U.S. PATENT AND TRADEMARK OFFICE
PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 C.F.R. §1.53(c).	Docket Number	GEVO-041/03US	Type a plus sign (+) inside this box→	+

INVENTOR(s)/APPLICANT(s)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
URANO	Jun		Aurora, Colorado
ASLESON DUNDON	Catherine		Englewood, Colorado
MEINHOLD	Peter		Denver, Colorado
FELDMAN	M. Renny	Reid	Denver, Colorado
ARISTIDOU	Aristos		Highlands Ranch, Colorado
HAWKINS	Andrew		Parker, Colorado
BUELTER	Thomas		Denver, Colorado
PETERS	Matthew		Highlands Ranch, Colorado
LIES	Doug		Parker, Colorado
PORTER-SCHEINMAN	Stephanie		Conifer, Colorado
SMITH	Christopher		Englewood, Colorado
DEY	Melissa		Aurora, Colorado

TITLE OF INVENTION

CYTOSOLICALLY ACTIVE DIHYDROXYACID DEHYDRATASES

CORRESPONDENCE ADDRESS

INDIVIDUAL AND FIRM NAME: **COOLEY GODWARD KRONISH LLP**
CUSTOMER NUMBER: **58249**

ADDRESS	ATTN: PATENT GROUP, 777 6th Street NW, Suite 1100 Tel: (202) 842-7800 Fax: (202) 842-7899

CITY Washington	STATE DC	ZIP CODE 20001	COUNTRY U.S.A.

ENCLOSED APPLICATION PARTS (check all that apply)

[X]   Specification - No. of Pages **132**   [] Application Data Sheet

[X]   Drawing(s) – No. of Sheets **1**   [] CD(s), number   [X] Other (specify): Sequence Listing (11 pages)

Application Size Fee: If the specification and drawings exceed 100 sheets of paper, the application size fee due is $270 ($135 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s)

METHOD OF PAYMENT (check one)

[X] The Commissioner is hereby authorized to charge the filing fee of **$245.00** required by this paper, and to credit any overpayment, to Deposit Account No. 50-1283.	PROVISIONAL FILING FEE AMOUNT ($) [] $220.00 Large Entity [X] $110.00 Small Entity

The Invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
[X]   Yes, the names of the U.S. Government agency and the Government contract number are: The National Science Foundation, Grant Number IIP-0823122 and the United States Environmental Protection Agency, Grant Number EP-D-09-023.

Respectfully submitted,

SIGNATURE _*Paul Wickman*_   Dated: _11/24/2009_

TYPED or PRINTED NAME: Paul Wickman   REGISTRATION NO. 61,242
103209 v1/DC

CYTOSOLICALLY ACTIVE DIHYDROXYACID DEHYDRATASES

SPECIFICATION

TO WHOM IT MAY CONCERN:

Be it known that we, with names, residences, and citizenships listed below, have invented the inventions described in the following specification entitled:

CYTOSOLICALLY ACTIVE DIHYDROXYACID DEHYDRATASES

Jun Urano
    Residence: Aurora, Colorado
    Citizenship: USA

Catherine Asleson Dundon
    Residence: Englewood, Colorado
    Citizenship: USA

Peter Meinhold
    Residence: Denver, Colorado
    Citizenship: Germany

Renny Feldman
    Residence: Denver, Colorado
    Citizenship: USA

Aristos Aristidou
    Residence: Highlands Ranch, Colorado
    Citizenship: Cyprus

Andrew Hawkins
    Residence: Parker, Colorado
    Citizenship: USA

Thomas Buelter
    Residence: Denver, Colorado
    Citizenship: Germany

Matthew Peters
    Residence: Highlands Ranch, Colorado
    Citizenship: USA

Doug Lies
    Residence: Parker, Colorado

Citizenship: USA

Stephanie Porter-Scheinman
      Residence: Conifer, Colorado
      Citizenship: USA

Christopher Smith
      Residence: Englewood, Colorado
      Citizenship: USA

Melissa Dey
      Residence: Aurora, Colorado
      Citizenship: USA

CYTOSOLICALLY ACTIVE DIHYDROXYACID DEHYDRATASES

TECHNICAL FIELD

[0001]   Recombinant microorganisms and methods of producing such organisms are provided. Also provided are methods of producing metabolites that are biofuels by contacting a suitable substrate with recombinant microorganisms and enzymatic preparations therefrom.

BACKGROUND

[0002]   Biofuels have a long history ranging back to the beginning of the 20th century. As early as 1900, Rudolf Diesel demonstrated at the World Exhibition in Paris, France, an engine running on peanut oil. Soon thereafter, Henry Ford demonstrated his Model T running on ethanol derived from corn. Petroleum-derived fuels displaced biofuels in the 1930s and 1940s due to increased supply, and efficiency at a lower cost.

[0003]   Market fluctuations in the 1970s coupled to the decrease in US oil production led to an increase in crude oil prices and a renewed interest in biofuels. Today, many interest groups, including policy makers, industry planners, aware citizens, and the financial community, are interested in substituting petroleum-derived fuels with biomass-derived biofuels. The leading motivations for developing biofuels are of economical, political, and environmental nature.

[0004]   One is the threat of 'peak oil', the point at which the consumption rate of crude oil exceeds the supply rate, thus leading to significantly increased fuel cost results in an increased demand for alternative fuels. In addition, instability in the Middle East and other oil-rich regions has increased the demand for domestically produced biofuels. Also, environmental concerns relating to the possibility of carbon dioxide related climate change is an important social and ethical driving force which is starting to result in government regulations and policies such as caps on carbon dioxide emissions from automobiles, taxes on carbon dioxide emissions, and tax incentives for the use of biofuels.

[0005]   Ethanol is the most abundant fermentatively produced fuel today but has several drawbacks when compared to gasoline. Butanol, in comparison, has several advantages over ethanol as a fuel: it can be made from the same feedstocks as ethanol but, unlike ethanol, it is compatible with gasoline at any ratio and can also be used as a pure fuel in existing combustion engines without modifications. Unlike ethanol, butanol does not absorb water and can thus be stored and distributed in the existing petrochemical infrastructure. Due to its higher energy content which is close to that of gasoline, the fuel economy (miles per gallon) is better than that of ethanol. Also, butanol-gasoline blends have lower vapor pressure than ethanol-gasoline blends, which is important in reducing evaporative hydrocarbon emissions.

[0006]   Isobutanol has the same advantages as butanol with the additional advantage of having a higher octane number due to its branched carbon chain. Isobutanol is also useful as a commodity chemical and is also a precursor to MTBE.

Isobutanol can be produced in microorganisms expressing a heterologous metabolic pathway, but these microorganisms are not of commercial relevance due to their inherent low performance characteristics, which include low productivity, low titer, low yield, and the requirement for oxygen during the fermentation process.

[0007] The present inventors have overcome these problems by developing metabolically engineered microorganisms that exhibit increased isobutanol productivity, titer, and/or yield.

SUMMARY OF THE INVENTION

[0008] The present invention provides cytosolically active dihydroxyacid dehydratase (DHAD) enzymes and recombinant microorganisms comprising said cytosolically active DHAD enzymes. In some embodiments, said recombinant microorganisms may further comprise one or more additional enzymes catalyzing a reaction in an isobutanol producing metabolic pathway. As described herein, the recombinant microorganisms of the present invention are useful for the production of several beneficial metabolites, including, but not limited to isobutanol.

[0009] In a first aspect, the invention provides cytosolically active dihydroxyacid dehydratase (DHAD) enzymes. These cytosolically active DHAD enzymes will generally exhibit the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol. The cytosolically active DHAD enzymes of the present invention, as described herein, can include modified or alternative dihydroxyacid dehydratase (DHAD) enzymes, wherein said DHAD enzymes exhibit increased cytosolic activity as compared to the parental or native DHAD enzyme.

[0010] In various embodiments described herein, the DHAD enzymes may be derived from a prokaryotic organism. In one embodiment, the prokaryotic organism is a bacterial organism. In another embodiment, the bacterial organism is *Lactococcus lactis*. In a specific embodiment, the DHAD enzyme from *L. lactis* comprises the amino acid sequence of SEQ ID NO: 9. In another embodiment, the bacterial organism is *Escherichia coli*. In a specific embodiment, the DHAD enzyme from *E. coli* comprises the amino acid sequence of SEQ ID NO: 129.

[0011] In alternative embodiments described herein, the DHAD enzyme may be derived from a eukaryotic organism. In one embodiment, the eukaryotic organism is a fungal organism. In an exemplary embodiment, the fungal organism is *Piromyces sp.* E2. In another embodiment, the eukaryotic organism is a yeast organism, such as *S. cerevisiae*. In another embodiment, the eukaryotic organism is selected from the group consisting of the genera Enamoeba and Giardia.

[0012] In some embodiments, the invention provides modified or mutated DHAD enzymes, wherein said DHAD enzymes exhibit increased cytosolic activity as compared to their parental DHAD enzymes. In another embodiment, the invention provides modified or mutated DHAD enzymes, wherein said DHAD enzymes exhibit increased cytosolic activity as compared to the DHAD enzyme comprised by the amino acid sequence of SEQ ID NO: 11.

[0013] In some embodiments, the invention provides modified or mutated DHAD enzymes have one or more amino acid deletions at the N-terminus. In one embodiment, said modified or mutated DHAD enzyme has at least about 10 amino

acid deletions at the N-terminus. In another embodiment, said modified or mutated DHAD enzyme has at least about 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or amino acid deletions at the N-terminus. In a specific embodiment, said modified or mutated DHAD has 19 amino acid deletions at the N-terminus. In another specific embodiment, said modified or mutated DHAD has 23 amino acid deletions at the N-terminus.

[0014] In further embodiments, the invention provides DHAD enzymes comprising the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid, and wherein said DHAD enzymes exhibit the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

[0015] In additional embodiments, the invention provides DHAD enzymes comprising the amino acid sequence CPGXGXC (SEQ ID NO: 123), wherein X is any amino acid, and wherein said DHAD enzymes exhibit the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

[0016] In another embodiment, the invention provides DHAD enzymes comprising the amino acid sequence CPGXG(A/S)C (SEQ ID NO: 124), wherein X is any amino acid, and wherein said DHAD enzymes exhibit the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

[0017] In yet another embodiment, the invention provides DHAD enzymes comprising the amino acid sequence CXXXPGXGXC (SEQ ID NO: 125), wherein X is any amino acid, and wherein said DHAD enzymes exhibit the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

[0018] In some embodiments, the DHAD enzymes of the present invention exhibit a properly folded iron-sulfur cluster domain and/or redox active domain in the cytosol. In one embodiment, the DHAD enzymes comprise a mutated or modified iron-sulfur cluster domain and/or redox active domain.

[0019] In another aspect, the present invention provides recombinant microorganisms comprising a cytosolically active DHAD enzyme. In one embodiment, the invention provides recombinant microorganisms comprising a DHAD enzyme derived from a prokaryotic organism, wherein said DHAD enzyme exhibits activity in the cytosol. In one embodiment, the DHAD enzyme is derived from a bacterial organism. In a specific embodiment, the DHAD enzyme is derived from *L. lactis* and comprises the amino acid sequence of SEQ ID NO: 9. In another embodiment, the invention provides recombinant microorganisms comprising a DHAD enzyme derived from a eukaryotic organism, wherein said DHAD enzyme exhibits activity in the cytosol. In one embodiment, the DHAD enzyme is derived from a fungal organism. In an alternative embodiment, the DHAD enzyme is derived from a yeast organism.

[0020] In one embodiment, the invention provides recombinant microorganisms comprising a modified or mutated DHAD enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the parental DHAD enzyme. In another embodiment, the invention provides recombinant microorganisms comprising a modified or mutated DHAD enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the DHAD enzyme comprised by the amino acid sequence of SEQ ID NO: 11.

[0021] In another embodiment, the invention provides recombinant microorganisms comprising a DHAD enzyme comprising the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

[0022] In some embodiments, the invention provides recombinant microorganisms comprising a DHAD enzyme fused to a peptide tag, whereby said DHAD enzyme exhibits increased cytosolic localization and/or cytosolic DHAD activity as compared to the parental microorganism. In one embodiment, the peptide tag is non-cleavable. In another embodiment, the peptide tag is fused at the N-terminus of the DHAD enzyme. In another embodiment, the peptide tag is fused at the C-terminus of the DHAD enzyme. In certain embodiments, the peptide tag may be selected from the group consisting of ubiquitin, ubiquitin-like (UBL) proteins, myc, HA-tag, green fluorescent protein (GFP), and the maltose binding protein (MBP).

[0023] In various embodiments described herein, the recombinant microorganisms may further comprise a nucleic acid encoding a chaperone protein, wherein said chaperone protein assists the folding of a protein exhibiting cytosolic activity. In a preferred embodiment, the protein exhibiting cytosolic activity is DHAD. In one embodiment, the chaperone may be a native protein. In another embodiment, the chaperone protein may be an exogenous protein. In some embodiments, the chaperone protein may be selected from the group consisting of: endoplasmic reticulum oxidoreductin 1 (Ero1, accession no. NP_013576.1), including variants of Ero1 that have been suitably altered to reduce or prevent its normal localization to the endoplasmic reticulum; thioredoxins (which includes Trx1, accession no. NP_013144.1; and Trx2, accession no. NP_011725.1), thioredoxin reductase (Trr1, accession no. NP_010640.1); glutaredoxins (which includes Grx1, accession no. NP_009895.1; Grx2, accession no. NP_010801.1; Grx3, accession no. NP_010383.1; Grx4, accession no. NP_01101.1; Grx5, accession no. NP_015266.1; Grx6, accession no. NP_010274.1; Grx7, accession no. NP_009570.1; Grx8, accession no. NP_013468.1); glutathione reductase Glr1 (accession no. NP_015234.1); and Jac1 (accession no. NP_011497.1), including variants of Jac1 that have been suitably altered to reduce or prevent its normal mitochondrial localization; Hsp10, Hsp60, GroEL, and GroES and homologs or variants thereof.

[0024] In some embodiments, the recombinant microorganisms may further comprise one or more genes encoding an iron-sulfur cluster assembly protein. In one embodiment, the iron-sulfur cluster assembly protein encoding genes may be derived from prokaryotic organisms. In one embodiment, the iron-sulfur cluster assembly protein encoding genes are derived from a bacterial organism, including, but not limited to *Escherichia coli*, *L. lactis*, *Helicobacter pylori*, and *Entamoeba histolytica*. In specific embodiments, the bacterially derived iron-sulfur cluster assembly protein encoding genes are selected from the group consisting of *cyaY*, *iscS*, *iscU*, *iscA*, *hscB*, *hscA*, *fdx*, *isuX*, *sufA*, *sufB*, *sufC*, *sufD*, *sufS*, *sufE*, *apbC*, and homologs or variants thereof.

[0025] In another embodiment, the iron-sulfur cluster assembly protein encoding genes may be derived from eukaryotic organisms, including, but not limited to yeasts

and plants. In one embodiment, the iron-sulfur cluster protein encoding genes are derived from a yeast organism, including, but not limited to *S. cerevisiae*. In specific embodiments, the yeast derived genes encoding iron-sulfur cluster assembly proteins are selected from the group consisting of Cfd1 (accession no. NP_012263.1), Nbp35 (accession no. NP_011424.1), Nar1 (accession no. NP_014159.1), Cia1 (accession no. NP_010553.1), and homologs or variants thereof. In a further embodiment, the iron-sulfur cluster assembly protein encoding genes may be derived from plant nuclear genes which encode proteins translocated to chloroplast or plant genes found in the chloroplast genome itself.

[0026] In some embodiments, one or more genes encoding an iron-sulfur cluster assembly protein may be mutated or modified to remove a signal peptide, whereby localization of the product of said one or more genes to the mitochondria or other subcellular compartment is prevented. In certain embodiments, it may be preferable to overexpress one or more genes encoding an iron-sulfur cluster assembly protein.

[0027] In certain embodiments described herein, it may be desirable to reduce or eliminate the activity and/or proteins levels of one or more iron-sulfur cluster containing cytosolic proteins. In a specific embodiment, the iron-sulfur cluster containing cytosolic protein is 3-isoprpylmalate dehydratase (Leu1p). In one embodiment, the recombinant microorganism comprises a mutation in the *LEU1* gene resulting in the reduction of Leu1p protein levels. In another embodiment, the recombinant microorganism comprises a partial deletion in the *LEU1* gene resulting in the reduction of Leu1p protein levels. In another embodiment, the recombinant microorganism comprises a complete deletion in the *LEU1* gene resulting in the reduction of Leu1p protein levels. In another embodiment, the recombinant microorganism comprises a modification of the regulatory region associated with the *LEU1* gene resulting in the reduction of Leu1p protein levels. In yet another embodiment, the recombinant microorganism comprises a modification of a transcriptional regulator for the *LEU1* gene resulting in the reduction of Leu1p protein levels.

[0028] In certain embodiments described herein, it may be desirable to increase the levels of iron within the yeast cytosol and mitochondria, such that this iron is more available for the production of iron-sulfur cluster-containing proteins in the cytosol. Thus, in certain embodiments, the recombinant microorganism may further been engineered to overexpress one or more genes selected from the group consisting of *AFT1*, *AFT2*, *GRX3*, and *GRX4*, or homologs thereof. In alternative embodiments, the microorganism may be engineered to delete and/or attenuate one or more genes selected from the group consisting of *GRX3* and *GRX4*, or homologs thereof.

[0029] In various embodiments described herein, it may be desirable to reduce the concentration of reactive oxygen species (ROS) in said cytosol, as DHAD enzymes may be susceptible to inactivation by ROS. Thus, the recombinant microorganisms of the present invention may further be engineered to express one or more proteins in the cytosol that reduce the concentration of reactive oxygen species (ROS) in said cytosol. The proteins to be expressed in the cytosol for reducing the concentration of reactive oxygen species in the cytosol may be selected

from catalases, superoxide dismutases, metallothioneins, and methionine sulphoxide reductases. In a specific embodiment, said catalase may be encoded by one of more of the genes selected from the group consisting of the *E. coli* genes *katG* and *katE*, the *S. cerevisiae* genes *CTT1* and *CTA1*, or homologs thereof. In another specific embodiment, said superoxide dismutase is encoded by one of more of the genes selected from the group consisting of the *E. coli* genes *sodA*, *sodB*, *sodC*, the *S. cerevisiae* genes *SOD1* and *SOD2*, or homologs thereof. In another specific embodiment, said metallothionein is encoded by one of more of the genes selected from the group consisting of the *S. cerevisiae* *CUP1-1* and *CUP1-2* genes or homologs thereof. In another specific embodiment, said metallothionein is encoded by one or more genes selected from the group consisting of the Mycobacterium tuberculosis *MymT* gene and the Synechococcus PCC 7942 *SmtA* gene or homologs thereof. In another specific embodiment, said methionine sulphoxide reductase is encoded by one or more genes selected from the group consisting of the *S. cerevisiae* genes *MXR1* and *MXR2*, or homologs thereof.

[0030] In some embodiments, it may be desirable to increase the level of available glutathione in the cytosol, which is essential for FeS cluster biogenesis. Thus, the recombinant microorganisms of the present invention may further be engineered to express one or more enzymes that increase the level of available glutathione in the cytosol. The proteins to be expressed to increase the level of available glutathione in the cytosol can be selected from glutaredoxin, glutathione reductase, and glutathione synthase. In a specific embodiment, said glutaredoxin is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GRX2*, *GRX4*, *GRX6*, and *GRX7*, or homologs thereof. In another specific embodiment, said glutathione reductase is encoded by the *S. cerevisiae* genes *GLR1* or homologs thereof. In another specific embodiment, said glutathione synthase is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GSH1* and *GSH2*, or homologs thereof. In some embodiments, two enzymes are expressed in and targeted to the cytosol of yeast to increase the level of available glutathione in the cytosol. In one embodiment, the enzymes are enzymes are γ-glutamyl cysteine synthase and glutathione synthase. In a specific embodiment, said glutathione synthase is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GSH1* and *GSH2*, or homologs thereof.

[0031] In some embodiments, it may be desirable to overexpress one or more cytosolic functional components of the thioredoxin system, as overexpression of the essential cytosolic functional components of the thioredoxin system is can increase the amount of bioavailable cytosolic thioredoxin, resulting in a significant increase in cellular redox buffering potential and concomitant increase in stable, active cytosolic FeS clusters and DHAD activity. In one embodiment, the functional components of the thioredoxin system may be selected from a thioredoxin and a thioredoxin reductase. In a specific embodiment, said thioredoxin is encoded by the *S. cerevisiae* *TRX1* and *TRX2* genes or homologs thereof. In another specific embodiment, said thioredoxin reductase is encoded by *S. cerevisiae* *TRR1* gene or homologs thereof. In additional embodiments, the recombinant microorganism may

further be engineered to overexpress the mitochondrial thioredoxin system. In one embodiment, the mitochondrial thioredoxin system is comprised of the mitochondrial thioredoxin and mitochondrial thioredoxin reductase. In a specific embodiment, said mitochondrial thioredoxin is encoded by the *S. cerevisiae TRX3* gene or homologs thereof. In another specific embodiment, said mitochondrial thioredoxin reductase is encoded by the *S. cerevisiae TRR2* gene or homologs thereof.

[0032] In various embodiments described herein, it may be desirable to engineer the recombinant microorganism to overexpress one or more mitochondrial export proteins. In a specific embodiment, said mitochondrial export protein may be selected from the group consisting of the *S. cerevisiae ATM1*, the *S. cerevisiae ERV1*, and the *S. cerevisiae BAT1*, or homologs thereof.

[0033] In addition, the present invention provides recombinant microorganisms that have further been engineered to increase the inner mitochondrial membrane electrical potential, $\Delta\Psi_M$. In one embodiment, this is accomplished via overexpression of an ATP/ADP carrier protein, wherein said overexpression increases ATP^{4-} import into the mitochondrial matrix in exchange for ADP^{3-}. In a specific embodiment, said ATP/ADP carrier protein is encoded by the *S. cerevisiae AAC1, AAC2*, and/or *AAC3* genes or homologs thereof. In another embodiment, the inner mitochondrial membrane electrical potential, $\Delta\Psi_M$ is increased via a mutation in the mitochondrial ATP synthase complex that increases ATP hydrolysis activity. In a specific embodiment, said mutation is an ATP1-111 suppressor mutation or a corresponding mutation in a homologous protein.

[0034] In various embodiments described herein, it may further be desirable to engineer the recombinant microorganism to express one or more enzymes in the cytosol that reduce the concentration of reactive nitrogen species (RNS) and/or nitric oxide (NO) in said cytosol. In one embodiment, said one or more enzymes are selected from the group consisting of nitric oxide reductases and glutathione-S-nitrosothiol reductase. In a specific embodiment, said nitric oxide reductase is encoded by one of more of the genes selected from the group consisting of the *E. coli* gene *norV* and the *Fusarium oxysporum* gene *P-450dNIR*, or homologs thereof. In another specific embodiment, said glutathione-S-nitrosothiol reductase is encoded by the *S. cerevisiae* gene *SFA1* or homologs thereof. In one embodiment, said glutathione-S-nitrosothiol reductase gene *SFA1* is overexpressed. In another specific embodiment, said one or more enzymes is encoded by a gene selected from the group consisting of the *E. coli* gene *ytfE*, the *Staphylococcus aureus* gene *scdA*, and *Neisseria gonorrhoeae* gene *dnrN*, or homologs thereof.

[0035] Also provided herein are recombinant microorganisms that demonstrate increased the levels of sulfur-containing compounds within yeast cells, including the amino acid cysteine, such that this sulfur is more available for the production of iron-sulfur cluster-containing proteins in the yeast cytosol. In one embodiment, the recombinant microorganism has been engineered to overexpress one or more of the genes selected from the *S. cerevisiae* genes *MET1, MET2, MET3, MET5, MET8, MET10, MET14, MET16, MET17, HOM2, HOM3, HOM6, CYS3, CYS4, SUL1*, and *SUL2*, or homologs thereof. The recombinant microorganism may additionally or optionally also overexpress one or more of the genes selected from the *S. cerevisiae*

genes *YCT1*, *MUP1*, *GAP1*, *AGP1*, *GNP1*, *BAP1*, *BAP2*, *TAT1*, and *TAT2*, or homologs thereof.

[0036] In various embodiments described herein, the recombinant microorganism may exhibit at least about 5 percent greater dihydroxyacid dehydratase (DHAD) activity in the cytosol as compared to the parental microorganism. In another embodiment, the recombinant microorganism may exhibit at least about 10 percent, at least about 15 percent, about least about 20 percent, at least about 25 percent, at least about 30 percent, at least about 35 percent, at least about 40 percent, at least about 45 percent, at least about 50 percent, at least about 55 percent, at least about 60 percent, at least about 65 percent, at least about 70 percent, at least about 75 percent, at least about 80 percent, at least about 100 percent, at least about 200 percent, or at least about 500 percent greater dihydroxyacid dehydratase (DHAD) activity in the cytosol as compared to the parental microorganism.

[0037] In certain embodiments described herein, it may be desirable to further overexpress an additional enzyme that converts 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol. In a specific embodiment, the enzyme may be selected from the group consisting of 3-isopropylmalate (Leu1p) and imidazoleglycerol-phosphate dehydrogenase (His3p).

[0038] In various embodiments described herein, the recombinant microorganisms may be further engineered to express an isobutanol producing metabolic pathway comprising at least one exogenous gene that catalyzes a step in the conversion of pyruvate to isobutanol. In one embodiment, the recombinant microorganism may be engineered to express an isobutanol producing metabolic pathway comprising at least two exogenous genes. In another embodiment, the recombinant microorganism may be engineered to express an isobutanol producing metabolic pathway comprising at least three exogenous genes. In another embodiment, the recombinant microorganism may be engineered to express an isobutanol producing metabolic pathway comprising at least four exogenous genes. In another embodiment, the recombinant microorganism may be engineered to express an isobutanol producing metabolic pathway comprising five exogenous genes.

[0039] In various embodiments described herein, the isobutanol pathway enzyme(s) is/are selected from the group consisting of acetolactate synthase (ALS), ketol-acid reductoisomerase (KARI), dihydroxyacid dehydratase (DHAD), 2-keto-acid decarboxylase (KIVD), and isobutyraldehyde dehydrogenase (IDH). In a preferred embodiment, said dihydroxyacid dehydratase (DHAD) is a cytosolically active (DHAD) enzyme.

[0040] In various embodiments described herein, the recombinant microorganisms may be engineered to express native genes that catalyze a step in the conversion of pyruvate to isobutanol. In one embodiment, the recombinant microorganism is engineered to increase the activity of a native metabolic pathway gene for conversion of pyruvate to isobutanol. In another embodiment, the recombinant microorganism is further engineered to include at least one enzyme encoded by a heterologous gene and at least one enzyme encoded by a native gene. In yet another embodiment, the recombinant microorganism comprises a

reduction in the activity of a native metabolic pathway as compared to a parental microorganism.

[0041] In various embodiments described herein, one or more of the enzymes catalyzing the conversion of pyruvate to isobutanol is/are localized in the cytosol. In a preferred embodiment, the enzyme is dihydroxyacid dehydratase (DHAD).

[0042] In some embodiments, the present invention provides recombinant microorganisms that have been engineered to express a heterologous metabolic pathway for conversion of pyruvate to isobutanol. In another embodiment, the recombinant microorganism further comprises a pathway for the fermentation of isobutanol from a pentose sugar. In one embodiment, the pentose sugar is xylose. In one embodiment, the recombinant microorganism is engineered to express a functional xylose isomerase (XI). In another embodiment, the recombinant microorganism further comprises a deletion or disruption of a native gene encoding for an enzyme that catalyzes the conversion of xylose to xylitol. In one embodiment, the native gene is xylose reductase (XR). In another embodiment, the native gene is xylitol dehydrogenase (XDH). In yet another embodiment, both native genes are deleted or disrupted. In yet another embodiment, the recombinant microorganism is engineered to express a xylulose kinase enzyme.

[0043] In another aspect, the present invention provides a recombinant microorganism engineered to include reduced pyruvate decarboxylase (PDC) activity as compared to a parental microorganism. In one embodiment, PDC activity is eliminated. PDC catalyzes the decarboxylation of pyruvate to acetaldehyde, which is reduced to ethanol by alcohol dehydrogenases via the oxidation of NADH to NAD+. In one embodiment, the recombinant microorganism includes a mutation in at least one *PDC* gene resulting in a reduction of PDC activity of a polypeptide encoded by said gene. In another embodiment, the recombinant microorganism includes a partial deletion of a *PDC* gene resulting in a reduction of PDC activity of a polypeptide encoded by said gene. In another embodiment, the recombinant microorganism comprises a complete deletion of a *PDC* gene resulting in a reduction of PDC activity of a polypeptide encoded by said gene. In yet another embodiment, the recombinant microorganism includes a modification of the regulatory region associated with at least one *PDC* gene resulting in a reduction of PDC activity of a polypeptide encoded by said gene. In yet another embodiment, the recombinant microorganism comprises a modification of the transcriptional regulator resulting in a reduction of *PDC* gene transcription. In yet another embodiment, the recombinant microorganism comprises mutations in all *PDC* genes resulting in a reduction of PDC activity of the polypeptides encoded by said genes.

[0044] In another aspect, the present invention provides a recombinant microorganism engineered to include reduced glycerol-3-phosphate dehydrogenase (GPD) activity as compared to a parental microorganism. In one embodiment, GPD activity is eliminated. GPD catalyzes the reduction of dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate (G3P) via the oxidation of NADH to NAD$^+$. Glycerol is produced from G3P by Glycerol-3-phosphatase (GPP). In one embodiment, the recombinant microorganism includes a mutation in at least one *GPD* gene resulting in a reduction of GPD activity of a polypeptide encoded by said gene. In another

embodiment, the recombinant microorganism includes a partial deletion of a *GPD* gene resulting in a reduction of GPD activity of a polypeptide encoded by the gene. In another embodiment, the recombinant microorganism comprises a complete deletion of a *GPD* gene resulting in a reduction of GPD activity of a polypeptide encoded by the gene. In yet another embodiment, the recombinant microorganism includes a modification of the regulatory region associated with at least one *GPD* gene resulting in a reduction of GPD activity of a polypeptide encoded by said gene. In yet another embodiment, the recombinant microorganism comprises a modification of the transcriptional regulator resulting in a reduction of *GPD* gene transcription. In yet another embodiment, the recombinant microorganism comprises mutations in all *GPD* genes resulting in a reduction of GPD activity of a polypeptide encoded by the gene.

[0045] In various embodiments described herein, the recombinant microorganisms may be microorganisms of the Saccharomyces clade, *Saccharomyces sensu stricto* microorganisms, Crabtree-negative yeast microorganisms, Crabtree-positive yeast microorganisms, post-WGD (whole genome duplication) yeast microorganisms, pre-WGD (whole genome duplication) yeast microorganisms, and non-fermenting yeast microorganisms.

[0046] In some embodiments, the recombinant microorganisms may be yeast recombinant microorganisms of the Saccharomyces clade.

[0047] In some embodiments, the recombinant microorganisms may be *Saccharomyces sensu stricto* microorganisms. In one embodiment, the *Saccharomyces sensu stricto* is selected from the group consisting of *S. cerevisiae, S. kudriavzevii, S. mikatae, S. bayanus, S. uvarum. S. carocanis* and hybrids thereof.

[0048] In some embodiments, the recombinant microorganisms may be Crabtree-negative recombinant yeast microorganisms. In one embodiment, the Crabtree-negative yeast microorganism is classified into a genera selected from the group consisting of *Kluyveromyces, Pichia, Hansenula,* or *Candida*. In additional embodiments, the Crabtree-negative yeast microorganism is selected from *Kluyveromyces lactis, Kluyveromyces marxianus, Pichia anomala, Pichia stipitis, Hansenula anomala, Candida utilis* and *Kluyveromyces waltii*.

[0049] In some embodiments, the recombinant microorganisms may be Crabtree-positive recombinant yeast microorganisms. In one embodiment, the Crabtree-positive yeast microorganism is classified into a genera selected from the group consisting of *Saccharomyces, Kluyveromyces, Zygosaccharomyces, Debaryomyces, Candida, Pichia* and *Schizosaccharomyces*. In additional embodiments, the Crabtree-positive yeast microorganism is selected from the group consisting of *Saccharomyces cerevisiae, Saccharomyces uvarum, Saccharomyces bayanus, Saccharomyces paradoxus, Saccharomyces castelli, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Candida glabrata, Z. bailli, Z. rouxii, Debaryomyces hansenii, Pichia pastorius, Schizosaccharomyces pombe,* and *Saccharomyces uvarum.*

[0050] In some embodiments, the recombinant microorganisms may be post-WGD (whole genome duplication) yeast recombinant microorganisms. In one embodiment, the post-WGD yeast recombinant microorganism is classified into a

genera selected from the group consisting of *Saccharomyces* or *Candida*. In additional embodiments, the post-WGD yeast is selected from the group consisting of *Saccharomyces cerevisiae, Saccharomyces uvarum, Saccharomyces bayanus, Saccharomyces paradoxus, Saccharomyces castelli,* and *Candida glabrata.*

[0051]   In some embodiments, the recombinant microorganisms may be pre-WGD (whole genome duplication) yeast recombinant microorganisms. In one embodiment, the pre-WGD yeast recombinant microorganism is classified into a genera selected from the group consisting of *Saccharomyces, Kluyveromyces, Candida, Pichia, Debaryomyces, Hansenula, Pachysolen, Yarrowia* and *Schizosaccharomyces.* In additional embodiments, the pre-WGD yeast is selected from the group consisting of *Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces marxianus, Kluyveromyces waltii, Kluyveromyces lactis, Candida tropicalis, Pichia pastoris, Pichia anomala, Pichia stipitis, Debaryomyces hansenii, Hansenula anomala, Pachysolen tannophilis, Yarrowia lipolytica,* and *Schizosaccharomyces pombe.*

[0052]   In some embodiments, the recombinant microorganisms may be microorganisms that are non-fermenting yeast microorganisms, including, but not limited to those, classified into a genera selected from the group consisting of *Tricosporon, Rhodotorula,* or *Myxozyma.*

[0053]   In another aspect, the present invention provides methods of producing isobutanol using a recombinant microorganism of the invention. In one embodiment, the method includes cultivating the recombinant microorganism in a culture medium containing a feedstock providing the carbon source until a recoverable quantity of the isobutanol is produced and optionally, recovering the isobutanol. In one embodiment, the microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 5 percent theoretical. In another embodiment, the microorganism is selected to produce isobutanol at a yield of at least about 10 percent, at least about 15 percent, about least about 20 percent, at least about 25 percent, at least about 30 percent, at least about 35 percent, at least about 40 percent, at least about 45 percent, at least about 50 percent, at least about 55 percent, at least about 60 percent, at least about 65 percent, at least about 70 percent, at least about 75 percent, or at least about 80 percent theoretical.

[0054]   In one embodiment, the microorganism is selected to produce isobutanol from a carbon source at a specific productivity of at least about 0.7 mg/L/hr per OD. In another embodiment, the microorganism is selected produce isobutanol from a carbon source at a specific productivity of at least about 1 mg/L/hr per OD, at least about 10 mg/L/hr per OD, at least about 50 mg/L/hr per OD, at least about 100 mg/L/hr per OD, at least about 250 mg/L/hr per OD, or at least about 500 g/L/hr per OD.

BRIEF DESCRIPTION OF DRAWINGS

[0055]   Illustrative embodiments of the invention are illustrated in the drawings, in which:

[0056]   Figure 1 illustrates an exemplary embodiment of an isobutanol pathway.

DETAILED DESCRIPTION

[0057] As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a polynucleotide" includes a plurality of such polynucleotides and reference to "the microorganism" includes reference to one or more microorganisms, and so forth.

[0058] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice of the disclosed methods and compositions, the exemplary methods, devices and materials are described herein.

[0059] Any publications discussed above and throughout the text are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior disclosure.

[0060] The term "microorganism" includes prokaryotic and eukaryotic microbial species from the Domains Archaea, Bacteria and Eucarya, the latter including yeast and filamentous fungi, protozoa, algae, or higher Protista. The terms "microbial cells" and "microbes" are used interchangeably with the term microorganism.

[0061] The term "genus" is defined as a taxonomic group of related species according to the Taxonomic Outline of Bacteria and Archaea (Garrity, G.M., Lilburn, T.G., Cole, J.R., Harrison, S.H., Euzeby, J., and Tindall, B.J. (2007) The Taxonomic Outline of Bacteria and Archaea. TOBA Release 7.7, March 2007. Michigan State University Board of Trustees. [http://www.taxonomicoutline.org/]).

[0062] The term "species" is defined as a collection of closely related organisms with greater than 97% 16S ribosomal RNA sequence homology and greater than 70% genomic hybridization and sufficiently different from all other organisms so as to be recognized as a distinct unit.

[0063] The terms "recombinant microorganism," "modified microorganism'" and "recombinant host cell" are used interchangeably herein and refer to microorganisms that have been genetically modified to express or over-express endogenous polynucleotides, or to express heterologous polynucleotides, such as those included in a vector, or which have an alteration in expression of an endogenous gene. By "alteration" it is meant that the expression of the gene, or level of a RNA molecule or equivalent RNA molecules encoding one or more polypeptides or polypeptide subunits, or activity of one or more polypeptides or polypeptide subunits is up regulated or down regulated, such that expression, level, or activity is greater than or less than that observed in the absence of the alteration. For example, the term "alter" can mean "inhibit," but the use of the word "alter" is not limited to this definition.

[0064] The term "expression" with respect to a gene sequence refers to transcription of the gene and, as appropriate, translation of the resulting mRNA transcript to a protein. Thus, as will be clear from the context, expression of a protein results from transcription and translation of the open reading frame

sequence. The level of expression of a desired product in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell, or the amount of the desired product encoded by the selected sequence. For example, mRNA transcribed from a selected sequence can be quantitated by qRT-PCR or by Northern hybridization (see Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989)). Protein encoded by a selected sequence can be quantitated by various methods, e.g., by ELISA, by assaying for the biological activity of the protein, or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay, using antibodies that recognize and bind the protein. See Sambrook et al., 1989, supra. The polynucleotide generally encodes a target enzyme involved in a metabolic pathway for producing a desired metabolite. It is understood that the terms "recombinant microorganism" and "recombinant host cell" refer not only to the particular recombinant microorganism but to the progeny or potential progeny of such a microorganism. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

[0065] The term "wild-type microorganism" describes a cell that occurs in nature, i.e. a cell that has not been genetically modified. A wild-type microorganism can be genetically modified to express or overexpress a first target enzyme. This microorganism can act as a parental microorganism in the generation of a microorganism modified to express or overexpress a second target enzyme. In turn, the microorganism modified to express or overexpress a first and a second target enzyme can be modified to express or overexpress a third target enzyme.

[0066] Accordingly, a "parental microorganism" functions as a reference cell for successive genetic modification events. Each modification event can be accomplished by introducing a nucleic acid molecule in to the reference cell. The introduction facilitates the expression or overexpression of a target enzyme. It is understood that the term "facilitates" encompasses the activation of endogenous polynucleotides encoding a target enzyme through genetic modification of e.g., a promoter sequence in a parental microorganism. It is further understood that the term "facilitates" encompasses the introduction of heterologous polynucleotides encoding a target enzyme in to a parental microorganism

[0067] The term "engineer" refers to any manipulation of a microorganism that results in a detectable change in the microorganism, wherein the manipulation includes but is not limited to inserting a polynucleotide and/or polypeptide heterologous to the microorganism and mutating a polynucleotide and/or polypeptide native to the microorganism.

[0068] The term "mutation" as used herein indicates any modification of a nucleic acid and/or polypeptide which results in an altered nucleic acid or polypeptide. Mutations include, for example, point mutations, deletions, or insertions of single or multiple residues in a polynucleotide, which includes alterations arising within a protein-encoding region of a gene as well as alterations in regions outside of a protein-encoding sequence, such as, but not limited to, regulatory or promoter

sequences. A genetic alteration may be a mutation of any type. For instance, the mutation may constitute a point mutation, a frame-shift mutation, an insertion, or a deletion of part or all of a gene. In addition, in some embodiments of the modified microorganism, a portion of the microorganism genome has been replaced with a heterologous polynucleotide. In some embodiments, the mutations are naturally-occurring. In other embodiments, the mutations are the results of artificial selection pressure. In still other embodiments, the mutations in the microorganism genome are the result of genetic engineering.

[0069] The term "biosynthetic pathway", also referred to as "metabolic pathway", refers to a set of anabolic or catabolic biochemical reactions for converting one chemical species into another. Gene products belong to the same "metabolic pathway" if they, in parallel or in series, act on the same substrate, produce the same product, or act on or produce a metabolic intermediate (*i.e.*, metabolite) between the same substrate and metabolite end product.

[0070] The term "heterologous" as used herein with reference to molecules and in particular enzymes and polynucleotides, indicates molecules that are expressed in an organism other than the organism from which they originated or are found in nature, independently of the level of expression that can be lower, equal or higher than the level of expression of the molecule in the native microorganism.

[0071] On the other hand, the term "native" or "endogenous" as used herein with reference to molecules, and in particular enzymes and polynucleotides, indicates molecules that are expressed in the organism in which they originated or are found in nature, independently of the level of expression that can be lower equal or higher than the level of expression of the molecule in the native microorganism. It is understood that expression of native enzymes or polynucleotides may be modified in recombinant microorganisms.

[0072] The term "feedstock" is defined as a raw material or mixture of raw materials supplied to a microorganism or fermentation process from which other products can be made. For example, a carbon source, such as biomass or the carbon compounds derived from biomass are a feedstock for a microorganism that produces a biofuel in a fermentation process. However, a feedstock may contain nutrients other than a carbon source.

[0073] The term "substrate" or "suitable substrate" refers to any substance or compound that is converted or meant to be converted into another compound by the action of an enzyme. The term includes not only a single compound, but also combinations of compounds, such as solutions, mixtures and other materials which contain at least one substrate, or derivatives thereof. Further, the term "substrate" encompasses not only compounds that provide a carbon source suitable for use as a starting material, such as any biomass derived sugar, but also intermediate and end product metabolites used in a pathway associated with a recombinant microorganism as described herein.

[0074] The term "C2-compound" as used as a carbon source for engineered yeast microorganisms with mutations in all pyruvate decarboxylase (PDC) genes resulting in a reduction of pyruvate decarboxylase activity of said genes refers to organic compounds comprised of two carbon atoms, including but not limited to

ethanol and acetate.

[0075]   The term "fermentation" or "fermentation process" is defined as a process in which a microorganism is cultivated in a culture medium containing raw materials, such as feedstock and nutrients, wherein the microorganism converts raw materials, such as a feedstock, into products.

[0076]   The term "volumetric productivity" or "production rate" is defined as the amount of product formed per volume of medium per unit of time. Volumetric productivity is reported in gram per liter per hour (g/L/h).

[0077]   The term "specific productivity" or "specific production rate" is defined as the amount of product formed per volume of medium per unit of time per amount of cells. Volumetric productivity is reported in gram or milligram per liter per hour per OD (g/L/h/OD).

[0078]   The term "yield" is defined as the amount of product obtained per unit weight of raw material and may be expressed as g product per g substrate (g/g). Yield may be expressed as a percentage of the theoretical yield. "Theoretical yield" is defined as the maximum amount of product that can be generated per a given amount of substrate as dictated by the stoichiometry of the metabolic pathway used to make the product. For example, the theoretical yield for one typical conversion of glucose to isobutanol is 0.41 g/g. As such, a yield of isobutanol from glucose of 0.39 g/g would be expressed as 95% of theoretical or 95% theoretical yield.

[0079]   The term "titer" is defined as the strength of a solution or the concentration of a substance in solution. For example, the titer of a biofuel in a fermentation broth is described as g of biofuel in solution per liter of fermentation broth (g/L).

[0080]   "Aerobic conditions" are defined as conditions under which the oxygen concentration in the fermentation medium is sufficiently high for an aerobic or facultative anaerobic microorganism to use as a terminal electron acceptor.

[0081]   In contrast, "anaerobic conditions" are defined as conditions under which the oxygen concentration in the fermentation medium is too low for the microorganism to use as a terminal electron acceptor. Anaerobic conditions may be achieved by sparging a fermentation medium with an inert gas such as nitrogen until oxygen is no longer available to the microorganism as a terminal electron acceptor. Alternatively, anaerobic conditions may be achieved by the microorganism consuming the available oxygen of the fermentation until oxygen is unavailable to the microorganism as a terminal electron acceptor.

[0082]   "Aerobic metabolism" refers to a biochemical process in which oxygen is used as a terminal electron acceptor to make energy, typically in the form of ATP, from carbohydrates. Aerobic metabolism occurs e.g. via glycolysis and the TCA cycle, wherein a single glucose molecule is metabolized completely into carbon dioxide in the presence of oxygen.

[0083]   In contrast, "anaerobic metabolism" refers to a biochemical process in which oxygen is not the final acceptor of electrons contained in NADH. Anaerobic metabolism can be divided into anaerobic respiration, in which compounds other than oxygen serve as the terminal electron acceptor, and substrate level phosphorylation, in which the electrons from NADH are utilized to generate a reduced product via a "fermentative pathway."

[0084] In "fermentative pathways", NAD(P)H donates its electrons to a molecule produced by the same metabolic pathway that produced the electrons carried in NAD(P)H. For example, in one of the fermentative pathways of certain yeast strains, NAD(P)H generated through glycolysis transfers its electrons to pyruvate, yielding ethanol. Fermentative pathways are usually active under anaerobic conditions but may also occur under aerobic conditions, under conditions where NADH is not fully oxidized via the respiratory chain. For example, above certain glucose concentrations, Crabtree positive yeasts produce large amounts of ethanol under aerobic conditions.

[0085] The term "byproduct" means an undesired product related to the production of a biofuel or biofuel precursor. Byproducts are generally disposed as waste, adding cost to a production process.

[0086] The term "non-fermenting yeast" is a yeast species that fails to demonstrate an anaerobic metabolism in which the electrons from NADH are utilized to generate a reduced product via a fermentative pathway such as the production of ethanol and CO_2 from glucose. Non-fermentative yeast can be identified by the "Durham Tube Test" (J.A. Barnett, R.W. Payne, and D. Yarrow. 2000. Yeasts Characteristics and Identification. 3rd edition. p. 28-29. Cambridge University Press, Cambridge, UK.) or by monitoring the production of fermentation productions such as ethanol and CO_2.

[0087] The term "polynucleotide" is used herein interchangeably with the term "nucleic acid" and refers to an organic polymer composed of two or more monomers including nucleotides, nucleosides or analogs thereof, including but not limited to single stranded or double stranded, sense or antisense deoxyribonucleic acid (DNA) of any length and, where appropriate, single stranded or double stranded, sense or antisense ribonucleic acid (RNA) of any length, including siRNA. The term "nucleotide" refers to any of several compounds that consist of a ribose or deoxyribose sugar joined to a purine or a pyrimidine base and to a phosphate group, and that are the basic structural units of nucleic acids. The term "nucleoside" refers to a compound (as guanosine or adenosine) that consists of a purine or pyrimidine base combined with deoxyribose or ribose and is found especially in nucleic acids. The term "nucleotide analog" or "nucleoside analog" refers, respectively, to a nucleotide or nucleoside in which one or more individual atoms have been replaced with a different atom or with a different functional group. Accordingly, the term polynucleotide includes nucleic acids of any length, DNA, RNA, analogs and fragments thereof. A polynucleotide of three or more nucleotides is also called nucleotidic oligomer or oligonucleotide.

[0088] It is understood that the polynucleotides described herein include "genes" and that the nucleic acid molecules described herein include "vectors" or "plasmids." Accordingly, the term "gene", also called a "structural gene" refers to a polynucleotide that codes for a particular sequence of amino acids, which comprise all or part of one or more proteins or enzymes, and may include regulatory (non-transcribed) DNA sequences, such as promoter sequences, which determine for example the conditions under which the gene is expressed. The transcribed region

of the gene may include untranslated regions, including introns, 5'-untranslated region (UTR), and 3'-UTR, as well as the coding sequence.

[0089]   The term "operon" refers to two or more genes which are transcribed as a single transcriptional unit from a common promoter. In some embodiments, the genes comprising the operon are contiguous genes.   It is understood that transcription of an entire operon can be modified (i.e., increased, decreased, or eliminated) by modifying the common promoter.   Alternatively, any gene or combination of genes in an operon can be modified to alter the function or activity of the encoded polypeptide. The modification can result in an increase in the activity of the encoded polypeptide. Further, the modification can impart new activities on the encoded polypeptide. Exemplary new activities include the use of alternative substrates and/or the ability to function in alternative environmental conditions.

[0090]   A "vector" is any means by which a nucleic acid can be propagated and/or transferred between organisms, cells, or cellular components. Vectors include viruses, bacteriophage, pro-viruses, plasmids, phagemids, transposons, and artificial chromosomes such as YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), and PLACs (plant artificial chromosomes), and the like, that are "episomes," that is, that replicate autonomously or can integrate into a chromosome of a host cell. A vector can also be a naked RNA polynucleotide, a naked DNA polynucleotide, a polynucleotide composed of both DNA and RNA within the same strand, a poly-lysine -conjugated DNA or RNA, a peptide-conjugated DNA or RNA, a liposome-conjugated DNA, or the like, that are not episomal in nature, or it can be an organism which comprises one or more of the above polynucleotide constructs such as an agrobacterium or a bacterium.

[0091]   "Transformation" refers to the process by which a vector is introduced into a host cell. Transformation (or transduction, or transfection), can be achieved by any one of a number of means including chemical transformation (e.g. lithium acetate transformation), electroporation, microinjection, biolistics (or particle bombardment-mediated delivery), or agrobacterium mediated transformation.

[0092]   The term "enzyme" as used herein refers to any substance that catalyzes or promotes one or more chemical or biochemical reactions, which usually includes enzymes totally or partially composed of a polypeptide, but can include enzymes composed of a different molecule including polynucleotides.

[0093]   The term "protein," "peptide," or "polypeptide" as used herein indicates an organic polymer composed of two or more amino acidic monomers and/or analogs thereof. As used herein, the term "amino acid" or "amino acidic monomer" refers to any natural and/or synthetic amino acids including glycine and both D or L optical isomers. The term "amino acid analog" refers to an amino acid in which one or more individual atoms have been replaced, either with a different atom, or with a different functional group. Accordingly, the term polypeptide includes amino acidic polymer of any length including full length proteins, and peptides as well as analogs and fragments thereof. A polypeptide of three or more amino acids is also called a protein oligomer or oligopeptide

[0094]   The term "homolog," used with respect to an original enzyme or gene of a first family or species, refers to distinct enzymes or genes of a second family or

species which are determined by functional, structural or genomic analyses to be an enzyme or gene of the second family or species which corresponds to the original enzyme or gene of the first family or species. Most often, homologs will have functional, structural or genomic similarities. Techniques are known by which homologs of an enzyme or gene can readily be cloned using genetic probes and PCR. Identity of cloned sequences as homolog can be confirmed using functional assays and/or by genomic mapping of the genes.

[0095] A protein has "homology" or is "homologous" to a second protein if the amino acid sequence encoded by a gene has a similar amino acid sequence to that of the second gene. Alternatively, a protein has homology to a second protein if the two proteins have "similar" amino acid sequences. (Thus, the term "homologous proteins" is defined to mean that the two proteins have similar amino acid sequences).

[0096] The term "analog" or "analogous" refers to nucleic acid or protein sequences or protein structures that are related to one another in function only and are not from common descent or do not share a common ancestral sequence. Analogs may differ in sequence but may share a similar structure, due to convergent evolution. For example, two enzymes are analogs or analogous if the enzymes catalyze the same reaction of conversion of a substrate to a product, are unrelated in sequence, and irrespective of whether the two enzymes are related in structure.

Cytosolically Active Dihydroxyacid Dehydratases (DHADs) and Recombinant Microorganisms Comprising the Same

[0097] The present inventors describe herein cytosolically active dihydroxyacid dehydratases (DHADs) and their use in the production of various beneficial metabolites, such as isobutanol and 2-methyl-1-butanol. Biosynthetic pathways for the production of isobutanol and 2-methyl-1-butanol, are described by Atsumi *et al.* (Atsumi *et al.*, 2008, *Nature* 451: 86-89). In these biosynthetic pathways, DHAD catalyzes the conversion of 2,3-dihydroxyisovalerate to 2-ketoisovalerate, and 2,3-dihydroxy-3-methylvalerate to 2-keto-3-methylvarate, respectively. Using a combination of genetic selection and biochemical analyses, the present inventors have identified a number of DHAD homologs that have activity in the cytosol. Cytosolic DHAD activity is a highly desirable characteristic, especially for the production of isobutanol since the ideal biocatalyst (*e.g.* recombinant microorganism) will have the entire isobutanol pathway functionally expressed in the same compartment (*e.g.* preferably in the cytosol). In addition, this localization allows the pathway to utilize pyruvate and NAD(P)H that is generated in the cytosol by glycolysis and/or the pentose phosphate pathway without the need for transfer of these metabolites to an alternative compartment (*i.e.* the mitochondria).

[0098] As used herein, the term "cytosolically active" or "active in the cytosol" means the enzyme exhibits enzymatic activity in the cytosol of a eukaryotic organism. Cytosolically active DHAD enzymes may further be additionally and/or independently characterized as DHAD enzymes that generally exhibit a specific cytosolic activity which is greater than the specific mitochondrial activity. In certain respects, "cytosolically active" DHAD enzymes of the present invention exhibit a ratio

of the specific activity of the mitochondrial fraction over the specific activity of the whole cell fraction of less than 1, as determined by the method disclosed in Example 4 herein.

[0099] As used herein, the term "cytosolically localized" or "cytosolic localization" means the enzyme is localized in the cytosol of a eukaryotic organism. Cytosolically localized DHAD enzymes may further be additionally and/or independently characterized as DHAD enzymes that exhibit a cytosolic protein level which is greater than the mitochondrial protein level.

<u>Identification of Cytosolically Active DHADs</u>

[00100] Among the many strategies for identifying cytosolically active DHADs, the present inventors performed multiway-protein alignments between several DHAD homologs. Using this analysis, a protein motif was identified that is unique to the subset of DHAD homologs exhibiting cytosolically activity. This protein motif, P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), was found in DHAD homologs demonstrating positive cytosolic activity. Therefore, in one embodiment, the present invention provides DHAD enzymes comprising the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol. DHAD enzymes harboring this sequence include those derived from *L. lactis*, *Grammella forsetii*, *Acidobacteria bacterium Ellin345*, *Saccharopolyspora erythraea*, *Yarrowia lipolytica*, *Francisella tularensis*, *Arabidopsis thaliana*, *Thermotoga petrophila*, and *Victivallis vadensis*. Also encompassed herein are DHAD enzymes that comprise a motif that is at least about 70% similar, at least about 80% similar, or at least about 90% similar to the motif shown in SEQ ID NO: 19.

[00101] As described herein, an even more specific version of this motif has been identified by the present inventors. Thus, in a further embodiment, the present invention provides DHAD enzymes comprising the amino acid sequence PIKXXGX(I/L)XIL (SEQ ID NO: 20), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol. DHAD enzymes harboring this sequence include those derived from *L. lactis*, *Grammella forsetii*, *Acidobacteria bacterium Ellin345*, *Yarrowia lipolytica*, *Francisella tularensis*, *Arabidopsis thaliana*, *Thermotoga petrophila*, and *Victivallis vadensis*. Also encompassed herein are DHAD enzymes that comprise a motif that is at least about 70% similar, at least about 80% similar, or at least about 90% similar to the motif shown in SEQ ID NO: 20.

[00102] As noted above, one such cytosolically active DHAD identified herein is exemplified by the *L. lactis* DHAD amino acid sequence of SEQ ID NO: 9 which is encoded by the *L. lactis ilvD* gene. As described herein, the present inventors have discovered that yeast strains expressing the cytosolically active *L. lactis ilvD* (DHAD) exhibit higher isobutanol production than yeast strains expressing the *S. cerevisiae* *ILV3* (DHAD), even when the *ILV3* from *S. cerevisiae* is truncated at its N-terminus to remove a putative mitochondrial targeting sequence. In addition to the use and identification of the cytosolically active DHAD homolog from *L. lactis*, the present

invention encompasses a number of different strategies for identifying DHAD enzymes that exhibit cytosolic activity and/or cytosolic localization, as well as methods for modifiying DHADs to increase their ability to exhibit cytosolic activity and/or cytosolic localization.

[00103] In various embodiments described herein, the DHAD enzymes may be derived from a prokaryotic organism. In one embodiment, the prokaryotic organism is a bacterial organism. In another embodiment, the bacterial organism is *Lactococcus lactis*. In a specific embodiment, the DHAD enzyme from *L. lactis* comprises the amino acid sequence of SEQ ID NO: 9. In other embodiments, the bacterial organisms are of the genus *Lactococcus*, *Grammella*, *Acidobacteria*, *Francisella*, *Thermotoga* and *Victivallis*.

[00104] In alternative embodiments described herein, the DHAD enzyme may be derived from a eukaryotic organism. In one embodiment, the eukaryotic organism is a fungal organism. As described herein, the present inventors have found that in general, an enzyme from a fungal source is more likely to show activity in yeast than a bacterial enzyme expressed in yeast. In addition, homologs that are normally expressed in the cytosol are desired, as a normally cytoplasmic enzyme is likely to show higher activity in the cytosol as compared to an enzyme that is relocalized to the cytosol from other organelles, such as the mitochondria. Fungal homologs of various isobutanol pathway enzymes, including DHAD, are often localized to the mitochondria. The present inventors have found that fungal homologs of DHAD that are cytosolically localized will generally be expected to exhibit higher activity in the cytosol of yeast than those of wild-type yeast strains. Thus, in one embodiment, the present invention provides fungal DHAD homologs that are cytosolically active and/or cytosolically localized.

[00105] In addition, the present inventors tested several different DHAD homologs using biochemical assays. Among the DHAD homologs showing superior cytosolic activity include the *S. cerevisiae ilv3ΔN* (e.g. the *S. cerevisiae* DHAD with N-terminal deletions), *the Lactococcus lactis DHAD*, *DHAD from Grammella, DHAD from Acidobacteria*, and the *Saccharopolyspora DHAD*.

[00106] As described herein, the *E. coli* IlvD protein comprises the sequence **VERSACPTCGSC** (SEQ ID NO. 121), which further comprises the 4Fe-4S cluster-binding motif **CXXCXXC** (SEQ ID NO. 122). The present inventors have observed the following pattern with respect to cytosolic activity in yeast: enzymes with either one of the following two-cysteine containing motifs exhibit activity in the yeast cytosol: (1) **CPGXGXC** (SEQ ID NO. 123), for example, CPGXG(A/S)C (SEQ ID NO. 124); and 2) **CXXXPGXGXC** (SEQ ID NO. 125). Accordingly, in one embodiment, the invention provides DHAD enzymes comprising the amino acid sequence **CPGXGXC** (SEQ ID NO 123), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the yeast cytosol. In another embodiment, the invention provides DHAD enzymes comprising the amino acid sequence CPGXG(A/S)C (SEQ ID NO 124), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the yeast cytosol. In yet another embodiment, the present invention provides DHAD enzymes comprising the amino

acid sequence CXXXPGXGXC (SEQ ID NO. 125), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the yeast cytosol. Also encompassed herein are DHAD enzymes that comprise a motif that is at least about 70% similar, at least about 80% similar, or at least about 90% similar to the motif shown in SEQ ID NOs: 123-125.

[00107] In various embodiments described herein, the DHAD enzyme may be a native protein. Also encompassed by the invention are modified (e.g. mutated) and natural variants of said native proteins. For example, said variant DHAD may be derived from a natural protein by an amino acid substitution. In additional embodiments, the DHAD enzyme may be a chimeric protein.

Removal and/or Modification of N-Terminal Mitochondrial Targeting Sequences

[00108] Among the mechanisms for cytosolic localization of DHAD enzymes described herein involves the removal and/or modification of N-terminal mitochondrial targeting sequences (MTS). Nuclear genome-encoded proteins destined to reside in the mitochondria often contain a Mitochondrial Targeting Sequence (MTS) which is often but not exclusively found at the N-terminal end of the protein. The MTS is recognized by a set of proteins collectively known as mitochondrial import machinery. Following recognition and import, the MTS is then physically cleaved off of the imported protein. In eukaryotes, homologs of two of the isobutanol pathway enzymes, ketol-acid reductoisomerase (KARI, e.g. S. cerevisiae ILV5) and dihydroxy acid dehydratase (DHAD, e.g. S. cerevisiae ILV3), are predicted to be mitochondrial, based upon the presence of an N-terminal MTS as well as several in vivo functional and mutational studies (see e.g., Omura, F., Applied Genetics and Molecular Biotechnology (2008), 78:503-513). As described herein, the present inventors have designed DHADs, whereby the predicted MTS is removed or modified. A computer algorithm for identifying the predicted MTS can be found at the MITOPROT website: http://mips.helmholtz-muenchen.de/cgi-bin/proj/medgen/mitofilter.

[00109] One example of an enzyme normally targeted to the mitochondria is the native S. cerevisiae DHAD. The native S. cerevisiae DHAD, encoded by ILV3, is involved in valine biosynthesis, and is thought to be targeted to the mitochondria by an N-terminal mitochondrial targeting sequence (MTS). However, for the purpose of isobutanol production, DHAD activity is required in the cytosol where the remainder of the pathway is expressed. To generate a cytosolic Ilv3, the present inventors performed a series of experiments in which amino acids were removed from the N-terminus of native S. cerevisiae ILV3 to disrupt the MTS. As described in Example 10, Ilv3ΔN19 and 23 exhibited cytosolic DHAD activity.

[00110] Accordingly, in one embodiment, the present invention provides a modified dihydroxyacid dehydratase (DHAD) enzyme having one or more amino acid deletions at the N-terminus. In various embodiments described herein, the DHAD enzyme may have at least about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more amino acid deletions at the N-terminus. In a preferred embodiment, the modified DHAD enzyme comprising one or more amino acid deletions at the N-terminus is derived from the S. cerevisiae Ilv3. As described herein, the advantage

of the *S. cerevisiae* Ilv3 is that it is a native yeast enzyme, which is generally expected to exhibit regulatory advantages over the use of a heterologous protein. In a specific embodiment, the modified DHAD has 19 amino acids deleted at the N-terminus of the *S. cerevisiae* Ilv3 and may be encoded by the nucleic acid sequence set forth in SEQ ID NO: 127. In yet another specific embodiment, the modified DHAD has 23 amino acids deleted at the N-terminus of the *S. cerevisiae* Ilv3 and may be encoded by the nucleic acid sequence set forth in SEQ ID NO: 128.

Peptide Tags to Augment Cytosolic Localization of DHADs

[00111] In additional embodiments described herein, the mitochondrially imported DHAD enzymes can be expressed as a chimeric fusion protein to augment cytosolic localization. In one embodiment, the DHAD enzyme is fused to a peptide tag, whereby said DHAD enzyme exhibits increased cytosolic localization and/or cytosolic DHAD activity in yeast as compared to the parental DHAD enzyme. In one embodiment, the DHAD enzyme is fused to a peptide tag following removal of the N-terminal Mitochondrial Targeting Sequence (MTS). In one embodiment, the peptide tag is non-cleavable. In a preferred embodiment, the peptide tag is fused at the N-terminus of the DHAD enzyme. Peptide tags useful in the present invention preferably have the following properties: (1) they do not significantly hinder the normal enzymatic function of the DHAD; (2) they fold in such as a way as to block recognition of an N-terminal MTS by the normal mitochondrial import machinery; (3) they promote the stable expression and/or folding of the DHAD it precedes; (4)they can be detected, for example, by Western blotting or SDS-PAGE plus Coomassie staining to facilitate analysis of the overexpressed chimeric protein.

[00112] Suitable peptide tags for use in the present invention include, but are not limited to, ubiquitin, ubiquitin-like (UBL) proteins, myc, HA-tag, green fluorescent protein (GFP), and the maltose binding protein (MBP). Ubiquitin, and the Ubiquitin-like protein (Ubl's) offer several advantages. For instance, the use of Ubiquitin or similar Ubl's (e.g, SUMO) as a solubility- and expression-enhancing fusion partner has been well documented (see, for example: Ecker, D.J. *et al.*, *Journal of Biological Chemistry* (1989), 264(5): 7715-7719; Marblestone et al., *Protein Science* (2006), 15: 182-189) . In fact, in *S.cerevisiae*, several ribosomal proteins are expressed as C-terminal fusions to ubiquitin. Following translation and protein folding, ubiqutin is cleaved from its co-expressed partner by a highly specific ubiquitin hydrolase, which recognizes and requires the extreme C-terminal Gly-Gly motif present in ubiquitin and cleaves immediately following this sequence; a similar pathway removes Ubl proteins from their fusion partners.

[00113] The invention described herein provides a method to re-localize a normally mitochondrial protein or enzyme by expressing it as fusion with an N-terminal, non-cleavable ubiquitin or ubiquitin-like molecule. In doing so, the re-targeted enzyme enjoys enhanced expression, solubility, and function in the cytosol. In another embodiment, the sequence encoding the MTS can be replaced with a sequence encoding one or more copies of the c-myc epitope tag (amino acids EQKLISEEDL, SEQ ID NO: 18), which will generally not target a protein into the mitochondria and can easily be detected by commercially available antibodies.

Altering the Iron-Sulfur Cluster Domain and/or Redox Active Domain

[00114] In general, the yeast cytosol demonstrates a different redox potential than a bacterial cell, as well as the yeast mitochondria. As a result, isobutanol pathway enzymes such as DHAD which exhibit an iron sulfur (FeS) domain and/or redox active domain, may require the redox potential of the native environments to be folded or expressed in a functional form. Expressing the protein in the yeast cytosol, which can harbor unfavorable redox potential, has the propensity to result in an inactive protein, even if the protein is expressed. The present inventors have identified a number of different strategies to overcome this problem, which can arise when an isobutanol pathway enzyme such as DHAD which is suited to a particular environment with a specific redox potential is expressed in the yeast cytosol.

[00115] In one embodiment, the present invention provides DHAD enzymes that exhibit a properly folded iron-sulfur cluster domain and/or redox active domain in the cytosol. Such DHAD enzymes may either be native or heterologous DHAD homologs or functional analogs or comprise a mutated or modified iron-sulfur cluster domain and/or redox active domain, allowing for a DHAD enzyme to be expressed in the yeast cytosol in a functional form. Thus, if an enzyme in the isobutanol production pathway was identified that was fully soluble and active in the cytosol of said recombinant microorganism, such enzyme can be used without addition of chaperone proteins not already present in the cytosol or without increased expression of chaperone proteins already present in the cytosol. However, some DHAD proteins may need the assistance of additional chaperones or increased chaperone levels to exhibit optimal cytosolic activity.

[00116] Therefore, in various embodiments described herein, the recombinant microorganisms may further comprise a nucleic acid encoding a chaperone protein, wherein said chaperone protein assists the folding of a protein exhibiting cytosolic activity. Addition of the chaperone protein can lead to improved activity, solubility, and/or correct folding of the DHAD enzyme. In one embodiment, the chaperone may be a native protein. In another embodiment, the chaperone protein may be an exogenous protein. In some embodiments, the chaperone protein may be selected from the group consisting of: endoplasmic reticulum oxidoreductin 1 (Ero1, accession no. NP_013576.1), including variants of Ero1 that have been suitably altered to reduce or prevent its normal localization to the endoplasmic reticulum; thioredoxins (which includes Trx1, accession no. NP_013144.1; and Trx2, accession no. NP_011725.1), thioredoxin reductase (Trr1, accession no. NP_010640.1); glutaredoxins (which includes Grx1, accession no. NP_009895.1; Grx2, accession no. NP_010801.1; Grx3, accession no. NP_010383.1; Grx4, accession no. NP_01101.1; Grx5, accession no. NP_015266.1; Grx6, accession no. NP_010274.1; Grx7, accession no. NP_009570.1; Grx8, accession no. NP_013468.1); glutathione reductase Glr1 (accession no. NP_015234.1); Jac1 (accession no. NP_011497.1), including variants of Jac1 that have been suitably altered to reduce or prevent its normal mitochondrial localization; Hsp60 and Hsp10 proteins (e.g., yeast Hsp 60 and Hsp10 proteins, or other eukaryotic Hsp60 and Hsp10 homologs), bacterial

chaperonin homologs (e.g., GroEL and GroES proteins from *Lactococcus lactis*); homologs or active variants thereof, and combinations thereof.

[00117]   As described herein, it is preferred that the DHAD enzymes are properly assembled and folded, thus allowing for said DHADs to exhibit maximal activity in the cytosol.   In yeast, the DHAD Ilv3 is involved in biosynthesis of the amino acids leucine, isoleucine and valine.   Ilv3 is typically localized to the mitochondria, where the chaperonin proteins Hsp60 and Hsp10 aid in the proper folding of the protein (Dubaquie *et. al.* The EMBO Journal 1998 17: 5868-5876).   In wild type yeast cells, Ilv3 is found in the soluble fraction of cell lysates.   In extracts from an *hsp60* temperature-sensitive mutant, at the non-permissive temperature, there is no detectable soluble Ilv3.   All of the protein is found in the insoluble fraction, in a presumably inactivated state.   In an *hsp10* temperature-sensitive mutant, at the non-permissive temperature, about half of the Ilv3 is found in the insoluble portion, indicating that Hsp10 is also important for proper folding of Ilv3, but that Hsp60 is required. (Dubaquie *et. al.* The EMBO Journal 1998 17: 5868-5876).

[00118]   Thus, in one embodiment of the present invention, wherein the yeast DHAD encoded by *ILV3* gene is used in the cytosol of a isobutanol-producing recombinant microorganism (e.g., a yeast microorganism), Hsp60 and/or Hsp10 from the same yeast, homologs thereof from other microorganisms, or active variants thereof can be overexpressed in said microorganism to increase the activity, solubility, and/or correct folding of DHAD encoded by *ILV3* gene to increase the productivity, titer, and/or yield of isobutanol produced.   Alternatively, if said microorganism is a yeast and it naturally expresses chaperonin proteins homologous to Hsp60 and/or Hsp10 in its cytosol, DHAD encoded by *ILV3* can be expressed in said yeast without the overexpression of the Hsp60 and/or the Hsp10 proteins.   In another embodiment, wherein the DHAD derived from an organism other than yeast is used for isobutanol production, chaperonin homologs, or active variants thereof derived from said non-yeast organism or related non-yeast organism can be overexpressed together with the DHAD derived from said non-yeast organism.   In one embodiment, said non-yeast organism is an eukaryotic organism.   In another embodiment, said non-yeast organism is a prokaryotic organism.   In a further embodiment, said non-yeast organism is a bacterium (e.g., *E. coli.*, or *Lactococcus lactis*).   For example, the *Lactococcus lactis* GroEL and GroES chaperonin proteins are expressed in the yeast cytosol in conjunction with the IlvD from *Lactococcus lactis*.   Overexpression of these genes may be accomplished by methods as described herein.

[00119]   Also disclosed herein are recombinant microorganisms comprising one or more genes encoding an iron-sulfur cluster assembly protein. Iron-sulfur cluster assembly for insertion into yeast apo-iron-sulfur proteins begins in yeast mitochondria. To assemble in yeast the active iron-sulfur proteins containing the cluster, either the apo-iron-sulfur protein is imported into the mitochondria from the cytosol and the iron-sulfur cluster is inserted into the protein and the active protein remains localized in the mitochondria; or the iron-sulfur clusters or precursors thereof are exported from the mitochondria to the cytosol and the active protein is assembled in the cytosol or other cellular compartments.

103182 v1/DC

[00120] Targeting of yeast mitochondrial iron-sulfur proteins or non-yeast iron-sulfur proteins to the yeast cytosol can result in such proteins not being properly assembled with their iron-sulfur clusters. This present invention overcomes this problem by co-expression and cytosolic targeting in yeast of proteins for iron-sulfur cluster assembly and cluster insertion into apo-iron-sulfur proteins, including iron-sulfur cluster assembly and insertion proteins from organisms other than yeast, together with the apo-iron-sulfur protein to provide assembly of active iron-sulfur proteins in the yeast cytosol.

[00121] In some embodiments, the present invention provides methods of using Fe-S cluster containing protein in the eukaryotic cytosol for improved isobutanol production in a microorganism, comprising overexpression of a Fe-S cluster-containing protein in the isobutanol production pathway in an microorganism. In a preferred embodiment, said microorganism is a yeast microorganism. In one embodiment, said Fe-S cluster-containing protein is a endogenous protein. In another embodiment, said Fe-S cluster-containing protein is an exogenous protein. In one embodiment, said Fe-S cluster-containing protein is derived from a eukaryotic organism. In another embodiment, said Fe-S cluster-containing protein is derived from a prokaryotic organism. In one embodiment, said Fe-S cluster-containing protein is DHAD. In one embodiment, said Fe-S cluster is a 2Fe-2S cluster. In another embodiment, said Fe-S cluster is a 4Fe-4S cluster.

[00122] All known DHAD enzymes contain an iron sulfur cluster, which is assembled *in vivo* by a multi-component pathway. DHADs contain one of at least two types of iron sulfur clusters, a 2Fe-2S cluster as typified by the spinach enzyme (Flint and Emptage, *JBC* 1988 263(8): 3558) or a 4Fe-4S cluster as typified by the *E. coli* enzyme (Flint *et. al., JBC* 1993 268(20): 14732). In eukaryotic cells, iron-sulfur cluster proteins can be found in either the cytosol or, more commonly, in the mitochondria. Within the mitochondria, a set of proteins, collectively similar to the ISC and/or SUF systems of *E.coli*, are present and participate in the assembly, maturation, and proper insertion of Fe-S clusters into mitochondrial target proteins. (Lill and Muhlenhoff, Ann. Rev. Biochem. 2008 77: 669). In addition, a cytosolic iron sulfur assembly system is present and is collectively termed the CIA machinery. The CIA system promotes proper Fe-S cluster maturation and loading into cytosolically-localized iron sulfur proteins such as Leu1. Importantly, function of the CIA system is dependent on a critical (but still uncharacterized) factor exported from the mitochondria. In the yeast *S.cerevisiae*, the native DHAD, encoded by *ILV3*, is a mitochondrially-localized protein, where it is presumably properly recognized and activated by Fe-S cluster insertion by the endogenous machinery. Accordingly, ectopic expression of a DHAD in the yeast cytosol might be not expected to be functional due to its presence in a non-native compartment and the concomitant lack of appropriate Fe-S cluster assembly machinery.

[00123] The *E. coli* DHAD (encoded by *ilvD*) is sensitive to oxygen, becoming quickly inactivated when isolated under aerobic conditions (Flint *et. al., JBC* 1993 268(20): 14732; Brown *et. al. Archives Biochem. Biophysics* 1995 319(1): 10). It is thought that this oxygen sensitivity is due to the presence of a labile 4Fe-4S cluster, which is unstable in the presence of oxygen and reactive oxygen species, such as

oxygen radicals and hydrogen peroxide. In yeast and other eukaryotes, the mitochondrial environment is reducing, *i.e.* it is a low oxygen environment, in contrast to the more oxygen-rich environment of the cytosol. The redox state of the cytosol is thus expected to be a problem for expressing mitochondrially localized DHADs, which are natively located in the mitochondria, or in expressing DHADs from many bacterial species which typically have an intracellular reducing environment. The spinach DHAD has been shown to be more oxygen resistant than the *E. coli* enzyme in *in vitro* assays (Flint and Emptage, JBC 1988 263(8):3558), which may be due to its endogenous localization to the plastid, where it would normally encounter a relatively high-oxygen environment. It has been suggested that DHADs with 2Fe-2S clusters are inherently more resistant to oxidative damage and they are therefore an attractive possibility for inclusion in the cytosolically localized isobutanol pathway.

[00124] An additional complication to the oxygen sensitivity of DHADs is that the iron sulfur clusters must be properly assembled and inserted into the enzyme such that an active enzyme results. There are several types of machinery that produce iron sulfur clusters and properly assemble them into proteins, including the NIF system found in bacteria and in some eukaryotes, the ISC system found in bacteria and mitochondria, the SUF system found in bacteria and plastids, and the CIA system found in the cytosol of eukaryotes.

[00125] Thus, the methods of using Fe-S cluster in the eukaryotic cytosol for improved enzymatic activity in isobutanol production pathway as described above may further comprise the co-expression a heterologous Fe-S cluster-containing DHAD with the NIF assembly system in the yeast cytosol to aid in assembling said heterologous DHADs. The NIF system found in the parasite *Entamoeba histolytica* has been shown to complement the double deletion of the *E. coli* ISC and SUF assembly systems (Ali et. al. JBC 2004 279(16): 16863). The critical components of the *Entamoeba* assembly system comprise only two genes, NifS and NifU. In one embodiment, these two components are overexpressed in the yeast cytosol to increase activity and/or stability of cytosolic DHADs. In one embodiment, the NIF system is the *E. hisotlytica* NIF system; in another embodiment, the NIF system is from other organisms (e.g. *Lactococcus lactis*). An advantage of using the *E. hisotlytica* assembly system is that it has already been demonstrated to work in a heterologous organism, *E. coli*.

[00126] A 2Fe-2S cluster-containing DHAD can be used in the present invention. In one embodiment, the 2Fe-2S cluster DHADs includes all known 2Fe-2S cluster dehydratase enzymes identified biochemically. In another embodiment, the 2Fe-2S cluster DHADs include those predicted to be 2Fe-2S cluster dehydratases containing some version of the consensus motif for 2Fe-2S cluster proteins, *e.g.*, the motif $CX_4CX_2CX_{-30}C$ (SEQ ID NO. 126, Lill and Muhlenhoff, *Ann. Rev. Biochem* 2008 77: 669). For example, based on the extremely highly conserved DHAD gene sequences shared amongst plant species, the inventors have synthesized a likely 2Fe-2S DHAD from *Arabidopsis* (and rice, *Oryza sativa japonica*) which can be used to improve isobutanol production *in vivo* in the cytosolic isobutanol pathway.

[00127] Alternatively, a DHAD may be determined to be a 2Fe-2S protein or a 4Fe-4S protein based on a phylogenetic tree, such as Figure 2, below. Sequences not

present on the example phylogenetic tree disclosed here could be added to the tree by one skilled in the art. Furthermore, once a new sequence was added to the DHAD phylogenetic tree, one skilled in the art may be able to determine if it is a 2Fe-2S or a 4Fe-4S cluster containing protein based on the phylogenetic relationship to known 2Fe-2S or a 4Fe-4S cluster containing DHADs.
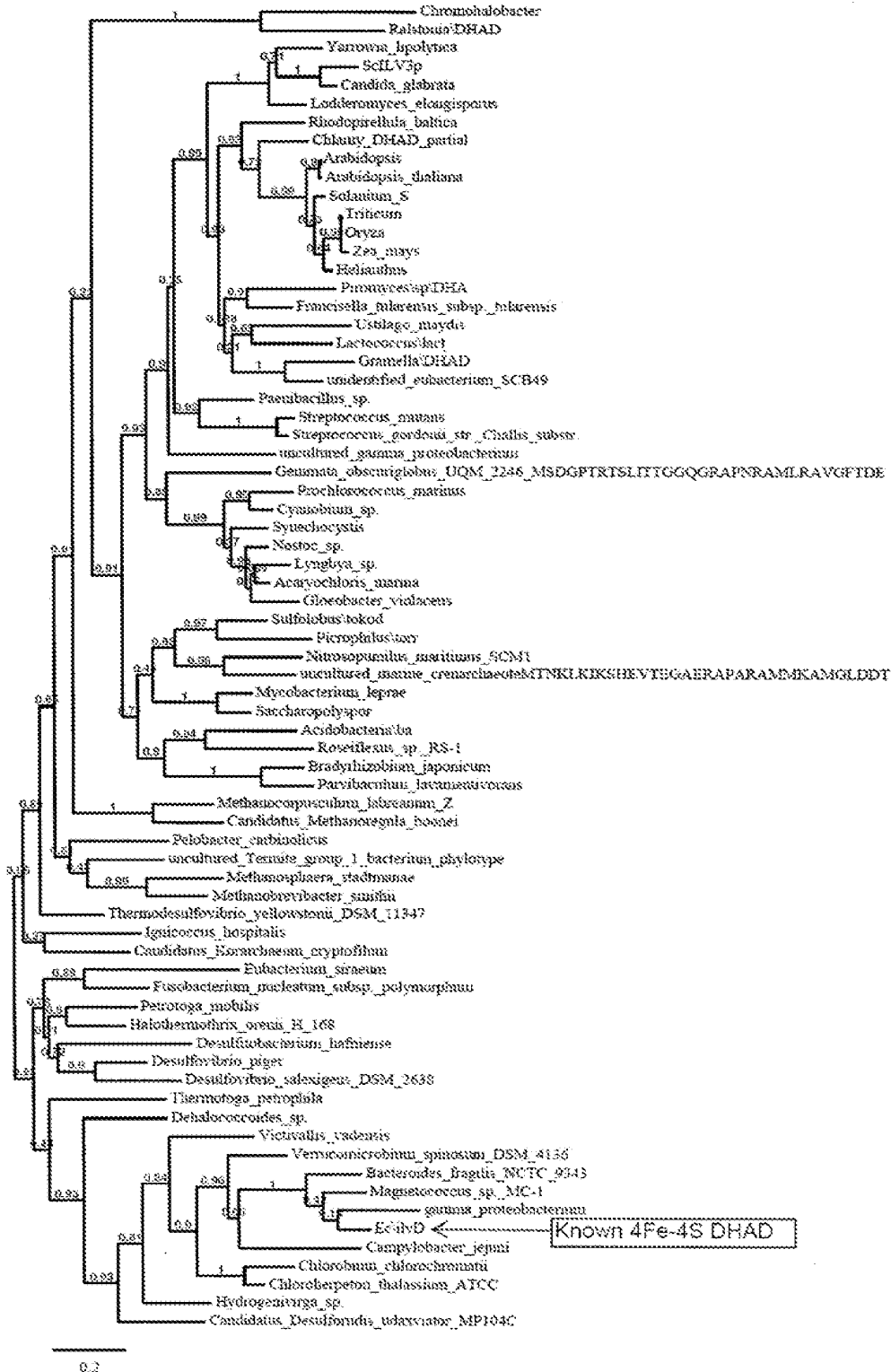

Figure 2. Phylogenetic tree of DHAD proteins. Numbers at nodes indicate bootstrap values. Ec\ilvD is a known 4Fe-4S DHAD enzyme from *Escherichia coli*.

[00128] In another embodiment, a 4Fe-4S cluster-containing DHAD could substitute for the 2Fe-2S cluster-containing DHAD in the cytosol. In one embodiment, said 4Fe-4S cluster DHAD is engineered to be oxygen resistant, and therefore more active in the cytosol of cells grown under aerobic conditions.

[00129] In one embodiment of this invention, the apo-iron-sulfur protein DHAD enzyme encoded by the *E. coli ilvD* gene is expressed in yeast together with *E. coli* iron-sulfur cluster assembly and insertion genes comprising either the *cyaY, iscS, iscU, iscA, hscB, hscA, fdx* and *isuX* genes or the *sufA, sufB, sufC, sufD, sufS* and *sufE* genes. This strategy allows for both the apo-iron-sulfur protein (DHAD) and the iron-sulfur cluster assembly and insertion components (the products of the *isc* or *suf* genes) to come from the same organism, causing assembly of the active DHAD iron-sulfur protein in the yeast cytosol. As a modification of this embodiment, for those *E. coli* iron-sulfur cluster assembly and insertion components that localize to or are predicted to localize to the yeast mitochondria upon expression in yeast, the genes for these components are engineered to eliminate such targeting signals to ensure localization of the components in the yeast cytoplasm. Thus, in some embodiments, one or more genes encoding an iron-sulfur cluster assembly protein may be mutated or modified to remove a signal peptide, whereby localization of the product of said one or more genes to the mitochondria is prevented. In certain embodiments, it may be preferable to overexpress one or more genes encoding an iron-sulfur cluster assembly protein.

[00130] In additional embodiments, iron-sulfur cluster assembly and insertion components from other than *E. coli* can be co-expressed with the *E. coli* DHAD protein to provide assembly of the active DHAD iron-sulfur cluster protein. Such iron-sulfur cluster assembly and insertion components from other organisms can consist of the products of the *Helicobacter pylori nifS* and *nifU* genes or the *Entamoeba histolytica nifS* and *nifU* genes. As a modification of this embodiment, for those non-*E. coli* iron-sulfur cluster assembly and insertion components that localize to or are predicted to localize to the yeast mitochondria upon expression in yeast, the genes for these components can be engineered to eliminate such targeting signals to ensure localization of the components in the yeast cytoplasm.

[00131] As a further modification of this embodiment, in addition to co-expression of these proteins in aerobically-grown yeast, these proteins may be co-expressed in anaerobically-grown yeast to lower the redox state of the yeast cytoplasm to improve assembly of the active iron-sulfur protein.

[00132] In another embodiment, the above iron-sulfur cluster assembly and insertion components can be co-expressed with DHAD apo-iron-sulfur enzymes other than the *E. coli* IlvD gene product to generate active DHAD enzymes in the yeast cytoplasm. As a modification of this embodiment, for those DHAD enzymes that localize to or are predicted to localize to the yeast mitochondria upon expression in yeast, then the genes for these enzymes can be engineered to eliminate such targeting signals to ensure localization of the enzymes in the yeast cytoplasm.

[00133] In additional embodiments, the above methods used to generate active DHAD enzymes localized to yeast cytoplasm may be combined with methods to

generate active acetolactate synthase, KARI, KIVD and ADH enzymes in the same yeast for the production of isobutanol by yeast.

[00134] In another embodiment, production of active iron-sulfur proteins other than DHAD enzymes in yeast cytoplasm can be accomplished by co-expression with iron-sulfur cluster assembly and insertion proteins from organisms other than yeast, with proper targeting of the proteins to the yeast cytoplasm if necessary and expression in anaerobically growing yeast if needed to improve assembly of the active proteins.

[00135] In another embodiment, the iron-sulfur cluster assembly protein encoding genes may be derived from eukaryotic organisms, including, but not limited to yeasts and plants. In one embodiment, the iron-sulfur cluster protein encoding genes are derived from a yeast organism, including, but not limited to S. cerevisiae. In specific embodiments, the yeast derived genes encoding iron-sulfur cluster assembly proteins are selected from the group consisting of Cfd1 (accession no. NP_012263.1), Nbp35 (accession no. NP_011424.1), Nar1 (accession no. NP_014159.1), Cia1 (accession no. NP_010553.1), and homologs or variants thereof. In a further embodiment, the iron-sulfur cluster assembly protein encoding genes may be derived from plant nuclear genes which encode proteins translocated to chloroplast or plant genes found in the chloroplast genome itself.

[00136] As noted above, the iron-sulfur cluster assembly genes may be derived from eukaryotic organisms, including, but not limited to yeasts and plants. In one embodiment, the iron-sulfur cluster genes are derived from a yeast organism, including, but not limited to S. cerevisiae. In specific embodiments, the yeast derived iron-sulfur cluster assembly genes are selected from the group consisting of CFD1, NBP35, NAR1, CIA1, and homologs or variants thereof. In a further embodiment, the iron-sulfur cluster assembly genes may be derived from a plant chloroplast.

[00137] In certain embodiments described herein, it may be desirable to reduce or eliminate the activity and/or proteins levels of one or more iron-sulfur cluster containing cytosolic proteins. This modification increases the capacity of a yeast to incorporate [Fe-S] clusters into cytosolically expressed proteins wherein said proteins can be native proteins that are expressed in a non-native compartment or heterologous proteins. This is achieved by deletion of a highly expressed native cytoplasmic [Fe-S]-dependent protein. More specifically, the gene *LEU1* is deleted coding for the 3-isopropylmalate dehydratase which catalyses the conversion of 3-isopropylmalate into 2-isopropylmaleate as part of the leucine biosynthetic pathway in yeast. Leu1p contains an 4Fe-4S cluster which takes part in the catalysis of the dehydratase. Some DHAD enzymes also contain a 4Fe-4S cluster involved in its dehydratase activity. Therefore, although the two enzymes have different substrate preferences the process of incorporation of the Fe-S cluster is generally similar for the two proteins. Given that Leu1p is present in yeast at 10000 molecules per cell (Ghaemmaghami S. *et al. Nature* 2003 425: 737), deletion of *LEU1* therefore ensures that the cell has enough spare capacity to incorporate [Fe-S] clusters into at least 10000 molecules of cytosolically expressed DHAD. Taking into account the specific activity of DHAD (*E. coli* DHAD is reported to have a specific activity of 63 U/mg (Flint, D.H. *et al., JBC* 1993 268: 14732), the *LEU1* deletion yeast strain would generally exhibit an increased capacity for DHAD activity in the cytosol as measured

in cell lysate.

[00138] In alternative embodiments, it may be desirable to further overexpress an additional enzyme that converts 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol. In a specific embodiment, the enzyme may be selected from the group consisting of 3-isopropylmalate dehydratase (Leu1p) and imidazoleglycerol-phosphate dehydrogenase (His3p) or other dehydratases listed in Table 1.

Table 1. Dehydratases with putative activity towards 2,3-dihydroxyisovalerate.

Gene	Species	Native Substrate	Comments
dgoD	E. coli	D-galactonate	Acid-sugar dehydratases
rspA	E. coli	D-mannonate, D-altronate	
yfaW	E. coli	L-rhamnonate	
fucD	X. campestris	L-fuconate	
LGD1	H. jecorina	L-galactonate	
pdd	K. oxytoca	diols	Other non-Fe-S dehydratases
ENO1/2, ERR1/2/3	S. cerevisiae	2-phosphoglycerate	
HIS3	S. cerevisiae	Imidazoleglycerol-phosphate	

[00139] Because DHAD activity is limited in the cytosol, alternative dehydratases that convert dihydroxyisovalerate (DHIV) to 2-ketoisovalerate (KIV) and are physiologically localized to the yeast cytosol may be utilized. Leu1p and His3p and other enzymes encoded by genes listed in Table 1 are dehydratases that potentially may exhibit affinity for DHIV. Leu1p is an Fe-S binding protein that is involved in leucine biosynthesis and is also normally localized to the cytosol. His3p is involved in histidine biosynthesis and is similar to Leu1p, it is generally localized to the cytosol or predicted to be localized to the cytosol. This modification overcomes the problem of a DHAD that is limiting isobutanol production in the cytosol of yeast. The use of an alternative dehydratase that has activity in the cytosol with a low activity towards DHIV may thus be used in place of the DHAD in the isobutanol pathway. As described herein, such enzyme may be further engineered to increase activity with DHIV.

Increased Mitochondrial Export of Essential Components for Iron Sulfur Protein Assembly in the Cytosol

[00140] As noted herein, the third step in the engineered isobutanol pathway is the conversion of dihydroxy isovalerate (DHIV) to ketosisovalerate (KIV) by a dihydroxyacid dehydratase (DHAD). DHADs often require iron sulfur clusters for activity, and the native yeast DHAD acquires its iron sulfur cluster via the mitochondria ISC machinery, remaining within the mitochondria as an active enzyme. However, isobutanol production by the engineered pathway requires DHAD

to be functionally expressed within the cytosol, and such a DHAD presumably requires iron sulfur clusters to be added in the cytosol. One of the inventions disclosed herein addresses possible genetic or chemical approaches to increase the functional activity of cytosol DHADs. The present invention provides ways to increase the export of an essential compound that is generated in mitochondria, thereby increasing the amount of the compound available for use by the cytosolic iron sulfur assembly machinery (e.g. CIA) to effectively increase the functional expression of cytosolic DHADs.

Overexpressing Mitochondrial Iron Sulfur Cluster (ISC) Machinery

[00141] The compound generated within the mitochondrial matrix that is essential for iron sulfur protein assembly in the cytosol is subsequently exported through the ABC transporter, Atm1, and is chaperoned across the intermembrane space of the mitochondria to the cytosol by Erv1 (reviewed in Lill Annual Review of Biochemistry 2008). Sc_BAT1 was identified as a third putative component of the compound's mitochondrial export machinery by a genetic selection of suppressors of a Sc_atm1 temperature sensitive allele (Kispal et al, 1996, JBC, 271:24458-24464). It is also suggested that a further strong indication for a direct functional relationship between Atm1p and Bat1p is the leucine auxotrophy associated with the deletion of the ATM1 gene.

[00142] To faciliate export of the essential compound, the present invention provides recombinant microorganisms that have been engineered to overexpress one or more mitochondrial export proteins. In various embodiments described herein, the mitochondrial export protein may be selected from the group consisting of the S. cerevisiae ATM1, the S. cerevisiae ERV1, and the S. cerevisiae BAT1, or homologs thereof. Such manipulations can increase the export of the essential compound out of the mitochondria to increase the amount available for use by the cytosolic iron sulfur assembly machinery (e.g. CIA) to effectively increase the functional expression of cytosolic DHADs.

Increasing Inner Mitochondrial Membrane Electrical Potential

[00143] In one embodiment, the present invention provides recombinant microorganisms that have further been engineered to increase inner mitochondrial membrane potential, $\Delta\Psi_M$. As described herein, although yeast cells require a function mitochondrial compartment, they are viable without the mitochondrial genome (mtDNA). However, loss of mtDNA has been linked to destabilization of the nuclear genome (Veatch et al., 2009, Cell, 137(7):1179-1181). Nuclear genome stability was restored in yeast lacking mtDNA when a suppressor mutation (ATP1-111) was introduced (Veatch et al., 2009, Cell, 137(7):1179-1181, Francis et al., 2007, J. Bioenerg. Biomembr. 39(2):149-157). The mutation has been shown to increase ATP hydrolysis activity of the mitochondrial ATP synthase, and similar mutations in the ATP synthase complex have also been shown to increase the electrical potential across the inner membrane of mitochondria, $\Delta\Psi_M$, in cells lacking mtDNA (Smith et al., 2005, Euk Cell, 4(12):2057-2065; Kominsky et al., 2002, Genetics, 162:1595-1604). Generation of $\Delta\Psi_M$ is required for efficient import of

proteins into the mitochondrial matrix, including those involved in assembly and export of a complex required for the assembly of iron sulfur clusters into proteins in the cytosol. The link between $\Delta\Psi_M$ and iron sulfur cluster assembly in the cytosol is supported by microarray data that indicate that the transcriptional profile of cells lacking mtDNA (decreased $\Delta\Psi_M$) is similar to yeast grown under iron depletion conditions (Veatch *et al.*, 2009, *Cell*, 137(7):1179-1181). Introduction of the *ATP1-111* suppressor mutation restores the transcriptional profile to one resembling a wild-type cell's transcriptional profile (Veatch *et al.*, 2009, *Cell*, 137(7):1179-1181). Taken together, these data indicate that $\Delta\Psi_M$ must be sufficient to support assembly of cytosolic iron sulfur proteins, particularly those involved in nuclear genome stability (Veatch et al., *Cell* 2009, 137(7):1247-1258).

[00144] Thus, the present invention aims to generate the highest possible $\Delta\Psi_M$ in a yeast with an intact mitochondrial genome, allowing for the maximization the export of the complex required for assembly of cytosolic iron sulfur proteins, which can in turn increase the amount available for use by the cytosolic iron sulfur assembly machinery (*e.g.* CIA) to effectively increase the functional expression of cytosolic DHADs. $\Delta\Psi_M$ can be maximized several different ways, including, but not limited to: (1) Introducing mutations in the mitochondrial ATP synthase complex that increase ATP hydrolysis activity, or active variants thereof; (2) Overexpressing an ATP/ADP carrier protein that leads to an increase ATP^{4-} import into the mitochondrial matrix in exchange for ADP^{3-}, contributing to generation of $\Delta\Psi_M$; (3) Removal and/or overexpression of additional gene(s) involved in generation of $\Delta\Psi_M$; and (4) Addition of chemical reagents that lead to an increase in $\Delta\Psi_M$.

[00145] In various embodiments described herein, the recombinant microorganism may comprise a mutation in the mitochondrial ATP synthase complex that increases ATP hydrolysis activity. In one embodiment, said mutant mitochondrial is an ATP synthase which can increase ATP hydrolysis activity is from a eukaryotic organism (*e.g.*, a yeast ATP1, ATP2, ATP3). In another embodiment, said mutant mitochondrial ATP synthase is from a prokaryotic organism (*e.g.*, bacteria). Non-limiting examples of said mutant mitochondrial ATP synthase include, mutant ATPase from the *ATP1-111* strain in Francis *et al.*, *J Bioenerg Biomembr*, 2007, 39(2):127-144), a mutant ATPase from the *atp2-227* strain in Smith et al., 2005, Euk Cell, 4(12):2057-2065, or a mutant ATPase from the *yme1* strain in Kominsky *et al.*, 2002, Genetics, 162:1595-1604). In another embodiment, active variants, or homologs of the mutant mitochondrial ATP synthases described above can be applied. In one embodiment, an ATP synthase having a homology to any of ATP1, ATP2, and ATP3 of at least about 70%, at least about 80%, or at least about 90% similarity can be used for a similar purpose.

[00146] In one embodiment, the inner mitochondrial membrane electrical potential can be increased by overexpressing an ATP/ADP carrier protein. Overexpression of the ATP/ADP carrier protein increases ATP^{4-} import into the mitochondrial matrix in exchange for ADP^{3-}. Non-limiting examples of ATP/ADP carrier proteins include the *S. cerevisiae*_AAC1 or the *S. cerevisiae*_AAC3, and active variants or homologs thereof. In one embodiment, an ATP/ADP carrier protein having a homology to either the *S. cerevisiae*_AAC1 or *S. cerevisiae*_AAC3 of at least about 70%, at least about

80%, or at least about 90% similarity can be used for a similar purpose.

[00147] In another embodiment, the inner mitochondrial membrane electrical potential can be increased by removal and/or overexpression of additional gene(s) involved in the generation of $\Delta\Psi_M$. A person skilled in the art will be familiar with proteins encoded by such genes. Non-limiting examples include the protein complexes in the mitochondrial electron transport chain which are responsible for establishing H^+ ions gradient. For examples, complexes on the inner membrane of mitochondria that are involved in conversion of NADH to NAD^+ (Complex I, NADH dehydrogenase), succinate to fumarate (Complex II, cytochrome bc_1 complex), and oxygen to water (Complex IV, cytochrome c oxidase), which are responsible for the transfer of H^+ ions. In another embodiment, enzymes in the citric acid cycle in the matrix of mitochondria can be overexpressed to increase NADH and succinate production, such that more H^+ ions are available. These enzymes include, citrate synthase, aconitase, isocitrate dehydrogenase, α-Ketoglutarate dehydrogenase, succinyl-CoA synthetase, succinate dehydrogenase, fumarase, and malate dehydrogenase.

[00148] In another embodiment, the inner mitochondrial membrane electrical potential can be increased by the addition of chemical reagents that lead to an increase in $\Delta\Psi_M$. In one embodiment, said chemical reagents are substrates in the citric acid cycle in the matrix of mitochondria, wherein when added into the culture, more NADH and succinate can be produced which in turn increase $\Delta\Psi_M$ in the mitochondria. Non-limiting examples of said substrates include, oxaloacetate, acetyl CoA, citrate, cis-Aconitate, isocitrate, oxalosuccinate, α-Ketoglutarate, succinyl-CoA, succinate, fumarate and L-Malate.

Enhancing Cytosolic DHADs Activity by Increasing Cytosol Iron Levels

[00149] Iron plays an essential role in many biological processes, often functioning as a redox active element. Specifically, iron is an important component of iron-sulfur cluster (FeS cluster)-containing proteins. Despite its abundance, the bioavailability of iron is limited by its solubility in aerobic and neutral pH environments and, therefore, cells require an efficient system to acquire this element.

[00150] In the budding yeast *Saccharomyces cerevisiae*, homeostatic regulation of iron uptake occurs (Eide et al., 1992). Iron deprivation induces activity of a high affinity iron uptake system. This induction is mediated by increased transcript levels for genes involved in the iron uptake system, and the regulatory gene *AFT1* plays a critical role in this process (Yamaguchi-Iwai et al., 1995). Mutant strains lacking *AFT1*, due to gene deletion, are unable to induce the high-affinity iron uptake system. On the other hand, mutant strains carrying the *AFT1 -1* UP allele exhibit a gain-of-function phenotype in which iron uptake cannot be repressed by available iron in the environment. The result is an accumulation of iron inside the cell and specifically within the yeast cytosol.

[00151] Iron-dependent gene regulation in *S. cerevisiae* is mediated by two transcription factors: Aft1 and Aft2 (for "activator of ferrous transport") activate gene expression when iron is scarce. Consequently, strains that lack both these factors exhibit reduced expression of the iron regulon. The genes that code for these

103182 v1/DC

factors are thought to have arisen from a genome duplication event. As with many other paralogous genes within *S. cerevisiae*, *AFT1* and *AFT2* code for proteins that have significant regions of identity and overlapping functions. The DNA-binding domain of each protein is in a highly conserved N-terminal region, and a conserved cysteine-to-phenylalanine mutation in either protein generates a factor that activates the high expression of the iron regulon irrespective of iron concentrations. There are clear phenotypic differences in strains that separately lack Aft1 or Aft2. An *aft1* null strain exhibits low ferrous iron uptake and grows poorly under low-iron conditions or on a respiratory carbon source. No phenotype has been attributed to an *aft2* null strain. An *aft1 aft2* double null strain is, however, more sensitive to low-iron growth than a single *aft1* null strain, which is consistent with the functional similarity of these factors. The partial redundancy of these factors allows *AFT2* to complement an *aft1* null strain when it is overexpressed from a plasmid. The properties of Aft1 and Aft2 that distinguish them from each other have not been fully identified. Both factors mediate gene regulation via an iron-responsive element that contains the core sequence 5'-CACCC-3'. It is likely that sequences adjacent to this element influence the ability of each factor to mediate regulation via a particular iron-responsive element. The differential regulation of individual genes by Aft1 and Aft2 results in each factor generating a distinct global transcriptional profile.

[00152] In *S. cerevisiae*, the Aft1 regulon consists of many genes that are involved in the acquisition, compartmentalization, and utilization of iron. These include genes involved in iron uptake (*FET3*, *FTR1*, and *FRE1,2*), siderophore uptake (*ARN1-4* and *FIT1-3*), iron transport across the vacuole membrane (*FTH1*), and iron-sulfur cluster formation (*ISU1,2*). Aft1 binds to a conserved promoter sequence in an iron-dependent manner and activates transcription under low-iron conditions. The Aft2 regulator controls the expression of several distinct genes (See, Table 1 in Rutherford, Eukaryotic Cell, 2004). The initial step in iron acquisition requires reduction of ferric iron chelates in the environment by externally directed reductases encoded by the *FRE1* and *FRE2* genes, thereby generating the ferrous iron substrate for the transport process (Dancis et al., 1992; Georgatsou and Alexandraki, 1994). *FET3* encodes a multi-copper oxidase (Askwith *et al.*, 1994; De Silva et al., 1995) that forms a molecular complex with the iron permease encoded by *FTR1*. This complex, located in the yeast plasma membrane, mediates the high-affinity transport of iron into the cell (Stearman *et al.*, 1996).

[00153] Herein are disclosed methods to increase the levels of iron within the yeast cytosol and mitochondria, such that this iron is more available for the production of iron-sulfur cluster-containing proteins in the cytosol. In one embodiment, the levels of iron within the yeast cytosol are increased. In another embodiment, an increase in iron in the yeast cytosol leads to an increase in activity of a cytosolically-expressed dihydroxyacid dehydratase (DHAD), an FeS cluster-containing protein that catalyzes the reaction of 2,3-dihydroxyisovalerate to 2-ketoisovalerate. In another embodiment, an increase in iron in the yeast cytosol leads to an increase in activity of a cytosolically-expressed DHAD and a subsequent increase in the productivity, titer, and/or yield of isobutanol produced by the DHAD-containing strain.

[00154] Specifically, increasing the expression of the genes *AFT1, AFT2, GRX3* and/or *GRX4* alone or in combination will modulate the amount and availability of iron in the yeast cytosol or mitochondria. Alternatively, deletion or attenuation of the genes *AFT1, AFT2, GRX3* and/or *GRX4* alone or in combination will modulate the amount and availability of iron in the yeast cytosol or mitochondria. Since Aft1 activates the expression of target genes in response to changes in iron availability, Overexpression of *AFT1* is predicted to increase the machinery to import more iron into the cytosol. In one embodiment, *AFT1* is overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, native Aft1 is replaced with a mutant version that is constitutively active (e.g. a conserved cysteine-to-phenylalanine mutation). Likewise, *AFT2* is predicted to result in increased expression of the machinery to import more iron into the mitochondria. In one embodiment, *AFT2* is overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, native *AFT2* is replaced with a mutant version that is constitutively active (e.g. a conserved cysteine-to-phenylalanine mutation). These embodiments can also be combined with increases in the extracellular iron concentration to provide increased iron in the cytosol or mitochondria of the cell. Increase in iron in either the cytosol or the mitochondria by this method may make iron more available for use in FeS cluster-containing proteins, such as DHAD. An increase in iron may lead to a corresponding increase in DHAD activity. An increase in DHAD activity in the yeast cytosol, in the context of an isobutanol production pathway, may lead to an increase in isobutanol productivity, titer, and/or yield by the cell.

[00155] Grx3 and Grx4 are monothiol glutaredoxins that have been shown to be involved in cellular Fe content modulation and delivery in yeast. Glutaredoxins are glutathione-dependent thiol-disulfide oxidoreductases that function in maintaining the cellular redox homeostasis. *S. cerevisiae* has two dithiol glutaredoxins (Grx1 and Grx2) and three monothiol glutaredoxins (Grx3, Grx4, and Grx5). The monothiol glutaredoxins are believed to reduce mixed disulfides formed between a protein and glutathione in a process known as deglutathionylation. In contrast, dithiol glutaredoxins can participate in deglutathionylation as well as in the direct reduction of disulfides. Grx5, the most studied monothiol glutaredoxin, is localized to the mitochondrial matrix, where it participates in the maturation of Fe-S clusters. Grx3 and Grx4 are predominantly localized to the nucleus. These proteins can substitute for Grx5 when overexpressed and targeted to the mitochondrial matrix; no information on their natural function has been reported. In addition to the reported interaction between Grx3 and Aft1, iron inhibition of Aft1 requires glutathione. It has been shown that iron sensing is dependent on the presence of the redundant Grx3 and Grx4 proteins. One report indicated that removal of both Grx3 and Grx4 resulted in constitutive expression of the genes regulated by Aft1/Aft2. This result suggested that the cells accumulated Fe at levels greater than normal.In one embodiment, Grx3 is overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter.

[00156] In another embodiment, Grx4 is overexpressed from a plasmid or by

inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, Grx3 and Grx4 are overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, Grx3, Grx4, or Grx3 and Grx4 are deleted or attenuated. In another embodiment, Grx3 and Aft1are overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, Grx4 and Aft1are overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, Grx3 and Aft2are overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. In another embodiment, Grx4 and Aft2 are overexpressed from a plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter. These embodiments can also be combined with increases in the extracellular iron concentration to provide increased iron in the cytosol or mitochondria of the cell. One or both of:Atf1, Atf2 is overexpressed either alone or in combination with: Grx3 or Grx4. Such overexpression can be accomplished by plasmid or by inserting multiple copies of the gene into the chromosome under the control of a constitutive promoter.

[00157] Thus, the present invention provides methods of increasing iron concentration and availability in the yeast cytosol to increase the amount of functional iron-sulfur cluster-containing proteins in the yeast cytosol. As noted above, genes that can modulate the amount and availability of iron in the cytosol include the genes *AFT1*, *AFT2*, *GRX3*, and *GRX4*. Alternatively, deletion or attenuation of the genes *GRX3* and/or *GRX4* will increase the amount of iron within the yeast cells.

[00158] Accordingly, in one embodiment, the present invention provides recombinant microorganisms that have further been engineered to overexpress one or more genes selected from the group consisting of *AFT1*, *AFT2*, *GRX3*, and *GRX4*, or homologs thereof. In an alternative embodiment, the present invention provides recombinant microorganisms that have been engineered to delete and/or attenuate one or more genes selected from the group consisting of *GRX3* and *GRX4*, or homologs thereof. In alternative embodiments, proteins having a homology to any one of *AFT1*, *AFT2*, *GRX3*, and *GRX4* of at least about 70%, at least about 80%, or at least about 90% similarity can be used for a similar purpose.

[00159] Overexpression of the genes may be accomplished by any number of plasmid vectors that function in yeast. Further, expression of genes from single or multiple copy integrations into the chromosome of the cell may be useful. Use of a number of constitutive promoters, such as *TDH3*, *TEF1*, *CCW12*, *PGK1*, and *ENO2*. Different levels of expression of the genes may be achieved by using promoters with different levels of activity, either in single or multiple copy integrations or on plasmids. Alternatively promoters that are active under desired conditions, such as growth on glucose, may be used. For example a promoter from one of the glycolytic genes, the *PDC1* promoter, and a promoter from one of the *ADH* genes in *S. cerevisiae* may all be useful. Also, embodiments are exemplified using the yeast

Saccharomyces cerevisiae. However, other yeasts, such as those from the genera listed herein may also be used.

Enhancing Cytosolic DHADs Activity by Increasing Cytosol Sulfur Levels

[00160] The present invention also describe methods of increasing the levels of sulfur-containing compounds within yeast cells, including the amino acid cysteine, such that this sulfur is more available for the production of iron-sulfur cluster-containing proteins in the yeast cytosol. Specifically, by increasing the concentration of sulfur-containing compounds in the cell such, the activity of a functional DHAD is enhanced in the yeast cytosol.

[00161] Accordingly, the present invention provides recombinant microorganisms that have been engineered to overexpress one or more genes to increase biosynthesis of cysteine or uptake of exogenous cysteine by the cell in order to increase the amount and availability of sulfur-containing compounds for the production of active iron-sulfur cluster-containing proteins in the yeast cytosol. In one embodiment, the recombinant microorganisms have been engineered to increase the expression of one or more proteins to increase cysteine biosynthesis by the cell, including, but not limited to *MET3, MET14, MET16, MET10, MET5, MET1, MET8, MET2, MET17, HOM3, HOM2, HOM6, CYS3, CYS4, SUL1, SUL2,* active variants thereof, homologs thereof, and combination thereof, to increase cysteine biosynthesis by the cell. In another embodiment, the recombinant microorganisms have been engineered to increase the expression of one or more transport proteins, including, but not limited to *YCT1, MUP1, GAP1, AGP1, GNP1, BAP1, BAP2, TAT1,* active variants thereof, homologs thereof, and combination thereof.

[00162] As noted above, increasing uptake of exogenous cysteine by the cell will increase the amount and availability of sulfur-containing compounds for the production of active iron-sulfur cluster containing proteins in the cytosol of the cell. Addition of increased exogenous cysteine to yeast cells, separately from or in addition to increased expression of the transport protein-encoding genes as described above, can also increase the level and availability of sulfur-containing compounds within the cell such that the sulfur is more available for the production of iron-sulfur cluster-containing proteins in the cell cytosol.

[00163] Sulfur is a necessary element for the biogenesis of iron-sulfur cluster (FeS cluster)-containing protein *in vivo.* Sulfur is a component of the FeS clusters that are incorporated into such proteins and is also a component of compounds such as glutathiones, which are essential for FeS cluster biogenesis in many organisms as well as being involved in cellular redox homeostasis. The direct source of the sulfur for these processes in many organisms is the amino acid cysteine. The sulfur from cysteine is mobilized into FeS clusters during FeS cluster biogenesis using cysteine desulfurase proteins identified in many organisms such as IscS, SufS (together with SufE), NifS and Nfs1 (together with Isd11). Additionally, glutathione biosynthesis requires cysteine.

[00164] Increased expression Fe-S cluster-containing proteins in organisms such as the budding yeast *Saccharomyces cerevisiae* results in an increased demand for sulfur, in the form of cysteine, in the cell. Such an increased demand for cysteine

may possibly be met by natural induction of the endogenous cysteine biosynthetic pathway but maximal natural induction of this pathway may be insufficient to provide enough cysteine for the proper assemble and maintenance of increased levels of FeS cluster-containing proteins in the cell. Such cells with an increased demand for cysteine may also induce cysteine and/or sulfate transport pathways to bring in exogenous cysteine for or sulfate, which is the sulfur donor for cysteine biosynthesis. However, maximal natural induction of these transport systems may also be insufficient to meet the sulfur requirement of such cells.

[00165] Assembly of active FeS cluster-containing proteins in the native yeast cytosol requires the production and export to the cytosol by the mitochondria of an unidentified sulfur-containing compound derived from the mitochondrial FeS cluster biogenesis pathway and the amino acid cysteine and requiring glutathione for export. Overexpression of an FeS cluster-containing protein in the yeast cytosol or the localization of a previously non-cytosolic FeS cluster-containing protein to the yeast cytosol may result in the decreased availability of this unidentified sulfur-containing compound in the yeast cytosol and low activity of the cytosolic FeS cluster-containing protein or proteins. Increased availability of cysteine to the cell may prevent this limitation by providing increased sulfur for the biosynthesis of this compound and sufficient glutathione for its export from the mitochondria.

[00166] Sulfur for the assembly of FeS cluster-containing proteins expressed in the yeast cytosol may also be provided by localization of cysteine desulfurase proteins to the yeast cytosol. Expression of such proteins in the yeast cytosol may result in an increased demand for cysteine by such cells, especially in the cytosol. Additionally, damage to the FeS cluster of FeS cluster-containing proteins expressed in the yeast cytosol, due to the oxic nature of the yeast cytosol or due to reactive oxygen or nitrogen species, may require additional sulfur derived from cysteine for repair or regeneration of the damaged clusters. As well, additional sulfur derived from cysteine may modulate the redox balance of the yeast cytosol through the production of increased levels of compounds such as glutathione which may positively affect the assembly or activity of FeS cluster-containing proteins in the yeast cytosol.

[00167] Increased cellular sulfur in the form of cysteine can be provided by increasing the biosynthesis of cysteine in the cell or by increasing cellular uptake of exogenous cysteine. Increasing the cellular level of cysteine in these ways is expected to increase the level of other sulfur-containing compounds in the cell that derive their sulfur from cysteine or the cysteine biosynthesis pathway. Cysteine biosynthesis in *S. cerevisiae* involves the uptake of exogenous sulfate by transport proteins encoded by the *SUL1* and/or *SUL2* genes and the action of the proteins encoded by the *MET3, MET14, MET16, MET10, MET5, MET1, MET8, MET2, MET17, HOM3, HOM2, HOM6, CYS4* and *CYS4* genes. Exogenous cysteine is taken up into *S. cerevisiae* by the high-affinity transport system encoded by the *YCT1* gene but also by the broader-specificity transport proteins encoded by the *MUP1, GAP1, AGP1, GNP1, BAP1, BAP2, TAT1* and *TAT2* genes.

[00168] Thus, the present invention present methods to increase the levels of sulfur-containing compounds within the yeast cytosol and/or mitochondria, such that sulfur is more available for the production of iron-sulfur cluster-containing proteins in

the cytosol. In one embodiment, the levels of sulfur-containing compounds within the yeast cytosol and/or mitochondria are increased. In another embodiment, an increase in sulfur-containing compounds in the yeast cytosol or mitochondria leads to an increase in activity of a cytosolically expressed FeS cluster-containing protein DHAD, which catalyzes the reaction of 2,3-dihydroxyisovalerate to 2-ketoisovalerate. In another embodiment, an increase in sulfur-containing compounds in the yeast cytosol or mitochondria leads to an increase in activity of a cytosolically expressed DHAD. In another embodiment, an increase in sulfur-containing compounds in the yeast cytosol and/or mitochondria leads to an increase in activity of a cytosolically expressed DHAD and a subsequent increase in the productivity, titer, and/or yield of isobutanol produced by the DHAD-containing strain.

[00169] In another embodiment, the genes *YCT1, MUP1, GAP1, AGP1, GNP1, BAP1, BAP2, TAT1,* and *TAT2,* active variants thereof, homologs thereof or combination thereof are overexpressed from a plasmid or by inserting multiple copies of the gene or genes into the chromosome under the control of a constitutive promoter. This embodiment can also be combined with providing increased extracellular cysteine to the yeast cells to provide increased sulfur-containing compounds in the cytosol and/or mitochondria of the cells. Overexpression of these genes may be accomplished by methods as described above.

[00170] In another embodiment, providing increased extracellular cysteine to the yeast cells in the absence of any additional engineered expression of transport porteins will provide increased sulfur containing compounds in the cytosol and/or mitochondria of the cells for the improved production of active FeS cluster-containing proteins in the yeast cytosol, which leads to increased isobutanol productivity, titer, and/or yield by the cell.

Enhancing Cytosolic DHAD Activity by Mitigating Oxidative Species or Oxidative Stress

[00171] The present inventors also describe herein methods of protecting enzymes in an isobutanol production pathway (specifically DHAD) in a microorganism to increase isobutanol production by mitigating oxidative species or oxidative stress induced damage in the cytosol of said microorganism. Non-limiting examples of oxidative species include, nitric oxide (NO), reactive nitrogen species (RNS), reactive oxygen species (ROS), hydroxyl radical species, organic hydroperoxide, hypochlorous acids, and combinations thereof. As used herein, the phrase "reactive oxygen species" or "ROS" refers to free radicals that contain the oxygen atom. ROS are very small molecules that include oxygen ions and peroxides and can be either inorganic or organic. They are highly reactive due to the presence of unpaired valence shell electrons. During times of environmental stress (e.g. UV or heat exposure) ROS levels can increase dramatically, which can result in significant damage to cell structures. This cumulates into a situation known as oxidative stress. ROS are also generated by exogenous sources such as ionizing radiation.

[00172] Oxidative stress is caused by an imbalance between the production of reactive oxygen and a biological system's ability to readily detoxify the reactive intermediates or easily repair the resulting damage. All forms of life maintain a

reducing environment within their cells. This reducing environment is preserved by enzymes that maintain the reduced state through a constant input of metabolic energy. Disturbances in this normal redox state can cause toxic effects through the production of peroxides and free radicals that damage all components of the cell, including proteins, lipids, and DNA.

[00173] In chemical terms, oxidative stress is a large rise (becoming less negative) in the cellular reduction potential, or a large decrease in the reducing capacity of the cellular redox couples, such as glutathione. The effects of oxidative stress depend upon the size of these changes, with a cell being able to overcome small perturbations and regain its original state. However, more severe oxidative stress can cause cell death and even moderate oxidation can trigger apoptosis, while more intense stresses may cause necrosis.

[00174] A particularly destructive aspect of oxidative stress is the production of reactive oxygen species, which include free radicals and peroxides, and/or other reactive species. Some of the less reactive of these species (such as superoxide) can be converted by oxidoreduction reactions with transition metals or other redox cycling compounds (including quinones) into more aggressive radical species that can cause extensive cellular damage. The major portion of long term effects is inflicted by damage on DNA. Most of these oxygen-derived species are produced at a low level by normal aerobic metabolism and the damage they cause to cells is constantly repaired. However, under the severe levels of oxidative stress that cause necrosis, the damage causes ATP depletion, preventing controlled apoptotic death and causing the cell to simply fall apart. Non-limiting example of oxidants include, superoxide anion ($\cdot O_2-$, formed in many autoxidation reactions and by the electron transport chain), hydrogen peroxide (H_2O_2, formed by disputation of $\cdot O_2-$ or by direct reduction of O_2) , organic hydroperoxide (ROOH, formed by radical reactions with cellular components such as lipids and/or nucleobases), oxygen centered organic radicals (e.g., RO$\cdot$ alkoxy and ROO$\cdot$, peroxy radicals, formed in the presence of oxygen by radical addition to double bonds or hydrogen abstraction), hypochlorous acid (HOCl, formed from H_2O_2 by myeloperoxidase, and peroxynitrite (ONOO-, formed in a rapid reaction between $\cdot O_2-$ and NO$\cdot$).

[00175] Biological defenses against oxidative damage include protective proteins that remove reactive oxygen species, molecules that sequester metal ions, and enzymes that repair damaged cellular components. Oxidative stress can be defined as a disturbance in the prooxidant-antioxidant balance in favor of prooxidants. One such class of prooxidants are reactive oxygen species, or ROS. ROS are highly reactive species of oxygen, such as superoxide ($O2\cdot^-$), hydrogen peroxide (H_2O_2), and hydroxyl radicals (OH$\cdot$), produced within the cell, usually as side products of aerobic respiration. By some reports, as much as 2% of the oxygen that enters the respiratory chain is converted to superoxide through a one-electron reduction of oxygen. A small amount of superoxide radical is always released from the enzyme when oxygen is reduced by electron carriers such as flavoproteins or cytochromes. This is because the electrons are transferred to oxygen one at a time. The hydroxyl radical and hydrogen peroxide are derived from the superoxide radical.

[00176] Many microbes possess native enzymes to detoxify these ROS. One

example of such a system is superoxide dismutase (SOD) plus catalase. SOD catalyzes a reaction where one superoxide radical transfers its extra electron to the second radical, which is then reduced to hydrogen peroxide. Catalase catalyzes the transfer of two electrons from one hydrogen peroxide molecule to the second, oxidizing the first to oxygen and reducing the second to two molecules of water. If the hydrogen peroxide is not disposed of, then it can oxidize transition metals, such as free iron(II) in the Fenton reaction, and form the free hydroxyl radical, OH·. No known mechanisms exists to detoxify hydroxyl radicals, and thus protection from toxic forms of oxygen must rely on eliminating superoxide and hydrogen peroxide.

[00177] In yeast, to counteract damage of oxidative stress, there are several antioxidant systems with an apparent functional redundancy. For example, there are detoxifying enzymes such as catalases, cytochrome c peroxidase, glutathione peroxidases, glytaredoxins and peroxiredoxins, and many isoforms in distince cellular compartments (Jamieson *et al.*, 1998, *Yeast.* 14:1511-1527; Grant *et al.*, 2001, *Mol. Microbiol* 39:533-541; Collinson *et al.*, 2003, *J. Biol. Chem.* 278:22492-22497; Park *et al.*, 2000, *J. Biol. Chem.* 275:5723-5732).

[00178] As described above, an enzyme involved in the isobutanol production pathway, dihydroxyacid dehydratase (DHAD), contains an iron-sulfur (FeS) cluster domain. This iron-sulfur (FeS) cluster domain is sensitive to damage by ROS, which can lead to inactive enzyme. Both 2Fe-2S and 4Fe-4S DHAD enzymes may be susceptible to inactivation by ROS, however direct evidence exists for inactivation of 4Fe-4S cluster containing proteins, such as homoaconitase and isopropylmalate dehydratase in yeast and DHAD and fumarase from *E. coli*. Therefore, to achieve a functional DHAD expressed in the yeast cytosol in an environment where a substantial amount of ROS may exist from respiration, it may be beneficial to protect the DHAD enzyme from ROS inactivation or oxidative stress through expression of on or more enzymes that reduce or eliminate ROS from the cell.

[00179] To mitigate the potential harmful effects of reactive oxygen species (ROS) or oxidative stress on DHAD in the yeast cytosol, the present inventors have devised several strategies to protect or repair the DHAD from ROS damage. In various embodiments described herein, the invention provides recombinant microorganisms that have been engineered to express one or more proteins in the cytosol that reduce the concentration of reactive oxygen species (ROS) in said cytosol.

[00180] In one embodiment, enzymes that reduce or eliminate the amount of ROS in the cytosol are expressed and targeted to the yeast cytosol. Specifically, enzymes such as catalase, superoxide dismutase (SOD), cytochrome c peroxidase, glutathione peroxidases, glytaredoxins, peroxiredoxins, metallothioneins, and methionine sulphoxide reductases, or any isoforms thereof are expressed, such that they lead to reduction in ROS such as hydrogen peroxide, superoxide, peroxide radicals, and other ROS in the yeast cytosol.

[00181] In one embodiment, a catalase is expressed to reduce the concentration of ROS in the cytosol. In another embodiment, a superoxide dismutase (SOD) is expressed to reduce the concentration of ROS in the cytosol. Usually, microbes that grow by aerobic respiration possess one or both of SOD and catalase. For example, the bacterium *E. coli* and the yeast *Saccharomyces cerevisiae* each

possesses at least one native SOD and catalase (e.g, SOD1 or SOD2 from yeast). In *E. coli*, the genes *kat*G and *kat*E encode catalase enzymes, and the genes *sod*A, *sod*B and *sod*C encode SodA, SodB, and SodC superoxide dismutase enzymes. respectively. In *S. cerevisiae*, the genes *CTT1* and *CTA1* encode catalase CTT1 and CTA1 enzymes, and the genes *SOD1* and *SOD2* encode SOD1 and SOD2 superoxide dismutase enzymes. Many other organisms posess catalase and SOD enzymes and these genes may also be useful for reduction of ROS in the yeast cytosol. In one embodiment, SOD homologs from species other than *E. coli* or yeast can be expressed in yeast cytosol to reduce oxidative stress. In one embodiment, said other species is a plant or a fungus. For example, *SOD1* from *N. crassa* (fungus) may be functionally expressed in the yeast cytosol. In various embodiments described herein, active variants or homologs of the above-described catalases and SODs can be functionally expressed in the yeast cytosol. In another embodiment, protein having a homology to any one of the catalases or SODs described above possessing at least about 70%, at least about 80%, or at least about 90% similarity can be functionally expressed in the yeast cytosol.

[00182]   In one embodiment, the catalase genes from *E. coli* are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, the catalase genes from *S. cerevisiae* are overexpressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In one embodiment, the SOD genes from *E. coli* are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, the *SOD* genes from *S. cerevisiae* are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, promoters of native genes are altered, such that the level of SOD or catalase in the *S. cerevisiae* cytosol is increased. In yet another embodiment, expression of SOD or catalase in the yeast cytosol is mediated by a plasmid. In yet another embodiment, expression of SOD or catalase in the yeast cytosol is mediated by expression of one or more copies of the gene from the chromosome. Other homologs of catalase or SOD may be identified by one skilled in the art through tools such as BLAST and sequence alignment. These other homologs may be expressed in a similar manner described above to achieve a functional catalase or SOD in the yeast cytosol.

[00183]   In another embodiment, a methionine sulphoxide reductase enzyme is expressed to reduce the amount of ROS and protect DHAD from ROS damage and inactivation. In one embodiment, the methionine sulphoxide reductase may be derived from a eukaryotic organism (*e.g.*, a yeast, fungus, or plant). In another embodiment, the methionine sulphoxide reductases may be derived from a prokaryotic organism (e.g., *E. coli*). The principal enzymatic mechanism for reversing protein oxidation acts on the oxidation product of just one amino acid residue, methionine. This specificity for Met reflects the fact that Met is particularly susceptible to oxidation compared with other amino acids. Methionine sulphoxide

reductases (MSRs) are conserved across nearly all organisms from bacteria to humans, and have been the focus of considerable attention in recent years. Two MSR activities have been characterized in the yeast *Saccharomyces cerevisiae*: MsrA (encoded by MXR1) reduces the S stereoisomer of methionine sulphoxide (MetO), while MsrB (encoded by the YCL033c ORF), which we term here MXR2) reduces the R stereoisomer of MetO. Consistent with defence against oxidative damage, mutants deficient in MSR activity are hypersensitive to pro-oxidants such as H_2O_2, paraquat and Cr, while MSR overexpression enhances resistance. Besides methionine residues, iron–sulphur (FeS) clusters are exquisitely ROS-sensitive components of many cellular proteins. It has been reported that MSR activity helps to preserve the function of cellular FeS clusters.

[00184] In one embodiment, the methionine sulphoxide reductase genes from *S. cerevisiae* are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. Specifically, the *S. cerevisiae* methionine sulphoxide reductase genes MsrA (encoded by *MXR1*) and MsrB (encoded by the *YCL033c* ORF) are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. The resulting methionine sulphoxide reductase expressing strain will generally demonstrate improved isobutanol productivity, titer, and/or yield compared to the parental strain that does not comprise methionine sulphoxide reductase genes that are expressed in and targeted to the cytosol. Methionine sulphoxide reductases from other organisms, such as bacteria, may be identified by sequence homology using tools such as BLAST and pairwise sequence alignments by one skilled in the art.

[00185] In yet another embodiment, expression or overexpression of glutathione synthesis enzymes, for example GSH1, leads to increased glutathione in the cell and protection of the DHAD enzyme in the yeast cytosol. In one embodiment, said enzymes are derived from a bacteria (*e.g.*, *E. coli.*). In another emdoiment, said enzymes are derived from yeast (*e.g.*, *S.cerevisiae*). In yet another embodiment, said enzymes are derived from a yeast species different from the yeast used for isobutanol production.

[00186] In one embodiment, one or more metallothionein proteins are expressed in the yeast cytosol to mitigate oxidative stress. Metallothioneins are a family of proteins found in many organisms including yeast and mammals. The biologic function of metallothionein (MT) has been a perplexing topic ever since the discovery of this protein. Many studies have suggested that MT plays a role in the homeostasis of essential metals such as zinc and copper, detoxification of toxic metals such as cadmium, and protection against oxidative stress. MT contains high levels of sulfur. The mutual affinity of sulfur for transition metals makes the binding of these metals to MT thermodynamically stable. Under physiologic conditions, zinc-MT is the predominant form of the metal-binding protein. However, other metals such as copper (Cu) are also bound by MT. Oxidation of the thiolate cluster by a number of mild cellular oxidants causes metal release and formation of MT-disulfide (or thionin if all metals are released from MT, but this is unlikely to occur in vivo), which have

been demonstrated in vivo. MT-disulfide can be reduced by glutathione in the presence of selenium catalyst, restoring the capacity of the protein to bind metals like Zn and Cu. This MT redox cycle may play a crucial role in MT biologic function. It may link to the homeostasis of essential metals, detoxification of toxic metals and protection against oxidative stress. In fact, MT has been shown to substitute for superoxide dismutase in yeast cells in the presence of Cu to protect cells and proteins from oxidative stress.

[00187] In one embodiment, said metallothuineins are derived from a eukaryotic organism (e.g., a yeast, fungus, or plant). In another embodiment, said metallothuineins are derived from a prokaryotic organism (e.g., *E. coli*, *Mycobacterium tuberculosis*). For example, the metallothionein genes *CUP1-1* and *CUP1-2* encoding metallothionein CUP1 from *S. cerevisiae*, active variants thereof, homologs thereof, or combination thereof are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, *S. cerevisiae* metallothionein genes *CUP1-1* and *CUP1-2* are expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, Mycobacterium tuberculosis metallothionein gene *MymT* encoding metallothionein is expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD that is also expressed in and targeted to the yeast cytosol. In another embodiment, *Synechococcus PCC 7942* metallothionein gene *SmtA* is expressed in and targeted to the cytosol of yeast to reduce the amount of ROS and increase the activity of DHAD that is also expressed in and targeted to the yeast cytosol. The resulting metallothionein expressing strain has improved isobutanol productivity, titer, and/or yield compared to the parental strain. Metallothioneins from other organisms, such as bacteria, may be identified by sequence homology using tools such as BLAST and pairwise sequence alignments by one skilled in the art.

[00188] In another embodiment, one or more proteins in the thioredoxin system and/or the glutathione/glutaredoxin system, active variants thereof, homologs thereof, or combination thereof are expressed in the yeast cytosol to mitigate oxidative stress. In one embodiment, said proteins in the thioredoxin system and/or the glutathione/glutaredoxin system are derived from a eukaryotic organism (e.g., a yeast, fungus, or plant). In another embodiment, said said proteins in the thioredoxin system and/or the glutathione/glutaredoxin system are derived from a prokaryotic organism (e.g., *E. coli*). The thioredoxin system and the glutathione/glutaredoxin system help maintain the reduced environment of the cell and play significant roles in defending the cell against oxidative stress. Glutathione is the major protective small molecule against oxidative stress in *Saccharomyces cerevisiae*. Glutathione, the tripeptide γ-glutamyl-cysteinyl-glycine, makes up the major free thiol pool present in millimolar concentrations in aerobic cells. The biosynthesis of glutathione requires γ-glutamyl cysteine synthase (termed Gsh1p) glutathione synthase (Gsh2p) and ATP. Glutathione is essential for viability of yeast but not of bacteria such as *E. coli*. Yeast cells lacking Gsh1p (genotype *gsh1Δ*) are

able to survive in the presence of an external source of glutathione. Deletion of the GSH1 gene encoding the enzyme that catalyzes the first step of glutathione biosynthesis leads to growth arrest, which can be relieved by either glutathione or reducing agents such as dithiothreitol. Evidence suggests that glutathione, in addition to its protective role against oxidative damage, performs a novel and specific function in the maturation of cytosolic Fe/S proteins. Therefore, increasing the levels of glutathione in the yeast cytosol is predicted to protect or increase the steady-state levels of active FeS cluster containing proteins expressed in the yeast cytosol. Specifically, increasing glutathione within the yeast cytosol may increase the amount of active DHAD enzyme expressed in the yeast cytosol, thereby leading to an increase in the titer, productivity, and/or yield of isobutanol produced from the pathway within which DHAD participates (e.g. the isobutanol pathway in Figure 1).

[00189] Thioredoxins and glutaredoxins are small heat-stable proteins with redox-active cysteines that facilitate the reduction of other proteins by catalyzing cysteine thiol-disulfide exchange reactions. The glutathione/glutaredoxin system consists of glutaredoxin, glutathione (produced by glutathione synthase), glutathione reductase and NADPH (as an electron donor). Thus, to increase the effective levels of available glutathione, one or a combination of each of the following enzymes is functionally overexpressed in the yeast cytosol: glutaredoxin (encoded in S.cerevisiae by GRX2, GRX4, GRX6, and GRX7), glutathione reductase (encoded in S.cerevisiae by GLR1); and glutathione synthase (encoded in S.cerevisiae by GSH1 and GSH2). In one embodiment, homologs thereof, active variants thereof, or combination thereof can be expressed in the yeast cytosol to mitigate oxidative stress.

[00190] In another embodiment, the γ-glutamyl cysteine synthase and glutathione synthase genes from S. cerevisiae are expressed in and targeted to the cytosol of yeast to increase the amount of glutathione and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. In another embodiment, S. cerevisiae γ-glutamyl cysteine synthase and glutathione synthase genes Gsh1 and Gsh2 are expressed in and targeted to the cytosol of yeast to increase the amount of glutathione and increase the activity of DHAD also expressed in and targeted to the yeast cytosol. The resulting γ-glutamyl cysteine synthase and glutathione synthase expressing strain has improved isobutanol productivity, titer, and/or yield compared to the parental strain. Homologous genes encoding γ-glutamyl cysteine synthase and glutathione synthase from other organisms, such as other yeast strains, may be identified by sequence homology using tools such as BLAST and pairwise sequence alignments by one skilled in the art.

[00191] Thioredoxins contain two conserved cysteines that exist in either a reduced form as in thioredoxin-$(SH)_2$) or in an oxidized form as in thioredoxin-S_2) when they form an intramolecular disulfide bridge. Thioredoxins donate electrons from their active center dithiol to protein disulfide bonds (Protein-S_2) that are then reduced to dithiols (Protein-$(SH)_2$). The resulting oxidized thioredoxin disulfide is reduced directly by thioredoxin reductase with electrons donated by NADPH. Hence the thioredoxin reduction system consists of thioredoxin, thioredoxin reductase, and NADPH. Oxidized glutaredoxins, on the other hand, are reduced by the tripeptide

glutathione (gamma-Glu-Cys-Gly, known as GSH) using electrons donated by NADPH. Hence the glutathione/glutaredoxin system consists of glutaredoxin, glutathione, glutathione reductase and NADPH.

[00192] *S. cerevisiae* contains a cytoplasmic thioredoxin system comprised of the thioredoxins Trx1p and Trx2p and the thioredoxin reductase Trr1p, and a complete mitochondrial thioredoxin system comprised of the thioredoxin Trx3p and the thioredoxin reductase Trr2p. Evidence suggests that the cytoplasmic thioredoxin system may have overlapping function with the glutathione/glutaredoxin system. The mitochondrial thioredoxin system, on the other hand, does not appear to be able to substitute for either the cytoplasmic thioredoxin or glutathione/glutaredoxin systems. Instead, the mitochondrial thioredoxn proteins, thioredoxin (Trx3p) and thioredoxin reductase (Trr2p) have been implicated in the defense against oxidative stress generated during respiratory metabolism.

[00193] Overexpression of the essential cytosolic functional components of the thioredoxin system is thus predicted to increase the amount of bioavailable cytosolic thioredoxin, resulting in a significant increase in cellular redox buffering potential and concomitant increase in stable, active cytosolic FeS clusters and DHAD activity. Thus, one or more of the following genes are expressed either singly or in combination, thereby resulting in a functional increase in available thioredoxin: a thioredoxin (encoded in *S.cerevisiae* by *TRX1* and *TRX2*) and a thioredoxin reductase (encoded in *S.cerevisiae* by *TRR1*). Separately, or in combination with the aforementioned genes, the mitochondrial thioredoxin system (encoded by thioredoxin gene *TRX3* and thioredoxin reductase gene *TRR2*) are overexpressed, and, although functional in the mitochondria, provide an added or synergistic effect on FeS cluster assembly or stability, as assayed by increased DHAD activity and/or output of isobutanol in a fermentation. Overexpression of these genes may be accomplished by methods as described above. In one embodiment, active variants of any one of the aforementioned thioredoxins or thioredoxin reducatses, homologs thereof, or combination thereof are expressed in the yeast cytosol to mitigate oxidative stress.

<u>Enhancing Cytosolic DHAD Activity by Mitigating Stress Mediated by Reactive Nitrogen Species (RNS)</u>

[00194] Nitric oxide and reactive nitrogen species are highly reactive, short-lived molecules that can be generated during periods of cellular stress. The exact mechanisms by which these molecules are created, or their downstream targets, is not completely understood and is the subject of intense investigation. However, the functional groups present in many proteins -- for example, FeS clusters -- are readily attacked and inactivated by NO/RNS. Loss of these labile functional groups usually results in an inactive enzyme.

[00195] Nitric oxide and reactive nitrogen species are highly reactive, short-lived molecules that can be generated during normal cellular function, respiration, and during periods of cellular or redox stress. RNS are produced in eukaryotic cells starting with the reaction of nitric oxide ($\cdot NO$) with superoxide ($O_2 \cdot -$) to form peroxynitrite ($ONOO-$):

•NO (nitric oxide) + O2•– (super oxide) → ONOO– (peroxynitrite)
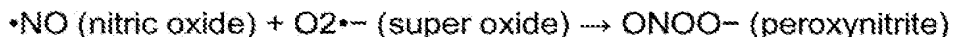
[00196] Peroxynitrite itself is a highly reactive species which can directly react with various components of the cell. Alternatively peroxynitrite can react with other molecules to form additional types of RNS including nitrogen dioxide ($•NO_2$) and dinitrogen trioxide (N_2O_3) as well as other types of chemically reactive radicals. Important reactions involving RNS include:

ONOO– + H+ → ONOOH (peroxynitrous acid) → $•NO_2$ (nitrogen dioxide) + •OH (hydroxyl radical)

ONOO– + CO_2 (carbon dioxide) → $ONOOCO_2$– (nitrosoperoxycarbonate)

$ONOOCO_2$– → $•NO_2$ (nitrogen dioxide) + O=C(O•)O– (carbonate radical)

•NO + $•NO_2$ is in equilibrium with N_2O_3 (dinitrogen trioxide)
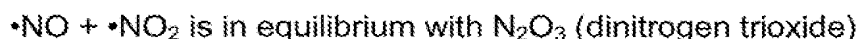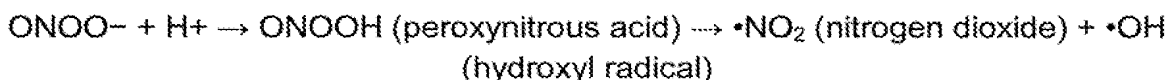
[00197] NO exhibits other types of interaction that are candidates for mediating aspects of its physiological action. Notably, in a process known as nitrosylation, or nitrosation, NO can modify free sulfydryl (thiol) groups of cysteines in proteins to produce nitrosothiols, SNOs. Transfer of the NO adduct from one sulfydryl to another transnitrosylation) is likely to play a signal transduction role (reviewed in Stamler *et al.*,2001). Study of this post-translational modification, which is proposed to be a widespread mediator of signaling, is a relatively new field, and the list of proteins that are modified through nitrosylation is expanding rapidly. Because NO is highly reactive, transport of an NO signal in tissues can be facilitated through reaction with glutathione and movement of the resulting *S*-nitrosoglutathione (GSNO), which can subsequently signal by modifying thiol groups on target proteins by transnitrosylation (Lipton *et al.*, 2001; Foster *et al.*, 2003). The discovery of GSNO reductase (GSNOR), which reduces GSNO to restore GSH and to eliminate the NO adduct as NH^{4+} (Jensen *et al.*, 1998), revealed the importance of the control of this NO metabolite.

[00198] The exact mechanisms by which the aforementioned molecules are generated, or their downstream targets, are not completely understood and are the subject of intense investigation. However, the functional groups present in many proteins -- for example, FeS clusters -- are readily attacked by NO/RNS. The enzyme dihydroxyacid dehydratase (DHAD) contains an iron-sulfur (FeS) cluster cofactor that is sensitive to damage by NO or RNS. As an example of the biological sensitivity of this class of enzyme to attack by NO/RNS, inactivation of the *E.coli* DHAD (encoded by *ilvD*) and subsequent bacterial cell death resulting from macrophage-generated NO is a major component of the mammalian humoral immune response.

[00199] The present invention provides methods of mitigating the potentially harmful effects of oxidative and nitrosative stress (e.g., NO and/or or RNS) on

enzymes involved in the production of isobutanol in the yeast cytosol. Specifically, the enzyme dihydroxyacid dehydratase (DHAD) contains an iron-sulfur (Fe-S) cluster that is sensitive to damage by NO and/or RNS, leading to inactive enzyme. Strategies of mitigating such harmful effects include, but are not limited to, increasing repair of iron-sulfur clusters damaged by oxidative and nitrosative stress conditions; reducing nitric oxide levels by introduction of a nitric oxide reductase (NOR) activity in the cell; reducing the levels of SNO's by overexpression of a GSNO-reductase; or combination thereof.

[00200] Strategies disclosed herein are intended to protect or repair DHAD from NO/RNS damage. Accordingly, in one embodiment, the present invention provides recombinant microorganisms that have been engineered to express one or more enzymes in the cytosol that reduce the concentration of reactive nitrogen species (RNS) and/or nitric oxide in said cytosol.

[00201] In one embodiment, the present invention provides recombinant microorganisms that have been engineered to express a nitric oxide reductase that reduce the concentration of reactive nitrogen species (RNS) and/or nitric oxide in said cytosol. To reduce nitric oxide levels in the yeast cytosol, one or more nitric oxide reductases (NORs) or active variants thereof can be introduced into the cell by overexpression. Genes present in several microbial species have been shown to encode a nitric oxide reductase activity. For example, in *E.coli* the gene for a flavorubredoxin, *norV*, encodes a flavo-diiron NO reductase that is one of the most highly induced genes when *E.coli* cells are exposed to NO or GSNO. Previous work has identified a gene present in the microbe *Fusarium oxysporum* as encoding a cytochrome P-450 55A1 (P-450dNIR) that encodes a nitric oxide reductase (Nakahara *et al.*, 1993, *J. Biol. Chem.* 268:8350–8355). When expressed in a eukaryotic cell, this gene product appears to be cytosolically localized and exhibits effects consistent with its reducing intracellular NO levels (Dijkers *et al.*, 2009, *Molecular Biology of the Cell*, 20: 4083-4090). Thus, in one embodiment, homologs of any above-described nitric oxide reductases, active variants thereof, or combinations thereof are expressed in the yeast cytosol to mitigate nitric oxide.

[00202] In contrast to *E. coli* and *F. oxysporum*, *S. cerevisiae* lacks an endogenous NOR activity (and no homologs of either NOR protein is found in the *S. cerevisiae* genome). Thus, to provide such an activity, the *F. oxysporum* NOR gene is synthesized or amplified from genomic DNA, or the *E. coli norV* gene is amplified from genomic DNA, and either (or both) cloned into a suitable yeast expression vector. Such a vector could either be high copy (e.g., 2micron origin) or low copy (CEN/ARSH), or a single or multiple copies of the gene could be stably integrated into the genome of a host organism, specifically a yeast containing a cytosolic isobutanol pathway. In each case, methods to clone a gene into a plasmid so that it is expressed at a desired level under the control of a known yeast promoter (including those steps required to transform a host yeast cell) are well known to those skilled in the art. In those cases where the NOR gene is expressed from an episomal plasmid, it can be advantageous to simultaneous overexpress a desired DHAD gene, either from the same or from another plasmid, thereby allowing one to assay the resulting output in DHAD activity. Similar approaches are undertaken to

express the NOR gene in the presence of a plasmid(s) encoding an isobutanol production pathway, where the resuls of NOR expression are manifested in changes in isobutanol productivity, titer, or yield. It is understood by one skilled in the art that expression of all genes, both NOR and genes encoding the isobutanol pathway may be integrated into the genome of a host organism in a single or multiple copies of the gene(s), specifically a yeast containing a cytosolic isobutanol pathway.

[00203] In another embodiment, the present invention provides recombinant microorganisms that have been engineered to express a glutathione-S-nitrosothiol reductase (GSNO-reductase) that reduces the concentration of reactive nitrogen species (RNS) and/or nitric oxide in said cytosol. To reduce the levels of SNO's, one or more GSNO-reductases or active variants thereof can be introduced into the cell by overexpression. In *S. cerevisiae*, the gene *SFA1* has been shown to encode a formaldehyde dehydrogenase that possesses GSNO reductase activity (Liu *et al.*, 2001, *Nature 410*:490–494). Sfa1p is a member of the class III alcohol dehydrogenases (EC:1.1.1.284), which are bifunctional enzymes containing both alcohol dehydrogenase and glutathione-dependent formaldehyde dehydrogenase activities. The glutathione-dependent formaldehyde dehydrogenase activity of Sfa1p is required for the detoxification of formaldehyde, and the alcohol dehydrogenase activity of Sfa1p can catalyze the final reactions in phenylalanine and tryptophan degradation. Sfa1p is also able to act as a hydroxymethylfurfural (HMF) reductase and catabolize HMF, a compound formed in the production of certain biofuels. Sfa1p has been localized to the cytoplasm and the mitochondria, and can act on a variety of substrates, including S-hydroxymethylglutathione, phenylacetaldehyde, indole acetaldehyde, octanol, 10-hydroxydecanoic acid, 12-hydroxydodecanoic acid, and S-nitrosoglutathione.

[00204] Sfa1 protein levels are reported as being low-to-moderate from proteome-wide analyses (Ghaemmaghami *et al.*, 2003, *Nature 425*(6959):737-41). Thus, in an analagous fashion to the approach described for overexpression of NOR, the gene *SFA1* is overexpressed, thereby decoupling it from its normal regulatory control and permitting significant increase in Sfa1 activity in the cell, which results in measureable increases in DHAD activity and/or fermentation output, as described above. Overexpression of these genes may be accomplished by methods as described above. In one embodiment, homologs of *SFA1*, active variants thereof, or combinations thereof are expressed in the yeast cytosol to mitigate stresses brought on by reactive nitrogen species.

[00205] In additional embodiments, alternative enzymes may be expressed and targeted to the yeast cytosol containing the isobutanol pathway to mitigate the effects of reactive nitrogen species. Specifically, the enzyme YtfE encoded by *E.coli ytfE*, homologs thereof, active variants thereof, may be expressed, such that they lead to reduction in NO/RNS in the yeast cytosol and/or a concomitant increase in DHAD function. Such an increase is detected by in vitro assay of DHAD activity, and/or by an increase in productivity, titer, or yield of isobutanol produced by isobutanol pathway-containing cells.

[00206] To increase repairment of iron-sulfur clusters, in one embodiment, the gene *ytfE* from *E.coli* is expressed in the yeast cytosol which contains a functional

isobutanol pathway and DHAD such that DHAD activity and/or isobutanol productivity, titer, or yield are increased from the yeast cells. In *E. coli*, the gene *ytfE* has been shown to play an important role in maintaining active Fe-S clusters. A recent report (Justino et al., (2009). Escherichia coli Di-iron YtfE protein is necessary for the repair of stress-damaged Iron-Sulfur Clusters. *JBC* 282(14): 10352-10359) showed that Δ*ytfE* strains have several phenotypes, including enhanced susceptibility to nitrosative stress and are defective in the activity of several iron-sulfur-containing proteins. For example, the damage of the $[4Fe-4S]^{2+}$ clusters of aconitase B and fumarase A caused by exposure to hydrogen peroxide and nitric oxide stress occurs at higher rates in the absence of *ytfE*. The *ytfE* null mutation also abolished the recovery of aconitase and fumarase activities, which is observed in wild type *E. coli* once the stress is scavenged. Notably, upon the addition of purified holo-YtfE protein to mutant cell extracts, the enzymatic activities of fumarase and aconitase were fully recovered, and at rates similar to the wild type strain. Thus, YtfE is critical for the repair of iron-sulfur clusters damaged by oxidative and nitrosative stress conditions, and presents an attractive candidate for overexpression in a host cell that normally lacks this activity, such as *S. cerevisiae*, where Fe-S cluster proteins are also being overexpressed as part of the isobutanol pathway.

[00207] To provide such an activity, the *E.coli ytfE* gene can be amplified from genomic DNA by PCR with appropriate primers, and cloned into a suitable yeast expression vector. Such a vector could either be high copy (e.g., 2micron origin) or low copy (CEN/ARSH), or a single or multiple copies of the gene could be stably integrated into the genome of a host organism. In each case, methods to clone a gene into a plasmid so that it is expressed at a desired level under the control of a known yeast promoter (including those steps required to transform a host yeast cell) are well known to those skilled in the art. In those cases where the *ytfE* gene is expressed from an episomal plasmid, it can be advantageous to simultaneous overexpress a desired DHAD gene, either from the same or from another plasmid, thereby allowing one to assay the resulting output in DHAD activity. Similar approaches are undertaken to express the *ytfE* gene in the presence of a plasmid(s) encoding an isobutanol production pathway, where the results of *ytfE* expression are manifested in changes in isobutanol productivity, titer, or yield. More specifically, ytfE is expressed in the yeast cytosol which contains a functional isobutanol pathway and DHAD such that DHAD activity and/or isobutanol productivity, titer, or yield are increased from the yeast cells.

[00208] In addition, functional homologs of *E.coli* ytfE have been identified and characterized. For example, genes from two pathogenic prokaryotes—*scdA* from *Staphylococcus aureus*, and *dnrN* from *Neisseria gonorrhoeae*, have been shown to have properties similar to that of ytfE (Overton, T.W., et al (2008). Widespread distribution in pathogenic bacteria of di-iron proteins that repair oxidative and nitrosative damage to iron-sulfur centers. *J. Bacteriology* 190(6): 2004-2013). Thus, similar approaches to overexpress either of these genes are employed, as described for *E.coli ytfE*, above. Overexpression of these genes may be accomplished by methods as described above.

The Microorganism in General

[00209] Native producers of 1-butanol, such as *Clostridium acetobutylicum*, are known, but these organisms also generate byproducts such as acetone, ethanol, and butyrate during fermentations. Furthermore, these microorganisms are relatively difficult to manipulate, with significantly fewer tools available than in more commonly used production hosts such as *S. cerevisiae* or *E. coli*. Additionally, the physiology and metabolic regulation of these native producers are much less well understood, impeding rapid progress towards high-efficiency production. Furthermore, no native microorganisms have been identified that can metabolize glucose into isobutanol in industrially relevant quantities.

[00210] The production of isobutanol and other fusel alcohols by various yeast species, including *S. cerevisiae* is of special interest to the distillers of alcoholic beverages, for whom fusel alcohols constitute often undesirable off-notes. Production of isobutanol in wild-type yeasts has been documented on various growth media, ranging from grape must from winemaking (Romano, *et al.*, Metabolic diversity of *Saccharomyces cerevisiae* strains from spontaneously fermented grape musts, World Journal of Microbiology and Biotechnology. 19:311-315, 2003), in which 12-219 mg/L isobutanol were produced, to supplemented minimal media (Oliviera, *et al.* (2005) World Journal of Microbiology and Biotechnology 21:1569-1576), producing 16-34 mg/L isobutanol. Work from Dickinson, *et al.* (J Biol Chem. 272(43):26871-8, 1997) has identified the enzymatic steps utilized in an endogenous *S. cerevisiae* pathway converting branch-chain amino acids (e.g., valine or leucine) to isobutanol.

[00211] Recombinant microorganisms provided herein can express a plurality of heterologous and/or native target enzymes involved in pathways for the production of isobutanol from a suitable carbon source.

[00212] Accordingly, "engineered" or "modified" microorganisms are produced via the introduction of genetic material into a host or parental microorganism of choice and/or by modification of the expression of native genes, thereby modifying or altering the cellular physiology and biochemistry of the microorganism. Through the introduction of genetic material and/or the modification of the expression of native genes the parental microorganism acquires new properties, *e.g.* the ability to produce a new, or greater quantities of, an intracellular metabolite. As described herein, the introduction of genetic material into and/or the modification of the expression of native genes in a parental microorganism results in a new or modified ability to produce isobutanol. The genetic material introduced into and/or the genes modified for expression in the parental microorganism contains gene(s), or parts of genes, coding for one or more of the enzymes involved in a biosynthetic pathway for the production of isobutanol and may also include additional elements for the expression and/or regulation of expression of these genes, *e.g.* promoter sequences.

[00213] In addition to the introduction of a genetic material into a host or parental microorganism, an engineered or modified microorganism can also include alteration, disruption, deletion or knocking-out of a gene or polynucleotide to alter the cellular physiology and biochemistry of the microorganism. Through the alteration,

disruption, deletion or knocking-out of a gene or polynucleotide the microorganism acquires new or improved properties (e.g., the ability to produce a new metabolite or greater quantities of an intracellular metabolite, improve the flux of a metabolite down a desired pathway, and/or reduce the production of byproducts).

[00214] Recombinant microorganisms provided herein may also produce metabolites in quantities not available in the parental microorganism. A "metabolite" refers to any substance produced by metabolism or a substance necessary for or taking part in a particular metabolic process. A metabolite can be an organic compound that is a starting material (e.g., glucose or pyruvate), an intermediate (e.g., 2-ketoisovalerate), or an end product (e.g., isobutanol) of metabolism. Metabolites can be used to construct more complex molecules, or they can be broken down into simpler ones. Intermediate metabolites may be synthesized from other metabolites, perhaps used to make more complex substances, or broken down into simpler compounds, often with the release of chemical energy.

[00215] Exemplary metabolites include glucose, pyruvate, and isobutanol. The metabolite isobutanol can be produced by a recombinant microorganism which expresses or over-expresses a metabolic pathway that converts pyruvate to isobutanol. An exemplary metabolic pathway that converts pyruvate to isobutanol may be comprised of an acetohydroxy acid synthase (ALS), a ketolacid reductoisomerase (KARI), a dihyroxy-acid dehydratase (DHAD), a 2-keto-acid decarboxylase (KIVD), and an alcohol dehydrogenase (ADH).

[00216] Accordingly, provided herein are recombinant microorganisms that produce isobutanol and in some aspects may include the elevated expression of target enzymes such as ALS, KARI, DHAD, KIVD, and ADH

[00217] The disclosure identifies specific genes useful in the methods, compositions and organisms of the disclosure; however it will be recognized that absolute identity to such genes is not necessary. For example, changes in a particular gene or polynucleotide comprising a sequence encoding a polypeptide or enzyme can be performed and screened for activity. Typically such changes comprise conservative mutation and silent mutations. Such modified or mutated polynucleotides and polypeptides can be screened for expression of a functional enzyme using methods known in the art.

[00218] Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or functionally equivalent polypeptides can also be used to clone and express the polynucleotides encoding such enzymes.

[00219] As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The genetic code is redundant with 64 possible codons, but most organisms typically use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons. Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias."

[00220] Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host (see also, Murray et al. (1989) Nucl. Acids Res.

17:477-508) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, typical stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The typical stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* commonly use UAA as the stop codon (Dalphin *et al.* (1996) Nucl. Acids Res. 24: 216-218). Methodology for optimizing a nucleotide sequence for expression in a plant is provided, for example, in U.S. Pat. No. 6,015,891, and the references cited therein.

[00221] Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of DNA compounds differing in their nucleotide sequences can be used to encode a given enzyme of the disclosure. The native DNA sequence encoding the biosynthetic enzymes described above are referenced herein merely to illustrate an embodiment of the disclosure, and the disclosure includes DNA compounds of any sequence that encode the amino acid sequences of the polypeptides and proteins of the enzymes utilized in the methods of the disclosure. In similar fashion, a polypeptide can typically tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or significant loss of a desired activity. The disclosure includes such polypeptides with different amino acid sequences than the specific proteins described herein so long as they modified or variant polypeptides have the enzymatic anabolic or catabolic activity of the reference polypeptide. Furthermore, the amino acid sequences encoded by the DNA sequences shown herein merely illustrate embodiments of the disclosure.

[00222] In addition, homologs of enzymes useful for generating metabolites are encompassed by the microorganisms and methods provided herein.

[00223] As used herein, two proteins (or a region of the proteins) are substantially homologous when the amino acid sequences have at least about 30%, 40%, 50% 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity. To determine the per*cent* identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In one embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, typically at least 40%, more typically at least 50%, even more typically at least 60%, and even more typically at least 70%, 80%, 90%, 100% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The per*cent* identity between the two sequences is a function of the number of identical positions shared by the

sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

[00224] When "homologous" is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of homology may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art (see, e.g., Pearson W.R. Using the FASTA program to search protein and DNA sequence databases, Methods in Molecular Biology,1994, 25:365-89, hereby incorporated herein by reference).

[00225] The following six groups each contain amino acids that are conservative substitutions for one another: 1) Serine (S), Threonine (T); 2) Aspartic Acid (D), Glutamic Acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

[00226] Sequence homology for polypeptides, which is also referred to as percent sequence identity, is typically measured using sequence analysis software. See, e.g., the Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wis. 53705. Protein analysis software matches similar sequences using measure of homology assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutant protein thereof. See, e.g., GCG Version 6.1.

[00227] A typical algorithm used comparing a molecule sequence to a database containing a large number of sequences from different organisms is the computer program BLAST (Altschul, S.F., et al. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410; Gish, W. and States, D.J. (1993) "Identification of protein coding regions by database similarity search." Nature Genet. 3:266-272; Madden, T.L., et al. (1996) "Applications of network BLAST server" Meth. Enzymol. 266:131-141; Altschul, S.F., et al. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25:3389-3402; Zhang, J. and Madden, T.L. (1997) "PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation." Genome Res. 7:649-656), especially blastp or tblastn (Altschul, S.F., et al. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25:3389-3402). Typical parameters for BLASTp are: Expectation value: 10

(default); Filter: seg (default); Cost to open a gap: 11 (default); Cost to extend a gap: 1 (default); Max. alignments: 100 (default); Word size: 11 (default); No. of descriptions: 100 (default); Penalty Matrix: BLOSUM62.

[00228] When searching a database containing sequences from a large number of different organisms, it is typical to compare amino acid sequences. Database searching using amino acid sequences can be measured by algorithms other than blastp known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, W.R. (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA" Meth. Enzymol. 183:63-98). For example, a percent sequence identity between amino acid sequences can be determined using FASTA with its default parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1, hereby incorporated herein by reference.

[00229] The disclosure provides recombinant microorganisms comprising a biochemical pathway for the production of isobutanol from a suitable substrate at a high yield. A recombinant microorganism of the disclosure comprises one or more recombinant polynucleotides within the genome of the organism or external to the genome within the organism. The microorganism can comprise a reduction in expression, disruption or knockout of a gene found in the wild-type organism and/or introduction of a heterologous polynucleotide and/or expression or overexpression of an endogenous polynucleotide.

[00230] In one aspect, the disclosure provides a recombinant microorganism comprising elevated expression of at least one target enzyme as compared to a parental microorganism or encodes an enzyme not found in the parental organism. In another or further aspect, the microorganism comprises a reduction in expression, disruption or knockout of at least one gene encoding an enzyme that competes with a metabolite necessary for the production of isobutanol. The recombinant microorganism produces at least one metabolite involved in a biosynthetic pathway for the production of isobutanol. In general, the recombinant microorganisms comprises at least one recombinant metabolic pathway that comprises a target enzyme and may further include a reduction in activity or expression of an enzyme in a competitive biosynthetic pathway. The pathway acts to modify a substrate or metabolic intermediate in the production of isobutanol. The target enzyme is encoded by, and expressed from, a polynucleotide derived from a suitable biological source. In some embodiments, the polynucleotide comprises a gene derived from a prokaryotic or eukaryotic source and recombinantly engineered into the microorganism of the disclosure. In other embodiments, the polynucleotide comprises a gene that is native to the host organism.

[00231] It is understood that a range of microorganisms can be modified to include a recombinant metabolic pathway suitable for the production of isobutanol. In various embodiments, microorganisms may be selected from yeast microorganisms. Yeast microorganisms for the production of isobutanol may be selected based on certain characteristics:

[00232] One characteristic may include the property that the microorganism is

selected to convert various carbon sources into isobutanol. The term "carbon source" generally refers to a substance suitable to be used as a source of carbon for prokaryotic or eukaryotic cell growth. Carbon sources include, but are not limited to, biomass hydrolysates, starch, sucrose, cellulose, hemicellulose, xylose, and lignin, as well as monomeric components of these substrates. Carbon sources can comprise various organic compounds in various forms, including, but not limited to polymers, carbohydrates, acids, alcohols, aldehydes, ketones, amino acids, peptides, etc. These include, for example, various monosaccharides such as glucose, dextrose (D-glucose), maltose, oligosaccharides, polysaccharides, saturated or unsaturated fatty acids, succinate, lactate, acetate, ethanol, etc., or mixtures thereof. Photosynthetic organisms can additionally produce a carbon source as a product of photosynthesis. In some embodiments, carbon sources may be selected from biomass hydrolysates and glucose. The term "biomass" as used herein refers primarily to the stems, leaves, and starch-containing portions of green plants, and is mainly comprised of starch, lignin, cellulose, hemicellulose, and/or pectin. Biomass can be decomposed by either chemical or enzymatic treatment to the monomeric sugars and phenols of which it is composed (Wyman, C.E. 2003 Biotechnological Progress 19:254-62). This resulting material, called biomass hydrolysate, is neutralized and treated to remove trace amounts of organic material that may adversely affect the biocatalyst, and is then used as a feed stock for fermentations using a biocatalyst.

[00233] Accordingly, in one embodiment, the recombinant microorganism herein disclosed can convert a variety of carbon sources to products, including but not limited to glucose, galactose, mannose, xylose, arabinose, lactose, sucrose, and mixtures thereof.

[00234] The recombinant microorganism may thus further include a pathway for the fermentation of isobutanol from five-carbon (pentose) sugars including xylose. Most yeast species metabolize xylose via a complex route, in which xylose is first reduced to xylitol via a xylose reductase (XR) enzyme. The xylitol is then oxidized to xylulose via a xylitol dehydrogenase (XDH) enzyme. The xylulose is then phosphorylated via an xylulokinase (XK) enzyme. This pathway operates inefficiently in yeast species because it introduces a redox imbalance in the cell. The xylose-to-xylitol step uses NADH as a cofactor, whereas the xylitol-to-xylulose step uses NADPH as a cofactor. Other processes must operate to restore the redox imbalance within the cell. This often means that the organism cannot grow anaerobically on xylose or other pentose sugar. Accordingly, a yeast species that can efficiently ferment xylose and other pentose sugars into a desired fermentation product is therefore very desirable.

[00235] Thus, in one aspect, the recombinant is engineered to express a functional exogenous xylose isomerase. Exogenous xylose isomerases functional in yeast are known in the art. See, e.g., Rajgarhia *et al*, US20060234364, which is herein incorporated by reference in its entirety. In an embodiment according to this aspect, the exogenous xylose isomerase gene is operatively linked to promoter and terminator sequences that are functional in the yeast cell. In a preferred embodiment, the recombinant microorganism further has a deletion or disruption of a

native gene that encodes for an enzyme (e.g. XR and/or XDH) that catalyzes the conversion of xylose to xylitol. In a further preferred embodiment, the recombinant microorganism also contains a functional, exogenous xylulokinase (XK) gene operatively linked to promoter and terminator sequences that are functional in the yeast cell. In one embodiment, the xylulokinase (XK) gene is overexpressed.

[00236] In one embodiment, the microorganism has reduced or no pyruvate decarboxylase (PDC) activity. PDC catalyzes the decarboxylation of pyruvate to acetaldehyde, which is then reduced to ethanol by ADH via an oxidation of NADH to NAD+. Ethanol production is the main pathway to oxidize the NADH from glycolysis. Deletion of this pathway increases the pyruvate and the reducing equivalents (NADH) available for the isobutanol pathway. Accordingly, deletion of *PDC* genes can further increase the yield of isobutanol.

[00237] In another embodiment, the microorganism has reduced or no glycerol-3-phosphate dehydrogenase (GPD) activity. GPD catalyzes the reduction of dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate (G3P) via the oxidation of NADH to NAD+. Glycerol is then produced from G3P by Glycerol-3-phosphatase (GPP). Glycerol production is a secondary pathway to oxidize excess NADH from glycolysis. Reduction or elimination of this pathway would increase the pyruvate and reducing equivalents (NADH) available for the isobutanol pathway. Thus, deletion of *GPD* genes can further increase the yield of isobutanol.

[00238] In yet another embodiment, the microorganism has reduced or no PDC activity and reduced or no GPD activity.

[00239] In one embodiment, the yeast microorganisms may be selected from the "*Saccharomyces* Yeast Clade", defined as an ascomycetous yeast taxonomic class by Kurtzman and Robnett in 1998 ("Identification and phylogeny of ascomycetous yeast from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences." Antonie van Leeuwenhoek 73: 331-371, See Figure 2 of Leeuwenhoek reference). They were able to determine the relatedness of approximately 500 yeast species by comparing the nucleotide sequence of the D1/D2 domain at the 5' end of the gene encoding the large ribosomal subunit 26S. In pair-wise comparisons of the D1/D2 nucleotide sequences of *S. cerevisiae* and of the two most distant yeast from this *Saccharomyces* yeast clade, *K. lactis* and *K. marxianus*, share greater than 80% identity.

[00240] The term "Saccharomyces *sensu stricto*" taxonomy group is a cluster of yeast species that are highly related to *S. cerevisiae* (Rainieri, S. *et al* 2003. Saccharomyces *Sensu Stricto*: Systematics, Genetic Diversity and Evolution. J. Biosci Bioengin 96(1)1-9. Saccharomyces *sensu stricto* yeast species include but are not limited to *S. cerevisiae, S. cerevisiae, S. kudriavzevii, S. mikatae, S. bayanus, S. uvarum, S. carocanis* and hybrids derived from these species (Masneuf *et al.* 1998. New Hybrids between Saccharomyces *Sensu Stricto* Yeast Species Found Among Wine and Cider Production Strains. Yeast 7(1)61-72).

[00241] An ancient whole genome duplication (WGD) event occurred during the evolution of the hemiascomycete yeast and was discovered using comparative genomic tools (Kellis *et al* 2004 "Proof and evolutionary analysis of ancient genome duplication in the yeast *S. cerevisiae.*" Nature 428:617-624. Dujon *et al* 2004

103182 v1/DC

"Genome evolution in yeasts." Nature 430:35-44. Langkjaer *et al* 2003 "Yeast genome duplication was followed by asynchronous differentiation of duplicated genes." Nature 428:848-852. Wolfe and Shields 1997 "Molecular evidence for an ancient duplication of the entire yeast genome." Nature 387:708-713.) Using this major evolutionary event, yeast can be divided into species that diverged from a common ancestor following the WGD event (termed "post-WGD yeast" herein) and species that diverged from the yeast lineage prior to the WGD event (termed "pre-WGD yeast" herein).

[00242] Accordingly, in one embodiment, the yeast microorganism may be selected from a post-WGD yeast genus, including but not limited to Saccharomyces and Candida. The favored post-WGD yeast species include: *S. cerevisiae, S. uvarum, S. bayanus, S. paradoxus, S. castelli,* and *C. glabrata.*

[00243] In another embodiment, the yeast microorganism may be selected from a pre-whole genome duplication (pre-WGD) yeast genus including but not limited to Saccharomyces, Kluyveromyces, Candida, Pichia, Issatchenkia, Debaryomyces, Hansenula, Yarrowia and, Schizosaccharomyces. Representative pre-WGD yeast species include: *S. kluyveri, K. thermotolerans, K. marxianus, K. waltii, K. lactis, C. tropicalis, P. pastoris, P. anomala, P. stipitis, I. orientalis, I. occidentalis, I. scutulata, D. hansenii, H. anomala, Y. lipolytica,* and *S. pombe.*

[00244] A yeast microorganism may be either Crabtree-negative or Crabtree-positive. A yeast cell having a Crabtree-negative phenotype is any yeast cell that does not exhibit the Crabtree effect. The term "Crabtree-negative" refers to both naturally occurring and genetically modified organisms. Briefly, the Crabtree effect is defined as the inhibition of oxygen consumption by a microorganism when cultured under aerobic conditions due to the presence of a high concentration of glucose (e.g., 50 g-glucose L^{-1}). In other words, a yeast cell having a Crabtree-positive phenotype continues to ferment irrespective of oxygen availability due to the presence of glucose, while a yeast cell having a Crabtree-negative phenotype does not exhibit glucose mediated inhibition of oxygen consumption.

[00245] Accordingly, in one embodiment the yeast microorgnanism may be selected from yeast with a Crabtree-negative phenotype including but not limited to the following genera: Kluyveromyces, Pichia, Issatchenkia, Hansenula, and Candida. Crabtree-negative species include but are not limited to: *K. lactis, K. marxianus, P. anomala, P. stipitis, I. orientalis, I. occidentalis, I. scutulata,, H. anomala,* and *C. utilis.*

[00246] In another embodiment, the yeast microorganism may be selected from a yeast with a Crabtree-positive phenotype, including but not limited to Saccharomyces, Kluyveromyces, Zygosaccharomyces, Debaryomyces, Pichia and Schizosaccharomyces. Crabtree-positive yeast species include but are not limited to: *S. cerevisiae, S. uvarum, S. bayanus, S. paradoxus, S. castelli, S. kluyveri, K. thermotolerans, C. glabrata, Z. bailli, Z. rouxii, D. hansenii, P. pastorius,* and *S. pombe.*

[00247] Another characteristic may include the property that the microorganism is that it is non-fermenting. In other words, it cannot metabolize a carbon source anaerobically while the yeast is able to metabolize a carbon source in the presence

of oxygen. Nonfermenting yeast refers to both naturally occurring yeasts as well as genetically modified yeast. During anaerobic fermentation with fermentative yeast, the main pathway to oxidize the NADH from glycolysis is through the production of ethanol. Ethanol is produced by alcohol dehydrogenase (ADH) via the reduction of acetaldehyde, which is generated from pyruvate by pyruvate decarboxylase (PDC). In one embodiment, a fermentative yeast can be engineered to be non-fermenative by the reduction or elimination of the native PDC activity. Thus, most of the pyruvate produced by glycolysis is not consumed by PDC and is available for the isobutanol pathway. Deletion of this pathway increases the pyruvate and the reducing equivalents available for the isobutanol pathway. Fermentative pathways contribute to low yield and low productivity of isobutanol. Accordingly, deletion of *PDC* may increase yield and productivity of isobutanol.

[00248] In some embodiments, the recombinant microorganisms may be microorganisms that are non-fermenting yeast microorganisms, including, but not limited to those, classified into a genera selected from the group consisting of *Tricosporon, Rhodotorula,* or *Myxozyma.*

[00249] In one embodiment, a yeast microorganism is engineered to convert a carbon source, such as glucose, to pyruvate by glycolysis and the pyruvate is converted to isobutanol via an engineered isobutanol pathway (PCT/US2006/041602, PCT/US2008/053514, Atsumi *et al., Nature,* 2008 Jan 3;451(7174):86-9). Alternative pathways for the production of isobutanol have been described in International Patent Application No PCT/US2006/041602 and in Dickinson *et al., Journal of Biological Chemistry* 273:25751-15756 (1998).

[00250] Accordingly, the engineered isobutanol pathway to convert pyruvate to isobutanol can be comprised of the following reactions:

1. 2 pyruvate $\rightarrow$ acetolactate + CO_2
2. acetolactate + NADPH $\rightarrow$ 2,3-dihydroxyisovalerate + $NADP^+$
3. 2,3-dihydroxyisovalerate $\rightarrow$ alpha-ketoisovalerate
4. alpha-ketoisovalerate $\rightarrow$ isobutyraldehyde + CO_2
5. isobutyraldehyde +NADPH $\rightarrow$ isobutanol + NADP

[00251] These reactions are carried out by the enzymes 1) Acetolactate Synthase (ALS, EC4.1.3.18), 2) Keto-acid Reducto-Isomerase (KARI, EC1.1.1.86), 3) Dihydroxy-acid dehydratase (DHAD, EC4.2.1.9), 4) Keto-isovalerate decarboxylase (KIVD, EC4.1.1.1), and 5) an Alcohol dehydrogenase (ADH, EC1.1.1.1 or 1.1.1.2) (Figure 1).

[00252] In another embodiment, the yeast microorganism is engineered to overexpress these enzymes. For example, these enzymes can be encoded by native genes. Alternatively, these enzymes can be encoded by heterologous genes. For example, ALS can be encoded by the *alsS* gene of *B. subtilis, alsS* of *L. lactis,* or the *ilvK* gene of *K. pneumonia.* For example, KARI can be encoded by the *ilvC* genes of *E. coli, C. glutamicum, M. maripaludis,* or *Piromyces sp E2.* For example, DHAD can be encoded by the *ilvD* genes of *E. coli* or *C. glutamicum.* For example, KIVD can be encoded by the *kivD* gene of *L. lactis.* ADH can be encoded by *ADH2, ADH6,* or ADH7 of *S. cerevisiae.*

[00253] The yeast microorganism of the invention may be engineered to have increased ability to convert pyruvate to isobutanol. In one embodiment, the yeast microorganism may be engineered to have increased ability to convert pyruvate to isobutyraldehyde. In another embodiment, the yeast microorganism may be engineered to have increased ability to convert pyruvate to keto-isovalerate. In another embodiment, the yeast microorganism may be engineered to have increased ability to convert pyruvate to 2,3-dihydroxyisovalerate. In another embodiment, the yeast microorganism may be engineered to have increased ability to convert pyruvate to acetolactate.

[00254] Furthermore, any of the genes encoding the foregoing enzymes (or any others mentioned herein (or any of the regulatory elements that control or modulate expression thereof)) may be optimized by genetic/protein engineering techniques, such as directed evolution or rational mutagenesis, which are known to those of ordinary skill in the art. Such action allows those of ordinary skill in the art to optimize the enzymes for expression and activity in yeast.

[00255] In addition, genes encoding these enzymes can be identified from other fungal and bacterial species and can be expressed for the modulation of this pathway. A variety of organisms could serve as sources for these enzymes, including, but not limited to, *Saccharomyces* spp., including *S. cerevisiae* and *S. uvarum*, *Kluyveromyces* spp., including *K. thermotolerans*, *K. lactis*, and *K. marxianus*, *Pichia* spp., *Hansenula* spp., including *H. polymorpha*, *Candida* spp., *Trichosporon* spp., *Yamadazyma* spp., including *Y.* spp. *stipitis*, *Torulaspora pretoriensis*, *Schizosaccharomyces* spp., including *S. pombe*, *Cryptococcus* spp., *Aspergillus* spp., *Neurospora* spp., or *Ustilago* spp. Sources of genes from anaerobic fungi include, but not limited to, *Piromyces* spp., *Orpinomyces* spp., or *Neocallimastix* spp. Sources of prokaryotic enzymes that are useful include, but not limited to, *Escherichia. coli*, *Zymomonas mobilis*, *Staphylococcus aureus*, *Bacillus* spp., *Clostridium* spp., *Corynebacterium* spp., *Pseudomonas* spp., *Lactococcus* spp., *Enterobacter* spp., and *Salmonella* spp.

Methods in General

Identification of DHAD in a microorganism

Any method can be used to identify genes that encode for enzymes with dihydroxyacid dehydratase (DHAD) activity. DHAD catalyzes the conversion of 2,3-dihydroxyisovalerate to ketoisovalerate. Generally, genes that are homologous or similar to a known DHAD gene, e.g. *S. cerevisiae ILV3* (encoding for SEQ ID NO. 11) or *L. lactis ilvD* (encoding for SEQ ID NO. 9) enzymes can be identified by functional, structural, and/or genetic analysis. In most cases, homologous or similar *DHAD* genes and/or homologous or similar DHAD enzymes will have functional, structural, or genetic similarities. Techniques known to those skilled in the art may be suitable to identify homologous genes and homologous enzymes. Generally, analogous genes and/or analogous enzymes can be identified by functional analysis and will have functional similarities. Techniques known to those skilled in the art may be suitable to identify analogous genes and analogous enzymes. For example, to

identify homologous or analogous genes, proteins, or enzymes, techniques may include, but not limited to, cloning a *DHAD* gene by PCR using primers based on a published sequence of a gene/enzyme or by degenerate PCR using degenerate primers designed to amplify a conserved region among *DHAD* genes. Further, one skilled in the art can use techniques to identify homologous or analogous genes, proteins, or enzymes with functional homology or similarity. Techniques include examining a cell or cell culture for the catalytic activity of an enzyme through *in vitro* enzyme assays for said activity (e.g. as described herein or in Kiritani, K. *Branched-Chain Amino Acids* Methods Enzymology, 1970), then isolating the enzyme with said activity through purification, determining the protein sequence of the enzyme through techniques such as Edman degradation, design of PCR primers to the likely nucleic acid sequence, amplification of said DNA sequence through PCR, and cloning of said nucleic acid sequence. To identify homologous or similar genes and/or homologous or similar enzymes, analogous genes and/or analogous enzymes or proteins, techniques also include comparison of data concerning a candidate gene or enzyme with databases such as BRENDA, KEGG, or MetaCYC. The candidate gene or enzyme may be identified within the above mentioned databases in accordance with the teachings herein.

Identification of PDC in a yeast microorganism

[00256] Any method can be used to identify genes that encode for enzymes with pyruvate decarboxylase (PDC) activity. PDC catalyzes the decarboxylation of pyruvate to form acetaldehyde. Generally, homologous or similar *PDC* genes *and/or homologo*us or similar PDC enzymes can be identified by functional, structural, and/or genetic analysis. In most cases, homologous or similar *PDC* genes and/or homologous or similar PDC enzymes will have functional, structural, or genetic similarities. Techniques known to those skilled in the art may be suitable to identify homologous genes and homologous enzymes. Generally, analogous genes and/or analogous enzymes can be identified by functional analysis and will have functional similarities. Techniques known to those skilled in the art may be suitable to identify analogous genes and analogous enzymes. For example, to identify homologous or analogous genes, proteins, or enzymes, techniques may include, but not limited to, cloning a *PDC* gene by PCR using primers based on a published sequence of a gene/enzyme or by degenerate PCR using degenerate primers designed to amplify a conserved region among *PDC* genes. Further, one skilled in the art can use techniques to identify homologous or analogous genes, proteins, or enzymes with functional homology or similarity. Techniques include examining a cell or cell culture for the catalytic activity of an enzyme through *in vitro* enzyme assays for said activity, then isolating the enzyme with said activity through purification, determining the protein sequence of the enzyme through techniques such as Edman degradation, design of PCR primers to the likely nucleic acid sequence, amplification of said DNA sequence through PCR, and cloning of said nucleic acid sequence. To identify homologous or similar genes and/or homologous or similar enzymes, analogous genes and/or analogous enzymes or proteins, techniques also include comparison of

data concerning a candidate gene or enzyme with databases such as BRENDA, KEGG, or MetaCYC. The candidate gene or enzyme may be identified within the above mentioned databases in accordance with the teachings herein. Furthermore, PDC activity can be determined phenotypically. For example, ethanol production under fermentative conditions can be assessed. A lack of ethanol production may be indicative of a yeast microorganism with no PDC activity.

Identification of GPD in a yeast microorganism

[00257] Any method can be used to identify genes that encode for enzymes with glycerol-3-phosphate dehydrogenase (GPD) activity. GPD catalyzes the reduction of dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate (G3P) with the corresponding oxidation of NADH to NAD+. Generally, homologous or similar *GPD* genes and/or homologous or similar GPD enzymes can be identified by functional, structural, and/or genetic analysis. In most cases, homologous or similar *GPD* genes and/or homologous or similar GPD enzymes will have functional, structural, or genetic similarities. Techniques known to those skilled in the art may be suitable to identify homologous genes and homologous enzymes. Generally, analogous genes and/or analogous enzymes can be identified by functional analysis and will have functional similarities. Techniques known to those skilled in the art may be suitable to identify analogous genes and analogous enzymes. For example, to identify homologous or analogous genes, proteins, or enzymes, techniques may include, but not limited to, cloning a *GPD* gene by PCR using primers based on a published sequence of a gene/enzyme or by degenerate PCR using degenerate primers designed to amplify a conserved region among *GPD* genes. Further, one skilled in the art can use techniques to identify homologous or analogous genes, proteins, or enzymes with functional homology or similarity. Techniques include examining a cell or cell culture for the catalytic activity of an enzyme through *in vitro* enzyme assays for said activity, then isolating the enzyme with said activity through purification, determining the protein sequence of the enzyme through techniques such as Edman degradation, design of PCR primers to the likely nucleic acid sequence, amplification of said DNA sequence through PCR, and cloning of said nucleic acid sequence. To identify homologous or similar genes and/or homologous or similar enzymes, analogous genes and/or analogous enzymes or proteins, techniques also include comparison of data concerning a candidate gene or enzyme with databases such as BRENDA, KEGG, or MetaCYC. The candidate gene or enzyme may be identified within the above mentioned databases in accordance with the teachings herein. Furthermore, GPD activity can be determined phenotypically. For example, glycerol production under fermentative conditions can be assessed. A lack of glycerol production may be indicative of a yeast microorganism with no GPD activity.

Genetic insertions and deletions

[00258] Any method can be used to introduce a nucleic acid molecule into yeast and many such methods are well known. For example, transformation and electroporation are common methods for introducing nucleic acid into yeast cells.

See, e.g., Gietz *et al., Nucleic Acids Res.* **27**:69-74 (1992); Ito *et al., J. Bacterol.* **153**:163-168 (1983); and Becker and Guarente, *Methods in Enzymology* **194**:182-187 (1991).

[00259] In an embodiment, the integration of a gene of interest into a DNA fragment or target gene of a yeast microorganism occurs according to the principle of homologous recombination. According to this embodiment, an integration cassette containing a module comprising at least one yeast marker gene and/or the gene to be integrated (internal module) is flanked on either side by DNA fragments homologous to those of the ends of the targeted integration site (recombinogenic sequences). After transforming the yeast with the cassette by appropriate methods, a homologous recombination between the recombinogenic sequences may result in the internal module replacing the chromosomal region in between the two sites of the genome corresponding to the recombinogenic sequences of the integration cassette. (Orr-Weaver *et al., Proc Natl Acad Sci U S A* **78**:6354-6358 (1981))

[00260] In an embodiment, the integration cassette for integration of a gene of interest into a yeast microorganism includes the heterologous gene under the control of an appropriate promoter and terminator together with the selectable marker flanked by recombinogenic sequences for integration of a heterologous gene into the yeast chromosome. In an embodiment, the heterologous gene includes an appropriate native gene desired to increase the copy number of a native gene(s). The selectable marker gene can be any marker gene used in yeast, including but not limited to, *HIS3, TRP1, LEU2, URA3, bar, ble, hph,* and *kan.* The recombinogenic sequences can be chosen at will, depending on the desired integration site suitable for the desired application.

[00261] In another embodiment, integration of a gene into the chromosome of the yeast microorganism may occur via random integration (Kooistra, R., Hooykaas, P.J.J., Steensma, H.Y. 2004. Yeast **21**: 781-792).

[00262] Additionally, in an embodiment, certain introduced marker genes are removed from the genome using techniques well known to those skilled in the art. For example, URA3 marker loss can be obtained by plating *URA3* containing cells in FOA (5-fluoro-orotic acid) containing medium and selecting for FOA resistant colonies (Boeke, J. *et al*, 1984, *Mol. Gen. Genet,* **197**, 345-47).

[00263] The exogenous nucleic acid molecule contained within a yeast cell of the disclosure can be maintained within that cell in any form. For example, exogenous nucleic acid molecules can be integrated into the genome of the cell or maintained in an episomal state that can stably be passed on ("inherited") to daughter cells. Such extra-chromosomal genetic elements (such as plasmids, *etc.*) can additionally contain selection markers that ensure the presence of such genetic elements in daughter cells. Moreover, the yeast cells can be stably or transiently transformed. In addition, the yeast cells described herein can contain a single copy, or multiple copies of a particular exogenous nucleic acid molecule as described above.

Reduction of enzymatic activity

[00264] Yeast microorganisms within the scope of the invention may have reduced enzymatic activity such as reduced glycerol-3-phosphate dehydrogenase activity.

The term "reduced" as used herein with respect to a particular enzymatic activity refers to a lower level of enzymatic activity than that measured in a comparable yeast cell of the same species. The term reduced also refers to the elimination of enzymatic activity than that measured in a comparable yeast cell of the same species. Thus, yeast cells lacking glycerol-3-phosphate dehydrogenase activity are considered to have reduced glycerol-3-phosphate dehydrogenase activity since most, if not all, comparable yeast strains have at least some glycerol-3-phosphate dehydrogenase activity. Such reduced enzymatic activities can be the result of lower enzyme concentration, lower specific activity of an enzyme, or a combination thereof. Many different methods can be used to make yeast having reduced enzymatic activity. For example, a yeast cell can be engineered to have a disrupted enzyme-encoding locus using common mutagenesis or knock-out technology. *See, e.g.,* Methods in Yeast Genetics (1997 edition), Adams, Gottschling, Kaiser, and Stems, Cold Spring Harbor Press (1998). In addition, certain point-mutation(s) can be introduced which results in an enzyme with reduced activity.

[00265] Alternatively, antisense technology can be used to reduce enzymatic activity. For example, yeast can be engineered to contain a cDNA that encodes an antisense molecule that prevents an enzyme from being made. The term "antisense molecule" as used herein encompasses any nucleic acid molecule that contains sequences that correspond to the coding strand of an endogenous polypeptide. An antisense molecule also can have flanking sequences (*e.g.,* regulatory sequences). Thus antisense molecules can be ribozymes or antisense oligonucleotides. A ribozyme can have any general structure including, without limitation, hairpin, hammerhead, or axhead structures, provided the molecule cleaves RNA.

[00266] Yeast having a reduced enzymatic activity can be identified using many methods. For example, yeast having reduced glycerol-3-phosphate dehydrogenase activity can be easily identified using common methods, which may include, for example, measuring glycerol formation via liquid chromatography.

Overexpression of heterologous genes

[00267] Methods for overexpressing a polypeptide from a native or heterologous nucleic acid molecule are well known. Such methods include, without limitation, constructing a nucleic acid sequence such that a regulatory element promotes the expression of a nucleic acid sequence that encodes the desired polypeptide. Typically, regulatory elements are DNA sequences that regulate the expression of other DNA sequences at the level of transcription. Thus, regulatory elements include, without limitation, promoters, enhancers, and the like. For example, the exogenous genes can be under the control of an inducible promoter or a constitutive promoter. Moreover, methods for expressing a polypeptide from an exogenous nucleic acid molecule in yeast are well known. For example, nucleic acid constructs that are used for the expression of exogenous polypeptides within Kluyveromyces and Saccharomyces are well known (*see, e.g.,* U.S. Pat. Nos. 4,859,596 and 4,943,529, for Kluyveromyces and, *e.g.,* Gellissen *et al.,* Gene 190(1):87-97 (1997) for *Saccharomyces*). Yeast plasmids have a selectable marker and an origin of replication. In addition certain plasmids may also contain a *centromeric* sequence.

These *centromeric* plasmids are generally a single or low copy plasmid. Plasmids without a *centromeric* sequence and utilizing either a 2 micron (*S. cerevisiae*) or 1.6 micron (*K. lactis*) replication origin are high copy plasmids. The selectable marker can be either prototrophic, such as *HIS3, TRP1, LEU2, URA3* or *ADE2*, or antibiotic resistance, such as, *bar, ble, hph,* or *kan*.

[00268] In another embodiment, heterologous control elements can be used to activate or repress expression of endogenous genes. Additionally, when expression is to be repressed or eliminated, the gene for the relevant enzyme, protein or RNA can be eliminated by known deletion techniques.

[00269] As described herein, any yeast within the scope of the disclosure can be identified by selection techniques specific to the particular enzyme being expressed, over-expressed or repressed. Methods of identifying the strains with the desired phenotype are well known to those skilled in the art. Such methods include, without limitation, PCR, RT-PCR, and nucleic acid hybridization techniques such as Northern and Southern analysis, altered growth capabilities on a particular substrate or in the presence of a particular substrate, a chemical compound, a selection agent and the like. In some cases, immunohistochemistry and biochemical techniques can be used to determine if a cell contains a particular nucleic acid by detecting the expression of the encoded polypeptide. For example, an antibody having specificity for an encoded enzyme can be used to determine whether or not a particular yeast cell contains that encoded enzyme. Further, biochemical techniques can be used to determine if a cell contains a particular nucleic acid molecule encoding an enzymatic polypeptide by detecting a product produced as a result of the expression of the enzymatic polypeptide. For example, transforming a cell with a vector encoding acetolactate synthase and detecting increased acetolactate concentrations compared to a cell without the vector indicates that the vector is both present and that the gene product is active. Methods for detecting specific enzymatic activities or the presence of particular products are well known to those skilled in the art. For example, the presence of acetolactate can be determined as described by Hugenholtz and Starrenburg, *Appl. Microbiol. Biotechnol.* **38**:17-22 (1992).

Increase of enzymatic activity

[00270] Yeast microorganisms of the invention may be further engineered to have increased activity of enzymes. The term "increased" as used herein with respect to a particular enzymatic activity refers to a higher level of enzymatic activity than that measured in a comparable yeast cell of the same species. For example, overexpression of a specific enzyme can lead to an increased level of activity in the cells for that enzyme. Increased activities for enzymes involved in glycolysis or the isobutanol pathway would result in increased productivity and yield of isobutanol.

[00271] Methods to increase enzymatic activity are known to those skilled in the art. Such techniques may include increasing the expression of the enzyme by increased copy number and/or use of a strong promoter, introduction of mutations to relieve negative regulation of the enzyme, introduction of specific mutations to increase specific activity and/or decrease the Km for the substrate, or by directed

evolution. See, e.g., Methods in Molecular Biology (vol. 231), ed. Arnold and Georgiou, Humana Press (2003).

Microorganism characterized by producing isobutanol at high yield

[00272] For a biocatalyst to produce isobutanol most economically, it is desired to produce a high yield. Preferably, the only product produced is isobutanol. Extra products lead to a reduction in product yield and an increase in capital and operating costs, particularly if the extra products have little or no value. Extra products also require additional capital and operating costs to separate these products from isobutanol.

[00273] The microorganism may convert one or more carbon sources derived from biomass into isobutanol with a yield of greater than 5% of theoretical. In one embodiment, the yield is greater than 10%. In one embodiment, the yield is greater than 50% of theoretical. In one embodiment, the yield is greater than 60% of theoretical. In another embodiment, the yield is greater than 70% of theoretical. In yet another embodiment, the yield is greater than 80% of theoretical. In yet another embodiment, the yield is greater than 85% of theoretical. In yet another embodiment, the yield is greater than 90% of theoretical. In yet another embodiment, the yield is greater than 95% of theoretical. In still another embodiment, the yield is greater than 97.5% of theoretical.

[00274] More specifically, the microorganism converts glucose, which can be derived from biomass into isobutanol with a yield of greater than 5% of theoretical. In one embodiment, the yield is greater than 10% of theoretical. In one embodiment, the yield is greater than 50% of theoretical. In one embodiment the yield is greater than 60% of theoretical. In another embodiment, the yield is greater than 70% of theoretical. In yet another embodiment, the yield is greater than 80% of theoretical. In yet another embodiment, the yield is greater than 85% of theoretical. In yet another embodiment the yield is greater than 90% of theoretical. In yet another embodiment, the yield is greater than 95% of theoretical. In still another embodiment, the yield is greater than 97.5% of theoretical

Microorganism characterized by production of isobutanol from pyruvate via an overexpressed isobutanol pathway and a Pdc-minus and Gpd-minus phenotype

[00275] In yeast, the conversion of pyruvate to acetaldehyde is a major drain on the pyruvate pool, and, hence, a major source of competition with the isobutanol pathway. This reaction is catalyzed by the pyruvate decarboxylase (PDC) enzyme. Reduction of this enzymatic activity in the yeast microorganism results in an increased availability of pyruvate and reducing equivalents to the isobutanol pathway and may improve isobutanol production and yield in a yeast microorganism that expresses a pyruvate-dependent isobutanol pathway.

[00276] Reduction of PDC activity can be accomplished by 1) mutation or deletion of a positive transcriptional regulator for the structural genes encoding for *PDC* or 2) mutation or deletion of all *PDC* genes in a given organism. The term "transcriptional regulator" can specify a protein or nucleic acid that works *in trans* to increase or to decrease the transcription of a different locus in the genome. For example, in

S.cerevisiae, the *PDC2* gene, which encodes for a positive transcriptional regulator of *PDC1,5,6* genes can be deleted; a *S. cerevisiae* in which the *PDC2* gene is deleted is reported to have only ~10% of wildtype PDC activity (Hohmann, *Mol Gen Genet*, **241**:657-666 (1993)). Alternatively, for example, all structural genes for PDC (e.g. in *S. cerevisiae*, *PDC1*, *PDC5*, and *PDC6*,or in *K. lactis*, *PDC1*) are deleted.

[00277] Crabtree-positive yeast strains such as *Saccharomyces.cerevisiae* strain that contains disruptions in all three of the *PDC* alleles no longer produce ethanol by fermentation. However, a downstream product of the reaction catalyzed by PDC, acetyl-CoA, is needed for anabolic production of necessary molecules. Therefore, the Pdc- mutant is unable to grow solely on glucose, and requires a two-carbon carbon source, either ethanol or acetate, to synthesize acetyl-CoA. (Flikweert MT, de Swaaf M, van Dijken JP, Pronk JT. FEMS Microbiol Lett. 1999 May 1;174(1):73-9. PMID:10234824 and van Maris AJ, Geertman JM, Vermeulen A, Groothuizen MK, Winkler AA, Piper MD, van Dijken JP, Pronk JT. Appl Environ Microbiol. 2004 Jan;70(1):159-66. PMID: 14711638).

[00278] Thus, in an embodiment, such a Crabtree-positive yeast strain may be evolved to generate variants of the PDC mutant yeast that do not have the requirement for a two-carbon molecule and has a growth rate similar to wild type on glucose. Any method, including chemostat evolution or serial dilution may be utilized to generate variants of strains with deletion of three *PDC* alleles that can grow on glucose as the sole carbon source at a rate similar to wild type (van Maris *et al.*, Directed Evolution of Pyruvate Decarboxylase-Negative *Saccharomyces cerevisiae*, Yielding a C2-Independent, Glucose-Tolerant, and Pyruvate-Hyperproducing Yeast, Applied and Environmental Microbiology, 2004, 70(1), 159-166).

[00279] Another byproduct that would decrease yield of isobutanol is glycerol. Glycerol is produced by 1) the reduction of the glycolysis intermediate, dihydroxyacetone phosphate (DHAP), to glycerol-3-phosphate (G3P) via the oxidation of NADH to NAD^+ by Glycerol-3-phosphate dehydrogenase (GPD) followed by 2) the dephosphorylation of glycerol-3-phophate to glycerol by glycerol-3-phosphatase (GPP). Production of glycerol results in loss of carbons as well as reducing equivalents. Reduction of GPD activity would increase yield of isobutanol. Reduction of GPD activity in addition to PDC activity would further increase yield of isobutanol. Reduction of glycerol production has been reported to increase yield of ethanol production (Nissen *et al.*, Anaerobic and aerobic batch cultivation of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis, Yeast, 2000, 16, 463-474; Nevoigt *et al.*, Method of modifying a yeast cell for the production of ethanol, WO 2009/056984). Disruption of this pathway has also been reported to increase yield of lactate in a yeast engineered to produce lactate instead of ethanol (Dundon *et al.*, Yeast cells having disrupted pathway from dihydroxyacetone phosphate to glycerol, US 2009/0053782).

[00280] In one embodiment, the microorganism is a crab-tree positive yeast with reduced or no GPD activity. In another embodiment, the microorganism is a crab-tree positive yeast with reduced or no GPD activity, and expresses an isobutanol biosynthetic pathway and produces isobutanol. In yet another embodiment, the microorganism is a crab-tree positive yeast with reduced or no GPD activity and with

reduced or no PDC activity. In another embodiment, the microorganism is a crab-tree positive yeast with reduced or no GPD activity,with reduced or no PDC activity, and expresses an isobutanol biosynthetic pathway and produces isobutanol.

[00281] In another embodiment, the microorganism is a crab-tree negative yeast with reduced or no GPD activity. In another embodiment, the microorganism is a crab-tree negative yeast with reduced or no GPD activity, expresses the isobutanol biosynthetic pathway and produces isobutanol. In yet another embodiment, the microorganism is a crab-tree negative yeast with reduced or no GPD activity and with reduced or no PDC activity. In another embodiment, the microorganism is a crab-tree negative yeast with reduced or no GPD activity,with reduced or no PDC activity, expresses an an isobutanol biosynthetic pathway and produces isobutanol.

Method of using microorganism for high-yield isobutanol fermentation

[00282] In a method to produce isobutanol from a carbon source at high yield, the yeast microorganism is cultured in an appropriate culture medium containing a carbon source.

[00283] Another exemplary embodiment provides a method for producing isobutanol comprising a recombinant yeast microorganism of the invention in a suitable culture medium containing a carbon source that can be converted to isobutanol by the yeast microorganism of the invention.

[00284] In certain embodiments, the method further includes isolating isobutanol from the culture medium. For example, isobutanol may be isolated from the culture medium by any method known to those skilled in the art, such as distillation, pervaporation, or liquid-liquid extraction.

[00285] This invention is further illustrated by the following examples that should not be construed as limiting. The contents of all references, patents, and published patent applications cited throughout this application, as well as the Figure and the Sequence Listing, are incorporated herein by reference for all purposes.

EXAMPLES

Table 2. Amino acid sequences disclosed herein

SEQ ID NO	Protein, Accession No.
1	*Thermotoga petrophila* RKU-1 dihydroxyacid dehydratase (DHAD), YP_001243973.1
2	*Victivallis vadensis* ATCC BAA-548 dihydroxyacid dehydratase (DHAD), ZP_01924101.1
3	Termite group 1 bacterium phylotype Rs-D17 dihydroxyacid dehydratase (DHAD), YP_001956631.1
4	*Yarrowia lipolytica* dihydroxyacid dehydratase (DHAD), XP_502180.2
5	*Francisella tularensis* subsp. tularensis WY96-3418 dihydroxyacid dehydratase (DHAD), YP_001122023.1
6	*Arabidopsis thaliana* dihydroxyacid dehydratase (DHAD), AAK64025.1
7	*Candidatus Koribacter versatilis* Ellin345 dihydroxyacid dehydratase (DHAD), YP_592184.1 (Acidobacter)
8	*Gramella forsetii* KT0803 dihydroxyacid dehydratase (DHAD), YP_862145.1
9	*Lactococcus lactis* subsp. lactis Il1403 dihydroxyacid dehydratase (DHAD), NP_267379.1
10	*Saccharopolyspora erythraea* NRRL 2338 dihydroxyacid dehydratase (DHAD), YP_001103528.2
11	*Saccharomyces cerevisiae* ILV3, NP_012550.1
12	*Piromcyes sp E2 ilvD*
13	*Ralstonia eutropha* JMP134 *ilvD*, YP_298150.1
14	*Chromohalobacter salexigens ilvD*, YP_573197.1
15	*Picrophilus torridus* DSM9790 *ilvD*, YP_024215.1
16	*Sulfolobus tokodaii* str. 7 dihydroxyacid dehydratase (DHAD), NP_378168.1
17	*Saccharomyces cerevisiae* ILV3ΔN
18	*c-myc* epitope tag
19	P(I/L)XXXGX(I/L)XIL (conserved motif described in Example 9)
20	PIKXXGX(I/L)XIL (conserved motif described in Example 9)

Table 3. Nucleic acid sequences disclosed herein

SEQ ID NO	Gene, Accession No.
107	*Candidatus Koribacter versatilis* Ellin345, (Acidobacter)
108	*Gramella forsetii* KT0803 (*Gf_ilvD*)
109	*Lactococcus lactis* subsp. lactis Il1403 (*Ll_ilvD*)
110	*Saccharopolyspora erythraea* NRRL 2338 (*Se_ilvD*)
111	*Saccharomyces cerevisiae* ILV3 (*ScILV3*(FL))
112	*Piromcyes sp E2 ilvD* (*Piromyces ilvD*)
113	*Ralstonia eutropha* JMP134 *ilvD*, (*Re_ilvD*)
114	*Chromohalobacter salexigens ilvD*, (*Cs_ilvD*)
115	*Picrophilus torridus* DSM9790 *ilvD*, (*Pt_ilvD*)

116	*Sulfolobus tokodaii* str. 7 *ilvD*, (*St_ilvD*)
117	*Saccharomyces cerevisiae* ILV3ΔN (*ScILV3ΔN*)
118	*E. coli ilvC^{Q110V}*,(*Ec_ilvC(Q110V)*)
119	*Lactococcus lactis* kivD, (*Ll_kivD*)
120	*S. cerevisiae* ILV5, (*ScILV5*)

[00286] **Determination of optical density.** The optical density of the yeast cultures is determined at 600 nm using a DU 800 spectrophotometer (Beckman-Coulter, Fullerton, CA, USA). Samples are diluted as necessary to yield an optical density of between 0.1 and 0.8.

[00287] **Gas Chromatography.** Analysis of volatile organic compounds, including ethanol and isobutanol was performed on a HP 5890 gas chromatograph fitted with an HP 7673 Autosampler, a DB-FFAP column (J&W; 30 m length, 0.32 mm ID, 0.25_µM film thickness) or equivalent connected to a flame ionization detector (FID). The temperature program was as follows: 200°C for the injector, 300°C for the detector, 50°C oven for 1 minute, 31°C/minute gradient to 140°C, and then hold for 2.5 min. Analysis was performed using authentic standards (>99%, obtained from Sigma-Aldrich), and a 5-point calibration curve with 1-pentanol as the internal standard.

[00288] **High Performance Liquid Chromatography for quantitativie analysis of glucose and organic acids.** Analysis of glucose and organic acids was performed on a HP-1100 High Performance Liquid Chromatography system equipped with an Aminex HPX-87H Ion Exclusion column (Bio-Rad, 300x7.8 mm) or equivalent and an H$^+$ cation guard column (Bio-Rad) or equivalent. Organic acids were detected using an HP-1100 UV detector (210 nm, 8 nm 360 nm reference) while glucose was detected using an HP-1100 refractive index detector. The column temperature was 60°C. This method was Isocratic with 0.008 N sulfuric acid in water as the mobile phase. Flow was set at 1 mL/min. Injection volume was 20 µL and the run time was 30 minutes.

[00289] **High Performance Liquid Chromatography for quantitativie analysis of ketoisovalerate and isobutyraldehyde.** Analysis of the DNPH derivatives of ketoisovalerate and isobutyraldehyde was performed on a HP-1100 High Performance Liquid Chromatography system equipped with a Hewlett Packard 1200 HPLC stack column (Agilent Eclipse XDB-18, 150 X 4.0 mm; 5 µm particles [P/N #993967-902] and C18 Guard cartridge). The analytes were detected using an HP-1100 UV detector at 360 nm The column temperature was 50°C. This method was isocratic with 0.1 % H$_3$PO$_4$ and 70% acetonitrile in water as mobile phase. Flow was set at 3 mL/min. Injection size was 10 µL and the run time was 2 minutes.

[00290] **Molecular biology and bacterial cell culture.** Standard molecular biology methods for cloning and plasmid construction are generally used, unless otherwise noted (Sambrook, J., Russel, D.W. *Molecular Cloning, A Laboratory Manual*. 3 ed. 2001, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

[00291] Standard recombinant DNA and molecular biology techniques used in the Examples are well known in the art and are described by Sambrook, J., Russel,

D.W. *Molecular Cloning, A Laboratory Manual*. 3 ed. 2001, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; and by T.J. Silhavy, M.L. Bennan, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984) and by Ausubel, F.M. *et al., Current Protocols in Molecular Biology*, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987).

[00292] General materials and methods suitable for the routine maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R.G.E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds.), American Society for Microbiology, Washington, D.C. (1994)) or by Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989).

[00293] **Yeast transformations – *S. cerevisiae.*** *S. cerevisiae* strains were transformed by the Lithium Acetate method (Gietz *et al., Nucleic Acids Res.* **27**:69-74 (1992): Cells from 50 mL YPD cultures (YPGal for valine auxotrophs) were collected by centrifugation (2700 rcf, 2 minutes, 25°C) once the cultures reached an OD_{600} of 1.0. The cells were washed cells with 50 mL sterile water and collected by centrifugation at 2700 rcf for 2 minutes at 25°C. The cells were washed again with 25 mL sterile water and collected cells by centrifugation at 2700 rcf for 2 minutes at 25°C. The cells were resuspended in 1 mL of 100 mM lithium acetate and transferred to a 1.5 mL eppendorf tube. The cells were collected cells by centrifugation for 20 sec at 18,000 rcf, 25°C. The cells were resuspended cells in a volume of 100 mM lithium acetate that was approximately 4x the volume of the cell pellet. A mixture of DNA (final volume of 15 µl with sterile water), 72 µl 50% PEG, 10 µl 1 M lithium acetate, and 3 µl denatured salmon sperm DNA was prepared for each transformation. In a 1.5 mL tube, 15 µl of the cell suspension was added to the DNA mixture (85 µl), and the transformation suspension was vortexed with 5 short pulses. The transformation was incubated at 30 minutes at 30°C, followed by incubation for 22 minutes at 42°C. The cells were collected by centrifugation for 20 sec at 18,000 rcf, 25°C. The cells were resuspended in 100 µl SOS (1 M sorbitol, 34% (v/v) YP (1% yeast extract, 2% peptone), 6.5 mM $CaCl_2$) or 100 µl YP (1% yeast extract, 2% peptone) and spread over an appropriate selective plate.

[00294] **Yeast colony PCR with FailSafe™ PCR System(EPICENTRE® Biotechnologies**, Madison, WI; Catalog #FS99250): Cells from each colony were added to 20 µl of colony PCR mix (per reaction mix contains 6.8 µl water, 1.5 µl of each primer, 0.2 µl of FailSafe PCR Enzyme Mix and 10 µl 2x FailSafe Master Mix). Unless otherwise noted, 2x FailSafe Master Mix E was used. The PCR reactions were incubated in a thermocycler using the following touchdown PCR conditions: 1 cycle of 94°Cx2 min, 10 cycles of 94°C x 20s, 63° - 54°C x 20s (decrease 1°C per cycle), 72°C x 60s, 40 cycles of 94°C x 20s, 53°C x 20s, 72°C x 60s and 1 cycle of 72°C x 5 min.

[00295]   **Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, CA;
Catalog #D4002) Protocol:** DNA fragments were recovered from agarose gels
according to manufacturer's protocol.

[00296]   **Zymo Research DNA Clean and Concentrator Kit (Zymo Research,
Orange, CA; Catalog #D4004) Protocol:** DNA fragments were purified according to
manufacturer's protocol.

Composition of Culture Media

[00297]   **Drugs:** When indicated, G418 (Calbiochem, Gibbstown, NJ) was
added at 0.2 g/L, Phleomycin (InvivoGen, San Diego, CA) was added at 7.5 mg/L,
Hygromycin (InvivoGen, San Diego, CA) was added at 0.2 g/L, and 5-fluoro-orotic
acid (FOA; Toronto Research Chemicals, North York, Ontario, Canada) was added
at 1 g/L.

[00298]   **YP:** 1% (w/v) yeast extract, 2% (w/v) peptone.

[00299]   **YPD:** YP containing 2% (w/v) glucose unless otherwise noted.

[00300]   **YPGal:** YP containing 2% (w/v) galactose

[00301]   **YPE:** YP containing 2% (w/v) Ethanol.

[00302]   **SC media:** 6.7 g/L Difco™ Yeast Nitrogen Base, 14g/L Sigma™
Synthetic Dropout Media supplement (includes amino acids and nutrients *excluding*
histidine, tryptophan, uracil, and leucine; Sigma-Aldrich, St. Louis, MO), 0.076 g/L
histidine, 0.076 g/L tryptophan, 0.380 g/L leucine, and 0.076 g/L uracil. Drop-out
versions of SC media is made by omitting one or more of histidine (H), tryptophan
(W), leucine (L), or uracil (U or Ura). When indicated, SC media are supplemented
with additional isoleucine (9xI; 0.684 g/L), valine (9xV; 0.684 g/L) or both isoleucine
and valine (9xIV). SCD is SC containing 2% (w/v) glucose unless otherwise noted,
SCGal is SC containing 2% (w/v) galactose and SCE is SC containing 2% (w/v)
ethanol. For example, SCD-Ura+9xIV would be composed of 6.7 g/L Difco™ Yeast
Nitrogen Base, 14g/L Sigma™ Synthetic Dropout Media supplement (includes amino
acids and nutrients *excluding* histidine, tryptophan, uracil, and leucine), 0.076 g/L
histidine, 0.076 g/L tryptophan, 0.380 g/L leucine, 0.684 g/L isoleucine, 0.684 g/L
valine, and 20 g/L glucose.

[00303]   **SCD-V+9xI:** 6.7 g/L Difco™ Yeast Nitrogen Base, 0.076 g/L Adenine
hemisulfate, 0.076 g/L Alanine 0.076 g/L,,Arginine hydrochloride, 0.076 g/L
Asparagine monohydrate, 0.076 g/L Aspartic acid, 0.076 g/L Cysteine hydrochloride
monohydrate, 0.076 g/L Glutamic acid monosodium salt, 0.076 g/L Glutamine, 0.076
g/L Glycine, 0.076 g/L myo-Inositol, 0.76 g/L Isoleucine, 0.076 g/L Lysine
monohydrochloride, 0.076 g/L Methionine, 0.008 g/L p-Aminobenzoic acid potassium
salt, 0.076 g/L Phenylalanine, 0.076 g/L Proline, 0.076 g/L Serine, 0.076 g/L
Threonine, 0.076 g/L Tyrosine disodium salt, and 20g/L glucose.

[00304]   **YNB:** 6.7 g/L Difco™ Yeast Nitrogen Base supplemented with
indicated nutrients as follows: histidine (H; 0.076 g/L), tryptophan (W; 0.076 g/L),
leucine (L; 0.380 g/L), uracil (U or Ura; 0.076 g/L), isoleucine (I; 0.076 g/L), valine (V;
0.076 g/L), and casamino acids (CAA; 10 g/L). When indicated, YNB media are
supplemented with higher amounts of isoleucine (10xI = 0.76 g/L), valine (10xV =
0.76 g/L) or both isoleucine and valine (10xIV). YNBD is YNB containing 2% (w/v)

glucose unless otherwise noted, YNBGal is YNB containing 2% (w/v) galactose and YNBE is YNB containing 2% (w/v) ethanol. For example, YNBGal+HWLU+10xl+G418 would be composed of 6.7 g/L Difco™ Yeast Nitrogen Base, 0.076 g/L histidine, 0.076 g/L tryptophan, 0.380 g/L leucine, 0.076 g/L uracil, 0.76 g/L isoleucine, 0.2 g/L G418, and 20 g/L galactose.

[00305] **Plates**: Solid versions of the above described media contain 2% (w/v) agar.

Example 1. Construction of an ILV3 deletion mutant

[00306] The purpose of this Example is to describe the construction of an ILV3 deletion mutant of *Saccharomyces cerevisiae*, GEVO2244.

Table 1-1 details the genotype of strains disclosed in this example:

#	Genotype
GEVO1147	*K. lactis*, NRRL Y-1140, (obtained from USDA)
GEVO1188	*S. cerevisiae*, CEN.PK, (obtained from Euroscarf); *MATα ura3 leu2 his3 trp1*
GEVO2145	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3::Kl_URA3*
GEVO2244	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ*

Table 1-2 outlines the plasmids disclosed in this example

Plasmid name	Genotype
pUC19	*bla, pUC-ori (obtained from Invitrogen)*
pGV1299	*K. lactis URA3, bla, pUC-ori (GEVO)*

Table 1-3 outlines the primers sequences disclosed in this example

GEVO #	PRIMER SEQUENCE
575	TTTTGAATTCTGGTTCTATCGAGGAGAAAAAGCGACAAG
576	TTTTGGATCCGGATGTGAAGTCGTTGACACAGTCG
1623	GTCTCTGATAAGGAAATGGCTC
1886	TCAAGAAGCCTCAAGTCGGGGTTGGTTCCTGTTGGTGGTCCGGTAACCCATGTAACATGC
1887	CGGTAACCCATGTAACATGCATCTATTGGACTTGAATAACATTCTGGTTCTATCGAGGAG
1888	CTTTCGTTAACAAGCCCATCTCTACTTTTTTCTTGGCTGTATCCGGATGTGAAGTCGTTG
1889	GATGGGCTTGTTAACGAAAGTTGCTACATCTAGACAATTCTGCATTATAGGCCCCAATCG
1890	TTAGTGGCAGCAAAGCAGAG
1892	ACATGATGCCCGTTCACAAC
1916	CAGGATGACAGTCGATGAG
1917	TGTCAACGACTTCACATCCG
1920	TGCAGCCTAGCTTTGAAGAC
1921	TACGTTAGGACCCCAGTATC

[00307] Plasmid pGV1299 was constructed by cloning the *K. lactis URA3* gene into pUC19. The *K. lactis URA3* was obtained by PCR using primers 575 and 576 from *K. lactis* genomic DNA. The PCR product was digested with *Eco*RI and *Bam*HI and cloned into pUC19 which was similarly digested. The *K. lactis URA3* insert was sequenced (Laragen Inc., Los Angeles, CA) to confirm correct sequence.

[00308] The integration cassette contained, from 5' to 3', the following: 1) a 80 bp homology to *ILV3* (position +158 to 237) that functions as the 5' targeting sequence for the integration, 2) the *K. lactis URA3* marker gene, 3) a 60 bp homology to a region *ILV3* (position -21 to +39) that is further upstream of the 5' targeting sequence to facilitate loop-out of the *K. lactis URA3* marker, and 4) a 221 bp homology to the 5' region of ILV3 (position +1759 to 1979) that functions as the 3' targeting sequence for the integration. This cassette was generated by SOE-PCR. The *K. lactis URA3* gene was amplified from pGV1299 using primers 1887 and 1888. Only the 3' region of *ILV3* was initially amplified using primers 1623 and 1892 from genomic DNA and this product was used as template to amplify the 3' region of *ILV3* using primers 1889 and 1890. The *K. lactis URA3* and the 3' region of *ILV3* were combined by SOE-PCR using primers 1886 and 1890.

[00309] GEVO1188 was transformed with the *ilv3::KI_URA3* cassette described above and plated onto YNBD+W+CAA (-Ura) plates. Initially, eight colonies (#1-8) were patched onto YNBD+HUWLIV plates and then replica plated onto YNBD+HUWLI (-V) plates to test for valine auxotrophy. As none of these exhibited valine auxotrophy, an additional eight colonies (#9-16) were streaked out for single colonies and 3 or 4 isolates (A – C or D) from each streak were tested for valine auxotrophy. Isolates A-C from clone #12 exhibited valine auxotrophy.

[00310] These isolates were tested for the correct integrations by colony PCR using primer pairs 1916 + 1920 and 1917 + 1921 for the 5' and 3' junctions, respectively. Correct sized bands were observed with clones #12A-C with primer pair 1916 + 1920. Correct sized bands were observed with clone 12A when FailSafe Master Mix A or C was used with primer pair 1917 + 1921. Clone #12A was designated as GEVO2145. The valine auxotrophies of GEVO2145 were reconfirmed by streaking them onto SCD+9xIV and SCD-V+9xI plates. GEVO2145 exhibited no growth on the medium lacking valine (SCD-V+9xI) while it grew on medium supplemented with valine (SCD+9xIV). The parent strain, GEVO1188, grew on both media.

[00311] GEVO2145 was streaked onto YNBE+W+CAA+FOA to isolate strains in which the *K. lactis URA3* had been excised through homologous recombination, i.e. "looped out". Five FOA resistant clones (A-E) were tested for auxotrophies for valine and uracil. All five clones exhibited auxotrophies to both nutrients. Clones A was designated GEVO2244. Colony PCRs using primers 1891 and 1892 with FailSafe Buffer C were performed and the loss of the *KI_URA3* cassette was confirmed.

Example 2: Dihydroxy acid dehydratase limits isobutanol production in yeast

[00312] This example illustrates the specific activity of various DHAD homologs in yeast. The example also illustrates that high specific activity of the *Lactococcus*

lactis IlvD enzyme (SEQ ID NO. 9) correlates with an increase in isobutanol production.

Table 2-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
GEVO1187	*S. cerevisiae*, CEN.PK; *MATa ura3 leu2 his3 trp1*
GEVO2244	*S. cerevisiae*, CEN.PK; *MATa ura3 leu2 his3 trp1 ilv3Δ*

Table 2-2 outlines the plasmids disclosed in this example:

pGV No.	Figure	Genotype
pGV1106	2-1	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TDH3* promoter-Myc tag-polylinker-*CYC1* terminator
pGV1662	2-2, 2-3	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-(kivD)
pGV1851		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Gramella forsetti ilvD*
pGV1852		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Chromohalobacter salexigens ilvD*
pGV1853		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Ralstonia eutropha ilvD*
pGV1854		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Saccharopolyspora erythraea ilvD*
pGV1855		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Ll_ilvD*
pGV1900		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*ScILV3(FL)*
pGV1904		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Acidobacteria bacterium Ellin345 ilvD*
pGV1905		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Picrophilus torridus DSM 9790 ilvD*
pGV1906		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Piromyces species E2 ilvD*
pGV1907		pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Sulfolobus tokodaii strain 7 ilvD*

Table 2-3 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence
271	CTAGCATGGAACAAAAACTCATCTCAGAAGAAGATGGTGTCGAC GAATTCCCGGGATCCGCGGCCGC
272	TCGAGCGGCCGCGGATCCCGGGAATTCGTCGACACCATCTTCTT CTGAGATGAGTTTTTGTTCCATG
1617	CGTTGA*GTCGAC*ATGGGCTTGTTAACGAAAGTTGC
1618	GCCAAC*GGATCC*TCAAGCATCTAAAACACAACCG

[00313]   Plasmid pGV1106 (Figure 2-1) is a variant of the plasmid p426GPD (which is described in Mumberg *et al*, Gene (1995), 119-122). To obtain pGV1106, annealed oligos 271 and 272 (see Table X-3) were ligated into p426GPD that had been digested with *Spe*I and *Xho*I, and the inserted DNA confirmed by sequencing.
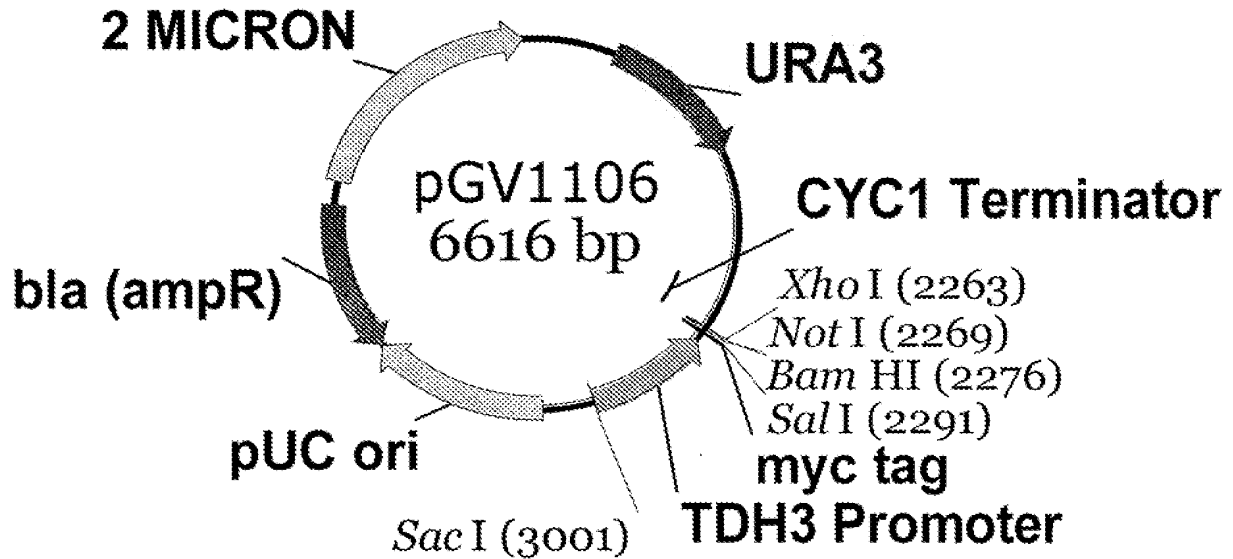
Figure 2-1. Schematic map of pGV1106.

[00314]   Plasmid pGV1662 (Figures 2-2, 2-3) served as the parental plasmid of pGV1855, pGV1900, and pGV2019. A schematic map of pGV1662 is shown in Figure 2-2, and its complete nucleotide sequence is shown in Figure 2-3. The salient features of pGV1662 include the yeast 2μm origin of replication, the *URA3* selectable marker, and the *ScTEF1* promoter sequence followed by restriction sites into which an ORF can be cloned to permit its expression under the regulation of the *TEF1* promoter.

Figure 2-2. Schematic map of plasmid pGV1662.

ttggatcatactaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggcccttttcgtctcgcgcgt
ttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagacttgtctgtaagcggatgccggggagcag
acaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactga
gagtgcaccataccacagcttttcaattcaattcatcattttttttttattcttttttttttgatttcggtttctttgaaatttt
tttgattcggtaatctccgaacagaaggaagaacgaaggaaggagcacagacttagattggtatatatacgcatatgtagtgt
tgaagaaacatgaaattgcccagtattcttaacccaactgcacagaacaaaaacctgcaggaaacgaagataaatcatgtcga
aagctacatataaggaacgtgctgctactcatcctagtcctgttgctgccaagctatttaatatcatgcacgaaaagcaaaca
aacttgtgtgcttcattggatgttcgtaccaccaaggaattactggagttagttgaagcattaggtcccaaaatttgtttact
aaaaacacatgtggatatcttgactgattttttccatggagggcacagttaagccgctaaaggcattatccgccaagtacaatt
ttttactcttcgaagacagaaaatttgctgacattggtaatacagtcaaattgcagtactctgcgggtgtatacagaatagca
gaatgggcagacattacgaatgcacacggtgtggtgggcccaggtattgttagcggtttgaagcaggcggcagaagaagtaac
aaaggaacctagaggcctttttgatgttagcagaattgtcatgcaaggctccctatctactggagaatatactaagggtactg
ttgacattgcgaagagcgacaaagattttgttatcggctttattgctcaaagagacatgggtggaagagatgaaggttacgat
tggttgattatgacacccggtgtgggtttagatgacaagggagacgcattgggtcaacagtatagaaccgtggatgatgtggt
ctctacaggatctgacattattattgttggaagaggactatttgcaaagggaagggatgctaaggtagaggggtgaacgttaca
gaaaagcaggctgggaagcatatttgagaagatgcggccagcaaaactaaaaaactgtattataagtaaatgcatgtatacta
aactcacaaattagagcttcaatttaattatatcagttattaccctatgcggtgtgaaataccgcacagatgcgtaaggagaa
aataccgcatcaggaaattgtaaacgttaatattttgttaaaattcgcgttaaattttttgttaaatcagctcattttttaacc
aataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttgttccagtttggaacaag
agtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaaccatc
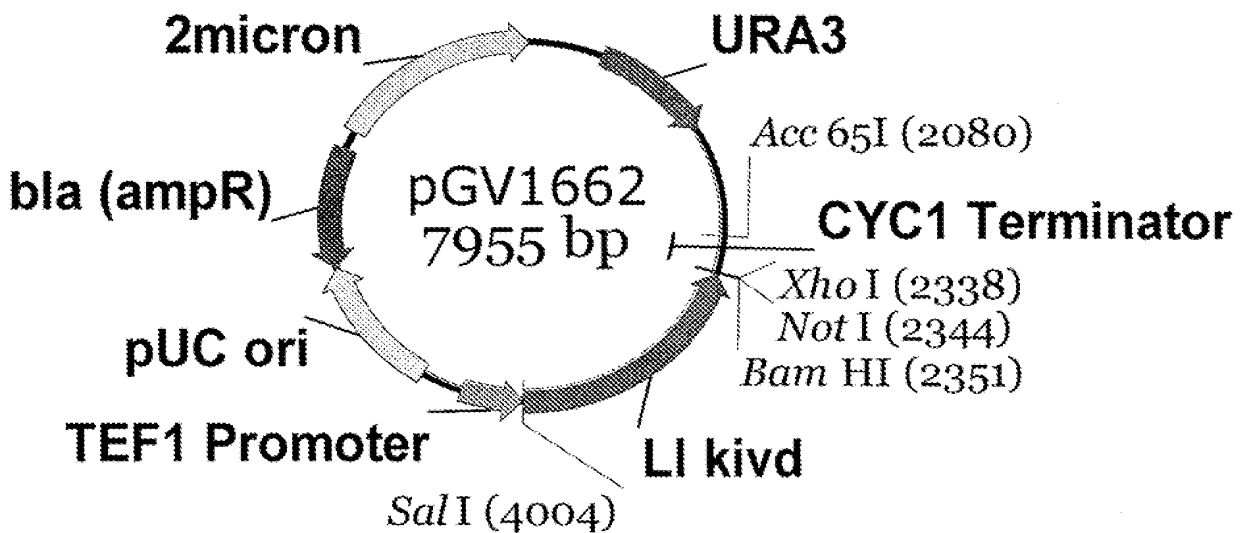accctaatcaagttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagggagcccccgatttagagcttgac
ggggaaagccggcgaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtc
acgctgcgcgtaaccaccacacccgccgcgcttaatgcgccgctacagggcgcgtcgcgccattcgccattcaggctgcgcaa
ctgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagggggatgtgctgcaaggcgattaagtt
gggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgagcgcgcgtaatacgactcactatagggcgaa
ttgggtaccggccgcaaattaaagccttcgagcgtcccaaaaccttctcaagcaaggttttcagtataatgttacatgcgtac
acgcgtctgtacagaaaaaaaagaaaaatttgaaatataaataacgttcttaatactaacataactataaaaaaataaatagg
gacctagacttcaggttgtctaactccttcctttcggttagagcggatgtggggggagggcgtgaatgtaagcgtgacataa
ctaattacatgactcgagcggccgcgggatccttaggatttattctgttcagcaaacagcttgcccattttcttcagtaccttc
ggtgcgccttctttcgccaggatcagttcgatccagtacatacggttcggatcggcctgggcctctttcatcacgctcacaaa
ttcgttttcggtacgcacaatttttagacacaacacggtcctcagttgcgccgaaggactccggcagtttagagtagttccaca
tagggatatcgttgtaagactggttcggaccgtggatctcacgctcaacggtgtagccgtcattgttaataatgaagcaaatc
gggttgatcttttcacgaattgccagacccagttcctgtacggtcagctgcagggaaccgtcaccgatgaacagcagatgacg
agattctttatcagcgatctgagagcccagcgctgccgggaaagtatagccaatgctaccccacagcggctgaccgataaaat
ggcttttggatttcagaaagatagaagacgcgccgaaaaagctcgtaccttgttccgccacgatggtttcattgctctgggtc
aggttctccacggcctgccacaggcgatcctgggacagcagtgcgttagatggtacgaaatcttcttgcttttgtcaatgta
tttgcctttatactcgatttcggacaggtccagcagagagctgatcaggctttcgaagtcgaagttctggatacgctcgttga
agattttaccctcgtcgatgttcaggctaatcattttgttttcgttcagatggtgagtgaatgcaccggtagaagagtcggtc
agtttaacgccagcatcaggatgaagtcgcagattcaacaaattctttcaggttcggttcgtcagagtaccgttgtagat
gcccaggaaagacggcagagcctcgtcaacagaggacttgccgaagtcagggtggtaatcggcagtttggtttgctgatga
attgggtcacggtcttctccagaccaaaagaaatgatttcgtggccggtgatcacgattggtttctttgcgttttcagagac
tcctggattttgttcaggattcctggtcgctagtgttagaagtggagttttcttttcagcggcaggctcggttttttcggc
tttagctgccgcaacatccacaggcaggttgatgtaaactggtttgcgttctttcagcagcgcagacagaacgcggtcgattt
ccacagtagcgttctctgcagtcagcagcgtacgtgccgcagtcacaggttcatgcattttcatgaagtgtttgaaatcgccg
tcagccagagtgtggtggacgaatttaccttcgttctgaactttgctcgttgggctgcctacgatctccaccacggcaggtt
ttcggcgtaggagccgccagacccgttgacggcgctcagttcgccaacaccgaaagtggtcagaaatgccgcggctttcttgg
tacgtgcataaccatctgccatgtagcttgcgttcagttcgttagcgttacccacccatttcatgtctttatgagagatgatc
tgatccaggaactgcagattgtaatcacccggaacgccgaagattcttcgataccagttcatgcagacggtccagcagata
atcaccaacagtatacatgtcgacaaacttagattagattgctatgctttcttttctaatgagcaagaagtaaaaaagttgta
atagaacaagaaaatgaaactgaaacttgagaaattgaagaccgtttattaacttaaatatcaatgggaggtcatcgaaaga
gaaaaaaatcaaaaaaaaaatttttcaagaaaaagaaacgtgataaaaattttattgccttttttcgacgaagaaaaagaaacg
aggcggtctcttttttcttttccaaacctttagtacgggtaattaacgacaccctagaggaagaaagaggggaaatttagtat
gctgtgcttgggtgttttgaagtggtacggcgatgcgcggagtccgagaaaatctggaagagtaaaaaaggagtagaaacatt
ttgaagctatgagctccagcttttgttccctttagtgagggttaattgcgcgcttggcgtaatcatggtcatagctgtttcct
gtgtgaaattgttatccgctcacaattccacacaacataggagccggaagcataaagtgtaaagcctggggtgcctaatgagt
gagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcg
gccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcgg
ctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtg
agcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagc
atcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctcc
ctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca

```
tagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccg
accgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggt
aacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggac
agtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccg
ctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttttct
acggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagat
cctttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatca
gtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgg
gagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaacca
gccagccggaaggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagcta
gagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggt
atggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctcctt
cggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactg
tcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagt
tgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttc
gggtgcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcat
cttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaa
tgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttga
atgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgaacgaagcatctgtgcttcat
tttgtagaacaaaaatgcaacgcgagagcgctaattttttcaaacaaagaatctgagctgcattttttacagaacagaaatgcaa
cgcgaaagcgctattttaccaacgaagaatctgtgcttcatttttgtaaaacaaaaatgcaacgcgagagcgctaattttttca
aacaaagaatctgagctgcatttttacagaacagaaatgcaacgcgagagcgctattttaccaacaaagaatctatacttctt
ttttgttctacaaaaatgcatcccgagagcgctattttttctaacaaagcatcttagattactttttttctcctttgtgcgctc
tataatgcagtctcttgataactttttttgcactgtaggtccgttaaggttagaagaaggctactttggtgtctattttctcttc
cataaaaaaagcctgactccacttcccgcgtttactgattactagcgaagctgcgggtgcatttttttcaagataaaggcatcc
ccgattatattctataccgatgtggattgcgcatactttgtgaacagaaagtgatagcgttgatgattcttcattggtcagaa
aattatgaacggtttcttctattttgtctctatatactacgtataggaaatgtttacatttttcgtattgttttcgattcactc
tatgaatagttcttactacaatttttttgtctaaagagtaatactagagataaacataaaaaatgtagaggtcgagtttagat
gcaagttcaaggagcgaaaggtggatgggtaggttatatagggatatagcacagagatatatagcaaagagatactttt gagc
aatgtttgtggaagcggtattcgcaatattttagtagctcgttacagtccggtgcgtttttggttttttgaaagtgcgtcttc
agagcgcttttggttttcaaaagcgctctgaagttcctatactttctagagaataggaacttcggaataggaacttcaaagcg
tttccgaaaacgagcgcttccgaaaatgcaacgcgagctgcgcacatacagctcactgttcacgtcgcacctatatctgcgtg
ttgcctgtatatatatatacatgagaagaacggcatagtgcgtgtttatgcttaaatgcgtacttatatgcgtctatttatgt
aggatgaaaggtagtctagtacctcctgtgatattatcccattccatgcggggtatcgtatgcttccttcagcactacccttt
agctgttctatatgctgccactcctcaattggattagtctcatccttcaatgctatcatttcctttgata
```

Figure 2-3. Nucleotide sequence of plasmid pGV1662.

[00315]   Plasmids pGV1851-1855 and pGV1904-1907 are all variants of pGV1662, in which the kivD ORF sequence present in pGV1662 was excised and replaced with a sequence encoding a DHAD homolog, as indicated below.

[00316]   Plasmid pGV1851 contains the *Gramella forsetti ilvD* gene sequence (SEQ ID NO. 108). Plasmid pGV1852 contains the *Chromohalobacter salexigens ilvD* gene sequence (SEQ ID NO. 114). Plasmid pGV1853 contains the *Ralstonia eutropha ilvD* gene sequence (SEQ ID NO. 113). Plasmid pGV1854 contains the *Saccharopolyspora erythraea ilvD* sequence (SEQ ID NO. 110). Plasmid pGV1855 contains the *Lactococcus lactis ilvD* sequence (SEQ ID NO. 109). Plasmid pGV1900 contains the *Saccharomyces cerevisiae ILV3* (*ScILV3*(FL)) sequence. Plasmid pGV1904 contains the *Acidobacteria bacterium Ellin345 ilvD* sequence (SEQ ID NO. 107). Plasmid pGV1905 contains the *Picrophilus torridus DSM 9790 ilvD* sequence (SEQ ID NO. 115). Plasmid pGV1906 contains the *Piromyces species E2 ilvD* sequence (SEQ ID NO. 112). Plasmid pGV1907 contains the *Sulfolobus tokodaii strain 7 ilvD* sequence (SEQ ID NO. 116). All sequences (except that of *ScILV3*(FL)) were synthesized with 5'*SalI* and 3'*NotI* sites by DNA2.0 (Menlo Park, CA), digested with *SalI* and *NotI*, and ligated into pGV1662 which had also been digested with *SalI*

and *NotI*. For plasmid pGV1900, the sequence containing the open reading frame of *ScILV3(FL)* was amplified from *S.cerevisiae* genomic DNA in a PCR reaction using primers 1617 and 1618, and the resulting 1.8kb fragment was digested with SaII plus BamHI and cloned into pGV1662 that had been digested with SaII plus BamHI.Various DHADs were tested for *in vitro* activity using whole cell lysates. In this case, the DHADs were expressed in a yeast deficient for DHAD activity (GEVO2244; *ilv3Δ*) to minimize endogenous background activity.

[00317] To grow cultures for cell lysates, triplicate independent cultures of each desired strain were grown overnight in 3 mL SCD-Ura+9xIV at 30°C, 250rpm. The following day, the overnight cultures were diluted 1:50 into 50mL fresh SCD-Ura in a 250 mL baffle-bottomed Erlenmeyer flask and incubated at 30°C at 250rpm. After approximately 10 hours, the OD_{600} of all cultures were measured, and the cells of each culture were collected by centrifugation (2700xg, 5 min). The cell pellets were washed by resuspending in 1mL of water, and the suspension was placed in a 1.5mL tube and the cells were collected by centrifugation (16,000xg, 30 seconds). All supernatant was removed from each tube and the tubes were frozen at -80°C until use.

[00318] Lysates were prepared by resuspending each cell pellet in 0.7mL of lysis buffer. Lysate lysis buffer consisted of: 0.1M Tris-HCl pH 8.0, 5mM $MgSO_4$, with 10 µL of of Yeast/Fungal Protease Arrest solution (G Biosciences, catalog #788-333) per 1mL of lysis buffer. Eight hundred microliters of cell suspension were added to 1mL of 0.5mm glass beads that had been placed in a chilled 1.5mL tube. Cells were lysed by bead beating (6 rounds, 1 minute per round, 30 beats per second) with 2 minutes chilling on ice in between rounds. The tubes were then centrifuged (20,000xg, 15 min) to pellet debris and the supernatant (cell lysates) were retained in fresh tubes on ice. The protein concentration of each lysate was measured using the BioRad Bradford protein assay reagent (BioRad, Hercules, CA) according to manufacturer's instructions.

[00319] The DHAD activity of each lysate was ascertained as follows. In a fresh 1.5mL centrifuge tube, 50µL of each lysate was mixed with 50µL of 0.1M 2,3-dihydroxyisovalerate (DHIV), 25µL of 0.1M $MgSO_4$, and 375µL of 0.05M Tris-HCl pH 8.0, and the mixture was incubated for 30min at 35°C. Each tube was then heated to 95°C for 5min to inactivate any enzymatic activity, and the solution was centrifuged (16,000xg for 5min) to pellet insoluble debris. To prepare samples for analysis, 100µL of each reaction were mixed with 100µL of a solution consisting of 4 parts 15mM dinitrophenyl hydrazine (DNPH) in acetonitrile with 1 part 50mM citric acid, pH 3.0, and the mixture was heated to 70°C for 30min in a thermocycler. The solution was then analyzed by HPLC as described above in General Methods to quantitate the concentration of ketoisovalerate (KIV) present in the sample.

Table 2-4. Specific activities (KIV generation) from lysates of *S.cerevisiae* strain GEVO2244 carrying plasmids to overexpress the indicated DHAD homolog. Each data point is the result of triplicate samples.

Plasmid	Gene	Specific activity (U/mg total protein)
pGV1106	Control (i.e. no DHAD)	n.d.
pGV1851	*Gramella forsetti ilvD*	0.012
pGV1852	*Chromohalobacter salexigens ilvD*	n.d.
pGV1853	*Ralstonia eutropha ilvD*	n.d.
pGV1854	*Saccharopolyspora erythraea ilvD*	0.002
pGV1855	*Lactococcus lactis ilvD*	0.027
pGV1900	*Saccharomyces cerevisiae ILV3(FL)*	0.148
pGV1904	*Acidobacteria bacterium Ellin345 DHAD*	0.004
pGV1905	*Picrophilus torridus DSM 9790 DHAD*	n.d.
pGV1906	*Piromyces Sp E2 DHAD*	0.016
pGV1907	*Sulfolobus tokodaii str. 7 DHAD*	0.001

* n.d., not detectable

Example 3: Dihydroxy acid dehydratase limits isobutanol production in yeast

[00320] This example illustrates that high specific DHAD activity, and in particular the high specific activity of the *Lactococcus lactis* IlvD enzyme (SEQ ID NO. 9) correlates with an increase in isobutanol production.

Table 3-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
GEVO1186	*S.cerevisiae, CEN.PK; MATa/α ura3/ura3 leu2/leu2 his3/his3 trp1/trp1*
GEVO1188	*S.cerevisiae, CEN.PK; MATα ura3 leu2 his3 trp1*
GEVO1803	*MATa/α ura3/ura3 leu2/leu2 his3/his3 trp1/trp1 pdc1::Bs-alsS,TRP1/PDC1*
GEVO2107	*MATa/α ura3/ura3 leu2/leu2 his3/his3 trp1/trp1 pdc1::Bs-alsS,TRP1/PDC1 pdc6::{ScTEF1p-Ll_kivd ScTDH3p-Dm_ADH URA3}/PDC6*

Table 3-2 outlines the plasmids disclosed in this example:

pGV No.	Genotype
p423GPD	P_{TDH3}:MCS:T_{CYC1},HIS3,2-micron, *bla*, pUC ori (Mumberg, D. et al. (1995) Gene **156**:119-122; obtained from ATCC)
pGV1103	P_{TDH3}:myc-tag:MCS:T_{CYC1},HIS3,2micron, *bla*, pUC ori
pGV1730	P_{CUP1}:Bs-alsS:T_{PDC1}/PDC1-3' region:PDC1-5' region, *TRP1*, *bla*, pUC ori
pGV1974	P_{TEF1}:Sc_ILV3ΔN:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, HIS3, 2micron, *bla*, pUC ori bla(ampR)
pGV1914	P_{TEF1}:Ll_kivD P_{TDH3}:Dm_ADH PDC6 5',3' targeting homology URA3 pUC ori bla(ampR)
pGV1981	P_{TEF1}:Lactococcus lactis ilvD-coSc:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, HIS3, 2micron, *bla*, pUC ori
pGV2001	P_{TEF1}:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, HIS3, 2micron, *bla*, pUC ori

Table 3-3 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence (5' to 3')
271	CTAGCATGGAACAAAAACTCATCTCAGAAGAAGATGGTGTCGACGAAT TCCCGGGATCCGCGGCCGC
272	TCGAGCGGCCGCGGATCCCGGGAATTCGTCGACACCATCTTCTTCTGA GATGAGTTTTGTTCCATG
637	TTTTGAGCTCGCCGATCCCATTACCGACATTTGGG
638	AAAGTCGACACCGATATACCTGTATGTGTCACCACCAATGTATCTATAA GTATCCATGCTAGCCCTAGGTTTATGTGATGATTGATTGATTGATTG
697	GAGTACGGATCCCTAGAGAGCTTTCGTTTTCATGAG
767	CAAGAAGTCGACATGTTGACAAAAGCAACAAAGAAC
1321	AATCATATCGAACACGATGC
1322	TCAGAAAGGATCTTCTGCTC
1323	ATCGATATCGTGAAATACGC
1324	AGCTGGTCTGGTGATTCTAC
1409	ATTGATGCGGCCGCGATTTAATCTCTAATTATTAGTTA
1410	CACCCAGTCGCGACATCCAATTTATAGAAATCAG
1411	ATTGGATGTCGCGACTGGGTGAGCATATGTTC
1412	GAGAAAGCCGGCAGGAGAGTGAAAGAGCCTTG
1440	ATCGTACATCTTCCAAGCATC
1441	AATCGGAACCCTAAAGGGAG
1443	TGCAGATGCAGATGTGAGAC
1587	CGGCTGCCAGAACTCTACTAACTG
1588	GCGACGTCTACTGGCAGGTTAAT
1633	TCCGTCACTGGATTCAATGCCATC
1634	TTCGCCAGGGAGCTGGTGAA

[00321]　Plasmid pGV1103 was generated by inserting a linker (primers 271 annealed to primer 272) containing a myc-tag and a new MCS (SalI-EcoRI-SmaI-BamHI-NotI) into the SpeI and XhoI sites of p423GPD.

[00322]　Plasmid pGV1730 is a yeast integration plasmid used to replace the PDC1 gene in S. cerevisiae with the Bs_alsS gene expressed using the S. cerevisiae CUP1 promoter. The CUP1 promoter originated as a PCR product from S. cerevisiae genomic DNA using primers 637 and 638. The B. subtilis alsS originated as a PCR product from B. subtilis genomic DNA using primers 767 and 697. Plasmid pGV1730 also carries the S. cerevisiae TRP1 gene as a selection marker. Plasmid pGV1730 also contains a targeting sequence suitable for directing the homologous integration of the plasmid to the S.cerevisiae PDC1 locus. This targeting sequence consists of, 5' to 3', the PDC1 terminator region, a unique NruI restriction site, and the PDC1 promoter region. This fragment was generated by SOE PCR. The PDC1 terminator sequence was amplified from S. cerevisiae genomic DNA using primers 1409 and 1410 and the PDC1 promoter sequence was amplified from S. cerevisiae genomic DNA using primers 1411 and 1412. Primers 1410 and 1411 have overlapping sequence and the sequences were chosen so that the junction between 3' end of the PDC1 terminator and the 5' end of PDC1 promoter created a unique NruI restriction site. These two PCR products were fused by SOE PCR using primers 1409 and 1411. Primer 1409 and 1411 introduced a

*Not*I and *Ngo*MIV sites, respectively. A schematic map of plasmid pGV1730 is shown in Figure 3-1, and the complete nucleotide sequence of pGV1730 is shown in Figure 3-2.
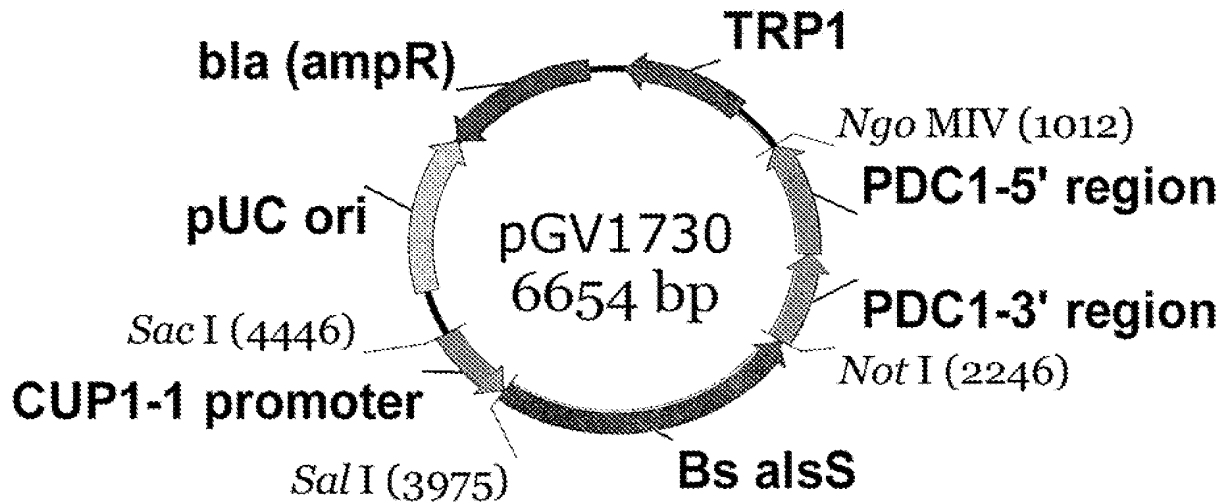

Figure 3-1. Schematic map of plasmid pGV1730.

CAGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTTTGACGAAATT
TGCTATTTTGTTAGAGTCTTTTACACCATTTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTAA
TCTAAGCGCATCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAAGCTTTCGGGGGCTCTCTTG
CCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCACCTGTCCCACCTGCTTCTGAATCAAACAAGGG
AATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCAGTCTTTTGGAAATACGAGTCTTTTAAT
AACTGGCAAACCGAGGAACTCTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGTA
ATCATTGACCAGAGCCAAAACATCCTCCTTAGGTTGATTACGAAACACGCCAACCAAGTATTTCGGAGTGCCTGA
ACTATTTTTATATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATTGTTCTCTTTCTATTGGG
CACACATATAATACCCAGCAAGTCAGCATCGGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCA
AACTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATACTCCAAGCTGCCTTTGTGTGCTTAAT
CACGTATACTCACGTGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACCGAATT
AATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGTACGTAAGGTGACAAGCTATTTTTCAATAAAG
AATATCTTCCACTACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGCTGTCTATTAAATGCT
TCCTATATTATATATATAGTAATGTCGTTGACGTCGCCGGCAGGAGAGTGAAAGAGCCTTGTTTATATATTTTTT
TTTCCTATGTTCAACGAGGACAGCTAGGTTTATGCAAAAATGTGCCATCACCATAAGCTGATTCAAATGAGCTAA
AAAAAAAATAGTTAGAAAATAAGGTGGTGTTGAACGATAGCAAGTAGATCAAGACACCGTCTAACAGAAAAAGGG
GCAGCGGACAATATTATGCAATTATGAAGAAAAGTACTCAAAGGGTCGGAAAAATATTCAAACGATATTTGCATT
AAATCCTCAATTGATTGATTATTCCATAGTAAAATACCGTAACAACACAAAATTGTTCTCAAATTCATAAATTAT
TCATTTTTTCCACGAGCCTCATCACACGAAAAGTCAGAAGAGCATACATAATCTTTTAAATGCATAGGTTATGCA
TTTTGCAAATGCCACCAGGCAACAAAAATATGCGTTTAGCGGGCGGAATCGGGAAGGAAGCCGGAACCACCAAAA
ACTGGAAGCTACGTTTTTAAGGAAGGTATGGGTGCAGTGTGCTTATCTCAAGAAATATTAGTTATGATATAAGGT
GTTGAAGTTTAGAGATAGGTAAATAAACGCGGGGTGTGTTTATTACATGAAGAAGAAGTTAGTTTCTGCCTTGCT
TGTTTATCTTGCACATCACATCAGCGGAACATATGCTCACCCAGTCGCGACATCCAATTTATAGAAATCAGCTTG
TGGGTATTGTTCAGAGAATTTTTCAATCATTGGAGCAATCATTTACATGGACCGCACCAAGTGGCGTAGAAATC
TACGACAACTAGCTTGTCTTGAGCAATTGCAGAGTCGAATTCGCTGGCAGTTTTGAATTGAGTAACCATTATTTG
TATCGAGGTGTCTAGTCTTCTATTACACTAATGCAGTTTCAGGGTTTTGGAAACCACACTGTTTAAACAGTGTTC
CTTAATCAAGGATACCTCTTTTTTTTTCCTTGGTTCCACTAATTCATCGGTTTTTTTTTTGGAAGACATCTTTTC
CAACGAAAAGAATATACATATCGTTTAAGAGAAATTCTCCAAATTTGTAAAGAAGCGGACCCAGACTTAAGCCTA
ACCAGGCCAATTCAACAGACTGTCGGCAACTTCTTGTCTGGTCTTTCCATGGTAAGTGACAGTGCAGTAATAATA
TGAACCAATTTATTTTTCGTTACATAAAAATGCTTATAAAACTTTAACTAATAATTAGAGATTAAATCGCGGCCG
CGGATCCCTAGAGAGCTTTCGTTTTCATGAGTTCCCCGAATTCTTTCGGAAGCTTGTCACTTGCTAAATTAACGT
TATCACTGTAGTCAACCGGGACATCAATGATGACAGGCCCCTCAGCGTTCATGCCTTGACGCAGAACATCTGCCA

GCTGGTCTGGTGATTCTACGCGTAAGCCAGTTGCTCCGAAGCTTTCCGCGTATTTCACGATATCGATATTTCCGA
AATCGACCGCAGATGTACGATTATATTTTTTCAATTGCTGGAATGCAACCATGTCATATGTGCTGTCGTTCCATA
CAATGTGTACAATTGGTGCTTTTAAACGAACTGCTGTCTCTAATTCCATAGCTGAGAATAAGAAACCGCCATCAC
CGGAGACTGATACTACTTTTTCTCCCGGTTTCACCAATGAAGCGCCGATTGCCCAAGGAAGCGCAACGCCGAGTG
TTTGCATACCGTTACTAATCATTAATGTTAACGGCTCGTAGCTGCGGAAATAACGTGACATCCAAATCGCGTGTG
AACCGATATCGCAAGTCACTGTAACATGATCATCGACTGCGTTTCGCAATTCTTTAACGATTTCAAGAGGATGCA
CTCTGTCTGATTTCCAATCTGCAGGCACCTGCTCACCCTCATGCATATATTGTTTAAATCAGAAAGGATCTTCT
GCTCACGTTCCGCAAAGTCTACTTTCACAGCATCGTGTTCGATATGATTGATCGTAGATGGAATATCACCGATCA
GTTCAAGATCCGGCTGGTAAGCATGATCAATGTCAGCCAGAATCTCGTCTAAATGGATGATCGTCCGGTCTCCAT
TGACATTCCAGAATTTCGGATCATATTCAATTGGGTCATAGCCGATTGTCAGAACAACATCAGCCTGCTCAAGCA
GCAGATCGCCAGGCTGGTTGCGGAATAAACCGATCCGGCCAAAATACTGATCCTCTAAATCTCTCGTAAGAGTAC
CGGCAGCTTGATATGTTTCAACGAATGGAAGCTGCACTTTTTTCAATAGCTTGCGAACCGCTTTAATCGCTTCCG
GTCTTCCGCCCTTCATGCCGACTAAAACGACAGGAAGTTTTGCTGTTTGAATTTTTGCAATGGCCATACTGATTG
CGTCATCTGCTGCGGGACCAAGTTTTGGCGCTGCGACAGCACGTACGTTTTTTGTATTTGTGACTTCATTCACAA
CATCTTGCGGAAAACTCACAAAAGCGGCCCCAGCCTGCCCTGCTGACGCTATCCTAAACGCATTTGTAACAGCTT
CCGGTATATTTTTTACATCTTGAACTTCTACACTGTATTTTGTAATCGGCTGGAATAGCGCCGCATTATCCAAAG
ATTGATGTGTCCGTTTTAAACGATCTGCACGGATCACGTTCCCAGCAAGCGCAACGACAGGGTCACCTTCAGTGT
TTGCTGTCAGCAGTCCTGTTGCCAAGTTCGAAGCACCTGGTCCTGATGTGACTAACACGACTCCCGGTTTTCCAG
TTAAACGGCCGACTGCTTGCGCCATAAATGCTGCATTTTGTTCATGCCGGGCAACGATAATTTCAGGCCCTTTAT
CTTGTAAAGCGTCAAATACCGCATCAATTTTTGCACCTGGAATGCCAAATACATGTGTGACACCTTGCTCCGCTA
AGCAATCAACAACAAGCTCCGCCCCTCTGCTTTTCACAAGGGATTTTTGTTCTTTTGTTGCTTTTGTCAACATGT
CGACTTTATGTGATGATTGATTGATTGATTGTACAGTTTGTTTTTCTTAATATCTATTTCGATGACTTCTATATG
ATATTGCACTAACAAGAAGATATTATAATGCAATTGATACAAGACAAGGAGTTATTTGCTTCTCTTTTATATGAT
TCTGACAATCCATATTGCGTTGGTAGTCTTTTTTGCTGGAACGGTTCAGCGGAAAAGACGCATCGCTCTTTTTGC
TTCTAGAAGAAATGCCAGCAAAAGAATCTCTTGACAGTGACTGACAGCAAAAATGTCTTTTTCTAACTAGTAACA
AGGCTAAGATATCAGCCTGAAATAAAGGGTGGTGAAGTAATAATTAAATCATCCGTATAAACCTATACACATATA
TGAGGAAAAATAATACAAAAGTGTTTTAAATACAGATACATACATGAACATATGCACGTATAGCGCCCAAATGTC
GGTAATGGGATCGGCGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTC
ATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAA
AGCCTGGGGTGCCTAATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA
CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC
TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG
TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC
ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCC
CGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG
CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT
CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG
TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGG
AAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG
AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAA
AGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGC
AGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG
CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAG
ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTG
AGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT

CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA
AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGT

Figure 3-2. Nucleotide sequence of plasmid pGV1730.

[00323] pGV1914 is a yeast integrating vector (YIp) that includes the *S. cerevisiae URA3* gene as a selection marker and contains homologous sequence for targeting the HpaI-digested, linearized plasmid for integration at the *PDC6* locus of *S. cerevisiae.*. pGV1914 carries the *D.melanogaster adh* (*Dm_ADH*) and *L.lactis kivd* (*Ll_kivD*) genes, expressed under the control of the *S. cerevisiae TDH3* and *TEF1* promoters, respectively. The open reading frame sequence of *DmADH* was originally amplified by PCR from clone RH54514 (available from the Drosophila Genome Resource Center). The nucleotide sequence of the *Dm_ADH* open reading frame is shown in Figure 3-3. A schematic map of pGV1914 is shown in Figure 3-4, and the complete nucleotide sequence of plasmid pGV1914 is shown in Figure 3-5.

ATGTCGTTTACTTTGACCAACAAGAACGTGATTTTCGTTGCCGGTCTGGGAGGCATTGGTCTGGACACCAGCAAG
GAGCTGCTCAAGCGCGATCTGAAGAACCTGGTGATCCTCGACCGCATTGAGAACCCGGCTGCCATTGCCGAGCTG
AAGGCAATCAATCCAAAGGTGACCGTCACCTTCTACCCCTATGATGTGACCGTGCCCATTGCCGAGACCACCAAG
CTGCTGAAGACCATCTTCGCCCAGCTGAAGACCGTCGATGTCCTGATCAACGGAGCTGGTATCCTGGACGATCAC
CAGATCGAGCGCACCATTGCCGTCAACTACACTGGCCTGGTCAACACCACGACGGCCATTCTGGACTTCTGGGAC
AAGCGCAAGGGCGGTCCCGGTGGTATCATCTGCAACATTGGATCCGTCACTGGATTCAATGCCATCTACCAGGTG
CCCGTCTACTCCGGCACCAAGGCCGCCGTGGTCAACTTCACCAGCTCCCTGGCGAAACTGGCCCCCATTACCGGC
GTGACGGCTTACACTGTGAACCCCGGCATCACCCGCACCACCCTGGTGCACACGTTCAACTCCTGGTTGGATGTT
GAGCCTCAGGTTGCCGAGAAGCTCCTGGCTCATCCCACCCAGCCCTCGTTGGCCTGCGCCGAGAACTTCGTCAAG
GCTATCGAGCTGAACCAGAACGGAGCCATCTGGAAACTGGACTTGGGCACCCTGGAGGCCATCCAGTGGACCAAG
CACTGGGACTCCGGCATCTAA

Figure 3-3. Nucleotide sequence of *Dm_ADH*.
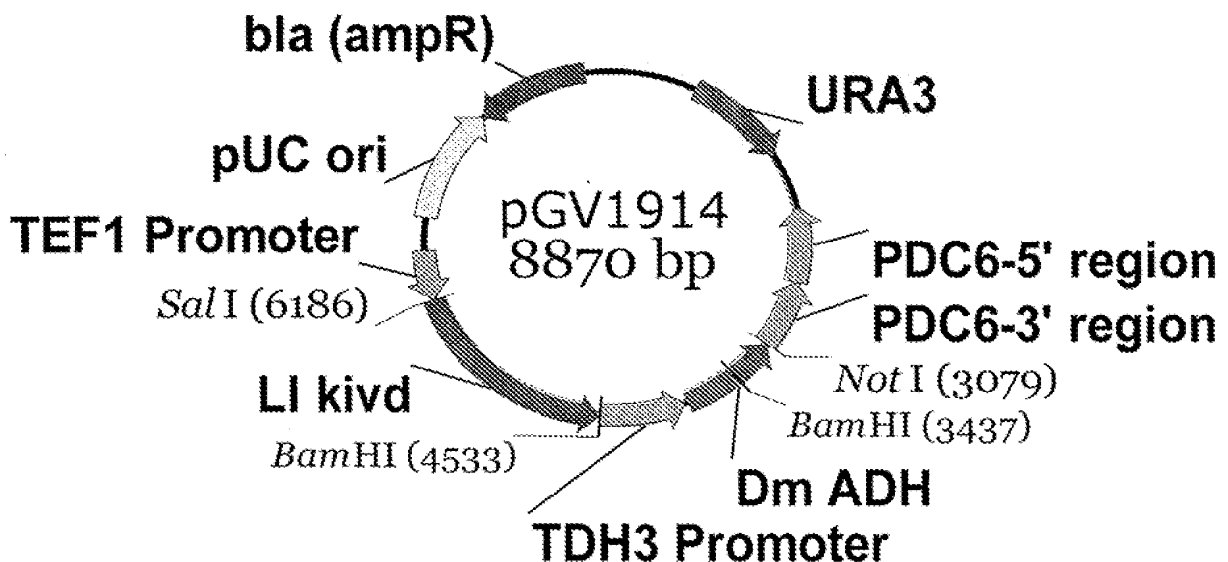
Figure 3-4. Schematic map of plasmid pGV1914.

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAG
CGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATG
CGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAT

103182 v1/DC

```
ACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT
TACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGAC
GTTGTAAAACGACGGCCAGTGAATTCATACCACAGCTTTTTCAATTCAATTCATCATTTTTTTTTTTATTCTTTTTT
TTGATTTCGGTTTCCTTGAAATTTTTTTGATTCGGTAATCTCCGAACAGAAGGAAGAACGAAGGAAGGAGCACAG
ACTTAGATTGGTATATATACGCATATGTAGTGTTGAAGAAACATGAAATTGCCCAGTATTCTTAACCCAACTGCA
CAGAACAAAAACCTGCAGGAAACGAAGATAAATCATGTCGAAAGCTACATATAAGGAACGTGCTGCTACTCATCC
TAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGAAAAGCAAACAAACTTGTGTGCTTCATTGGATGTTCG
TACCACCAAGGAATTACTGGAGTTAGTTGAAGCATTAGGTCCCAAAATTTGTTACTAAAAACACATGTGGATAT
CTTGACTGATTTTTCCATGGAGGGCACAGTTAAGCCGCTAAAGGCATTATCCGCCAAGTACAATTTTTTACTCTT
CGAAGACAGAAAATTTGCTGACATTGGTAATACAGTCAAATTGCAGTACTCTGCGGGTGTATACAGAATAGCAGA
ATGGGCAGACATTACGAATGCACACGGTGTGGTGGGCCCAGGTATTGTTAGCGGTTTGAAGCAGGCGGCAGAAGA
AGTAACAAAGGAACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCATGCAAGGGCTCCCTATCTACTGGAGAATA
TACTAAGGGTACTGTTGACATTGCGAAGAGCGACAAAGATTTTGTTATCGGCTTTATTGCTCAAAGAGACATGGG
TGGAAGAGATGAAGGTTACGATTGGTTGATTATGACACCCGGTGTGGGGTTTAGATGACAAGGGAGACGCATTGGG
TCAACAGTATAGAACCGTGGATGATGTGGTCTCTACAGGATCTGACATTATTATTGTTGGAAGAGGACTATTTGC
AAAGGGAAGGGATGCTAAGGTAGAGGGTGAACGTTACAGAAAAGCAGGCTGGGAAGCATATTTGAGAAGATGCGG
CCAGCAAAAACTAAAAAACTGTATTATAAGTAAATGCATGTATACTAAACTCACAAATTAGAGCTTCAATTTAATT
ATATCAGTTATTACCCTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGT
AAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCC
ACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACC
ATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATT
TAGAGCTTGACGGGGAAAGCCGGCGAGGACTGCAATAGCACAAGATTAAGATAGAATGGCTTCAAACAGCCGCCT
TTTATACATATTGGTAAAAGCTCGCGAATCGCACCATATCCCTTATCCTGTAATCAAATCGATCTAGGTGCAGAT
ACAGATCAATTCATAAAAAGAAATTGAAGCACCAGTTTATCACTACTACACTATCTTTTTCTTTTTTTTTTTTTTTT
TTGCGCAGTTTCGCCCTTTGTTCAATATCACTTGATAAGTTGTGGGCTTTTTCTGTCACTCATTCGGCTTAAAAA
GTATTCGTTCTTTTGTGTGTTTATGAAAAGGGAACGTGATATAAAAAAACATCCTTTGGTGTGGGACATGGGCTTT
TGTTTAGAGAATGGTTATCACTACCGCCCCCACCCTTGAAAGCCACAGAAAATGAAAAAGTATGTGAATAAGGTG
TGAACTCTATAACATTTTGGCCAAATGCCACAGCCGATCTGCATATTCCAATGGACATGATGCAACAACAATTGA
TGTCACATTCTCTTACACACTTCGATTGGTCCGTACGTAGTACTTTTTACATAACTGACTCAGGCGTTTCCTTCA
TTGAAATGCTCATCTATTGCCAAGTACATAGAATCCACAGTGCATAGGTTAACGCATTGTACCCAAACGACGGGA
AACAAGGAAGGATGCAGAATGAGCACTTGTTATTTATAAAAAGACACGGGAGGGGGAATCCCGTCTTTCGTCCGT
CGGAGCCAAAGAGATGAGCCAAAGCAGAAAAACAGGGGACGCCGCCCTTCTTCCGTCCCGTGCGTGAGGGGGGCG
CGGCCATTCGGTTTTTGCAATATGACCTGTGGGCCAAAAATCGAAAAAAAAAAAAAAAAAATAAGAGGCGGCTGCGG
AATTTTATAAGACAAGCGCAGGGCCAAAGAAAAAATAATAATTGACGTGGCTGAACAACAGTCTCTCCCCACCCC
TTTCCAAAAAGGGGAATGAAATACGAGTTCTTTTTCCCAATTGGTAGATATTCAACAAGAGACGCGCAGTACGTA
ACATGCGAATTGCGTAATTCACGGCGATAACGTAGTATTTAGATTTAGTATAATTTGAACCGATGTATTTATTTG
TCTGATTGATTTATGTATTCAAACTGTGTAAGTTTATTTATTTGCAACAATAATTCGTTTGAGTACACTACTAAT
GGCGGCCGCTTAGATGCCGGAGTCCCAGTGCTTGGTCCACTGGATGGCCTCCAGGGTGCCCAAGTCCAGTTTCCA
GATGGCTCCGTTCTGGTTCAGCTCGATAGCCTTGACGAAGTTCTCGGCGCAGGCCAACGAGGGCTGGGTGGGATG
AGCCAGGAGCTTCTCGGCAACCTGAGGCTCAACATCCAACCAGGAGTTGAACGTGTGCACCAGGGTGGTGCGGGT
GATGCCGGGGTTCACAGTGTAAGCCGTCACGCCGGTAATGGGGGCCAGTTTCGCCAGGGAGCTGGTGAAGTTGAC
CACGGCGGCCTTGGTGCCGGAGTAGACGGGCACCTGGTAGATGGCATTGAATCCAGTGACGGATCCAATGTTGCA
GATGATACCACCGGGACCGCCCTTGCGCTTGTCCCAGAAGTCCAGAATGGCCGTCGTGGTGTTGACCAGGCCAGT
GTAGTTGACGGCAATGGTGCGCTCGATCTGGTGATCGTCCAGGATACCAGCTCCGTTGATCAGGACATCGACGGT
CTTCAGCTGGGCGAAGATGGTCTTCAGCAGCTTGGTGGTCTCGGCAATGGGCACGGTCACATCATAGGGGTAGAA
GGTGACGGTCACCTTTGGATTGATTGCCTTCAGCTCGGCAATGGCAGCCGGGTTCTCAATGCGGTCGAGGATCAC
CAGGTTCTTCAGATCGCGCTTGAGCAGCTCCTTGCTGGTGTCCAGACCAATGCCTCCCAGACCGGCAACGAAAAT
CACGTTCTTGTTGGTCAAAGTAAACGACATACCGGTATCTCCTAGATCCGTCGAAGTCGAAACTAAGTTCTGGTG
TTTTAAAACTAAAAAAAAGACTAACTATAAAAGTAGAATTTAAGAAGTTTAAGAAATAGATTACAGAATTACAA
TCAATACCTACCGTCTTTATATACTTATTAGTCAAGTAGGGGAATAATTTCAGGGAACTGGTTTCAACCTTTTTTT
TTCAGCTTTTTTCCAAATCAGAGAGAGCAGAAGGTAATAGAAGGTGTAAGAAAATGAGATAGATACATGCGTGGGT
CAATTGCCTTGTGTCATCATTTACTCCAGGCAGGTTGCATCACTCCATTGAGGTTGTGCCCGTTTTTTGCCTGTT
TGTGCCCTGTTCTCTGTAGTTGCGCTAAGAGAATGGACCTATGAACTGATGGTTGGTGAAGAAAACAATATTTT
GGTGCTGGGATTCTTTTTTTTTCTGGATGCCAGCTTAAAAAGCGGGCTCCATTATATTTAGTGGATGCCAGGAAT
```

```
AAACTGTTCACCCAGACACCTACGATGTTATATATTCTGTGTAACCCGCCCCCTATTTTGGGCATGTACGGGTTA
CAGCAGAATTAAAAGGCTAATTTTTTGACTAAATAAAGTTAGGAAAATCACTACTATTAATTATTTACGTATTCT
TTGAAATGGCGAGTATTGATAATGATAAACTGGATCCTTAGGATTTATTCTGTTCAGCAAACAGCTTGCCCATTT
TCTTCAGTACCTTCGGTGCGCCTTCTTTCGCCAGGATCAGTTCGATCCAGTACATACGGTTCGGATCGGCCTGGG
CCTCTTTCATCACGCTCACAAATTCGTTTTCGGTACGCACAATTTTAGACACAACACGGTCCTCAGTTGCGCCGA
AGGACTCCGGCAGTTTAGAGTAGTTCCACATAGGGATATCGTTGTAAGACTGGTTCGGACCGTGGATCTCACGCT
CAACGGTGTAGCCGTCATTGTTAATAATGAAGCAAATCGGGTTGATCTTTTCACGAATTGCCAGACCCAGTTCCT
GTACGGTCAGCTGCAGGGAACCGTCACCGATGAACAGCAGATGACGAGATTCTTTATCAGCGATCTGAGAGCCCA
GCGCTGCCGGGAAAGTATAGCCAATGCTACCCCACAGCGGCTGACCGATAAAATGGCTTTTGGATTTCAGAAAGA
TAGAAGACGCGCCGAAAAAGCTCGTACCTTGTTCCGCCACGATGGTTTCATTGCTCTGGGTCAGGTTCTCCACGG
CCTGCCACAGGCGATCCTGGGACAGCAGTGCGTTAGATGGTACGAAATCTTCTTGCTTTTTGTCAATGTATTTGC
CTTTATACTCGATTTCGGACAGGTCCAGCAGAGAGCTGATCAGGCTTTCGAAGTCGAAGTTCTGGATACGCTCGT
TGAAGATTTTACCCTCGTCGATGTTCAGGCTAATCATTTTGTTTTCGTTCAGATGGTGAGTGAATGCACCGGTAG
AAGAGTCGGTCAGTTTAACGCCCAGCATCAGGATGAAGTCCGCAGATTCAACAAATTCTTTCAGGTTCGGTTCGC
TCAGAGTACCGTTGTAGATGCCCAGGAAAGACGGCAGAGCCTCGTCAACAGAGGACTTGCCGAAGTTCAGGGTGG
TAATCGGCAGTTTGGTTTTGCTGATGAATTGGGTCACGGTCTTCTCCAGACCAAAAGAAATGATTCGTGGCCGG
TGATCACGATTGGTTTCTTTGCGTTTTTCAGAGACTCCTGGATTTTGTTCAGGATTTCCTGGTCGCTAGTGTTAG
AAGTGGAGTTTTCTTTCTTCAGCGGCAGGCTCGGTTTTTCCGCTTTAGCTGCCGCAACATCCACAGGCAGGTTGA
TGTAAACTGGTTTGCGTTCTTTCAGCAGCGCAGACAGAACGCGGTCGATTTCCACAGTAGCGTTCTCTGCAGTCA
GCAGCGTACGTGCCGCAGTCACAGGTTCATGCATTTTCATGAAGTGTTTGAAATCGCCGTCAGCCAGAGTGTGGT
GGACGAATTTACCTTCGTTCTGAACTTTGCTCGTTGGGCTGCCTACGATCTCCACCACCGGCAGGTTTTCGGCGT
AGGAGCCCGCCAGACCGTTGACGGCGCTCAGTTCGCCAACACCGAAAGTGGTCAGAAATGCCGCGGCTTTCTTGG
TACGTGCATAACCATCTGCCATGTAGCTTGCGTTCAGTTCGTTAGCGTTACCCACCCATTTCATGTCTTTATGAG
AGATGATCTGATCCAGGAACTGCAGATTGTAATCACCCGGAACGCCGAAGATTTCTTCGATACCCAGTTCATGCA
GACGGTCCAGCAGATAATCACCAACAGTATACATGTCGACAAACTTAGATTAGATTGCTATGCTTTCTTTCTAAT
GAGCAAGAAGTAAAAAAAGTTGTAATAGAACAAGAAAAATGAAACTGAAACTTGAGAAATTGAAGACCGTTTATT
AACTTAAATATCAATGGGAGGTCATCGAAAGAGAAAAAAATCAAAAAAAAAATTTTCAAGAAAAAGAAACGTGAT
AAAAATTTTTATTGCCTTTTTCGACGAAGAAAAAGAAACGAGGCGGTCTCTTTTTTCTTTTCCAAACCTTTAGTA
CGGGTAATTAACGACACCTAGAGGAAGAAAGAGGGGAAATTTAGTATGCTGTGCTTGGGTGTTTTGAAGTGGTA
CGGCGATGCGCGGAGTCCGAGAAAATCTGGAAGAGTAAAAAAGGAGTAGAAACATTTTGAAGCTATGAGCTCCAG
CTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG
TTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAG
GTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT
GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA
CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG
AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT
GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA
AAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTG
GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC
CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG
AGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCC
TGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAG
TTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC
CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCC
GATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGT
CATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
```

TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG
TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC
CGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCAT
TTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG
CACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCG
TATCACGAGGCCCTTTCGTC

Figure 3-5. Complete nucleotide sequence of plasmid pGV1914.

[00324] Plasmid pGV1974 (Fig. 3-6) is a yeast high copy plasmid with *HIS3* as a marker for the expression of *E. coli* ilvC^{Q110V} and *S. cerevisiae* ILV3ΔN (SEQ ID NO. 117). pGV1974 was generated by cloning a *SacI-NotI* fragment (4.9 kb; Fig. 3-7) carrying the *S. cerevisiae TEF1* promoter:*S. cerevisiae ilv3ΔN*:*S. cerevisiae TDH3* promoter:*E. coli* ilvC^{Q110V} into the *SacI-NotI* sites of pGV1103 (5.4 kb), a yeast expression plasmid carrying the *HIS3* marker.
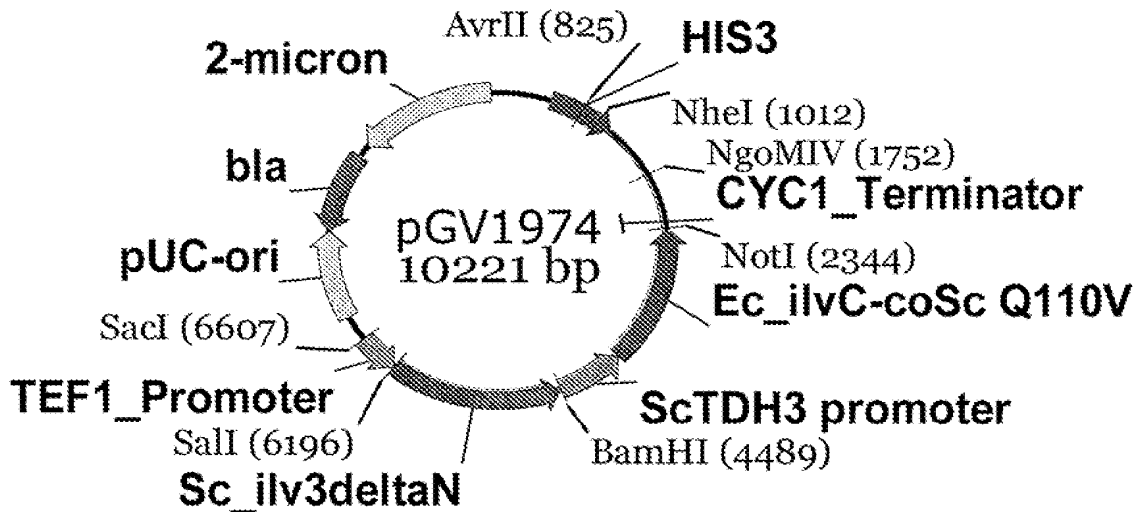
Figure 3-6. Schematic map of plasmid pGV1974.

GAGCTCATAGCTTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTAC
CACTTCAAAACACCCAAGCACAGCATACTAAATTTCCCCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTAC
TAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTCTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATC
ACGTTTCTTTTTCTTGAAAATTTTTTTTTTTGATTTTTTTCTCTTTCGATGACCTCCCATTGATATTTAAGTTAAT
AAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGCTCATT
AGAAAGAAAGCATAGCAATCTAATCTAAGTTTGTCGACATGAAGAAGCTCAACAAGTACTCGTATATCATCACTG
AACCTAAGGGCCAAGGTGCGTCCCAGGCCATGCTTTATGCCACCGGTTTCAAGAAGGAAGATTTCAAGAAGCCTC
AAGTCGGGGGTTGGTTCCTGTTGGTGGTCCGGTAACCCATGTAACATGCATCTATTGGACTTGAATAACAGATGTT
CTCAATCCATTGAAAAAGCGGGTTTGAAAGCTATGCAGTTCAACACCATCGGTGTTTCAGACGGTATCTCTATGG
GTACTAAAGGTATGAGATACTCGTTACAAAGTAGAGAAATCATTGCAGACTCCTTTGAAACCATCATGATGGCAC
AACACTACGATGCTAACATCGCCATCCCATCATGTGACAAAAACATGCCCGGTGTCATGATGGCCATGGGTAGAC
ATAACAGACCTTCCATCATGGTATATGGTGGTACTATCTTGCCCGGTCATCCAACATGTGGTTCTTCGAAGATCT
CTAAAAACATCGATATCGTCTCTGCGTTCCAATCCTACGGTGAATATATTTCCAAGCAATTCACTGAAGAAGAAA
GAGAAGATGTTGTGGAACATGCATGCCCAGGTCCTGGTTCTTGTGGTGGTATGTATACTGCCAACACAATGGCTT
CTGCCGCTGAAGTGCTAGGTTTGACCATTCCAAACTCCTCTTCCTTCCCAGCCGTTTCCAAGGAGAAGTTAGCTG
AGTGTGACAACATTGGTGAATACATCAAGAAGACAATGGAATTGGGTATTTTACCTCGTGATATCCTCACAAAAG
AGGCTTTTGAAAACGCCATTACTTATGTCGTTGCAACCGGTGGGTCCACTAATGCTGTTTTGCATTTGGTGGCTG
TTGCTCACTCTGCGGGTGTCAAGTTGTCACCAGATGATTTCCAAAGAATCAGTGATACTACACCATTGATCGGTG
ACTTCAAACCTTCTGGTAAATACGTCATGGCCGATTTGATTAACGTTGGTGGTACCCAATCTGTGATTAAGTATC
TATATGAAAACAACATGTTGCACGGTAACACAATGACTGTTACCGGTGACACTTTGGCAGAACGTGCAAAGAAAG
CACCAAGCCTACCTGAAGGACAAGAGATTATTAAGCCACTCTCCCACCCAATCAAGGCCAACGGTCACTTGCAAA

TTCTGTACGGTTCATTGGCACCAGGTGGAGCTGTGGGTAAAATTACCGGTAAGGAAGGTACTTACTTCAAGGGTA
GAGCACGTGTGTTCGAAGAGGAAGGTGCCTTTATTGAAGCCTTGGAAAGAGGTGAAATCAAGAAGGGTGAAAAAA
CCGTTGTTGTTATCAGATATGAAGGTCCAAGAGGTGCACCAGGTATGCCTGAAATGCTAAAGCCTTCCTCTGCTC
TGATGGGTTACGGTTTGGGTAAAGATGTTGCATTGTTGACTGATGGTAGATTCTCTGGTGGTTCTCACGGGTTCT
TAATCGGCCACATTGTTCCCGAAGCCGCTGAAGGTGGTCCTATCGGGTTGGTCAGAGACGGCGATGAGATTATCA
TTGATGCTGATAATAACAAGATTGACCTATTAGTCTCTGATAAGGAAATGGCTCAACGTAAACAAAGTTGGGTTG
CACCTCCACCTCGTTACACAAGAGGTACTCTATCCAAGTATGCTAAGTTGGTTTCCAACGCTTCCAACGGTTGTG
TTTTAGATGCTTGAGGATCCAGTTTATCATTATCAATACTCGCCATTTCAAAGAATACGTAAATAATTAATAGTA
GTGATTTTCCTAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGG
GGGCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCACTAAATATAA
TGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACC
ATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAACC
TCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATT
TTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCT
GAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCT
TAAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTAGTTTTAAAACACCAGAACTTAGTTTCGACTCGA
GATGGCCAACTATTTTAACACATTAAATTTGAGACAACAATTGGCTCAACTGGGTAAGTGCAGATTTATGGGAAG
GGACGAGTTTGCTGATGGTGCTTCTTATCTGCAAGGAAAGAAAGTAGTAATTGTTGGCTGCGGTGCTCAGGGTCT
AAACCAAGGTTTAAACATGAGAGATTCAGGTCTGGATATTTCGTATGCATTGAGGAAAGAGGCAATTGCAGAAAA
GAGGGCCTCCTGGCGTAAAGCGACGGAAAATGGGTTCAAAGTTGGTACTTACGAAGAACTGATCCCTCAGGCAGA
TTTAGTGATTAACCTAACACCAGATAAGGTTCACTCAGACGTAGTAAGAACAGTTCAACCGCTGATGAAGGATGG
GGCAGCTTTAGGTTACTCTCATGGCTTTAATATCGTTGAAGTGGGCGAGCAGATCAGAAAAGATATAACAGTCGT
AATGGTTGCACCAAAGTGCCCAGGTACGGAAGTCAGAGAGGAGTACAAGAGGGGTTTTGGTGTACCTACATTGAT
CGCCGTACATCCTGAAAATGACCCCAAAGGTGAAGGTATGGCAATTGCGAAGGCATGGGCAGCCGCAACCGGAGG
TCATAGAGCGGGTGTGTTAGAGAGTTCTTTCGTAGCTGAGGTCAAGAGTGACTTAATGGGTGAACAAACCATTCT
GTGCGGAATGTTGCAGGCAGGGTCTTTACTATGCTTTGATAAATTGGTCGAAGAGGGTACAGATCCTGCCTATGC
TGAAAAGTTGATACAATTTGGTTGGGAGACAATCACCGAGGCACTTAAACAAGGTGGCATAACATTGATGATGGA
TAGACTTTCAAATCCGGCCAAGCTAAGAGCCTACGCCTTATCTGAGCAACTAAAAGAGATCATGGCACCATTATT
CCAAAAGCACATGGACGATATTATCTCCGGTGAGTTTTCCTCAGGAATGATGGCAGATTGGGCAAACGATGATAA
AAAGTTATTGACGTGGAGAGAAGAAACCGGCAAGACGGCATTCGAGACAGCCCCACAATACGAAGGTAAAATTGG
TGAACAAGAATACTTTGATAAGGGAGTATTGATGATAGCTATGGTGAAGGCAGGGGTAGAACTTGCATTCGAAAC
TATGGTTGACTCCGGTATCATTGAAGAATCTGCATACTATGAGTCTTTGCATGAATTGCCTTTGATAGCAAATAC
TATTGCAAGAAAAAGACTTTACGAGATGAATGTTGTCATATCAGACACTGCAGAATATGGTAATTACTTATTTAG
CTACGCATGTGTCCCGTTGTTAAAGCCCTTCATGGCCGAGTTACAACCTGGTGATTTGGGGAAGGCTATTCCGGA
AGGAGCGGTTGACAATGGCCAACTGAGAGACGTAAATGAAGCTATTCGTTCACATGCTATAGAACAGGTGGGTAA
AAAGCTGAGAGGATATATGACCGATATGAAAAGAATTGCAGTGGCAGGATGAAGATCCGCGGCCGCTCGAGTCAT
GTAATTAGTTATGTCACGCTTACATTCACGCCCTCCCCCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACA
ACCTGAAGTCTAGGTCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTT
TCTTTTTTTTCTGTACAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACG
CTCGAAGGCTTTAATTTGCGGCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTC
GTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCC
AGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGC
GACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGC

Figure 3-7. Nucleotide sequence of *SacI-NotI* fragment carrying the *S. cerevisiae* *TEF1* promoter:*S. cerevisiae ilv3ΔN*:*S. cerevisiae TDH3* promoter:*E. coli ilvC^{Q110V}*.

[00325]   Plasmid pGV1981 is a yeast high copy plasmid with *HIS3* as a marker for the expression of *E. coli* ilvC^{Q110V} and *Lactococcus lactis* ilvD. pGV1981 was generated by cloning a *SalI-BamHI* fragment (1.7 kb) carrying the *Lactococcus lactis* ilvD ORF (SEQ ID NO. 109 with a *SalI* and *BamHI* sites introduces at the 5' and 3' ends, respectively) into the *SalI-BamHI* of pGV1974 (8.5 kb), replacing the *S. cerevisiae Ilv3ΔN* ORF.

[00326]   Plasmid pGV2001 is a yeast high copy plasmid with *HIS3* as a marker for the expression of *E. coli* ilvC^{Q110V}. pGV2001 was generated by digesting pGV1974 with *SalI-BamHI* to remove the *S. cerevisiae Ilv3ΔN* ORF. The digest was treated with Klenow to fill-in the 5' overhangs, the larger 8.5 kb fragment was isolated and self-ligated.

[00327]   GEVO1803 was made by transforming GEVO1186 with the 6.7 kb pGV1730 (contains *S. cerevisiae TRP1* marker and the *CUP1* promoter-driven *B. subtilis alsS*) that had been linearized by digestion with NruI. Completion of the digest was confirmed by running a small sample on a gel. The digested DNA was then purified using Zymo Research DNA Clean and Concentrator and used in the transformation. Trp+ clones were confirmed for the correct integration into the *PDC1* locus by colony PCR using primer pairs 1440+1441 and 1442+1443 for the 5' and 3' junctions, respectively. Expression of *B. subtilis alsS* was confirmed by qRT-PCR using primer pairs 1323+1324

[00328]   GEVO2107 was made by transforming GEVO1803 with linearized, HpaI-digested pGV1914. Correct integration of pGV1914 at the *PDC6* locus was confirmed by analyzing candidate Ura+ colonies by colony PCR using primers 1440 plus 1441, or 1443 plus 1633, to detect the 5' and 3' junctions of the integrated construct, respectively. Expression of all transgenes were confirmed by qRT-PCR using primer pairs 1321 plus 1322, 1587 plus 1588, and 1633 plus 1634 to examine *Bs_alsS*, *Ll_kivD*, and *Dm_ADH* transcript levels, respectively.

[00329]   GEVO 2107 was transformed with plasmids that contained either a KARI alone (pGV2001 with *E. coli* ilvC^{Q110V}) or the same KARI with a DHAD (pGV1974 with the *S. cerevisiae Ilv3ΔN* or pGV1981 with the *L. lactis ilvD*). Fermentations were carried out with three independent transformants for each DHAD homolog being tested, as well as the no DHAD control plasmid. Seed cultures were grown in SCD-H medium to mid-log phase. The fermentations were initiated by collecting cells and resuspending in 25 mL of SCD-H (5% glucose) medium to an OD$_{600}$ of 1. Fermentations were performed aerobically in 125 mL unbaffled flasks shaken at 250 rpm at 30°C. At t = 0, 24, 48 and 72 hours OD$_{600}$'s were checked and 2 mL samples were taken. These samples were centrifuged at 18,000 x g in a microcentrifuge and 1.5 mL of the clarified media was transferred to a 1.5 mL Eppendorf tube. The clarified media was stored at 4°C until analyzed by GC and HPLC as described in General Methods. At 24 and 48 hours 2.5 mL of glucose from a 400 g/L stock solution was added to the cultures. Figure 3-8 shows the production of isobutanol in these fermentations. All values were adjusted for the dilution caused by the volume change from adding glucose. An increased amount of isobutanol was produced from the cells expressing Ll_ilvD.
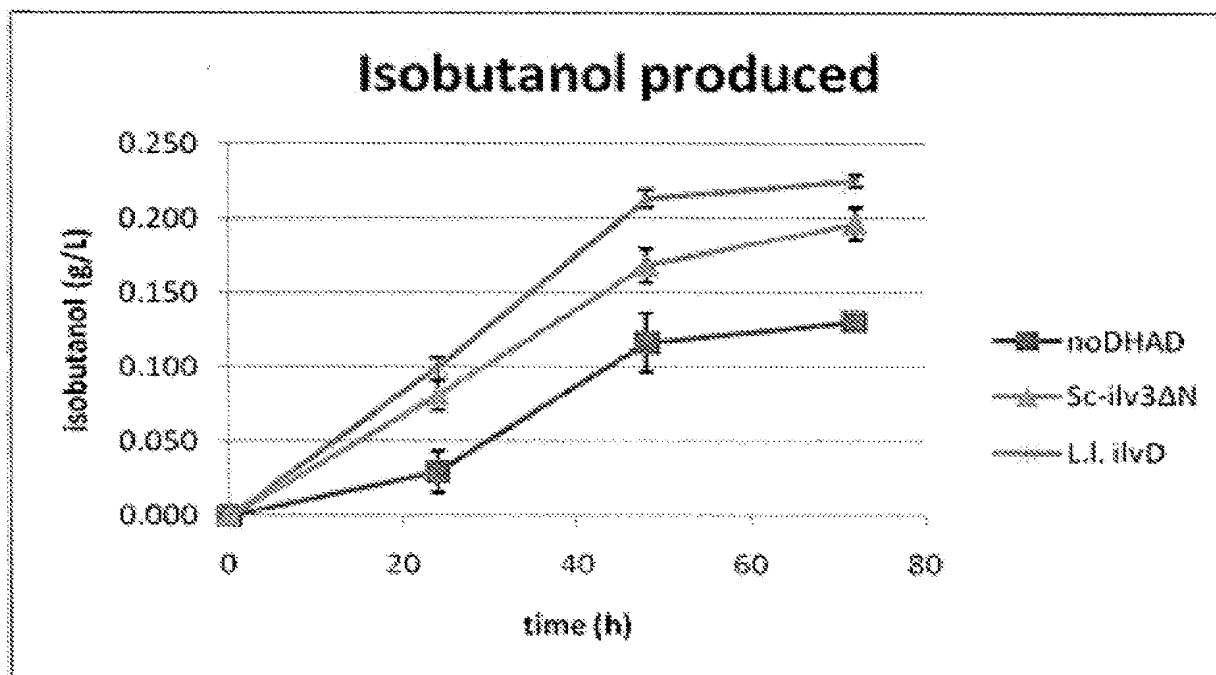
Figure 3-8. Results from fermentations of GEVO2107 (*pdc1::P_CUP1:Bs-alsS,TRP1/PDC1 pdc6::{P_ScTEF1:Ll_kivd P_ScTDH3:Dm_ADH URA3}/PDC6*) transformed with plasmids for expression of KARI and different DHAD homologs (shown in legend).

Example 4. Assaying DHAD activity in fractionated cell extracts

[00330] The purpose of this Example is to describe how DHAD activity can be measured in fractionated cellular extracts that are enriched for either mitochondrial or soluble cytosolic components.

Table 4-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
Gevo2244	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ*

Table 4-2 outlines the plasmids disclosed in this example:

pGV No.	Figure	Genotype
pGV1106	2-1	pUC ori, bla (AmpR), 2µm ori, URA3, *TDH3* promoter-Myc tag-polylinker-*CYC1* terminator
pGV1662	2-2, 2-3	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-(kivD)
pGV1855		pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Ll_ilvD*
pGV1900		pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*ScILV3*(FL)

Table 4-3 outlines the oligonucleotide primers disclosed in this example:

GEVO #	Sequence
271	CTAGCATGGAACAAAAACTCATCTCAGAAGAAGATGGTGTCGACGAA TTCCCGGGATCCGCGGCCGC
272	TCGAGCGGCCGCGGATCCCGGGAATTCGTCGACACCATCTTCTTCT GAGATGAGTTTTTGTTCCATG
421	GCCAACGGATCCTCAAGCATCTAAAACACAACCG
551	GCTCATGTCGACATGAAGAAGCTCAACAAGTACTCG
1617	CGTTGAGTCGACATGGGCTTGTTAACGAAAGTTGC
1618	GCCAACGGATCCTCAAGCATCTAAAACACAACCG

[00331] Plasmids pGV1106, pGV1662, and pGV1855 are described in Example 2 above.

[00332] Plasmid pGV1900 was generated by amplifying the full-length, native *ScILV3* nucleotide sequence from *S.cerevisiae* strain CEN.PK genomic DNA in a PCR reaction using primers 1617 and 1618. The resulting 1.76kb fragment, which contained the complete ScILV3 coding sequence (ScILV3(FL); SEQ ID NO.:111) flanked by 5' *Sal*I and 3' *Bam*HI restriction site sequences was digested with *Sal*I and *Bam*HI and ligated into pGV1662 which had been digested with *Sal*I and *Bam*HI.

[00333] To measure the DHAD activities present in fractionated cell extracts, the strain GEVO2244 was transformed singly with either pGV1106, which served as an empty vector control, or with one of: pGV1855, pGV1900, or pGV2019, which are expression plasmids for LI_ilvD and ScILV3(FL), respectively.

[00334] An independent clonal transformant of each plasmid was isolated, and a 1L culture of each strain was grown in SCGal-Ura+9xIV at 30°C at 250rpm. The OD$_{600}$ was noted, the cells were collected by centrifugation (1600xg, 2 min) and the culture medium was decanted. The cell pellets were resuspended in 50mL sterile deionized water, collected by centrifugation (1600xg, 2min), and the supernatant was discarded. The OD$_{600}$ and total wet cell pellet weight of each culture are listed in Table 4-4, below:

Table 4-4. OD$_{600}$ and pellet mass (g) of strain GEVO2244 transformed with the indicated plasmids.

Plasmid	OD$_{600}$	Pellet mass (g)
pGV1106	2.2	7.6
pGV1855	2.3	7.7
pGV1900	1.3	3.8
pGV2019	2.6	8.4

[00335] To obtain spheroplasts, the cell pellets were resuspnded in 0.1M Tris-SO$_4$, pH 9.3, to a final concentration of 0.1 g/mL, and DTT was added to a final concentration of 10mM. Cells were incubated with gentle (60 rev/min) agitation on an

orbital shaker for 20 min at 30°C, and the cells were then collect by centrifugation (1600xg, 2min) and the supernatant discarded. Each cell pellet was resuspended in spheroplasting buffer, which consists of (final concentrations): 1.2M sorbitol (Amresco, catalog #0691), 20mM potassium phosphate pH 7.4) and then collected by centrifugation (1600xg, 10min). Each cell pellet was resuspended in spheroplasting buffer to a final concentration of 0.1g cells/mL in a 500mL centrifuge bottle, and 50mg of Zymolyase 20T (Seikagaku Biobusiness, Code#120491) was added to each cell suspension. The suspensions were incubated overnight (approximately 16hrs) at 30°C with gentle agitation (60 rev/min) on an orbital shaker. The efficacy of spheroplasting was ascertained by diluting an aliquot of each cell suspension 1:10 in either sterile water or in spheroplasting buffer, and comparing the aliquots microscopically (under 40x magnification). In all cases, >90% of the water-diluted cells lysed, indicating efficient spheroplasting. The spheroplasts were centrifuged (3000xg, 10min, 20°C), and the supernatant was discarded. Each cell pellet was resuspended in 50mL spheroplast buffer without Zymolyase, and cells were collected by centrifugation (3000xg, 10min, 20°C).

[00336]   To fractionate spheroplasts, the cells were resuspended to a final concentration of 0.5 g/mL in ice cold mitochondrial isolation buffer (MIB), consisting of (final concentration): 0.6M D-mannitol (BD Difco Cat#217020), 20mM HEPES-KOH, pH 7.4. For each 1mL of resulting cell suspension, 0.01mL of Yeast/Fungal Protease Arrest solution (G Biosciences, catalog #788-333) was added. The cell suspension was subjected to 35 strokes of a Dounce homogenizer with the B (tight) pestle, and the resulting cell suspension was centrifuged (2500xg, 10min, 4°C) to collect cell debris and unbroken cells and spheroplasts. Following centrifugation, 2mL of each sample (1mL of the pGV1900 transformed cells) were saved in a 2mL centrifuge tube on ice and designated the "W" (for Whole cell extract) fraction, while the remaining supernatant was transferred to a clean, ice-cold 35mL Oakridge screw-cap tube and centrifuged (12,000xg, 20min, 4°C) to pellet mitochondria and other organellar structures. Following centrifugation, 5mL of each resulting supernatant was transferred to a clean tube on ice, being careful to avoid the small, loose pellet, and labelled the "S" (soluble cytosol) fraction. The resulting pellets were resuspended in MIB containing Protease Arrest solution, and were labelled the "P" ("pellet") fractions. Protein from the "P" fraction was released after dilution 1:5 in DHAD assay buffer (see above) by rapid mixing in a 1.5 mL tube with a Retsch Ball Mill MM301 in the presence of 0.1 mM glass beads. The bead-beating was performed 4 times for 1 minute, 30 beats per second, after which insoluble debris was removed by centrifugation (20,000xg, 10min, 4°C) and the soluble portion retained for use.

[00337]   The BioRad Protein Assay reagant (BioRad, Hercules, CA) was used according to manufacturer's instructions to determine the protein concentration of each fraction; the data are summarized in Table 4-5, below:

Table 4-5. Protein concentrations of mitochondrial/organellar (P) and cytosolic (S) fractions and whole cell (W) lysates, prepared as described in the text.

plasmid/fraction	protein [μg/μL]
1106 P	20.3
1855 P	17.7
1900 P	9.2
2019 P	19.7
1106 S	12.3
1855 S	12.9
1900 S	7.9
2019 S	12.4
1106 W	14.0
1855 W	15.0
1900 W	7.9
2019 W	14.7

[00338]   The DHAD activity of each fraction was ascertained as follows. In a fresh 1.5mL centrifuge tube, 50μL of each fraction was mixed with 50μL of 0.1M 2,3-dihydroxyisovalerate (DHIV), 25μL of 0.1M $MgSO_4$, and 375μL of 0.05M Tris-HCl pH 8.0, and the mixture was incubated for 30min at 35°C. Each reaction was carried out in triplicate. Each tube was then heated to 95°C for 5min to inactivate any enzymatic activity, and the solution was centrifuged (16,000xg for 5min) to pellet insoluble debris. To prepare samples for analysis, 100μL of each reaction were mixed with 100μL of a solution consisting of 4 parts 15mM dinitrophenyl hydrazine (DNPH) in acetonitrile with 1 part 50mM citric acid, pH 3.0, and the mixture was heated to 70°C for 30min in a thermocycler. Analysis of ketoisovalerate via HPLC was carried out as described in General Methods. Data from the experiment are summarized below in Table 4-6.

Table 4-6. Specific activities (KIV generation) and ratios of specific activities from fractionated lysates of *S.cerevisiae* strain GEVO2244 carrying plasmids to overexpress the indicated DHAD homolog. Each data point is the result of triplicate samples.

Lysate (pGV# and fraction*)	DHAD	Sp. Activity [U/mg protein in fraction]	Std. Dev.	Ratio of Sp. Activities (Cyto or Mito to Whole-Cell)
1106 WCL	---	n.d.		
1106 cyto	---	n.d.		
1106 mito	---	n.d.		
1855 WCL	LI_ilvD	0.0006	4.7E-05	
1855 cyto	LI_ilvD	0.0011	0.0001	1.76

1855 mito	LI_ilvD	2E-05	3.5E-05	0.03
1900 WCL	ScILV3(FL)	0.0096	0.0018	
1900 cyto	ScILV3(FL)	0.0052	0.0004	0.54
1900 mito	ScILV3(FL)	0.0340	0.0029	3.53

* WCL, whole cell lysate; cyto, cytosolic-enriched fraction; mito, mitochondrial (organellar)-enriched fraction

[00339]   Cells overexpressing LI_ilvD generated significantly greater proportion of DHAD activity in the cytosolic fraction versus the mitochondrial fraction, whereas cells overexpressing the full-length, native (mitochondrial) ScIlv3(FL) resulted in a greater proportion of the specific activity residing in the mitochondrial fraction.

Example 5: Targeting dihydroxy acid dehydratase to the yeast cytosol.

[00340]   The purpose of this is example is to demonstrate that DHAD enzymes can be targeted to the yeast cytosol.

Table 5-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
Gevo2242	*S.cerevisiae, CEN.PK; MAT-alpha ura3 leu2 his3 trp1 ilv5^{D255E} pdc1::Bs-alsS,TRP1*
Gevo2244	*S. cerevisiae, CEN.PK; MATα ura3 leu2 his3 trp1 ilv3Δ*

Table 5-2 outlines the plasmids disclosed in this example:

pGV No.	Genotype
pGV1106	pUC ori, bla (AmpR), 2µm ori, *URA3, TDH3* promoter-Myc tag-polylinker-*CYC1* terminator
pGV1662	pUC ori, bla (AmpR), 2µm ori, *URA3, TEF1* promoter-*(kivD)*
pGV1784	pUC ori, kanR, *Mm_ubiquitin* coding sequence
pGV1855	pUC ori, bla (AmpR), 2µm ori, *URA3, TEF1* promoter-*LI_ilvD*
pGV1897	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)
pGV1900	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*ScILV3*(FL)
pGV2019	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*ScILV3ΔN*
pGV2052	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)-*ScIlv3*(FL)
pGV2053	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)-*ScIlv3ΔN*
pGV2054	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)-*LI_ilvD*
pGV2055	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)-*Gf_ilvD*
pGV2056	pUC ori, bla (AmpR), 2µm ori, URA3, *TEF1* promoter-*Mm_ubiquitin*(Gly-X)-*Se_ilvD*

Table 5-3 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence (5' to 3')
1792	TTTTCTCGAGATGCAGATTTTTGTGAAGACCCTCACTG
1794	TTTTGCGGCCGCGGATCCGTCGACACCTCGCAGGCGCAACACCAGGT GCAG

[00341] To develop the constructs required to express DHAD as a fusion with an N-terminal ubiquitin, plasmid pGV1784 was synthesized by DNA2.0. This plasmid contained the synthesized sequence for the *Mus musculus* ubiqutiin gene, codon-optimized for expression in *S.cerevisiae* (*Mm_ubiquitin*, Figure 5-1).

ATGCAGATTTTTGTGAAGACCCTCACTGGCAAAACCATCACCCTTGAGGTCGAGCCCAGTGA
CACCATTGAGAATGTCAAAGCCAAAATTCAAGACAAGGAGGGTATCCCACCTGACCAGCAGC
GTCTGATATTTGCCGGCAAACAGCTGGAGGATGGCCGCACTCTCTCAGACTACAACATCCAG
AAAGAGTCCACCCTGCACCTGGTGTTGCGCCTGCGAGGTGGA

Figure 5-1. *Mus musculus* ubiquitin gene coding sequence

[00342] Using this plasmid as the template, the *Mm_ubiquitin* gene was amplified via PCR using primers 1792 and 1794 to generate a PCR product containing the *Mm_ubiquitin* gene codon sequence flanked by restriction sites *Xhol* and *Notl* at its 5' and 3' ends, respectively, and altered so as to lack the codon for its endogenous C-terminal most glycine residue (denoted as Gly-X). This PCR product was cloned into pGV1662 (described in Example 2), yielding pGV1897.
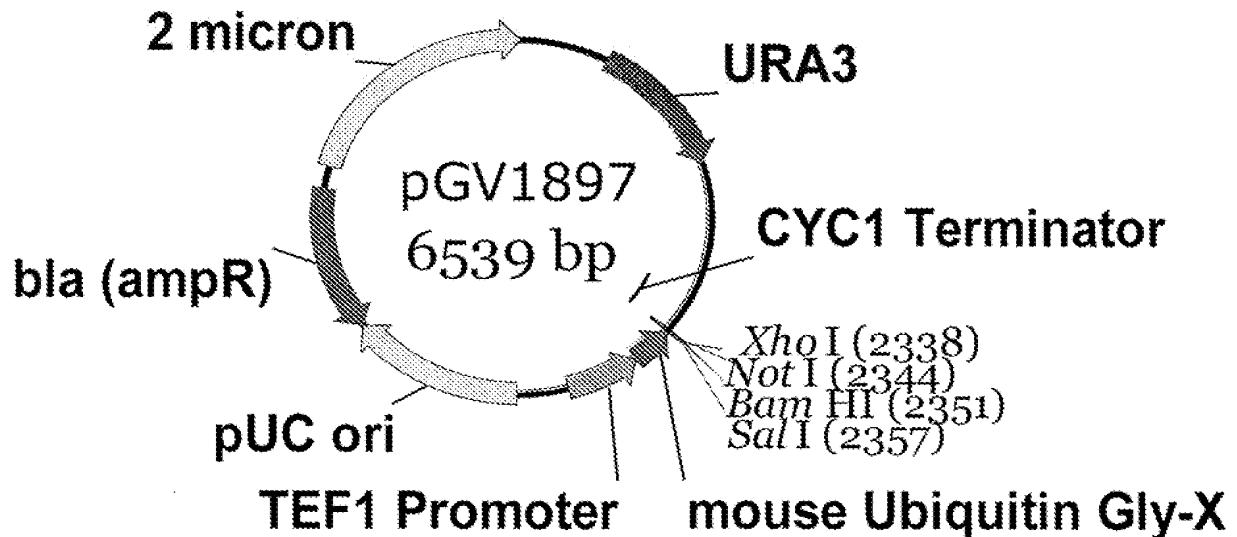
Figure 5-2. Schematic map of plasmid pGV1897

[00343] Plasmid pGV1897 was then used as a recipient cloning vector for sequences encoding *S.cerevisiae ILV3* (ScIlv3(FL), SEQ ID NO. 111), *S.cerevisiae Ilv3ΔN* (ScIlv3ΔN, SEQ ID NO. 117), *Lactococcus lactis ilvD* (Ll_ilvD, SEQ ID NO. 109), *Gramella forsetti ilvD* (Gf_ilvD, SEQ ID NO. 108), and *Saccharopolyspora*

erythraea ilvD (*Se_ilvD*, SEQ ID NO. 110), yielding plasmids pGV2052-2056, respectively.

[00344] The DHAD activity exhibited by cells transformed with each of the resulting constructs is ascertained by *in vitro* assay. GEVO2244 is transformed (singly) with pGV2052-2056, pGV1106 (empty control vector), pGV1855 (expressing native, unfused *Ll_ilvD*), pGV1900 (expressing native, full-length *Sc_ILV3*(FL)), or pGV2019 (expressing unfused *Sc_ILV3ΔN*). Lysates of transformants are prepared and DHAD activity is assayed as described in Example 2. DHAD activity in mitochondrial/organellar (P) and cytosolic (S) fractions and whole cell (W) lysates is assayed as described in Example 4.

[00345] In an analagous manner, a desired ALS (e.g., Bs_alsS) or KARI gene whose product is known or predicted to be mitochondrial can be re-targeted to the cytosol by means of the methods detailed in this example. The nucleotide sequence encoding for a full-length, or variant, ALS or KARI is amplified by PCR using primers that introduce restriction sites convenient for cloning the final product as an in-frame fusion of the Mm_ubiquitin gene. The resulting construct is transformed into a host S.cerevisiae cell suitable for assaying the in vitro activity of the expressed Mm_ubiquitin-gene chimeric fusion protein, using methods described in Examples 2 and 4.

Example 6: Alternative, native dehydratases with DHAD activity.

[00346] This example describes how the overexpression of native dehydratases in *S. cerevisiae* for the conversion of 2,3-dihydroxyisovalerate to ketoisovalerate is measured.

Table 6-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
Gevo2244	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ*

Table 6-2 outlines the plasmids disclosed in this example:

pGV No.	Genotype
p426TEF	P_{TEF1}:MCS:T_{CYC1}, URA3, 2-micron, bla, pUC-ori (Mumberg, D. et al. (1995) *Gene* 156:119-122; obtained from ATCC)
1102	P_{TEF1}:HA-tag:MCS:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
1106	P_{TDH3}:myc-tag:MCS:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
1662	P_{TEF1}:Ll_kivd:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
1894	P_{TEF1}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
2000	P_{TEF1}:Sc_ILV3ΔN:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
2111	P_{TEF1}:Ll_ilvD:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
2112	P_{TEF1}:Sc_LEU1:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, URA3, 2-micron, bla, pUC-ori
2113	P_{TEF1}:Sc_HIS3:P_{TDH3}:Ec_ilvC^{Q110V}-coSc:T_{CYC1}, URA3, 2-micron, bla, pUC-ori

Table 6-3 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence (5' to 3')
269	CTAGCATGTACCCATACGATGTTCCTGACTATGCGGGTGTCGACGAAT TCCCGGGATCCGCGGCCGC
270	TCGAGCGGCCGCGGATCCCGGGAATTCGTCGACACCCGCATAGTCAG GAACATCGTATGGGTACATG
271	CTAGCATGGAACAAAAACTCATCTCAGAAGAAGATGGTGTCGACGAAT TCCCGGGATCCGCGGCCGC
272	TCGAGCGGCCGCGGATCCCGGGAATTCGTCGACACCATCTTCTTCTGA GATGAGTTTTTGTTCCATG
1842	TTTTGGATCCCTACCAATCCTGGTGGACTTTATCG
2163	TTGGTAGTCGACATGGTTTACACTCCATCCAAGGGTC
2183	ACAGTAGTCGACATGACAGAGCAGAAAGCCCT
2184	TACATCGGATCCCTACATAAGAACACCTTTGGTG

[00347] Plasmid pGV1102 was generated by inserting a linker (primers 269 annealed to primer 270) containing a HA-tag and a new MCS (*SalI-EcoRI-SmaI-BamHI-NotI*) into the *SpeI* and *XhoI* sites of p426TEF.

[00348] Plasmids pGV1106, pGV1662 are described in Example 2.

[00349] Plasmid pGV1894 is a yeast high copy plasmid with *URA3* as a marker for the expression of *E. coli* ilvC^{Q110V} and was generated by cloning a *XhoI-NotI* fragment (1.5 kb) carrying the *E. coli* ilvC^{Q110V} ORF (SEQ ID NO. 118) into the *SalI-NotI* of pGV1662 (6.3 kb), replacing the *Ll_kivd* ORF.

[00350] Plasmids pGV2000, pGV2111, pGV2112, and pGV2113 are yeast high copy plasmids with *URA3* as a marker for the expression of *E. coli* ilvC^{Q110V} and a DHAD.

[00351] pGV2000 is generated by cloning a *SacI-NotI* fragment (4.9 kb) from pGV1974 (described in Example 3) carrying the *S. cerevisiae TEF1* promoter:*S. cerevisiae Ilv3ΔN*:*S. cerevisiae TDH3* promoter:*E. coli* ilvC^{Q110V} into the *SacI-NotI* sites of pGV1106 (6.6 kb), a yeast expression plasmid carrying the *URA3* marker.

[00352] pGV2111 is generated by cloning a *SalI-BamHI* fragment (1.7 kb) carrying the *Lactococcus lactis ilvD* ORF (SEQ ID NO. 109 with a *SalI* and *BamHI* sites introduces at the 5' and 3' ends, respectively) into the *SalI-BamHI* of pGV2000 (8.4 kb), replacing the *S. cerevisiae Ilv3ΔN* ORF.

[00353] pGV2112 is generated by cloning the *S. cerevisiae LEU1* gene as a *SalI-BamHI* fragment (2.3 kb), generated by PCR using primers 2163 and 1842 using genomic DNA as template, into the *SalI-BamHI* of pGV2000 (8.4 kb), replacing the *S. cerevisiae Ilv3ΔN* ORF.

[00354] pGV2113 is generated by cloning the *S. cerevisiae HIS3* gene as a *SalI-BamHI* fragment (0.7 kb), generated by PCR using primers 2183 and 2184 using genomic DNA as template, into the *SalI-BamHI* of pGV2000 (8.4 kb), replacing the *S. cerevisiae Ilv3ΔN* ORF.

[00355] DHADs are tested for *in vitro* activity using whole cell lysates. The DHADs as well as *LEU1* and *HIS3* are expressed from pGV2000, pGV2112, and pGV2113 GEVO2244 to minimize endogenous DHAD background activity. A

plasmid that does not express DHAD, pGV1894, and a plasmid that expresses *LI_ilvD*, pGV2111, are used as negative and positive controls, repectively.

[00356] To grow cultures for cell lysates, triplicate independent cultures of each desired strain are grown overnight in 3 mL YNBD+HLW+10xIV at 30°C, 250rpm. The following day, the overnight cultures are diluted 1:50 into 50mL fresh YNBD+HLW+10xIV in a 250 mL baffle-bottomed Erlenmeyer flask and incubated at 30°C at 250rpm. After approximately 10 hours, the OD_{600} of all cultures are measured, and the cells of each culture are collected by centrifugation (2700x*g*, 5 min). The cell pellets are washed by resuspending in 1mL of water, and the suspension is placed in a 1.5mL tube and the cells are collected by centrifugation (16,000x*g*, 30 seconds). All supernatant is removed from each tube and the tubes are frozen at -80°C until use.

[00357] Lysates are prepared by resuspending each cell pellet in 0.7mL of lysis buffer. Lysate lysis buffer consisted of: 0.1M Tris-HCl pH 8.0, 5mM MgSO$_4$, with 10 μL of of Yeast/Fungal Protease Arrest solution (G Biosciences, catalog #788-333) per 1mL of lysis buffer. Eight hundred microliters of cell suspension are added to 1mL of 0.5mm glass beads that had been placed in a chilled 1.5mL tube. Cells are lysed by bead beating (6 rounds, 1 minute per round, 30 beats per second) with 2 minutes chilling on ice in between rounds. The tubes are then centrifuged (20,000x*g*, 15 min) to pellet debris and the supernatant (cell lysates) are retained in fresh tubes on ice. The protein concentration of each lysate is measured using the BioRad Bradford protein assay reagent (BioRad, Hercules, CA) according to manufacturer's instructions.

[00358] The DHAD activity of each lysate is ascertained as follows. In a fresh 1.5mL centrifuge tube, 50μL of each lysate is mixed with 50μL of 0.1M 2,3-dihydroxyisovalerate (DHIV), 25μL of 0.1M MgSO$_4$, and 375μL of 0.05M Tris-HCl pH 8.0, and the mixture is incubated for 30min at 35°C. Each tube is then heated to 95°C for 5min to inactivate any enzymatic activity, and the solution is centrifuged (16,000xg for 5min) to pellet insoluble debris. To prepare samples for analysis, 100μL of each reaction are mixed with 100μL of a solution consisting of 4 parts 15mM dinitrophenyl hydrazine (DNPH) in acetonitrile with 1 part 50mM citric acid, pH 3.0, and the mixture is heated to 70°C for 30min in a thermocycler. The solution is then analyzed by HPLC as described above in General Methods to quantitate the concentration of ketoisovalerate (KIV) present in the sample.

[00359] DHADs are tested for *in vitro* activity using whole cell lysates. The DHADs are expressed in a yeast deficient for DHAD activity (GEVO2244; *ilv3Δ*) to minimize endogenous background activity.

<u>Example 7A: Cloning of low-abundance, endogenous cytosolic iron-sulfur cluster assembly machinery for overexpression in *S.cerevisiae.*</u>

[00360] The purpose of this example is to describe how three known components of the *S.cerevisiae* cytosolic iron-sulfur assembly machinery were cloned to permit their overexpression in *S.cerevisiae*, to increase cytosolic DHAD activity.

[00361] In the yeast *S.cerevisiae*, at four least genes—*CIA1*, *CFD1*, *NAR1*, and *NBP35*—encode activities that contribute to the proper assembly and/or transfer of iron-sulfur [Fe-S] clusters of cytosolic proteins. Of these four genes, three-- *CFD1*, *NAR1*, and *NBP35*—have been shown to be expressed at very low levels during aerobic growth on glucose (Ghaemmaghami, S., *et al.*, *Nature* (2003):425(16), 737-741). These three genes thus represent attractive candidates for overexpression, to increase the cellular capacity for proper cytosolic [Fe-S] cluster protein assembly.

Table 7A-1 outlines the plasmids disclosed in this example:

pGV No.	Fig. No.	Genotype
pGV2074	7A-1	pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter, *TEF1* promoter, *TDH3* promoter
pGV2127		pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter, *TEF1* promoter, *TDH3* promoter-*CFD1*
pGV2138		pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter, *TEF1* promoter-*NAR1*, *TDH3* promoter-*CFD1*
pGV2144		pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter-*NBP35*, *TEF1* promoter, *TDH3* promoter
pGV2147	7A-2	pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter-*NBP35*, *TEF1* promoter-*NAR1*, *TDH3* promoter-*CFD1*

Table 7A-2 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence (5' to 3')
GEVO2195	TTGTTCCTCGAGATGGAGGAACAGGAGATAGGCGTTCCTGC
GEVO2196	GTTCTTGCGGCCGCTTATTTTGGAGATTCTATCTGGGGTTGC
GEVO2197	TTCTTGGTCGACATGAGTGCTCTACTGTCCGAGTCTGACC
GEVO2198	TTGTTCGGATCCTTACCAGGTGCTCCCAACAGAGACGAGATCC
GEVO2259	TCAGTAAGATCTATGACTGAGATACTACCACATGTAAACGAC
GEVO2260	CATATCCTCGAGGTACCCTATACATCCCCCACAGCATCTCGCAG

[00362] To clone the sequences for *CFD1*, *NAR1*, and *NBP35* into an appropriate *S.cerevisiae* expression vector, the following steps were carried out: Vector pGV2074 was used as a parental plasmid for subsequent cloning steps described below. A schematic map of pGV2074 is shown in Figure 7A-1, and the nucleotide sequence is given in Figure 7A-3. The salient features of pGV2074 include a bacterial origin of replication (pUC) and selectable marker (*bla*), an *S.cerevisiae* 2µm origin of replication and selectable marker (the *hph* gene, conferring resistance to hygromycin, operably linked to the *TPI1* promoter region), and sequences containing the *S.cerevisiae* promoters for the *PGK1*, *TDH3* and

TEF1 genes, each followed by one or more unique restriction sites to facilitate the introduction of coding sequences.

[00363] First, the *CFD1* coding sequence was amplified from *S.cerevisiae* genomic DNA by PCR, using primers 2195 and 2196, which also added 5' *XhoI* and 3' *NotI* sites, respectively. The resulting ~890bp PCR product was digested with *XhoI* plus *NotI* and ligated into pGV2074 that had been digested with *XhoI* plus *NotI*, yielding the plasmid pGV2127. All sequences amplified by PCR were confirmed by DNA sequencing. Next, the *NAR1* coding sequence was amplified from *S.cerevisiae* genomic DNA by PCR, using primers 2197 and 2198, which added 5' *SalI* and 3' *BamHI* sites, respectively. The resulting ~1485bp product was digested with *SalI* plus *BamHI* and cloned into pGV2127 which had also been digested with *SalI* plus *BamHI*, thereby yielding pGV2138. Next, the *NBP35* coding sequence was amplified *S.cerevisiae* genomic DNA by PCR, using primers 2259 and 2260, which added 5' *BglII* and 3' *KpnI* and *XhoI* (from 5' to 3') sites, respectively. The resulting ~995 bp product was digested with *BglII* plus *XhoI* and ligated into pGV2074 that had been digested with *BglII* plus *SalI*, yielding pGV2144. Finally, pGV2144 was digested with *AvrII* plus *BamHI*, and the resulting 1.78kb fragment (which contained the *PGK1* promoter and the *NBP35* ORF sequence) was gel purified and ligated into the vector pGV2138 that had digested with *AvrII* plus *BglII*, yielding pGV2147. A schematic map of plasmid pGV2147 is shown in Figure 7A-2.
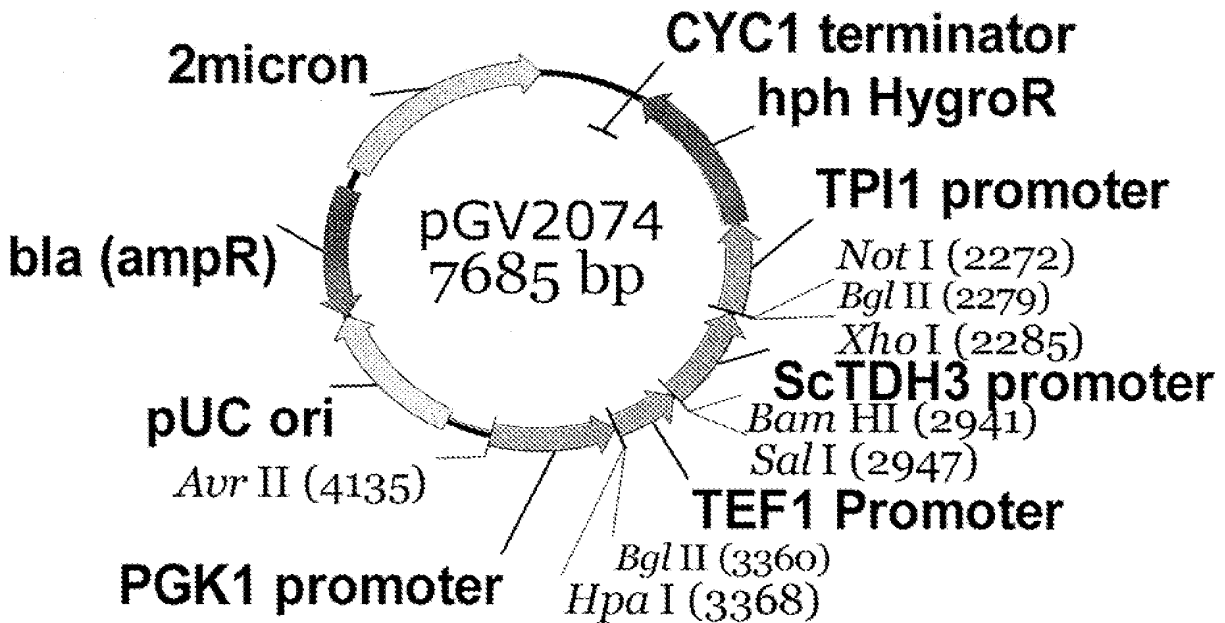
Figure 7A-1: Schematic map of plasmid pGV2074.
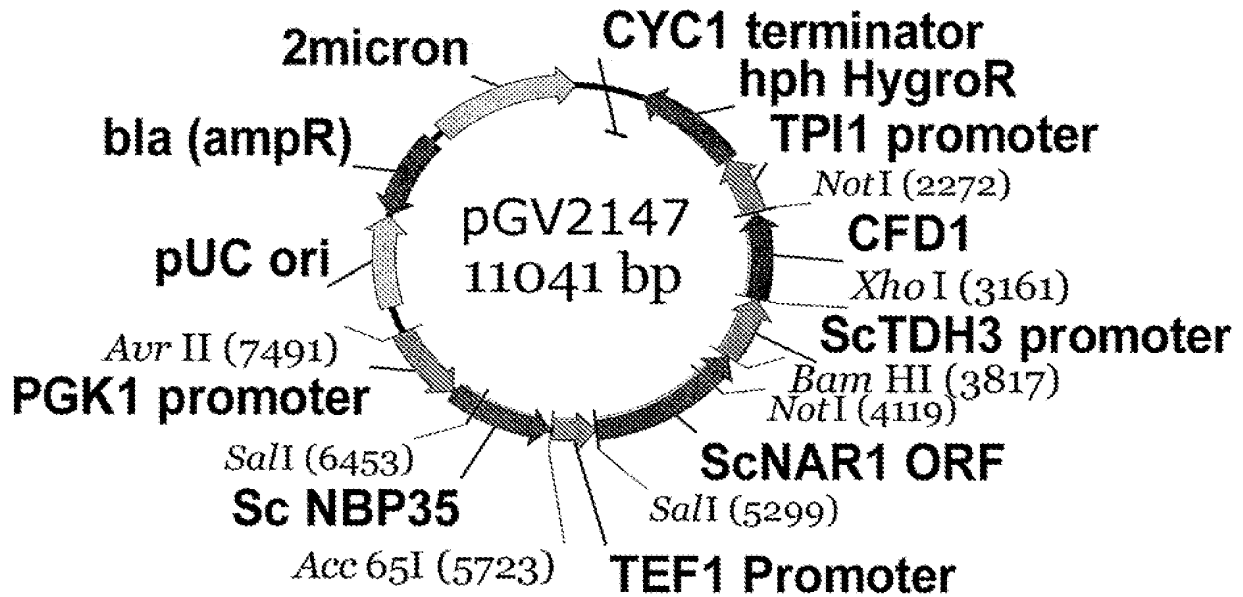
Figure 7A-2: Schematic map of plasmid pGV2147.

```
ttggatcatactaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc
tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaag
cggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatg
cggcatcagagcagattgtactgagagtgcaccataccacagcttttcaattcaattcatcattttttttttatt
ctttttttttgatttcggtttctttgaaattttttgattcggtaatctccgaacagaaggaagaacgaaggaagg
agcacagacttagattggtatatatacgcatatggcaaattaaagccttcgagcgtcccaaaaccttctcaagca
aggttttcagtataatgttacatgcgtacacgcgtctgtacagaaaaaaaagaaaaatttgaaatataaataacg
ttcttaatactaacataactataaaaaaataaatagggacctagacttcaggttgtctaactccttccttttcgg
ttagagcggatgtggggggaggggcgtgaatgtaagcgtgacataagaattcttattcctttgccctcggacgagt
gctggggcgtcggtttccactatcggcgagtacttctacacagccatcggtccagacggccgcgcttctgcgggc
gatttgtgtacgcccgacagtcccggctccggatcggacgattgcgtcgcatcgaccctgcgcccaagctgcatc
atcgaaattgccgtcaaccaagctctgatagagttggtcaagaccaatgcggagcatatacgcccggaggcgcgg
cgatcctgcaagctccggatgcctccgctcgaagtagcgcgtctgctgctccatacaagccaaccacggcctcca
gaagaggatgttggcgacctcgtattgggaatccccgaacatcgcctcgctccagtcaatgaccgctgttatgcg
gccattgtccgtcaggacattgttggagccgaaatccgcatgcacgaggtgccggacttcgggcagtcctcggc
ccaaagcatcagctcatcgagagcctgcgcgacggacgcactgacggtgtcgtccatcacagtttgccagtgata
cacatggggatcagcaatcgcgcatatgaaatcacgccatgtagtgtattgaccgattccttgcggtccgaatgg
gccgaaccgctcgtctggctaagatcggccgcagcgatcgcatccatggcctccgcgaccggctggagaacagc
gggcagttcggtttcaggcaggtcttgcaacgtgacaccctgtgcacggcgggagatgcaataggtcaggctctc
gctgaactccccaatgtcaagcacttccggaatcgggagcgcggccgatgcaaagtgccgataaacataacgatc
tttgtagaaaccatcggcgcagctatttacccgcaggacatatccacgccctcctacatcgaagctgaaagcacg
agattcttcgccctccgagagctgcatcaggtcggagacgctgtcgaacttttcgatcagaaacttctcgacaga
cgtcgcggtgagttcaggctttttacccatactagttttttagttatgtatgtgttttttgtagttatagatta
agcaagaaaagaatacaaacaaaaaattgaaaaagattgatttagaattaaaaagaaaaatatttacgtaagaag
ggaaaatagtaaatgttgcaagttcactaaactcctaaattatgctgcctttatattccctgttacagcagccg
agccaaaggtatataggctcctttgcattagcatgcgtaacaaaccacctgtcagtttcaaccgaggtggtatcc
gagagaattgtgtgattgctttaattaatttcggagaatctcacatgccactgaagattaaaaactggatgccag
aaaagggtgtccaggtgtaacatcaatagaggaagctgaaaagtcttagaacgggtaatcttccaccaacctga
tgggttcctagatataatctcgaagggaataagtaggtgataccgcagaagtgtctgaatgtattaaggtcctc
acagtttaaatcccgctcacactaacgtaggattattataactcaaaaaaatggcattattctaagtaagttaaa
tatccgtaatctttaaacagcggccgcagatctctcgagtcgaaactaagttctggtgttttaaaactaaaaaa
agactaactataaaagtagaatttaagaagtttaagaaatagatttacagaattacaatcaatacctaccgtctt
tatatacttattagtcaagtaggggaataatttcagggaactggtttcaacctttttttttcagcttttttccaaat
cagagagagcagaaggtaatagaaggtgtaagaaaaatgagatagatacatgcgtgggtcaattgccttgtgtcat
```

```
catttactccaggcaggttgcatcactccattgaggttgtgcccgttttttgcctgtttgtgcccctgttctctg
tagttgcgctaagagaatggacctatgaactgatggttggtgaagaaaacaatattttggtgctgggattctttt
tttttctggatgccagcttaaaaagcgggctccattatatttagtggatgccaggaataaaactgttcacccagac
acctacgatgttatatattctgtgtaacccgccccctattttgggcatgtacgggttacagcagaattaaaaggc
taattttttgactaaataaagttaggaaaatcactactattaattatttacgtattctttgaaatggcgagtatt
gataatgataaactggatccgtcgacaaacttagattagattgctatgctttctttctaatgagcaagaagtaaa
aaaagttgtaatagaacaagaaaaatgaaactgaaacttgagaaattgaagaccgtttattaacttaaatatcaa
tgggaggtcatcgaaagagaaaaaatcaaaaaaaaaattttcaagaaaaagaaacgtgataaaaattttattg
cctttttcgacgaagaaaaagaaacgaggcggtctcttttttcttttccaaacctttagtacgggtaattaacga
caccctagaggaagaaagagggaaatttagtatgctgtgcttggtgtgttttgaagtggtacggcgatgcgcgga
gtccgagaaatctggaagagtaaaaaaggagtagaaacattttgaagctatgagctcagatctgttaaccttgt
tttatatttgttgtaaaaagtagataattacttccttgatgatctgtaaaaaagagaaaaagaaagcatctaaga
acttgaaaaactacgaattagaaaagaccaaatatgtatttcttgcattgaccaatttatgcaagtttatatata
tgtaaatgtaagtttcacgaggttctactaaactaaaccacccccttggttagaagaaaagagtgtgtgagaaca
ggctgttgttgtcacacgattcggacaattctgtttgaaagagagagagtaacagtacgatcgaacgaactttgc
tctggagatcacagtgggcatcatagcatgtggtactaaaccctttcccgccattccagaaccttcgattgcttg
ttacaaaacctgtgagccgtcgctaggaccttgttgtgtgacgaaattggaagctgcaatcaataggaagacagg
aagtcgagcgtgtctgggtttttttcagttttgttctttttgcaaacaaatcacgagcgacggtaatttcttctc
gataagaggccacgtgctttatgagggtaacatcaattcaagaaggagggaaacacttccttttctggccctga
taatagtatgagggtgaagccaaaataaaggattcgcgcccaaatcggcatctttaaatgcaggtatgcgatagt
tcctcactctttccttactcacgagtaattcttgcaaatgcctattatgcagatgttataatatctgtgcgtctt
gagttgagcctagggagctccagctttttgttccctttagtgagggttaattgcgcgcttggcgtaatcatggtca
tagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacataggagccggaagcataaagtgtaaa
gcctggggtgcctaatgagtgaggtaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaac
ctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgct
tcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaata
cggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgt
aaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagt
cagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcct
gttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctca
cgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagccc
gaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagca
gccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactac
ggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagc
tcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggg
attttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatc
taaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgt
ctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggc
cccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccgga
agggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctaga
gtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcg
tttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaa
gcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggca
gcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtca
ttctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagc
agaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgaga
tccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtga
gcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttc
ctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaa
aataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgaacgaagcatctgtgcttcattttgt
agaacaaaaatgcaacgcgagagcgctaatttttcaaacaaagaatctgagctgcatttttacagaacagaaatg
caacgcgaaagcgctattttaccaacgaagaatctgtgcttcattttttgtaaaacaaaaatgcaacgcgagagcg
ctaattttttcaaacaaagaatctgagctgcattttttacagaacagaaatgcaacgcgagagcgctattttaccaa
caaagaatctatacttcttttttgttctacaaaaatgcatcccgagagcgctattttttctaacaaagcatcttag
```

```
attacttttttttctcctttgtgcgctctataatgcagtctcttgataactttttgcactgtaggtccgttaaggt
tagaagaaggctactttggtgtctattttctcttccataaaaaaagcctgactccacttcccgcgtttactgatt
actagcgaagctgcgggtgcattttttcaagataaaggcatcccgattatattctataccgatgtggattgcgc
atactttgtgaacagaaagtgatagcgttgatgattcttcattggtcagaaaattatgaacggtttcttctattt
tgtctctatatactacgtataggaaatgtttacattttcgtattgttttcgattcactctatgaatagttcttac
tacaattttttgtctaaagagtaatactagagataaacataaaaaatgtagaggtcgagtttagatgcaagttc
aaggagcgaaaggtggatgggtaggttatatagggatatagcacagagatatatagcaaagagatacttttgagc
aatgtttgtggaagcggtattcgcaatattttagtagctcgttacagtccggtgcgttttggttttttgaaagt
gcgtcttcagagcgctttggttttcaaaagcgctctgaagttcctatactttctagagaataggaacttcggaa
taggaacttcaaagcgtttccgaaaacgagcgcttccgaaatgcaacgcgagctgcgcacatacagctcactgt
tcacgtcgcacctatatctgcgtgttgcctgtatatatatatacatgagaagaacggcatagtgcgtgtttatgc
ttaaatgcgtacttatatgcgtctatttatgtaggatgaaaggtagtctagtacctcctgtgatattatcccatt
ccatgcggggtatcgtatgcttccttcagcactaccctttagctgttctatatgctgccactcctcaattggatt
agtctcatccttcaatgctatcatttcctttgata
```

Figure 7A-3: Nucleotide sequence of plasmid pGV2074.

Example 7B: Cloning of heterologous cytosolic iron-sulfur cluster assembly
machinery for overexpression in *S.cerevisiae*.

[00364] The purpose of this example is to describe how one or more cytosolic
iron-sulfur assembly machinery components, from various species, can be cloned to
permit their overexpression in *S.cerevisiae*, thereby increasing cytosolic DHAD
activity.

[00365] In addition to the endogenous cytosolic iron-sulfur assembly machinery
found in *S.cerevisiae*, homologous sequences and activities have been identified in
other microbial and eukaryotic species. In one example, the ApbC protein of
Salmonella enterica serovar Typhimurium has been shown, *in vitro*, to bind and
effectively transfer iron-sulfur clusters to a known cytosolic [Fe-S] cluster-containing
S.cerevisiae substrate, Leu1 (Boyd, J.M., et al., Biochemistry (2008), 47(31):8195-
202). Thus, a number of other useful homologs of the known *S.cerevisiae* cytosolic
iron-sulfur assembly machinery components exist and present attractive candidates
for overexpression in *S.cerevisiae*. Table 7B-1 lists several exemplary homologs
and their GenBank accession numbers, as identified by previous homology searches
(Boyd, J.M., et al., Journal of Biological Chemistry (2009), 284(1):110-118). Also
included in the table are two closely related *S.cerevisiae* homologs, Nbp35 and Cfd1.
Of note, Ind1 is reported to be localized to and functional in the mitochondria (Bych,
K., et al., EMBO J. (2008), 27(12):1736-46) whereas Hcf101 is reported to
participate in iron-sulfur cluster assembly in *Arabidopsis* chloroplasts (Lezhneva, L.,
et al., Plant Journal for Cell and Molecular Biology (2004); 37, 174-185).

Table 7B-1. Functionally homologous proteins involved in iron-sulfur cluster
formation.

Gene	Source, Accession Number
ApbC	*Salmonella enterica* serovar Typhimurium LT2, NP_461098
Ind1	*Yarrowia lypolytica*, YALI0B18590g
Hcf101	*Arabidopsis thaliana*, AAR97892.1
Nubp1	*Homo sapiens*, NP_002475.2
Nbp35	*S.cerevisiae*, CAA96797.1

Cfd1	S.cerevisiae, AAS56623

[00366] The cloning of one or more of these genes is carried out using techniques well known to one skilled in the art. Oligonucleotide primers are designed that are homologous to the 5' and 3' ends of each desired reading, and which furthermore incorporate a restriction site sequence convenient for the cloning of each reading frame into vector pGV2074. A standard PCR reaction is used to amplify each gene, either from the genome of each host organism, or from an in vitro synthesized DNA fragment, and the resulting PCR product is cloned into an expression vector (pGV2074). In the case of a protein known to be targeted to the mitochondira, such as *Yarrowia lypolytica* Ind1, PCR primers are designed to amplify the majority of the coding sequence while excluding the known N-terminal mitochondrial targeting sequence (Bych, K., et al., EMBO Journal (2008) **27**:1736-1746)

Example 7C: Overexpression of *S.cerevisiae* cytosolic iron-sulfur assembly machinery to increase cytosolic DHAD activity

[00367] The purpose of this example is to describe how a plasmid expressing one or more iron-sulfur assembly machinery components is co-expressed with a DHAD, thereby increasing the cytosolic activity of the DHAD.

Table 7C-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
GEVO2244	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ*

Table 7C-2 details the relevant features of plasmids disclosed in this example:

pGV No.	Genotype
pGV1851	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Gramella forsetti ilvD*
pGV1852	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Chromohalobacter salexigens ilvD*
pGV1853	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Ralstonia eutropha ilvD*
pGV1854	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Saccharopolyspora erythraea ilvD*
pGV1855	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter-*Ll_ilvD*
pGV1904	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *ScILV3(FL)*
pGV1905	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Acidobacteria bacterium Ellin345 ilvD*
pGV1906	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Picrophilus torridus DSM 9790 ilvD*
pGV1907	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Piromyces species E2 ilvD*
pGV2074	pUC ori, bla (AmpR), 2μm ori, *URA3*, *TEF1* promoter- *Sulfolobus tokodaii strain 7 ilvD*

pGV2147	pUC ori, bla (AmpR), 2µm ori, *TPI1* promoter-*hph* (HygroR), *PGK1* promoter-*NBP35*, *TEF1* promoter-*NAR1*, *TDH3* promoter-*CFD1*

[00368] Strain GEVO2244 is simultaneously co-transformed with one of: pGV1851, pGV1852, pGV1853, pGV1854, pGV1855, pGV1904, pGV1905, pGV1906, or pGV1907; plus, one of either: pGV2074 (which serves as an empty-vector control) or pGV2147 (which serves as the cytosolic Fe-S cluster machinery overexpression plasmid), and doubly-transformed cells are selected by plating onto SCD-Ura+9xIV containing 0.1g/L Hygromycin B.

[00369] Three independent isolates from each transformation are cultured in SCD-Ura+9xIV containing 0.1g/L Hygromycin B to obtain a cell mass suitable for preparation of a lysate, as described in Example 2. Lysates are prepared from each culture, and the resulting lysates are assayed for DHAD activity as described in Example 2. To further confirm that the increased DHAD activity is due specifically to increased cytosolic activity, cultures of GEVO2244 containing pGV1855 plus either pGV2074 or pGV2147 are grown in SCD-Ura+9xIV containing 0.1g/L Hygromycin B as otherwise described in Example 4. Fractionated lysates are prepared and in vitro assays to measure DHAD activity are further carried out as described in Example 4.

Example 8: Deletion of *LEU1*.

[00370] The purpose of this example is to describe the deletion of *LEU1* to increase the iron-sulfur cluster availability in the yeast cytosol.

Table 8-1 details the genotype of strains disclosed in this example:

GEVO No.	Genotype / Source
GEVO2244	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ*
GEVO2570	*S. cerevisiae*, CEN.PK; *MATα ura3 leu2 his3 trp1 ilv3Δ leu1::Kl_URA3*

Table 8-2 outlines the plasmids disclosed in this example:

pGV No.	Genotype
pGV1299	*K. lactis URA3, bla, pUC-ori (GEVO)*
pGV1981	P_{TEF1}:*Lactococcus lactis ilvD-coSc*:P_{TDH3}:*Ec_ilvC^{Q110V}-coSc*:T_{CYC1}, *HIS3, 2-micron, bla, pUC-ori*
pGV2001	P_{TEF1}:P_{TDH3}:*Ec_ilvC^{Q110V}-coSc*:T_{CYC1}, *HIS3, 2-micron, bla, pUC-ori*

Table 8-3 outlines the primers sequences disclosed in this example:

Gevo No.	Sequence (5' to 3')
587	CCAATGCAGACCGATCTTCTACCC
588	GATCACGTGATCTGTTGTATTG
2167	TACATGGGGTACTTCTCCTC
2170	CAGTCAACAAATATAAAGAATATTGAAATTGACAGTTTTTGTCGCTATCG ATTTTTATTA
2171	TTTTGTCGCTATCGATTTTTATTATTTGCTGTTTTAAATCATTCTGGTTCT ATCGAGGAG
2172	CATGTTATTGACGCCAGGTTTGGACGTTGTTTTTCACTGTATCCGGATG TGAAGTCGTTG

2173	TGGTTTTAGAAAAGGATGGTGTGCTTGTCGCTGAGACACATGTTATTGA CGCCAGGTTTG
2175	TCTAGTTCAGAGCTTGGTGC
2226	TGCTCCATTTGGAAGTCTCG
2227	TATCTACGAAGTGACCTGCG

[00371] The *LEU1* gene was deleted by transforming cells with a *leu1::KI_URA3* deletion cassette that was generated by two rounds of PCR. Initially, the *KI_URA3* gene was amplified with primers 2171 and 2172 from pGV1299 (described in Example 1). These primers add 40bp of the *LEU1* promoter and terminator sequences to the 5' and 3' ends of the *KI_URA3* gene. This PCR product was then used as a template for a PCR using primers 2170 and 2173. Primer 2170 adds an additional 36 bp of the *LEU1* promoter sequence at the 5' end and primer 2173 adds an additional 38 bp of the *LEU1* terminator sequence at the 3' end. This PCR product was transformed into GEVO2244 (described in Example 1) to generate GEVO2570. The 5' junction of the integrations were confirmed by colony PCR using primers 2226 and 587. The 3' junction of the integrations were confirmed by colony PCR using primers 588 and 2175. The loss of the *LEU1* gene was confirmed by a lack of PCR product using primers 2167 and 2227.

[00372] GEVO2570 has a deletion in ILV3. GEVO2570 is used to measure DHAD activity in the presence of *L. lactis ilvD* overexpressed as described in Examples 2 and 4. A plasmid (pGV2001) with no DHAD is used as a negative control.

<u>Example 9: Conserved motif amongst cytosolically active DHAD enzymes</u>

[00373] This example illustrates that a DHAD enzymes with a specific amino acid sequence motif are more likely to be functional when expressed in the yeast cytosol.

[00374] Based on the data from biochemical assays (see Example 2), several DHAD homologs were identified that exhibit at least some cytosolic activity. A total of ten different homologs were tested using biochemical assays. The DHADs were expressed from 2 micron yeast vectors and transformed into GEVO2244. The homologs were then ranked based on their measured specific activity in both whole cell lysates and in cytosolic fractions. .

[00375] Based on these data, four DHAD homologs (*Lactococcus lactis* (SEQ ID NO. 9), Grammella (SEQ ID NO.8), Acidobacteria (SEQ ID NO.7), and Saccharopolyspora (SEQ ID NO. 10)) exhibit cytosolic activity. Four homologs exhibit no cytosolic activity (Ralstonia (SEQ ID NO. 13), Chromohalobacter (SEQ ID NO. 14), Picrophilus (SEQ ID NO. 15), and Sulfolobus (SEQ ID NO. 16)). One motif-containing homolog was inconclusive (Piromyces, SEQ ID NO. 12) - it did not complement the Gevo2242 valine auxotrophy and had detectable biochemical activity. Since, this homolog has a putative organellar targeting sequence, the protein is likely to be mitochondrially located explaining its inability to complement the Gevo2242 auxotrophy despite containing the motif.

[00376] A multiple sequence alignment (MSA) was created using the Align Multiple Sequences tool of Clone Manger 9 Professional Addition Software using the "Multi-Way" function. This function will do exhaustive pairwise global alignments of all

sequences and progressive assembly of alignments using Neighbor-Joining phylogeny. A total of 53 representative DHAD homologs (Figure 9-1) were aligned using the following using the BLOSUM62 scoring matrix setting. This alignment generated the tree in Figure 9-1.

[00377] Many of the DHAD homologs exhibiting cytosolic activity (highlighted in yellow) are related by overall homology (>40%) homology when compared to the *S. cerevisiae* DHAD encoded by *ScILV3*. However, the 40% homology cut-off still includes several DHAD homologs that do not exhibit cytosolic activity (highlighted in blue). The gray highlighted DHAD failed to complement in the genetic/biochemistry assay but this result is still consistent with our motif hypothesis since the protein still retained its mitochondrial localization signal. Therefore, a common sequence motif, unique to DHAD homologs that are cytosolically active, was identified: **P(I/L)XXXGX(I/L)XIL** (SEQ ID NO. 19), where (I/L) indicates an isoleucine or leucine at that position, and X indicates any amino acid. This motif can be found in all five DHAD homologs that are cytosolically active, as well as the inconclusive *Piromyces* homolog.

[00378] An even more specific version of this motif was identified that is conserved in all of the DHAD homologs that are cytosolically active except for the Saccharopolyspora DHAD: **PIKXXGX(I/L)XIL** (SEQ ID NO. 20). This motif is conserved amongst the majority if not all eukaryotic homologs of DHAD.

[00379] Six additional DHAD homologs were identified: SEQ ID NOs 1-6 as specified in Table 2. These DHAD homologs (SEQ ID NOs 1-6) contain the motifs PYHKEGGLGIL, PYSEKGGLAIL, PYKPEGGIAIL, PLKPSGHLQIL, PIKKTGHLQIL, and PIKETGHIQIL, respectively.
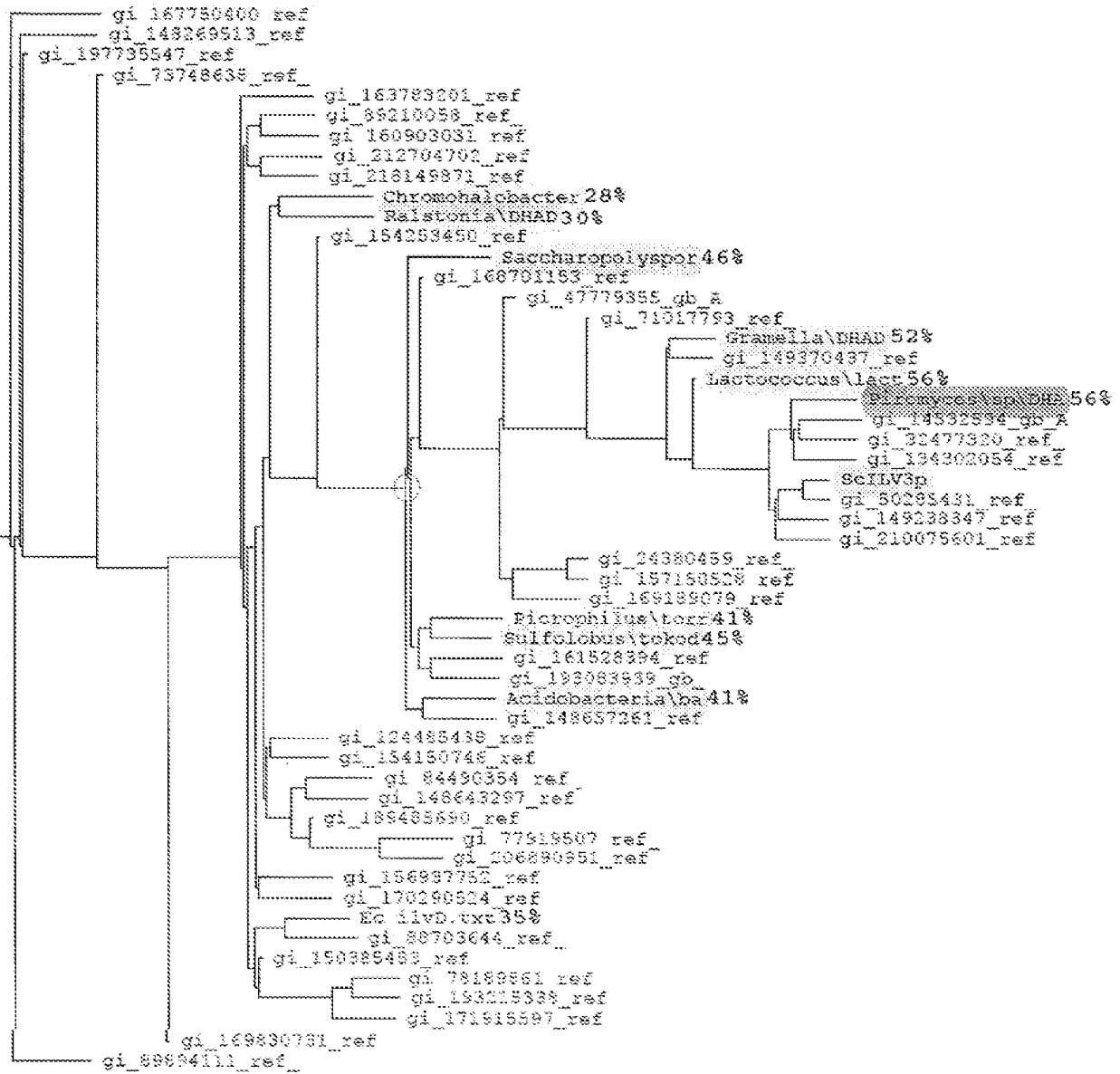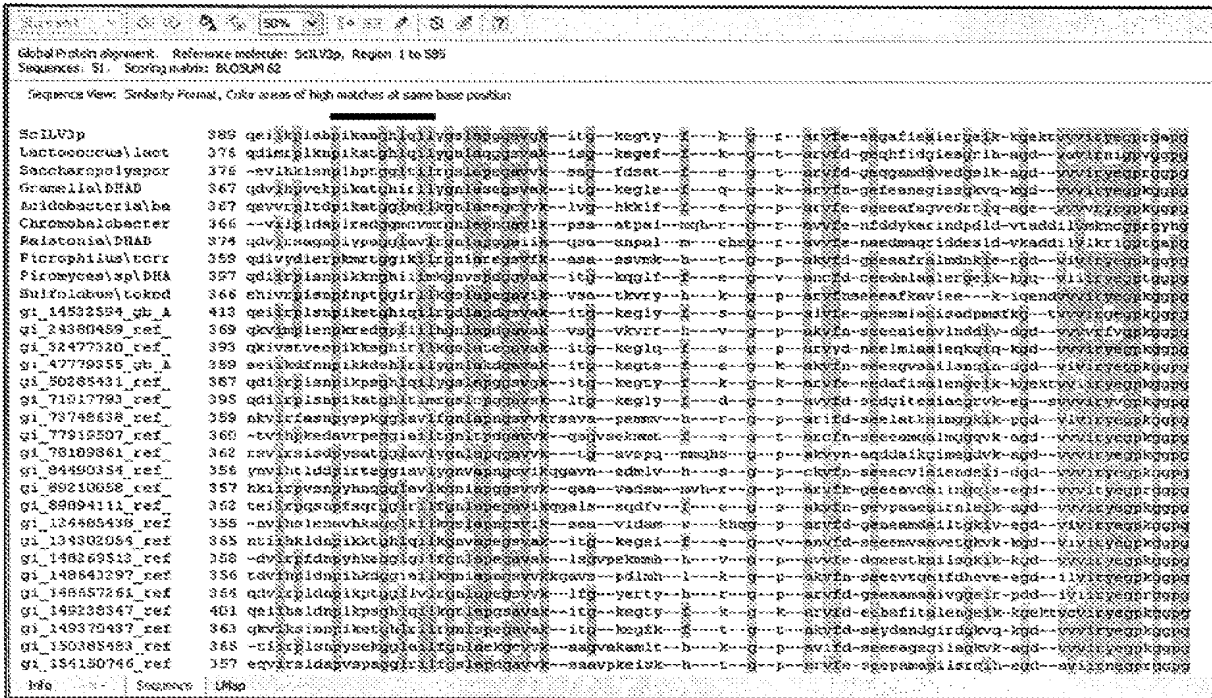
Figure 9-1.

Figure 9-2.

Example 10: N-terminal deletion to generate cytosolically active DHAD enzymes

[00380] To generate a cytosolic DHAD, either 19 or 23 amino acids were removed from the N terminus of native *S. cerevisiae* ILV3 to disrupt the MTS. These Ilv3s were assayed for in vitro DHAD activity.

[00381] To construct the N-terminal Ilv3 deletions, pGV1817 was used as the template for PCR amplification of *S. cerevisiae* ILV3 with the appropriate forward primer (2390 or 2394) and 2468 as the reverse primer. The forward primers added 40 nucleotides matching the 3' end of the *TDH3* promoter to the 5' end of *ILV3*, and the reverse primer added 45 nucleotides matching the 5' end of the *CYC1* terminator to the 3' end of *ILV3*. These regions of homology allow for homologous recombination to occur between the *ILV3* PCR products and pGV2080 (CEN) linearized with *Xho*I and *Not*I. GEVO2244 was transformed as described above using 100 ng of *ILV3* PCR product and 100 ng of linearized or pGV2080.

Construction of strains and plasmids

[00382] Plasmid pGV1817 was constructed to create a vector to be used in the bipartite integration system which contained *S. cerevisiae* ILV3 and ILV5. It was constructed by PCR amplification of *ILV3* from *S. cerevisiae* genomic DNA with primers 1617 and 1618. Primer 1617 added a *Sal*I site to the 5' end of *ILV3*, and primer 1618 added a *BamH*I site to the 3' end. The resulting PCR product and pGV1810 were digested with *Sal*I and *BamH*I and ligated to yield pGV1817.
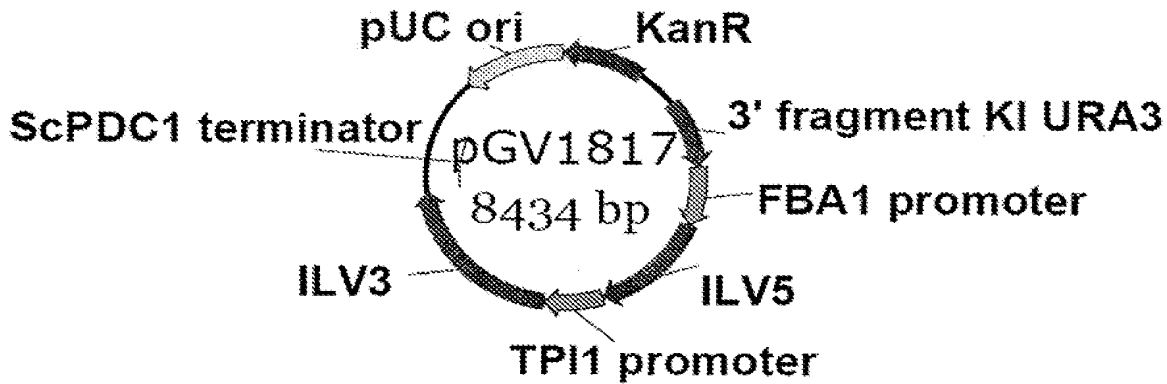
Figure 10-1: Plasmid pGV1817

[00383]   The CEN plasmid pGV2080 is used for expressing *Ec_ilvC_Q110V_coSc* under *TDH3* promoter control. The *Ec_ilvC_Q110V_coSc* gene was PCR amplified from pGV1981 using primer pairs 2123/2124 followed by gel extraction and purification of the PCR product using Zymoclean Gel DNA Recovery Kit. The purified PCR product was then digested with SalI and XbaI and ligated into XhoI/XbaI digested pGV1056 vector (~5.8 kb fragment, containing P_{TDH3}, T_{CYC1}, *HIS3*, *CEN*, *ARS*, *bla*, *pUC-ori*). Purified plasmid DNA was confirmed by the generation of ~5.6 kb and ~1.7 kb fragments by XbaI and MluI restriction enzyme digest and sequenced with primers 350/352/1911 to ensure sequence identity of the insert.
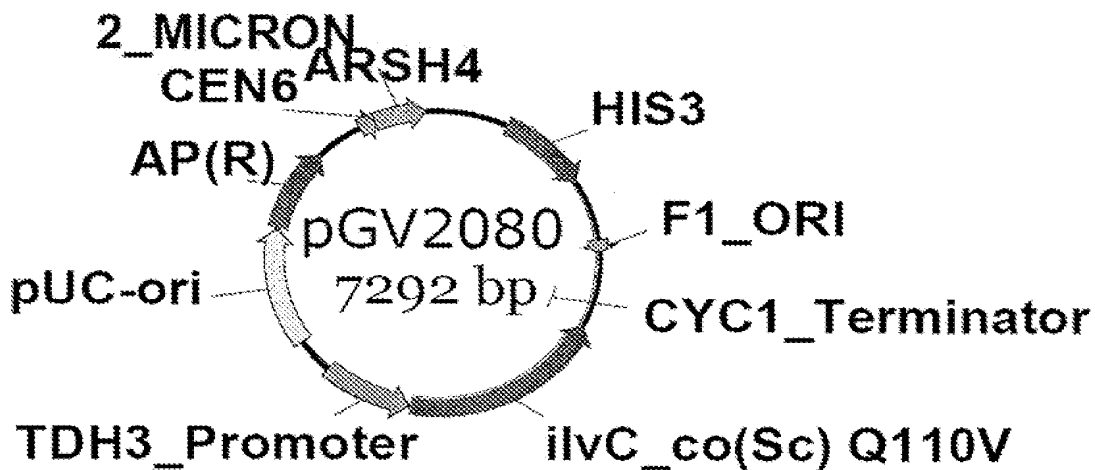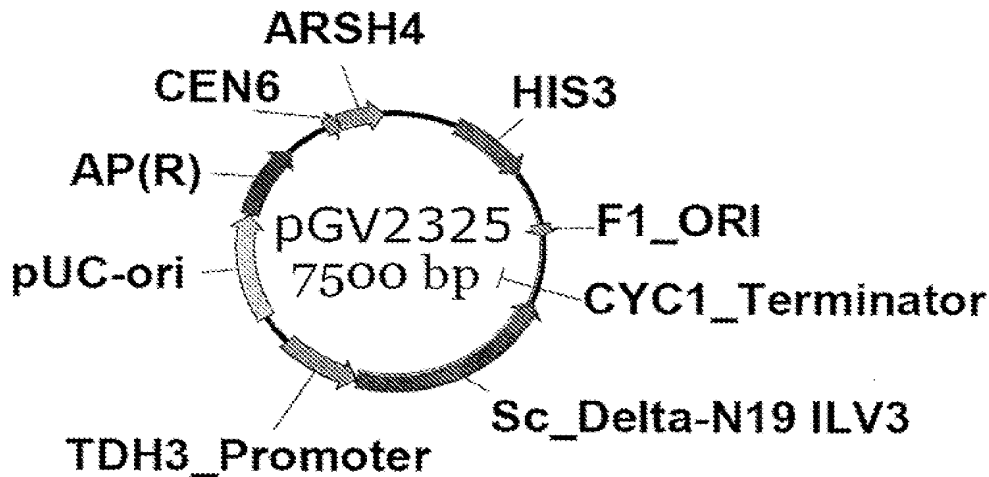

Figure 10-2: Plasmid pGV2080

[00384]   Plasmid pGV2325 was constructed by amplification of *S. cerevisiae* *ILV3ΔN19* from pGV1817 with primers 2390 and 2468. Primer 2390 added 40 nucleotides encoding part of the 3' end of the *TDH3* promoter to the 5' end of *ILV3*, and primer 2468 added 45 nucleotides encoding part of the 5' end of the *CYC1* terminator to the 3' end of *ILV3*. These regions of homology allowed for homologous

recombination to occur between the *ILV3* PCR product and pGV2080 (CEN) linearized with *XhoI* and *NotI*. GEVO2624 was transformed as described above using 100 ng of *ILV3* PCR product and 100 ng of linearized pGV2080. Plasmid was recovered from GEVO2624 as described above, transformed into *E. coli*, and then purified. The plasmid was verified by restriction digest with *SalI* and *NotI* (1711bp and 5789bp). The construct was then sequence verified with primers 592, 1620, 1621, 1622, and 1623.

Figure 10-3: Plasmid pGV2325

[00385] Plasmid pGV2326 was constructed by amplification of *S. cerevisiae* *ILV3ΔN23* from pGV1817 with primers 2394 and 2468. Primer 2394 added 40 nucleotides encoding part of the 3' end of the *TDH3* promoter to the 5' end of *ILV3*, and primer 2468 added 45 nucleotides encoding part of the 5' end of the *CYC1* terminator to the 3' end of *ILV3*. These regions of homology allowed for homologous recombination to occur between the *ILV3* PCR product and pGV2080 (CEN) linearized with *XhoI* and *NotI*. GEVO2624 was transformed as described above using 100 ng of *ILV3* PCR product and 100 ng of linearized pGV2080. Plasmid was recovered from GEVO2624 was described above, transformed into *E. coli*, and then purified. The plasmid was verified by restriction digest with *SalI* and *NotI* (1699bp and 5789bp). The construct was then sequence verified with primers 592, 1620, 1621, 1622, and 1623.
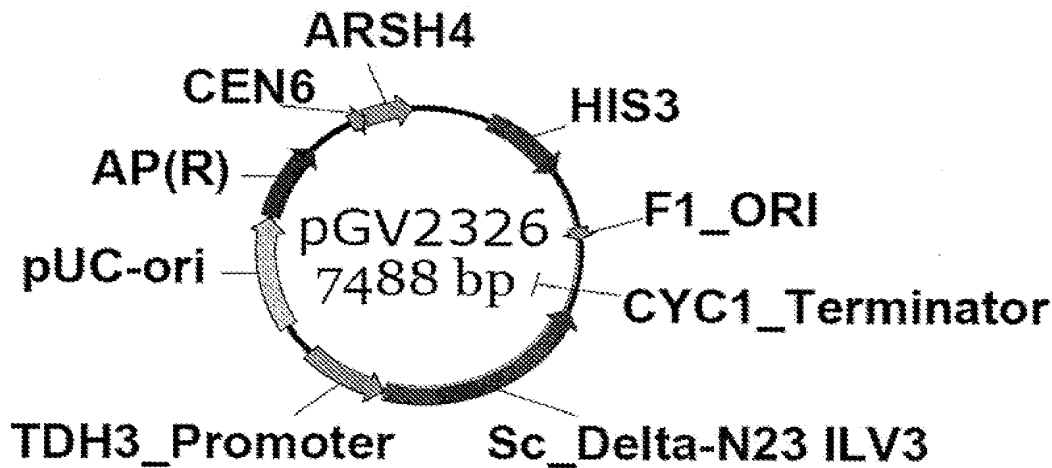
Figure 10-4: Plasmid pGV2326

Table 10-1 details the genotype of strains disclosed herein:

GEVO No.	Genotype / Source
GEVO2244	CEN.PK2, MAT-alpha ura3 leu2 his3 trp1, ilv3
GEVO2624	CEN.PK2, MAT-alpha ura3 leu2 his3 trp1, ilv5-D255E, pdc1::(P_{ScCUP1}:Bs_alsS:P_{ScTDH3}:Ec_ilvC co Q110V)

Table 10-2 details the plasmids disclosed herein:

GEVO No.	Figure	Genotype
pGV1817	10-1	
pGV2080	10-2	P_{Sc_TDH3}:Ec_ilvC_Q110V_co:T_{Sc_CYC1}, HIS3, CEN, ARS, bla, pUC-ori
pGV2325	10-3	P_{Sc_TDH3}: Sc_ΔN19 ILV3:T_{Sc_CYC1}, HIS3, CEN, ARS, bla, pUC-ori
pGV2326	10-4	P_{Sc_TDH3}: Sc_ΔN23 ILV3:T_{Sc_CYC1}, HIS3, CEN, ARS, bla, pUC-ori

Table 10-3 details the primer sequences disclosed herein:

#	Sequence
2390	tttttttttagttttaaaacaccagaacttagtttcgacgg**gtcgacatg**gcaaagaagctcaacaagtactcgt
2394	tttttttttagttttaaaacaccagaacttagtttcgacgg**gtcgacatg**aacaagtactcgtatatcatcactg
2468	tgggggggagggcgtgaatgtaagcgtgacataactaattacatga**gcggccgc**tcaagcatctaaaacacaaccg

[00386] The DHAD enzymes *S. cerevisiae* ILV3ΔN19 and ILV3ΔN23 (encoded by SEQ ID NO: 127 and 128, respectively) were expressed in strain GEVO2244 from plasmids pGV2325 and pGV2326, respectively. The strains were cultured in SCD-His+9xIV medium and harvested at an OD of 3-4. Frozen pellets were lysed by bead beating andthe enzymatic activity was assessed as follows. In a 0.2 mL PCR tube,

20 µL of each sample was mixed with 20 µL of 0.1 M 2,3-dihydroxyisovalerate (DHIV), 10 µL of 0.1 M MgSO₄, and 150 µL of 0.05 M Tris-HCl pH 8.0, and the mixture was incubated for 30 min at 35°C. Each reaction was carried out in triplicate. Each tube was then heated to 95°C for 5 min to inactivate any enzymatic activity, and the solution was centrifuged (3000xg for 5min) to pellet insoluble debris. To prepare samples for analysis, 100 µL of each reaction were mixed with 100 µL of a solution consisting of 4 parts 15 mM dinitrophenyl hydrazine (DNPH) in acetonitrile with 1 part 50 mM citric acid, pH 3.0, and the mixture was heated to 70°C for 30 min in a thermocycler. Analysis of DNPH derivatized ketoisovalerate via HPLC was carried out as described in General Methods. Results are shown in Table 11-4. One unit of DHAD activity is defined as the activity converting 1 µmol of DHIV to KIV in 1 minute.

Table 11-4. Specific enzyme activity in U/mg lysate for Ilv3ΔN19 and Ilv3ΔN23

Enzyme	Activity [U/mg]
Ilv3ΔN19	0.035
Ilv3ΔN23	0.009
Empty vector control	0.000

Example 11: Overexpression of Fe-S assembly machinery

[00387] To ascertain the effects of overexpressing a cytosolic 2Fe-2S or 4Fe-4S cluster-containing DHAD with candidate assembly machinery, the following steps, or equivalent steps can be carried out. First, the coding sequence for the open reading of the DHAD from spinach or other 2Fe-2S or 4Fe-4S cluster-containing DHAD is cloned into the high-copy (2micron origin) S.cerevisiae expression vector pGV2074, such that expression of the coding sequence is directed by the PGK1 promoter sequence, yielding plasmid pGV2074-1. Next, the NifU and NifS genes from Entamoeba histolytica or the homologous NIF genes from Lactococcus lactis are successively introduced into the aforementioned vector, eventually yielding a single plasmid (pGV2074-2) where the expression of all 3 genes is directed by strong constitutive S.cerevisiae promoter sequences. Plasmids pGV2074-1 and pGV2074-2 are transformed into S. cerevisiae strain GEVO2244 (relevant genotype, ilv3Δ) and transformants selected by resistance to Hygromycin B (0.1 g/L). At least 3 individual colonies arising from each transformation are cultured, a cell lysate produced, and the DHAD activity present therein measured, all according to previously-described methods.

[00388] The foregoing detailed description has been given for clearness of understanding only and no unnecessary limitations should be understood there from as modifications will be obvious to those skilled in the art.

[00389] While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice

within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

[00390]   The disclosures, including the claims, figures and/or drawings, of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

1. A recombinant microorganism for producing isobutanol, said recombinant microorganism comprising:

an isobutanol producing metabolic pathway comprising at least one exogenous gene, wherein said recombinant microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 5 percent theoretical and/or a specific productivity of at least about 0.7 mg/L/hr per OD.

2. The recombinant microorganism of claim 1, wherein said recombinant microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 10 percent theoretical and/or a specific productivity of at least about 1 mg/L/hr per OD.

3. The recombinant microorganism of claim 1, wherein said recombinant microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 20 percent theoretical and/or a specific productivity of at least about 10 mg/L/hr per OD.

4. The recombinant microorganism of claim 1, wherein said recombinant microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 50 percent theoretical and/or a specific productivity of at least about 50 mg/L/hr per OD.

5. The recombinant microorganism of claim 1, wherein said recombinant microorganism is selected to produce isobutanol from a carbon source at a yield of at least about 80 percent theoretical and/or a specific productivity of at least about 100 mg/L/hr per OD.

6. The recombinant microorganism of any of claims 1-5, wherein said recombinant microorganism comprises a cytosolically active dihydroxyacid dehydratase (DHAD) enzyme.

7. The recombinant microorganism of any of claims 1-5, wherein said recombinant microorganism comprises a modified or alternative dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the parental or native DHAD enzyme.

8. The recombinant microorganism of any of claims 1-5, wherein said recombinant microorganism comprises a modified dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the DHAD enzyme comprised of the amino acid sequence of SEQ ID NO: 11.

9. The recombinant microorganism of any of claims 1-8, wherein said recombinant microorganism comprises a dihydroxacid dehydratase (DHAD) enzyme that is fused

103182 v1/DC

to a peptide tag, whereby said dihydroxacid dehydratase (DHAD) enzyme exhibits increased cytosolic DHAD localization and/or cytosolic DHAD activity as compared to the parental enzyme.

10. The recombinant microorganism of claim 9, wherein said peptide tag is non-cleavable.

11. The recombinant microorganism of any of claims 9-10, wherein said peptide tag is fused at the N-terminus of said dihydroxacid dehydratase (DHAD) enzyme.

12. The recombinant microorganism of any of claims 9-11, wherein said peptide tag is selected from the group consisting of ubiquitin, ubiquitin-like (UBL) proteins, myc, HA-tag, green fluorescent protein (GFP), and the maltose binding protein (MBP).

13. The recombinant microorganism of any of claims 1-12, wherein said recombinant microorganism comprises a modified dihydroxyacid dehydratase (DHAD) enzyme having one or more amino acid deletions at the N-terminus.

14. The recombinant microorganism of claim 13, wherein DHAD enzyme has at least about 10 amino acid deletions at the N-terminus.

15. The recombinant microorganism of claim 14, wherein said DHAD enzyme has at least about 17 amino acid deletions at the N-terminus.

16. The recombinant microorganism of claim 15, wherein said DHAD enzyme has 19 amino acid deletions at the N-terminus.

17. The recombinant microorganism of claim 15, wherein said DHAD enzyme has 23 amino acid deletions before the N-terminus.

18. The recombinant microorganism of any of claims 6-17, wherein said DHAD enzyme comprises the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid.

19. The recombinant microorganism of any of claims 6-18, wherein said DHAD enzyme is derived from a bacterial organism.

20. The recombinant microorganism of claim 19, wherein said bacterial organism is *L. lactis* or *E. coli*.

21. The recombinant microorganism of claim 20, wherein said DHAD enzyme is from *L. lactis* and comprises the amino acid sequence of SEQ ID NO: 9.

22. The recombinant microorganism of claim 20, wherein said DHAD enzyme is from *E. coli* and comprises the amino acid sequence of SEQ ID NO: 129.

23. The recombinant microorganism of any of claims 6-18, wherein said DHAD enzyme is derived from a eukaryotic organism.

24. The recombinant microorganism of claim 23, wherein said eukaryotic organism is *Piromyces* or *Saccharomyces.*

25. The recombinant microorganism of any of claims 6-17, wherein said DHAD enzyme comprises the amino acid sequence CPGXGXC (SEQ ID NO: 123), wherein X is any amino acid.

26. The recombinant microorganism of any of claims 6-17, wherein said DHAD enzyme comprises the amino acid sequence CPGXG(A/S)C (SEQ ID NO: 124), wherein X is any amino acid.

27. The recombinant microorganism of any of claims 6-17, wherein said DHAD enzyme comprises the amino acid sequence CXXXPGXGXC (SEQ ID NO: 125), wherein X is any amino acid.

28. The recombinant microorganism of any of claims 6-27, wherein said DHAD enzyme exhibits a properly folded iron-sulfur cluster domain and/or redox active domain in the cytosol, thereby rendering the enzyme cytosolically active.

29. The recombinant microorganism of claim 28, wherein said DHAD enzyme comprises a mutated or modified iron-sulfur cluster domain.

30. The recombinant microorganism of any of claims 1-29, wherein said recombinant microorganism further comprises a nucleic acid encoding a chaperone protein, wherein said chaperone protein assists the folding of a protein exhibiting cytosolic activity.

31. The recombinant microorganism of claim 30, wherein said chaperone protein is a native protein.

32. The recombinant microorganism of claim 30, wherein said chaperone protein is an exogenous protein.

33. The recombinant microorganism of any of claims 30-32, wherein said chaperone protein is selected from the group consisting of: endoplasmic reticulum oxidoreductin 1 (Ero1), including variants of Ero1 that have been suitably altered to reduce or prevent its localization to the endoplasmic reticulum; thioredoxin (Trx1 and Trx2), thioredoxin reductase (Trr1), glutaredoxin (Grx1, Grx2, Grx3, Grx4, Grx5, Grx6, Grx7, and Grx8,); glutathione reductase (Glr1), and Jac1, including variants of Jac1 that have been suitably altered to reduce or prevent its mitochondrial localization; Hsp10, Hsp60, GroEL, and GroES and homologs or variants thereof.

34. The recombinant microorganism of any of claims 1-33, wherein said recombinant microorganism further comprises one or more genes encoding an iron-sulfur cluster assembly protein.

35. The recombinant microorganism of claim 34, wherein said genes encoding an iron-sulfur cluster assembly protein are selected from the group consisting of *cyaY*, *iscS*, *iscU*, *iscA*, *hscB*, *hscA*, *fdx*, *isuX*, *sufA*, *sufB*, *sufC*, *sufD*, *sufS*, *sufE*, *apbC*, and homologs or variants thereof.

36. The recombinant microorganism of claim 35, wherein said genes encoding an iron-sulfur cluster assembly are derived from an organism selected from the group consisting of *Escherichia coli* and *Lactococcus lactis*.

37. The recombinant microorganism of claim 34, wherein said genes encoding an iron-sulfur cluster assembly are selected from the group consisting of *nifS*, *nifU*, and homologs or variants thereof.

38. The recombinant microorganism of claim 37, wherein said genes encoding an iron-sulfur cluster assembly are derived from an organism selected from the group consisting of *Helicobacter pylori* and *Entamoeba histolytica*.

39. The recombinant microorganism of claim 34, wherein said genes encoding an iron-sulfur cluster assembly are selected from the group consisting of CFD1, NBP35, NAR1, CIA1, and homologs or variants thereof.

40. The recombinant microorganism of claim 39, wherein said genes encoding an iron-sulfur cluster assembly are derived from yeast.

41. The recombinant microorganism of any of claims 34-40, wherein said one or more genes encoding an iron-sulfur cluster assembly is mutated or modified to remove a signal peptide, whereby localization of the product of said one or more genes to the mitochondria is prevented.

42. The recombinant microorganism of any of claims 34-41, wherein the genes encoding an iron-sulfur cluster assembly are overexpressed.

43. The recombinant microorganism of any of claims 1-42, wherein said recombinant microorganism has been engineered to reduce the activity of one or more iron-sulfur cluster containing cytosolic proteins.

44. The recombinant microorganism of any of claims 1-42, wherein said recombinant microorganism has been engineered to eliminate the activity of one or more iron-sulfur cluster containing cytosolic proteins.

45. The recombinant microorganism of any of claims 1-42, wherein said recombinant microorganism has been engineered to reduce to the protein levels of one or more iron-sulfer cluster containing cytosolic proteins.

46. The recombinant microorganism of claim 45, wherein said iron-sulfur cluster containing cytosolic protein is 3-isopropylmalate dehydratase (LEU1).

47. The recombinant microorganism of claim 46, wherein said recombinant microorganism comprises a mutation in the 3-isopropylmalate dehydratase (LEU1) gene resulting in a reduction of LEU1 protein levels.

48. The recombinant microorganism of claim 46, wherein said recombinant microorganism comprises a partial deletion of a 3-isopropylmalate dehydratase (LEU1) gene resulting in a reduction of LEU1 protein levels.

49. The recombinant microorganism of claim 46, wherein said recombinant microorganism comprises a complete deletion of a 3-isopropylmalate dehydratase (LEU1) gene resulting in a reduction of LEU1 protein levels.

50. The recombinant microorganism of claim 46, wherein said recombinant microorganism comprises a modification of the regulatory region associated with a 3-isopropylmalate dehydratase (LEU1) gene resulting in reduction of LEU1 protein levels.

51. The recombinant microorganism of claim 46, wherein said recombinant microorganism comprises a modification of the transcriptional regulator resulting in a reduction of LEU1 protein levels.

52. The recombinant microorganism of any of claims 1-51, wherein said recombinant microorganism exhibits at least about 50 percent greater dihydroxyacid dehydratase (DHAD) activity in the cytosol as compared to the parental microorganism, and wherein said parental microorganism comprises an unmodified or native DHAD enzyme.

53. The recombinant microorganism of any of claims 1-52, wherein said recombinant microorganism further comprises a 3-isopropylmalate dehydratase (LEU1) and/or imidazoleglycerol-phosphate dehydrogenase (HIS3) that converts 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

54. The recombinant microorganism of claim 53, wherein said 3-isopropylmalate dehydratase (LEU1) and/or said imidazoleglycerol-phosphate dehydrogenase (HIS3) is overexpressed in the cytosol.

55. The recombinant microorganism of any of claims 1-54, wherein said recombinant microorganism has further been engineered to overexpress one or more genes

selected from the group consisting of *AFT1*, *AFT2*, *GRX3*, and *GRX4*, or homologs thereof.

56. The recombinant microorganism of any of claims 1-54, wherein said recombinant microorganism has further been enginereered to delete and/or attenuate one or more genes selected from the group consisting of *GRX3* and *GRX4*, or homologs thereof.

57. The recombinant microorganism of any of claims 1-56, wherein said recombinant microorganism has further been engineered to express one or more proteins in the cytosol that reduce the concentration of reactive oxygen species (ROS) in said cytosol.

58. The recombinant microorganism of claim 57, wherein said one or more proteins are selected from the group consisting of catalases, superoxide dismutases, metallothioneins, and methionine sulphoxide reductases.

59. The recombinant microorganism of claim 58, wherein said catalase is encoded by one of more of the genes selected from the group consisting of the *E. coli* genes *katG* and *katE*, the *S. cerevisiae* genes *CTT1* and *CTA1*, or homologs thereof.

60. The recombinant microorganism of claim 58, wherein said superoxide dismutase is encoded by one of more of the genes selected from the group consisting of the *E. coli* genes *sodA*, *sodB*, *sodC*, the *S. cerevisiae* genes *SOD1* and *SOD2*, or homologs thereof.

61. The recombinant microorganism of claim 58, wherein said metallothionein is encoded by one of more of the genes selected from the group consisting of the *S. cerevisiae CUP1-1* and *CUP1-2* genes or homologs thereof.

62. The recombinant microorganism of claim 58, wherein said metallothionein is encoded by one or more genes selected from the group consisting of the Mycobacterium tuberculosis *MymT* gene and the Synechococcus PCC 7942 *SmtA* gene or homologs thereof.

63. The recombinant microorganism of claim 58, wherein said methionine sulphoxide reductase is encoded by one or more genes selected from the group consisting of the *S. cerevisiae* genes *MXR1* and *MXR2*, or homologs thereof.

64. The recombinant microorganism of any of claims 1-63, wherein said recombinant microorganism has further been engineered to express one or more enzymes that increase the level of available glutathione in the cytosol.

65. The recombinant microorganism of claim 64, wherein said one or more enzymes are selected from the group consisting of glutaredoxin, glutathione reductase, and glutathione synthase.

66. The recombinant microorganism of claim 65, wherein said glutaredoxin is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GRX2, GRX4, GRX6*, and *GRX7*, or homologs thereof.

67. The recombinant microorganism of claim 65, wherein said glutathione reductase is encoded by the *S. cerevisiae* genes *GLR1* or homologs thereof.

68. The recombinant microorganism of claim 65, wherein said glutathione synthase is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GSH1* and *GSH2*, or homologs thereof.

69. The recombinant microorganism of claim 64, wherein two enzymes are expressed in and targeted to the cytosol of yeast to increase the level of available glutathione in the cytosol.

70. The recombinant microorganism of claim 69, wherein said enzymes are γ-glutamyl cysteine synthase and glutathione synthase.

71. The recombinant microorganism of claim 70, wherein said glutathione synthase is encoded by one of more of the genes selected from the group the *S. cerevisiae* genes *GSH1* and *GSH2*, or homologs thereof.

72. The recombinant microorganism of any of claims 1-71, wherein said recombinant microorganism has further been engineered to overexpress one or more cytosolic functional components of the thioredoxin system.

73. The recombinant microorganism of claim 72, wherein said one or more cytosolic functional components of the thioredoxin system are selected from a thioredoxin and a thioredoxin reductase.

74. The recombinant microorganism of claim 73, wherein said thioredoxin is encoded by the *S. cerevisiae TRX1* and *TRX2* genes or homologs thereof.

75. The recombinant microorganism of claim 73 or 74, wherein said thioredoxin reductase is encoded by S. cerevisiae *TRR1* gene or homologs thereof.

76. The recombinant microorganism of any of claims 73-75, wherein said recombinant microorganism further overexpresses the mitochondrial thioredoxin system.

77. The recombinant microorganism of claim 76, wherein said mitochondrial thioredoxin system is comprised of the mitochondrial thioredoxin and mitochondrial thioredoxin reductase.

78. The recombinant microorganism of claim 77, wherein said mitochondrial thioredoxin is encoded by the *S. cerevisiae TRX3* gene or homologs thereof.

79. The recombinant microorganism of claim 77 or 78, wherein said mitochondrial thioredoxin reductase is encoded by the *S. cerevisiae TRR2* gene or homologs thereof.

80. The recombinant microorganism of any of claims 1-79, wherein said recombinant microorganism is further engineered to overexpress one or more mitochondrial export proteins.

81. The recombinant microorganism of claim 80, wherein said mitochondrial export protein is selected from the group consisting of the *S. cerevisiae ATM1*, the *S. cerevisiae ERV1*, and the *S. cerevisiae BAT1*, or homologs thereof.

82. The recombinant microorganism of any of claims 1-81, wherein said recombinant microorganism is further engineered to increase inner mitochondrial membrane electrical potential, $\Delta\Psi_M$.

83. The recombinant microorganism of claim 82, wherein said recombinant microorganism comprises a mutation in the mitochondrial ATP synthase complex that increases ATP hydrolysis activity.

84. The recombinant microorganism of claim 83, wherein said mutation is an ATP1-111 suppressor mutation or a corresponding mutation in a homologous protein.

85. The recombinant microorganism of claim 82, wherein said recombinant microorganism has been engineered to overexpress an ATP/ADP carrier protein, wherein said overexpression increases ATP^{4-} import into the mitochondrial matrix in exchange for ADP^{3-}.

86. The recombinant microorganism of claim 85, wherein said ATP/ADP carrier protein is encoded by the *S. cerevisiae AAC1*, *AAC2*, and/or *AAC3* genes or homologs thereof.

87. The recombinant microorganism of any of claims 1-86, wherein said recombinant microorganism has further been engineered to express one or more enzymes in the cytosol that reduce the concentration of reactive nitrogen species (RNS) and/or nitric oxide (NO) in said cytosol.

88. The recombinant microorganism of claim 87, wherein said one or more enzymes are selected from the group consisting of nitric oxide reductases and glutathione-S-nitrosothiol reductase.

89. The recombinant microorganism of claim 88, wherein said nitric oxide reductase is encoded by one of more of the genes selected from the group consisting of the *E. coli* gene *norV* and the *Fusarium oxysporum* gene *P-450dNIR*, or homologs thereof.

90. The recombinant microorganism of claim 88, wherein said glutathione-S-nitrosothiol reductase is encoded by the *S. cerevisiae* gene *SFA1* or homologs thereof.

91. The recombinant microorganism of claim 90, wherein said glutathione-S-nitrosothiol reductase gene *SFA1* is overexpressed.

92. The recombinant microorganism of claim 87, wherein said one or more enzymes is encoded by a gene selected from the group consisting of the *E. coli* gene *ytfE*, the *Staphylococcus aureus* gene *scdA*, and *Neisseria gonorrhoeae* gene *dnrN*, or homologs thereof.

93. The recombinant microorganism of any of claims 1-92, wherein said recombinant microorganism has further been engineered to overexpress one or more of the genes selected from the *S. cerevisiae* genes *MET1*, *MET2*, *MET3*, *MET5*, *MET8*, *MET10*, *MET14*, *MET16*, *MET17*, *HOM2*, *HOM3*, *HOM6*, *CYS3*, *CYS4*, *SUL1*, and *SUL2*, or homologs thereof.

94. The recombinant microorganism of any of claims 1-93, wherein said recombinant microorganism has further been engineered to overexpress one or more of the genes selected from the *S. cerevisiae* genes *YCT1*, *MUP1*, *GAP1*, *AGP1*, *GNP1*, *BAP1*, *BAP2*, *TAT1*, and *TAT2*.

95. The recombinant microorganism of any of claims 1-94, wherein said recombinant microorganism further comprises a pathway for the fermentation of isobutanol from a pentose sugar.

96. The recombinant microorganism of claim 95, wherein said pentose sugar is xylose.

97. The recombinant microorganism of any of claims 1-96, wherein said recombinant microorganism is engineered to express a functional xylose isomerase (XI).

98. The recombinant microorganism of claim 97, wherein said recombinant microorganism further comprises a deletion or disruption of a native gene encoding for an enzyme that catalyzes the conversion of xylose to xylitol.

99. The recombinant microorganism of claim 98, wherein said native gene encodes for a xylose reductase (XR).

100. The recombinant microorganism of any of claims 97-99, wherein said recombinant microorganism further comprises a deletion or disruption of a native gene encoding for an enzyme that catalyzes the conversion of xylitol to xylulose.

101. The recombinant microorganism of claim 100, wherein said native gene encodes a xylitol dehydrogenase (XDH).

102. The recombinant microorganism of any of claims 96-101, wherein said recombinant microorganism further comprises the overexpression of a heterologous or native gene encoding for an enzyme that catalyzes the conversion of xylulose to xylulose-5-phosphate.

103. The recombinant microorganism of claim 192, wherein said native gene encodes a xylulose kinase (XK).

104. The recombinant microorganism of any of claims 1-103, wherein said recombinant microorganism is a yeast microorganism of the Saccharomyces clade.

105. The recombinant microorganism of any of claims 1-104, wherein said recombinant microorganism is a *Saccharomyces sensu stricto* microorganism.

106. The recombinant microorganism of claim 105, wherein said *Saccharomyces sensu stricto* microorganism is selected from the group consisting of *S. cerevisiae, S. kudriavzevii, S. mikatae, S. bayanus, S. uvarum, S. carocanis* and hybrids thereof.

107. The recombinant microorganism of any of claims 1-104, wherein said recombinant microorganism is a Crabtree-negative yeast microorganism.

108. The recombinant microorganism of claim 107, wherein said Crabtree-negative yeast microorganism is classified into a genera selected from a group consisting of *Kluyveromyces, Pichia, Hansenula*, and *Candida*.

109. The recombinant microorganism of claim 108, wherein said Crabtree-negative yeast microorganism is selected from the group consisting of *Kluyveromyces lactis, Kluyveromyces marxianus, Pichia anomala, Pichia stipitis, Hansenula anomala, Candida utilis* and *Kluyveromyces waltii*.

110. The recombinant microorganism of any of claims 1-104, wherein said recombinant microorganism is a Crabtree-positive yeast microorganism.

111. The recombinant microorganism of claim 110, wherein said Crabtree-positive yeast microorganism is classified into a genera selected from a group consisting of

Saccharomyces, Kluyveromyces, Zygosaccharomyces, Debaryomyces, Pichia, Candida, and *Schizosaccharomyces.*

112. The recombinant microorganism of claim 111, wherein said Crabtree-positive yeast microorganism is selected from the group consisting of *Saccharomyces cerevisiae, Saccharomyces uvarum, Saccharomyces bayanus, Saccharomyces paradoxus, Saccharomyces castelli, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Candida glabrata, Z. bailli, Z. rouxii, Debaryomyces hansenii, Pichia pastorius, Schizosaccharomyces pombe,* and *Saccharomyces uvarum.*

113. The recombinant microorganism of any of claims 1-104, wherein said recombinant microorganism is a post-WGD (whole genome duplication) yeast microorganism.

114. The recombinant microorganism of claim 113, wherein said post-WGD yeast microorganism is classified into a genera selected from a group consisting of *Saccharomyces* or *Candida.*

115. The recombinant microorganism of claim 114, wherein said post-WGD yeast microorganism is selected from the group consisting of *Saccharomyces cerevisiae, Saccharomyces uvarum, Saccharomyces bayanus, Saccharomyces paradoxus, Saccharomyces castelli,* and *Candida glabrata.*

116. The recombinant microorganism of any of claims 1-104, wherein said recombinant microorganism is a pre-WGD (whole genome duplication) yeast microorganism.

117. The recombinant microorganism of claim 116, wherein said pre-WGD yeast microorganism is classified into a genera selected from a group consisting of *Saccharomyces, Kluyveromyces, Candida, Pichia, Debaryomyces, Hansenula, Pachysolen, Yarrowia* and *Schizosaccharomyces.*

118. The recombinant microorganism of claim 117, wherein said pre-WGD yeast microorganism is selected from the group consisting of *Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces marxianus, Kluyveromyces waltii, Kluyveromyces lactis, Candida tropicalis, Pichia pastoris, Pichia anomala, Pichia stipitis, Debaryomyces hansenii, H. anomala, Pachysolen tannophilis, Yarrowia lipolytica,* and *Schizosaccharomyces pombe.*

119. A method of producing isobutanol, comprising:

    providing a recombinant microorganism comprising an isobutanol producing metabolic pathway according to any one of claims 1-118, and

    cultivating said recombinant microorganism in a culture medium containing a feedstock providing the carbon source, until a recoverable quantity of the isobutanol is produced.

120. The method of claim 119, further comprising the step of recovering the isobutanol.

121. A recombinant microorganism comprising a cytosolically active dihydroxyacid dehydratase (DHAD) enzyme.

122. A recombinant microorganism comprising a modified or mutated dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the parental DHAD enzyme.

123. A recombinant microorganism comprising a modified or mutated dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the DHAD enzyme encoded by the amino acid sequence of SEQ ID NO: 11.

124. The recombinant microorganism of any of claims 121-123, wherein said DHAD enzyme is fused to a peptide tag, whereby the DHAD enzyme is localized to the cytosol.

125. The recombinant microorganism of any of claims 121-124, wherein said DHAD enzyme comprises the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid.

126. The recombinant microorganism of any of claims 121-125, wherein said DHAD enzyme is derived from a bacterial organism.

127. The recombinant microorganism of claim 126, wherein said bacterial organism is *L. lactis* or *E. coli*.

128. The recombinant microorganism of claim 127, wherein said DHAD enzyme is from *L. lactis* and comprises the amino acid sequence of SEQ ID NO: 9.

129. The recombinant microorganism of claim 127, wherein said DHAD enzyme is from *E. coli* and comprises the amino acid sequence of SEQ ID NO: 129.

130. The recombinant microorganism of any of claims 121-125, wherein said DHAD enzyme is derived from a eukaryotic organism.

131. The recombinant microorganism of claim 130, wherein said eukaryotic organism is *Piromyces or S. cerevisiae*

132. The recombinant microorganism of claim 131, wherein said DHAD enzyme is from *Piromyces* and comprises the amino acid sequence of SEQ ID NO: 12.

133. The recombinant microorganism of any of claims 121-124, wherein said DHAD enzyme comprises the amino acid sequence CPGXGXC (SEQ ID NO: 123), wherein X is any amino acid.

134. The recombinant microorganism of any of claims 121-124, wherein said DHAD enzyme comprises the amino acid sequence CPGXG(A/S)C (SEQ ID NO: 124), wherein X is any amino acid.

135. The recombinant microorganism of any of claims 121-124, wherein said DHAD enzyme comprises the amino acid sequence CXXXPGXGXC (SEQ ID NO: 125), wherein X is any amino acid.

136. The recombinant microorganism of any of claims 121-135, wherein said DHAD enzyme exhibits a properly folded iron-sulfur cluster domain and/or redox active domain in the cytosol.

137. The recombinant microorganism of claim 136, wherein said DHAD enzyme comprises a mutated or modified iron-sulfur cluster domain and/or redox active domain.

138. The recombinant microorganism of any of claims 121-137, wherein said recombinant microorganism further comprises a nucleic acid encoding a chaperone protein, wherein said chaperone protein assists the folding of a protein exhibiting cytosolic activity.

139. The recombinant microorganism of any of claims 121-138, wherein said recombinant microorganism further comprises one or more genes encoding an iron-sulfur cluster assembly protein.

140. The recombinant microorganism of claim 139, wherein said one or more genes encoding an iron-sulfur cluster assembly protein is mutated or modified to remove a signal peptide, whereby localization of the product of said one or more genes to the mitochondria is prevented.

141. The recombinant microorganism of any of claims 139-140, wherein said one or more genes encoding an iron-sulfur cluster assembly protein is overexpressed.

142. The recombinant microorganism of any of claims 121-141, wherein said recombinant microorganism has been engineered to reduce the activity of one or more iron-sulfur cluster containing cytosolic proteins.

143. The recombinant microorganism of any of claims 121-141, wherein said recombinant microorganism has been engineered to eliminate the activity of one or more iron-sulfur cluster containing cytosolic proteins.

144. The recombinant microorganism of any of claims 121-141, wherein said recombinant microorganism has been engineered to reduce to the protein levels of one or more iron-sulfur cluster containing cytosolic proteins.

145. The recombinant microorganism of claim 144, wherein said iron-sulfur cluster containing cytosolic protein is 3-isopropylmalate dehydratase (LEU1).

146. The recombinant microorganism of any of claims 121-145, wherein said recombinant microorganism exhibits at least about 50% greater dihydroxyacid dehydratase (DHAD) activity in the cytosol as compared to the parental microorganism.

147. The recombinant microorganism of any of claims 121-146, further comprising an isobutanol producing metabolic pathway comprising at least one exogenous gene.

148. A method of producing isobutanol, comprising:
    providing a recombinant microorganism comprising an isobutanol producing metabolic pathway according to claims 118-147, and
    cultivating said recombinant microorganism in a culture medium containing a feedstock providing the carbon source, until a recoverable quantity of the isobutanol is produced.

149. The method of claim 148, further comprising the step of recovering the isobutanol.

150. A cytosolically active dihydroxyacid dehydratase (DHAD) enzyme.

151. A modified or mutated dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the parental DHAD enzyme.

152. A modified or mutated dihydroxyacid dehydratase (DHAD) enzyme, wherein said DHAD enzyme exhibits increased cytosolic activity as compared to the DHAD enzyme encoded by the amino acid sequence of SEQ ID NO: 11.

153. A dihydroxyacid dehydratase (DHAD) enzyme comprising the amino acid sequence P(I/L)XXXGX(I/L)XIL (SEQ ID NO: 19), wherein X is any amino acid, and wherein said DHAD enzyme exhibits the ability to convert 2,3-dihydroxyisovalerate to ketoisovalerate in the cytosol.

154. The DHAD enzyme of claim 153, wherein said DHAD enzyme is derived from a bacterial organism.

155. The DHAD enzyme of claim 154, wherein said bacterial organism is *L. lactis* or *E. coli.*

156. The DHAD enzyme of claim 155, wherein said DHAD enzyme is from *L. lactis* and comprises the amino acid sequence of SEQ ID NO: 9.

157. The DHAD enzyme of claim 155, wherein said DHAD enzyme is from *E. coli* and comprises the amino acid sequence of SEQ ID NO: 129.

158. The DHAD enzyme of claim 153, wherein said DHAD enzyme is derived from a eukaryotic organism.

159. The DHAD enzyme of claim 158, wherein said eukaryotic organism is *Piromyces or S. cerevisiae*

160. The DHAD enzyme of claim 158, wherein said DHAD enzyme is from *Piromyces* and comprises the amino acid sequence of SEQ ID NO: 12.

161. A dihydroxyacid dehydratase (DHAD) enzyme comprising the amino acid sequence CPGXGXC (SEQ ID NO: 123), wherein X is any amino acid.

162. A dihydroxyacid dehydratase (DHAD) enzyme comprising the amino acid sequence CPGXG(A/S)C (SEQ ID NO: 124), wherein X is any amino acid.

163. A dihydroxyacid dehydratase (DHAD) enzyme comprising the amino acid sequence CXXXPGXGXC (SEQ ID NO: 125), wherein X is any amino acid.

164. A dihydroxyacid dehydratase (DHAD) enzyme having one or more amino acid deletions at the N-terminus.

165. The DHAD enzyme of claim 153, wherein said DHAD enzyme has at least about 10 amino acid deletions at the N-terminus.

166. The DHAD enzyme of claim 154, wherein said DHAD enzyme has at least about 17 amino acid deletions at the N-terminus.

167. The DHAD enzyme of claim 166, wherein said DHAD enzyme has 19 amino acid deletions at the N-terminus.

168. The DHAD enzyme of claim 155, wherein said DHAD enzyme has 23 amino acid deletions at the N-terminus.
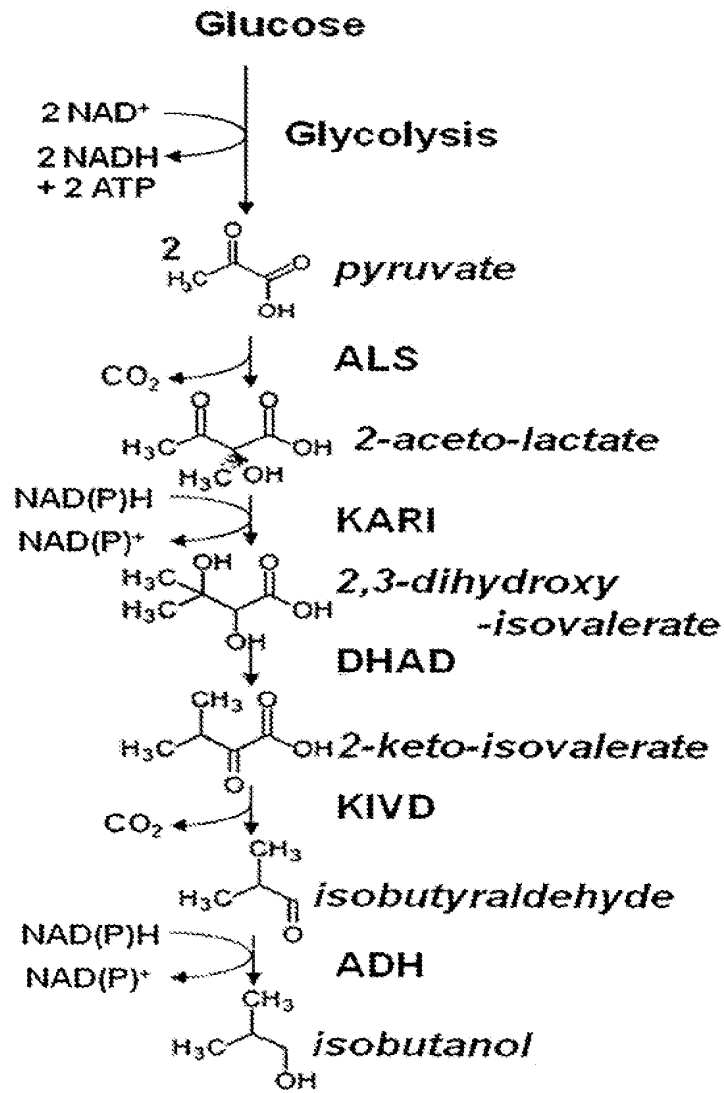
FIGURES

FIGURE 1