The present invention relates to microorganisms genetically engineered to increase yield and/or efficiency of biomass production from a carbon source, such as e.g. glucose. Processes for generating such microorganisms are also provided by the present invention. The invention also relates to polynucleotide sequences comprising genes that encode proteins that are involved in the bioconversion of a carbon source such as e.g. glucose into biomass. The invention also features polynucleotides comprising the full-length polynucleotide sequences of the novel genes and fragments thereof, the novel polypeptides encoded by the polynucleotides and fragments thereof, as well as their functional equivalents. Also included are processes of using the polynucleotides and modified polynucleotide sequences to transform host microorganisms leading to a microorganism with reduced carbon source diversion, i.e. higher yield and/or efficiency of biomass production from a carbon source such as e.g. glucose.
The present invention relates to microorganisms genetically engineered to increase yield and/or efficiency of biomass production from a carbon source, such as e.g. glucose. Processes for generating such microorganisms are also provided by the present invention. The invention also relates to polynucleotide sequences comprising genes that encode proteins that are involved in the bioconversion of a carbon source such as e.g. glucose into biomass. The invention also features polynucleotides comprising the full-length polynucleotide sequences of the novel genes and fragments thereof, the novel polypeptides encoded by the polynucleotides and fragments thereof, as well as their functional equivalents. Also included are processes of using the polynucleotides and modified polynucleotide sequences to transform host microorganisms leading to a microorganism with a reduced carbon source diversion, i.e. higher yield and/or efficiency of biomass production from a carbon source such as e.g. glucose.

The bioconversion of a carbon source may involve many different metabolic routes, and involve several enzymatic steps to generate biomass, wherein the enzymes may be located in the cytosol, on the membrane or in the periplasmic space of a host cell. Furthermore, transporters may also play an important role in the efficient conversion of a carbon source into biomass.

For instance, in the case of acetol-plastic bacteria, which are obligate aerobe, gram-negative microorganisms belonging to the genus Acetobacter, Gluconobacter, and Gluconacetobacter, these microorganisms are able to oxidize D-glucose at the periplasmic membrane level to D-gluconate by means of a membrane-bound D-glucose dehydrogenase, and transport D-gluconate into the cytosol. D-gluconate can then be phosphorylated by glucokinase to D-gluconate-6-phosphate. In addition to that, it is believed that D-glucose can be directly transported into the cytosol and then converted into D-gluconate by means of a cytosolic NAD(P)-dependent D-glucose dehydrogenase. Furthermore, D-glucose can be transported into the cytosol and, in the cytosol, also be first phosphorylated to D-ribose-6-phosphate by glucokinase and then be dehydrogenated to D-gluconate-6-phosphate. D-gluconate-6-phosphate may enter the pentose phosphate pathway, being further metabolized to produce reducing power in the form of NAD(P)H and tricarboxylic compounds necessary for growth and maintenance.

Proteins, in particular enzymes, that are active in the metabolism of glucose are herein referred to as being involved in the Glucose Metabolism System. Such proteins are abbreviated herein as GMS proteins and function in the direct metabolism or bioconversion of a carbon source such as e.g. glucose into biomass.

One disadvantage of such bioconversion processes, however, is the diversion of carbon sources or intermediates throughout said bioconversion process such that for instance the next (enzymatic) step cannot be performed in an optimal way, leading to the loss of available carbon substrate material for conversion into biomass and resultant energy losses, in form of e.g. ATP or NADPH. In the case of for instance the bioconversion of glucose into biomass, this loss may be due to transporting the glucose or an intermediate thereof out of the cytosol or by using the glucose or the resulting intermediates as substrates for other pathways not leading to the production of biomass.

It is an object of the present invention to provide microorganisms which are engineered in such a way that the carbon source diversion throughout the bioconversion of carbon sources, such as e.g. glucose, is altered, e.g. via reduction of carbon source diversion, leading to higher production and/or yield of biomass produced from such carbon sources.

Surprisingly, it has now been found that GMS proteins or subunits of such proteins having activity towards or which are involved in the bioconversion of a carbon source such as e.g. glucose play an important role in the production of biomass.

In one embodiment, GMS proteins of the present invention are selected from oxidoreductases [EC 1], preferably oxidoreductases acting on the CH—OH group of donors [EC 1.1], more preferably oxidoreductases with NAD* or NADP* as acceptor [EC 1.1.1] and oxidoreductases with a quinone or similar compound as acceptor [EC 1.1.1.5], most preferably NADP-glucose dehydrogenase [EC 1.1.1.47] and PQQ-dependent glucose dehydrogenase [EC 1.1.5.2] or from transferases [EC 2], preferably transferases transferring phosphorus-containing groups [EC 2.7], more preferably phosphotransferases with an alcohol group as acceptor [EC 2.7.1], most preferably glucoaminase [EC 2.7.1.12].

Furthermore, the GMS proteins or subunits of such proteins having activity towards or which are involved in the bioconversion of a carbon source such as e.g. glucose into biomass are selected from the group consisting of membrane-bound PQQ-dependent D-glucose dehydrogenase, NAD(P)-dependent D-glucose dehydrogenase, cytosolic D-glucose kinase, enzymes or enzyme subunits having activity towards or involved in the assimilation of D-glucate such as membrane-bound FAD-dependent D-glucose dehydrogenase (2-keto-D-gluconate-forming), membrane-bound PQQ-dependent D-glucate dehydrogenase (5-keto-D-glucate-forming), cytosolic D-glucose kinase, enzymes or enzyme subunits having activity towards or involved in the assimilation of 2 KD such as membrane-bound FAD-dependent 2-keto-D-glucate dehydrogenase, NAD(P)-dependent glucose-1-dehydrogenase, flavin containing glucose-2,5-dehydrogenase, glucarate-5-dehydrogenase (5-keto-D-glucurate reductase), and cytosolic NAD(P)-dependent 2-keto-D-glucurate reductase.

In particular, it has now been found that GMS proteins encoded by polynucleotides having a nucleotide sequence that hybridizes preferably under highly stringent conditions to a sequence shown in SEQ ID NO:1 play an important role in the bioconversion of a carbon source such as e.g. glucose to biomass. It has also been found, that by genetically altering such nucleotides in a microorganism, such as for example Gluconobacter, the efficiency of said bioconversion within said microorganism can be even greatly improved leading e.g. to higher production and/or yield of biomass from a carbon source such as e.g. glucose.

Consequently, the invention relates to a polynucleotide selected from the group consisting of:
(a) polynucleotides encoding a polypeptide comprising the amino acid sequence according to SEQ ID NO:2;
(b) polynucleotides comprising the nucleotide sequence according to SEQ ID NO:1;
(c) polynucleotides comprising a nucleotide sequence obtainable by nucleic acid amplification such as polymerase chain...
reaction, using genomic DNA from a microorganism as a template and a primer set according to SEQ ID NO:3 and SEQ ID NO:4;

(d) polynucleotides comprising a nucleotide sequence encoding a fragment or derivative of a polypeptide encoded by a polynucleotide of any of (a) to (c) wherein in said derivative one or more amino acid residues are conservatively substituted compared to said polypeptide, and said fragment or derivative has the activity of an oxidoreductase [EC 1], preferably an oxidoreductase acting on the CH—OH group of donors [EC 1.1] (GMS 01);

(e) polynucleotides the complementary strand of which hybridizes under stringent conditions to a polynucleotide as defined in any one of (a) to (d) and which encode an oxidoreductase [EC 1], preferably an oxidoreductase acting on the CH—OH group of donors [EC 1.1] (GMS 01); and

(f) polynucleotides which are at least 70%, such as 85, 90 or 95% homologous to a polynucleotide as defined in any one of (a) to (d) and which encode an oxidoreductase [EC 1], preferably an oxidoreductase acting on the CH—OH group of donors [EC 1.1] (GMS 01);

or

the complementary strand of such a polynucleotide.

[0012] The GMS protein as isolated from *Glucanobacter oxydans* DSM 17078 shown in SEQ ID NO:2 and described herein was found to be particularly useful, since the gene encoding said GMS protein appeared to perform a crucial function in the bioconversion of a carbon source such as e.g. glucose to biomass in microorganisms, in particular in bacteria, such as acetate acid bacteria, such as *Glucanobacter, Acetobacter* and *Glucanacetobacter*. Accordingly, the invention relates to a polynucleotide encoding a polypeptide according to SEQ ID NO:2. This polypeptide may be encoded by a nucleotide sequence as shown in SEQ ID NO:1. The invention therefore also relates to polynucleotides comprising the nucleotide sequence according to SEQ ID NO:1.

[0013] The nucleotide and amino acid sequences determined above were used as a “query sequence” to perform a search with the Blast2 program (version 2 or BLAST from National Center for Biotechnology [NCBI] against the database PRO SW-SwissProt (full release plus incremental updates). From the searches, the GMS 01 polynucleotide according to SEQ ID NO:1 was annotated as encoding a protein having membrane-bound PQQ-dependent glucose dehydrogenase activity.

[0014] A nucleic acid according to the invention may be obtained by nucleic acid amplification using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers such as the nucleotide primers according to SEQ ID NO:3 and SEQ ID NO:4 according to standard PCR amplification techniques. The nucleic acid thus amplified may be cloned into an appropriate vector and characterized by DNA sequence analysis.

[0015] The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from strains known or suspected to comprise a polynucleotide according to the invention. The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a new nucleic acid sequence as described herein, or a functional equivalent thereof.

[0016] The PCR fragment may then be used to isolate a full length cDNA clone by a variety of known methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage or cosmid cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

[0017] Accordingly, the invention relates to polynucleotides comprising a nucleotide sequence obtainable by nucleic acid amplification such as polymerase chain reaction, using DNA such as genomic DNA from a microorganism as a template and a primer set according to SEQ ID NO:3 and SEQ ID NO:4.

[0018] The invention also relates to polynucleotides comprising a nucleotide sequence encoding a fragment or derivative of a polypeptide encoded by a polynucleotide as described herein wherein in said derivative one or more amino acid residues are conservatively substituted compared to said polypeptide, and said fragment or derivative has the activity of a GMS polypeptide, preferably a GMS 01 polypeptide.

[0019] The invention also relates to polynucleotides the complementary strand of which hybridizes under stringent conditions to a polynucleotide as defined herein and which encode a GMS polypeptide, preferably a GMS 01 polypeptide.

[0020] The invention also relates to polynucleotides which are at least 70% identical to a polynucleotide as defined herein and which encode a GMS polypeptide, and the invention also relates to polynucleotides being the complementary strand of a polynucleotide as defined herein above.

[0021] The invention also relates to primers, probes and fragments that may be used to amplify or detect a DNA according to the invention and to identify related species or families of microorganisms also carrying such genes.

[0022] The present invention also relates to vectors which include polynucleotides of the invention and microorganisms which are genetically engineered with the polynucleotides or said vectors.

[0023] The invention also relates to processes for producing microorganisms capable of expressing a polypeptide encoded by the above defined polynucleotide and a polypeptide encoded by a polynucleotide as defined above.

[0024] The invention also relates to microorganisms wherein the activity of a GMS polypeptide, preferably a GMS 01 polypeptide, is reduced or abolished so that the yield and/or production of biomass which is produced through bioconversion of a carbon source such as e.g. glucose is increased.

[0025] A suitable carbon source that can be converted into biomass may be for instance glucose or selected from carbon sources the assimilation of which results in the formation of glucose, in particular D-glucose, such as sucrose, maltose, starch, cellulose, cellulobiose, lactose, isomaltose, dextran, trehalose or mixtures thereof.

[0026] The skilled person will know how to reduce or abolish the activity of a GMS protein, preferably a GMS 01 protein. Such may be for instance accomplished by either genetically modifying the host organism in such a way that it produces less or no copies of the GMS protein, preferably the GMS 01 protein, than the wild type organism or by decreasing or abolishing the specific activity of the GMS protein, preferably the GMS 01 protein.

[0027] In the following description, procedures are detailed to achieve this goal, i.e. the increase in the yield and/or production of biomass which is produced through bioconversion of a carbon source such as e.g. glucose by reducing or abol-
ishing the activity of a GMS 01 protein. These procedures apply mutatis mutandis for other GMS proteins.

[0028] Modifications in order to have the organism produce less or no copies of the GMS 01 gene and/or protein may include the use of a weak promoter, or the mutation (e.g. insertion, deletion or point mutation) of (parts of) the GMS 01 gene or its regulatory elements. Decreasing or abolishing the specific activity of a GMS 01 protein may also be accomplished by methods known in the art. Such methods may include the mutation (e.g. insertion, deletion or point mutation) of (parts of) the GMS 01 gene.

[0029] Also known in the art are methods of reducing or abolishing the activity of a given protein by contacting the GMS 01 protein with specific inhibitors or other substances that specifically interact with the GMS 01 protein. In order to identify such specific inhibitors, the GMS 01 protein may be expressed and tested for activity in the presence of compounds suspected to inhibit the activity of the GMS 01 protein. Potential inhibiting compounds may for instance be monoclonal or polyclonal antibodies against the GMS 01 protein. Such antibodies may be obtained by routine immunization protocols of suitable laboratory animals.

[0030] The invention may be performed in any microorganism carrying a GMS 01 gene or equivalent or homologue thereof. Suitable microorganisms may be selected from the group consisting of yeast, algae and bacteria, either as wild type strains, mutant strains derived by classic mutagenesis and selection methods or as recombinant strains. Examples of such yeasts may be, e.g., Candida, Saccharomyces, Zygosaccharomyces, Schizosaccharomyces, or Kluveromyces. An example of such algae may be, e.g., Chlorella. Examples of such bacteria may be, e.g., Gluconobacter, Acetobacter, Gluconacetobacter, Ketogulonicigenium, Pantoea, Pseudomonas, such as, e.g., Pseudomonas putida, and Escherichia, such as, e.g., Escherichia coli. Preferred are Gluconobacter or Acetobacter acetii, such as for instance G. oxydans, G. cerinus, G. fraterculi, A. acetii subsp. xylinum or A. acetii subsp. orleanus, preferably G. oxydans DSM 17078. Gluconobacter oxydans DSM 17078 (formerly known as Gluconobacter oxydans N44-1) has been deposited at Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Mascheroder Weg 1B, D-38124 Braunschweig, Germany according to the Budapest Treaty on 26. Jan. 2005.

[0031] Microorganisms which can be used for the present invention may be publicly available from different sources, e.g., Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), Mascheroder Weg 1B, D-38124 Braunschweig, Germany; American Type Culture Collection (ATCC), P.O. Box 1549, Manassas, Va. 20108 USA; or Culture Collection Division, NITE Biological Resource Center, 2-5-8, Kazusakamatari, Kisanaru-shi, Chiba, 292-0818, Japan (formerly: Institute for Fermentation, Osaka (IFO), 17-85, Juso-hommachi 2-chome, Yodogawa-ku, Osaka 532-8686, Japan). Examples of preferred bacteria deposited with IFO are for instance Gluconobacter oxydans (formerly known as G. melanogenus) IFO 3293, Gluconobacter oxydans (formerly known as G. melanogenus) IFO 3292, Gluconobacter oxydans (formerly known as G. rubiginosus) IFO 3244, Gluconobacter fraterculi (formerly known as G. industrius) IFO 3260, Gluconobacter cerinus IFO 3266, Gluconobacter oxydans IFO 3287, and Acetobacter acetii subsp. orleanus IFO 3259, which were all deposited on Apr. 5, 1954; Acetobacter acetii subsp. xylinum IFO 13693 deposited on Oct. 22, 1975, and Acetobacter acetii subsp. xylinum IFO 13773 deposited on Dec. 8, 1977. Strain Acetobacter sp. ATCC 15164, which is also an example of a preferred bacterium, was deposited with ATCC. Strain Gluconobacter oxydans (formerly known as G. melanogenus) N 44-1 as another example of a preferred bacterium is a derivative of the strain IFO 3293 and is described in Sugisawa et al., Agric. Biol. Chem. 54: 1201-1209, 1990.

[0032] A microorganism as of the present invention may carry further modifications either on the DNA or protein level (see above), as long as such modification(s) has/have a direct impact on the yield, production and/or efficiency of biomass from a substrate such as e.g. glucose. Such further modifications may for instance affect other genes encoding oxidoreductases or isomerases as described above, in particular genes encoding NAD(P)-glucose dehydrogenase, POQ-dependent glucose dehydrogenase, FAD-dependent glucose-2-dehydrogenase, NADP+-glucose-2-dehydrogenase, gluconate-5-dehydrogenase, 2-ketogluconate dehydrogenase, gluconate oxidase, phosphoglucoisomerase, or further genes encoding enzymes involved in the pentose phosphate pathway, such as for instance 6-phosphogluconate dehydrogenase, 6-phosphogluconolactonase, ribulose-5-phosphate 3-epimerase, transaldolase or transketolase, or further genes encoding enzymes linking the pentose phosphate pathway to other pathways of carbon metabolism, such as for instance 6-phosphogluconate dehydrogenase, fructose-1,6-bisphosphate or phosphofructokinase. Methods of performing such modifications are known in the art, with some examples further described herein.

[0033] In accordance with a further object of the present invention there is provided the use of a polynucleotide as defined above or a microorganism which is genetically engineered using such polynucleotides in the efficient production of biomass from a carbon source.

[0034] The invention also relates to processes for the expression of endogenous genes in a microorganism, to processes for the production of polypeptides as defined above in a microorganism and to processes for the production of microorganisms capable of growth on a given carbon source such as e.g. glucose. All these processes comprise the step of altering a microorganism, wherein “altering” as used herein encompasses the process for “genetically altering” or “altering the composition of the cell culture media and/or methods used for culturing” in such a way that the yield and/or productivity of biomass can be improved compared to the wild-type organism. As used herein, “improved yield of biomass” means an increase of at least 5%, 15%, 25%, 30%, 40%, 50%, 75%, 100%, 200% or even more than 500%, depending on the carbon source used for growth, compared to a wild-type microorganism, i.e. a microorganism which is not genetically altered.

[0035] The term “genetically engineered” or “genetically altered” means the scientific alteration of the structure of genetic material in a living organism. It involves the production and use of recombinant DNA. More in particular it is used to delineate the genetically engineered or modified organism from the naturally occurring organism. Genetic engineering may be done by a number of techniques known in the art, such as e.g. gene replacement, gene amplification, gene disruption, transfection, transformation using plasmids, viruses, or other vectors. A genetically modified organism, e.g. genetically modified microorganism, is also often referred to as a recombinant organism, e.g. recombinant microorganism.
Advantageous embodiments of the invention become evident from the dependent claims. These and other aspects and embodiments of the present invention should now be apparent to those skilled in the art based on the teachings herein.

The sequence of the gene comprising a nucleotide sequence according to SEQ ID NO:1 encoding a GMS 01 protein was determined by sequencing a genomic clone obtained from Gluconobacter oxydans DSM 17078.

The invention also relates to a polynucleotide encoding at least a biologically active fragment or derivative of a GMS 01 polypeptide as shown in SEQ ID NO:2.

As used herein, "biologically active fragment or derivative" means a polypeptide which retains essentially the same biological function or activity as the polypeptide shown in SEQ ID NO:2. Examples of biological activity may for instance be enzymatic activity, signaling activity or antibody reactivity activity. The term "biological function" or "functional equivalent" as used herein means that the protein has essentially the same biological activity, e.g. enzymatic, signaling or antibody reactivity activity, as a polypeptide shown in SEQ ID NO:2.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living microorganism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition and still be isolated in that such vector or composition is not part of its natural environment.

An isolated polynucleotide or nucleic acid as used herein may be a DNA or RNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5'-end and one on the 3'-end) in the naturally occurring genome of the organism from which it is derived. Thus, in one embodiment, a nucleic acid includes some or all of the 5'-non-coding (e.g., promoter) sequences that are immediately contiguous to the coding sequence. The term "isolated polynucleotide" therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide that is substantially free of cellular material, viral material, or culture medium (when produced by recombinant DNA techniques), or chemical precursors or other chemicals (when chemically synthesized). Moreover, an "isolated nucleic acid fragment" is a nucleic acid fragment that is not naturally occurring as a fragment and would not be found in the natural state.

As used herein, the terms "polynucleotide", "gene" and "recombinant gene" refer to nucleic acid molecules which may be isolated from chromosomal DNA, which include an open reading frame encoding a protein, e.g. G. oxydans DSM 17078 GMS proteins. A polynucleotide may include a polynucleotide sequence as shown in SEQ ID NO:1 or fragments thereof and regions upstream and downstream of the gene sequences which may include, for example, promoter regions, regulator regions and terminator regions important for the appropriate expression and stabilization of the polypeptide derived thereof.

A gene may include coding sequences, non-coding sequences such as for instance untranslated sequences located at the 3'- and 5'-ends of the coding region of a gene, and regulatory sequences. Moreover, a gene refers to an isolated nucleic acid molecule as defined herein. It is furthermore appreciated by the skilled person that DNA sequence polymorphisms that lead to changes in the amino acid sequences of GMS proteins may exist within a population, e.g., the Gluconobacter oxydans population. Such genetic polymorphism in the GMS 01 gene may exist among individuals within a population due to natural variation or in cells from different populations. Such natural variations can typically result in 1%-5% variance in the nucleotide sequence of the GMS 01 gene. Any and all such nucleotide variations and the resulting amino acid polymorphism in GMS 01 are the result of natural variation and that do not alter the function or biological activity of GMS proteins are intended to be within the scope of the invention.

As used herein, the terms "polynucleotide" or "nucleic acid molecule" are intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is double-stranded DNA. The nucleic acid may be synthesized using oligonucleotide analogs or derivatives (e.g., inosine or phosphorothioate nucleotides). Such oligonucleotides may be used, for example, to prepare nucleic acids that have altered base-pairing abilities or increased resistance to nucleases.

The sequence information as provided herein should not be so narrowly construed as to require inclusion of erroneously identified bases. The specific sequences disclosed herein may be readily used to isolate the complete gene from a microorganism capable of converting a given carbon source such as e.g. glucose into biomass, in particular Gluconobacter oxydans, preferably Gluconobacter oxydans DSM 17078 which in turn may easily be subjected to further sequence analyses thereby identifying sequencing errors.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence may be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the
amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

[0048] The person skilled in the art is capable of identifying such erroneously identified bases and knows how to correct for such errors.

[0049] A nucleic acid molecule according to the invention may comprise only a portion or a fragment of the nucleic acid sequence provided by the present invention, such as for instance the sequence shown in SEQ ID NO:1, for example a fragment which may be used as a probe or primer such as for instance SEQ ID NO:3 or SEQ ID NO:4 or a fragment encoding a portion of a protein according to the invention. The nucleotide sequence determined from the cloning of the GMS 01 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other GMS 01 family members, as well as GMS 01 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotides which typically comprises a region of nucleotide sequence that hybridizes preferably under highly stringent conditions to at least about 12 or 15, preferably about 18 or 20, more preferably about 22 or 25, even more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 or more consecutive nucleotides of a nucleotide sequence shown in SEQ ID NO:1 or a fragment or derivative thereof.

[0050] A nucleic acid molecule encompassing all or a portion of the nucleic acid sequence of SEQ ID NO:1 may be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence information contained herein.

[0051] A nucleic acid of the invention may be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid thus amplified may be cloned into an appropriate vector and characterized by DNA sequence analysis.

[0052] Fragments of a polynucleotide according to the invention may also comprise polynucleotides not encoding functional polypeptides. Such polynucleotides may function as probes or primers for a PCR reaction.

[0053] Nucleic acids according to the invention irrespective of whether they encode functional or non-functional polypeptides, may be used as hybridization probes or polymerase chain reaction (PCR) primers. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having a GMS 01 activity include, inter alia, (1) isolating the gene encoding the protein of the present invention, or allelic variants thereof from a cDNA library, e.g., from other organisms than *Glucanobacter oxydans* and (2) Northern blot analysis for detecting expression of mRNA of said protein in specific cells or (3) use in abolishing or altering the function or activity of homologous GMS 01 genes in said other organisms.

[0054] Probes based on the nucleotide sequences provided herein may be used to detect transcripts or genomic sequences encoding the same or homologous proteins for instance in other organisms. Nucleic acid molecules corresponding to natural variants and non-G. oxydans homologues of the *G. oxydans* GMS 01 DNA of the invention which are also embraced by the present invention may be isolated based on their homology to the *G. oxydans* GMS 01 nucleic acid disclosed herein using the *G. oxydans* DNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques, preferably under highly stringent hybridization conditions.

[0055] In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme cofactor.

[0056] Homologous gene sequences may be isolated, for example, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of nucleotide sequences as taught herein.

[0057] The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from strains known or suspected to express a polynucleotide according to the invention. The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a new nucleic acid sequence as described herein, or a functional equivalent thereof.

[0058] The PCR fragment may then be used to isolate a full length cDNA clone by a variety of known methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage or cosmID cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

[0059] PCR technology can also be used to isolate full-length cDNA sequences from other organisms. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5'-end of the amplified fragment for the priming of first strand synthesis.

[0060] The resulting RNA/DNA hybrid may then be “tailed” (e.g., with guanines) using a standard terminal transferase reaction, the hybrid may be digested with RNaseH, and second strand synthesis may then be primed (e.g., with a poly-C primer). Thus, cDNA sequences upstream of the amplified fragment may easily be isolated. For a review of useful cloning strategies, see e.g., Sambrook et al., supra; and Ausubel et al., supra.

[0061] Also, nucleic acids encoding other GMS 01 family members, which thus have a nucleotide sequence that differs from a nucleotide sequence according to SEQ ID NO:1, are within the scope of the invention. Moreover, nucleic acids encoding GMS 01 proteins from different species which thus may have a nucleotide sequence which differs from a nucleotide sequence shown in SEQ ID NO:1 are within the scope of the invention.

[0062] The invention also relates to an isolated polynucleotide hybridizable under stringent conditions, preferably under highly stringent conditions, to a polynucleotide according to the present invention, such as for instance a polynucleotide shown in SEQ ID NO:1. Advantageously, such polynucleotide may be obtained from a microorganism capable of converting a given carbon source such as e.g. glucose into biomass, in particular *Glucanobacter oxydans*, preferably *Glucanobacter oxydans* DSM 17078.

[0063] As used herein, the term “hybridizing” is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about 50%, at least about 60%, at least about 70%, more preferably at least about 80%, even more preferably at least about 85% to 90%, most preferably at least 95% homologous to each other typically remain hybridized to each other.
In one embodiment, a nucleic acid of the invention is at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to a nucleic acid sequence shown in SEQ ID NO:1 or the complement thereof.

A preferred, non-limiting example of such hybridization conditions are hybridization in 6x sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 1xSSC, 0.1% SDS at 50°C, preferably at 55°C, more preferably at 60°C, and even more preferably at 65°C.

Highly stringent conditions include, for example, 2 h to 4 days incubation at 42°C using a digoxigenin (DIG)-labeled DNA probe (prepared by using a DIG labeling system; Roche Diagnostics GmbH, 68298 Mannheim, Germany) in a solution such as DigEasyHyb solution (Roche Diagnostics GmbH) with or without 100 μg/ml salmon sperm DNA, or a solution comprising 50% formamide, 5xSSC (150 mM NaCl, 15 mM trisodium citrate), 0.02% sodium dodecyl sulfate, 0.1% N-lauroylsarcosine, and 2% blocking reagent (Roche Diagnostics GmbH), followed by washing the filters twice for 5 to 15 minutes in 2xSSC and 0.1% SDS at room temperature and then washing twice for 15-30 minutes in 0.5xSSC and 0.1% SDS or 0.1xSSC and 0.1% SDS at 65-68°C.

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under preferably highly stringent conditions to a nucleotide sequence of the invention corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural G. oxydans GMS 01 protein.

The skilled artisan will know which conditions to apply for stringent and highly stringent hybridization conditions. Additional guidance regarding such conditions is readily available in the art, for example, in Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.).

Of course, a polynucleotide which hybridizes only to a poly(A) sequence (such as the 3'-terminal poly(A) tract of mRNAs), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to specifically hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly(A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

In a typical approach, genomic DNA or cDNA libraries constructed from other organisms, e.g. microorganisms capable of converting of a given carbon source such as e.g. glucose into biomass, in particular other Gluconobacter species may be screened.

For example, Gluconobacter strains may be screened for homologous polynucleotides by Northern blot analysis. Upon detection of transcripts homologous to polynucleotides according to the invention, DNA libraries may be constructed from RNA isolated from the appropriate strain, utilizing standard techniques well known to those of skill in the art. Alternatively, a total genomic DNA library may be screened using a probe hybridisable to a polynucleotide according to the invention. A preferred nucleic acid molecule of the present invention, such as for instance a nucleic acid molecule shown in SEQ ID NO:1 or a fragment or derivative thereof, may be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or portion of the nucleic acid sequence shown in SEQ ID NO:1 as a hybridization probe, nucleic acid molecules according to the invention may be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Frith, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

Furthermore, oligonucleotides corresponding to or hybridisable to nucleotide sequences according to the invention may be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

The terms "homology" or "percent identity" are used interchangeably herein. For the purpose of this invention, it is defined here that in order to determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps may be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid 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. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity=number of identical positions/total number of positions×100). Preferably, the two sequences are the same length.

The skilled person will be aware of the fact that several different computer programs are available to determine the homology between two sequences. For instance, a comparison of sequences and determination of percent identity between two sequences may be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48): 444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.accelrys.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6 or 4 and a length weight of 1, 2, 3, 4, 5 or 6. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.accelrys.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70 or 80 and a length weight of 1, 2, 3, 4, 5 or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4: 11-17 (1989) which has been incorporated into the ALIGN program (version 2.0) (available at http://vega.
igh.cnrS.fr/bin/align-guess.cgi) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention may further be used as a “query sequence” to perform a search against public databases to, for example, identify other family members or related sequences. Such searches may be performed using the BLASTN and BLASTX programs (version 2.0) of Altschul et al. (1990) J. Mol. Biol. 215:403–10. BLAST nucleotide searches may be performed with the BLASTN program, score=100, word length=12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the present invention. BLAST protein searches may be performed with the BLASTX program, score=50, word length=3 to obtain amino acid sequences homologous to the protein molecules of the present invention. To obtain gapped alignments for comparison purposes, Gapped BLAST may be utilized as described in Altschul et al. (1997) Nucleic Acids Res. 25 (17): 3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) may be used. See http://www.ncbi.nlm.nih.gov.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is the complement of a nucleotide sequence as of the present invention, such as for instance the sequence shown in SEQ ID NO:1. A nucleic acid molecule, which is complementary to a nucleotide sequence disclosed herein, is one that is sufficiently complementary to a nucleotide sequence shown in SEQ ID NO:1 such that it may hybridize to said nucleotide sequence thereby forming a stable duplex.

In a further preferred embodiment, a nucleic acid of the invention as shown in SEQ ID NO:1 or the complement thereof contains at least one mutation leading to a gene product with modified function/activity. The at least one mutation may be introduced by methods described herein. In one aspect, the at least one mutation leads to a GMS 01 protein whose function compared to the wild type counterpart is completely or partially destroyed. Methods for introducing such mutations are well known in the art.

The term “reduction” of activity as used herein encompasses decreasing activity of one or more polypeptides in the producing organism, which in turn are encoded by the corresponding polynucleotides described herein. There are a number of methods available in the art to accomplish reduction of activity of a given protein, in this case the GMS 01 protein. In general, the specific activity of a protein may be decreased or the copy number of the protein may be decreased.

To facilitate such a decrease, the copy number of the genes corresponding to the polynucleotides described herein may be decreased, such as for instance by underexpression or disruption of a gene. A gene is said to be “underexpressed” if the level of transcription of said gene is reduced in comparison to the wild type gene. This may be measured by, for instance Northern blot analysis quantifying the amount of mRNA as an indication for gene expression. As used herein, a gene is underexpressed if the amount of generated mRNA is decreased by at least 1%, 2%, 5% 10%, 25%, 50%, 75%, 100%, 200% or even more than 500%, compared to the amount of mRNA generated from a wild-type gene. Alternatively, a weak promoter may be used to direct the expression of the polynucleotide. In another embodiment, the promoter, regulatory region and/or the ribosome binding site upstream of the gene can be altered to achieve the down-expression. The expression may also be reduced by decreasing the relative half-life of the messenger RNA. In another embodiment, the activity of the polypeptide itself may be decreased by employing one or more mutations in the polypeptide amino acid sequence, which decrease the activity. For example, altering the affinity of the polypeptide for its corresponding substrate may result in reduced activity. Likewise, the relative half-life of the polypeptide may be decreased. In either scenario, that being reduced gene expression or reduced activity, the reduction may be achieved by altering the composition of the cell culture media and/or methods used for cultivating.

“Reduced expression” or “reduced activity” as used herein means a decrease of at least 5%, 10%, 25%, 50%, 75%, 100%, 200% or even more than 500%, compared to a wild-type protein, polynucleotide, gene, or the activity and/or the concentration of the protein present before the polynucleotides or polypeptides are reduced. The activity of the GMS 01 protein may also be reduced by contacting the protein with a specific or general inhibitor of its activity. The terms “reduced activity”, “decreased or abolished activity” are used interchangeably herein.

Another aspect of the invention pertains to vectors, containing a nucleic acid encoding a protein according to the invention or a functional equivalent or portion thereof. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacteriophage origin of replication). Other vectors are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome.

Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “expression vectors”. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. The terms “plasmid” and “vector” can be used interchangeably herein as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vector includes one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, “operatively linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term “regulatory sequence” is intended to include promoters, enhancers and other expression control elements (e.g., attenuror). Such regulatory sequences are described, for example, in Goeddel; Gene Expression Tech-
ology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those which direct constitutive or inducible expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in a certain host cell (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention may be introduced into host cells to thereby produce proteins or peptides, encoded by nucleic acids as described herein, including, but not limited to, mutant proteins, fragments thereof, variants or functional equivalents thereof, and fusion proteins, encoded by a nucleic acid as described herein, e.g., GMS 01 proteins, mutant forms of GMS 01 proteins, fusion proteins and the like.

The recombinant expression vectors of the invention may be designed for expression of GMS 01 proteins in a suitable microorganism. For example, a protein according to the invention may be expressed in bacterial cells such as strains belonging to the genus *Glucobacter*, *Glucocetobacter* or *Acetobacter*. Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

The DNA insert may be operatively linked to an appropriate promoter, which may be either a constitutive or inducible promoter. The skilled person will know how to select suitable promoters. The expression constructs may contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcript expressed by the constructs may preferably include an initiation codon at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

Vector DNA may be introduced into suitable host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation", "transfection" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipid-mediated transfection or electroporation. Suitable methods for transforming or transfecting host cells may be found in Sambrook, et al. (supra), Davis et al., Basic Methods in Molecular Biology (1986) and other laboratory manuals.

In order to identify and select cells which have integrated the foreign DNA into their genome, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those that confer resistance to drugs, such as kanamycin, tetracycline, ampicillin and streptomycin. A nucleic acid encoding a selectable marker is preferably introduced into a host cell on the same vector as that encoding a protein according to the invention or can be introduced on a separate vector such as, for example, a suicide vector, which cannot replicate in the host cells. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

The invention provides also an isolated polypeptide having the amino acid sequence shown in SEQ ID NO:2 or an amino acid sequence obtainable by expressing a polynucleotide of the present invention, such as for instance a polynucleotide sequence shown in SEQ ID NO:1 in an appropriate host.

Polypeptides according to the invention may contain only conservative substitutions of one or more amino acids in the amino acid sequence represented by SEQ ID NO:2 or substitutions, insertions or deletions of non-essential amino acids. Accordingly, a non-essential amino acid is a residue that may be altered in the amino acid sequences shown in SEQ ID NO:2 without substantially altering the biological function. For example, amino acid residues that are conserved among the proteins of the present invention, are predicted to be particularly unamenable to alteration. Furthermore, amino acids conserved among the proteins according to the present invention and other GMS 01 proteins are not likely to be amenable to alteration.

The term "conservative substitution" is intended to mean that a substitution in which the amino acid residue is replaced with an amino acid residue having a similar side chain. These families are known in the art and include amino acids with basic side chains (e.g., lysine, arginine and histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

As mentioned above, the polynucleotides of the invention may be utilized in the genetic engineering of a suitable host cell to make it better and more efficient in the bioconversion of a carbon source such as e.g. glucose to biomass, i.e. reduction of carbon source diversion throughout the bioconversion process.

According to the invention a genetically engineered/recombinant host cell (also referred to as recombinant cell or transformed cell) may be produced carrying such a modified polynucleotide wherein the function of the linked protein is significantly modified in comparison to a wild-type cell such that the yield and/or productivity of biomass from a carbon source such as e.g. glucose is improved. The host cell may be selected from a microorganism capable of converting a given carbon source such as e.g. glucose into biomass, in particular *Glucobacter oxydans*, preferably *G. oxydans* DSM 17078.

A "transformed cell" or "recombinant cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid according to the invention, or wherein the activity of the GMS 01 protein has been decreased or abolished. Suitable host cells include cells of microorganisms capable of converting a given carbon source such as e.g. glucose into biomass. In particular, these include strains from the genera *Pseudomonas*, *Pantoaea*, *Escherichia*, *Corynebacterium*, *Ketogluconigenium* and acetic acid bacteria like e.g., *Glucobacter*, *Acetobacter* or *Glucocetobacter*, preferably *Acetobacter sp.*, *Acetobacter acetii*, *Glucobacter flavescens*, *Gluconobacter cerinus*, *Gluconobacter thailandicus*, Glu-
conobacter oxydans, more preferably G. oxydans, most preferably G. oxydans DSM 17078.

[0094] To improve the biomass production of a certain recombinant host cell, i.e. reduce the diversion of carbon substrates throughout the bioconversion process, GMS 01 gene expression may be inhibited in that organism for instance by targeting nucleotide sequences complementary to the regulatory region of a GMS 01 nucleotide sequence (e.g., a GMS 01 promoter and/or enhancers) to form triple helical structures that prevent transcription of a GMS 01 gene in target cells. See generally, Helene, C. (1991) Anticancer-DrugDes. 6 (6): 569-84; Helene, C. et al. (1992) Ann. N.Y. Acad. Sci. 660: 27-36; and Maher, L. J. (1992) Bioassays 14 (12): 807-15.

[0095] Inhibition or prevention of gene expression may also be achieved by modifying the GMS 01 gene, e.g., by introducing one or more mutations into the GMS 01 gene wherein said modification leads to a GMS 01 protein with a function which is significantly decreased in comparison to the wild-type protein.

[0096] Therefore, in one other embodiment, the polynucleotide carrying the at least one mutation is derived from a polynucleotide as represented by SEQ ID NO:1 or equivalents thereof.

[0097] A mutation as used herein may be any mutation leading to a less functional or unstable polypeptide, e.g. less functional or unstable GMS 01 gene products. This may include for instance an alteration in the genome of a microorganism, which interferes with the synthesis of GMS 01 or leads to the expression of a GMS 01 protein with an altered amino acid sequence whose function compared with the wild type counterpart having a non-altered amino acid sequence is completely or partially destroyed. The interference may occur at the transcriptional, translational or post-translational level.

[0098] The alteration in the genome of the microorganism may be obtained e.g. by replacing through a single or double crossover recombination a wild type DNA sequence by a DNA sequence containing the alteration. For convenient selection of transformants of the microorganism with the alteration in its genome the alteration may, e.g. be a DNA sequence encoding an antibiotic resistance marker or a gene complementing a possible auxotrophy of the microorganism. Mutations include, but are not limited to, deletion-insertion mutations.

[0099] An alteration in the genome of the microorganism leading to a less or non-functional polypeptide may also be obtained by randomly mutating the genome of the microorganism using e.g. chemical mutagens, radiation or transposons and selecting or screening for mutants which are better or more efficient producers of biomass. Standard methods for screening and selection are known to the skilled person.

[0100] In a specific embodiment, it is desired to knockout the GMS 01 gene of the present invention, i.e., wherein its gene expression is artificially suppressed in order to improve the yield and/or efficiency of biomass production when introduced into a suitable host cell. Methods of providing knockouts as well as microorganisms carrying such suppressed genes are well known in the art. The suppression of the endogenous GMS 01 gene may be induced by deleting at least a part of the gene or the regulatory region thereof. As used herein, “suppression of the gene expression” includes complete and partial suppression, as well as suppression under specific conditions and also suppression of the expression of either one of the two alleles.

[0101] In order to create a knockout microorganism in which the expression of the GMS 01 gene is artificially suppressed, first the GMS 01 gene may be cloned and then a vector for homologous recombination may be constructed by using the gene to inactivate the endogenous GMS 01 gene in the target microorganism. The vector for homologous recombination then contains a nucleic acid sequence designed to inactivate the endogenous GMS 01 gene in the target microorganism. Such a nucleic acid may be for instance a nucleic acid sequence of the GMS 01 gene or the regulatory region thereof, such as the existing flanking region of the gene to be inactivated (in cis), or existing separately (in trans), containing at least a partial deletion, or alternatively it may be a nucleic acid sequence of the GMS 01 gene or the regulatory region thereof containing other genes. A gene which can also function as a marker is preferably selected as the gene to be inserted into the GMS 01 gene or the regulatory region thereof. The insert genes to be used include for instance drug-resistance genes as defined above. There is no particular limitation on the position where the genes may be inserted in the GMS 01 gene, as long as the insertion at that position results in the suppression of the expression of the endogenous GMS 01 gene in the target. To avoid polar effects of the insertion, in-frame silent deletions can be introduced by using, for example, the sacB system or long-flanking homology PCR. These techniques are well known to the person skilled in the art.

[0102] The aforementioned mutagenesis strategies for GMS 01 proteins may result in increased yield and/or production of biomass. This list is not meant to be limiting; variations on these mutagenesis strategies will be readily apparent to one of ordinary skill in the art. By these mechanisms, the nucleic acid and protein molecules of the invention may be utilized to generate microorganisms such as Gluconoconobacter oxydans or related strains of bacteria expressing mutated GMS 01 nucleic acid and protein molecules such that the yield, efficiency and/or productivity of biomass production from a carbon source such as e.g. glucose is improved.

[0103] Biomass concentration can be measured using several methods known to the person skilled in the art, such as e.g. measuring the optical density of the respective cell suspensions at for instance a wavelength of 600 nm (OD600), or measuring either the wet- or dry-cell mass concentration or by counting the numbers of cells of the respective cell suspension. Such methods of measurement are known in the art. Thus, the cell dry weight (CDW) [g/l] may be for instance measured on a dried and tared nitrocellulose filter with a pore size of e.g. 0.45 μm on which the culture broth is filtrated. After washing the filter with water, the weight of the dried filter is again determined and the CDW calculated as follows:

\[\text{CDW} = \frac{m(\text{dried filter after broth filtration}) - m(\text{dried filter before broth filtration})}{V}\]

[0104] In one aspect of the invention, microorganisms (in particular from the genera Gluconoobacter, Gluconacetobacter and Acetobacter) are provided that are able to perform such improved bioconversion. When measured by a method as described herein these organisms were found to have an improved capability to produce biomass from a carbon source such as e.g. glucose. Such may be achieved by increasing the activity of a GMS polypeptide, preferably a GMS 01 polypeptide. The yield of biomass produced from a carbon
source, such as for instance glucose, when measured according to this method after an incubation period of 24 hours may be about 0.1 g or more dry biomass/g carbon source, preferably about 0.2 g or more dry biomass/g carbon source, or even about 0.3 g or more dry biomass/g carbon source, such as for instance about 0.4 g, 0.5 g, 0.6 g, 0.7 g, 0.8 g or 1.0 g or more dry biomass/g carbon source. Preferably, the carbon source is glucose.

[0105] The recombinant microorganism carrying e.g. a modified GMS 01 gene and which is able to produce biomass from a carbon source such as e.g. glucose in significantly higher yield, productivity, and/or efficiency may be cultured in an aqueous medium supplemented with appropriate nutrients under aerobic conditions. The cultivation may be conducted in batch, fed-batch, semi-continuous or continuous mode. The cultivation period may vary depending on for instance the host, pH, temperature and nutrient medium to be used, and is preferably about 1 to about 10 days when run in batch or fed-batch mode. The cultivation may be conducted at for instance a pH of about 4.0 to about 9.0, preferably about 5.0 to about 8.0. The preferred temperature range for carrying out the cultivation is from about 13°C to about 36°C, preferably from about 18°C to about 33°C. Usually, the culture medium may contain besides e.g. D-glucose as main carbon source other assimilable carbon sources, e.g., glycerol, D-mannitol, D-sorbitol, L-sorbose, erythritol, ribitol, xylitol, arabinol, inositol, dulcitol, D-ribose, D-fructose, sucrose and ethanol, preferably D-sorbitol, D-mannitol, D-fructose, glyceraldehyde and ethanol; and digestible nitrogen sources such as organic substances, e.g., peptone, yeast extract, baker’s yeast, urea, amino acids, and corn steep liquor. Various inorganic substances may also be used as nitrogen sources, e.g., nitrates and ammonium salts. Furthermore, the culture medium usually may contain inorganic salts, e.g., magnesium sulfate, manganese sulfate, potassium phosphate, and calcium carbonate.

[0106] The nucleic acid molecules, polypeptides, vectors, primers, and recombinant microorganisms described herein may be used in one or more of the following methods: identification of *Gluconobacter oxydans* and related organisms; mapping of genomes of organisms related to *Gluconobacter oxydans*; identification and localization of *Gluconobacter oxydans* sequences of interest; evolutionary studies; determination of GMS 01 protein regions required for function; modulation of a GMS 01 protein activity or function; modulation of the activity of a GMS pathway; and modulation of cellular production of biomass from a carbon source such as e.g. glucose.

[0107] The invention provides methods for screening molecules which modulate the activity of a GMS 01 protein, either by interacting with the protein itself or a substrate or binding partner of the GMS 01 protein, or by modulating the transcription or translation of a GMS 01 nucleic acid molecule of the invention. In such methods, a microorganism expressing one or more GMS 01 proteins of the invention is contacted with one or more test compounds, and the effect of each test compound on the activity or level of expression of the GMS 01 protein is assessed.

[0108] The biological, enzymatic or other activity of GMS proteins can be measured by methods well known to a skilled person, such as, for example, by incubating a cell fraction containing the GMS 01 protein in the presence of its substrate, electron acceptor(s) or donor(s) including phenazine methosulfate (PMS), dichlorophenol-indophenol (DCIP), NAD, NADH, NADP, NADPH, which consumption can be directly or indirectly measured by photometric, calorimetric or fluorimetric methods, and other inorganic components which might be relevant for the development of the activity. Thus, for example, the activity of membrane-bound D-glucose dehydrogenase can be measured in an assay where membrane fractions containing this enzyme are incubated in the presence of phosphate buffer at pH 6, D-glucose and the artificial electron acceptors DCIP and PMS. The rate of consumption of DCIP can be measured at 600 nm, and is directly proportional to the D-glucose dehydrogenase activity present in the membrane fraction.

[0109] Thus, the present invention is directed to the use of a polynucleotide, polypeptide, vector, primer and recombinant microorganism as described herein in the production of biomass, i.e., the bioconversion of a carbon source such as e.g. glucose into biomass. In a preferred embodiment, a modified polynucleotide, polypeptide, vector and recombinant microorganism as described herein is used for improving the yield, productivity, and/or efficiency of said biomass production.

[0110] The terms “production” or “productivity” are art-recognized and include the amount of biomass formed within a given time and a given cultivation volume (e.g., kg product per hour per liter) from a given amount or concentration of a carbon source such as e.g. glucose. The term “efficiency of production” includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a product). The term “yield” is art-recognized and includes the efficiency of the conversion of the carbon source into biomass. This is generally written as, for example, kg biomass per kg carbon source. By “increasing the yield and/or production of biomass” it is meant that in a given amount of culture over a given amount of time the biomass is increased. By measuring the increase in biomass and decrease in carbon source concentration in a growing culture, one can calculate the yield of biomass on consumed carbon source; i.e. the weight of biomass obtained by bioconverting a determined weight of carbon source such as e.g. glucose, defined as kg biomass per kg carbon source consumed. To determine the consumption of carbon source, the residual amount of carbon source in the growth medium can be measured by methods well known in the art (e.g., HPLC). The increase of biomass can be measured by methods well known in the art (e.g. optical density at 600 nm, or measurement of dry weight of a sample).

[0111] The terms “biosynthesis” or a “biosynthetic pathway” are art-recognized and include the synthesis of a compound/product, preferably an organic compound, by a cell from intermediate compounds in what may be a multistep and highly regulated process. The term “bioconversion” is art-recognized and includes the conversion of a carbon source such as e.g. glucose into a product, i.e. biomass, by means of one or more biosynthetic step(s) which involve one or more enzyme(s) and/or transporter(s). The language “metabolism” is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (e.g., the metabolism of an amino acid such as glycine) comprises the overall biosynthetic, modification, and degradation pathways in the cell related to this compound. The language “transport” or “import” is art-recognized and includes the facilitated movement of one or
more molecules across a cellular membrane through which the molecule would otherwise either be unable to pass or be passed inefficiently.

[0112] In one preferred embodiment, the present invention is related to a process for the production of biomass from a carbon source such as e.g. glucose wherein a modified polynucleotide sequence as described above is introduced into a suitable microorganism and the recombinant microorganism is cultured under conditions that allow the production of biomass in high productivity, yield, and/or efficiency.

[0113] Recombinant microorganisms according to the present invention carrying e.g. a modified GMS gene, in particular a modified GMS 01 gene, and which are able to produce biomass from a carbon source such as e.g. glucose in significantly higher yield, productivity, and/or efficiency may advantageously be used in a number of applications. One particularly suited application is a method for the production of Vitamin C and/or intermediates thereof such like 2-keto-L-gulonic acid (2-KGA).

[0114] Thus, it is an aspect of the present invention to provide a process for the production of Vitamin C and/or 2-KGA wherein a nucleotide according to the invention or a modified polynucleotide sequence as described above is introduced into a suitable microorganism as described above wherein a suitable carbon source such as e.g. glucose is efficiently converted into biomass, the recombinant microorganism is cultured under conditions that allow the production of Vitamin C and/or 2-KGA in high productivity, yield, and/or efficiency, the produced fermentation product is isolated from the culture medium and optionally further purified.

[0115] The carbon sources used for the production of (1) biomass and (2) Vitamin C and/or 2-KGA may be the same or may differ. In one preferred embodiment, the carbon source(s) used for the production of biomass is different from the carbon source(s) used for the direct production of Vitamin C and/or 2-KGA. These two different carbon sources may be used simultaneously or sequentially, wherein a first carbon source would be used for the production of biomass, and this biomass would then be used to convert a second carbon source into Vitamin C and/or 2-KGA.

[0116] A suitable carbon source that can be converted directly into Vitamin C and/or 2-KGA may be selected from the D-glucose or D-sorbitol metabolism pathway such as, for example, D-glucose, D-sorbitol, L-sorbose, L-sorbosone, D-glucurate, 2-keto-D-glucurate or 2,5-diketo-glucorate or mixtures thereof. Preferably, the substrate is selected from D-glucose, D-sorbitol, L-sorbose or L-sorbosone mixtures thereof, more preferably from D-glucose, D-sorbitol, L-sorbose mixtures thereof. Preferably from D-glucose, D-sorbitol, L-sorbose, or mixtures thereof.

[0117] Vitamin C as used herein may be any chemical form of L-ascorbic acid found in aqueous solutions, such as for instance undissociated, in its free acid form or dissociated as an anion. The solubilized salt form of L-ascorbic acid may be characterized as the anion in the presence of any kind of cations usually found in fermentation supernatants, such as for instance potassium, sodium, ammonium, or calcium. Also included may be isolated crystals of the free acid form of L-ascorbic acid. On the other hand, isolated crystals of a salt form of L-ascorbic acid are called by their corresponding salt name, i.e. sodium ascorbate, potassium ascorbate, calcium ascorbate and the like.

[0118] Conversion of a suitable carbon source into Vitamin C and/or 2-KGA in connection with the above process using a microorganism means that the conversion of the carbon source resulting in Vitamin C and/or 2-KGA is performed by the microorganism, i.e. the carbon source may be directly converted into Vitamin C and/or 2-KGA. Said microorganism is cultured under conditions which allow such conversion from said carbon source as it is known for the skilled person.

[0119] A medium as used herein for the above process using a microorganism may be any suitable medium for the production of Vitamin C and/or 2-KGA. Typically, the medium is an aqueous medium comprising for instance salts, substrate(s), and a certain pH.

[0120] The term "direct conversion" and the like is intended to mean that a microorganism is capable of the conversion of a certain carbon source into the specified product by means of one or more biological conversion steps, without the need of any additional chemical conversion step. For instance, the term "direct conversion of D-sorbitol into Vitamin C" is intended to describe a process wherein a microorganism is producing Vitamin C and wherein D-sorbitol is offered as a carbon source without the need of an intermediate chemical conversion step. A single microorganism capable of directly fermenting Vitamin C is preferred.

[0121] This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents and published patent applications, cited throughout this application are hereby incorporated by reference.

EXAMPLES

Example 1

Analytical Methods for Biomass and Glucose Determination

[0122] For determination of biomass production, cells of G. oxydans were grown at 27°C. for 3 days on MB agar, containing 25 g/l mannitol, 5.0 g/l yeast extract (Difco), 3.0 g/l Bactopeptone (Difco) and 15 g/l agar. Cells were resuspended in 10% glycerol (10 ml per plate) and 1 ml aliquots were stored at ~80°C.

[0123] A baffled 500 ml shake flask with 100 ml medium No. 5 containing 50 g/l D-glucose, 0.5 g/l glycerol, 15 g/l yeast extract (Difco), 2.5 g/l MgSO₄·7H₂O, 0.3 g/l K₂HPO₄, 15 g/l CaCO₃ and 1 drop antifoam was inoculated with a 1 ml aliquot and incubated at 30°C. and 180 rpm in a shaker for 48 h. The optical density at 600 nm (OD₆₀₀) was determined using a Spectronic 4001/N spectrophotometer and an identical medium No. 5 flask was then inoculated with this culture, such that the start OD₆₀₀ in the second flask was 0.25. This flask was then further incubated at 30°C., 180 rpm for another 60 h with samples of 3 ml taken every 12 h for analysis of biomass and glucose concentrations as follows:

[0124] Production of biomass was determined via measurement of the dry weight of the sample. Reaction tubes were pre-dried at 40°C. under vacuum for 48 h and the weight of the empty tubes measured [w1]. The sample was added and the weight of the full tube determined [w2]. 0.1 ml 18.5% HCl was added to dissolve the CaCO₃. The tubes were subsequently centrifuged for 3 min at 13000 rpm. The supernatant was discarded, 0.5 ml of water was added and the pellet was resuspended. The tube was centrifuged again for 3 min at 13000 rpm and the supernatant was discarded. The pellet was
dried at 40°C under vacuum for 48 h and the tube containing the pellet again weighted [w3]. The concentration of biomass [B] in the sample was calculated as follows:

\[ \frac{(w_3-w_1)}{(w_2-w_1)} \times 1000 = \text{concentration of biomass} \ [\text{g/kg}] \]

[0125] The amount or concentration of glucose [G] was measured by HPLC on an Hewlett-Packard 1100 instrument using an Aminex-HPX-78H (300x7.8 mm) column (Biorad, Reinach, Switzerland) combined with a LiChrospher-100-RP18, 5 μm precolumn (Merck, Darmstadt, Germany). The mobile phase was 0.004 M sulfuric acid, pumped at a flow rate of 0.6 ml/min. A refractive index detector was used to monitor the signal. External standard calibration was applied based on peak areas.

[0126] The yield, i.e. the amount of biomass (dry weight) obtained from conversion of D-glucose [g biomass/g D-glucose] after a given time was calculated as follows, wherein (t1) defines the concentration of biomass and D-glucose, respectively, measured during or at the end of cultivation, for instance after 24 or 48 h, and (t0) defines the respective concentrations at the beginning of the cultivation (see above):

\[ \frac{[B(t_1)] - [B(t_0)]}{[G(t_1)] - [G(t_0)]} = \text{yield of biomass on converted glucose} \ [\text{g/kg}] \]

[0127] The volumetric biomass productivity [g/kg/h], i.e. the amount of biomass produced per liter of culture and per time unit was calculated as follows:

\[ \frac{[B(t_1)] - [B(t_0)]}{(t_1 - t_0)} = \text{volumetric biomass productivity} \ [\text{g/kg/h}] \]

Example 2
Preparation of Chromosomal DNA and Amplification of DNA Fragment by PCR

[0128] Chromosomal DNA of *Glucanobacter oxydans* DSM 17078 was prepared from the cells cultivated at 30°C for 1 day in mannitol broth (MB) liquid medium consisting of 25 g/l mannitol, 5 g/l yeast extract (Difco), and 3 g/l of Bactopeptone (Difco) by the method described by Sambrook et al. (1989) “Molecular Cloning: A Laboratory Manual, Second Edition”, Cold Spring Harbor Laboratory Press).

[0129] A DNA fragment was prepared by PCR with the chromosomal DNA prepared above and a set of primers, Pf (SEQ ID NO:3) and Pr (SEQ ID NO:4). For the reaction, the Expand High Fidelity PCR kit (Roche Diagnostics) and 10 μg of the chromosomal DNA was used in total volume of 100 μl according to the supplier’s instruction to have the PCR product containing GMS 01 DNA sequence (SEQ ID NO:1). The PCR product was recovered from the reaction and its correct sequence confirmed.

Example 3
Disruption of the GMS 01 Gene in *G. oxydans* DSM 17078

[0130] The PCR product obtained in Example 2 was cloned in an *E. coli* vector pCR2.1-TOP and used to transform *E. coli* TG1 to have a Apr transformant carrying pCR2.1-GMS 01.

[0131] Then, Km cassette isolated from pUC-4K (Amersham Bioscience, accession No. X06404) was inserted into one of the restriction site of the target gene with ligase and the resulting ligation product was used to transform *E. coli* TG1 to have Apr Km transformant carrying pCR2.1-GMS 01:

Km. The pCR2.1-GMS 01::Km plasmid prepared from the transformant was digested by two restriction enzymes selected from the multi-cloning site of the vector part to isolate a DNA fragment containing GMS 01::Km. The resulting DNA fragment was used to transform *G. oxydans* DSM 17078 by electroporation to have the gene disruptant, *G. oxydans* DSM 17078-GMS 01::Km.

Example 4
Production of Biomass from D-Glucose in Liquid Cultures

[0132] Cells of *G. oxydans* DSM 17078 and *G. oxydans* DSM 17078-GMS 01::Km were cultivated as described above and the yield of produced biomass and volumetric biomass productivity determined (see Example 1), wherein (t1) = 24 h. The results are depicted in Table 1.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Yield of biomass on consumed glucose [g/kg]</th>
<th>Volumetric biomass productivity [g/kg/h]</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>G. oxydans</em> DSM 17078</td>
<td>0.032</td>
<td>0.052</td>
</tr>
<tr>
<td><em>G. oxydans</em> DSM 17078-GMS 01::Km</td>
<td>0.24</td>
<td>0.109</td>
</tr>
</tbody>
</table>

Example 5
Presence of the GMS 01 Gene and Equivalents in Other Organisms

[0133] The presence of SEQ ID NO:1 and/or equivalents in other organisms than the ones disclosed herein before, e.g. organisms as mentioned in Table 2, may be determined by a simple DNA hybridization experiment.

[0134] Strains of *Acetobacter aceti* subsp. *xylinum* IFO 13693 and IFO 13773 are grown at 27°C for 3 days on No. 350 medium containing 5 μl Bactopeptone (Difco), 5 g/l yeast extract (Difco), 5 g/l glucose, 5 g/l mannitol, 1 μl MgSO4, 7H2O, 5 ml/l ethanol, and 15 μl agar. All other *Acetobacter, Glucanacetobacter* and all *Glucanobacter* strains are grown at 27°C for 3 days on mannitol broth (MB) agar medium containing 25 g/l mannitol, 5 μl yeast extract (Difco), 3 g/l Bactopeptone (Difco), and 18 g/l agar (Difco). *E. coli* K-12 is grown on Luria Broth agar medium. The other strains are grown on medium recommended by the suppliers or according to methods known in the art. Genomic DNA is extracted as described by e.g. Sambrook et al., 1989, “Molecular Cloning: A Laboratory Manual, Second Edition”, Cold Spring Harbor Laboratory Press) from a suitable organism as, e.g. mentioned in Table 2.

[0135] Genomic DNA preparations are digested with restriction enzymes such as EcoRI or HindIII, and 1 μg of the DNA fragments are separated by agarose gel electrophoresis (1% agarose). The gel is treated with 0.25 N HCl for 15 min and then 0.5 N NaOH for 30 min, and then blotted onto nitrocellulose or a nylon membrane with Vacuum Blotter Model 785 (BIO-RAD Laboratories AG, Switzerland) according to the instruction of the supplier. The resulting blot is then brought into contact/hybridized with a solution wherein the probe, such as e.g. a DNA fragment with SEQ ID NO:1 sequence or a DNA fragment containing the part or
whole of the SEQ ID NO: 1 sequence to detect positive DNA fragment(s) from a test organism. A DIG-labeled probe, e.g. SEQ ID NO: 1, may be prepared according to Example 1 by using the PCR-DIG labeling kit (Roche Diagnostics) and a set of primers, SEQ ID NO:3 and SEQ ID NO:4. A result of such a blot is depicted in Table 2.

[0136] The hybridization may be performed under stringent or highly stringent conditions. A preferred, non-limiting example of such conditions are hybridization in 6x sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 1xSSC, 0.1% SDS at 50°C, preferably at 55°C, more preferably at 60°C and even more preferably at 65°C. Highly stringent conditions include, for example, 2 hours to 4 days incubation at 42°C in a solution such as DigEasyHyb solution (Roche Diagnostics GmbH) with or without 100 µg/ml salmon sperm DNA, or a solution comprising 50% formamide, 5xSSC (150 mM NaCl, 15 mM trisodium citrate), 0.02% sodium dodecyl sulfate, 0.1% N-lauroylsarcosine, and 2% blocking reagent (Roche Diagnostics GmbH), followed by washing the filters twice for 5 to 15 min in 2xSSC and 0.1% SDS at room temperature and then washing twice for 15-30 min in 0.5xSSC and 0.1% SDS or 0.1xSSC and 0.1% SDS at 65-68°C. To detect DNA fragments with lower identity to the probe DNA, final washing steps can be done at lower temperatures such as 50-65°C and for shorter washing time such as 1-15 min.

[0137] The genes corresponding to the positive signals within the respective organisms shown in Table 2 can be cloned by a PCR method well known in the art using genomic DNA of such an organism together with a suitable primer set, such as e.g. SEQ ID NO:3 and SEQ ID NO:4 under conditions as described in Example 1 or as follows: 5 to 100 ng of genomic DNA is used per reaction (total volume 50 µl). Expand High Fidelity PCR system (Roche Diagnostics) can be used with reaction conditions consisting of 94°C for 2 min; 30 cycles of (i) denaturation step at 94°C for 15 sec, (ii) annealing step at 60°C for 30 sec, (iii) synthesis step at 72°C for 30 sec or 0.5 to 5 min depending on the target DNA length (1 min/1 kb); extension at 72°C for 7 min. Alternatively, one can perform a PCR with degenerate primers, which can be synthesized based on SEQ ID NO:2 or amino acid sequences as consensus sequences selected by aligning several amino acid sequences obtained by a sequence search program such as BLASTP (or BLASTX when nucleotide sequence is used as a "query sequence") to find proteins having a similarity to the protein of SEQ ID NO:2. For PCR using degenerate primers, temperature of the second annealing step (see above) can be lowered to 55°C, or even to 50-45°C. A result of such an experiment is shown in Table 2.

[0138] Samples of the PCR reactions are separated by agarose gel electrophoresis and the bands are visualized with a transilluminator after staining with e.g. ethidium bromide, isolated from the gel and the correct sequence is confirmed.

[0139] Consensus sequences mentioned above might be amino acid sequences belonging to certain categories of several protein domain/family databases such as PROSITE (database of protein families and domains), COGs (Cluster of Ortholog Groups), CDD (Conserved Domain Databases), pfam (large collection of multiple sequence alignments and hidden Markov models covering many common protein domains and families). Once one can select certain protein with identical/similar function to the protein of this invention from proteins containing domain or family of such databases, corresponding DNA encoding the protein can be amplified by PCR using the protein sequence or its nucleotide sequence when it is available in public databases. Following organisms may further provide genes, which can be used as an alternative gene of this invention: Burkholderia xenovorans LB400, Acinetobacter calscoeticus LMD79.41, Mesorhizobium loti MAFF303099, Sinorhizobium melloti 1021.

Example 6

Disruption of the GMS 01 Gene and Equivalents in Other Organisms for Production of Biomass

[0140] In order to improve biomass production in a suitable microorganism from a carbon source such as e.g. glucose, the GMS 01 gene and equivalents as e.g. a PCR product obtained in Example 5, referred to hereafter as gene X, can be disrupted in accordance to the GMS 01 gene in G. oxydans DSM 17078 (see Example 2) to generate a knockout mutant carrying GMS 01 equivalent gene::Kn. Suitable host strains for generation of such knockout mutants may be selected from e.g. Gluconobacter strains listed in Table 2, in particular e.g. G. oxydans IFO 3292, G. oxydans ATCC 621H, G. oxydans IFO 12528, G. oxydans IFO 3291, G. oxydans IFO 3255, G. oxydans IFO 3244, G. cerinus IFO 3266, G. frateurii IFO 3269, G. oxydans IFO 3287, Acetobacter acetii subsp. orleaus IFO 3259, Acetobacter acetii subsp. xylinum IFO 13693, Acetobacter acetii subsp. xylinum IFO 13773 and Acetobacter sp. ATCC 15164.

[0141] The knockout mutant such as, e.g. a knockout mutant G. oxydans IFO 3292-GMS 01 equivalent gene::Kn can be generated as follows: the PCR product obtained from G. oxydans IFO 3292 described in Example 5 is cloned in an E. coli vector pCR2.1-TOPO and used to transform E. coli TG1 to have a Ap’ transformant carrying pCR2.1-gene X. Then, Kn’ cassette isolated from pUC-4K (Amersham BioScience, accession No. X06404) is inserted into one of the restriction site of the target gene with ligase and the resulting ligation product is used to transform E. coli TG1 to have Ap’ Kn’ transformant carrying pCR2.1-gene X::Kn. The pCR2.1-gene X::Kn plasmid prepared from the transformant is digested by two restriction enzymes selected from the multiple cloning site of the vector part to isolate a DNA fragment containing gene X::Kn. The resulting DNA fragment is used to transform the host strain carrying the GMS 01 equivalent gene by electroporation to have the gene disruptant carrying GMS 01 equivalent gene::Kn.

[0142] Further modifications including genes involved in the conversion of a carbon source such as e.g. glucose into biomass within said strains may be generated to improve biomass production within such strains.

[0143] Production of biomass from a carbon source such as e.g. glucose using the cells of the knockout mutant, e.g. G. oxydans IFO 3292-GMS 01 equivalent gene::Kn, and the corresponding wild-type strain, e.g. G. oxydans IFO 3292, are performed according to Example 4.

[0144] In a reaction with 5% glucose as carbon source, the mutant strain can produce at least more than 250% biomass compared to the wild-type strain.
TABLE 2

<table>
<thead>
<tr>
<th>Strain</th>
<th>Signal 1</th>
<th>Signal 2</th>
<th>Signal 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. oxydans DSM 17078</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 3293</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>G. oxydans IPO 3292</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans ATCC 621H</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 12528</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>G. oxydans G 624</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans T-100</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 3291</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 3255</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans ATCC 9937</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 3244</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. cerinus IPO 3206</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. frateurii IPO 3260</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>G. oxydans IPO 3287</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter acetii subsp. oleumus IPO 3259</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter acetii subsp. xylinum IPO 13693</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter acetii subsp. xylinum IPO 13773</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter sp. ATCC 15164</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G. thailandensis NBRC 100600</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gluconacetobacter liquefaciens ATCC 14835</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gluconacetobacter polyoxogenes NBI 1028</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

TABLE 2-continued

<table>
<thead>
<tr>
<th>Strain</th>
<th>Signal 1</th>
<th>Signal 2</th>
<th>Signal 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gluconacetobacter diazotrophicus ATCC 49037</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gluconacetobacter europaeus DSM 6160</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter acetii 1023</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Acetobacter passeurianus NC1193</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pseudomonas putida ATCC 21812</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa PA01</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pseudomonas fluorescens DSM 50106</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pseudomonas syringae B728a</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Rhodopseudomonas palustris CGA009</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pantoea citrea 10856R</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E. coli K12</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Aspergillus niger</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mouse</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Signal 1: Detection of DNA on a blot with genomic DNA of different strains and SEQ ID NO: 1 as labeled probe. Signal 2: Detection of DNA of different strains in a PCR reaction using primer pair SEQ ID NO: 2 and SEQ ID NO: 3. Signal 3: Detection of DNA of different strains in a PCR reaction using degenerate primers. For more explanation refer to the text.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 4
<210> SEQ ID NO 1
<211> LENGTH: 2427
<212> TYPE: DNA
<213> ORGANISM: Gluconobacter oxydans DSM 17078

<400> SEQUENCE: 1
atgaagcaac acactccgcc gggctcagt ggccgctcgg accgccgctt attccggctt 60
tgcccggag tccatcctt tggcgccgca tggctcgctt ccactgggccc cctctcttt 120	tagtcatcct tgggctgctg actctggcgc acggctctc ccctctcttc ggcgaaccgg 180	gcctgctcc actctgcgc actctgcgc tcctgaaaaag cctctgcggg aaccaccatt 240
tgctgctgg cactctggcg cccctgctc ggctctggct cagcctctctt caccagcggc 300	tggctgtctg tgcggcttgct tccgcaagcc atgggagggc cctggaaccgc gctcgtcgg 360
tgcgctggcg cctggtctcg gcgctgtgct cgggctctt ctccctctct cccgaccggc 420
cagcagatc gcggcagatt gcgacagac cttggcaacc cttccgcgcc ccgaccgccg 480
cggccggtcg cgccgagtc gcggcagctt cttggcaacc cttccgcgcc ccgaccgccg 540
cggcggggcc gcggcagctt cttggcaacc cttccgcgcc ccgaccgccg 600
acccgggtg ttgtaactt acggcagcgg gcgcgccgag ccagcgccga ccgaccgccg 660
ggcatcacc gcagcgttta ttcgtcggc ctcgccctgg ctcgcttgcc ggtggtcgtt 720
gctgcgctcc gcgctggacg gcgctggagct ttcgcttcct tggcagcttc gctcacccccc 780
cctgctggct cgccgcggtc cgccgcggtc ccggcgcggtc ccggcgcggtc gctccgtgg 840
gggctcggcg ctcgcaagcc ccctgctgct ctcgcaagcc ccctgctgct ctcgcaagcc 900
ttcgggtcag tcctgtgcgc ctcgcaagcc ctcgcaagcc ccctgctgct ctcgcaagcc 960
-continued

tcgatctgc gccctccga cccgctctac acgacgcccc gcagcatctga ggcgaactgcc
1020
cgcgccgca tcaaggcaca ggtgatcate gcgaacagcgc ctcaacgca taacggttctg
1080
gtcagcaggg ctcgggagcc accgagccgta ttgcaagcttc acaacgccca ggcggcttgg
1140
gtcgctgatc gtcctcgcgc gatcctcggc atgacgatcc cctcgcttctc
1200
cacccgacat cccccgactc tctgatcggtc tcgctctcat acaggaagtaa gaaacctcgctg
1260
tacatccgca cggggtccttg actcccccagc cggggtggcg gcagcaccac gaaggacttc
1320
gacgttttcg ctcggtgtat cggcgcctgtg aacgagcata cggggaagct cggctggttc
1380
taccagacgc tctatacagcg tctctgaggac agtagacgctc cgctccagcgc gacccctttg
1440
gacgttaagc gccagaaggg cccgctgtttgc cgcggccact atcgctcgcac gcgaacgcgc
1500
gcacattctgg ctcgtcagcc gcagctcgcct ccggccttggc gggagaccccctc
1560
gttccaggg gcggctgctgc ggccgatcae accagccgga ccagcagcat gtcgcagctg
1620
acctctgctg cgaagccacc gcgtcaccgc ctgctctacat gcggggtgac gcgtcctgac
1680
cagatgcttc gcagcagctc tttccagcaga cgctcctgctg aacgagccctt cggcgccgog
1740
tcgctcaaggt cgtctcctat cttctgcgggc gcagctgagag gggtggtctg
1800
gcgcgatggt gcgctggttgc gggcgcctaca ttcctggctgc gtcgtgttccc
1860
cagcttgctgcc cccgctcgac gcggacctgg cccgctgtggc aaaaaagcgc gcgggagcatc
1920
ggtggcggaa cggcgcctgc gcacattctg atcgctgctgc cgccgcttggc gcgcgtcttgc
1980
ttccttggct gcgctgtgctg gcgtcctgcgc ttcacagtcgc gcgcgtgctgc gcgcgtcttgc
2040
ggcgtagctc cggtgatgta cctgcaagcc aacaaggtgct ctgtgcagcgc gcgcctgcgc
2100
acccgctgctgcc acgcgaagatgt atgcggctgct gcgctgctgc gcgcgtcttgc
2160
sgttgctgcgc gcgctgctgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc
2220
tccatggct acctcagctcg tcgctgctgcc gcgcctgcgc gcgcgtcttgc gcgcgtcttgc
2280
gcgctgccgc gcgctgctgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc
2340
atggctgctgcc gcgcctgcgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc
2400
gctgctgcgc gcgctgctgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc gcgcgtcttgc
2467

<210> SEQ ID NO 2
<211> LENGTH: 808
<212> TYPE: PRT
<213> ORGANISM: Gluconobacter oxydans DSM 17076
<400> SEQUENCE: 2

Mets Ser Thr Thr Ser Arg Pro Gly Leu Trp Ala Leu Ile Thr Ala Ala
1    5    10    15
Val Phe Ala Leu Cys Gly Ala Ile Leu Thr Val Gly Gly Ala Trp Val
20   25    30
Ala Ala Ile Gly Gly Pro Leu Tyr Tyr Val Ile Leu Gly Leu Ala Leu
35   40    45
Leu Ala Thr Ala Phe Leu Ser Phe Arg Arg Asn Pro Ala Ala Leu Tyr
50   55    60
Leu Phe Ala Val Val Phe Gly Thr Val Ile Trp Glu Leu Thr Ile
65   70    75    80
Val Gly Leu Arg Ile Trp Ala Leu Ile Pro Ser Arg Ser Ile Val Ile
85    90    95
<table>
<thead>
<tr>
<th>Ile</th>
<th>Leu</th>
<th>Gly</th>
<th>Ile</th>
<th>Trp</th>
<th>Leu</th>
<th>Leu</th>
<th>Leu</th>
<th>Pro</th>
<th>Phe</th>
<th>Val</th>
<th>Ser</th>
<th>Arg</th>
<th>Gln</th>
<th>Ile</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Thr</td>
<td>Arg</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Leu</td>
<td>Pro</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
<td>Val</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td>Ala</td>
<td>Leu</td>
<td>Phe</td>
<td>Ala</td>
<td>Ser</td>
<td>Leu</td>
<td>Phe</td>
<td>Thr</td>
<td>Asp</td>
<td>Pro</td>
<td>His</td>
<td>Asp</td>
<td>Ile</td>
<td>Ser</td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Glu</td>
<td>Leu</td>
<td>Pro</td>
<td>Thr</td>
<td>Gln</td>
<td>Ile</td>
<td>Ala</td>
<td>Asn</td>
<td>Ala</td>
<td>Ser</td>
<td>Pro</td>
<td>Ala</td>
<td>Asp</td>
<td>Pro</td>
<td>Asp</td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
<td>Glu</td>
<td>Trp</td>
<td>His</td>
<td>Ala</td>
<td>Tyr</td>
<td>Gly</td>
<td>Arg</td>
<td>Thr</td>
<td>Gln</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Arg</td>
<td>Trp</td>
<td>Ser</td>
<td>Pro</td>
<td>Leu</td>
<td>Asn</td>
<td>Gln</td>
<td>Ile</td>
<td>Asn</td>
<td>Ala</td>
<td>Ser</td>
<td>Asn</td>
<td>Val</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Lys</td>
<td>Val</td>
<td>Ala</td>
<td>Trp</td>
<td>His</td>
<td>Ile</td>
<td>His</td>
<td>Thr</td>
<td>Lys</td>
<td>Asp</td>
<td>Met</td>
<td>Met</td>
<td>Asn</td>
<td>Ser</td>
<td>Asn</td>
</tr>
<tr>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Pro</td>
<td>Gly</td>
<td>Glu</td>
<td>Ala</td>
<td>Thr</td>
<td>Asn</td>
<td>Glu</td>
<td>Ala</td>
<td>Thr</td>
<td>Pro</td>
<td>Ile</td>
<td>Glu</td>
<td>Phe</td>
<td>Asn</td>
<td>Asn</td>
</tr>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Leu</td>
<td>Tyr</td>
<td>Met</td>
<td>Cys</td>
<td>Ser</td>
<td>Leu</td>
<td>His</td>
<td>Gln</td>
<td>Lys</td>
<td>Leu</td>
<td>Phe</td>
<td>Ala</td>
<td>Val</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Thr</td>
<td>Gly</td>
<td>Asn</td>
<td>Val</td>
<td>Lys</td>
<td>Trp</td>
<td>Val</td>
<td>Tyr</td>
<td>Asp</td>
<td>Pro</td>
<td>Lys</td>
<td>Leu</td>
<td>Gln</td>
<td>Ile</td>
<td>Asn</td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Gly</td>
<td>Phe</td>
<td>Glu</td>
<td>His</td>
<td>Leu</td>
<td>Thr</td>
<td>Cys</td>
<td>Arg</td>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
<td>Phe</td>
<td>His</td>
<td>Glu</td>
<td>Thr</td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td>Asn</td>
<td>Ala</td>
<td>Thr</td>
<td>Asp</td>
<td>Ser</td>
<td>Asp</td>
<td>Gly</td>
<td>Asn</td>
<td>Pro</td>
<td>Ala</td>
<td>Pro</td>
<td>Thr</td>
<td>Asp</td>
<td>Cys</td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td>285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Lys</td>
<td>Arg</td>
<td>Ile</td>
<td>Ile</td>
<td>Leu</td>
<td>Pro</td>
<td>Val</td>
<td>Asn</td>
<td>Am</td>
<td>Gly</td>
<td>Arg</td>
<td>Leu</td>
<td>Val</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td>Asp</td>
<td>Thr</td>
<td>Gly</td>
<td>Lys</td>
<td>Thr</td>
<td>Cys</td>
<td>Ser</td>
<td>Gly</td>
<td>Phe</td>
<td>Gly</td>
<td>Asn</td>
<td>Asn</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td>300</td>
<td>310</td>
<td>315</td>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Asp</td>
<td>Leu</td>
<td>Arg</td>
<td>Val</td>
<td>Pro</td>
<td>Asn</td>
<td>Gln</td>
<td>Pro</td>
<td>Tyr</td>
<td>Thr</td>
<td>Thr</td>
<td>Pro</td>
<td>Gly</td>
<td>Gln</td>
<td>Tyr</td>
</tr>
<tr>
<td>325</td>
<td>330</td>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Pro</td>
<td>Thr</td>
<td>Ser</td>
<td>Pro</td>
<td>Val</td>
<td>Ile</td>
<td>Thr</td>
<td>Asp</td>
<td>Lys</td>
<td>Leu</td>
<td>Ile</td>
<td>Ile</td>
<td>Ala</td>
<td>Ann</td>
<td></td>
</tr>
<tr>
<td>340</td>
<td>345</td>
<td>350</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ala</td>
<td>Ile</td>
<td>Thr</td>
<td>Asp</td>
<td>Asn</td>
<td>Gly</td>
<td>Ser</td>
<td>Val</td>
<td>Lys</td>
<td>Gln</td>
<td>Ala</td>
<td>Ser</td>
<td>Gly</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td>365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Ala</td>
<td>Phe</td>
<td>Asp</td>
<td>Val</td>
<td>Tyr</td>
<td>Thr</td>
<td>Gly</td>
<td>Arg</td>
<td>Val</td>
<td>Trp</td>
<td>Val</td>
<td>Phe</td>
<td>Asp</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td>380</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Asn</td>
<td>Pro</td>
<td>Asp</td>
<td>Pro</td>
<td>Asn</td>
<td>Gln</td>
<td>Leu</td>
<td>Pro</td>
<td>Asp</td>
<td>Asp</td>
<td>Ser</td>
<td>His</td>
<td>Pro</td>
<td>Val</td>
<td>Phe</td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td>395</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Pro</td>
<td>Asn</td>
<td>Ser</td>
<td>Pro</td>
<td>Asn</td>
<td>Ser</td>
<td>Tyr</td>
<td>Thr</td>
<td>Asp</td>
<td>Arg</td>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>405</td>
<td>410</td>
<td>415</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Leu</td>
<td>Val</td>
<td>Tyr</td>
<td>Ile</td>
<td>Pro</td>
<td>Met</td>
<td>Gly</td>
<td>Val</td>
<td>Gly</td>
<td>Thr</td>
<td>Pro</td>
<td>Asp</td>
<td>Gln</td>
<td>Trp</td>
</tr>
<tr>
<td>420</td>
<td>425</td>
<td>430</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Asp</td>
<td>Arg</td>
<td>Thr</td>
<td>Lys</td>
<td>Arg</td>
<td>Asp</td>
<td>Ser</td>
<td>Glu</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Pro</td>
<td>Gly</td>
<td>Ile</td>
</tr>
<tr>
<td>435</td>
<td>440</td>
<td>445</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Asn</td>
<td>Ala</td>
<td>Asp</td>
<td>Thr</td>
<td>Gly</td>
<td>Lys</td>
<td>Leu</td>
<td>Ala</td>
<td>Trp</td>
<td>Phe</td>
<td>Tyr</td>
<td>Gln</td>
<td>Thr</td>
<td>Val</td>
</tr>
<tr>
<td>450</td>
<td>455</td>
<td>460</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>His</td>
<td>Asp</td>
<td>Leu</td>
<td>Trp</td>
<td>Asp</td>
<td>Met</td>
<td>Asp</td>
<td>Pro</td>
<td>Ser</td>
<td>Gln</td>
<td>Pro</td>
<td>Ser</td>
<td>Leu</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>465</td>
<td>470</td>
<td>475</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Val</td>
<td>Thr</td>
<td>Gln</td>
<td>Lys</td>
<td>Asp</td>
<td>Gly</td>
<td>Thr</td>
<td>Leu</td>
<td>Val</td>
<td>Pro</td>
<td>Ala</td>
<td>Ile</td>
<td>Tyr</td>
<td>Ala</td>
<td>Pro</td>
</tr>
<tr>
<td>485</td>
<td>490</td>
<td>495</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Lys</td>
<td>Thr</td>
<td>Gly</td>
<td>Asp</td>
<td>Ile</td>
<td>Phe</td>
<td>Val</td>
<td>Leu</td>
<td>Asp</td>
<td>Arg</td>
<td>Arg</td>
<td>Thr</td>
<td>Gly</td>
<td>Lys</td>
<td>Glu</td>
</tr>
</tbody>
</table>
 Continued

```
500  505  510
Ile Val Pro Ala Pro Glu Thr Pro Val Pro Gln Gly Ala Ala Pro Gly
515  520  525
Asp His Thr Ser Pro Thr Gln Pro Met Ser Gln Leu Thr Leu Arg Pro
530  535  540
Lys Asn Pro Leu Asn Asp Ser Ile Trp Gly Gly Thr Ile Phe Asp
545  550  555  560
Gln Met Phe Cys Ser Ile Tyr Phe His Arg Leu Arg Tyr Glu Gly Pro
565  570  575
Phe Thr Pro Pro Ser Leu Lys Gly Ser Leu Ile Phe Pro Gly Asp Leu
580  585  590
Gly Met Phe Glu Trp Gly Gly Leu Ala Val Asp Pro Gln Arg Glu Val
595  600  605
Ala Phe Ala Asn Pro Ile Ser Leu Pro Phe Val Ser Gln Leu Val Pro
610  615  620
Arg Gly Pro Gly Asn Pro Leu Trp Pro Glu Lys Asp Ala Lys Gly Thr
625  630  635  640
Gly Gly Glu Thr Gly Leu Gln His Asn Tyr Gly Ile Pro Tyr Ala Val
645  650  655
Asn Leu His Pro Phe Leu Asp Pro Val Leu Pro Phe Gly Ile Lys
660  665  670
Met Pro Cys Arg Thr Pro Pro Trp Gly Tyr Val Ala Gly Ile Asp Leu
675  680  685
Lys Thr Asn Lys Val Val Trp Gin His Arg Asn Gly Thr Leu Arg Asp
690  695  700
Ser Met Tyr Gly Ser Ser Leu Pro Ile Pro Leu Pro Ile Lys Ile
705  710  715  720
Gly Val Pro Ser Leu Gly Gly Pro Leu Ser Thr Ala Gly Asn Leu Gly
725  730  735
Phe Leu Thr Ala Ser Met Asp Tyr Ile Arg Ala Tyr Asn Leu Thr
740  745  750
Thr Gly Lys Val Leu Trp Gin Asp Arg Leu Pro Ala Gly Ala Glu Ala
755  760  765
Thr Pro Ile Thr Tyr Ala Ile Asn Gly Lys Gin Tyr Ile Val Thr Tyr
770  775  780
Ala Gly Gly His Asn Ser Phe Pro Thr Arg Met Gly Asp Asp Ile Ile
785  790  795  800
Ala Tyr Ala Leu Pro Asp Gln Lys
805
```

<210> SEQ ID NO 3
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer

<400> SEQUENCE: 3
atgagcacaacatcccgagc

<210> SEQ ID NO 4
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
1. A polynucleotide selected from the group consisting of:
   (a) polynucleotides encoding a polypeptide comprising the
       amino acid sequence according to SEQ ID NO:2;
   (b) polynucleotides comprising the nucleotide sequence
       according to SEQ ID NO:1;
   (c) polynucleotides comprising a nucleotide sequence
       obtainable by nucleic acid amplification such as
       polymerase chain reaction, using genomic DNA from a
       microorganism as a template and a primer set according
       to SEQ ID NO:3 and SEQ ID NO:4;
   (d) polynucleotides comprising a nucleotide sequence
       encoding a fragment or derivative of a polypeptide
       encoded by a polynucleotide of any of (a) to (c) wherein
       in said derivative one or more amino acid residues are
       conservatively substituted compared to said polypeptide,
       and said fragment or derivative has the activity of an
       oxidoreductase [EC 1], preferably an oxidoreductase
       acting on the CH—OH group of donors [EC 1.1];
   (e) polynucleotides the complementary strand of which
       hybridizes under stringent conditions to a polynucleo-
       tide as defined in any one of (a) to (d) and which encode
       an oxidoreductase [EC 1], preferably an oxidoreductase
       acting on the CH—OH group of donors [EC 1.1]; and
   (f) polynucleotides which are at least 70%, such as 85, 90
       or 95% homologous to a polynucleotide as defined in
       any one of (a) to (d) and which encode an oxidoreductase
       [EC 1], preferably an oxidoreductase acting on the
       CH—OH group of donors [EC 1.1];
   or the complementary strand of such a polynucleotide.

2. A polypeptide encoded by a polynucleotide according to
   claim 1.

3. A microorganism comprising a polypeptide according to
   claim 2 having reduced or abolished oxidoreductase [EC 1]
   activity.

4. A microorganism according to claim 3 comprising a
   gene with reduced or abolished activity.

5. A microorganism according to claim 4 wherein the gene
   is disrupted.

6. A microorganism according to claim 3 capable of produc-
   ing biomass from glucose at a yield of at least about 0.2 g
   biomass/g glucose.

7. A microorganism according to claim 6 wherein the yield
   is at least about 0.3.

8. Process for producing cells according to claim 3, com-
   prising the step of genetically engineering cells comprising
   a polynucleotide according to claim 1.

9. Use of a disrupted polynucleotide according to claim 1
   for the production of biomass from a carbon source.

10. A microorganism according to claim 3, said microor-
    ganism being genetically altered in such a way that it leads
    to an improved yield and/or production of biomass from a
    carbon source produced by said microorganism.

11. A microorganism according to claim 3 selected from the
    group consisting of Pseudomonas, Pantoea, Escherichia,
    Corynebacterium, Ketogulonicigenium and acetic acid bac-
    teria like e.g., Gluconobacter, Acetobacter or Gluconacetobacter,
    preferably Acetobacter sp., Acetobacter aceti, Glu-
    conobacter frateurii, Gluconobacter cerinus, Gluconobacter
    thailandicus, Gluconobacter oxydans, preferably Glucono-
    bacter oxydans, more preferably Gluconobacter oxydans
    DSM 17078.

12. Use of a microorganism according to claim 3 for the
    production of biomass from a carbon source.

13. Process for the production of a disrupted endogenous
    oxidoreductase [EC 1], preferably oxidoreductase acting on
    the CH—OH group of donors [EC 1.1] gene in a microor-
    ganism, said microorganism comprising a polynucleotide
    according to claim 1, said process comprising the step of
    altering said polynucleotide in such a way that it leads to an
    improved yield and/or efficiency of biomass production
    produced from a carbon source by said microorganism.

14. Process for the production of a microorganism capable
    of growth on a given carbon source, comprising the step of
    altering said microorganism comprising a gene comprising a
    polynucleotide according to claim 1 so that the microor-
    ganism produces a polypeptide having reduced or abolished
    oxidoreductase [EC 1] activity, leading to an improved yield
    and/or production of biomass from said carbon source by said
    microorganism.

15. Process for the production of a microorganism contain-
    ing an endogenous gene comprising a polynucleotide accord-
    ing to claim 1, comprising the step of altering said microor-
    ganism so that the endogenous gene is underexpressed or
    disrupted, leading to an improved yield and/or production of
    biomass from a carbon source produced by said microorgan-
    ism.

16. Process according to claim 14 for the production of a
    microorganism.

17. Microorganism obtainable by a process according to
    claim 14.

18. Process for the production of biomass with a microor-
    ganism according to claim 3 wherein said microorganism is
    cultivated in an aqueous nutrient medium under conditions
    that allow production of biomass from a carbon source.

19. Process according to claim 8 wherein the carbon source
    is selected from the group consisting of glucose, sucrose,
    maltose, starch, cellulose, cellobiose, lactose, isomaltose,
    dextran, trehalose and mixtures thereof.

20. Use according to claim 9 wherein the carbon source is
    selected from the group consisting of glucose, sucrose, mal-
    tose, starch, cellulose, cellobiose, lactose, isomaltose, dext-
    ran, trehalose and mixtures thereof.

21. A microorganism according to claim 10 wherein the
    carbon source is selected from the group consisting of glu-
    cose, sucrose, maltose, starch, cellulose, cellobiose, lactose,
    isomaltose, dextran, trehalose and mixtures thereof.

* * * * *