



(51) International Patent Classification:

C12N 1/21 (2006.01) C12P 7/06 (2006.01)
C12N 15/52 (2006.01)

(21) International Application Number:

PCT/US2012/024713

(22) International Filing Date:

10 February 2012 (10.02.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/442,109 11 February 2011 (11.02.2011) US

(71) Applicant (for all designated States except US): **THE TRUSTEES OF DARTMOUTH COLLEGE** [US/US]; 11 Rope Ferry Road, Room 6210, Hanover, NH 03755 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GUSS, Adam** [US/US]; 12938 Lady Slipper Lane, Knoxville, TN 37934 (US). **LYND, Lee, R.** [US/US]; 181 Bean Road, Meriden, NH 03770 (US).

(74) Agent: **CHEN, Xiaoyue**; Lathrop & Gage LLP, 4845 Pearl East Circle, Suite 201, Boulder, CO 80301 (US).

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

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

Published:

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

(54) Title: CLOSTRIDIUM THERMOCELLUM STRAINS FOR ENHANCED ETHANOL PRODUCTION AND METHOD OF THEIR USE

(57) Abstract: Clostridium thermocellum is a thermophilic anaerobic bacterium that specializes in the rapid solubilization and fermentation of crystalline cellulose to products that include ethanol. C. thermocellum deletion mutants have been constructed to eliminate pyruvate-formate lyase (pfl) and hydrogenases activities. C. thermocellum Apfl no longer synthesizes formate. The pfl mutants make less acetate than the parental strain. The mutant that does not express one of the maturation proteins required for production of functional hydrogenases in C. thermocellum has reduced acetate production as compared to normal strains.



CLOSTRIDIUM THERMOCELLUM STRAINS FOR ENHANCED ETHANOL PRODUCTION AND METHOD OF THEIR USE

RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional application 61/442,109 filed February 11, 2011, the entire contents of which are hereby incorporated by reference into this disclosure.

SEQUENCE LISTING

[0002] This application is accompanied by a sequence listing in a computer readable form that accurately reproduces the sequences described herein.

BACKGROUND

1. Field of the Invention

[0003] The present disclosure pertains to the field of biomass processing to produce ethanol. In particular, new thermophilic organisms that can use a variety of biomass derived substrates and produce ethanol in high yield are disclosed.

2. Description of the Related Art

[0004] Lignocellulosic biomass represents one of the most abundant renewable resources on Earth. It is formed of three major components - cellulose, hemicellulose, and lignin - and includes, for example, agricultural and forestry residues, municipal solid waste (MSW), fiber resulting from grain operations, waste cellulose (e.g., paper and pulp operations), and energy crops. The cellulose and hemicellulose polymers of biomass may be hydrolyzed into their component sugars, such as glucose and xylose, which can then be fermented by microorganisms to produce ethanol. Conversion of even a small portion of the available biomass into ethanol could substantially reduce current gasoline consumption and dependence on petroleum.

[0005] Significant research has been performed in the areas of reactor design, pretreatment protocols and separation technologies, so that bioconversion processes are becoming economically competitive with petroleum fuel technologies. However, it is estimated that the largest cost savings may be achieved by combining two or more process steps. For example, simultaneous saccharification and fermentation (SSF) and simultaneous saccharification and co-fermentation (SSCF) processes combine an

enzymatic saccharification step with fermentation in a single reactor or continuous process apparatus. In an SSF process, end-product inhibition is removed as the soluble sugars are continually fermented into ethanol. When multiple sugar types are fermented by the same organism, the SSF process is usually referred to as a simultaneous saccharification and co-fermentation (SSCF) process.

[0006] In addition to savings associated with shorter reaction times and reduced capital costs, co-fermentation processes may also provide improved product yield because certain compounds that would otherwise accrue at levels that inhibit metabolism or hydrolysis are consumed by the co-fermenting organism(s). In one such example, β -glucosidase ceases to hydrolyze cellobiose in the presence of glucose and, in turn, the build-up of cellobiose impedes cellulose degradation. An SSCF process involving co-fermentation of cellulose and hemicellulose hydrolysis products may alleviate this problem by converting glucose into one or more products that do not inhibit the hydrolytic activity of β -glucosidase.

[0007] Consolidated bioprocessing (CBP) involves four biologically-mediated events: (1) enzyme production, (2) substrate hydrolysis, (3) hexose fermentation and (4) pentose fermentation. In contrast to conventional approaches, which perform each step independently, all four events may be performed simultaneously in a CBP configuration. This strategy requires a microorganism that utilizes both cellulose and hemicellulose. Otherwise, a CBP process that utilizes more than one organism to accomplish the four biologically-mediated events is referred to as a consolidated bioprocessing co-culture fermentation.

[0008] In SSF, SSCF and CBP processes, bacterial strains that have the ability to convert pentose sugars into hexose sugars, and to ferment the hexose sugars into a mixture of organic acids and other products via glycolysis perform a crucial function. The glycolytic pathway begins with conversion of a six-carbon glucose molecule into two three-carbon molecules of pyruvate. Pyruvate may then be converted to lactate by the action of lactate dehydrogenase ("*ldh*"), or to acetyl coenzyme A ("*acetyl-CoA*") by the action of pyruvate dehydrogenase or pyruvate-ferredoxin oxidoreductase. Acetyl-CoA is further converted to acetate by phosphotransacetylase ("*pta*") and acetate kinase ("*ack*"), or reduced to ethanol by acetaldehyde dehydrogenase ("*AcDH*") and alcohol dehydrogenase ("*adh*").

[0009] Carbohydrate metabolic pathways, such as those described above, may be altered by directing the flow of carbon to a desired end product, such as ethanol. *See generally*, Lynd, L. R., P. J. Weimer, W. H. van Zyl, and I. S. Pretorius (2002) Microbial cellulose utilization: Fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66: 506. A "carbon-centered" approach to metabolic engineering involves inactivating enzymatic pathways that direct carbon containing molecules away from ethanol or otherwise promoting the flow of carbon towards ethanol. For instance, Desai, S. G., M. L. Gueriot, L. R. Lynd (2002) Cloning of L-lactate dehydrogenase and elimination of lactic acid production via gene knockout in *Thermoanaerobacterium saccharolyticum* JW/SL-YS485. Appl. Microbiol. Biotechnol. 65: 600-605 and PCT/US07/67941, describe the inactivation of L-lactate dehydrogenase (*ldh*) alone and in combination with acetate kinase (*ack*) and/or phosphotransacetylase (*pta*), respectively, which results in strains that produce ethanol in higher yield than the native organisms.

[0010] Although a "carbon-centered" approach to producing knockout organisms represents an advance in the art, additional and/or alternative approaches to modifying the glycolytic pathway may result in more efficient biomass conversion.

SUMMARY

[0011] The present instrumentalities advance the art by providing methods for manipulating branched end-product metabolism of fermentative microorganisms. The end-product ratio between ethanol and organic acids is altered by eliminating one or more enzymatic activities associated with the formation of hydrogen. More specifically, one or more mutations in a maturation protein implicated in post-translational modification of hydrogenases are introduced into a organism to modify the hydrogenase activities within the organism. For purpose of this disclosure, the maturation proteins that facilitates the maturation of other proteins may be referred to as "maturase" or "maturation protein," and those maturases that facilitate maturation of hydrogenase may be referred to as "hydrogenase maturase," or "hydrogenase maturation protein."

[0012] In another aspect of the disclosure, one or more genes encoding enzymes implicated in the production of organic acids are mutated to reduce production of these organic acids and to direct the carbon flow towards ethanol. Example of organic acids may include but are not limited to lactic acid, acetic acid, formic acid or salts thereof. The resulting organism may utilize a variety substrates derived from biomass to

generate ethanol in higher yield as compared to ethanol yield by the parental strain. In one aspect, the gene encoding the pyruvate-formate lyase (*pfl*) which is required for the production of formate may be disrupted in order to obtain an modified organism with higher ethanol yield. Methods for generating such organisms by genetic engineering are also disclosed.

[0013] In one embodiment, mutations may be generated in various genes either singly or in combination, where such genes encode proteins that play some roles in the formation of hydrogen and/or organic acids in the native (host) organism. These genes may include but are not limited to: (a) hydrogenase genes, (b) genes encoding one or more maturation proteins that modify hydrogenase post-translationally, such as HydE, HydF and HydG, (c) any other genes encoding proteins that modulate the activities of hydrogenase, (d) pyruvate-formate lyase (*pfl*) gene, (e) *pfl* and/or one or more of HydE, HydF and HydG genes, (f) acetate kinase (*ack*) gene, (g) phosphotransacetylase (*pta*) gene, and (h) lactate dehydrogenase (*ldh*) gene; or (i) combination of two or more genes from (a)-(h) listed above.

[0014] For purpose of this disclosure, a mutation may include but not limited to knockout, deletion, insertion, substitution and so on. Mutations may affect the coding sequence which determine the sequence of the encoded proteins. Mutations may also occur in the non-coding regions of the target genes.

[0015] In another aspect, enzymatic activities may be eliminated by rendering non-expression of a target gene through methods other than mutations. For instance, expression of a gene can be eliminated by RNA interference, gene silencing, among others.

[0016] In an embodiment, an organism capable of fermenting a saccharification product of a substrate derived from a biomass is created to obtain a modified organism capable of producing ethanol in higher yield as compared to the ethanol yield of the unmodified organism. In one aspect, a first gene endogenous to the organism which encodes a post-translational modifying protein is inactivated through genetic engineering, which, in turn, leads to a decrease in hydrogenase activities in the organism. Preferably, the saccharification product is derived from a carbohydrate-rich biomass substrate.

[0017] In another embodiment, the post-translational modifying protein is a protein capable of facilitating maturation of at least one hydrogenase synthesized by said

organism in its native state. In one aspect, the post-translational modifying protein may be a protein that facilitates the assembly of protein complexes having hydrogenase activities. In another aspect, the first gene may be a gene selected from the group consisting of HydE (SEQ ID NO. 1), HydF (SEQ ID NO. 2) and HydG (SEQ ID NO. 3). In another aspect, the first gene may encode a protein having at least 80%, 90%, 95%, 99% or more preferably, 100% sequence identity with a protein encoded by a polynucleotide sequence selected from the group consisting of SEQ ID Nos. 1-3.

[0018] In another embodiment, a first gene endogenous to the organism which encodes a protein required for the generation of formate or formic acid is inactivated through genetic engineering, which, in turn, leads to a decrease in formate production in the organism. In one aspect, the first gene may encode a pyruvate-formate lyase (pfl). In another aspect, the first gene may encode a protein having at least 80%, 90%, 95%, 99%, or more preferably, 100% sequence identity with a protein encoded by the pflA gene (SEQ ID No. 4) or pflB gene (SEQ ID No. 5) of *C. thermocellum*. Note that in *C. thermocellum*, the the pflA and pflB genes are so close to each other that the pflA locus and the pflB locus may be referred to as the pflAB locus.

[0019] In another embodiment, the organism is a bacterium, preferably, a thermophilic, anaerobic bacterium, more preferably, a *Clostridium thermocellum*. In another embodiment, the organism is a bacterium strain deposited under Patent Deposit Designation No. PTA-11763 or PTA-11764.

[0020] In another embodiment, an isolated polynucleotide comprising a polynucleotide sequence having at least 90%, 95%, 99%, or more preferably, 100% sequence identity with a polynucleotide sequence selected from the group consisting of SEQ ID Nos. 1-5 is described.

[0021] In another embodiment, a method for producing ethanol includes generating a modified organism with a first gene encoding a post-translational modifying protein inactivated, and incubating the organism in a medium containing at least one substrate. Examples of the substrate may include but are not limited to glucose, xylose, mannose, arabinose, galactose, fructose, cellobiose, sucrose, maltose, xylan, mannan, starch, cellulose, pectin or combinations thereof. In one aspect, the post-translational modifying protein is capable of facilitating maturation of at least one hydrogenase synthesized by the organism in its native state. By inactivating the first gene encoding

such a post-translational modifying protein, the modified organism has decreased hydrogenase activities which, in turn, leads to enhanced ethanol production.

[0022] In another embodiment, a method for producing ethanol includes generating a modified organism with a first gene inactivated which encodes a protein required for the formation of formate or formic acid, and incubating the organism in a medium containing at least one substrate. Examples of the substrate may include but are not limited to glucose, xylose, mannose, arabinose, galactose, fructose, cellobiose, sucrose, maltose, xylan, mannan, starch, cellulose, pectin or combinations thereof.

[0023] In another embodiment, a method for producing ethanol includes providing within a reaction vessel, a reaction mixture comprising a carbohydrate-rich biomass substrate, a cellulolytic material, and a fermentation agent. In one aspect, the fermentation agent contains a bacterium that has been genetically modified to inactivate a first gene encoding a post-translational modifying protein. Preferably, the post-translational modifying protein is a maturation protein (or maturase) that facilitates the hydrogenase in the bacterium. The reaction mixture may be incubated under suitable conditions for a period of time sufficient to allow saccharification and fermentation of the carbohydrate-rich biomass substrate. In another aspect, the fermentation agent contains a bacterium that has been genetically modified to inactivate a first gene encoding a protein required for the production of formate or formic acid by the bacterium. In another aspect, the reaction mixture may be incubated under a suitable condition wherein the temperature is at least 50°C.

[0024] In another embodiment, one or more genes in addition to the *pfl* gene and/or the maturation genes may be inactivated in an organism. Preferably, these one or more genes encode one or more proteins selected from the group consisting of lactate dehydrogenase (*ldh*), acetate kinase (*ack*), phosphotransacetylase (*pta*), hydrogenase, and combination thereof.

[0025] In another embodiment, a modified organism may be created wherein both the *pfl* gene and at least one gene encoding a hydrogenase maturase are inactivated.

[0026] In another embodiment, a modified organism may be created wherein the *pfl* gene, the *ldh* gene and at least one gene encoding a hydrogenase maturase are inactivated.

[0027] In yet another embodiment, a modified organism may be created wherein the pfl gene, the ldh gene, the ack gene, the pta gene, and at least one gene encoding a hydrogenase maturase are inactivated. These various organisms may be used alone or in combination with other microorganisms in the conversion of biomass to ethanol.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] Fig. 1 shows a modified glycolytic pathway after hydrogenase inactivation, according to an embodiment.

[0029] Fig. 2 shows the sequences of the hydrogenase maturase genes HydE (SEQ ID NO. 1), HydF (SEQ ID NO. 2), HydG (SEQ ID NO. 3), and the sequences of the pyruvate-formate lyase pflA gene (SEQ ID No. 4), and pflB gene (SEQ ID No. 5) of *C. thermocellum*.

[0030] Fig. 3 shows a diagram of the plasmid pAMG258, along with its full sequence.

[0031] Fig. 4 shows a diagram of the plasmid pAMG278.

[0032] Fig. 5 shows a diagram of the plasmid pAMG281.

[0033] Fig. 6 panels A and B are eletrophoresis pictures showing deletion of the HydG gene.

[0034] Fig. 7 shows the end-product profile in the mutant hydG strain and mutant pfl strain.

DETAILED DESCRIPTION

[0035] There will now be shown and described methods for engineering and utilizing thermophilic, anaerobic, Gram-positive bacteria in the conversion of biomass to ethanol.

[0036] As used herein, an organism is "in a native state" if it has not been genetically engineered or otherwise manipulated by the hand of man in a manner that alters the genotype and/or phenotype of the organism. For example, a wild-type organism may be considered to be in a native state.

[0037] "Identity" refers to a comparison between sequences of polynucleotide or polypeptide molecules. Methods for determining sequence identity are commonly known. Computer programs typically employed for performing an identity comparison include, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8

for Unix, Genetics Computer Group, University Research Park, Madison Wisconsin), which uses the algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2: 482-489.

[0038] "Lignocellulosic substrate" generally refers to any lignocellulosic biomass suitable for use as a substrate to be converted into ethanol. A "Lignocellulosic substrate" may be referred to as a "substrate" in this disclosure.

[0039] "Saccharification" refers to the process of breaking a complex carbohydrate, such as starch or cellulose, into its monosaccharide or oligosaccharide components. For purposes of this disclosure, a complex carbohydrate is preferably processed into its monosaccharide components during a saccharification process.

[0040] The term "endogenous" is used to describe a molecule that exists naturally in an organism. A molecule that is introduced into an organism using molecular biology tools, such as transgenic techniques, is not endogenous to that organism.

[0041] The terms "inactivated", "inactivate", "gene inactivation", "knock-out (knockout)", "disrupt" or "non-expression" may be used interchangeably, and they refer to a process by which a gene is rendered substantially non-expressing and/or non-functional. The term "substantially" means more than seventy percent. Thus, for purposes of this disclosure, a gene is considered inactivated if its expression or its function has been reduced by more than seventy percent. Techniques for inactivation of a target gene may include, but are not limited to, deletion, insertion, substitution in the coding or non-coding regulatory sequences of the target gene, as well as the use of RNA interference to suppress gene expression. The process of inactivating a gene is also referred to as "knocking out" a gene. Thus, an organism that has one or more of its genes inactivated may be called a "knockout" (KO) strain.

[0042] For purposes of this disclosure, an organism that possesses the necessary biological and chemical components, including polynucleotides, polypeptides, carbohydrates, lipids and other molecules, as well as cellular or subcellular structures that may be required for performing or facilitating certain biological and/or chemical processes is deemed to be capable of performing said processes. Thus, an organism that contains certain inducible genes may be considered capable of performing the function attributable to the proteins encoded by those genes.

[0043] The term "genetic engineering" is used to refer to a process by which genetic materials, including DNA and/or RNA, are manipulated in a cell or introduced into a cell to affect expression of certain proteins in the cell. Manipulation may include

introduction of a foreign (or "exogenous") gene into the cell or inactivation or modification of an endogenous gene. Such a modified cell may be called a "genetically engineered cell" or a "genetically modified cell". If the original cell to be genetically engineered is a bacterial cell, said genetically engineered cell may be said to have been derived from a bacterial cell. A molecule that is introduced into a cell to genetically modify the cell may be called a genetic construct. A genetic construct typically carries one or more DNA or RNA sequences on a single molecule.

[0044] The expression of a protein is generally regulated by the non-coding region of a gene termed promoter. When a promoter controls the transcription of a gene, it can also be said that the expression of the gene (or the encoded protein) is driven by the promoter. When a promoter is placed in proximity of a coding sequence, such that transcription of the coding sequence is under control of the promoter, it can be said that the coding sequence is operably linked to the promoter. A promoter that is not normally associated with a gene is called a heterologous promoter.

[0045] A "cellulolytic material" is a material that may facilitate the breakdown of cellulose into its component oligosaccharides or monosaccharides. For example, cellulolytic material may comprise a cellulase or hemicellulase.

[0046] It is to be noted that, as used in this specification and the claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Capital letters may be used when referring to certain genes or proteins. For purpose of this disclosure, upper case or lower case letters may be used interchangeably in the names of genes or proteins. So for example, the terms "hydG" and "HydG" may be used to refer to the same gene.

[0047] As discussed above, carbohydrate metabolic pathways in a microorganism may be altered by directing the flow of carbon to a desired end product, such as ethanol, using a "carbon-centered" approach to metabolic engineering. An alternative, "electron-centered" approach, is disclosed herein where ethanol yield may be increased by inactivation of an enzymatic pathway that produces hydrogen. For example, Fig. 1 illustrates a portion of the glycolytic pathway, where a cross indicates blocking of hydrogenase activity that leads to hydrogen production. Based on stoichiometric equations, hydrogen production is related to acetic acid production. Therefore, disrupting the ability of an organism to produce hydrogen may result in decreased production of acetic acid and increased ethanol production.

[0048] The thermophilic bacterium, *Clostridium thermocellum*, is used by way of example to illustrate how maturases in an organism may be manipulated to affect hydrogenase activities and to increase ethanol production. The methods and materials disclosed herein may however apply to members of the *Clostridium*, *Thermoanaerobacter* and *Thermoanaerobacterium* genera, as well as other microorganisms. Members of the *Clostridium* genus may include but are not limited to, *Clostridium thermosulfurogenes*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium thermohydrosulfuricum*, *Clostridium thermoaceticum*, *Clostridium thermosaccharolyticum*, *Clostridium tartarivorum*, *Clostridium thermocellulaseum*. Members of the *Thermoanaerobacter* and *Thermoanaerobacterium* genera may include, for example, *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zeae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter brockii*, variants thereof, and/or progeny thereof. Both the carbon-centered and the electron-centered approaches for maximizing ethanol production from biomass may be applicable in metabolic engineering of other microorganisms, such as yeast or fungi.

[0049] Major groups of bacteria include eubacteria and archaeobacteria. Thermophilic eubacteria include: phototropic bacteria, such as cyanobacteria, purple bacteria and green bacteria; Gram-positive bacteria, such as *Bacillus*, *Clostridium*, lactic acid bacteria and Actinomyces; and other eubacteria, such as *Thiobacillus*, Spirochete, *Desulfotomaculum*, Gram-negative aerobes, Gram-negative anaerobes and *Thermotoga*. In certain embodiments, the present instrumentalities relate to Gram-negative organotrophic thermophiles of the genus *Thermus*; Gram-positive eubacteria, such as *Clostridium*, which comprise both rods and cocci; eubacteria, such as *Thermosipho* and *Thermotoga*; archaeobacteria, such as *Thermococcus*, *Thermoproteus* (rod-shaped), *Thermofilum* (rod-shaped), *Pyrodictium*, *Acidianus*, *Sulfolobus*, *Pyrobaculum*, *Pyrococcus*, *Thermodiscus*, *Staphylothermus*, *Desulfurococcus*, *Archaeoglobus* and *Methanopyrus*. Some examples of thermophilic or mesophilic organisms (including bacteria, prokaryotic microorganisms and fungi), which may be suitable for use with the disclosed instrumentalities include, but are not limited to: *Clostridium*

thermosulfurogenes, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium thermohydrosulfuricum*, *Clostridium thermoaceticum*, *Clostridium thermosaccharolyticum*, *Clostridium tartarivorum*, *Clostridium thermocellulaseum*, *Anaerocellum* sp., *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermobacteroides acetoethylicus*, *Thermoanaerobium brockii*, *Methanobacterium thermoautotrophicum*, *Pyrodictium occultum*, *Thermoproteus neutrophilus*, *Thermofilum librum*, *Thermothrix thioparus*, *Desulfovibrio thermophilus*, *Thermoplasma acidophilum*, *Hydrogenomonas thermophilus*, *Thermomicrobium roseum*, *Thermus flavas*, *Thermus ruber*, *Pyrococcus furiosus*, *Thermus aquaticus*, *Thermus thermophilus*, *Chloroflexus aurantiacus*, *Thermococcus litoralis*, *Pyrodictium abyssi*, *Bacillus stearothermophilus*, *Cyanidium caldarium*, *Mastigocladus laminosus*, *Chlamydothrix calidissima*, *Chlamydothrix penicillata*, *Thiothrix carnea*, *Phormidium tenuissimum*, *Phormidium geysericola*, *Phormidium subterraneum*, *Phormidium bijahensi*, *Oscillatoria filiformis*, *Synechococcus lividus*, *Chloroflexus aurantiacus*, *Pyrodictium brockii*, *Thiobacillus thiooxidans*, *Sulfolobus acidocaldarius*, *Thiobacillus thermophila*, *Bacillus stearothermophilus*, *Cercosulcifer hamathensis*, *Vahlkampfia reichi*, *Cyclidium citrullus*, *Dactylaria gallopava*, *Synechococcus lividus*, *Synechococcus elongatus*, *Synechococcus minervae*, *Synechocystis aquatilis*, *Aphanocapsa thermalis*, *Oscillatoria terebriformis*, *Oscillatoria amphibia*, *Oscillatoria germinata*, *Oscillatoria okenii*, *Phormidium laminosum*, *Phormidium parparasiensis*, *Symploca thermalis*, *Bacillus acidocaldarias*, *Bacillus coagulans*, *Bacillus thermocatenalatus*, *Bacillus licheniformis*, *Bacillus pamilas*, *Bacillus macerans*, *Bacillus circulans*, *Bacillus laterosporus*, *Bacillus brevis*, *Bacillus subtilis*, *Bacillus sphaericus*, *Desulfotomaculum nigrificans*, *Streptococcus thermophilus*, *Lactobacillus thermophilus*, *Lactobacillus bulgaricus*, *Bifidobacterium thermophilum*, *Streptomyces fragmentosporus*, *Streptomyces thermonitrificans*, *Streptomyces thermovulgaris*, *Pseudonocardia thermophila*, *Thermoactinomyces vulgaris*, *Thermoactinomyces sacchari*, *Thermoactinomyces candidas*, *Thermomonospora curvata*, *Thermomonospora viridis*, *Thermomonospora citrina*, *Microbispora thermodiastatica*, *Microbispora aerata*, *Microbispora bispore*, *Actinobifida dichotomica*, *Actinobifida chromogena*, *Micropolyspora caesia*, *Micropolyspora faeni*, *Micropolyspora cectivugida*, *Micropolyspora cabrobrunea*, *Micropolyspora thermovirida*, *Micropolyspora viridinigra*, *Methanobacterium thermoautotrophicum*, variants thereof, and/or progeny thereof.

[0050] In certain embodiments, thermophilic bacteria for use with the disclosed instrumentalities may be selected from the group consisting of *Fervidobacterium gondwanense*, *Clostridium thermolacticum*, *Moorella* sp. and *Rhodothermus marinus*.

[0051] In certain embodiments, the disclosed instrumentalities relate to microorganisms of the genera *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus* and *Anoxybacillus*, including but not limited to species selected from the group consisting of: *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldoxylosilyticus*, *Saccharococcus thermophilus*, *Paenibacillus campinasensis*, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*, *Anoxybacillus gonensis*, variants thereof, and/or progeny thereof.

[0052] In certain embodiments, the disclosed instrumentalities relate to mesophilic bacteria selected from the group consisting of *Saccharophagus degradans*; *Flavobacterium johnsoniae*; *Fibrobacter succinogenes*; *Clostridium hungatei*; *Clostridium phytofermentans*; *Clostridium cellulolyticum*; *Clostridium aldrichii*; *Clostridium termitididis*; *Acetivibrio cellulolyticus*; *Acetivibrio ethanoligignens*; *Acetivibrio multivorans*; *Bacteroides cellulosolvens*; and *Alkalibacter saccharofomentans*, variants thereof, and/or progeny thereof.

[0053] In certain preferred embodiments, the disclosed instrumentalities relate to organisms having a ferredoxin-linked hydrogenase (EC subclass 1.12.7.2), including but not limited to organisms selected from the groups of eubacteria and achaeobacteria, phototropic bacteria (such as cyanobacteria, purple bacteria and green bacteria), Gram-positive bacteria and lactic acid bacteria and Gram-negative anaerobes, as well as organisms selected from the genera including, but not limited to: *Bacillus*, *Clostridium*, *Thermotoga*, *Pyrococcus* and *Saccharococcus*. Such organisms include those selected from the group consisting of: *Thermotoga maritima*, *Clostridium acetobutylicum*, *Clostridium pasteurianum*, *Clostridium beijerinckii*, *Clostridium thermosulfurogenes*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium thermohydrosulfuricum*, *Clostridium thermosaccharolyticum*, *Clostridium tartarivorum*, *Clostridium thermocellulaseum*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermobacteroides acetoethylicus*, *Thermoanaerobium brockii*, *Pyrococcus furiosus*, *Bacillus coagulans*, *Clostridium thermolacticum*, *Clostridium hungatei*, *Clostridium phytofermentans*, *Clostridium*

cellulolyticum, *Clostridium aldrichii*, *Clostridium termitididis*, *Acetivibrio cellulolyticus*, *Acetivibrio ethanolgignens*, *Acetivibrio multivorans*, *Bacteroides cellulosolvens*, *Alkalibacter saccharofomentans*, variants thereof, and/or progeny thereof.

[0054] By way of example, Figure 1 shows the general pathways of sugar fermentation. During sugar fermentation, hexoses may be converted to pyruvate through the glycolytic pathway. Pyruvate is a branch point from which carbon and electron may flow to various fermentation end products. Examples of these end products include but are not limited to lactate, acetate, formate, carbon dioxide, hydrogen, and ethanol.

[0055] Because ethanol and acetate are the major fermentation end products in *Clostridium thermocellum*, reducing the production of acetate becomes a priority in an effort to enhance ethanol yield. In *C. thermocellum* and many other species, carbon flux to acetate is coupled with electron flux to either formate or hydrogen. It is disclosed here that disruption of the genes involved in the production of either formate or hydrogen helps divert carbon and electrons away from acetate, which, in turn, increases ethanol yield.

[0056] Three classes of enzymes are capable of producing hydrogen, namely, nitrogenases, alkaline phosphatases and hydrogenases. Due to their highly reactive and complex metallocenters, hydrogenases are considered the most efficient enzymes for hydrogen production. For review, see Calusinska *et al.*, Microbiology, 156:1575-88 (2010). The genome of *Clostridium thermocellum* contains four putative hydrogenase genes. Based on their predicted amino acid sequences, one hydrogenase gene encodes a Ni-Fe hydrogenase and other three encode Fe-Fe hydrogenases.

[0057] After being synthesized in a cell, hydrogenases typically require extensive post-translational modification (i.e., maturation) before gaining enzymatic activity. In *Clostridium thermocellum*, all Fe-Fe hydrogenases share one single dedicated protein maturation system, which contains at least maturation proteins: HydE, HydF and HydG, all of which are required for the maturation of Fe-Fe hydrogenases in *C. thermocellum*.

[0058] It is disclosed here the sequences of the genes encoding these three maturation proteins: HydE (SEQ ID NO: 1), HydF (SEQ ID NO: 2) and HydG (SEQ ID NO: 3). In another aspect, a polynucleotide having 70%, 80%, 90%, 99% sequence identity with the polynucleotides of SEQ ID Nos. 1-3 may encode a functional hydrogenase maturase.

[0059] *C. thermocellum* contains two putative pyruvate-formate lyase (pfl) genes which are predicted to be involved in formate production during pyruvate conversion to acetyl-CoA. It is disclosed here the sequence of the genes encoding the putative pyruvate-formate lyase, namely, pflA (SEQ ID NO: 4) and pflB (SEQ ID NO: 5). In another aspect, a polynucleotide having 70%, 80%, 90%, 99% sequence identity with the polynucleotide of SEQ ID No. 4 or SEQ ID No. 5 may encode a functional putative pyruvate-formate lyase. It is disclosed here that disruption of one or both of the pfl genes in *C. thermocellum* reduces formate formation and increases ethanol production.

[0060] In another aspect, one or both of the pfl genes as well as one or all of the hydrogenase maturation protein genes may be disrupted in one organism. In another aspect, one or more of the genes involved in the production of lactate, acetate, formate, or hydrogen may be disrupted in an organism in order to direct carbon flow to ethanol. For example, lactate dehydrogenase (*ldh*), the gene that confers the ability to produce lactic acid, and acetate kinase (*ack*) and/or phosphotransacetylase (*pta*), the genes that confer the ability to produce acetic acid, may be targeted for gene disruption as described in PCT/US07/67941, which is incorporated by reference herein.

[0061] In another aspect, an organism may be generated in which all hydrogenase activities leading to synthesis of hydrogen are disrupted in order to maximize ethanol production. For instance, maturation proteins required for enzymatic activities of Fe-Fe hydrogenase and Ni-Fe hydrogenase may all be inactivated to remove any residual hydrogen production.

[0062] The use of the maturation protein knockout strains and the pfl knockout strains, either alone or in combination with *ldh*, *ack* and *pta* knockout strains, may contribute significant cost savings to the conversion of biomass to ethanol due to their growth conditions, which are substantially optimal for cellulase activity in SSF and SSCF processes. For example, optimal cellulase activity parameters include a pH between 4-5 and temperature between 40-50°C, which are substantially similar to the optimal growth conditions of thermophilic bacteria. By way of comparison, the optimal growth temperature for *T. saccharolyticum* is about 50-60°C. (Esterbauer, H., W. Steiner, I. Labudova, A. Hermann, and M. Hayn. (1991) Production of *Trichoderma* Cellulase in Laboratory and Pilot Scale. Bioresource Technology 36: 51-65.) Thus, if the reaction is carried out within the temperature range of 40-60°C, the biocatalysts and cellulases may both achieve their maximal activities. One benefit of this overlap in optimal temperature

is that the amount of cellulase required for producing the same amount of ethanol may be lowered by as much as two-thirds resulting in a significant cost reduction. *See, e. g.,* Mabee, W. E. and J. N. Saddler (2005) Progress in Enzymatic Hydrolysis of Lignocellulosics. Additionally, it is unnecessary to adjust the pH of the fermentation broth when knockout organisms, which lack the ability to produce organic acids, are used. These knockout organisms may also be suitable for a consolidated bioprocessing co-culture fermentation where cellulose may be degraded by a cellulolytic organism such as *C. thermocellum* and these knockout organisms may convert pentoses to ethanol. *C. thermocellum* is capable of rapidly degrading cellulose, but it cannot ferment pentose sugars, which, in the form of xylan and other polysaccharides, may account for up to 30% of total carbohydrates in a typical saccharified biomass. By contrast, *T. saccharolyticum* is capable of fermenting and utilizing pentose sugars. A process utilizing both *C. thermocellum* and a knockout of *T. saccharolyticum* may therefore be an efficient way to improve cellulosic ethanol production, and reduce process costs. *See* Lynd, L. R., W. H. van Zyl, J. E. McBride, and M. Laser (2005) Consolidated bioprocessing of cellulosic biomass: an update. *Curr. Opin. Biotechnol.* 16: 577-583.

[0063] Operating either an SSF, SSCF or CBP process at thermophilic temperatures offers several important benefits over conventional mesophilic fermentation temperatures of 30-37°C. In particular, enzyme concentrations necessary to achieve a given amount of conversion may be reduced due to higher enzyme activity at thermophilic temperatures. As a result, costs for a process step dedicated to cellulase production are substantially reduced for thermophilic SSF and SSCF (e.g., 2-fold or more), and are eliminated for CBP. Costs associated with fermentor cooling and heat exchange before and after fermentation are also expected to be reduced for thermophilic SSF, SSCF and CBP. Finally, processes featuring thermophilic biocatalysts may be less susceptible to microbial contamination as compared to processes featuring conventional mesophilic biocatalysts.

[0064] In another aspect, a method for producing ethanol includes providing within a reaction vessel, a reaction mixture comprising lignocellulosic substrate, a cellulolytic material and a fermentation agent. The fermentation agent comprises an organism that has been transformed to eliminate expression of at least one gene encoding a hydrogenase maturase or the pyruvate-formate lyase. The reaction mixture is reacted under suitable conditions for a period of time sufficient to allow saccharification and

fermentation of the lignocellulosic substrate. Appropriate substrates for the production of ethanol include, for example, one or more of glucose, xylose, cellobiose, sucrose, xylan, starch, cellulose, pectin and combinations thereof. These substrates may, in some aspects, be produced during an SSF, SSCF or CBP process to achieve efficient conversion of biomass to ethanol.

[0065] It will be appreciated that carbohydrate-rich biomass material that is saccharified to produce one or more of glucose, xylose, mannose, arabinose, galactose, fructose, cellobiose, sucrose, maltose, xylan, mannan, starch cellulose and pectin may be utilized by the disclosed organisms. In various embodiments, the biomass may be lignocellulosic biomass that comprises wood, corn stover, sawdust, bark, leaves, agricultural and forestry residues, grasses such as switchgrass, ruminant digestion products, municipal wastes, paper mill effluent, newspaper, cardboard, or combinations thereof.

Deposit of Strains

[0066] The two strains disclosed herein where the pfl gene or the hydG gene have been inactivated have been deposited with the American Type Culture Collection, Manassas, VA 20110-2209, and received Patent Deposit Designation Number PTA-11764 (Δ hpt Δ pfl) and PTA-11763 (Δ hpt Δ hydG), respectively. The deposit was made on March 18, 2011. This deposit was made in compliance with the Budapest Treaty requirements that the duration of the deposit should be for thirty (30) years from the date of deposit or for five (5) years after the last request for the deposit at the depository or for the enforceable life of a U.S. Patent that matures from this application, whichever is longer. The strains will be replenished should it become non-viable at the depository.

EXAMPLES

[0067] The following examples are provided for purpose of illustration, but are not intended to be limiting. The reaction conditions such as pH, temperature, are presented as typical components, and various substitutions or modifications may be made in view of the foregoing disclosure by one of skills in the art without departing from the principle and spirit of the present invention.

EXAMPLE 1

Isolation and Deletion/Disruption of Target Genes in *Clostridium thermocellum*

[0068] A plasmid was constructed using yeast gap repair cloning (Burke, D., D. Dawson, and T. Stearns. 2000. Methods in yeast genetics: a Cold Spring Harbor Laboratory course manual. Cold Spring Harbor Laboratory Press, Plainview, N.Y.) to delete the *hydG* gene from the *C. thermocellum* chromosome. A region ca. 1 kilobase (kb) upstream of the *C. thermocellum hydG* gene was amplified by polymerase chain reaction (PCR) using primers:

AATGGGGCGCTACAGGGCGCGTGTTGGGATGATCCACTAGTAAGCTTTAA
TAAGGAGTTCTATCCAGCTTAGC (SEQ ID NO. 7)

and

ATTGCTTCTAATATCACTTAAATTTTTCAGCTTTCAATTTACAAACCTTG
TATCG (SEQ ID NO. 8)

[0069] Another region ca. 1 kilobase downstream of the *C. thermocellum hydG* gene was PCR amplified using primers:

TGGGTTATTGATACACCCATGTCAGAATACTGGCCGGCCTCGCGAAAT
TCTGATATTGTCAACTCCGAA (SEQ ID NO. 9)

and

ATATGTGCCGATACAAGGTTTGTAATTGAAAGCTGAAAAATTTAAGT
GATATTAGAAGC (SEQ ID NO. 10)

[0070] The PCRs were performed such that these two PCR products would fuse with each other and with the regions flanking the BamHI site in plasmid pAMG258 (Fig. 3; SEQ ID NO: 6) during gap repair cloning. Simultaneously, an internal region of the *C. thermocellum hydG* gene ca. 1 kilobase in size was PCR amplified using primers:

GGCCCTCTAGGATCAGCGGGTTTAAACGCTGAGGCGCGCCCCCGGTAC
GACATCTTCTGT (SEQ ID NO. 11)

and

AAAGGCTCAGTCGGAAGACTGGGCCTTTTGTGTTTGGTACCCCTGGGAG
AAGTTGCAAAGC (SEQ ID NO. 12)

[0071] This PCR was performed such that the PCR product would fuse with the regions flanking the EcoRI site in plasmid pAMG258. The resulting plasmid, pAMG278 (Fig. 4), has a thiamphenicol resistance gene (*cat*) and hypoxanthine

phosphoribosyltransferase gene (*hpt*) that are expressed together from the *C. thermocellum* *gapD* promoter for positive and negative selection in *C. thermocellum* flanked on one side by a fusion of the upstream and downstream region of *C. thermocellum* *hydG* and on the other side by an internal fragment of *hydG*. The plasmid also contains the *Thermoanaerobacterium saccharolyticum* *tdk* gene expressed from the *C. thermocellum* *cbp* promoter for a second negative selection in *C. thermocellum*, the pUC19 origin of replication and *bla* gene for replication and selection in *Escherichia coli*, the URA3+ gene and CEN6/ARSH4 origin of replication for selection and propagation of the plasmid in *Saccharomyces cerevisiae*, and the pNW33N origin of replication for plasmid replication in *C. thermocellum*.

[0072] The DNA was transformed into yeast via a modified Lazy Bones protocol (Elble, R. 1992. A simple and efficient procedure for transformation of yeasts. *Biotechniques* 13:18-20.) and was assembled into a contiguous piece of DNA via yeast homologous recombination. Plasmid DNA was isolated from yeast using Zymoprep Yeast Plasmid Miniprep II kit (Zymo, Orange, CA, USA.) and introduced via electroporation into *E. coli* Top10 (Invitrogen, Carlsbad, CA) and then via chemical competence into *E. coli* BL21 (DE3) (New England Biolabs, Ipswich, MA, USA). All PCR amplified regions were sequenced at the Dartmouth College Molecular Biology Core Facility to verify PCR fidelity.

[0073] A second plasmid, pAMG281 (Fig. 5), was constructed in an essentially identical manner to delete *pflAB* from the *C. thermocellum* genome, except that the upstream region of *pflAB* was PCR amplified using primers:

AATGGGGCGCTACAGGGCGCGTGGGGATGATCCACTAGTAAGCTTTGA
TACGATGGATATTTTGGATAA (SEQ ID NO. 13)

and

AAAAAAGCAAATAATAAAAGAAAAAAGCTTTTATTACCCAATAAACA
AGAACAATATCCA (SEQ ID NO. 14)

[0074] The downstream region of *pflAB* was PCR amplified using primers:

TTATTTTGGATATTGTTCTTGTTTATTGGGTAATAAAAGCTTTTTTCTTT
TATTATTGTC (SEQ ID NO. 15)

and

TGGGTTATTGATACACCCATGTCAGAATACTGGCCGGCCTCGCGATTTT
GTCATCATAATTTATATTGGAGC (SEQ ID NO. 16)

[0075] The internal fragment of pflAB was PCR amplified using primers:
 AAACAAAAGGCTCAGTCGGAAGACTGGGCCTTTTGT TTTTGGTACCCTG
 GTTACAAAATCTTCGTTTCAGG (SEQ ID NO. 17)

and

TGGGTTATTGATACACCCATGTCAGAATACTGGCCGGCCTCGCGATTTT
 GTCATCATAATTTATATTGGAGC (SEQ ID NO. 18)

[0076] *C. thermocellum* Δhpt strain was obtained from Mascoma Corporation and was transformed via electroporation as described (WO 2010/056450 ELECTRO-TRANSFORMATION OF GRAM-POSITIVE, ANAEROBIC, THERMOPHILIC BACTERIA and Tyurin, M. V., S. G. Desai, and L. R. Lynd. 2004. Electrotransformation of *Clostridium thermocellum*. Appl Environ Microbiol 70:883-90) with modifications. Briefly, 400 ml *C. thermocellum* Δhpt strain was grown inside a Coy anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) in modified DSM122 media supplemented with 50 mM MOPS and 10 mM sodium citrate at 51°C. to an OD between 0.8 and 1.0, centrifuged without measures to exclude oxygen at room temperature in a Beckman Coulter Avanti J-25 centrifuge with a JA-10 rotor at 5000 x g, and the supernatant was removed. Being careful to minimize disturbance, cell pellets were washed with 400 ml ice cold electroporation buffer prepared without measures to exclude oxygen and consisting of 250 mM sucrose, 10% glycerol, 100 μ M MOPS pH 7.0, 0.5 mM $MgCl_2$, 0.5 mM $MgSO_4$ and centrifuged at 4000 x g. The cells were rinsed and centrifuged a second time as above and brought on ice into a Coy anaerobic chamber. Cells were resuspended in an additional 500 μ l electroporation buffer and kept on ice until use. Plasmid DNA was diluted to 500 ng/ μ l, and 2 μ l DNA was mixed with 20 μ l washed cells in pre-chilled 1 mm gap electroporation cuvettes. The mixture was then subjected to a 1.2 kV, 1.5 msec square pulse using a BioRad GenePulser XCell. Cells were immediately resuspended in 1 ml room temperature growth medium and serial dilutions were plated with no recovery period (to ensure each colony represents a unique transformant) by mixing with 25 ml molten media + 0.8% agar + thiamphenicol (10 μ g/ml). Once plates had solidified, they were placed in 2.5 L AnaeroPack Rectangular Jars (bioMerieux, Durham, NC, USA) to minimize desiccation and incubated at 51°C for up to one week.

[0077] To select for a merodiploid that contains the *cat-hpt* genes integrated into the chromosome with concomitant loss of the plasmid, *C. thermocellum* Δhpt strains

containing plasmid pAMG281 or pAMG278 were grown in 5 ml liquid medium supplemented with thiamphenicol. Serial dilutions were plated via pour plating in medium supplemented with 0.8% agar, 10 µg/ml thiamphenicol, and 10 µg/ml 5-fluoro-2'-deoxuracil (FUdR) to select for the simultaneous integration of the *cat-hpt* genes at the *hydG* or *pflAB* locus and the loss of the plasmid. Thiamphenicol-resistant, FUdR-resistant colonies were purified via streak plating and isolated colonies were grown in 5 ml liquid medium supplemented with thiamphenicol. These cultures were diluted 1:100 in medium lacking thiamphenicol and serial dilutions were plated via pour plating in medium supplemented with 0.8% agar and 500 µg/ml 8-azahypoxanthine (8AZH) to select against the *hpt* gene, thus selecting for resolution of the merodiploid state and deletion of the *hydG* or *pflAB* genes. 8AZH-resistant colonies were purified via streak plating and isolated colonies were grown in 5 ml liquid medium without added antibiotics and were screened by PCR for the correct genomic structure.

EXAMPLE 2

Verification of gene disruption/deletion in the mutant *C. thermocellum* strains

[0078] Colony PCR was performed to screen for the correct genomic structure of putative deletion mutants (Fig. 6). The gel was loaded with primer sets grouped together and ordered from 1 – 5, flanked by 2-Log DNA ladder from NEB. For each primer pair, templates were, in order, (1) No template negative control, (2) *C. thermocellum* Δhpt parent strain (wild type at *hydG* locus) (3)-(5) *C. thermocellum* Δhpt $\Delta hydG$ clones #1-3.

[0079] Primer pair #1 targets an internal region of the *hydG* gene that should generate a ca. 1000 base pair band in strains wild type at this locus and no product in strains deleted for *hydG*. The PCR results using Primer pair #1 are shown in Fig. 6A, lanes 2-6. Lane 2: No template negative control; lane 3: *C. thermocellum* Δhpt parent strain (wild type at *hydG* locus); lanes 4-6: *C. thermocellum* Δhpt $\Delta hydG$ clones #1-3. The sequences of Primer pair #1 are:

GGCCCTCTAGGATCAGCGGGTTTAAACGCTGAGGCGCGCCCCCGGTAC
GACATCTTCTGT (SEQ ID NO. 19)

and

AAAGGCTCAGTCGGAAGACTGGGCCTTTTGT TTTTGGTACCCCTGGGAG

AAGTTGCAAAGC (SEQ ID NO. 20)

[0080] Primer pair #2 targets a ca. 700 base pair fragment of the plasmid backbone and should be absent in all strains. The PCR results using Primer pair #2 are shown in Fig. 6A, lanes 8-12. Lane 8: No template negative control; lane 9: *C. thermocellum* Δ hpt parent strain (wild type at *hydG* locus); lanes 10-12: *C. thermocellum* Δ hpt Δ hydG clones #1-3. Primer pair #2:

TTTCGGTCGAATCATTTGAAC (SEQ ID NO. 21)

and

GGGTTTTAGTGGACAAGACAAAA (SEQ ID NO. 22)

[0081] Primer pair #3 targets the upstream junction of the deleted region and should produce a ca. 2700 base pair band in the wild type and a ca. 1200 base pair band in the *hydG* deletion mutant. The PCR results using Primer pair #3 are shown in Fig. 6A, lanes 14-18. Lane 14: No template negative control; lane 15: *C. thermocellum* Δ hpt parent strain (wild type at *hydG* locus); lanes 16-18: *C. thermocellum* Δ hpt Δ hydG clones #1-3. Primer pair #3:

GCCGCATTGTCAAGATAAATC (SEQ ID NO. 23)

and

GCTGTAAGTCTTCGGTGAGAGTT (SEQ ID NO. 24)

[0082] Primer pair #4 targets the downstream junction of the deleted region and should produce a ca. 2900 base pair band in the wild type and a ca. 1300 base pair band in the *hydG* deletion mutant. The PCR results using Primer pair #4 are shown in Fig. 6B, lanes 2-6. Fig. 6B, lane 2: No template negative control; lane 3: *C. thermocellum* Δ hpt parent strain (wild type at *hydG* locus); lanes 4-6: *C. thermocellum* Δ hpt Δ hydG clones #1-3. Primer pair #4 sequences are:

TTTACCTGACGTGACGGATTG (SEQ ID NO. 25)

and

CTCACTTTTGTAGAATCCACACCT (SEQ ID NO. 26)

[0083] Primer pair #5 flank the upstream and downstream regions, and should result in a band ca. 3600 base pairs in the wild type and a ca. 2200 base pair band in the *hydG* deletion mutant. The PCR results using Primer pair #5 are shown in Fig. 6B, lanes 10-14. Fig. 6B, lane 10: No template negative control; lane 11: *C. thermocellum* Δ hpt

parent strain (wild type at *hydG* locus); lanes 12-14: *C. thermocellum* Δhpt $\Delta hydG$ clones #1-3. Primer pair #5:

GCTGTAAGTCTTCGGTGAGAGTT (SEQ ID NO. 27)

and

CTCACTTTTGTAGAATCCACACCT (SEQ ID NO. 28)

[0084] Taken together, the results shown in Fig. 6 confirmed that *hydG* was deleted in the mutant *C. thermocellum* isolates #1-3 as disclosed herein.

EXAMPLE 3

Fermentation Profiles of the parental Δhpt *C. thermocellum* strains and the *hydG* or *pfl* knockout Strains

[0085] To determine the fermentation products formed by these deletion mutants, *C. thermocellum* Δhpt Δpfl , *C. thermocellum* Δhpt $\Delta hydG$, and the parent strain *C. thermocellum* Δhpt were grown anaerobically in Balch tubes containing 10 ml medium with a ca. 17 ml headspace. After growth was complete, 1 ml samples were removed and centrifuged at 15000 x g to remove cells. Then, 700 μ l of the supernatant was combined with 100 μ l 10% sulfuric acid and filtered through a 0.22 μ m filter. Ethanol, acetic acid, lactic acid, and formic acid were then measured by HPLC using a Bio-Rad Aminex HPX-87H column with RI detection on a Waters HPLC system at 60°C, as shown in Fig. 7. The parent strain *C. thermocellum* Δhpt is shown as Dhpt in Fig. 7, the mutant strain Δhpt Δpfl is indicated as Dpfl, while the mutant Δhpt $\Delta hydG$ is shown as DhydG in Fig. 7.

EXAMPLE 4

Generation of strains with multiple genes deleted/disrupted

[0086] Additional mutants carrying mutations in more than one genes may be created by genetic engineering. These genes may include, for example, (a) hydrogenase genes, (b) genes encoding one or more maturation proteins that modify hydrogenase post-translationally, such as HydE, HydF and HydG, (c) any other genes encoding proteins that modulate the activities of hydrogenase, (d) pyruvate-formate lyase (*pfl*) genes, such as *pflA* and *pflB*, (e) acetate kinase (*ack*) gene, (f) phosphotransacetylase (*pta*) gene, and (g) lactate dehydrogenase (*ldh*) gene, among others. Double mutant can be generated wherein two genes that are known to play non-redundant roles in the fermentation

pathways are disrupted. For instance, the pfl gene may be disrupted in a mutant strain carrying a deletion to one or more of the hydrogenase maturase genes, such as HydE, HydF or HydG. Conversely, one or more of the hydrogenase maturase genes may be deleted/disrupted in a mutant strain already carrying a pfl gene deletion.

[0087] Mutations in three or more genes may also be introduced into a strain. For instance, the ldh gene may be disrupted in a double mutant strain which already has both the pfl gene and the hydG gene deleted. As indicated by the results shown in Fig. 7, such a triple mutant may produce much higher levels of ethanol as compared to the pfl-hydG double mutant because the blocking of lactate production may direct the more carbon flow to ethanol. Similarly, other multiple disruption mutants may be generated to maximize ethanol production.

[0088] The description of the specific embodiments reveals general concepts that others can modify and/or adapt for various applications or uses that do not depart from the general concepts. Therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not limitation.

[0089] All references mentioned in this application are incorporated by reference to the same extent as though fully replicated herein.

CLAIMS

WHAT IS CLAIMED IS:

1. An organism capable of fermenting a saccharification product of a substrate derived from a biomass, wherein a first gene endogenous to said organism has been inactivated by genetic engineering, said first gene encoding a post-translational modifying protein, wherein the inactivation of said first gene results in a decrease in hydrogenase activities in said organism.
2. The organism of claim 1, wherein said post-translational modifying protein is capable of facilitating maturation of at least one hydrogenase synthesized by said organism in its native state.
3. The organism of any one of the preceding claims, wherein said first gene has at least 90% sequence identity with a polynucleotide sequence selected from the group consisting of SEQ ID Nos. 1-3.
4. The organism of any one of the preceding claims, wherein the organism is a bacterium.
5. The organism of any one of the preceding claims, wherein the organism is a thermophilic, anaerobic bacterium.
6. The organism of any one of the preceding claims, wherein the bacterium is *Clostridium thermocellum*.
7. The organism of any one of the preceding claims, wherein a second gene encoding a protein other than a post-translational modifying protein is inactivated.
8. The organism of any one of the preceding claims, wherein the second gene encodes a protein that is required by the organism to produce lactic acid as a fermentation product.
9. The organism of any one of the preceding claims, wherein the second gene encodes a lactate dehydrogenase (*ldh*).

10. The organism of any one of the preceding claims, wherein the second gene encodes a protein that is required by the organism to produce acetic acid as a fermentation product.

11. The organism of any one of the preceding claims, wherein the second gene encodes a protein selected from the group consisting of acetate kinase (*ack*) and phosphotransacetylase (*pta*).

12. A bacterium capable of fermenting a saccharification product of a carbohydrate-rich biomass substrate, wherein the *hydG* and *ldh* genes are inactivated by genetic engineering.

13. A *Clostridium thermocellum* strain deposited under Patent Deposit Designation No. PTA-11763 or PTA-11764.

14. A method for producing ethanol, said method comprising:
creating an organism with at least one gene inactivated, said at least one gene encoding a post-translational modifying protein, wherein said post-translational modifying protein is capable of facilitating maturation of at least one hydrogenase synthesized by said organism in its native state; and incubating said organism in a medium containing at least one substrate selected from the group consisting of glucose, xylose, mannose, arabinose, galactose, fructose, cellobiose, sucrose, maltose, xylan, mannan, starch, cellulose, pectin and combinations thereof to allow for production of ethanol from the substrate.

15. The method of claim 14, wherein the organism is a member of the *Clostridium* genus.

16. The method of any one of the preceding claims, wherein the organism is *Clostridium thermocellum*.

17. A method for producing ethanol, said method comprising:
providing within a reaction vessel, a reaction mixture comprising a carbohydrate-rich biomass substrate, and a fermentation agent, the fermentation agent comprising a bacterium that has been genetically modified to inactivate a

first gene endogenous to said bacterium, said first gene encoding a post-translational modifying protein,

wherein the reaction mixture is incubated under suitable conditions for a period of time sufficient to allow saccharification and fermentation of the carbohydrate-rich biomass substrate.

18. The method of claim 17, wherein the suitable conditions comprise a temperature of at least 50°C.

19. The method of any one of the preceding claims, wherein the bacterium is a member of the *Clostridium* genus.

20. The method of any one of the preceding claims, wherein the bacterium is a *Clostridium thermocellum*.

21. The method of any one of the preceding claims, wherein the reaction mixture further comprises a cellulolytic material.

22. The method of any one of the preceding claims, wherein the cellulolytic material comprises a cellulase.

23. The method of any one of the preceding claims, wherein said first gene has at least 90% sequence identity with SEQ ID NO. 1-3.

24. The method of any one of the preceding claims, wherein one or more genes in addition to the first gene is inactivated in the bacterium, said one or more genes encoding one or more proteins selected from the group consisting of lactate dehydrogenase (*ldh*), acetate kinase (*ack*), phosphotransacetylase (*pta*) and combination thereof.

25. The method of any one of the preceding claims, wherein the organism is a *Clostridium thermocellum* strain deposited under Patent Deposit Designation No. PTA-11763 or PTA-11764.

26. An organism capable of fermenting a saccharification product of a substrate derived from a biomass, wherein a first gene endogenous to said organism has been inactivated by genetic engineering, said first gene encoding a protein required for

the generation of formate, wherein the inactivation of said first gene results in a decrease in formate production in said organism.

27. The organism of any one of the preceding claims, wherein said first gene has at least 90% sequence identity with a polynucleotide sequence selected from the group consisting of SEQ ID Nos. 4-5.

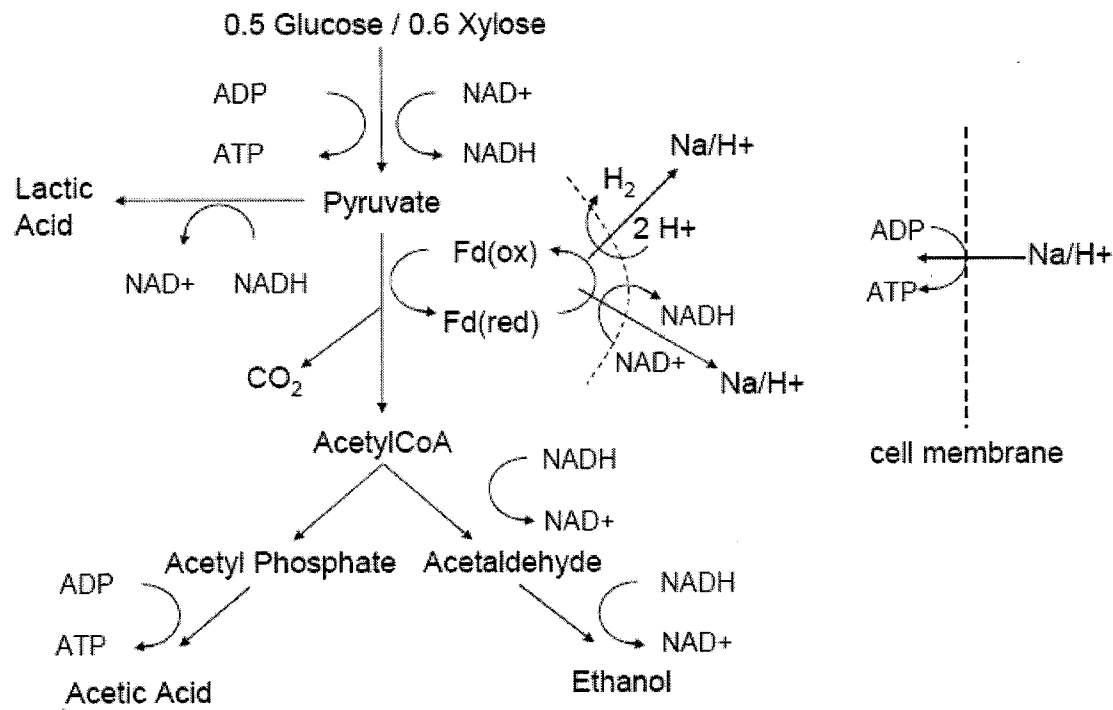
28. An organism capable of fermenting a saccharification product of a substrate derived from a biomass, wherein at least a first gene and a second gene have been inactivated by genetic engineering, said first gene encoding a protein required for the generation of formate, said second gene encoding a post-translational modifying protein capable of facilitating maturation of hydrogenase in said organism.

29. The organism of any one of the preceding claims, wherein a third gene is inactivated, said third gene encoding a protein required for the generation of lactate in said organism.

30. The organism of any one of the preceding claims, wherein said first gene is the *pfl* gene, said third gene is the *ldh* gene, and said second gene encode a post-translational modifying protein selected from the group consisting of HydE, HydF and HydG proteins.

31. The organism of any one of the preceding claims, wherein the acetate kinase (*ack*) and the phosphotransacetylase (*pta*) genes have been inactivated.

FIG. 1



2/11

FIG. 2

SEQ ID NO: 1

atgacaaacatgacaaatatgataaatttaattgacaaactttcaacaacacatactttgtcgtatgatgaaatgtatcagcttattgag
catagaaacgaggaacttgccaattatctgtttgaaaaggcaaggcaggtgcgtatcctctactacggccacgatgtctatatgcg
cgggtctcattgaattcaccattactgtcgaaatgactgtctactattgcggaataaggaaaagcaactgcaatgccgaaagatacc
gtcttacaaaagagcaaaacttgaatgctgtgacgtgggatatgagctgggttttcgcaccttcgtgcttcaagggggcgaagac
ggttattataccgacaaaattttggcggacatagtaagcagcatcaaggcaaaatatcccattgtgcgattactctctctttgggtg
aaaaaagttagaaagctataaattgctttatgaggctggagcggacagataccttctcgcctatgaacagcaaatgccagcac
tactcaaagcttcatccgctgttatgtcccttaaaaacagaaaacaatgtctttacaatctcaaagaaataggataccaggtaggtt
gcggttttatggtcggttcaccgtttcagaccacggaatgtctcgttgatgacttaatgtttataaaagaattgcagccccacatggtg
ggaatagggtccgtttatcccgacaaaggatacgcttttgccggcaaacccgccgtaccctggagctgacattgttcttctcgg
catcatacggctaattgcttccctacgttctgcttccggccaccacagcccttggcacaatccatcccaaggcagggaactgggt
attcttgcaggcgcaaacgtggtaatgccaaacctttgccgaaagaagtaagaagcaagtatcttttatgacaataaaatctgt
accgggggatgaagccgcagaatgcagaatgtgcctaaccaccgtattgaaagcatcggatacaaaactggttgtgtcaagagg
cgactgcaaaaagccaaattaa

SEQ ID NO: 2

atgggattgaatgaacaccatctgcaaaccgtcttcatataggattcttcggcaaaagaaacgcaggtaaatcaagtgttgtaaat
gctgttacgggtcagaatctggctattgttcggacgttaaagggaaccacagaccctgtttacaaagcaatggagcttttacctt
tagggccggttgatgataattgacactcccgaattgatgacgaggggaacattgggagaaatgcgtgtgaaaaggagccgccag
gtacttaacaaaaccgatattgcagtattggtaatcgatgccacgtgcggcaagtctgaagatgatgaaaaattgattgaattgttg
agaagaaggacataaagtatgttgttttacaacaaggcggatttgaaggacatgaagaaacggtgggcgataatgaaatttat
gtcagtgccaaaacgggttacaacattaataaattgaaggaaaaaatcgcatctcttctgttaccgatgatattaccacaaaata
gtgggtgacctgattagcttccgattttgtcgtacttgcgtacctattgacaaggcagctcccaaaggacgttatttgcctca
gcagcagaccattcgtgatattttggaatcggtgcggttgcgattgttggtaaaagagaacgaactaaaaatacgttagacagcc
ttggcaaaaagcctaagctgtcattaccgacagtcaggcctttgaaaaagtggcggcgacacgcccgcagatatttacttactt
cgttttccattcttttgcacgatacaagggtaatctggaaattgcggtaaaaggagcaaagacgctcgattctttacaagacggtga
taccgttctatctccgaagggtgtacacatcacagacagtgcgatgatataggcacgggtgaagcttcaaggtggataaataatt
acactaaaaagaatttaaaacttgaatttacgtccggtacggagttccggaggatttgacccgctataaattaattgtgcattgtggc
ggctgcatgctgaatgaaagggaatgaaataccgttacaatgtgcggtggaacagaatgttccattaccaattacggaatttt
gattgcataatgtgcatggaattttaaaaagaagccttcaaatcttcccgacatacttcagaaaattcttag

3/11

SEQ ID NO: 3

atggttgaaaaagttgattttataaaagaagatttgatttttctcttcttgaaaaaggtaaaactactgacaggaatgaaataagagaa
atattggcaaaagccagggagtgcaagggcataagcctgggagaagttgcaaagctgctttacctggaagacgaagagctgtt
ggaagaactctatgatgtggccaaatatataaaaaacaaaatatacggaaaaagagtggtttgttctcctttatatacaagcaat
gagtgtacaacaactgcctttactgcggttcaggcatgacaacaagagcttcacagaaagactttaagtcttgaagaaatagt
ggaagaagcaaaagctattgaaagacagggacataaaagattgctccttatttgcggagaagatcccagaaagactaatgtaaa
gcattttaccgatgcaatggaggcaatatataatccaccgatatcagaagaatcaatgtcgaaggccgcgcgatgacgggtgga
agattacagggagctgaagaaagccggtataggaacttatgtgatattccaggagacatatcacagggaaacctatagaataatg
catcctgtgggcaaaaaagcgaattatgactggcgtataacggcaatagacagagcccttgaaggcgggtattgacgatgtgggtg
tcggagcgtcttcggactttatgattatagattcgaagtttaggcctttgatgcattgtatgcatttgaagaaaaatatggtgtag
ggccgcacaccatactgttccgagacttcgtctgccttgggagctccttgaagagattccgtacaaagtaccgacaaggatt
tcaagaaaattgtggccatattccgaattgctgttccatatacgggaattattcttcaacaagagaaagagctgaattcagggatga
acttttaagtgttgagtatctcagataagtgacaggttccaaaaccaatcccggaggataccaggaggtatgacgacctgcggat
cagtttgaataagtgaacaacaggagtttgcgaaagtaatggaacaatatgtcaacaagggtatattccaagcttttgcaccgctt
gctacagaagatgtcgtaccggggagcatttcatggaatatgcaaaggccggggatatacatgaattctgtcagccgaatgccat
tctaactttcaaggaaaaacttaattgattatgccgatgagcctttgagaaaaatgggtgaggaagtaataacttaaggctctggaaga
aattgaagatgaaaaatgaagactcttactattgcaaaacttgaggaaattgaaaaaggaaagagagatatatttttttaa

SEQ ID NO: 4

atggatgcatggcgcggatttaataaaaggcaactgggtgccaggaaattgacgttcgtgattttataattagaattatactccttatga
aggcgtatgaaagctttctttagtagacctacggatagaacgcggaaactttgggagaaggtttccgaactgttaagaaagaacg
ggagaacggcgggggtattggatgttgataccatacaattcaacgattacgtctcataaaacctggatatatagataaaagaactga
agttattgtcgggcttcagacggatgagccttataaaagagccataatgccgtttggcgggtatagctatggtgattaaggagccg
aagcttatggccacagtgtggacctcaggttgtgaaatattcacaaagtacagaaagactcataaccagggagtttatgatgtat
atactcccgaatgagaaaaagccaaaaagccgggattattacaggacttcccgcacatagcgcagaggaagaataattggc
gattacagaagggttgcactttatggcgttgacaggtgattgtgaaaaagagaaagaaatggcaagtcttgaagagattacat
tgactatgagactgttcgagacagagaagaaataagcgagcagattaaactttaaaacaacttaagaaatggctttaaagttacg
gttttgacatatcttgcctgcaaaggatgccagagaagccttcaatgggtgtattttgcatatcttgcagcagtcagggaacagaa
cggcgcggcaatgagtattggaagaatttcactttccttgacatatacattgaaagggtatctcaaagaaggaaaactcacggag
gagttggctcaggaactgggtgaccagctggttataaagctgagaattgtgagattttgagaactcctgagtatgaaaagctcttca
gcggagacccacttgggttaaccgaaagtatcgaggtatggcgctggatggaagaacgctggttacaaaatcttcgttcaggtt
tttgacactcttttaacctgggacatgcaccggagcccaaccttacagtactttggtccgtcaatcttcccgaaggcttataaaag
tactgtgcaaaggtatcaattcattcaagctccatccagtatgaaagcgacgacataatgaggaaactggggagacgattatg
gaatagcatgctgtgtttctgctatgagaattggaacacagatgcagttcttcgggtgcaagatgcaatcttgcaaaagctcttcttac
gctattaacggcggaaaggatgaaatgacgggagaaacagattgtccgatgtttgcaccgggtggaaccgaataaccttgattac
gaggacgtaataagagggttgacatgggtgcttgactgggtggcaaggctttatatgaacacctcaatataattcactacatgcat
gacaaatatgccatgaggcgtgcagatggcattgcatgacaaagacgtgttcaggacgatggcatgcggaatagccgggttgt
ctgtgggtggcagactcccttagcgcgataaaatatgcaaagggttaaaccgatacgcgaatgaaaacaacctcgttgtgactacgaa
gttgagggtgattatcctaaattcgaaataacgacgaacgtgttgatgaaattgcagtgaagtagtaaaaatgttcatgaacaag
cttagaaagcaaagggttacagaagtccactccgacctttccatacttaccataacttcaaactggttatatggaagaaaac
cggaaacactcctgacggcagaaaagctggagaacctttggcgccgggagcaaatccgatgcatggaagggtatataaacgga
gcattggctgtactgaacagatttgcgaagcttccctatgaatatgccaggacggcatttcatatactttccataattccaaaagc
tctgggaagagacgaggaaccagaataaacaatcttaaatcaatgcttgacggatatttcaagcagggcggccaccacataaa
tgtaaatgtgttgaaaaagagacactgttagatgccatggaacatccggaaaaatatccacaacttaccataagagtgtccgggt
atgcagtgaactttataaagcttacacgggagcaacagctggatgttattaacagaacgattcacggaaagatttaa

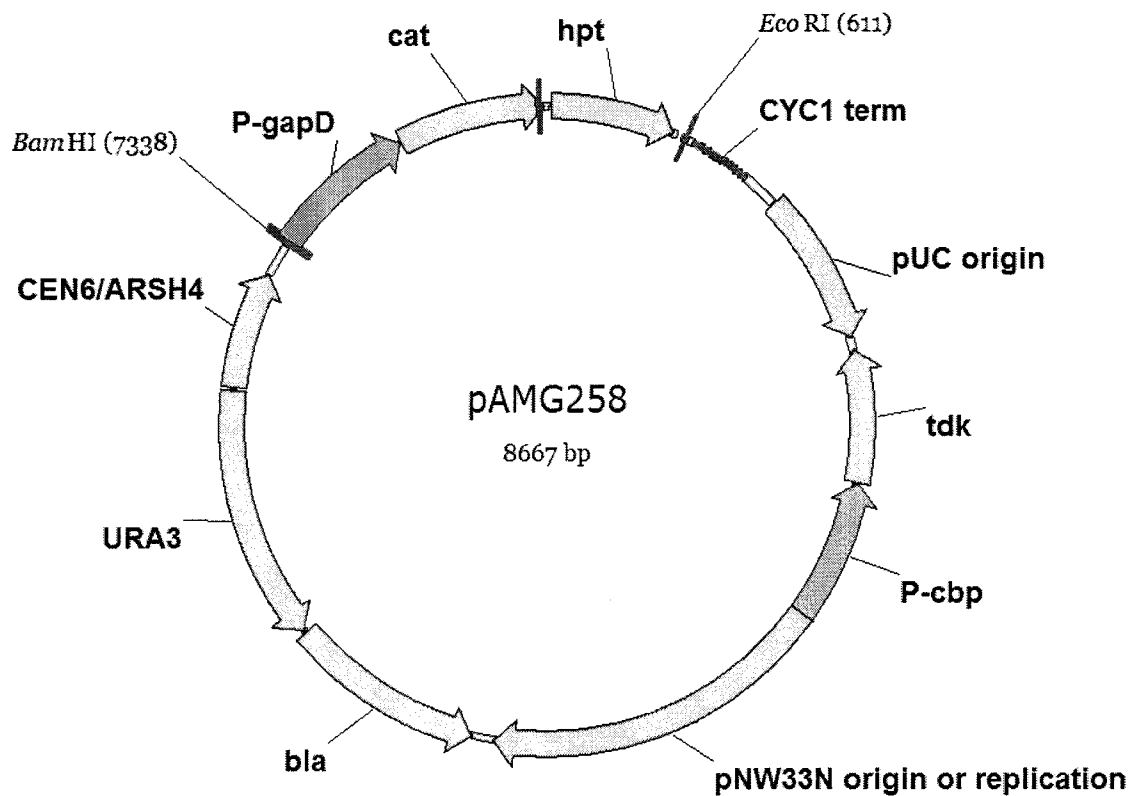
4/11

SEQ ID NO: 5

atgacattaaagggcaggatacactcattgaatcttttgggacactggacggaccgggtataagatttgggtttcatgcagggc
tgtcccttgcggtgtatatattgccacaacagggaacacgtgggatgttaatgcggggagtgagtacactccccggcaagtaattgat
gaaatgatgaaatacatagactatataaaggtctccggaggcgggaataactgttacggcggggagcctgttctccaggccgatt
ttgtggccgaggtgttcagactgcaaaagagcaggagtgcatacggcgtggataccaatggatttgctgacatagagaagg
ttgaaaggcttataaaatacaccgatcttgattgctggatataaagcatgcccgaggagataaacataagataattaccggtgtgt
ccaacgaaaaaatcaagcggtttgcgctgtatcttcggaccaggagtgccatctggataagatatgccttgccccggatata
ccgacgatgaagatgaccttaaaatggcggctgatttcataaaaaagcttaaacgggtggaaaaatcgaagttctccttatcaca
acatgggagcatacaaatgggaaaaacttggtcagaaatacatgctgaaggagtaaaggggccgagtgcgcaagaggtgga
aaaagcaaagaggattctgtcaggcaaataa

5/11

FIG. 3



SEQ ID. No. 6

gtcatggaaaatttatcaaaagacatc gatgaaattttgatcacagaagaagaacttaaggaaaagataaaagagcttgggagggc
 aaatcacaaaagactacaaagggaaaaatttgatgttgtaggaggttttaaaagggtgcttaattgtttatggctgattgtcaagaca
 catagatttgcctttatcacttgattttatggctgtttccagctatggaagctcaactcattcatcaggaatagtaaagataatcaaagat
 cttgatataagcatagaaggcaaagatgttctgattgtggaagacataattgacagcgggttgactttgtcttacttaagggaaacttt
 acttgggaagggaagccaaaaagcctgaaaatagcacaatattagacaaaccggagagaagagaagcatctgtaaaagtcgatta
 ttaggatttaagatacctgataagttgtcgtgggttatggattggactttgatgaaaagtacaggaaccttccctttatagggcgtttg
 aaacctgaaatgtacagctaagcccgggcctcgagaaaacaaaaggctcagtcggaagactgggcctttgtttggtaccgaat
 tcggcgcgcctcagcgtttaaacccgctgacctagaggccgcacatcatgtaattagttatgtcacgcttacattcacgccctccc
 ccacatccgctctaaccgaaaaggaaggagtagacaacctgaagtctaggtccctatttttttatagttatgtagtattaagaa
 cgttatttatatttcaaattttctttttctgtacagacgcgtgacgcattgtaacattatactgaaaaccttgcttgagaagggtttggg
 acgctcgaaggctttaatttgcaagctgcggccctgcattaatgaatcggccaacgcgcggggagaggcggttgcgtattggg
 cgctcttcgcttccctcgtcactgactcgtcgcctcggtcgttcggctgcggcgagcggtatcagctcactcaaaaggcggtaat
 acggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaagcccggaacccgtaaaaa
 ggccgcgttgctggcgttttccataggtccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa
 cccgacaggactataaagataaccaggcgtttcccccctggaagctccctcgtgcgtctcctgttccgacctgccgcttaccggat
 acctgtccgcttttcccttcgggaagcgtggcgttttcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgtc
 caagctgggctgtgtgcacgaacccccgttcagccgaccgctgcgccttatccggtaactatcgtcttgagccaacccggtgta

6/11

agacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcgggtctacagagtcttg
aagtgggtggcctaactacggctacactagaaggacagtatttggtagctgcgctctgctgaagccagttaccttcggaaaaagagt
tggtagctcttgatccggcaaaacaaccaccgctggtagcgggtggtttttgttgaagcagcagattacgcgcagaaaaaag
gatcacaagaagatcctttgatctttctacggggtctgacgctcagtggaacgaaaactcacgttaagggatttggctatgaaaaa
atccgcttaagtccgcgggacttcatggcactttctacaccttgcttcataagactcctttgctccaaccattataatgggagtcgctg
aattggcaggcttcccatttataagccttgcgctgcgagtagcaggattgccacacttcatgcatatagcagtaagcttatcgacaa
attcagctatggccatcaattctggagtcggaccacaaatggttcacctctaaagtccatgtcaagccctgcacatatgactctttccc
gctatcggaatctctttacgatgtcgactatttcagaatcaaaaaattgaacttcatctatagcaataacatccgtatcttcttcagca
taagccaatatgtcagaagccttactatggcaatggcgtgcatgttgcgccgttatgagatacgactttgtctatggaatacctatc
gtctatagccggttgaaaaacctgaacttttgcctggcaatcttagctctctttatccttcttataagttcctcacttttccactgaacat
gggaccagttacaacttctatgtacccgtggtctttagggccatacatcttctcctccttaaaatttctgtttatttggttttatttaggt
gcctaaaagccactcttatacttgcacttatactcttcttctgtatctcttacctcatatctcttattttaactaccgaagtattttt
cataacggaatcttatgacttatttttgcactgccgtttaccatacggtaatttatacggaaacagctaacttaaccttgcgaagggc
atgtttacaaaataattatctactttaccatattaattcaagtggctattttagattttacgcaaaatttttagcgttcttacttgata
atttatccaaaaaccttcatcgtcaaaaacaattatgatataaattatgatgatgattaattaattttacggataactaatcatttcaat
gcaatagatcaacatgcaaatcaatatccacctactccgaacattttgcaccaacgaatgccagcactatatcccggaagcaa
tataacgtctataatttttatgataacgataaaattaaattttatgaaggggtaaaagcatgagaaaaatagctgtattgttaattactt
tgacgttcttagtcacgactcgaatttaggaggttacttgcctgtcttctcattagaatcaatcctttttaaagtcattcccgttgtt
gaactactctttaataaaataattttccgttcccaattccacattgcaataatagaaaatccatcttcacggttttctgcacatctgt
atgaatcaaatcgccttcttctgtgtcatcaagggttaatttttatgtattctttaaacaaccaccataggagattaaccttttacggtg
taaaccttctccaaatcagacaaacgtttcaaatcttttctcatcagcgtcataaaatccgtatcctttacaggatattttgcagttt
cgtcaattgccgattgtatatccgatttatatttttccgtcgaatcattgaacttttacatttggatcatagtctaatcttgcctttt
tccaaaattgaatccattgttttgattcacgtagttttctgtattctttaaataagttggttcacacataccaatacatgcatgtgctgat
tataagaattatcttattatttattgtcacttccgttgcacgcataaaaccaagaatttttataattttttatattgcacattcggcga
aatccttgagccatactgcacaaacttatttaattcttcgccatcataaacatttttaactgttaattgtgagaacaaccaacgaact
gttggcttttgttaataacttcagcaacaaccttttgcactgaatgccatgttccattgctctcctccagttgcacattggacaaagcc
tggatttacaaaaccacactcgatacaactttcttcgctgttccagatttggttatactctaatatttcagcacaacttttactcttc
agcctttttaattcaagaatatgcagaagttcaaaagtaatcaacattagcgattttcttctcctatggtctcacttttccacttttgc
ttgtccactaaaaccttgattttcatctgaataaatgctactattaggacacataatattaaaagaaacccccatctatttagttatttg
tttggtcacttataactttaacagatgggggttttctgtgcaaccaattttaagggtttcaactttaaaacacatacataccaacactt
caacgcacctttcagcaactaaaataaaaatgacgttatttctatgtatcaagaatagaaagaactcgttttctgctacgtcaaaa
cgcaaaaaaagcactcattcgagtgttttcttatcgtctccaaatcatcgatttttcttcttcttcttgcacgaagttctcga
tcacgttgcaaaacatcttgaagcgaaaaagttattcttcttcttccgatcgtcatgctgacgcagaaaagccctctaggcgca
taggaacaactcctaaatgcatgtgaggggtttctcgtccatgtgaacagtcgcatacgcaatattttgttcccatagcggccgcg
attatcaaaaaggatcttcacatagatccttttaattaaaaatgaagtttaaatcaatctaaagtatatagtaaaacttgggtgac
agttaccaatgcttaatcagttaggcacatatctcagcgatctgtctatttgcctcatcattagttgcctgactccccgtcgttagata
actacgatacgggagcgttaccatctggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatca
gcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgctccatccagttatattgttgc
gggaagctagagtaagtagttcgccagttatagtttgcgcaacgttgttggcattgctacaggcatcgtggtgtcactctcgtcgt
tggtatggctcattcagctccggttcccaacgatcaaggcgagttacatgatcccccattgtgtgcaaaaaagcggttagctcctt
cggtcctccgatcgttgcagaagtaagttggccgcagtggtatcactcatggttatggcagcactgcataattcttactgtcatgc
catccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatcgggcgaccgagttgctcttggcc
ggcgtcaatacgggataatagtgtatcacatagcagaactttaaagtgtcatcattggaaaacgttcttcggggcgaaaactctc
aaggatcttaccgctgttgagatccagttcgatgaaccactcgtgcaccaactgatcttcagcatcttttactttaccagcggtt
ctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaagggaataagggcgacacggaaatgtgaatactcatactcttc
cttttcaattgggtaataactgatataaattgaagctctaatttgtgagtttagtatacatgcatttacttataatcagtttttagttt
gctggccgcactcttcaaatatgcttccagcctgcttttctgaacgttcacctctaccttagcatcccttcccttgcataatagtc

tttcccaacaataataatgtcagatcctgtagagaccacatcatccacgggttctatactgttgaccaatgctgtccttctgcatcta
aaccacaccgggtgtcataatcaaccaatcgtaaccttcatctctccacccatgtctctttagcaataaaggcgataacaaaat
ctttgtcgtcttctcgcaatgtcaacagtacccttagtatattctccagtagatagggagcccttgcataaattctgtaacatcaaa
aggcctctaggttcccttgttacttcttctgccgctgtcttcaaacggctaacaatacctggggccaccacaccgtgtgcattcgtaat
gtctgccattctgttattctgtatacaccgcagagtactgcaatttgactgtattaccaatgtcagcaaatttctgtcttcgaagag
taaaaaattgtacttggcggataatgcccttagcgggcttaactgtgccctccatggaaaaatcagtcaagatatccacatgtgtttta
gtaaacaaaattttgggacctaatgttcaactaactccagtaattccttgggtgtacgaacatccaatgaagcacacaagttgtttgc
tttctgtgcatgatataaataagcttggcagcaacaggactaggatgagtagcagcacgttcccttatatgtagctttcgacatgattat
cttcgttctcgcaggttttgttctgtgcagttgggttaagaatactgggcaatttcatgtttcttcaacactacatatgcgtatatatac
caatctaagtctgtgctccttcttctgttcttcttctgttcggagattaccgaatcaaaaaatttcaaagaaaccgaaatcaaaaaa
aagaataaaaaaaaaaatgatgaattgaattgaaaagctagcttcatgatgggtccttttcatcacgtgctataaaaaataattataat
aatttttaataataataataaataaaaaatagaaaagtaaaaaaagaaataaagaaaaaatagttttgtttccgaagatgtaaaag
actctaggggggatcgccaacaaatactaccttttcttctgtcttctcgtctcaggtattaatgccgaattgttcatcttctgtgtga
gaagaccacacacgaaaatcctgtgattttacattttacttategttaatcgaatgtatatctatttaactgtcttttctgtctaataaatat
atatgtaaagtagcgttttgttgaattttttaacctttgtttatttttttcttcttccgtaactcttctaccttcttattactttctaaaa
ccaaatacaaaacataaaaaataaataaacacagagtaaatcccaaattattccatcattaaaagatacgaggcgcgtgtaaattac
aggcaagcgatccgtccgccggcggaacgtggcgagaaaagggaagggaagaaagcgaaaaggagcgggggctagggcggtg
ggaagtgtaggggtcacgctggggcgtaaccaccacaccgccgcgcttaattggggcgctacagggcgcggtgggggatgatcc
actagtaagcttggatcctcgcgaggccggccagtattctgacatgggtgtatcaataacccatgcgtttccgtattgtatcggaat
ggtttcggacagggcggtgggaatagacatgaaaagattttttcccgaggatgcattgataaagtatttctttccgaaagaga
ggaaaagattctaaagagttttggaatactgatgaatattgtgtgcagagtacaattctatggacaagaaaagaggccttgtcaaa
acttttctgtctgggaatgaggatggattttaaaaagctggatactttggaggacgaggtgttttcaggaaacaaacaggggcgc
gtctgttttctttatatgcaataattactgtatctctctggcattgccagggtttaataaagattaaaattattgactagaaataaaaaat
gtccataatattaatggacaaaaaaacaaaagaattacatcaaaagggaagataaaaaatactttgttaaaaaataattatttttatctaaa
ctattgaaaaatgaaaaataaataataaaaaatgaatcatagtgcagagatacttggcagaggatgaatattttactgcattcatgctt
atggcagctaataagggcattaaattaaattttaatttacaataggaggcgatattaatgaactttaataaaattgatttagacaattgg
aagagaaaagagatatttaactattttgaaccaacaaacgacttttagtataaccacagaaattgatattagtgttttataccgaaa
cataaaacaagaaggatataaattttacctgtcatttattttcttagtgacaagggtgataaactcaaatacagcttttagaactggta
caatagcgcagcgagagttagggtattgggataagtttagagccactttatacaattttgatgggtgatctaaaacattctctggatttg
gactcctgtaaaagaatgacttcaaagagttttatgatttatacctttctgatgtagagaaatataatggttcggggaaattgttcccaa
aacacctatacctgaaaatgcttttcttcttattattccatggacttcatcttactgggtttaacttaaatatcaataataatagtaattac
cttctaccattattacagcaggaaaaattcattaataaaggtaattcaatatatttaccgctatctttacagggtacatcattctgtttgtgat
ggttatcatgcaggattgtttatgaactctattcaggaattgtcagataggcctaatactgctgttttataataaaggagggtcgac

FIG. 4

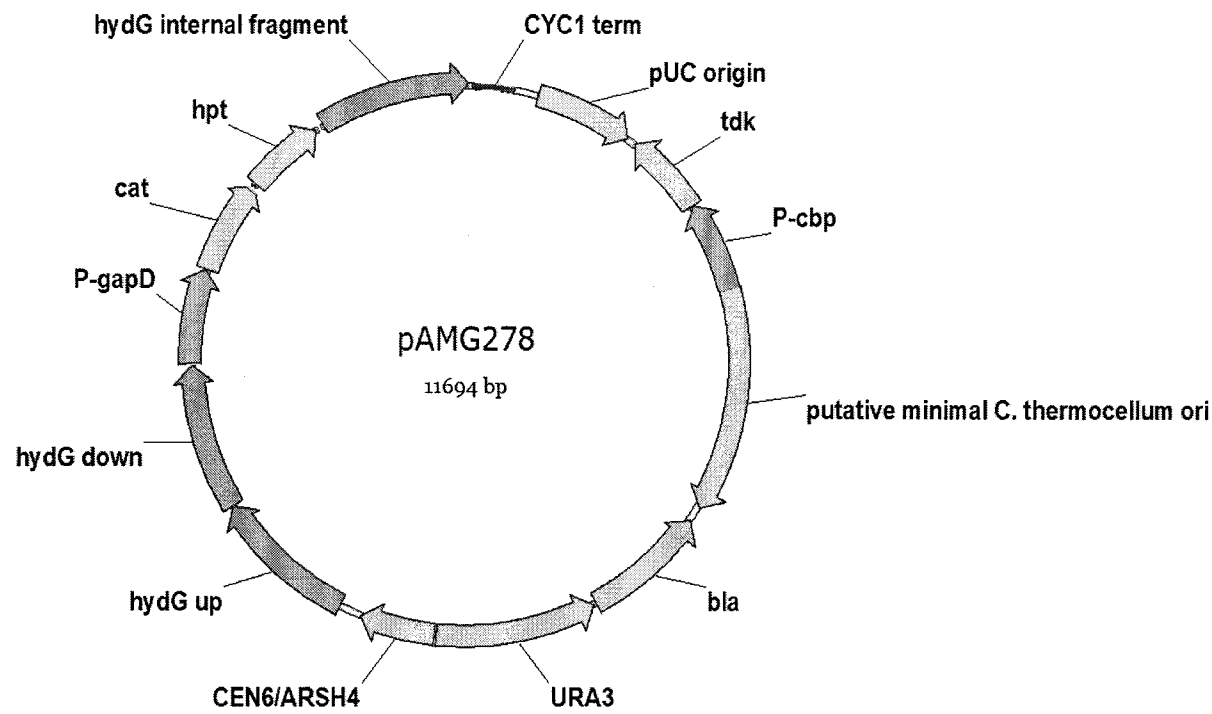
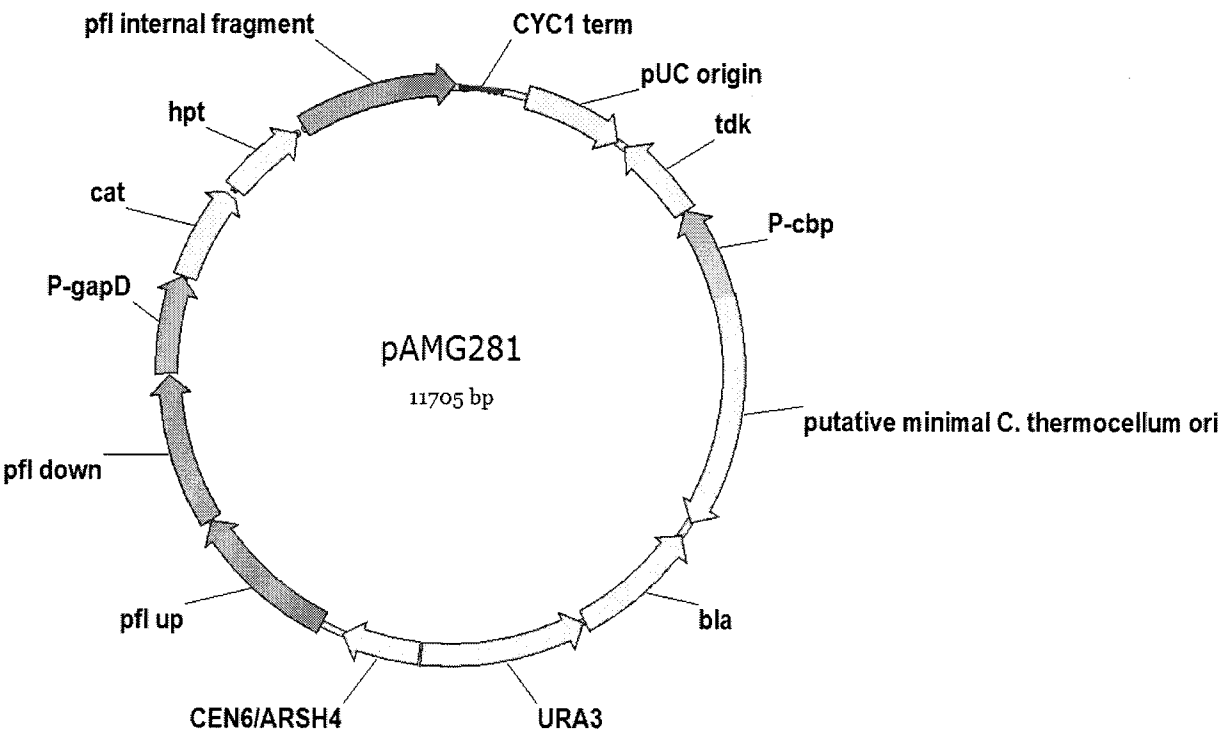


FIG. 5



10/11

FIG. 6A

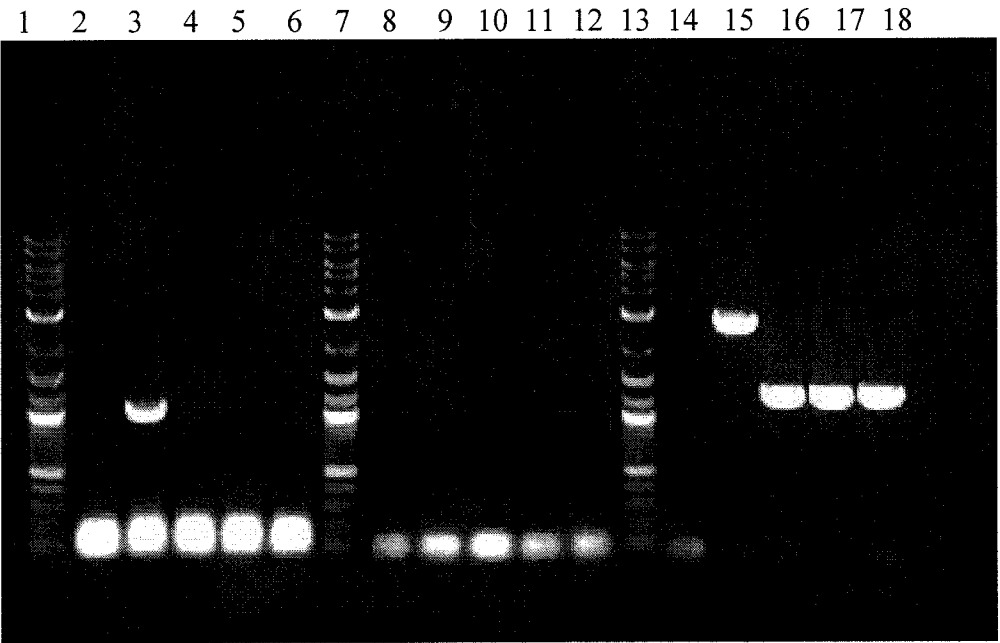
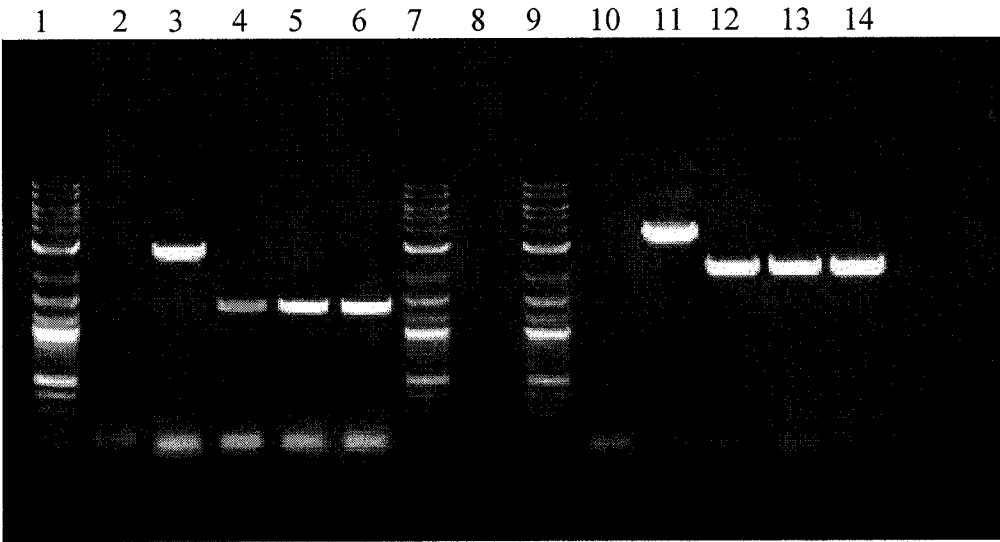


FIG. 6B



11/11

FIG. 7

End-Product profile in mutant and parent strains

