, the filamentous fungal cell can be any of *Trichoderma longibrachiatum*, *T. viride*, *T. koningii*, *T. harzianum*, *Penicillium* sp., *Humicola insolens*, *H. lanuginosa*, *H. grisea*, *Chrysosporium* sp., *C. lucknowense*, *Gliocladium* sp., *Aspergillus* sp., such as *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, or *A. awamori*, *Fusarium* sp., such as *F. roseum*, *F. gramineum*, *F. cerealis*, *F. oxysporum*, or *F. venenatum*, *Neurospora* sp., such as *N. crassa*, *Hypocrea* sp., *Mucor* sp., such as *M. miehei*, *Rhizopus* sp. or *Emericella* sp. In some aspects, the fungus is *A. nidulans*, *A. awamori*, *A. oryzae*, *A. aculeatus*, *A. niger*, *A. japonicus*, *T. reesei*, *T. viride*, *F. oxysporum*, or *F. solani*. In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. Patent Pub. No. US 2011/0045563.

[0106] The host cell can also be a yeast, such as *Saccharomyces* sp., *Schizosaccharomyces* sp., *Pichia* sp., or *Candida* sp. In some aspects, the *Saccharomyces* sp. is *Saccharomyces cerevisiae* (See, e.g., Romanos et al., *Yeast*, 8(6):423-488 (1992)). In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. pat. No. 7,659,097 and U.S. Patent Pub. No. US 2011/0045563.

[0107] The host cell can also be a species of plant, such as a plant from the family Fabaceae, such as the Faboideae subfamily. In some aspects, the host cell is kudzu, poplar (such as *Populus alba x tremula* CAC35696), aspen (such as *Populus tremuloides*), or *Quercus robur*.

[0108] The host cell can additionally be a species of algae, such as a green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates. (See, e.g., Saunders & Warmbrodt, "*Gene Expression in Algae and Fungi, Including Yeast*," (1993), National Agricultural Library, Beltsville, MD). In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. Patent Pub. No. US 2011/0045563. In some aspects, the host cell is a cyanobacterium, such as cyanobacterium classified into any of the following groups based on morphology: *Chlorococcales*, *Pleurocapsales*, *Oscillatoriales*, *Nostocales*, or *Stigonematales* (See, e.g., Lindberg et al., *Metab. Eng.*, 12(1):70-79 (2010)). In certain embodiments, plasmids or plasmid components for use herein

include those described in U.S. Patent Pub. No.: US 2010/0297749; US 2009/0282545 and PCT Pat. Appl. No. WO 2011/034863.

[0109] *E. coli* host cells can be used in any of the methods or systems for regulating the expression of a gene of interest disclosed herein. In one aspect, the host cell is a recombinant cell of an *Escherichia coli* (*E. coli*) strain, or progeny thereof, capable of producing one or more bio-products.

[0110] In other aspects, the host cell can be a species of yeast other than *S. cerevisiae* such as, but not limited to, a *Pichia* spp., a *Candida* spp., a *Hansenula* spp., a *Kluyveromyces* spp., a *Kluyveromyces* spp., or a *Schizosaccharomyces* spp. In still other aspects, the host cell can be a species of bacterium including, but not limited to, an *Arthrobacter* spp., a *Zymomonas* spp., a *Brevibacterium* spp., a *Clostridium* spp., an *Aerococcus* spp., a *Bacillus* spp., an *Actinobacillus* spp. (such as, but not limited to, *A. succinogens*), a *Carbobacterium* spp., a *Corynebacterium* spp., an *Enterococcus* spp., an *Erysipelothrix* spp., a *Gemella* spp., a *Geobacillus* spp., a *Globicatella* spp., a *Lactobacillus* spp. (such as, but not limited to, *L. lactis* and *L. rhammosus*), a *Lactococcus* spp., a *Leuconostoc* spp., a *Pediococcus* spp., a *Streptococcus* spp., a *Tetragenococcus* spp., an *Actinobacillus* spp., or a *Vagococcus* spp. In other aspects, the fermenting organism can be a fungus such as, but not limited to, a *Rhizopus* spp.

[0111] In other aspects, the host cell can be a lactic acid bacterium, such as those of the genera *Aerococcus*, *Bacillus*, *Carbobacterium*, *Enterococcus*, *Erysipelothrix*, *Gemella*, *Globicatella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, *Streptococcus*, *Tetragenococcus* or *Vagococcus*. For example, other bacteria of the genus *Lactobacillus* which may be substituted include, but are not limited to, *L. heiveticus*, *L. delbrueckii*, *L. casei*, *L. acidophilus*, *L. amylovorus*, *L. leichmanii*, *L. bulgaricus*, *L. amylovorus*, or *L. pentosus*.

V. Growth and/or Production Parameters

[0112] The host cells disclosed and compositions thereof can be engineered to produce industrial bio-product using the methods for regulating the expression of a gene of interest disclosed herein in a fermentation system. In one embodiment the system is substantially free of oxygen. In another embodiment, the system is an oxygen-containing system. In some embodiments, the fermentation system contains a carbohydrate as the energy and/or carbon

source. In some embodiments, the fermentation system contains carbohydrate and hydrogen as an energy and/or carbon source.

A. Small Molecules for Riboswitch Control

[0113] In some embodiments of the host cells disclosed and compositions thereof, the regulation of the expression of the gene of interest depends on the concentration of the small molecule that binds to the riboswitch. In some embodiments, the small molecule is present in an initial culture, or a seed stock, at a concentration of about 0.005 mg/L, 0.01 mg/L, 0.02 mg/L, 0.03 mg/L, 0.04 mg/L, or 0.05 mg/L. In some embodiments, the small molecule is present in the final culture, or production-scale bioreactor, at a concentration of about 0.0001 mg/L, 0.0002 mg/L, 0.0005 mg/L, 0.001 mg/L, 0.002 mg/L, 0.003 mg/L, or 0.004 mg/L. In some embodiments, the small molecule is present in an initial culture, or a seed stock, at a concentration of about 0.005 mg/L, 0.01 mg/L, 0.02 mg/L, 0.03 mg/L, 0.04 mg/L, or 0.05 mg/L, and the initial culture is diluted 5-fold, 10-fold, 15-fold, or 20-fold to provide the industrial scale bioreactor culture.

[0114] In some embodiments, the specific small molecule that binds to the riboswitch is selected from the group consisting of cobalamin (coenzyme of vitamin B12), cyclic di-MGMP, flavin mononucleotide (FMN), glucosamine-6-phosphate, glutamine, glycine, lysine, pre-queosine, purines, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), both SAH and SAM, and thiamine pyrophosphate (TPP). In some embodiments, the small molecule is cobalamin and is present in an initial culture, or a seed stock, at a concentration of about 0.005 mg/L, 0.01 mg/L, 0.02 mg/L, 0.03 mg/L, 0.04 mg/L, or 0.05 mg/L, and is present in the final culture, or production-scale bioreactor, at a concentration of about 0.0001 mg/L, 0.0002 mg/L, 0.0005 mg/L, 0.001 mg/L, 0.002 mg/L, 0.003 mg/L, or 0.004 mg/L. In some embodiments, the small molecule is present in an initial culture, or a seed stock, at a concentration of about 0.005 mg/L, 0.01 mg/L, 0.02 mg/L, 0.03 mg/L, 0.04 mg/L, or 0.05 mg/L, and the initial culture is diluted 5-fold, 10-fold, 15-fold, or 20-fold to provide the industrial scale bioreactor culture.

B. Feedstock

[0115] Various types of feedstock can be used for the recombinant microbial cells described herein. The feedstock can be a carbon source or syngas. Information regarding carbon sources available for use in exemplary feedstocks is provided below.

C. Carbon Source

[0116] Any carbon source can be used to cultivate the host cells. The term “carbon source” refers to one or more carbon-containing compounds capable of being metabolized by recombinant microbial cells described herein. For example, the cell medium used to cultivate the recombinant microbial cells described herein may include any carbon source suitable for maintaining the viability or growing the cells.

[0117] In some embodiments, the carbon source is a carbohydrate (such as monosaccharide, disaccharide, oligosaccharide, or polysaccharide), invert sugar (*e.g.*, enzymatically treated sucrose syrup), glycerol, glycerine (*e.g.*, a glycerine byproduct of a biodiesel or soap-making process), dihydroxyacetone, one-carbon source, oil (*e.g.*, a plant or vegetable oil such as corn, palm, or soybean oil), animal fat, animal oil, fatty acid (*e.g.*, a saturated fatty acid, unsaturated fatty acid, or polyunsaturated fatty acid), lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, polypeptide (*e.g.*, a microbial or plant protein or peptide), renewable carbon source (*e.g.*, a biomass carbon source such as a hydrolyzed biomass carbon source), yeast extract, component from a yeast extract, polymer, acid, alcohol, aldehyde, ketone, amino acid, succinate, lactate, acetate, ethanol, or any combination of two or more of the foregoing. In some embodiments, the carbon source is a product of photosynthesis, including, but not limited to, glucose.

[0118] Exemplary monosaccharides include glucose and fructose; exemplary oligosaccharides include lactose and sucrose; and exemplary polysaccharides include starch and cellulose. Exemplary carbohydrates include C6 sugars (*e.g.*, fructose, mannose, galactose, or glucose) and C5 sugars (*e.g.*, xylose or arabinose). In some embodiments, the cell medium includes a carbohydrate as well as a carbon source other than a carbohydrate (*e.g.*, glycerol, glycerine, dihydroxyacetone, one-carbon source, oil, animal fat, animal oil, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source,

or a component from a yeast extract). In some embodiments, the cell medium includes a carbohydrate as well as a polypeptide (*e.g.*, a microbial or plant protein or peptide). In some embodiments, the microbial polypeptide is a polypeptide from yeast or bacteria. In some embodiments, the plant polypeptide is a polypeptide from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

[0119] In some embodiments, the cells are cultured under limited glucose conditions. By “limited glucose conditions,” it is meant that the amount of glucose that is added is less than or about 105% (such as about 100%) of the amount of glucose that is consumed by the cells. In particular embodiments, the amount of glucose that is added to the culture medium is approximately the same as the amount of glucose that is consumed by the cells during a specific period of time. In some embodiments, the rate of cell growth is controlled by limiting the amount of added glucose such that the cells grow at the rate that can be supported by the amount of glucose in the cell medium. In some embodiments, glucose does not accumulate during the time the cells are cultured. In various embodiments, the cells are cultured under limited glucose conditions for greater than or about 1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, or 70 hours. In various embodiments, the cells are cultured under limited glucose conditions for greater than or about 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 95, or 100% of the total length of time the cells are cultured. While not intending to be bound by any particular theory, it is believed that limited glucose conditions may allow more favorable regulation of the cells.

[0120] In some embodiments, the cells are cultured in the presence of an excess of glucose. In particular embodiments, the amount of glucose that is added is greater than about 105% (such as about or greater than 110, 120, 150, 175, 200, 250, 300, 400, or 500%) or more of the amount of glucose that is consumed by the cells during a specific period of time. In some embodiments, glucose accumulates during the time the cells are cultured.

[0121] Exemplary lipids are any substance containing one or more fatty acids that are C4 and above fatty acids that are saturated, unsaturated, or branched.

[0122] Exemplary oils are lipids that are liquid at room temperature. In some embodiments, the lipid contains one or more C4 or above fatty acids (*e.g.*, contains one or

more saturated, unsaturated, or branched fatty acid with four or more carbons). In some embodiments, the oil is obtained from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, linseed, oleagineous microbial cells, Chinese tallow, or any combination of two or more of the foregoing.

[0123] Exemplary fatty acids include compounds of the formula RCOOH , where “R” is a hydrocarbon. Exemplary unsaturated fatty acids include compounds where “R” includes at least one carbon-carbon double bond. Exemplary unsaturated fatty acids include, but are not limited to, oleic acid, vaccenic acid, linoleic acid, palmitelaidic acid, and arachidonic acid. Exemplary polyunsaturated fatty acids include compounds where “R” includes a plurality of carbon-carbon double bonds. Exemplary saturated fatty acids include compounds where “R” is a saturated aliphatic group. In some embodiments, the carbon source includes one or more C_{12} - C_{22} fatty acids, such as a C_{12} saturated fatty acid, a C_{14} saturated fatty acid, a C_{16} saturated fatty acid, a C_{18} saturated fatty acid, a C_{20} saturated fatty acid, or a C_{22} saturated fatty acid. In an exemplary embodiment, the fatty acid is palmitic acid. In some embodiments, the carbon source is a salt of a fatty acid (*e.g.*, an unsaturated fatty acid), a derivative of a fatty acid (*e.g.*, an unsaturated fatty acid), or a salt of a derivative of fatty acid (*e.g.*, an unsaturated fatty acid). Suitable salts include, but are not limited to, lithium salts, potassium salts, sodium salts, and the like. Di- and triglycerols are fatty acid esters of glycerol.

[0124] In some embodiments, the concentration of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is at least or about 1 gram per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400, or more g/L. In some embodiments, the concentration of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 10 and about 400 g/L, such as between about 25 and about 300 g/L, between about 60 and about 180 g/L, or between about 75 and about 150 g/L. In some embodiments, the concentration includes the total amount of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both (i) a lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride and (ii) a carbohydrate, such as glucose. In some embodiments, the ratio of the lipid, oil, fat, fatty acid,

monoglyceride, diglyceride, or triglyceride to the carbohydrate is about 1:1 on a carbon basis (*i.e.*, one carbon in the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride per carbohydrate carbon). In particular embodiments, the amount of the lipid, oil, fat, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 60 and 180 g/L, and the amount of the carbohydrate is between about 120 and 360 g/L.

[0125] Exemplary microbial polypeptide carbon sources include one or more polypeptides from yeast or bacteria. Exemplary plant polypeptide carbon sources include one or more polypeptides from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

[0126] Exemplary renewable carbon sources include cheese whey permeate, cornsteep liquor, sugar beet molasses, barley malt, and components from any of the foregoing. Exemplary renewable carbon sources also include glucose, hexose, pentose and xylose present in biomass, such as corn, switchgrass, sugar cane, cell waste of fermentation processes, and protein by-product from the milling of soy, corn, or wheat. In some embodiments, the biomass carbon source is a lignocellulosic, hemicellulosic, or cellulosic material such as, but are not limited to, a grass, wheat, wheat straw, bagasse, sugar cane bagasse, soft wood pulp, corn, corn cob or husk, corn kernel, fiber from corn kernels, corn stover, switch grass, rice hull product, or a by-product from wet or dry milling of grains (*e.g.*, corn, sorghum, rye, triticate, barley, wheat, and/or distillers grains). Exemplary cellulosic materials include wood, paper and pulp waste, herbaceous plants, and fruit pulp. In some embodiments, the carbon source includes any plant part, such as stems, grains, roots, or tubers. In some embodiments, all or part of any of the following plants are used as a carbon source: corn, wheat, rye, sorghum, triticate, rice, millet, barley, cassava, legumes, such as beans and peas, potatoes, sweet potatoes, bananas, sugarcane, and/or tapioca. In some embodiments, the carbon source is a biomass hydrolysate, such as a biomass hydrolysate that includes both xylose and glucose or that includes both sucrose and glucose.

[0127] In some embodiments, the renewable carbon source (such as biomass) is pretreated before it is added to the cell culture medium. In some embodiments, the pretreatment includes enzymatic pretreatment, chemical pretreatment, or a combination of both enzymatic and chemical pretreatment (*see*, for example, Farzaneh *et al.*, *Bioresource Technology* 96 (18): 2014-

2018, 2005; U.S. Patent No. 6,176,176; U.S. Patent No. 6,106,888; which are each hereby incorporated by reference in their entireties, particularly with respect to the pretreatment of renewable carbon sources). In some embodiments, the renewable carbon source is partially or completely hydrolyzed before it is added to the cell culture medium.

[0128] In some embodiments, the renewable carbon source (such as corn stover) undergoes ammonia fiber expansion (AFEX) pretreatment before it is added to the cell culture medium (*see*, for example, Farzaneh *et al.*, *Bioresource Technology* 96 (18): 2014-2018, 2005). During AFEX pretreatment, a renewable carbon source is treated with liquid anhydrous ammonia at moderate temperatures (such as about 60 to about 100 °C) and high pressure (such as about 250 to about 300 psi) for about 5 minutes. Then, the pressure is rapidly released. In this process, the combined chemical and physical effects of lignin solubilization, hemicellulose hydrolysis, cellulose decrystallization, and increased surface area enables near complete enzymatic conversion of cellulose and hemicellulose to fermentable sugars. AFEX pretreatment has the advantage that nearly all of the ammonia can be recovered and reused, while the remaining serves as nitrogen source for microbes in downstream processes. Also, a wash stream is not required for AFEX pretreatment. Thus, dry matter recovery following the AFEX treatment is essentially 100%. AFEX is basically a dry to dry process. The treated renewable carbon source is stable for long periods and can be fed at very high solid loadings in enzymatic hydrolysis or fermentation processes. Cellulose and hemicellulose are well preserved in the AFEX process, with little or no degradation. There is no need for neutralization prior to the enzymatic hydrolysis of a renewable carbon source that has undergone AFEX pretreatment. Enzymatic hydrolysis of AFEX-treated carbon sources produces clean sugar streams for subsequent fermentation use.

[0129] In some embodiments, the concentration of the carbon source (*e.g.*, a renewable carbon source) is equivalent to at least or about 0.1, 0.5, 1, 1.5, 2, 3, 4, 5, 10, 15, 20, 30, 40, or 50% glucose (w/v). The equivalent amount of glucose can be determined by using standard HPLC methods with glucose as a reference to measure the amount of glucose generated from the carbon source. In some embodiments, the concentration of the carbon source (*e.g.*, a renewable carbon source) is equivalent to between about 0.1 and about 20% glucose, such as between about 0.1 and about 10% glucose, between about 0.5 and about 10% glucose, between

about 1 and about 10% glucose, between about 1 and about 5% glucose, or between about 1 and about 2% glucose.

[0130] In some embodiments, the carbon source includes yeast extract or one or more components of yeast extract. In some embodiments, the concentration of yeast extract is at least 1 gram of yeast extract per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, or more g/L. In some embodiments, the concentration of yeast extract is between about 1 and about 300 g/L, such as between about 1 and about 200 g/L, between about 5 and about 200 g/L, between about 5 and about 100 g/L, or between about 5 and about 60 g/L. In some embodiments, the concentration includes the total amount of yeast extract that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both yeast extract (or one or more components thereof) and another carbon source, such as glucose. In some embodiments, the ratio of yeast extract to the other carbon source is about 1:5, about 1:10, or about 1:20 (w/w).

[0131] Additionally the carbon source may also be one-carbon substrates such as carbon dioxide, or methanol. Glycerol production from single carbon sources (*e.g.*, methanol, formaldehyde, or formate) has been reported in methylotrophic yeasts (Yamada *et al.*, *Agric. Biol. Chem.*, 53(2) 541-543, 1989, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources) and in bacteria (Hunter *et. al.*, *Biochemistry*, 24, 4148-4155, 1985, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). These organisms can assimilate single carbon compounds, ranging in oxidation state from methane to formate, and produce glycerol. The pathway of carbon assimilation can be through ribulose monophosphate, through serine, or through xylulose-momophosphate (Gottschalk, Bacterial Metabolism, Second Edition, Springer-Verlag: New York, 1986, which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). The ribulose monophosphate pathway involves the condensation of formate with ribulose-5-phosphate to form a six carbon sugar that becomes fructose and eventually the three carbon product glyceraldehyde-3-phosphate. Likewise, the serine pathway assimilates the one-carbon compound into the glycolytic pathway via methylenetetrahydrofolate.

D. Syngas

[0132] Syngas (also referred to as synthesis gas) can be used as a source of energy and/or carbon for any of the recombinant host cells described herein. Syngas can include CO and H₂. In some aspects, the syngas comprises CO, CO₂, and H₂. In some aspects, the syngas further comprises H₂O and/or N₂. For example, the syngas may comprise CO, H₂, and H₂O (*e.g.*, CO, H₂, H₂O and N₂). The syngas may comprise CO, H₂, and N₂. The syngas may comprise CO, CO₂, H₂, and H₂O (*e.g.*, CO, CO₂, H₂, H₂O and N₂). The syngas may comprise CO, CO₂, H₂, and N₂. The CO and/or CO₂ in the syngas may be used as carbon source for cells.

[0133] In some aspects, the molar ratio of hydrogen to carbon monoxide in the syngas is about any of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 3.0, 4.0, 5.0, or 10.0. In some aspects, the syngas comprises about any of 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume carbon monoxide. In some aspects, the syngas comprises about any of 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume hydrogen. In some aspects, the syngas comprises about any of 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume carbon dioxide. In some aspects, the syngas comprises about any of 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume water. In some aspects, the syngas comprises about any of 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume nitrogen.

[0134] The syngas of the present invention may be derived from natural or synthetic sources. In some aspects, the syngas is derived from biomass (*e.g.*, wood, switch grass, agriculture waste, municipal waste) or carbohydrates (*e.g.*, sugars). In other aspects, the syngas is derived from coal, petroleum, kerogen, tar sands, oil shale, natural gas, or a mixture thereof. In other aspects, the syngas is derived from rubber, such as from rubber tires. In some aspects, the syngas is derived from a mixture (*e.g.*, blend) of biomass and coal. In some aspects, the mixture has about or at least about any of 1%, 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 90%, 95%, or 99% biomass. In some aspects, the mixture has about or at least about any of 1%, 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 90%, 95%, or 99% coal. In some aspects, the ratio of biomass to coal in the mixture is about any of 5:95, 10:90, 15:85, 20:80, 25:75, 30:70, 35:65, 40:60, 45:55, 50:50, 55:45, 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10, or 95:5.

[0135] Syngas can be derived from a feedstock by a variety of processes, including methane reforming, coal liquefaction, co-firing, fermentative reactions, enzymatic reactions, and biomass gasification. Biomass gasification is accomplished by subjecting biomass to partial oxidation in a reactor at temperatures above about 700 °C in the presence of less than a stoichiometric amount of oxygen. The oxygen is introduced into the bioreactor in the form of air, pure oxygen, or steam. Gasification can occur in three main steps: 1) initial heating to dry out any moisture embedded in the biomass; 2) pyrolysis, in which the biomass is heated to 300-500 °C in the absence of oxidizing agents to yield gas, tars, oils and solid char residue; and 3) gasification of solid char, tars and gas to yield the primary components of syngas. Co-firing is accomplished by gasification of a coal/biomass mixture. The composition of the syngas, such as the identity and molar ratios of the components of the syngas, can vary depending on the feedstock from which it is derived and the method by which the feedstock is converted to syngas.

[0136] Syngas can contain impurities, the nature and amount of which vary according to both the feedstock and the process used in production. Fermentations may be tolerant to some impurities, but there remains the need to remove from the syngas materials such as tars and particulates that might foul the fermentor and associated equipment. It is also advisable to remove compounds that might contaminate the isoprene product such as volatile organic compounds, acid gases, methane, benzene, toluene, ethylbenzene, xylenes, H₂S, COS, CS₂, HCl, O₃, organosulfur compounds, ammonia, nitrogen oxides, nitrogen-containing organic compounds, and heavy metal vapors. Removal of impurities from syngas can be achieved by one of several means, including gas scrubbing, treatment with solid-phase adsorbents, and purification using gas-permeable membranes.

[0137] Examples of other fermentation systems and culture conditions which can be used are described in International Patent Application Publication Nos. WO2009/076676, WO2010/003007, WO2009/132220, WO2010/031062, WO2010/031068, WO2010/031076, WO2010/013077, WO2010/031079, WO2010/148150, WO2010/078457, and WO2010/148256, which are hereby incorporated in their entirety, particularly with respect to fermentation systems and culture conditions for host cells as described herein.

[0138] In some aspects, the culture medium is prepared using anoxic techniques. In some aspects, the culture medium comprises one or more of NH_4Cl , NaCl , KCl , KH_2PO_4 , $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, NaHCO_3 , yeast extract, cysteine hydrochloride, $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$, trace metals, and vitamins. In some aspects, the culture medium contains, per liter, about 1.0 g NH_4Cl , about 0.8 g NaCl , about 0.1 g KCl , about 0.1 g KH_2PO_4 , about 0.2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, about 0.02 g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, about 1.0 g NaHCO_3 , about 1.0 g yeast extract, about 0.2 g cysteine hydrochloride, about 0.2 g $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$, about 10 mL trace metal solution, and about 10 mL vitamin solution. In some aspects, the culture condition comprises mevalonate.

[0139] The growth conditions, carbon sources, energy sources, and culture media may be according to any of the growth conditions, carbon sources, energy sources, and culture media described in the Examples of the present disclosure.

VI. Methods of Using Engineered Microorganisms for Production of Industrial Bio-Products

[0140] The invention provides for microbial expression systems for the production of one or more industrial bio-products (*e.g.*, isoprene, butadiene, or ethanol). In some embodiments, the system can include one or more of: a regulatable gene expression construct comprising (i) a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest.

[0141] In some embodiments, the system provides for the expression of one or more nucleic acids of interest (*e.g.*, nucleic acids encoding isoprene synthase or enzymes involved in the production of ethanol from acetyl-CoA). As described herein, microorganisms expressing one or more nucleic acids of interest can be engineered to produce various industrial bio-products under the control of a regulatable gene expression construct, such as any of those disclosed herein. These bio-products can include, but are not limited to, isoprene, butadiene, ethanol, propanediol (*e.g.*, 1,2-propanediol, 1,3-propanediol), hydrogen, acetate, microbial fuels, non-fermentative alcohols, fatty alcohols, fatty acid esters, isoprenoid alcohols, alkenes, alkanes, terpenoids, isoprenoids, carotenoids or other C5, C10, C15, C20, C25, C30, C35, or C40 product. The production of these industrial bio-products is described in further detail below and herein.

[0142] As described herein, the constructs, compositions, and methods for regulating and controlling gene expression can be used to engineer microorganism host cells responsive to one or more small molecules. When cultured in the presence of these small molecules, heterologous nucleic acid expression is decreased or absent. However, when the small molecules are removed from the culture medium, heterologous gene expression is no longer repressed or absent and the engineered microorganisms can produce various industrial bio-products, including but not limited to, isoprene, butadiene, ethanol, propanediol (*e.g.*, 1,2-propanediol, 1,3-propanediol), hydrogen, acetate, microbial fuels, non-fermentative alcohols, fatty alcohols, fatty acid esters, isoprenoid alcohols, alkenes, alkanes, terpenoids, isoprenoids, carotenoids or other C5, C10, C15, C20, C25, C30, C35, or C40 product. The production of these industrial bio-products is described in further detail below and herein.

A. Isoprene Production

[0143] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells that contain one or more pathways for the production of isoprene (*e.g.*, host cells as described herein that contain the pathways illustrated in FIG. 11 to FIG. 15) with one or more heterologous polynucleotides encoding one or more isoprene pathway enzymes expressed in a sufficient amount to produce isoprene.

1. Exemplary Isoprene Synthase Polypeptides and Nucleic Acids

[0144] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein with polynucleotides encoding an isoprene synthase polypeptide. Isoprene synthase polypeptides convert dimethylallyl diphosphate (DMAPP) into isoprene. Exemplary isoprene synthase polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an isoprene synthase polypeptide. Standard methods can be used to determine whether a polypeptide has isoprene synthase polypeptide activity by measuring the ability of the polypeptide to convert DMAPP into isoprene *in vitro*, in a cell extract, or *in vivo* (*e.g.*, as described in Example 1 of US 8420360 B2, which is incorporated herein in its entirety, particularly with respect to methods for assessing isoprene synthase activity). Isoprene synthase polypeptide activity in cell extracts can be measured, for example, as described in Silver et al., *J. Biol. Chem.*

270:13010-13016, 1995 and references therein, which are each hereby incorporated by reference in their entireties, particularly with respect to assays for isoprene synthase polypeptide activity.

[0145] In some embodiments, the isoprene synthase polypeptide or nucleic acid is from the family *Fabaceae*, such as the *Faboideae* subfamily. In some embodiments, the isoprene synthase polypeptide or nucleic acid is a naturally-occurring polypeptide or nucleic acid from *Pueraria montana* (kudzu) (Sharkey et al., *Plant Physiology* 137: 700-712, 2005), *Pueraria lobata*, poplar (such as *Populus alba* x *tremula* CAC35696) (Miller et al., *Planta* 213: 483-487, 2001) aspen (such as *Populus tremuloides*) (Silver et al, *JBC* 270(22): 13010-1316, 1995), or English Oak (*Quercus robur*) (Zimmer et al., WO 98/02550), which are each hereby incorporated by reference in their entireties, particularly with respect to isoprene synthase nucleic acids and the expression of isoprene synthase polypeptides. Suitable isoprene synthases include, but are not limited to, those identified by Genbank Accession Nos. AY341431, AY316691, AY279379, AJ457070, and AY1 82241, which are each hereby incorporated by reference in their entireties, particularly with respect to sequences of isoprene synthase nucleic acids and polypeptides. In some embodiments, the isoprene synthase polypeptide or nucleic acid is not a naturally-occurring polypeptide or nucleic acid from *Quercus robur* (i.e., the isoprene synthase polypeptide or nucleic acid is an isoprene synthase polypeptide or nucleic acid other than a naturally-occurring polypeptide or nucleic acid from *Quercus robur*). In some embodiments, the isoprene synthase nucleic acid or polypeptide is not a naturally-occurring polypeptide or nucleic acid from poplar (such as *Populus alba* x *tremula* CAC35696).

[0146] Exemplary isoprene synthase nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an isoprene synthase polypeptide. Exemplary isoprene synthase polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein, or as described in U.S. Patent No. 8,173,410, U.S. Patent No. 8,735,134 and U.S. Publ. No. 2013/0045891, the disclosures of which are hereby incorporated by reference in their entirety, particularly with respect to isoprene synthase polypeptides and nucleic acids.

2. Exemplary DXP Pathway Polypeptides and Nucleic Acids

[0147] In some aspects of the invention, the cells described in any of the compositions or methods described herein further comprise one or more heterologous nucleic acids encoding a DXS polypeptide or other DXP pathway polypeptides. In some aspects, the cells further comprise a chromosomal copy of an endogenous nucleic acid encoding a DXS polypeptide or other DXP pathway polypeptides. In some aspects, the cells further comprise one or more nucleic acids encoding an IDI polypeptide and a DXS polypeptide or other DXP pathway polypeptides. In some aspects, one nucleic acid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide or other DXP pathway polypeptides. In some aspects, one plasmid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide or other DXP pathway polypeptides. In some aspects, multiple plasmids encode the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide or other DXP pathway polypeptides.

[0148] Exemplary DXP pathways polypeptides include, but are not limited to any of the following polypeptides: DXS polypeptides, DXR polypeptides, MCT polypeptides, CMK polypeptides, MCS polypeptides, HDS polypeptides, HDR polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of one, two, or more of the DXP pathway polypeptides. In particular, DXP pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXP pathway polypeptide. Exemplary DXP pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a DXP pathway polypeptide. Exemplary DXP pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein. Exemplary DXP pathway polypeptides and nucleic acids and methods of measuring DXP pathway polypeptide activity are described in more detail in International Publication No. WO 2010/148150.

[0149] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein with polynucleotides encoding l-deoxy-D-xylulose-5-phosphate synthase (DXS) polypeptides. DXS polypeptides convert pyruvate and D-

glyceraldehyde-3-phosphate into l-deoxy-D-xylulose-5-phosphate. Exemplary DXS polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXS polypeptide. Standard methods can be used to determine whether a polypeptide has DXS polypeptide activity by measuring the ability of the polypeptide to convert pyruvate and D-glyceraldehyde-3-phosphate into l-deoxy-D-xylulose-5-phosphate *in vitro*, in a cell extract, or *in vivo* (see, *e.g.*, US 8420360 B2, which is hereby incorporated herein in its entirety, particularly with respect to methods of assessing DXS polypeptide activity). Exemplary DXS nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a DXS polypeptide. Exemplary DXS polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

[0150] DXR polypeptides convert 1-deoxy-D-xylulose 5-phosphate (DXP) into 2-C-methyl-D-erythritol 4-phosphate (MEP). Standard methods can be used to determine whether a polypeptide has DXR polypeptides activity by measuring the ability of the polypeptide to convert DXP *in vitro*, in a cell extract, or *in vivo*.

[0151] MCT polypeptides convert 2-C-methyl-D-erythritol 4-phosphate (MEP) into 4-(cytidine 5'-diphospho)-2-methyl-D-erythritol (CDP-ME). Standard methods can be used to determine whether a polypeptide has MCT polypeptides activity by measuring the ability of the polypeptide to convert MEP *in vitro*, in a cell extract, or *in vivo*.

[0152] CMK polypeptides convert 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-ME) into 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-MEP). Standard methods can be used to determine whether a polypeptide has CMK polypeptides activity by measuring the ability of the polypeptide to convert CDP-ME *in vitro*, in a cell extract, or *in vivo*.

[0153] MCS polypeptides convert 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-MEP) into 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate (ME-CPP or cMEPP). Standard methods can be used to determine whether a polypeptide has MCS

polypeptides activity by measuring the ability of the polypeptide to convert CDP-MEP *in vitro*, in a cell extract, or *in vivo*.

[0154] HDS polypeptides convert 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate into (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate (HMBPP or HDMAPP). Standard methods can be used to determine whether a polypeptide has HDS polypeptides activity by measuring the ability of the polypeptide to convert ME-CPP *in vitro*, in a cell extract, or *in vivo*.

[0155] HDR polypeptides convert (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Standard methods can be used to determine whether a polypeptide has HDR polypeptides activity by measuring the ability of the polypeptide to convert HMBPP *in vitro*, in a cell extract, or *in vivo*.

3. Exemplary IDI Polypeptides and Nucleic Acids

[0156] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein with polynucleotides encoding isopentenyl diphosphate isomerase polypeptides (isopentenyl-diphosphate delta-isomerase or IDI). IDI catalyzes the interconversion of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (*e.g.*, converting IPP into DMAPP and/or converting DMAPP into IPP). Exemplary IDI polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an IDI polypeptide. Standard methods can be used to determine whether a polypeptide has IDI polypeptide activity by measuring the ability of the polypeptide to interconvert IPP and DMAPP *in vitro*, in a cell extract, or *in vivo* (*see, e.g.*, US 8420360 B2, which is hereby incorporated by reference in its entirety, particularly with respect to assays for IDI activity). Exemplary IDI nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an IDI polypeptide. Exemplary IDI polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

4. Exemplary MVA Pathway Polypeptides and Nucleic Acids

[0157] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein with polynucleotides encoding MVA pathway polypeptides. MVA pathway polypeptides include acetyl-CoA acetyltransferase (AA-CoA thiolase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides, mevalonate kinase (MVK) polypeptides, phosphomevalonate kinase (PMK) polypeptides, diphosphomevalonate decarboxylase (MVD) polypeptides, IDI polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of two or more MVA pathway polypeptides. In a modified MVA pathway, phosphomevalonate decarboxylase (PMDC) and isopentenyl phosphate kinase (IPK) are included in place of phosphomevalonate kinase (PMK) and diphosphomevalonate decarboxylase (MVD). In particular, MVA pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

[0158] In particular, acetyl-CoA acetyltransferase polypeptides (AA-CoA thiolase or AACT) convert two molecules of acetyl-CoA into acetoacetyl-CoA. Standard methods (such as those described herein) can be used to determine whether a polypeptide has AA-CoA thiolase polypeptide activity by measuring the ability of the polypeptide to convert two molecules of acetyl-CoA into acetoacetyl-CoA *in vitro*, in a cell extract, or *in vivo*.

[0159] 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase or HMGS) polypeptides convert acetoacetyl-CoA into S-hydroxy-S-methylglutaryl-CoA. Standard methods (such as those described herein) can be used to determine whether a polypeptide has HMG-CoA synthase polypeptide activity by measuring the ability of the polypeptide to convert acetoacetyl-CoA into 3-hydroxy-3-methylglutaryl-CoA *in vitro*, in a cell extract, or *in vivo*.

[0160] 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase or HMGR) polypeptides convert 3-hydroxy-3-methylglutaryl-CoA into mevalonate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has HMG-CoA reductase polypeptide activity by measuring the ability of the polypeptide to convert 3-hydroxy-3-methylglutaryl-CoA into mevalonate *in vitro*, in a cell extract, or *in vivo*.

[0161] Mevalonate kinase (MVK) polypeptides phosphorylate mevalonate to form mevalonate-5-phosphate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has MVK polypeptide activity by measuring the ability of the polypeptide to convert mevalonate into mevalonate-5-phosphate *in vitro*, in a cell extract, or *in vivo*.

[0162] Phosphomevalonate kinase (PMK) polypeptides phosphorylate mevalonate-5-phosphate to form mevalonate-5-diphosphate. Standard methods (such as those described herein) can be used to determine whether a polypeptide has PMK polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-phosphate into mevalonate-5-diphosphate *in vitro*, in a cell extract, or *in vivo*.

[0163] Diphosphomevalonate decarboxylase (MVD or DPMDC) polypeptides convert mevalonate-5-diphosphate into isopentenyl diphosphate polypeptides (IPP). Standard methods (such as those described) can be used to determine whether a polypeptide has MVD polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-diphosphate into IPP *in vitro*, in a cell extract, or *in vivo*.

[0164] Phosphomevalonate decarboxylase (PMDC) polypeptides convert mevalonate-5-phosphate into isopentenyl phosphate (IP). Standard methods (such as those described) can be used to determine whether a polypeptide has PMDC polypeptide activity by measuring the ability of the polypeptide to convert mevalonate-5-phosphate into isopentenyl phosphate *in vitro*, in a cell extract, or *in vivo*.

[0165] Isopentenyl phosphate kinase (IPK) polypeptides convert isopentenyl phosphate into isopentenyl diphosphate. Standard methods (such as those described) can be used to determine whether a polypeptide has IPK polypeptide activity by measuring the ability of the

polypeptide to convert isopentenyl phosphate into isopentenyl diphosphate *in vitro*, in a cell extract, or *in vivo*.

[0166] In some embodiments, the compositions and methods described herein can be used to transform host cells as described herein that have been engineered to produce isoprene from syngas and/or from carbohydrates or mixtures thereof.

5. Exemplary Source Organisms

[0167] Isoprene synthase, DXP pathway, IDI, or MVA pathway nucleic acids (and their encoded polypeptides) can be obtained from any organism that naturally contains isoprene synthase, DXP pathway, IDI, and/or MVA pathway nucleic acids. As noted above, isoprene is formed naturally by a variety of organisms, such as bacteria, yeast, plants, and animals. Organisms contain the MVA pathway, DXP pathway, or both the MVA and DXP pathways for producing isoprene. Thus, DXS, DXR, MCT, CMK, MCS, HDS, or HDR nucleic acids can be obtained, *e.g.*, from any organism that contains the DXP pathway or contains both the MVA and DXP pathways. IDI and isoprene synthase nucleic acids can be obtained, *e.g.*, from any organism that contains the MVA pathway, DXP pathway, or both the MVA and DXP pathways. MVA pathway nucleic acids can be obtained, *e.g.*, from any organism that contains the MVA pathway or contains both the MVA and DXP pathways.

[0168] In some embodiments, the nucleic acid sequence of the isoprene synthase, DXP pathway, IDI, or MVA pathway nucleic acid is identical to the sequence of a nucleic acid that is produced by any of the following organisms in nature. In some embodiments, the amino acid sequence of the isoprene synthase, DXP pathway, IDI, or MVA pathway polypeptide is identical to the sequence of a polypeptide that is produced by any of the following organisms in nature. In some embodiments, the isoprene synthase, DXP pathway, IDI, or MVA pathway nucleic acid or polypeptide is a mutant nucleic acid or polypeptide derived from any of the organisms described herein. As used herein, “derived from” refers to the source of the nucleic acid or polypeptide into which one or more mutations is introduced. For example, a polypeptide that is “derived from a plant polypeptide” refers to polypeptide of interest that results from introducing one or more mutations into the sequence of a wild-type (*i.e.*, a sequence occurring in nature) plant polypeptide.

[0169] In some embodiments, the source organism is a fungus, examples of which are species of *Aspergillus* such as *A. oryzae* and *A. niger*, species of *Saccharomyces* such as *S. cerevisiae*, species of *Schizosaccharomyces* such as *S. pombe*, and species of *Trichoderma* such as *T. reesei*. In some embodiments, the source organism is a filamentous fungal cell. The term “filamentous fungi” refers to all filamentous forms of the subdivision Eumycotina (see, Alexopoulos, C. J. (1962), *Introductory Mycology*, Wiley, New York). These fungi are characterized by a vegetative mycelium with a cell wall composed of chitin, cellulose, and other complex polysaccharides. The filamentous fungi are morphologically, physiologically, and genetically distinct from yeasts. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism is obligatory aerobic. The filamentous fungal parent cell may be a cell of a species of, but not limited to, *Trichoderma*, (e.g., *Trichoderma reesei*, the asexual morph of *Hypocrea jecorina*, previously classified as *T. longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*) (Sheir-Neirs *et al.*, Appl. Microbiol. Biotechnol 20: 46-53, 1984; ATCC No. 56765 and ATCC No. 26921); *Penicillium* sp., *Humicola* sp. (e.g., *H. insolens*, *H. lanuginosa*, or *H. grisea*); *Chrysosporium* sp. (e.g., *C. lucknowense*), *Glucocladium* sp., *Aspergillus* sp. (e.g., *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, or *A. awamori*) (Ward *et al.*, Appl. Microbiol. Biotechnol. 39: 7380743, 1993 and Goedegebuur *et al.*, Genet 41: 89-98, 2002), *Fusarium* sp., (e.g., *F. roseum*, *F. gramineum*, *F. cerealis*, *F. oxysporum*, or *F. venenatum*), *Neurospora* sp., (e.g., *N. crassa*), *Hypocrea* sp., *Mucor* sp., (e.g., *M. miehei*), *Rhizopus* sp. and *Emericella* sp. (see also, Innis *et al.*, Sci. 228: 21-26, 1985). The term “*Trichoderma*” or “*Trichoderma* sp.” or “*Trichoderma* spp.” refer to any fungal genus previously or currently classified as *Trichoderma*.

[0170] In some embodiments, the fungus is *A. nidulans*, *A. awamori*, *A. oryzae*, *A. aculeatus*, *A. niger*, *A. japonicus*, *T. reesei*, *T. viride*, *F. oxysporum*, or *F. solani*. *Aspergillus* strains are disclosed in Ward *et al.*, Appl. Microbiol. Biotechnol. 39:738-743, 1993 and Goedegebuur *et al.*, Curr Gene 41:89-98, 2002, which are each hereby incorporated by reference in their entireties, particularly with respect to fungi. In particular embodiments, the fungus is a strain of *Trichoderma*, such as a strain of *T. reesei*. Strains of *T. reesei* are known and non-limiting examples include ATCC No. 13631, ATCC No. 26921, ATCC No. 56764, ATCC No. 56765, ATCC No. 56767, and NRRL 15709, which are each hereby incorporated by reference in their entireties, particularly with respect to strains of *T. reesei*. In some

embodiments, the host strain is a derivative of RL-P37. RL-P37 is disclosed in Sheir-Neiss *et al.*, Appl. Microbiol. Biotechnology 20:46–53, 1984, which is hereby incorporated by reference in its entirety, particularly with respect to strains of *T. reesei*.

[0171] In some embodiments, the source organism is a yeast, such as *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.*

[0172] In some embodiments, the source organism is a bacterium, such as strains of *Bacillus* such as *B. lichenformis* or *B. subtilis*, strains of *Pantoea* such as *P. citrea*, strains of *Pseudomonas* such as *P. alcaligenes*, strains of *Streptomyces* such as *S. lividans* or *S. rubiginosus*, strains of *Thermosynechococcus* such as *T. elongatus*, strains of *Sinorhizobium* such as *S. meliloti*, strains of *Helicobacter* such as *H. pylori*, strains of *Agrobacterium* such as *A. tumefaciens*, strains of *Deinococcus* such as *D. radiodurans*, strains of *Listeria* such as *L. monocytogenes*, strains of *Lactobacillus* such as *L. spp*, or strains of *Escherichia* such as *E. coli*.

[0173] In some embodiments, the source organism is a bacterium, such as strains of *Escherichia* (e.g., *E. coli*), or strains of *Bacillus* (e.g., *B. subtilis*).

[0174] As used herein, “the genus *Escherichia*” includes all species within the genus “*Escherichia*,” as known to those of skill in the art, including but not limited to *E. coli*, *E. adecarboxylata*, *E. albertii*, *E. blattae*, *E. fergusonii*, *E. hermannii*, *E. senegalensis*, and *E. vulneris*. The genus “*Escherichia*” is defined as Gram-negative, non-spore forming, facultatively anaerobic, rod-shaped bacteria are classified as members of the Family Enterobacteriaceae, Order Enterobacteriales, Class Gamma Proteobacteria.

[0175] As used herein, “the genus *Bacillus*” includes all species within the genus “*Bacillus*,” as known to those of skill in the art, including but not limited to *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, and *B. thuringiensis*. It is recognized that the genus *Bacillus* continues to undergo taxonomical reorganization. Thus, it is intended that the genus include species that have been reclassified, including but not limited to such organisms as *B. stearothermophilus*, which is now named “*Geobacillus stearothermophilus*.” The production of resistant endospores in the

presence of oxygen is considered the defining feature of the genus *Bacillus*, although this characteristic also applies to the recently named *Alicyclobacillus*, *Amphibacillus*, *Aneurinibacillus*, *Anoxybacillus*, *Brevibacillus*, *Filobacillus*, *Gracilibacillus*, *Halobacillus*, *Paenibacillus*, *Salibacillus*, *Thermobacillus*, *Ureibacillus*, and *Virgibacillus*.

[0176] In some embodiments, the source organism is a gram-positive bacterium. Non-limiting examples include strains of *Streptomyces* (e.g., *S. lividans*, *S. coelicolor*, or *S. griseus*), *Bacillus*, *Listeria* (e.g., *L. monocytogenes*) or *Lactobacillus* (e.g., *L. spp.*). In some embodiments, the source organism is a gram-negative bacterium, such as *E. coli*, *Pseudomonas sp.*, or *H. pylori*.

[0177] In some embodiments, the source organism is a plant, such as a plant from the family *Fabaceae*, such as the *Faboideae* subfamily. In some embodiments, the source organism is kudzu, poplar (such as *Populus alba x tremula* CAC35696), aspen (such as *Populus tremuloides*), *Quercus robur*, *Arabidopsis* (such as *A. thaliana*), or *Zea* (such as *Z. mays*).

[0178] In some embodiments, the source organism is an algae, such as a green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates.

[0179] In some embodiments, the source organism is a cyanobacterium, such as cyanobacteria classified into any of the following groups based on morphology: *Chroococcales*, *Pleurocapsales*, *Oscillatoriales*, *Nostocales*, or *Stigonematales*. In some embodiments, the cyanobacterium is *Thermosynechococcus elongates*.

B. Method for Using Engineered cells for Butadiene Production

[0180] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein that contain one or more pathways for the production of butadiene (shown in FIG. 16 to FIG. 18) with one or more heterologous polynucleotides encoding one or more butadiene pathway enzymes expressed in a sufficient amount to produce butadiene. The butadiene pathway includes an acetyl-CoA:acetyl-CoA acyltransferase, an acetoacetyl-CoA reductase, a 3-hydroxybutyryl-CoA dehydratase, a crotonyl-CoA reductase (aldehyde forming), a crotonaldehyde reductase (alcohol forming), a crotyl alcohol kinase, a 2-butenyl-4-phosphate kinase, a butadiene synthase, a crotonyl-CoA

hydrolase, a crotonyl-CoA synthetase, a crotonyl-CoA transferase, a crotonate reductase, a crotonyl-CoA reductase (alcohol forming), a glutaconyl-CoA decarboxylase, a glutaryl-CoA dehydrogenase, an 3-aminobutyryl-CoA deaminase, a 4-hydroxybutyryl-CoA dehydratase or a crotyl alcohol diphosphokinase. The production of butadiene from bacteria is described in WO 2011/140171 A2, hereby incorporated by reference in its entirety, particularly with respect to the pathways for production of butadiene from acetyl-CoA (FIG. 16), from erythrose-4-phosphate (FIG. 17), and from malonyl-CoA plus acetyl-CoA (FIG. 18).

C. Method of Using Engineered cells for Ethanol Production

[0181] Several host cells as described herein are known to produce ethanol through the acetyl-CoA pathway, which can utilize both carbon monoxide and hydrogen as carbon sources and as energy sources. For example, the production of ethanol from bacteria is described in Kopke et al., 2011, Fermentative production of ethanol from carbon monoxide, Current Opinion in Biotechnology, Vol. 22:320-323, and in Wilkins et al., 2011, Microbial production of ethanol from carbon monoxide, Current Opinion in Biotechnology, Vol. 22:326-330, both of which are hereby incorporated in their entirety, particularly with respect to their discussion of the pathway for the production of ethanol from acetyl-CoA in bacteria.

[0182] In some embodiments, the compositions and methods disclosed herein can be used to transform host cells as described herein that contain the ethanol pathway with one or more heterologous polynucleotides encoding one or more ethanol pathway enzymes expressed in sufficient amount to produce ethanol. In certain bacteria, the pathway for production of ethanol from acetyl-CoA includes the aldehyde dehydrogenase enzyme and the alcohol dehydrogenase enzyme (*see, e.g.*, FIG. 11).

D. Method of Using Engineered cells for Production of Other Industrial Bio-Products

[0183] In some aspects of the invention, any of the methods described herein may be used to produce products other than isoprene, butadiene, and ethanol. Such products may be excreted, secreted, or intracellular products. Any one of the methods described herein may be used to produce isoprene and/or one or more of the other products. The products described herein may be, for example, propanediol (*e.g.*, 1,2-propanediol, 1,3-propanediol), hydrogen, acetate, or microbial fuels. Exemplary microbial fuels are fermentative alcohols (*e.g.*, ethanol

or butanol), non-fermentative alcohols (*e.g.*, isobutanol, methyl butanol, 1-propanol, 1-butanol, methyl pentanol, or 1-hexanol), fatty alcohols, fatty acid esters, isoprenoid alcohols, alkenes, and alkanes. The products described herein may also be a terpenoid, isoprenoid (*e.g.*, farnesene), carotenoid or other C5, C10, C15, C20, C25, C30, C35, or C40 product.

[0184] In some aspects, the terpenoids are selected from the group consisting of hemiterpenoids, monoterpenoids, sesquiterpenoids, diterpenoids, sesterterpenoids, triterpenoids, tetraterpenoids, and higher polyterpenoids. In some aspects, the hemiterpenoid is prenol, isoprenol, or isovaleric acid. In some aspects, the monoterpenoid is geranyl pyrophosphate, eucalyptol, limonene, or pinene. In some aspects, the sesquiterpenoid is farnesyl pyrophosphate, artemisinin, or bisabolol. In some aspects, the diterpenoid is geranylgeranyl pyrophosphate, retinol, retinal, phytol, taxol, forskolin, or aphidicolin. In some aspects, the triterpenoid is squalene or lanosterol. In some aspects, the tetraterpenoid is lycopene or carotene. In some aspects, the carotenoids are selected from the group consisting of xanthophylls and carotenes. In some aspects, the xanthophyll is lutein or zeaxanthin. In some aspects, the carotene is α -carotene, β -carotene, γ -carotene, β -cryptoxanthin or lycopene.

[0185] The products described herein may be derived from Acetyl-CoA produced via syngas fermentation or via fermentation of other carbon sources such as fructose. In some aspects, the cell is grown under conditions suitable for the production of the product(s) other than isoprene.

[0186] The products described herein may be naturally produced by the cell. In some aspects, the cells naturally produce one or more products including excreted, secreted, or intracellular products. In some aspects, the cells naturally produce ethanol, propanediol, hydrogen, or acetate. In some aspects, production of a naturally occurring product is increased relative to wild-type cells. Any method known in the art to increase production of a metabolic cellular product may be used to increase the production of a naturally occurring product. In some aspects, the nucleic acid encoding all or a part of the pathway for production of a product described herein is operably linked to a promoter such as a strong promoter. In some aspects, the nucleic acid encoding all or a part of the pathway for production of a product described herein is operably linked to a constitutive promoter. In some aspects, the cell is engineered to comprise additional copies of an endogenous nucleic acid encoding a polypeptide for the

production of a product described herein. In some aspects, the product described herein is not naturally produced by the cell. In some aspects, the cell comprises one or more heterologous nucleic acids encoding one or more polypeptides for the production of a product described herein.

[0187] Under normal growth conditions, acetogens produce acetate and ethanol. Acetate is produced in a 2-step reaction in which acetyl-CoA is firstly converted to acetyl-phosphate by phosphotransacetylase (*pta*), and then acetyl-phosphate is dephosphorylated by acetate kinase (*ack*) to form acetate. Ethanol is formed by a two-step process in which acetyl-CoA is converted to acetaldehyde and then to ethanol by the multifunctional enzyme alcohol dehydrogenase (*adhE*). The production of acetate and ethanol may not be desirable in isoprene-producing cells, as it fluxes carbon away from isoprene and ultimately results in decreased yield of isoprene. Thus, some or all of the genes coding for phosphotransacetylase (*pta*), acetate kinase (*ack*), and alcohol dehydrogenase (*adhE*) may be disrupted or the expressions thereof are reduced in anaerobic cells for the purpose of redirecting carbon flux away from acetate and/or ethanol and increasing the production of isoprene.

[0188] In some aspects, the cells are deficient in at least one polypeptide involved in production of acetate, ethanol, succinate, and/or glycerol. In some aspects, one or more pathways for production of a metabolite other than isoprene (*e.g.*, lactate, acetate, ethanol (or other alcohol(s)), succinate, or glycerol) are blocked, for example, the production of a metabolite other than isoprene may be reduced by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%. In some aspects, one or more of the pathways for production of lactate, acetate, ethanol, succinate, or glycerol is blocked, for example, the production for lactate, acetate, ethanol, succinate, and/or glycerol is reduced by at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%. In some aspects, the cells are deficient in at least one polypeptide in pathway(s) of producing acetate, ethanol, succinate, and/or glycerol. Polypeptides in pathway(s) of producing acetate, ethanol, succinate, and/or glycerol may have reduced activities or the expressions thereof are reduced. Nucleic acids encoding polypeptides in pathway(s) of producing acetate, ethanol, succinate, and/or glycerol may be disrupted. The polypeptides involved in various pathways (*e.g.*, pathways for producing ethanol and/or acetate) are known to one skilled in the art, including, for example, those described in Misoph *et al.* 1996, *Journal of Bacteriology*, 178(11):3140-45, the contents

of which are expressly incorporated by reference in its entirety with respect to the polypeptides involved in pathways of producing succinate, acetate, lactate, and/or ethanol.

[0189] In some aspects, the cells are deficient in *pta*. In some aspects, the cells are deficient in *ack*. In some aspects, the cells are deficient in *adhE*. In some aspects, the cells are deficient in *pta*, *ack*, and/or *adhE*. In some aspects, the expressions of phosphotransacetylase, acetate kinase, and/or alcohol dehydrogenase are reduced. In some aspects, the activities of phosphotransacetylase, acetate kinase, and/or alcohol dehydrogenase are reduced. In some aspects, the cells are deficient in polypeptide(s) having similar activities as phosphotransacetylase, acetate kinase, and/or alcohol dehydrogenase. The expression of *pta*, *ack*, *adhE*, and/or polypeptide(s) having similar activities as phosphotransacetylase, acetate kinase, and/or alcohol dehydrogenase may be reduced by any of the methods known to one skilled in the art, for example, the expression may be reduced by antisense RNA(s) (*e.g.*, antisense RNA driven by any of the promoters described herein such as any of the inducible promoters). In some aspects, the antisense RNA(s) are operably linked to a suitable promoter such as any of the promoters described herein including inducible promoters.

[0190] In some aspects, isoprene and product(s) other than isoprene are both recovered from the gas phase. In some aspects, isoprene is recovered from the gas phase (*e.g.* from the fermentation of gas), and the other product(s) are recovered from the liquid phase (*e.g.* from the cell broth).

VII. Bioreactors

[0191] A variety of different types of reactors can be used for production of isoprene or other industrial bio-products. In some embodiments, a carbohydrate is used as energy and/or carbon source. In some embodiments, a carbohydrate and hydrogen are used as energy and/or carbon source. In some embodiments, syngas is used as energy and/or carbon source. There are a large number of different types of fermentation processes that are used commercially. Bioreactors for use in the present invention should be amenable to anaerobic conditions. The bioreactor can be designed to optimize the retention time of the cells, the residence time of liquid, and the sparging rate of syngas.

[0192] In various aspects, the cells are grown using any known mode of fermentation, such as batch, fed-batch, continuous, or continuous with recycle processes. In some aspects, a batch method of fermentation is used. Classical batch fermentation is a closed system where the composition of the media is set at the beginning of the fermentation and is not subject to artificial alterations during the fermentation. Thus, at the beginning of the fermentation the cell medium is inoculated with the desired host cells and fermentation is permitted to occur adding nothing to the system. Typically, however, “batch” fermentation is batch with respect to the addition of carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems, the metabolite and biomass compositions of the system change constantly until the time the fermentation is stopped. Within batch cultures, cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. In some aspects, cells in log phase are responsible for the bulk of the isoprene production. In some aspects, cells in stationary phase produce isoprene.

[0193] In some aspects, a variation on the standard batch system is used, such as the Fed-Batch system. Fed-Batch fermentation processes comprise a typical batch system with the exception that the carbon source (*e.g.* syngas, glucose, fructose) is added in increments as the fermentation progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of carbon source in the cell medium. Fed-batch fermentations may be performed with the carbon source (*e.g.*, syngas, glucose, fructose) in a limited or excess amount. Measurement of the actual carbon source concentration in Fed-Batch systems is difficult and is therefore estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen, and the partial pressure of waste gases such as CO₂. Batch and Fed-Batch fermentations are common and well known in the art and examples may be found in Brock, *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc.

[0194] In some aspects, continuous fermentation methods are used. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned medium is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth.

[0195] Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or isoprene production. For example, one method maintains a limiting nutrient such as the carbon source or nitrogen level at a fixed rate and allows all other parameters to moderate. In other systems, a number of factors affecting growth can be altered continuously while the cell concentration (*e.g.*, the concentration measured by media turbidity) is kept constant. Continuous systems strive to maintain steady state growth conditions. Thus, the cell loss due to media being drawn off is balanced against the cell growth rate in the fermentation. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., which is hereby incorporated by reference in its entirety, particularly with respect to cell culture and fermentation conditions.

[0196] A variation of the continuous fermentation method is the continuous with recycle method. This system is similar to the continuous bioreactor, with the difference being that cells removed with the liquid content are returned to the bioreactor by means of a cell mass separation device. Cross-filtration units, centrifuges, settling tanks, wood chips, hydrogels, and/or hollow fibers are used for cell mass separation or retention. This process is typically used to increase the productivity of the continuous bioreactor system, and may be particularly useful for anaerobes, which may grow more slowly and in lower concentrations than aerobes.

[0197] In one aspect, a membrane bioreactor can be used for the growth and/or fermentation of the anaerobic cells described herein, in particular, if the cells are expected to grow slowly. A membrane filter, such as a crossflow filter or a tangential flow filter, can be operated jointly with a liquid fermentation bioreactor that produces isoprene gas. Such a membrane bioreactor can enhance fermentative production of isoprene gas by combining fermentation with recycling of select broth components that would otherwise be discarded. The MBR filters fermentation broth and returns the non-permeating component (filter “retentate”) to the reactor, effectively increasing reactor concentration of cells, cell debris, and other broth solids, while maintaining specific productivity of the cells. This substantially improves titer, total production, and volumetric productivity of isoprene, leading to lower capital and operating costs.

[0198] The liquid filtrate (or permeate) is not returned to the reactor and thus provides a beneficial reduction in reactor volume, similar to collecting a broth draw-off. However, unlike a broth draw-off, the collected permeate is a clarified liquid that can be easily sterilized by filtration after storage in an ordinary vessel. Thus, the permeate can be readily reused as a nutrient and/or water recycle source. A permeate, which contains soluble spent medium, may be added to the same or another fermentation to enhance isoprene production.

VIII. Recovery Methods

[0199] Any of the methods described herein further include recovering the industrial bio-product (*e.g.*, isoprene, butadiene, ethanol, etc.). For example, the isoprene produced using the compositions and methods of the invention can be recovered using standard techniques, such as gas stripping, membrane enhanced separation, fractionation, adsorption/desorption, evaporation, thermal or vacuum desorption of isoprene from a solid phase, or extraction of isoprene immobilized or absorbed to a solid phase with a solvent (see, for example, U.S. Patent Nos. 4,703,007 and 4,570,029). In one aspect, the isoprene is recovered by absorption stripping (*see, e.g.*, International Patent Application No. PCT/US2010/060552 (WO 2011/075534)). In particular aspects, extractive distillation with an alcohol (such as ethanol, methanol, propanol, or a combination thereof) is used to recover the isoprene. In some aspects, the recovery of isoprene involves the isolation of isoprene in a liquid form (such as a neat solution of isoprene or a solution of isoprene in a solvent). Gas stripping involves the removal of isoprene vapor from the fermentation off-gas stream in a continuous manner. Such removal can be achieved in several different ways including, but not limited to, adsorption to a solid phase, partition into a liquid phase, or direct condensation (such as condensation due to exposure to a condensation coil or due to an increase in pressure). In some aspects, membrane enrichment of a dilute isoprene vapor stream above the dew point of the vapor resulting in the condensation of liquid isoprene. In some aspects, the isoprene is compressed and condensed.

[0200] The recovery of isoprene may involve one step or multiple steps. In some aspects, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed simultaneously. For example, isoprene can be directly condensed from the off-gas stream to form a liquid. In some aspects, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed

sequentially. For example, isoprene may be adsorbed to a solid phase and then extracted from the solid phase with a solvent.

[0201] In some aspects, any of the methods described herein further include a step of recovering the compounds produced. In some aspects, any of the methods described herein further include a step of recovering the isoprene. In some aspects, the isoprene is recovered by absorption stripping (See, *e.g.*, U.S. Publ. No. 2011/0178261).

[0202] Isoprene compositions recovered from fermentations contain impurities. The identities and levels of impurities in an isoprene composition can be analyzed by standard methods, such as GC/MS, GC/FID, and ¹H NMR. An impurity can be of microbial origin, or it can be a contaminant in the syngas feed or other fermentation raw materials.

[0203] In some aspects, the isoprene composition recovered from fermentation comprises one or more of the following impurities: hydrogen sulfide, carbonyl sulfide, carbon disulfide, ethanol, acetone, methanol, acetaldehyde, methacrolein, methyl vinyl ketone, 2-methyl-2-vinylloxirane, *cis*- and *trans*-3-methyl-1,3-pentadiene, a C5 prenyl alcohol (such as 3-methyl-3-buten-1-ol or 3-methyl-2-buten-1-ol), 2-heptanone, 6-methyl-5-hepten-2-one, 2,4,5-trimethylpyridine, 2,3,5-trimethylpyrazine, citronellal, methanethiol, ethanethiol, methyl acetate, 1-propanol, diacetyl, 2-butanone, 2-methyl-3-buten-2-ol, ethyl acetate, 2-methyl-1-propanol, 3-methyl-1-butanal, 3-methyl-2-butanone, 1-butanol, 2-pentanone, 3-methyl-1-butanol, ethyl isobutyrate, 3-methyl-2-butenal, butyl acetate, 3-methylbutyl acetate, 3-methyl-3-buten-1-yl acetate, 3-methyl-2-buten-1-yl acetate, (E)-3,7-dimethyl-1,3,6-octatriene, (Z)-3,7-dimethyl-1,3,6-octatriene, (E,E)-3,7,11-trimethyl-1,3,6,10-dodecatetraene, (E)-7,11-dimethyl-3-methylene-1,6,10-dodecatriene, 3-hexen-1-ol, 3-hexen-1-yl acetate, limonene, geraniol (trans-3,7-dimethyl-2,6-octadien-1-ol), citronellol (3,7-dimethyl-6-octen-1-ol), (E)-3-methyl-1,3-pentadiene, (Z)-3-methyl-1,3-pentadiene, thiol(s), mono and disulfide(s), or gas(es) such as CS₂ and COS. In one embodiment, the isoprene composition recovered from syngas under anaerobic fermentation conditions may comprise one or more of the components described in Rimbault A *et al.* 1986, *J. of Chromatography*, 375:11-25, the contents of which are expressly incorporated herein by reference in its entirety.

[0204] In some aspects, any of the methods described herein further include purifying the isoprene. For example, the isoprene produced using the compositions and methods of the

invention can be purified using standard techniques. Purification refers to a process through which isoprene is separated from one or more components that are present when the isoprene is produced. In some aspects, the isoprene is obtained as a substantially pure liquid. Examples of purification methods include (i) distillation from a solution in a liquid extractant and (ii) chromatography. As used herein, “purified isoprene” means isoprene that has been separated from one or more components that are present when the isoprene is produced. In some aspects, the isoprene is at least about 20%, by weight, free from other components that are present when the isoprene is produced. In various aspects, the isoprene is at least or about 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 95%, or 99%, by weight, pure. Purity can be assayed by any appropriate method, *e.g.*, by column chromatography, HPLC analysis, or GC-MS analysis.

[0205] In some aspects, at least a portion of the gas phase remaining after one or more recovery steps for the removal of isoprene is recycled by introducing the gas phase into a cell culture system (such as a fermentor) for the production of isoprene.

[0206] In some embodiments, recovery of industrial enzymes can use any method known to one of skill in the art and/or any of the exemplary protocols that are disclosed in U.S. Appl. Pub. Nos. 2009/0311764, 2009/0275080, 2009/0252828, 2009/0226569, 2007/0259397 and U.S. Patent Nos. 7,629,451; 7,604,974; 7,541,026; and 7,527,959 and for neutraceuticals (see, *e.g.*, U.S. Patent No. 7,622,290), and for antimicrobials (see, *e.g.*, U.S. Appl. Pub. No. 2009/0275103).

[0207] The following examples have been provided for illustrative purposes only and are not intended to limit the invention.

EXAMPLES

Example 1: Construction of a vitamin B12 riboswitched reporter gene construct and demonstration of derepression of gene expression by dilution

[0208] This example describes the construction of a vitamin B12 riboswitched reporter gene construct and demonstration of derepression of gene expression by dilution.

A. Bioluminescence reporter gene construct

[0209] Plasmid pUC57-btuB2 was purchased from GenScript and contains the following sequence, which comprises four *E. Coli* gapA promoters, the *E. Coli* btuB leader sequence, and first 99 bases of the btuB coding sequence:

gaattctacgaaacgtcagcggtcaacacccgccagcaatggactgtattgcgctcttcgtgcgtcgcgtctgttaaaaactggcgctaac
aatacaggctaaagtgaaccgccaggctagacttttagttccacaacactaaacctataagttggggaaatacaatgttcagcaagaag
ttaccattaccgctccgaacggcttgactgattctaacaaaacattaacaccaactggcaaaattttgtcctaaacttgatctcgacgaaatg
gctgcacctaaatcgtgatgaaaatcacatttttatcgttaattgccctttaaaattcggggcgccgaccccatgtggtctcaagcccaaagg
aagagtgaggcgagtcagtcgcgtaatgcttaggcacaggattgattgtcgcaatgattgacacgattccgcttgacgctgcgtaagggt
tttgaattttacaggcagccggctctgtgagttaatagggaatccagtcgcaatctggagctgacgcgacggtgaaggaaagggtgcga
tgattgcgttatgcggacactgccattcgggtgggaagtcacatctcttagtatcttagataccctccaagcccgaagacctgccggcca
acgtcgcacatcgtgtctcatcgcgtaattgatgaaacctgcggcaccctcttctattgttgatgctttacaatgattaaaaaagcttcg
ctgctgacggcggtgtccgtcacggcattttccgcttgggcacaggataaccagcccggatactctcgtcgttactgctaacactatcaaagt
aggtatcaacgggtttggccgtatcggtcgcattgtttccgtgctgctcagaaacgttctgacatcgagatcgttgcaatcaacgacctgtt
agacgctgattacatggcatacatgctgaaatatgactccactcacggccgtttcgacgggaattc (SEQ ID NO:1)

[0210] Plasmid pDEW201 carries a luxCDABE bioluminescent reporter gene without a promoter and has been fully described (Van Dyk, T. K. and Rosson, R. A. 1998. Photorhabdus luminescens luxCDABE promoter probe vectors. IN: R. A. LaRossa (Ed) Methods in Molecular Biology: Bioluminescence Methods and Protocols. Humana Press Inc. 102: 85-95).

[0211] Plasmid pDMWP156 that has the four gapA promoters, the B12 dependent riboswitch and a translational fusion of the first 33 amino acids of the btuB protein to luxC was constructed by Gibson assembly following the manufacturer's directions (New England Biolabs), as schematically represented in Figure 2.

[0212] The following primers were designed to amplify the desired fragments before carrying out the Gibson Assembly reactions, pDEW201 as template, ODMWP217/219 as primers + pUC57-btuB2 as template, ODMWP220/222 as primers.

ODMWP217: ATTTGCCCAACAGTTGCGCAGCCTG (SEQ ID NO:2)

ODMWP219: ACTAAAAAATTTTCATTATTATTAACG (SEQ ID NO:3)

ODMWP220: CAGGCTGCGCAACTGTTGGGGCAAATTGACTGATTCTAACAAA (SEQ ID NO:4)

ODMWP222 GTTAATAATGAATGAAATTTTTTTTAGTGTTAGCAGTAACGACGA (SEQ ID NO:5)

[0213] PCR reactions, using Thermo Scientific Phusion Flash PCR Master Mix were carried out, following the manufacturer's protocol, to amplify the above fragments. Using New England Biolab's Gibson Assembly Master Mix, the reactions were set up following the manufacturer's protocol. The Gibson Assembly reactions were transformed into *E. Coli* K12 strain FM5 with selection for ampicillin resistance.

[0214] Colonies from the transformation plates were picked into 100 μ l MOPS glucose medium (Teknova) in white 96-well microplates (Dynex). Bioluminescence was measured with a Luminoskan Ascent luminometer (Labsystems). A positive clone was identified by light production. The plasmid in this clone was named pDMWP156 and the strain containing it was named DP2016. Sequence analysis using primers ODMWP215 and ODMWP224 confirmed that plasmid pDMWP156 had the correct insert.

ODMWP215: GCAATGATTGACACGATTCCGCTTGACG (SEQ ID NO:6)

ODMWP224: CTGGCCGTTAATAATGAATGAAA (SEQ ID NO:7)

B. Demonstration of derepression by dilution

[0215] *E. coli* K12 strain DP2016, which carries pDMWP156 in FM5, was grown overnight in LB Amp100 at 37 °C. This culture was diluted 1:1000 into minimal MOPS glucose 2 g/L medium (Teknova) with a final concentration of 0.01 mg/L of vitamin B12 or the same medium lacking vitamin B12 in a final volume of 100 μ L in 4 wells of a white microplate (Dynex). After 5 hours growth, two of the wells with vitamin B12 were diluted 1:10 to reduce the vitamin B12 concentration to 0.001 mg/L. Two of the wells without vitamin B12 were also diluted 1:10. Bioluminescence was measured with a Luminoskan Ascent luminometer (Labsystems) every 7.5 minutes for 18 hours. During this time the bioluminescent signal increased as the cells grew, then decreased as the cultures went into stationary phase. The peak bioluminescence (RLU) was recorded and the average of the duplicate cultures was calculated. The plot of RLU as a function of time is provided in Figure 3A (undiluted) and Figure 3B (diluted). The repression ratio was calculated by dividing the peak bioluminescence of the wells that did not contain vitamin B12 by the vitamin B12 containing wells. The results are shown in the table 1-1 below:

Table 1-1: Results of derepression by dilution experiments

DP2016 Culture	Average Peak RLU	Repression Ratio
Undiluted 0.01 mg/L vitamin B12	32.3 +/- 3.4	17.3
Undiluted no vitamin B12	559 +/- 40	
Diluted to 0.001 mg/L vitamin B12	349 +/- 26	1.4
Diluted no vitamin B12	498 +/- 28	

[0216] Addition of 0.01 mg/L vitamin B12 repressed expression of the luxCDABE reporter by seventeen-fold and a simple 10 fold dilution allowed recovery of 70% of the expression level of cultures lacking vitamin B12, thereby demonstrating derepression of gene expression by dilution.

Example 2: Altering host cell uptake of Vitamin B12

[0217] This example shows that vitamin B12 repression of a reporter gene construct does not work in *E. coli* strain BL21 but can be made to work by repair of the host cells' vitamin B12 uptake machinery.

A. Lack of vitamin B12 repression of a reporter gene construct in E. coli BL2

[0218] The riboswitched reporter gene plasmid, pDMWP156 described in Example 1, was transformed into *E. coli* BL21 to make a strain named DP2022. Strain DP2022 and strain DP2016, which carries the same plasmid in *E. coli* K12 strain FM5 were grown overnight in

LB Amp100 at 37 °C. The overnight cultures were diluted 1:100 in duplicate into 100 µL of MOPS glucose 2 g/L with or without 1 mg/L vitamin B12 in wells of a white microplate (Dynex). Bioluminescence was measured with a Luminoskan Ascent luminometer (Labsystems) every 7.5 minutes for 2.4 hours. The table 2-1 below shows the average of the duplicate light production (RLU) at 2.4 hours.

Table 2-1: Bioluminescence in *E. coli* strains

Culture	RLU at 2.4 hours with 1 mg/L vitamin B12	RLU at 2.4 hours without vitamin B12
DP2016 (<i>E. Coli</i> K12)	1.59 +/- 0.04	7.62 +/- 0.01
DP2022 (<i>E. Coli</i> BL21)	4.04 +/- 0.35	2.48 +/-0.21

[0219] Thus, for strain DP2016, the reporter gene expression was repressed in the presence of vitamin B12, but for DP2022 the reporter gene expression was not repressed by vitamin B12.

B. Methods for lambda Red recombination and selection to generate a B strain capable of B12 uptake for use in system

[0220] The likely reason that *E. Coli* BL21 does not exhibit vitamin B12 repression of the riboswitched reporter construct is because of a deficiency in B12 uptake that has been reported in the literature (Studier, F. W. et al. Understanding the Differences between Genome Sequences of *Escherichia coli* B Strains REL606 and BL21 (DE3) and Comparison of the *E. Coli* B and K-12 Genomes. *J. Mol. Biol.* (2009) 394, 653–680).

[0221] To repair vitamin B12 transport, the *E. Coli* K12 *btuB* gene was used to replace the defective *btuB* gene in an *E. Coli* BL21 derived strain, HMB (BL21 *tpgl* PL2mKKDyl). Plasmid pKD46 (Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97, 6640-6645) encodes arabinose inducible lambda Red mediated recombination functions. This plasmid was transformed into strain HMB and the resultant strain was induced with

arabinose and transformed with a PCR product using *E. Coli* K12 strain MG1655 chromosomal DNA as a template and primers ODMWP231/232.

ODMWP231: ATGATTAAAAAAGCTTCGCTGC (SEQ ID NO:8)

ODMWP232: GATGATATTCACCACCCCGC (SEQ ID NO:9)

[0222] Selection was for growth on minimal medium plates with vitamin B12 and ethanolamine as the sole nitrogen source. The media recipe was 5 g/L glycerol, 7 g/L K₂HPO₄, 3 g/L KH₂PO₄, 1 g/L Na₂SO₄, 0.1 g/L MgSO₄•7H₂O, 1 g/L ethanolamine hydrochloride, 0.04 mg/L vitamin B12, and 2% agar. After incubation at 37 °C for 2 days, single colonies were streaked to LB Lennox plates. A purified single colony was then tested for growth on the glycerol/ethanolamine plates with vitamin B12. A strain that grew well was retained and named DP2035. PCR was done using DP2035 chromosomal DNA as a template and primers ODMWP231/232. The PCR product was submitted for DNA sequencing and the sequence data showed that the TAG stop codon in BL21 *btuB* was changed to a Gln codon CAG, which is the same as in MG1655.

C. Demonstration of repression of reporter gene expression by vitamin B12 in *btuB*+ host strain and derepression by dilution

[0223] Strain DP2035 was transformed with the vitamin riboswitched reporter plasmid, pDMWP156, with selection for carbenicillin resistance. The resultant strain was named DP2037. A single colony of this strain was suspended in 2.5 mL of minimal MOPS (Teknova) medium with 2 g/L glucose and 25 µg/mL of ampicillin and incubated at 37 °C overnight. This culture was diluted 1:1000 into minimal MOPS glucose 2 g/L medium (Teknova) with a final concentration of 0.01 mg/L of vitamin B12 or the same medium lacking vitamin B12 in a final volume of 100 µL in 4 wells of a white microplate (Dynex). After 5 hours growth, two of the wells with vitamin B12 were diluted 1:10 to reduce the vitamin B12 concentration to 0.001 mg/L. Two of the wells without vitamin B12 were also diluted 1:10. Bioluminescence was measured with a Luminoskan Ascent luminometer (Labsystems) every 7.5 minutes for 18 hours. During this time the bioluminescent signal increased as the cells grew, then decreased as the cultures went into stationary phase. The peak bioluminescence (RLU) was recorded and the average of the duplicate cultures was calculated. The repression ratio was calculated by

dividing the peak bioluminescence of the wells that did not contain vitamin B12 by the vitamin B12 containing wells. The results are shown in the table 2-2 below.

Table 2-2: Repression ratio calculated via Bioluminescence

DP2035 Culture	Average Peak RLU	Repression Ratio
Undiluted 0.01 mg/L vitamin B12	11.1 +/- 2.1	5.36
Undiluted no vitamin B12	59.5 +/- 6.4	
Diluted to 0.001 mg/L vitamin B12	41.4 +/- 1.1	1.04
Diluted no vitamin B12	42.9 +/- 0.3	

[0224] Addition of 0.01 mg/L vitamin B12 repressed expression of the luxCDABE reporter by five-fold and a simple 10 fold dilution allowed recovery of 96% of the expression level of cultures lacking vitamin B12, thereby demonstrating derepression of gene expression by dilution in this BL21-derived strain with repaired vitamin B12 transport.

Example 3: Construction a vitamin B12 riboswitched alternative sigma factor plasmid

[0225] The purpose of this example is to describe construction a vitamin B12 riboswitched alternative sigma factor plasmid and to show that the sigma factor FliA functions with N-terminal fusion of 33 BtuB amino acids.

[0226] FliA is the *E. coli* Sigma28 factor, which is an alternative sigma factor that drives motility related gene expression. In *E. Coli* K12, FliA can accumulate, under endogenous expression, to roughly half the level of Sigma70 and is capable of inducing very high levels of gene expression. FliA is not encoded in the *E. Coli* BL21 genome. Likewise, 20 FliA-dependent motility genes are also not encoded in the BL21 genome. Use of FliA-dependent promoters for expression of genes of interest in BL21-derived strains is a potential to

reprogram the cellular transcriptome upon induction of FliA expression. One such mechanism to regulate FliA expression is by control with a vitamin B12 dependent riboswitch. It has been reported in the literature that the vitamin B12 riboswitch from *E. Coli* btuB yields maximum repression with a N-terminal fusion of 33 amino acids of BtuB fused to the protein of interest (Ravnum, S. and Andersson, D.I. (1997), Vitamin B12 repression of the btuB gene in *Salmonella typhimurium* is mediated via a translational control which requires leader and coding sequences. *Mol Microbiol.* 23:35-42). Thus, it was important to show that FliA retains function with this N-terminal translation fusion.

A. Description of pDMWP165 construction

[0227] Plasmid pDMWP165 (see Figure 4) is a chloramphenicol-resistant single copy plasmid with a glucose isomerase promoter P1.20 driving expression of a btuB riboswitched fliA gene with an N-terminal translational fusion of the first 33 amino acids of BtuB. It was constructed as follows:

[0228] A promoter-multiple cloning site-terminator region was synthesized by IDT and cloned into their pIDTsmart vector, resulting in the construction of pDMWP3. The promoter in pDMWP3 is the glucose isomerase P1.20 promoter. The following is the sequence for the synthesized region:

aggaattccctagggcgatctgtgctgtttgccacggatgcagcaccagcgcgagattatgggctcgcacgctcgactgtcggacggggg
cactggaacgagaagtcaggcgagccgtcacgcccttgacgatgccacatcctgagcaaataattcaaccactaaacaaatcaaccgcgt
ttcccgaggtaaccaagcttggccggatccgcatgcgcggccgcgtcgactctagtttaaacccccgggtgatcgatagctcttaattaag
ttgtttgccaatgtaatgccgctgcaccaggcatcaaataaacgaaaggctcagtcgaaagactgggcctttcgtttatctgtgtttgtcg
gtgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttatacagctgtcgggtaccgccag (SEQ ID
NO:10)

[0229] pDMWP4, a backbone for subsequent constructs, is pBR322 with modified restriction sites. All mutagenesis reactions carried out to make pDMWP4 were done using Stratagene's Quikchange kit. With the exception of the addition of 3 µl of Quik Solution to each of the mutagenesis reactions, the manufacturer's protocol was followed.

[0230] pBR322 was mutagenized with ODMWP12/13 to add a sca1 site. The resulting plasmid was pBR322-sca1. pBR322-sca1 was mutagenized again with primers ODMWP14/15

to add a kpn1 site, resulting in pBR322-sca/kpn. This resulting plasmid was further mutagenized with primers ODMWP16/17 to delete an existing sca1 site. The new plasmid was named pBR322-sca/kpn (-sca.) One final mutagenesis reaction with pBR322-sca/kpn (-sca) was carried out with ODMWP18/19 to add a final kpn1 site. The plasmid was named pDMWP4 or pBR322 sca/kpn (-sca+kpn).

ODMWP12: TAA TGC GGT AGT TTA TCA CAG TAC TAT TGC TAA CGC AGT CAG GCA CC (SEQ ID NO:11)

ODMWP13: GGT GCC TGA CTG CGT TAG CAA TAG TAC TGT GAT AAA CTA CCG CAT TA (SEQ ID NO:12)

ODMWP14: TAT CAC AGT ACT ATT GCT AAC GGT ACC AGG CAC CGT GTA TGA AAT CTA (SEQ ID NO:13)

ODMWP15: TAG ATT TCA TAC ACG GTG CCT GGT ACC GTT AGC AAT AGT ACT GTG ATA (SEQ ID NO:14)

ODMWP16: CAGAATGACTTGGTTGAGAACTCACCAGTCACAGAAA (SEQ ID NO:15)

ODMWP17: TTTCTGTGACTGGTGAGTTCTCAACCAAGTCATTCTG (SEQ ID NO:16)

ODMWP18: TCGACCTGAATGGTACCCGGCGGCACCTC (SEQ ID NO:17)

ODMWP19: GAGGTGCCGCCGGGTACCATTTCAGGTCTGA (SEQ ID NO:18)

[0231] To construct pDMWP12 (pBR322*-Pgi1.20), the 438 bp EcoR1/Kpn1 fragment from pDMWP3 (pIDTSmart + P1.20-mcs-spacer-double terminator) was cloned into the 3076 bp EcoR1/Kpn1 digested pDMWP4.

[0232] pDMWP162 (pDMWP12 + BtuB2/fliA fusion) was constructed using a three fragment Gibson Assembly reaction, following the manufacturer's protocol. The vector fragment was amplified from pDMWP12 with primers ODMWP233/234. The second fragment was amplified from pUC57-btuB2 using primers ODMWP235/236. The third fragment was amplified from MG1655 chromosomal DNA using ODMWP237/238.

ODMWP233: TTAATTAAGTTGTTTGCCAATGTAATG (SEQ ID NO:19)

ODMWP234: GGGAAACGCGGTTGATTTGTTTAGTG (SEQ ID NO:20)

ODMWP235:CAAATCAACCGCGTTTCCCTGACTGATTCTAACAAAACATTAAC (SEQ ID NO:21)

ODMWP236: CCTTCAGCGGTATAGAGTGAATTGTTAGCAGTAACGACG (SEQ ID NO:22)

ODMWP237: AATTCACCTCTATACCGCTGAAGGTGTAATG (SEQ ID NO:23)

ODMWP238: CATTACATTGGCAAACAACCTTAATTAATTATAACTTACCCAGTT (SEQ ID NO:24)

[0233] Finally, pDMWP165 was constructed by NEB's Gibson Assembly kit following the manufacturer's protocol using two fragments; pCC2Fos (purchased from Epicentre) amplified with ODMWP265/266 and pDMWP162 amplified with ODMWP267/268

ODMWP265: CAGGGGCGGTCGACCAATTCTCATGTTTG (SEQ ID NO:25)

ODMWP266: CAATTCGCCCTATAGTGAGTCGTATTAC (SEQ ID NO:26)

ODMWP267: CTCACTATAGGGCGAATTGAATTCCCTAGGCGATCTGTG (SEQ ID NO:27)

ODMWP268: GAATTGGTCGACGGCCCCTGTATAAACGCAGAAAGGCCAC (SEQ ID NO:28)

B. Motility assay

[0234] The function of the riboswitched FliA with N-terminal fusion of 33 amino acids of BtuB was demonstrated by restoration of motility to an *E. Coli* K12 fliA mutant strain. *E. Coli* K12 lacking fliA is not motile because it cannot express flagellar genes. All other genes required for motility are encoded in the *E. Coli* K12 genome. The Keio collection (Baba, T. et al. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol.* Epub 2006 Feb 21) of *E. Coli* K12 deletion strains includes a deletion of fliA. This fliA deletion strain was retrieved from the Keio collection and was renamed TV3004. This strain was transformed with pDMWP165, which as described above, is a single copy plasmids expressing riboswitched fusion FliA protein, and named DP2076. Motility of BW25113, otherwise isogenic to TV3004 except fliA+, TV3004, and DP2076 was tested on motility agar plates (LB with 0.3 % agar). These strains were suspended in LB medium and the OD600 was adjusted to 0.1. 5 µL were spotted in the center of 25.0 mL LB motility agar plates +/- 0.1 mg/L vitamin B12. The plates were incubated at 37 °C for 18 hours at which time the diameter of the swarm was measured. The results are shown in the table 3-1 below:

Table 3-1: Results of motility assay

Strain	Motility without vitamin B12	Motility with 0.1 mg/L vitamin B12
BW25113 (fliA+)	38 mm swarm	27 mm swarm
TV3004 (Δ fliA)	No swarm	No swarm
DP2076 (Δ fliA + riboswitched FliA with N-terminal fusion)	39 mm swarm	6 mm

[0235] As expected, the fliA+ strain was motile in the presence and absence of vitamin B12 and the fliA deletion strain was not motile in either condition. In the absence of vitamin B12, motility was restored to the fliA deletion strain by FliA with the N-terminal fusion, thus demonstrating function.

Example 4: Construction of a plasmid for production of mevalonic acid controlled by a FliA-responsive promoter

[0236] This example describes the construction of a plasmid with mevalonic acid production genes controlled by a promoter responsive to FliA and demonstrates mevalonic acid production controlled by the presence of vitamin B12 addition in the culture media.

A. Description of pDMWP174 construction

[0237] Plasmid pDMWP174 (see Figure 5) is a spectinomycin-resistant medium copy plasmid with the *E. coli* fliC promoter driving expression of *Enterococcus gallinarum* mvaE and mvaS genes for mevalonic acid production. It was constructed as follows:

[0238] pDMWP174 was constructed by Gibson Assembly following the manufacturer's protocol.

[0239] The *E. coli* fliC promoter was amplified using *E. Coli* K12 MG1655 chromosomal DNA as a template and primers ODMWP305 and ODMWP306.

ODMWP305:

CATTAATGCAGATTCTGAAATGAGCAAAAATGGCTGTTTTTGAAAAAATTC (SEQ ID NO:29)

ODMWP306: GGCGCTGTTTCCTGTGTGACCACCCGTCGGCTCAATCGCCGTC (SEQ ID NO:30)

[0240] A vector containing the mvaE and mvaS genes, pMCM1225 (pCL-Ptrc-Upper_GcMM_163 (Enterococcus gallinarum EG2)), was amplified with primers ODMWP279 and ODMWP281.

ODMWP279 (forward primer to amplify vector, pMCM1225):

TCACACAGGAAACAGCGCCGCTGAG (SEQ ID NO:31)

ODMWP281 (reverse primer to amplify vector, pMCM1225):

GCTCATTTTCAGAATCTGCATTAATG (SEQ ID NO:32)

B. Demonstration of MVA production controlled by vitamin B12

[0241] A previously constructed BL21-derived strain appropriate for expression of MVA production genes was made btuB+ by P1 mediated generalized transduction. Strain MD12-778 (BL21 GI1.2gltA, yhfSFRTpyddVIspA yhfS, thiFRTtruncIspA) was the recipient strain and the donor strain was DP2035, described in Example 2, with selection for growth at 37 °C on minimal Glycerol Ethanolamine B12 plates, described in Example 2. A transductant colony was single colony purified on LB Lennox agar plates. A colony was positive for growth minimal Glycerol Ethanolamine B12 plate was retained and named TV3011.

[0242] E.coli strain DP2088 was made by transformation of strain TV3011 with plasmid pDMWP165, described in Example 3, with selection for chloramphenicol resistance followed by transformation of a chloramphenicol-resistant transformant with plasmid pDMWP174 with selection for chloramphenicol and spectinomycin resistance. Thus, strain DP2088 has expression of fliA controlled by a vitamin B12 dependent riboswitch and expression of mvaE and mvaS controlled by a fliA dependent promoter.

[0243] Overnight cultures of DP2088 were grown in LA medium with Chloramphenicol 5 µg/mL, Spectinomycin 50 µg/mL and 0.1 mg/L vitamin B12 at 34 °C. The following day, the culture was diluted 1:1000 into 25 mL TM3 medium with 8 mM MgSO₄, 0.02% yeast extract,

10 g/L glucose, Chloramphenicol 5 µg/mL, and Spectinomycin 50 µg/mL with or without 0.1 mg/L vitamin B12 in 125 mL plastic vented top flasks. The flasks were incubated for 28 hours at 34 °C. Samples were prepared for HPLC analyses by adding 54 µL of 10% w/v sulfuric acid to a 300 µL aliquot of flask broth. The acidified tubes were held at 4 °C for 5 minute, then spun in a centrifuge tube filter (Costar, Spin-X, 0.22 µm nylon) at 14,000 rpm for 5 minutes. The filtrate was put into HPLC vials. HPLC analysis was done on a Waters e2695 HPLC, using refractive index (RI) detection. Chromatographic separation was achieved using a Shodex SH-1011 column and 0.01 N (0.005 M) H₂SO₄ as the mobile phase with a flow rate of 0.5 mL/min and a column temperature of 50 °C. Eluted compounds were quantified by refractive index detection with reference to a standard curve prepared from commercially purchased pure compounds dissolved to known concentrations.

[0244] Results of MVA production in the presence and absence of vitamin B12 are shown in the table 4-1 below.

Table 4-1: MVA production

Sample Name	MVA g/L
DP2088+0.1 mg/L vitamin B12	0.223
DP2088 without vitamin B12	0.446

[0245] Twice as much MVA was produced in the absence of vitamin B12 as in the presence of vitamin B12, thus demonstrating vitamin B12 control of MVA production.

Example 5: Mevalonic acid production controlled by vitamin B12 addition using medium copy number plasmids with vitamin B12 riboswitched alternative sigma factor

[0246] This Example describes the construction of a plasmid with mevalonic acid production genes controlled by a promoter responsive to FliA, construction of medium copy number plasmids with vitamin B12 riboswitched alternative sigma factor and demonstration of mevalonic acid production controlled by vitamin B12 addition.

A. Description of pDMWP175 construction

[0247] Plasmid pDMWP175 (see Figure 8) is a spectinomycin-resistant medium copy plasmid with the *E. coli* fliC promoter driving expression of *Enterococcus gallinarum* mvaE and mvaS genes for mevalonic acid production. This plasmid differs from DMWP174 (in Example 4) by a lacO site not present in pDMWP174. Plasmid pDMWP175 was constructed as follows.

[0248] pDMWP175 was constructed by Gibson Assembly following the manufacturer's protocol.

[0249] The *E. coli* fliC promoter was amplified using *E. Coli* K12 MG1655 chromosomal DNA as a template and primers ODMWP305 (sequence in Example 4) and ODMWP307. ODMWP307 CCGCTCACAATTCCACACCCACCCGTCGGCTCAATCG (SEQ ID NO:33)

[0250] A vector containing the mvaE and mvaS genes, pMCM1225 (pCL-Ptrc-Upper_GcMM_163 (*Enterococcus gallinarum* EG2)), was amplified with primers ODMWP259/260.

ODMWP259 GTGTGGAATTGTGAGCGGATAACAATTTTC (SEQ ID NO:34)

ODMWP260 GCTCATTTCAGAATCTGCATTAATGAATC (SEQ ID NO:35)

B. Description of pDMWP180 and pDMWP181 construction

[0251] Plasmid pDMWP180 (see Figure 6) is a chloramphenicol-resistant medium copy number plasmid with a glucose isomerase promoter P1.20 driving expression of a btuB riboswitched fliA gene with an N-terminal translational fusion of the first 33 amino acids of BtuB. Plasmid pDMWP181 (see Figure 7) is like pDMWP180 with the exception that it has with a glucose isomerase promoter P1.5 which differs in one nucleotide in the -35 region from the P1.20 version and is known to be a stronger promoter. These plasmids were constructed as follows:

[0252] The construction of plasmid pDMWP162 is described in Example 3.

[0253] Plasmid pDMWP161 was constructed in a similar fashion as pDMWP162.

[0254] A promoter-multiple cloning site-terminator region was synthesized by IDT and cloned into their pIDTsmart vector, resulting in the construction of pDMWP1. The promoter in pDMWP1 is the glucose isomerase P1.5 promoter. The following is the sequence for the synthesized region.

aggaattccctagggcgatctgtgctgtttgccacggatgcagcaccagcgcgagattatgggctcgcacgctcgactgtcggacggggg
cactggaacgagaagtcaggcgagccgtcacgcccttgactatgccacatcctgagcaaataattcaaccactaaacaaatcaaccgcgtt
tcccgaggtaaccaagcttggccggatccgcatgcgcggccgcgtcgactctagttaaaccctgggtgatcgatagctcttaattaagt
tggttgccaatgtaatgccgctgcacccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgtttatctgttgttgcgg
tgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttatacagctgtcggtaccgccag (SEQ ID
NO:36)

[0255] Construction of pDMWP4, a pBR322 with modified restriction sites was described in Example 3.

[0256] To construct pDMWP10 (pBR322*-Pgi1.5), the 438 bp EcoR1/Kpn1 fragment from pDMWP1 (pIDTSmart + P1.5-mcs-spacer-double terminator) was cloned into the 3076 bp EcoR1/Kpn1 digested pDMWP4.

[0257] pDMWP161 (pDMWP10 + BtuB2/fliA fusion) was constructed using a three fragment Gibson Assembly reaction, following the manufacturer's protocol. The vector fragment was amplified from pDMWP10 with primers ODMWP233/234. The second fragment was amplified from pUC57-btuB2 using primers ODMWP235/236. The third fragment was amplified from MG1655 chromosomal DNA using ODMWP237/238 (sequences of these primers is found in Example 3).

[0258] Plasmid pBBR1 has been described in the literature (Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop II, R.M. and Peterson, K.M.: pBBR1MCS: a broad-host-range cloning vector. *BioTechniques* 16 (1994) 800-802).

[0259] pDMWP180 was constructed by Gibson Assembly reaction, following the manufacturer's protocol, with the following two fragments:

template pBBR1, primers ODMWP320/321

ODMWP320: CCTGCAGCCCGGGGGATCCACTAGTTC (SEQ ID NO:37)

ODMWP321: AATTCGATATCAAGCTTATCGATACCG (SEQ ID NO:38)

template pDMWP162, primers ODMWP322/323

ODMWP322: GATCCCCCGGGCTGCAGGAATTCCCTAGGCGATCTGTG (SEQ ID NO:39)

ODMWP323: GTATCGATAAGCTTGATATCGAATTTATAAACGCAGAAAGGCCC (SEQ ID NO:40)

[0260] pDMWP181 was constructed by Gibson Assembly reaction, following the manufacturer's protocol, with the following two fragments:

ODMWP320/321 with pBBR1 as template.

ODMWP322/323 with pDMWP161 as template.

C. Demonstration of very low MVA production driven by a Pflia-dependent promoter in absence of FliA

[0261] Strain TV3011 (described in Example 4) was transformed with plasmid pCL1920 with selection for Spectinomycin-resistance to make strain DP2101, a control strain lacking mvaE, mvaS, and fliA genes. Strain TV3011 was also transformed with plasmid pDMWP175 with selection for Spectinomycin-resistance to make strain DP2103, which thus has a FliA promoter driving expression of mvaE and mvaS but does not have a FliA gene.

[0262] Overnight cultures of DP2101 and DP2103 were grown in LA medium with Spectinomycin 50 µg/mL at 34 °C. The following day, the cultures were diluted 1:100 into 25 mL TM3 medium with 8 mM MgSO₄, 0.02% yeast extract, 10 g/L glucose and Spectinomycin 50 µg/mL in 125 mL plastic vented top flasks. The flasks were incubated for 24 hours at 34 °C. OD₆₀₀ was measured and samples were prepared for HPLC analyses by adding 54 µL of 10% w/v sulfuric acid to a 300 µL aliquot of flask broth. The acidified tubes were held at 4 °C for 5 minutes, then spun in a centrifuge tube filter (Costar, Spin-X, 0.22 µm nylon) at 14,000 rpm for 5 minutes. The filtrate was put into HPLC vials. The acidified supernatant samples were analyzed for acetic acid and MVA by liquid chromatography. The column was a crosslinked sulfonated polystyrene-divinylbenzene resin in the hydrogen form (Aminex HPX-87H, 300 x 7.8 mm, Bio-Rad Laboratories, Inc., Hercules, CA). The mobile phase was 5 mM sulfuric acid at a flow rate of 0.6 mL/min and the column temperature was isothermal at 50 °C. Detection was ultraviolet absorbance at a wavelength of 210 nm. Calibration standards were made by diluting glacial acetic acid and mevalanolactone with 5 mM sulfuric acid. Calibration

curves were constructed and results was calculated using linear regression. The results were reported as acetic acid and MVA equivalents. The results are shown in the table 5-1 below.

Table 5-1: Production of acetic acid and MVA

Strain name	plasmid	OD600	Acetate g/L	MVA g/L
DP2101	pCL1920 (no mvaE or mvaS, no fliA)	4.87	2.77	0.17
DP2103	PfliC-mvaE-mvaS w lacO (no fliA)	3.91	2.54	0.31

[0263] The extremely low level of MVA detected in strain DP2101 was likely an artifact of the HPLC detection. Thus strain DP2103 had detectable MVA production, but only slightly more than the signal from the control strain lacking MVA production genes, thus demonstrating very low expression of mvaE and mvaS from the flhC promoter in the absence of fliA.

D. Demonstration of MVA production controlled by vitamin B12

[0264] Strain TV3011 (described in Example 4) was transformed with pDMWP180 and pDMWP181 with selection for chloramphenicol resistance on plates containing 0.1 mg/L vitamin B12, the resultant chloramphenicol-resistant strains were further transformed with pDMWP175 with selection for Spectinomycin-resistance and chloramphenicol resistance on plates containing 0.1 mg/L vitamin B12 to make strains DP2108 and DP2113, respectively. Thus, both these strains have the plasmid with a FliA promoter driving expression of mvaE and mvaS, while strain DP2108 has a weaker promoter driving expression of a btuB riboswitched fliA gene with an N-terminal translational fusion of the first 33 amino acids of BtuB and strain DP2113 has a stronger promoter.

[0265] Overnight cultures of DP2108 and DP2113 were grown in LA medium with Spectinomycin 50 µg/mL at 34 °C. The following day, the cultures were diluted 1:100 into 25

mL TM3 medium with 8 mM MgSO₄, 0.02% yeast extract, 10 g/L glucose Spectinomycin 50 µg/mL, chloramphenicol 5 µg/mL + or – vitamin 0.1 mg/L vitamin B12 in 125 mL plastic vented top flasks. The flasks were incubated for 24 hours at 34 °C. OD600 was measured and samples were prepared for HPLC analyses by adding 54 µL of 10% w/v sulfuric acid to a 300 µL aliquot of flask broth. The acidified tubes were held at 4 °C for 5 minute, then spun in a centrifuge tube filter (Costar, Spin-X, 0.22 µm nylon) at 14,000 rpm for 5 minutes. The filtrate was put into HPLC vials. The acidified supernate samples were analyzed for acetic acid and MVA by liquid chromatography. The column was a crosslinked sulfonated polystyrene-divinylbenzene resin in the hydrogen form (Aminex HPX-87H, 300 x 7.8 mm, Bio-Rad Laboratories, Inc., Hercules, CA). The mobile phase was 5 mM sulfuric acid at a flow rate of 0.6 mL/min and the column temperature was isothermal at 50 °C. Detection was ultraviolet absorbance at a wavelength of 210 nm. Calibration standards were made by diluting glacial acetic acid and mevalanolactone with 5 mM sulfuric acid. Calibration curves were constructed and results was calculated using linear regression. The results were reported as acetic acid and MVA equivalents. The results are shown in the table 5-2 below.

Table 5-2: Acetic acid and MVA production

Strain name	plasmid(s)	OD600	Acetate g/L	MVA g/L
DP2108 without vitamin B12	PfliC-upper w lacO/ pBBR-P1.20 btuB2-fliA	4.03	0.80	1.53
DP2108+ vitamin B12	"	4.4	1.36	0.46
DP2113 without vitamin B12	PfliC-upper w lacO/pBBR-P1.5 btuB2-fliA	4.6	0.59	2.67
DP2113+ vitamin B12	"	4.77	1.56	1.01

[0266] The amount of MVA produced by these two strains in the absence of vitamin B12 is greater than that produced by the single copy *fliA* plasmid (Example 4) showing that increased *fliA* expression results in elevated expression of *mvaE* and *mvaS*. Both of these strains had substantially more MVA produced in the absence of vitamin B12 than in the presence of vitamin B12, thus demonstrating vitamin B12 control of MVA production.

Example 6: Vitamin B12 transcriptional regulation of *mvaE* and *mvaS* expression

[0267] This example shows vitamin B12 transcriptional regulation of *mvaE* and *mvaS* expression when expressed from a *FliA*-dependent promoter with a vitamin B12 riboswitched *FliA*.

A. RNA samples

[0268] The cultures of strains DP2108 and DP2113 described in Example 5 were sampled for RNA at 6 hours after inoculation growth by adding 5 mL of the samples into 2x RNAprotect (Qiagen, Germantown, MD). They were left at room temp for 5 min. The samples were then spun for 10 min and the supernatant was discarded. The pellets were then frozen at -80 °C until RNA isolation was performed. RNA isolation was done using the Qiagen RNeasy Mini kit (Qiagen, Germantown, MD). Lysozyme was used to break up the cells. To TE buffer, 1 mg/mL lysozyme was added. The pellets were then resuspended in 200 µL of the TE/lysozyme mix. The samples were allowed to sit at room temperature for 10 min. To help inhibit RNase activity, 7 µL β-mercaptoethanol were added to 693µL of RLT buffer supplied in the RNeasy kit and this mix was added to each sample. After complete mixing, 500 µL of 100% ethanol was added and mixed. The samples were then applied to an RNeasy column, 700 µL at a time. They were spun for 1 min at 10,000 rpm and flow through was discarded. Another 700 µL of each sample was then added to the column, spun and again flow through was discarded. To each column, 700 µL of RW1 buffer was added and the columns were spun for 1 min.

[0269] An on column DNase treatment was then performed. The RNase-free DNase kit from Qiagen (Germantown, MD) was used to perform this treatment. To each column, a mix of 10 µL of DNase and 70 µL of RDD buffer was added. The columns were allowed to stand at room temperature for 15 min. After DNase treatment, the columns were cleaned by adding 500 µL of RW1 buffer and spun for 1 min at 15,000 G. The filtrate was discarded. The

RNeasy column was transferred to a new 2 mL collection tube and 500 μ L of RPE buffer was added. After it was spun for 1 min and the filtrate discarded, this wash step was repeated but was centrifuged for 2 min. The column was again transferred to a new 2 mL collection tube and spun for 1 min to remove any residual buffer. Finally, the column was placed in a 1.5 mL collection tube and 50 μ L of RNase free water was pipetted directly onto the RNeasy membrane. The samples were allowed to sit at room temperature for 5 min, after which they were spun for 1 min to elute the RNA. The RNA concentration was determined by measuring 2 μ L of sample on a NanoDrop spectrophotometer (Wilmington, DE). The RNA samples were then stored at -80 °C until qPCR was performed.

B. q-PCR

[0270] Quantitative Reverse Transcription PCR (qRT-PCR) analysis was performed as follows. To remove any residual genomic DNA, 3 μ g of total RNA was treated with RNase-free DNase (Qiagen, Hilden, Germany). The DNase was then inactivated by 1 mM EDTA and heating to 75 °C for 5 minutes. 1 μ g of DNase-treated RNA was then converted to cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA) as per the manufacturer's instructions. cDNA was then diluted 1:10 in water for qPCR analysis.

[0271] qPCR was performed for the target genes. All primers and probes were designed utilizing Primer Express v 3.0.1 software (Applied Biosystems, Foster City, CA). The 5' end of the TaqMan fluorogenic probes have the 6-FAMTM (6-carboxyfluorescein) fluorescent reporter dye bound, while the 3' end includes the TAMRATM (Carboxytetramethylrhodamine) quencher dye. All primers and probes were obtained from Sigma-Genosys (Woodlands, TX). Primers were evaluated for specificity utilizing BLAST analysis (www.genolevures.org/yali) and validated for quantitation utilizing genomic DNA. Primers with PCR efficiencies between 0.85 – 1.15 were validated for quantitation (data not shown).

[0272] Real-time PCR reactions included 10 pmoles each of forward and reverse primers, 2.5 pmoles of TaqMan probe, 10 μ l TaqMan Universal PCR Master Mix-No AmpErase[®] Uracil-N-Glycosylase (UNG) (Catalog No. PN 4326614, Applied Biosystems), 1 μ l 1:10 diluted cDNA, and 8.5 μ l RNase-/ DNase-free water for a total volume of 20 μ l per reaction. Reactions were run on the ABI PRISM[®] 7900 Sequence Detection System under the following conditions: initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 95

°C for 15 sec and annealing at 60 °C for 1 min. Real time data was collected automatically during each cycle by monitoring 6-FAM™ fluorescence.

[0273] Relative expression (RQ) was calculated using Data Assist Software v3.01 and the $\Delta\Delta C_t$ method (Applied Biosystems, Foster City, Ca). The *rrsB* gene was utilized for data normalization. Relative expression was then calculated by comparing the gene expression in the different strains and growth conditions. The gene expression in the table 6-1 below is relative to the sample from DP2108 in the presence of vitamin B12.

Table 6-1: Gene expression as measured by q-PCR

	mvaE	mvaE SD	mvaS	mvaS SD
DP2108 without vitamin B12	25.69	1.56	13.60	0.79
DP2108 0.1 mg/L vitamin B12	1.00	0.05	1.00	0.13
DP2113 without vitamin B12	15.93	0.88	11.69	1.31
DP2113 0.1 mg/L vitamin B12	0.87	0.06	1.56	0.08

[0274] These results show strong regulation by vitamin B12 via the riboswitched FliA of *mvaE* and *mvaS* transcription at mid-log growth.

CLAIMS

What is claimed is:

1. A regulatable gene expression construct comprising:
 - (i) a nucleic acid encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and
 - (ii) a sigma factor-dependent promoter operably linked to one or more nucleic acids encoding a gene of interest.
2. The construct of claim 1, wherein the riboswitch comprises an aptamer domain and an expression platform domain.
3. The construct of any one of claims 1-2, wherein said riboswitch is responsive to a small molecule selected from the group consisting of cobalamin (coenzyme of vitamin B12), flavin mononucleotide (FMN), glutamine, glycine, lysine, purine, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), tetrahydrofolate, and thiamin pyrophosphate (TPP).
4. The construct of any of claims 1-3, wherein said riboswitch is responsive to cobalamin (coenzyme of vitamin B12).
5. The construct of any of one of claims 1-4, wherein said sigma factor is selected from the group consisting of sigma 70 (RpoD), sigma 19 (FecI), sigma 24 (RpoE), sigma 28 (RpoF), sigma 32 (RpoH), sigma 38 (RpoS) and sigma 54 (RpoN).
6. The construct of any one of claims 1-5, wherein said sigma factor is sigma 28 (RpoF).
7. The construct of any one of claims 1-6, where said sigma 28 is encoded by the *fliA* gene.
8. The construct of any one of claims 1-7, wherein said sigma factor-dependent promoter is activated by a sigma factor is selected from the group consisting of sigma 70 (RpoD), sigma 19 (FecI), sigma 24 (RpoE), sigma 28 (RpoF), sigma 32 (RpoH), sigma 38 (RpoS) and sigma 54 (RpoN).
9. The construct of any one of claims 1-8, wherein said sigma factor-dependent promoter is activated by sigma 28 (RpoF).

10. The construct of any one of claims 1-9, wherein said sigma factor-dependent promoter is the *pfliC* promoter.
11. A recombinant bacterial cell comprising the regulatable gene construct of any one of claims 1-10.
12. The recombinant bacterial cell of claim 11, wherein when said bacterial cell is cultured in the presence of said small molecule and the expression of said sigma factor is repressed, thereby inhibiting the expression of said gene of interest operably linked to said sigma factor-dependent promoter.
13. The recombinant bacterial cell of claim 11, wherein when said bacterial cell is cultured in the absence of said small molecule and the expression of said sigma factor is activated, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter.
14. The recombinant bacterial cell of any one of claims 11-13, wherein said cell comprises:
 - (i) a first extrachromosomal plasmid comprising a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and
 - (ii) a second extrachromosomal plasmid comprising a sigma factor-dependent promoter operably linked one or more nucleic acids encoding a gene of interest.
15. The recombinant bacterial cell of any one of claims 11-13, wherein said cell comprises a single extrachromosomal plasmid comprising:
 - (i) a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and
 - (ii) a sigma factor-dependent promoter operably linked one or more nucleic acids encoding a gene of interest.
16. The recombinant bacterial cell of any one of claims 11-13, comprising:
 - (i) a nucleic acid molecule encoding an RNA comprising a small molecule-responsive riboswitch operably linked to one or more nucleic acids encoding a sigma factor; and

(ii) a sigma factor-dependent promoter operably linked one or more nucleic acids encoding a gene of interest,

wherein at least one of (i) or (ii) is integrated into the chromosome of said bacterial cell.

17. The recombinant bacterial cell of any one of claims 11-16, wherein said bacterial cell further comprises a genetic modification to increase intracellular transport of a small molecule selected from the group consisting of cobalamin (coenzyme of vitamin B12), flavin mononucleotide (FMN), glutamine, glycine, lysine, purine, S-adenosylhomocysteine (SAH), S-adenosyl methionine (SAM), tetrahydrofolate, or thiamin pyrophosphate (TPP).
18. The recombinant bacterial cell of any one of claims 11-17, wherein said bacterial cell further comprises a genetic modification to increase intracellular transport of cobalamin (coenzyme of vitamin B12).
19. A method for regulating the expression of a gene of interest comprising:
 - (i) initially culturing the recombinant bacterial cell of any of claims 11-18 in a first culture condition comprising said small molecule in an amount sufficient to repress expression of said sigma factor, thereby inhibiting the expression of said gene of interest operably linked to said sigma factor-dependent promoter; and
 - (ii) subsequently culturing said recombinant bacterial cells under a second culture condition comprising an amount of said small molecule that is not sufficient to repress expression of said sigma factor, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in second culture conditions.
20. The method of claim 19, where said initial culture is added into said second culture, wherein the addition of the initial culture to the second culture results in the dilution of said small molecule such that the amount of said small molecule in said second culture is not sufficient to repress said sigma factor expression, thereby activating the expression of said gene of interest operably linked to said sigma factor-dependent promoter in second culture conditions.

FIG. 1A

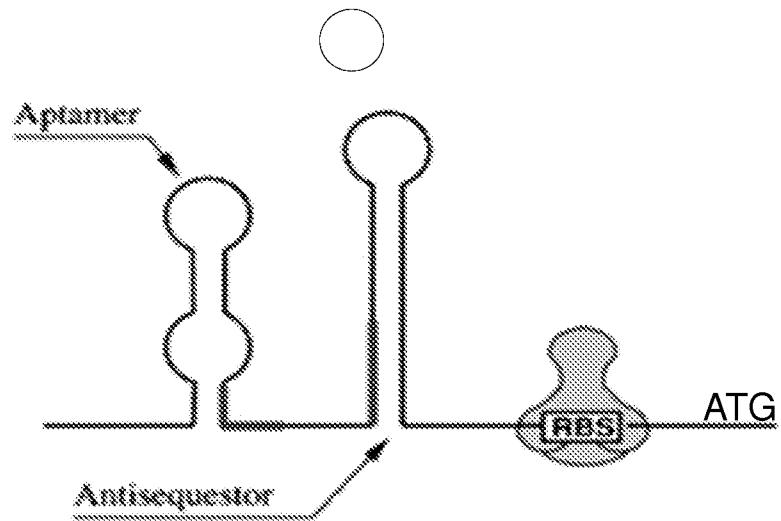


FIG. 1B

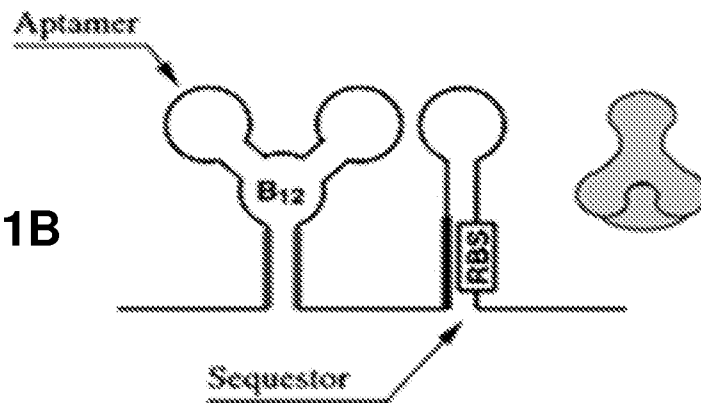


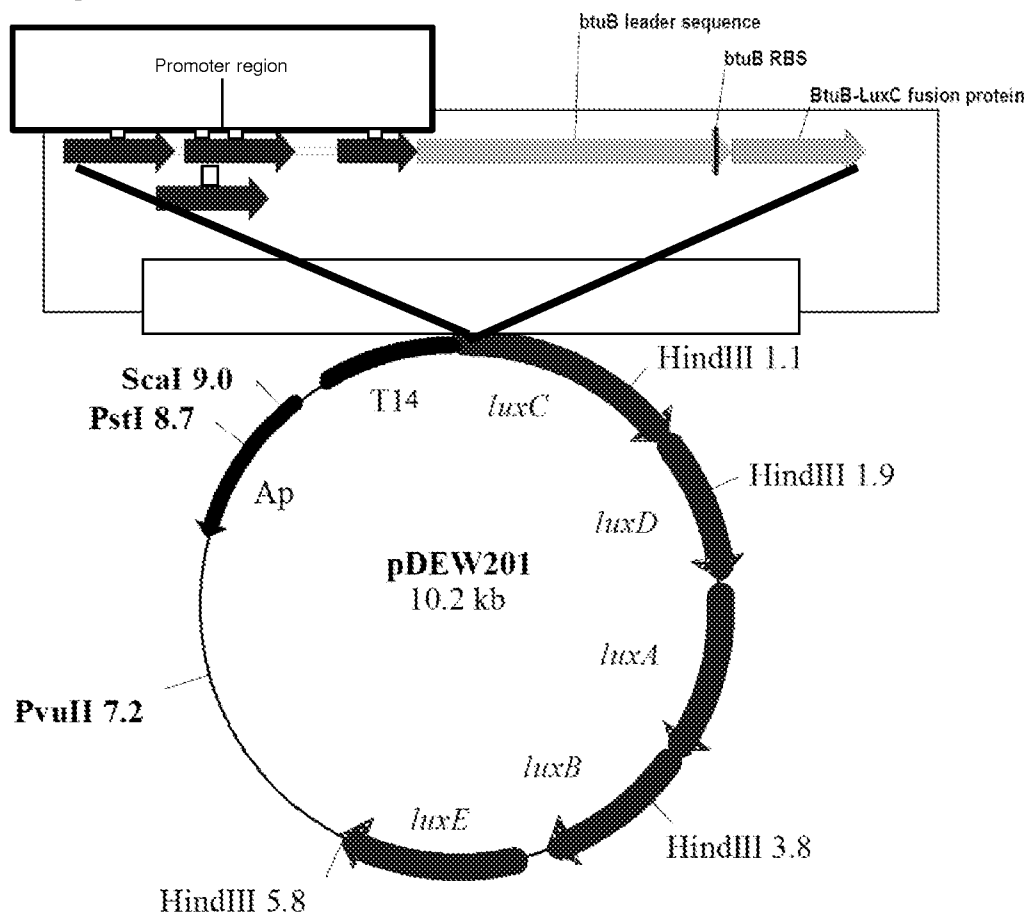
FIG. 2

FIG. 3B

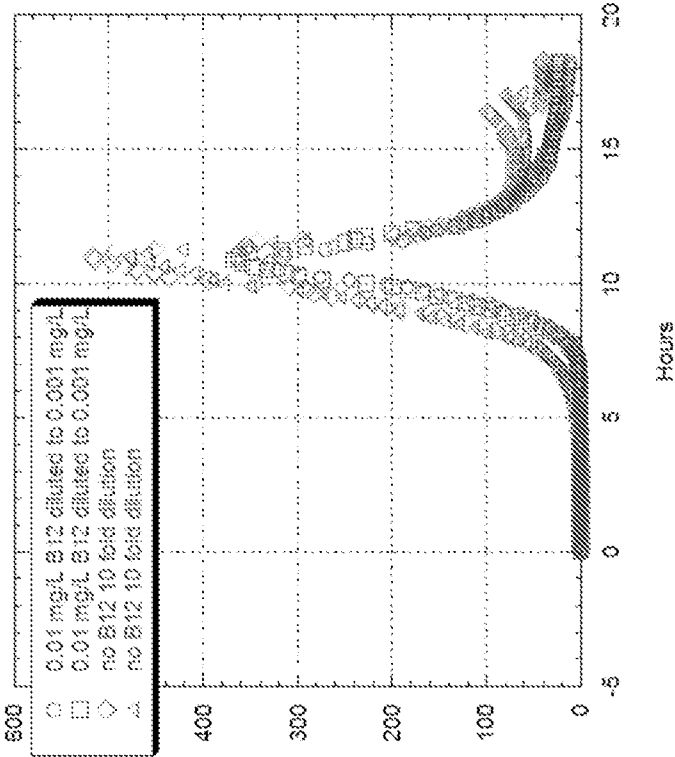
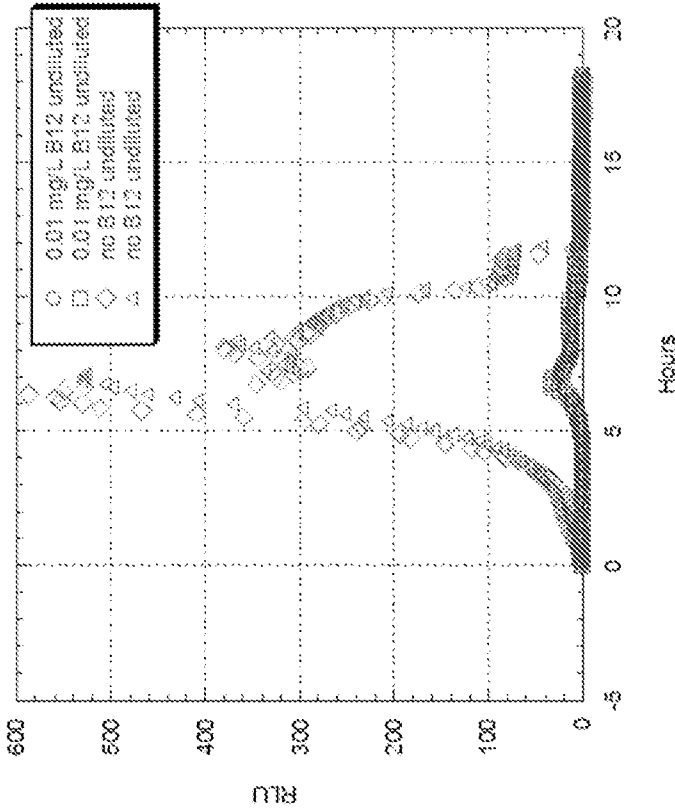
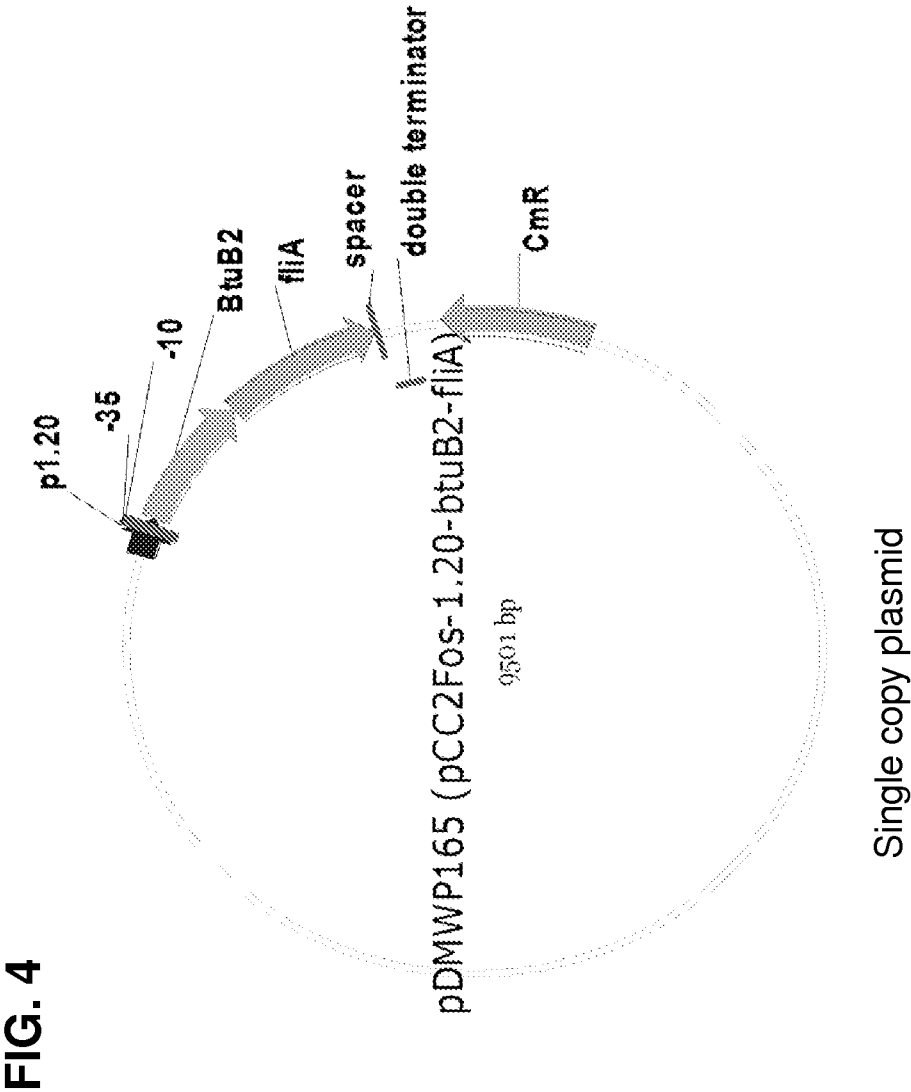
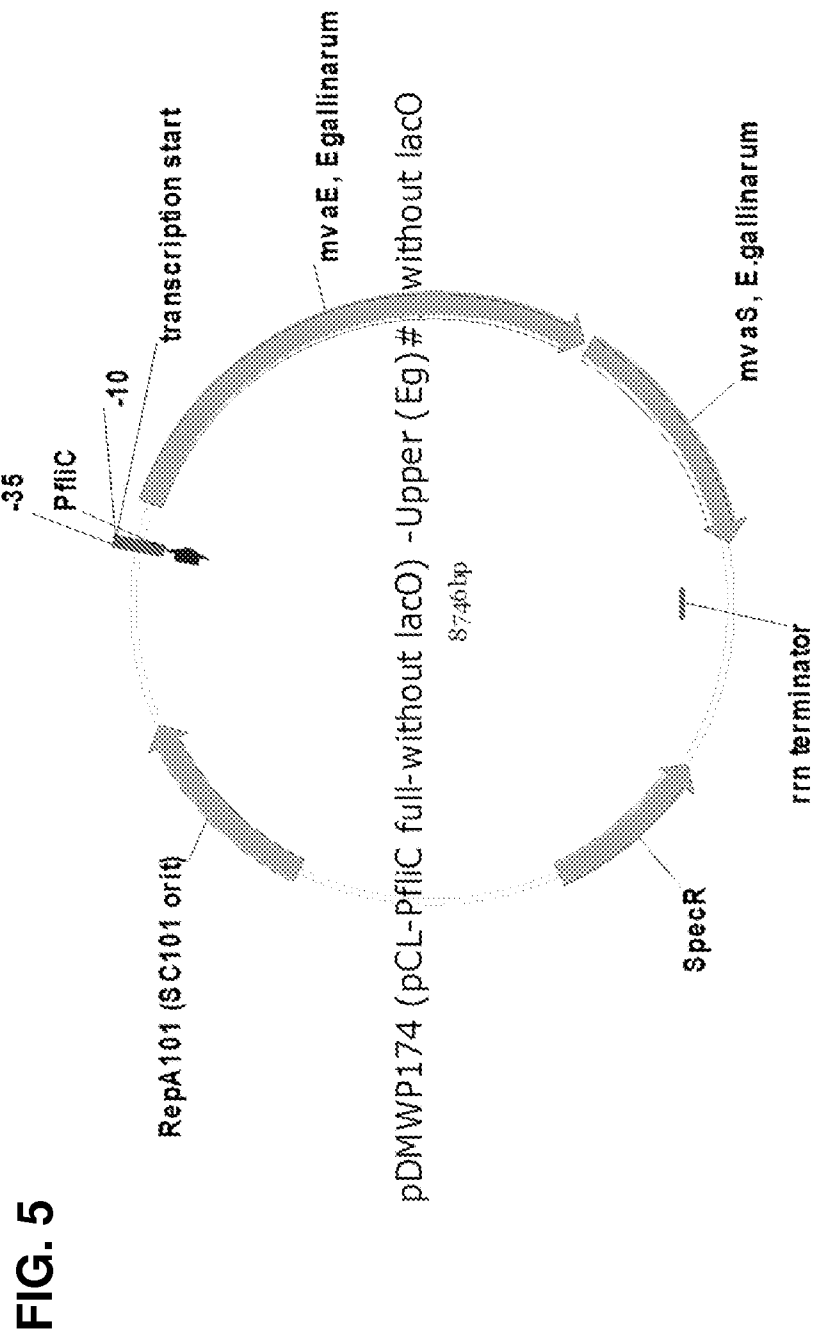


FIG. 3A







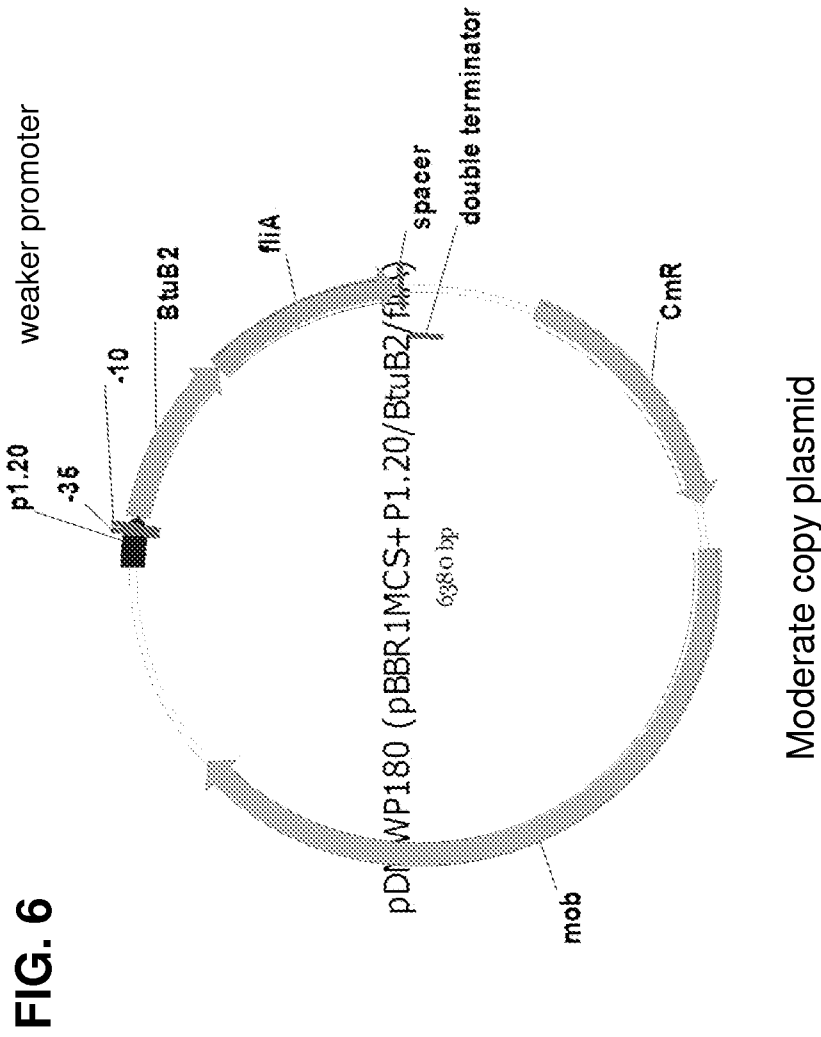
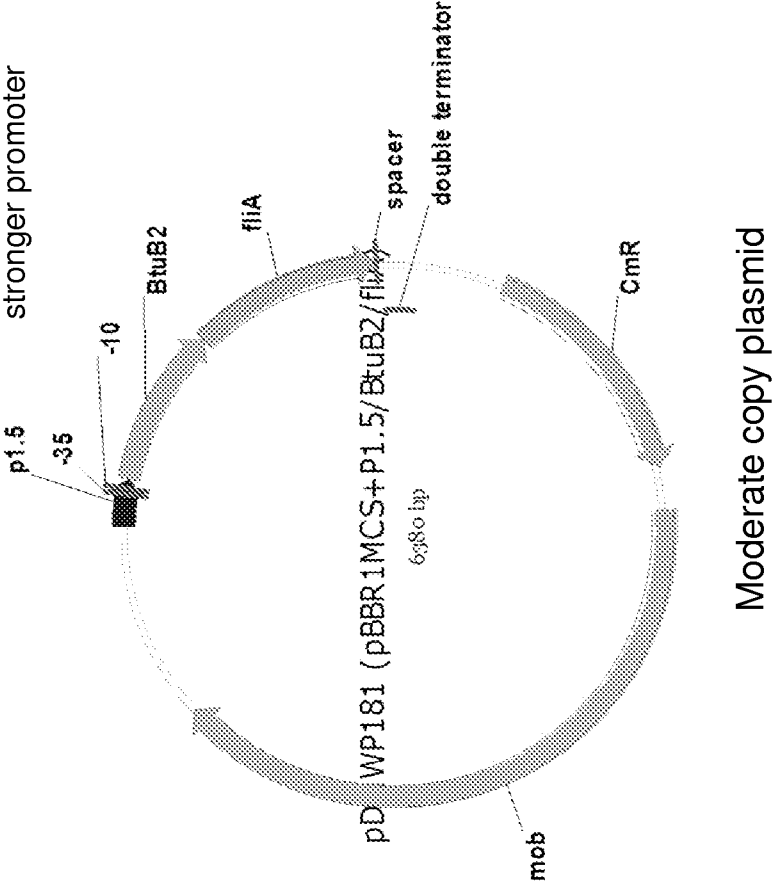


FIG. 7



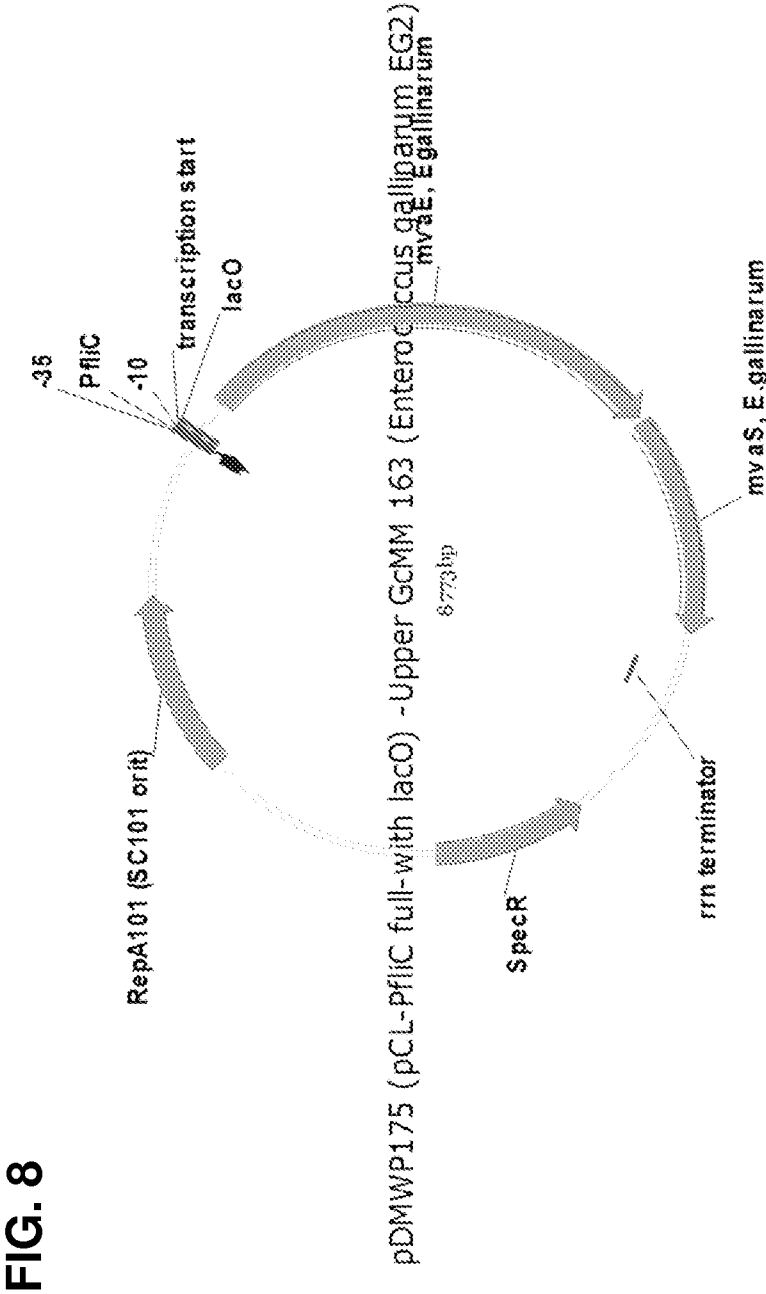


FIG. 9

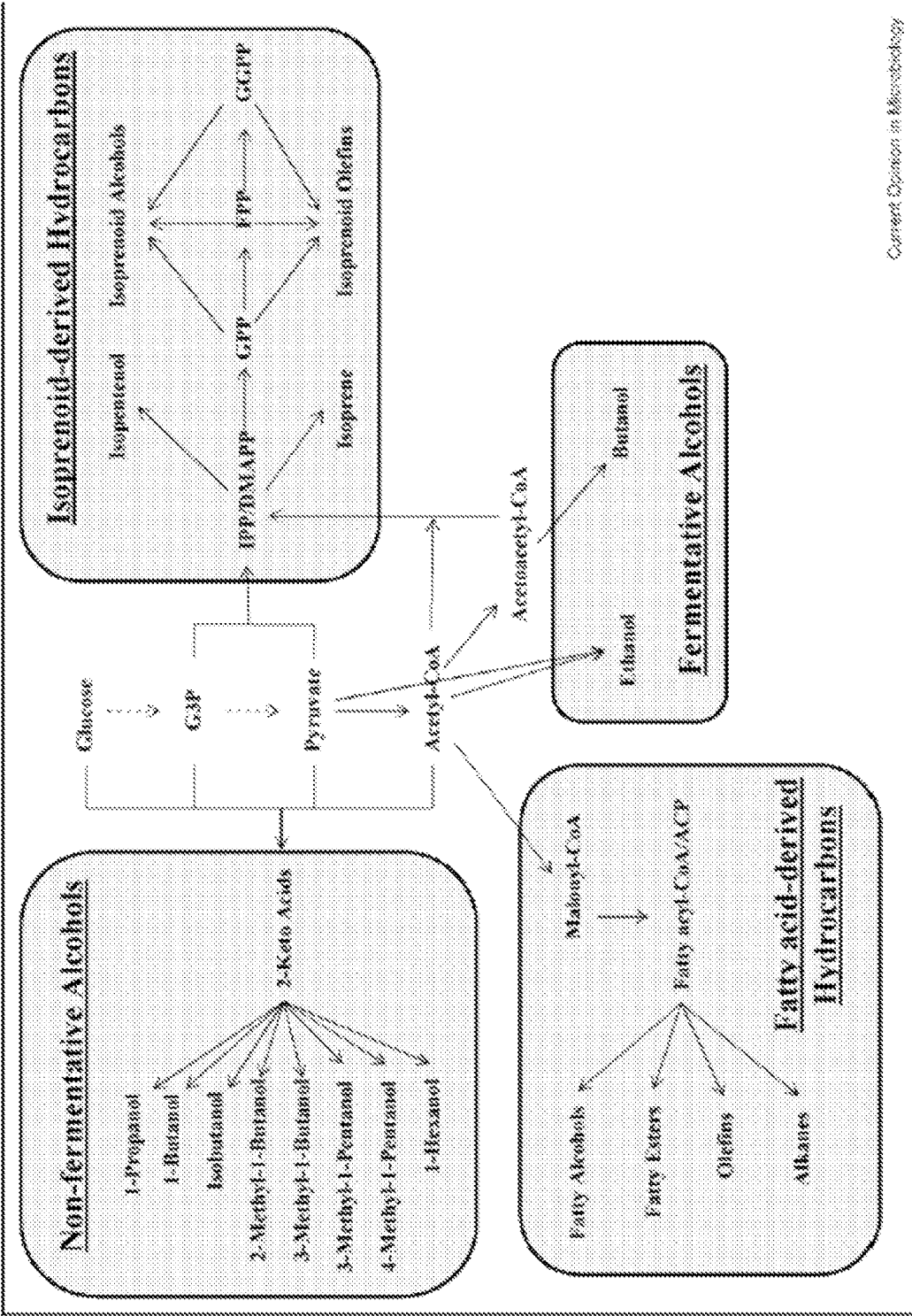
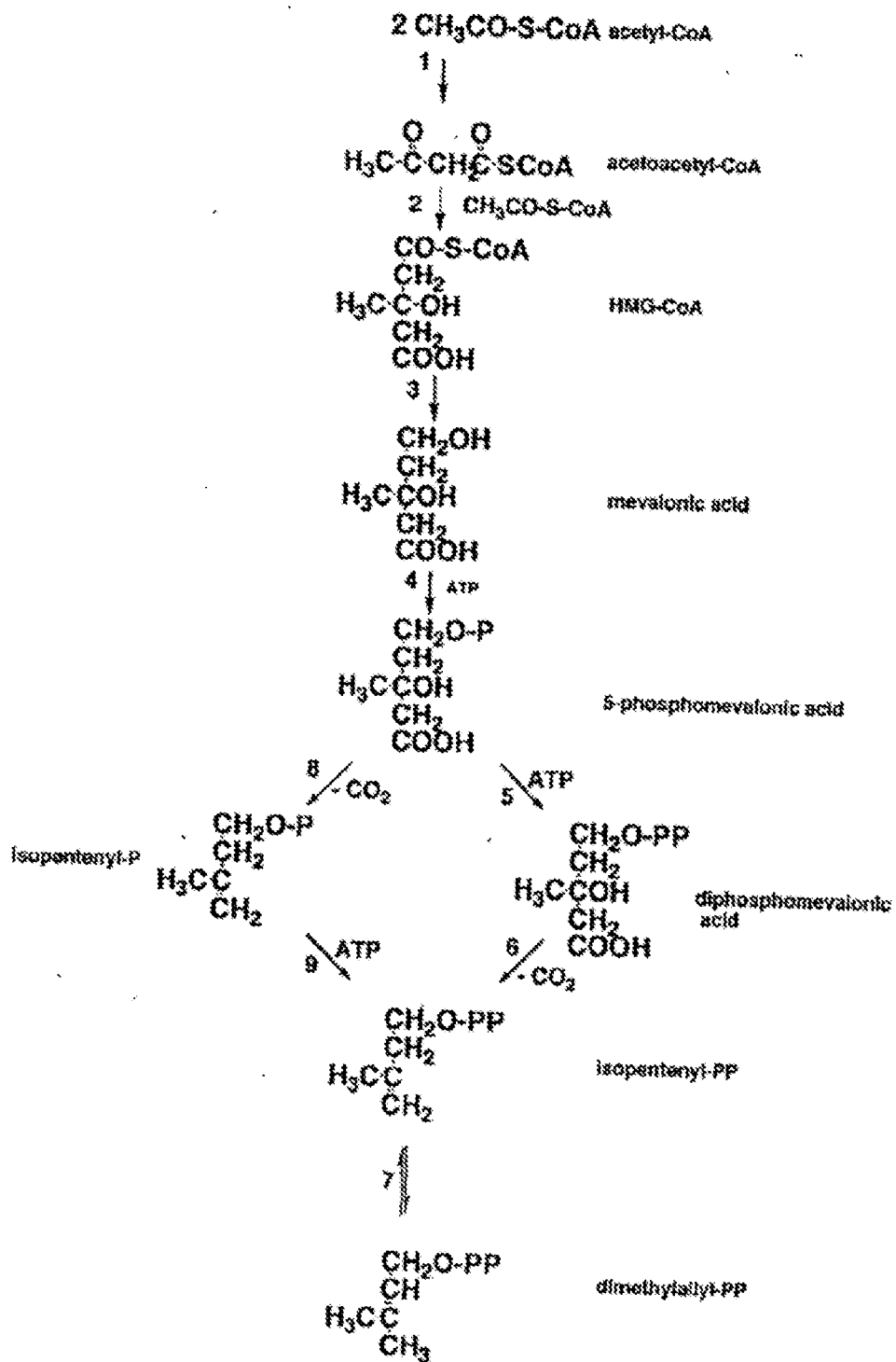


FIG. 10



Increase DXP Pathway Flux

FIG. 11

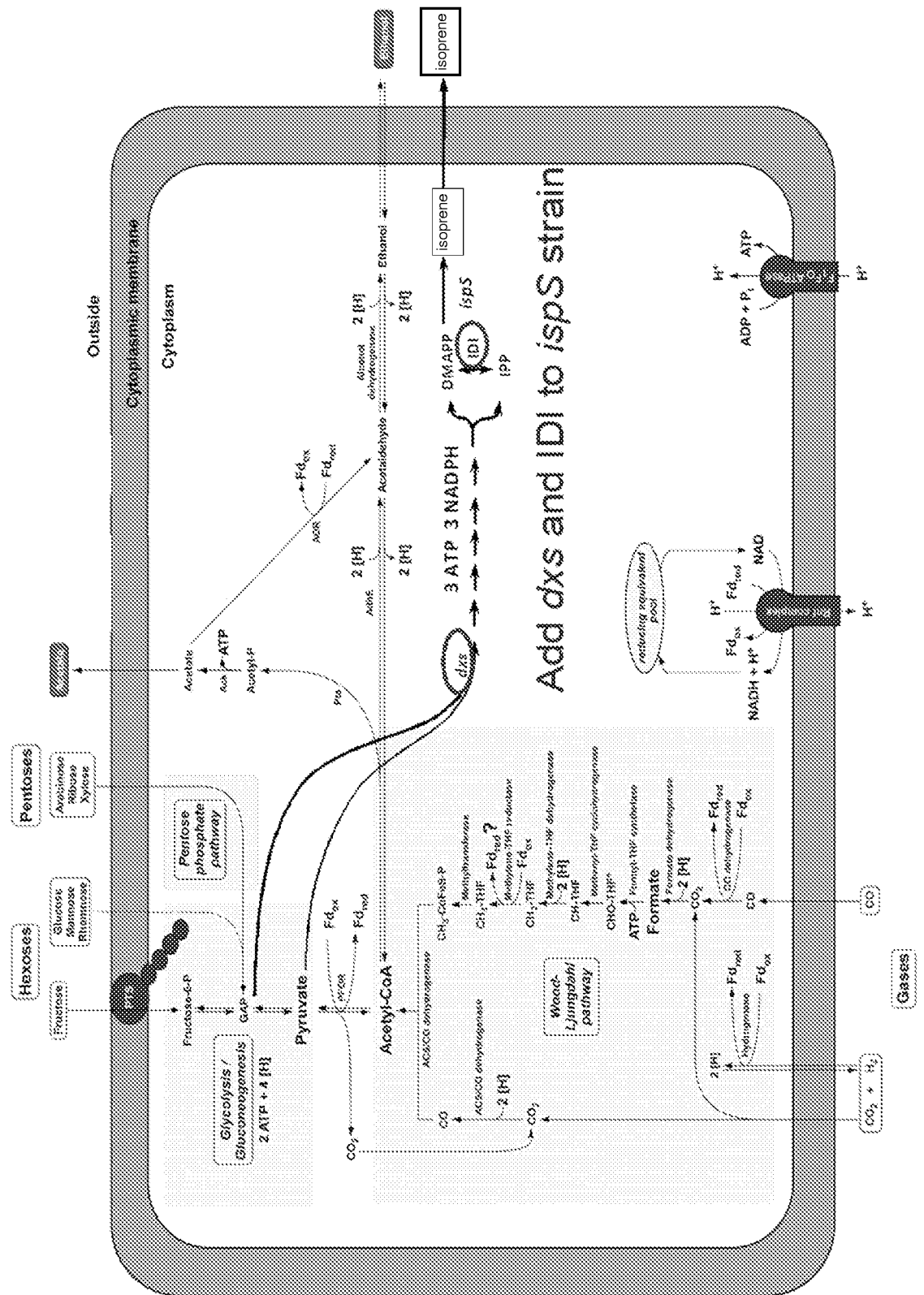
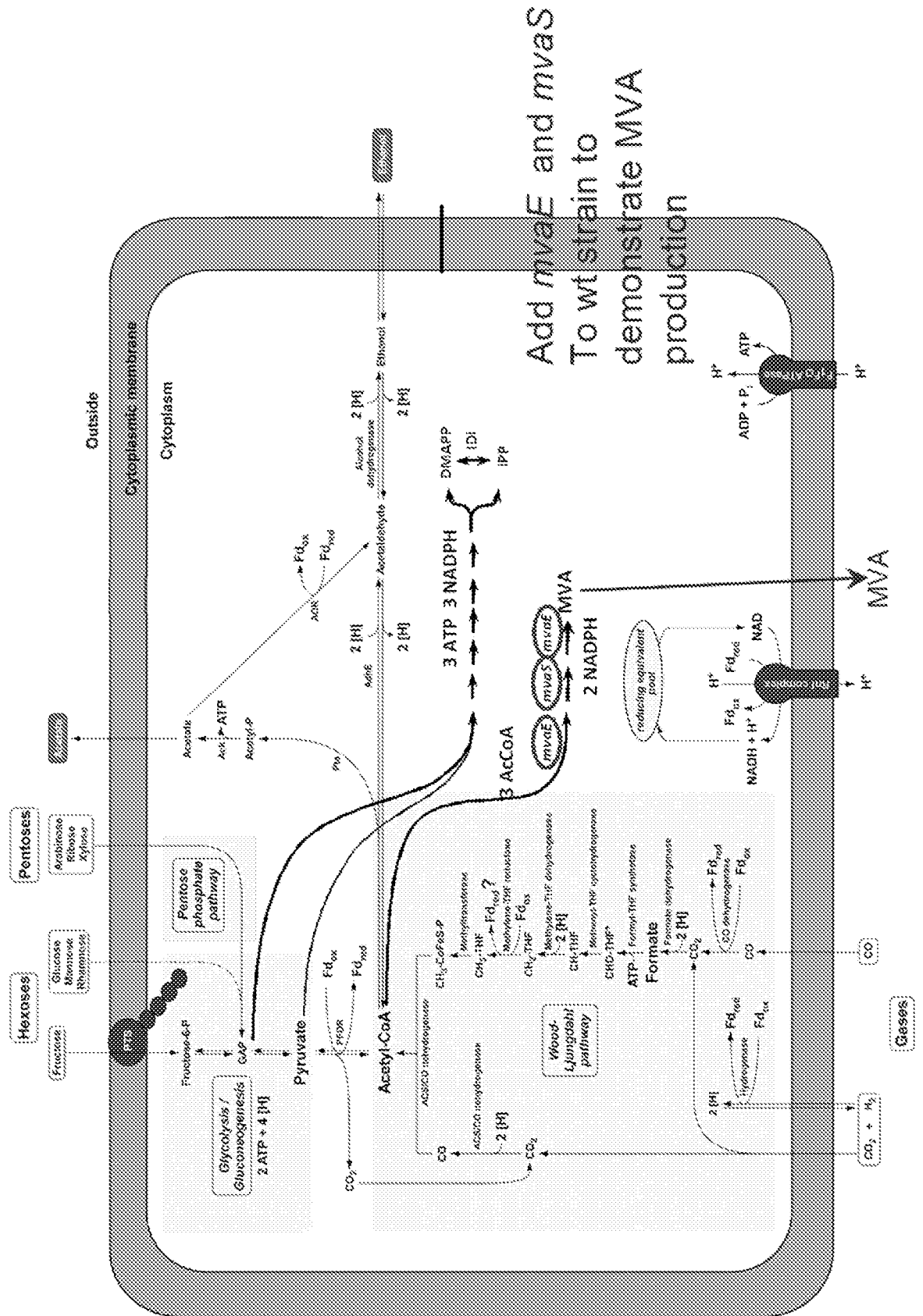
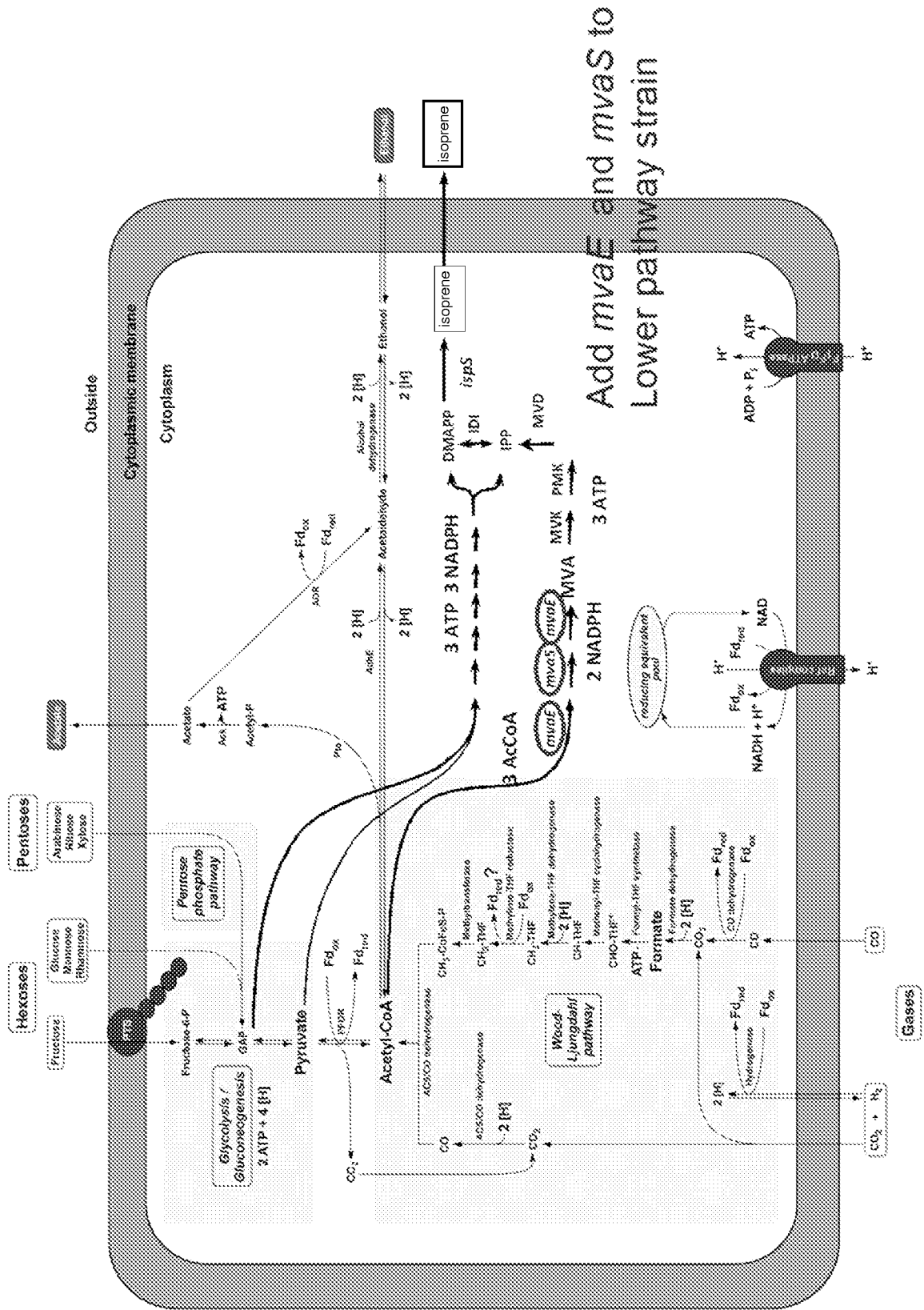


FIG. 12 Express Upper MVA Pathway



Express Full MVA Pathway



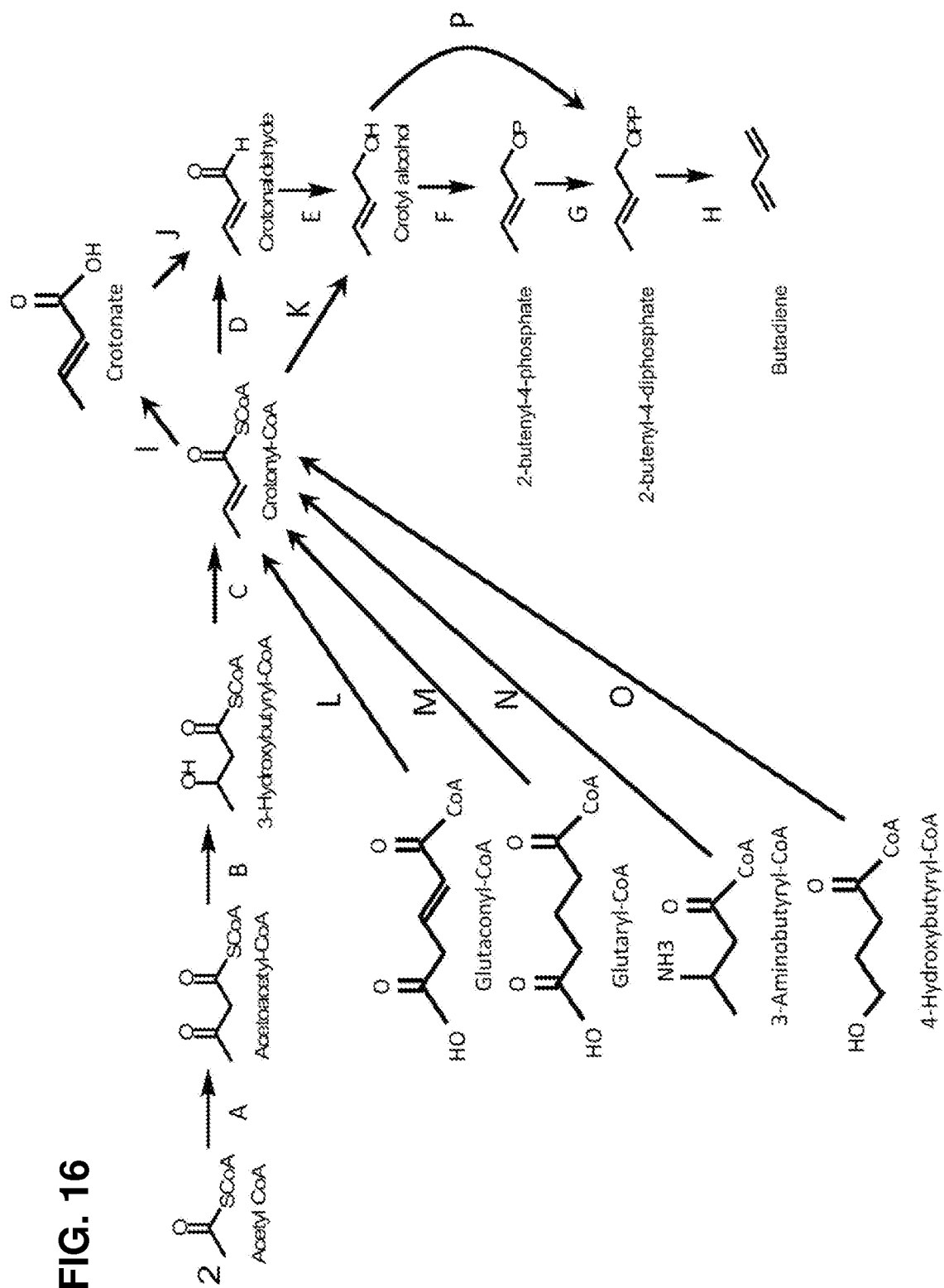
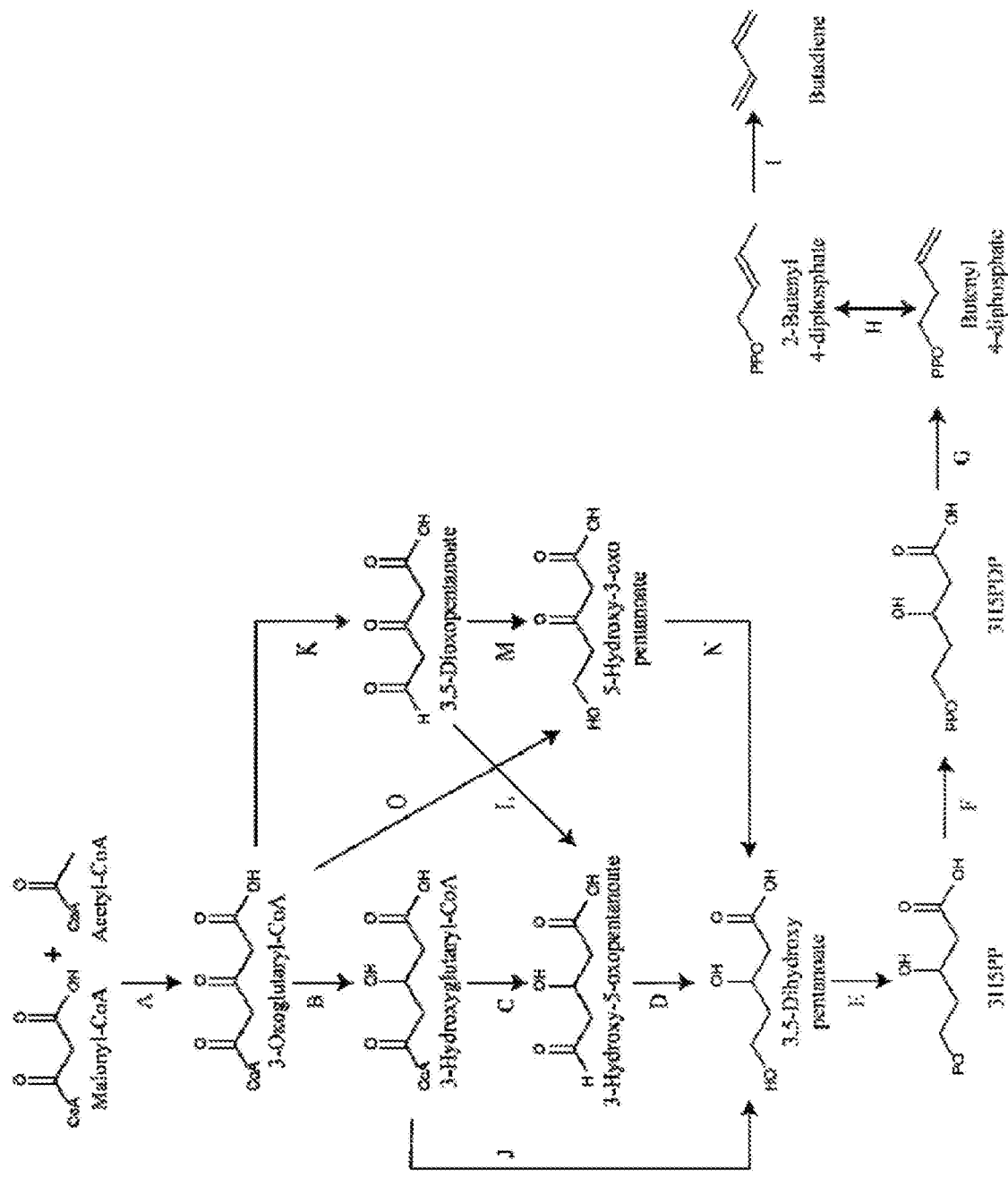


FIG. 18



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/71536

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 4-20
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/71536

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/09, C12N 15/113, C12N 15/63 (2015.01)

CPC - C12N 15/63, C12N 15/09, C12N 15/113

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - C12N 15/09, C12N 15/113, C12N 15/63 (2015.01)

CPC - C12N 15/63, C12N 15/09, C12N 15/113

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - C12N 15/11

(keyword limited; terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatBase, Google Patents, Google Scholar

Search terms: riboswitch, sigma factor, expression construct, aptamer, expression platform domain, cobalamin, flavin mononucleotide, FMN, glutamine, glycine, lysine, purine, Sadenosylhomocysteine, SAH, S-adenosyl methionine, SAM, tetrahydrofolate, thiamin pyrophosph

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2013/0004980 A1 (HUANG et al.) 03 January 2013 (03.01.2013) para [0016], [0027], [0101], [0103], [0156], [0157]	1-3
Y	GRUBER et al., Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol. 2003, Vol 57, pp 441-e66. Especially abstract; p 442, para 1	1-3
Y	US 2010/0041742 A1 (BREAKER et al.) 18 February 2010 (18.02.2010) para [0109], [0110]	3

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

25 February 2015 (25.02.2015)

Date of mailing of the international search report

18 MAR 2015

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774