

VARIANT BACTERIAL STRAINS AND PROCESSES FOR PROTEIN OR BIOMASS PRODUCTION

FIELD OF THE INVENTION

5 The present invention relates to the production of protein and/or other macromolecules using microorganisms. In particular, the invention relates to novel bacterial strains and continuous culture processes for the production of protein or biomass using bacteria wherein gases and minerals are supplied to the cells. The invention also relates to the products of these processes and use of these products in e.g., food or feed.

10

BACKGROUND OF THE INVENTION

Growing world population, climate change and shortage of water increasingly pose a threat to traditional agriculture and thus sufficient supply of food and feed. Therefore, alternative sources of organic molecules, such as proteins, are being investigated. A potential alternative is single cell production, i.e. the production of protein and/or other macromolecules using microorganisms.

Chemoautotrophic microorganisms have been described which are able to grow on minimal mineral medium with hydrogen gas as the energy source and carbon dioxide as the only carbon source. For a review of these microorganisms, see e.g. Shively et al. (1998) Annu Rev Microbiol 52:191. Patent application WO2018144965 describes various microorganisms and bioprocesses for converting gaseous substrates into high-protein biomass. Andersen et al. (1979) Biochim Biophys Acta 585:1-11 describes mutant strains of *Alcaligenes eutrophus*, a hydrogen bacterium that grows readily under heterotrophic and autotrophic conditions. Mutants having altered ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) activity were characterised. Ohmiya et al. (2003) J. Biosci. Bioeng. 95:549-561 reviews the application of microbial genes to recalcitrant biomass utilization. Yu Jian et al. (2013) Int J Hydrogen Ener 38:8683-8690 describes carbon dioxide fixation by a hydrogen-oxidizing bacterial isolate. A high energy efficiency of 50% was measured under a moderate oxygen concentration (10 mol%).

However, various chemoautotrophic microorganisms have different properties in terms of growth rate, yield, biomass composition as well as properties related to being used as a food ingredient such as safety in human consumption, taste, smell, mouth-feel, technical and functional properties in cooking, etc. Not every chemoautotrophic microorganism has sufficient growth rate and provides sufficient yield and not every process can realistically be upscaled to an economically viable large-scale process. In order to have sufficient output of functional protein, e.g. for food or feed applications,

it is important to find a suitable production organism and a suitable process which can be performed at large scale. This need is addressed by the present invention.

Under metabolic stress, such as nitrogen limitation, bacteria may store energy in the form of storage polymers called polyhydroxyalkanoic acids (PHAs) (Rehm and Steinbüchel, 1999 Int J Biol Macromol 25:3-19). When bacteria are grown in a large bioreactor in an industrial setting, gas exchange often becomes insufficient and leads to increased PHA production. PHAs can be utilized as bioplastics, but when the bacteria are grown for other products than PHAs, its production is an unwanted outcome as it reduces the carbon yield. This problem is also address by the present invention.

10

SUMMARY OF THE INVENTION

The inventors found that chemoautotrophic bacteria, for example of the genus *Xanthobacter*, under certain conditions store energy in the form of PHAs. It was found that variant strains comprising a gene disruption of the *phaC1* gene produced almost no PHA under the same conditions, but retained favourable properties and suitability for processes for the production of biomass and/or protein.

In a first main aspect, the invention relates to a variant of bacterial strain VTT-E-193585 comprising a genetic modification that reduces the bacterial production of polyhydroxyalkanoic acid (PHA) as compared to strain VTT-E-193585.

In a further aspect, the invention relates to a process for the production of biomass and/or protein, said process comprising culturing a variant chemoautotrophic bacterial strain in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide and wherein said variant chemoautotrophic strain comprises a gene disruption of one or more genes encoding a PHA synthase. In further main aspects, the invention relates to bulk protein, biomass or non-protein cellular or chemical components obtained or obtainable by the process of the invention, and to a food or feed product obtained or obtainable by a process of the invention.

In a further aspect, the invention relates to a variant *Xanthobacter* strain comprising a gene disruption of one or more genes encoding a PHA synthase.

The invention also relates to methods for genetic modification of bacterial strain VTT-E-193585 and to genetically-modified variants of strain VTT-E-193585.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Optical density measured at 600 nm (black circles) and optical density probe readings during chemoautotrophic 200-L cultivation of isolated bacterial strain deposited as VTT-E-193585.

Figure 2. Optical density measured at 600 nm during parallel chemoautotrophic 200-mL cultivations of isolated bacterial strain deposited as VTT-E-193585 on different nitrogen sources.

Figure 3. Growth curve of *SoF1* (VTT-E-193585) (solid line) and *SoF1-2.0* (VTT E-213595) (dashed line) in autotrophic conditions.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

When used herein, the term "isolated", e.g. in the context of a strain, means isolated from its natural environment. Preferably, an isolated strain is pure, i.e. free of other strains.

The term "variant", when used herein in the context of a strain, refers to a strain which is derived from a reference strain, i.e. generated using the reference strain as starting point, and contains a genetic modification as compared to said reference strain. Genetic modifications include modifications include point mutations, as well as disruptions, such as insertions or deletions, of entire loci or fragments thereof. The variant preferably has fewer than 10 genetic modifications, e.g. fewer than 5, such as 4, 3, 2 or 1 genetic modification(s) compared to the reference strain. Preferably, the genome sequence of the variant strain is more than 90%, such as more than 95%, for example more than 99% identical to the genome sequence of the reference strain.

The term "chemoautotrophic" when used herein, refers to the ability to grow on minimal mineral medium with hydrogen gas as the energy source and carbon dioxide as the only carbon source.

When used herein, the noun "culture" refers to a suspension of viable cells in a liquid medium.

The term "biomass" has its usual meaning in the field of bacterial fermentation and refers to cellular material.

The term "continuous culture", when used herein, refers to a culturing process wherein fresh media is added continuously to the culture and media with bacterial culture is removed continuously at essentially the same rate.

Aspects and embodiments of the invention

Strain VTT-E-193585 has been isolated from the seashore of the Baltic sea in Naantali, Finland. This organism is able to grow in suitable bioreactor conditions with minimal mineral medium with hydrogen as the energy source and carbon dioxide as the carbon source at limited oxygen conditions. 16S sequencing and Illumina metagenomics sequencing have shown that the strain most likely is a member of the genus

Xanthobacter, but is not a known species. The bacterial strain is highly suitable for food and feed applications, because the dried cell powder has a high protein content and contains all the essential amino acids. It also contains more unsaturated than saturated fatty acids and a high level of B-group vitamins. The levels of peptidoglycans and lipopolysaccharides, which may cause allergy or toxicity, are low. A toxicity analysis was performed and no genotoxicity or cytotoxicity was observed for the strain. In addition, the strain is generally sensitive to antibiotics.

Strain VTT-E-193585 (*SoF1*) has been deposited on June 11th, 2019 in the VTT Culture Collection at the VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044 VTT, Finland, an International Depositary Authority under the Budapest Treaty. Further information on the characteristics of the strain and methods for culturing the strain are provided in the Examples herein and in European patent application EP19205786.7 (incorporated herein by reference).

The present invention relates *inter alia* to variants of VTT-E-193585, in particular variants comprising a genetic modification that reduces the bacterial production of polyhydroxyalkanoic acid (PHA), as well as more generally to chemoautotrophic bacteria having reduced production of polyhydroxyalkanoic acid (PHA), such as strains comprising a gene disruption of one or more genes encoding a PHA synthase.

The inventors have constructed genetically-modified variants of VTT-E-193585 comprising disruptions of the *phaC1* and/or *phaC2* loci. The variant comprising a gene disruption of *phaC1* has been deposited on April 19th, 2021 in the VTT Culture Collection at the VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044 VTT, Finland, an International Depositary Authority under the Budapest Treaty. The accession number is VTT E-213595. Further information on the characteristics of the strain and methods for culturing the strain are provided in the Examples herein.

In a first main aspect, the invention relates to a variant of bacterial strain VTT-E-193585 comprising a genetic modification that reduces the bacterial production of polyhydroxyalkanoic acid (PHA) as compared to strain VTT-E-193585. Thus, the invention relates to a genetically-modified variant, i.e. derivative, of bacterial strain VTT-E-193585. In other words, strain VTT-E-193585 further characterized in that it comprises a genetic modification.

In one embodiment, the genetic modification reduces bacterial PHA synthase activity as compared to strain VTT-E-193585, preferably wherein PHA synthase activity has been reduced to less than 10%, such as less than 5%, for example less than 2%.

In a further embodiment, the genetic modification reduces bacterial PHB production under autotrophic growth conditions to less than 10%, such as less than 5%, for example less than 2%. This can e.g. be determined by measuring the PHB dry content as described in Example 5 herein.

5 In one embodiment, the variant comprises a genetic modification reducing the expression level of *phaC1* and/or the activity of the *phaC1* enzyme.

In another embodiment, the variant comprises a genetic modification reducing the expression level of *phaC2* and/or the activity of the *phaC2* enzyme.

10 In one embodiment, the genetic modification is a gene disruption, such as an insertion and/or a deletion of the gene or part thereof.

In one embodiment, the variant comprises gene disruptions of both *phaC1* and *phaC2*.

In another embodiment, the variant comprises a gene disruption of *phaC1* but not of *phaC2*.

15 In one embodiment, the variant is the bacterial strain deposited under number VTT-E-213595, in which the *phaC1* gene has been disrupted.

In a preferred embodiment, the variant has retained the ability to grow using hydrogen gas as energy source and carbon dioxide as the only carbon source.

20 In one embodiment, if the strain is a variant of strain VTT-E-193585, the variant comprises the 16S ribosomal RNA set forth in SEQ ID NO:1 or a 16S ribosomal RNA having up to 20 nucleotide differences with SEQ ID NO:1, e.g. 1 to 10, such as 1 to 5, e.g. one, two or three nucleotide differences with SEQ ID NO:1.

SEQ ID NO:1. 16S ribosomal RNA sequence of strain VTT-E-193585:

25
 CTTGAGAGTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGA
 GCGCCCAGCAATGGGAGCGGCAGACGGGTGAGTAACGCGTGGGGATGTGCCCAATGGTACG
 GAATAACCCAGGGAACTTGGACTAATACCGTATGAGCCCTTCGGGGGAAAGATTTATCGCCA
 TTGGATCAACCCGCGTCTGATTAGCTAGTTGGTGGGGTAACGGCCCCACCAAGGCGACGATCA
 30 GTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGG
 GAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGTG
 TGATGAAGGCCTTAGGGTTGTAAAGCACTTTCGCCGGTGAAGATAATGACGGTAACCGGAGA
 AGAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCTAGCGTTGCTCG
 GAATCACTGGGCGTAAAGCGCACGTAGGCGGATCGTTAAGTCAGGGGTGAAATCCTGGAGCT
 35 CAACTCCAGAACTGCCCTTGATACTGGCGACCTTGAGTTCGAGAGAGGTTGGTGGAAGTGGC
 AGTGTAGAGGTGAAATTCGTAGATATTCGCAAGAACACCAAGTGGCGAAGGCGGCCAACTGGC
 TCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC

CACGCCGTAAACGATGGATGCTAGCCGTTGGGCAGCTTGCTGTTTCAGTGGCGCAGCTAACGC
 ATTAAGCATCCCGCCTGGGGAGTACGGTCGCAAGATTAAGCAAGGAATTGACGGGGGC
 CCGCACAAGCGGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGCAGAACCTTACCAGCCTTT
 GACATGGCAGGACGATTTCAGAGATGGATCTCTTCCAGCAATGGACCTGCACACAGGTGCT
 5 GCATGGCTGTCGTCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCC
 TCGCCTCTAGTTGCCAGCATTAGTTGGGCACTCTAGAGGGACTGCCGGTGATAAGCCGAGA
 GGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGGCTGGGCTACACACGTGCTACA
 ATGGTGGTGACAGTGGGATGCGAAAGGGCGACCTCTAGCAAATCTCCAAAAGCCATCTCAGT
 TCGGATTGTACTCTGCAACTCGAGTGCATGAAGTTGGAATCGCTAGTAATCGTGGATCAGCAT
 10 GCCACGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGGAGTTGGCTT
 TACCCGAAGGCGCTGCGCTAACCCGCAAGGGAGGCAGGCGACACGGTAGGGTCAGCGACT
 GGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTT

In a further aspect, the invention relates to a variant *Xanthobacter* strain comprising a
 15 gene disruption of one or more genes encoding a PHA synthase. Preferably, the strain
 is selected from the group consisting of: *X. agilis*, *X. aminoxidans*, *X. autotrophicus*, *X.*
flavus, *X. taetidis*, *X. viscosus*, *Xanthobacter sp. 126*, *Xanthobacter sp. 91* and strain
 VTT-E-193585.

In one embodiment, the variant *Xanthobacter* strain is a variant *X. autotrophicus*
 20 strain wherein the *phaC* gene (UniProtKB - A0A6C1KXK2) has been disrupted.

In another embodiment, the variant *Xanthobacter* strain is a variant *X. taetidis*
 strain wherein the *phaC* gene (UniProtKB - A0A3L7AJD5) has been disrupted.

In another embodiment, the variant *Xanthobacter* strain is a variant of a
Xanthobacter species wherein the gene in the genome of said strain that has the highest
 25 sequence identity to *phaC1* (SEQ ID NO: 62) has been disrupted. Methods for
 determining sequence identity are well-known in the art. Preferably, the encoded protein
 has more than 50% sequence identity, such as more than 60%, for example more than
 70%, such as more than 80%, for example more than 90%, such as more than 95%
 sequence identity to SEQ ID NO: 62.

In another embodiment, the variant *Xanthobacter* strain is a variant of strain *X.*
 30 *agilis*, *X. aminoxidans*, *X. flavus*, *X. viscosus*, *Xanthobacter sp. 126* or *Xanthobacter sp.*
91 wherein the gene in the genome of said strain that encodes the protein that has the
 highest sequence identity to *phaC1* (SEQ ID NO:62) has been disrupted. For example,
 in one embodiment, the variant is a variant of *X. agilis* wherein the gene of *X. agilis*,
 35 that has the highest sequence identity to *phaC1* (SEQ ID NO:62) of all genes in the *X.*
agilis genome has been disrupted.

In one embodiment, the variant *Xanthobacter* strain is a variant *X. tagetidis* strain wherein the gene set forth in NCBI ref. MBB6309058.1 has been disrupted.

In one embodiment, the variant *Xanthobacter* strain is a variant *X. flavus* strain wherein the gene set forth in NCBI ref. MBP2147722.1 has been disrupted.

5 In one embodiment, the variant *Xanthobacter* strain is a variant *X. autotrophicus* wherein the gene set forth in NCBI ref. WP_138398147.1 has been disrupted.

In one embodiment, the variant *Xanthobacter* strain is a variant *X. tagetidis* strain wherein the gene set forth in NCBI ref. WP_210210858.1 has been disrupted.

10 In one embodiment, the variant *Xanthobacter* strain is a variant *X. flavus* strain wherein the gene set forth in NCBI ref. WP_209489961.1 has been disrupted.

In one embodiment, the variant *Xanthobacter* strain is a variant *X. autotrophicus* Py2 strain wherein the gene set forth in NCBI ref. ABS67253.1 has been disrupted.

15 In a further aspect, the invention relates to a culture comprising the variant bacterial strain of the invention. In a preferred embodiment, the volume of the culture is 100 mL or more, e.g. 1 L or more, such as 10 L or more, e.g. 1,000 L or more, such as 10,000 L or more, e.g. 50,000 L or more, such as 100,000 L or more, e.g. 200,000 L or more.

20 In a further aspect, the invention relates to a process for the production of biomass and/or protein, said process comprising culturing the variant bacterial strain of the invention. In one embodiment, the process is for the production of biomass. In another embodiment, the process is for the production of protein. In one embodiment, the process comprises culturing the bacteria in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide. In a further embodiment, the process is for the production of biomass and comprises culturing the bacteria in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide. Various further embodiments of the process are described herein below.

30 In a further main aspect, the invention relates to a process for the production of biomass and/or protein, said process comprising culturing a variant chemoautotrophic bacterial strain in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide and wherein said variant chemoautotrophic strain comprises a gene disruption of one or more genes encoding a PHA synthase, preferably the gene with most sequence identity to phaC1 (SEQ ID NO:62). In one embodiment, the process is for the production of biomass. In

35

another embodiment, the process is for the production of protein. Various further embodiments of the process are described herein below.

In one embodiment, the variant chemoautotrophic strain used in the process is of the genus *Xanthobacter*, preferably a variant of strain VTT-E-193585.

5 In another embodiment, the variant chemoautotrophic strain used in the process is of the species *Cupriavidus necator*.

According to the genome sequence, the strain deposited under number VTT-E-193585 uses most likely Calvin-Benson-Bassham cycle for the carbon fixation where
10 carbon dioxide molecule is connected to 5-carbon chain of ribulose 1,5-bisphosphate forming two molecules of glycerate 3-phosphate. This enables the strain to synthesise all the other organic molecules it requires for growth. Energy from hydrogen comes into the cell most likely through NAD⁺-reducing hydrogenases and/or NiFeSe-hydrogenases. In essence that is a redox reaction where hydrogen (H₂) is oxidized to H⁺ and NAD⁺ is
15 reduced to NADH. In addition to ATP, NADH is one of the main energy carriers inside living organisms. Alternatively, some other energy equivalent is reduced by another hydrogenase enzyme using H₂. The Calvin-Benson-Bassham cycle requires energy in the form of ATP and NADH/NADPH in order to fix CO₂. The strain most likely generates ATP through oxidative phosphorylation, which consists of four protein complexes
20 generating a proton gradient across a membrane. The proton gradient is generated using mainly energy from NADH. The proton gradient drives the ATP synthase complex generating ATP. According to the genome sequence, the strain has a bacterial F-type ATP synthase.

It is to be understood, when it is specified that the process comprises culturing
25 the strain with an inorganic carbon source, that the inorganic carbon source is the *main* carbon source in the culture. Thus, there may be minor amounts of organic carbon sources present in the culture, but the main metabolism and growth of the culture is based on the utilisation of the inorganic carbon source, preferably carbon dioxide, as carbon source. Preferably the proportion of the carbon supplied to the culture that is
30 organic is less than 5%, such as less than 1%, e.g. less than 0.1% of all carbon supplied to the culture during the process. Preferably, no organic carbon sources are supplied to the process.

Similarly, it is to be understood, when it is specified that the process comprises
culturing the strain with hydrogen (H₂) as energy source, that hydrogen is the *main*
35 energy source in the culture. Thus, there may be other minor energy sources present in the culture such as ammonia, which may be supplied as nitrogen source, or minor amounts of organic compounds, but the main metabolism and growth of the culture is

based on the utilisation of hydrogen as energy source. In the overall process hydrogen is preferably produced by water electrolysis; i.e. by splitting water with electricity to hydrogen and oxygen gases. Thus, the hydrogen and oxygen gases are provided to the bioreactor from an electrolyser nearby. Alternatively, electrodes may be placed inside
5 the bioreactor to produce hydrogen and oxygen in the bioreactor rather than in a separate electrolyser.

The inorganic carbon source comprising carbon dioxide may comprise other inorganic carbon sources, such as e.g. carbon monoxide. In one embodiment, only carbon sources in gaseous form are provided to the culture. In a preferred embodiment,
10 carbon dioxide is the only inorganic carbon source, and indeed the only carbon source, provided to the culture. In one embodiment, only gases and minerals are provided to the culture and the level of carbon dioxide in the gas provided is between 10% and 50%, e.g. between 15% and 45%, such as between 20% and 40%, e.g. between 25% and 35%, such as between 26% and 30%.

15 In another embodiment, gases and minerals are provided to the culture and the level of hydrogen (H_2) in the gas provided is between 30% and 80%, e.g. between 35% and 75%, such as between 40% and 70%, e.g. between 45% and 65%, such as between 50% and 60%.

In another embodiment, gases and minerals are provided to the culture and the
20 level of oxygen (O_2) in the gas provided is between 10% and 25%, e.g. between 15% and 20%, such as between 16% and 18%. In another embodiment, the level of oxygen provided is such that the level of dissolved oxygen in the culture is maintained at between 5% and 10%.

In a preferred embodiment, only gases and minerals are provided to the culture
25 and the gas provided comprising H_2 , CO_2 and O_2 , wherein the percentage of H_2 is between 40% and 70%, the percentage of CO_2 is between 18% and 28% and the percentage of O_2 is between 12% and 22%.

Typically, the process of the invention includes the addition of a nitrogen source. The nitrogen source may for example be provided in the form of ammonium hydroxide, an ammonium salt, such as ammonium sulphate or ammonium chloride, ammonia, urea
30 or nitrate, e.g. potassium nitrate. In other embodiments, nitrogen gas (N_2) is provided as nitrogen source. In a preferred embodiment, the nitrogen source is ammonium hydroxide or an ammonium salt, such as ammonium sulphate.

In one embodiment, the nitrogen source provided is ammonium hydroxide at a
35 concentration of between 100 mg/L and 10 g/L, such as between 250 mg/L and 4 g/L, e.g. between 0.5 g/L and 2 g/L, such as between 0.75 g/L and 1.5 g/L.

Typically, the process of the invention includes the addition of minerals, such as minerals containing ammonium, phosphate, potassium, sodium, vanadium, iron, sulphate, magnesium, calcium, molybdenum, manganese, boron, zinc, cobalt, selenium, iodine, copper and/or nickel. Suitable mineral media are well-known art, and have e.g. been described in *Thermophilic Bacteria*, CRC Press, Boca Raton, FL, Jacob K. Kristjansson, ed., 1992, for example on page 87, Table 4.

In one embodiment, the minerals added include one or more of the following: ammonia, ammonium (e.g., ammonium chloride (NH_4Cl), ammonium sulphate ($(\text{NH}_4)_2\text{SO}_4$)), nitrate (e.g., potassium nitrate (KNO_3)), urea or an organic nitrogen source; phosphate (e.g., disodium phosphate (Na_2HPO_4), potassium phosphate (KH_2PO_4), phosphoric acid (H_3PO_4), potassium dithiophosphate ($\text{K}_3\text{PS}_2\text{O}_2$), potassium orthophosphate (K_3PO_4), disodium phosphate ($\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$), dipotassium phosphate (K_2HPO_4) or monopotassium phosphate (KH_2PO_4); sulphate; yeast extract; chelated iron (chelated e.g. with EDTA or citric acid); potassium (e.g., potassium phosphate (KH_2PO_4), potassium nitrate (KNO_3), potassium iodide (KI), potassium bromide (KBr)); and other inorganic salts, minerals, and trace nutrients (e.g., sodium chloride (NaCl), magnesium sulphate ($\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$) or magnesium chloride (MgCl_2), calcium chloride (CaCl_2), calcium sulphate (CaSO_4) or calcium carbonate (CaCO_3), manganese sulphate ($\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$) or manganese chloride (MnCl_2), ferric chloride (FeCl_3), ferrous sulphate ($\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$) or ferrous chloride ($\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$), sodium bicarbonate (NaHCO_3) or sodium carbonate (Na_2CO_3), zinc sulphate (ZnSO_4) or zinc chloride (ZnCl_2), ammonium molybdate (NH_4MoO_4) or sodium molybdate ($\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$), cuprous sulphate (CuSO_4) or copper chloride ($\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$), cobalt chloride ($\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$) or cobalt sulphate (CoSO_4), aluminium chloride ($\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$), lithium chloride (LiCl), boric acid (H_3BO_3), nickel chloride ($\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$) or nickel sulphate (NiSO_4), tin chloride ($\text{SnCl}_2 \cdot \text{H}_2\text{O}$), barium chloride ($\text{BaCl}_2 \cdot 2\text{H}_2\text{O}$), copper selenate ($\text{CuSeO}_4 \cdot 5\text{H}_2\text{O}$), sodium selenate (Na_2SeO_4) or sodium selenite (Na_2SeO_3), sodium metavanadate (NaVO_3), chromium salts).

In a preferred embodiment, the process of the invention includes the addition of one, more or all of: NH_4OH , KH_2PO_4 , $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, $\text{NaVO}_3 \cdot \text{H}_2\text{O}$, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, CaSO_4 , $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, H_3BO_3 , CoSO_4 , CuSO_4 , NiSO_4 .

In one embodiment, the medium provided to the cells comprises less than 1 g/L of chloride salts, such as less than 0.25 g/L of chloride salts, e.g. less than 0.1 g/L of chloride salts, such as less than 0.025 g/L of chloride salts, e.g. less than 0.01 g/L of chloride. In one embodiment, no chloride salts are supplied to the culture.

In another embodiment, no vitamins are supplied during the process, i.e. the media provided to the culture does not contain vitamins.

In another embodiment, no amino acids are supplied during the process, i.e. the media provided to the culture does not contain amino acids.

5 In another embodiment, no organic compounds are supplied during the process, i.e. the media provided to the culture does not contain any organic compounds.

In certain embodiments, the pH of the bacterial culture is controlled at a certain level. In certain embodiments, pH is controlled within an optimal range for bacterial maintenance and/or growth and/or production of organic compounds. In one
10 embodiment, the pH in the culture is maintained between 5.5 and 8.0, e.g. between 6.5 and 7.0, such as at 6.8.

In certain embodiments, the temperature of the bacterial culture is controlled. In certain embodiments, temperature is controlled within an optimal range for bacterial maintenance and/or growth and/or production of organic compounds. In one
15 embodiment, the culture is grown at a temperature between 25°C and 40°C, e.g. between 28°C and 32°C, such as at 30°C.

Typically, the process of the invention is carried out in a bioreactor. A bioreactor is utilized for the cultivation of cells, which may be maintained at particular phases in their growth curve. The use of bioreactors is advantageous in many ways for cultivating
20 chemoautotrophic growth. Generally, the control of growth conditions, including control of dissolved carbon dioxide, oxygen, and other gases such as hydrogen, as well as other dissolved nutrients, trace elements, temperature and pH, is facilitated in a bioreactor. Nutrient media, as well as gases, can be added to the bioreactor as either a batch addition, or periodically, or in response to a detected depletion or programmed set point,
25 or continuously while the period the culture is grown and/or maintained. In a continuous culture process, nutrient media, as well as gases, are added to the bioreactor continuously. Furthermore, bacteria-containing medium is being removed from the bioreactor continuously.

In a preferred embodiment, the volume of the bacterial culture is 100 mL or
30 more, such as 1 L or more, e.g. 10 L or more, such as 100 L or more, e.g. 1,000 L or more, such as 10,000 L or more, e.g. 50,000 L or more, such as 100,000 L or more, e.g. 200,000 L or more.

In one embodiment, the productivity of the culture is more than 0.1 g cell dry weight per liter per hour, such as more than 0.2, e.g. more than 0.3, such as more than
35 0.4, e.g. more than 0.5, such as more than 0.6, e.g. more than 0.7, such as more than 0.8, e.g. more than 0.9, such as more than 1 g per liter per hour.

Bacteria can be inoculated directly from a cell bank, or via a seed culture at a smaller scale. Preferably, supply of fresh media to the culture and removal of used up media with bacteria is occurring at the same rate, such that the volume in the bioreactor remains the same.

5 In one embodiment, after an initial phase of reaching a suitable cell density, the bacteria grow at steady state or pseudo steady state, remaining continuously in their log phase, at an OD600 above 5, such as above 10, e.g. above 20, such as between 50 and 200, e.g. between 50 and 100.

In one embodiment of the process of the invention, the bacterial strain has a
10 growth rate of $0.001\text{--}0.12\text{ h}^{-1}$, such as $0.01\text{--}0.12\text{ h}^{-1}$, for example $0.04\text{--}0.12\text{ h}^{-1}$.

In another embodiment of the process of the invention, the liquid feed rate in the continuous phase is 50-80% of the growth rate.

Xanthobacter is a genus of Gram-negative bacteria from the *Xanthobacteraceae* family. As mentioned above, in one embodiment, the variant chemoautotrophic strain
15 used in the process of the invention (i.e. a variant that comprises a gene disruption of one or more genes encoding a PHA synthase) is of the genus *Xanthobacter*. Preferably, the strain is selected from the group consisting of: *X. agilis*, *X. aminoxidans*, *X. autotrophicus*, *X. flavus*, *X. tagetidis*, *X. viscosus*, *Xanthobacter sp. 126*, *Xanthobacter sp. 91* and strain VTT-E-193585.

20 In one embodiment, the variant chemoautotrophic strain used in the process of the invention is a strain which uses the Calvin Benson Bassham pathway to convert carbon dioxide into organic compounds, e.g. glucose, essential for living organisms.

In one embodiment, the variant chemoautotrophic strain used in the process of the invention is a strain which uses NiFeSe-hydrogenases for converting hydrogen (H_2)
25 into cellular energy equivalents.

In one embodiment, the variant chemoautotrophic strain used in the process of the invention is a strain which uses NAD^+ -reducing hydrogenases for converting hydrogen (H_2) into cellular energy equivalents.

In one embodiment, the variant chemoautotrophic strain used in the process of
30 the invention capable of nitrogen fixation.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises the 16S ribosomal RNA set forth in SEQ ID NO:1 or a 16S ribosomal RNA having up to 20 nucleotide differences with SEQ ID NO:1, e.g. 1 to 10, such as 1 to 5, e.g. one, two or three nucleotide differences with SEQ ID NO:1.

35 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) large chain having the sequence set forth in SEQ ID

NO:3 or a sequence having more than 93% identity, e.g. more than 95% identity, such as more than 96% identity, e.g. more than 97% identity, such as more than 98% identity, e.g. more than 99% sequence identity to the sequence set forth in SEQ ID NO:3.

- 5 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) small chain having the sequence set forth in SEQ ID NO:5 or a sequence having more than 83% sequence identity, e.g. more than 86%, identity such as more than 90% identity, e.g. more than 95% identity, such as more than 96% identity, e.g. more than 97% identity, such as more than 98% identity, e.g. more than 99% sequence identity to the sequence set forth in SEQ ID NO:5.

SEQ ID NO:2:

Nucleotide sequence of Ribulose bisphosphate carboxylase large chain:

- 15 ATGGGTGCCGAAGCAACCGTCGGGCAGATCACGGACGCCAAGAAGAGATACGCCGCCGGCG
TGCTGAAGTACGCCAGATGGGCTACTGGAACGGCGACTACGTTCCCAAGGACACCGACCTC
CTGGCGGTGTTCCGCATCAACCCCCAGGCGGGCGTGGACCCGGTGAAGCCGCCGCGGCGG
TCGCCGGCGAAAGCTCCACCGCTACCTGGACCGTGGTGTGGACCGACCGGCTCACCGCCGC
CGACGTCTACCGCGCCAAGGCCTACAAGGTGGAGCCGGTGCCGGGCCAGGAAGGCCAGTAT
20 TTCTGCTACATCGCCTATGATCTCGATTTGTTTCGAGGAAGGCTCCATCGCCAACCTCACGGCG
TCGATCATCGGCAACGTCTTCTCCTTCAAGCCGCTGAAGGCGGCGCGGCTGGAGGACATGCG
GCTTCCCGTCGCCTATGTGAAGACCTTCCGCGGCCCGCCACCGGCATCGTGGTCGAGCGCG
AGCGCCTGGACAAGTTCGGCCGCCCCCTTCTGGGCGCCACCACCAAGCCGAAGCTTGGCCTC
TCGGGCAAGAATTACGGCCGCGTGGTCTATGAGGCCCTCAAGGGCGGCCTCGACTTCGTGAA
25 GGACGACGAGAACATCAACTCGCAGCCCTTCATGCACTGGCGGATCGCTTCCTCTATTGCAT
GGAGGCCGTCAACAAGGCCAGGCCGAGACCGGCGAGGTGAAGGGGCACTATCTCAACATC
ACCGCCGGGACCATGGAGGAGATGTACCGCCGCGCCGAGTTCGCCAAGGAACTGGGCTCCG
TGGTGGTGATGGTGGATCTCATCATCGGCTGGACCGCCATCCAGTCCATGTCCAACCTGGTGC
CGCGAGAACGACATGATCCTGCACATGCACCGTGCGGGCCATGGCACCTACACGCGCCAGAA
30 GAGCCACGGCGTCTCCTTCCGCGTCATCGCCAAGTGGCTGCGGCTCGCCGGCGTCGACCACC
TGACACCGGCACCGCCGTGGGCAAGCTGGAAGGCGACCCCATGACCGTGACGGGCTTCTA
CAATGTCTGCCGCGAGACGACGACGACGACGAGGACCTCACCCGCGGCCTGTTCTTCGAGCAGG
ACTGGGGCGGCATCCGCAAGGTGATGCCGGTGGCCTCCGGCGGCATCCATGCGGGCCAGAT
GCACCAGCTCATCGACCTGTTGCGCGAGGACGTGGTGTCCAGTTCGGCGGCGGCACCATCG
35 GCCACCCGGACGGCATCCAGGCCGGCGCCACCGCCAACCGCGTGGCGCTGGAAACCATGAT
CCTCGCCCGCAACGAGGGCCGCGACATCAGGAACGAGGGCCCGAAATCCTGGTGAAGCC

GCCAAATGGTGCCGTCCGCTGCGCGCGGCGCTCGATACCTGGGGCGAGGTGACCTTCAACTA
CGCCTCCACCGACACGTCCGATTACGTGCCCACCGCGTCCGTGCGCTGA

SEQ ID NO: 3:

- 5 Amino acid sequence of Ribulose biphosphate carboxylase large chain
MGA EATV GQITDAKKRYAAGVLKYAQMGYWNGDYVPKDTDLLAVFRITPQAGVDPVEAAAAVA
GESSTATWTVVWTDRLTAADVYRAKAYKVEPVPGQEGQYFCYIAYDLDFEEGSIANLTASIIGN
VFSFKPLKAARLEDMRLPVAYVKTRGPPTGIVVERERLDKFGRP LLGATTKPKLGLSGKNYGRVV
YEALKGGLDFVKDDENINSQPFMHWDRFLYCM EAVNKAQAETGEVKGHYLNITAGTMEEMYRR
10 AEF AKELGSVVVMVDLIIGWTAIQSMSNWCRENDMILHMRAGHGT YTRQKSHGVSFRVIAKW
LRLAGVDHLHTGTAVGKLEGPMTVQG FYNVCRETTTQQDLTRGLFFEQDWGGIRKVM PVASG
GIHAGQMHLIDLFGEDVVLQFGGGTIGHPDGIQAGATANRVALETMILARNEGRDIRNEGPEIL
VEAAKWCRPLRAALDTWGEVTFNYASTDTS DYVPTASVA

- 15 SEQ ID NO: 4:

Nucleotide sequence of Ribulose biphosphate carboxylase small chain:

- ATGCGCATCACCCAAGGCTCCTTCTCCTTCTGCGGACCTCACCGACACGCAGATCAAGGCC
CAGGTGCAATATTGCCTGGACCAGGGCTGGGCGGTCTCGGTGGAGCACACCGACGATCCCCA
CCGCGCAACACCTATTGGGAGATGTGGGGCCCGCCCATGTTGATCTGCGCGACGCGGCC
20 GCGCTCTTCGGCGAGATCGAAGCCTGCCGGGCCGCAATCCCGAGCATTATGTGCGGGTGAA
CGCCTTCGATTCCAGCCGCGGATGGGAGACGATCCGCCTGTCCTTCATCGTTCAGCGGCCCA
CCGTGGAAGAGGGCTTCCGCCTCGACCGCACCGAAGGCAAGGGCCGCAACCAGAGCTACGC
CATGCGCTACCGGGCGCAGTTCGCGCCGCGCTGA

- 25 SEQ ID NO: 5:

Amino acid sequence of Ribulose biphosphate carboxylase small chain:

- MRITQGSFSFLPDLTDTQIKAQVQYCLDQGWAVSVEHTDDPHPRNTYWEMWGPPMFDLRDAAG
VFGEIEACRAANPEHYVRVNAFDSSRGWETIRLSFIVQRPTVEEGFRLDRTEGKGRNQSYAMRYR
AQFAPR

30

- In another embodiment, the variant chemoautotrophic bacterial strain used in the
process of the invention comprises a gene encoding a NAD⁺-reducing hydrogenase HoxS
subunit alpha having the sequence set forth in SEQ ID NO:7 or a sequence having more
than 70% sequence identity, such as more than 80% identity, e.g. more than 90%
35 identity, such as more than 95% identity, e.g. more than 96% identity, such as more
than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to
the sequence set forth in SEQ ID NO:7.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a NAD⁺-reducing hydrogenase HoxS subunit beta having the sequence set forth in SEQ ID NO:9 or a sequence having more than 77% sequence identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:9.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a NAD⁺-reducing hydrogenase HoxS subunit gamma having the sequence set forth in SEQ ID NO:11 or a sequence having more than 70% sequence identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:11.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a NAD⁺-reducing hydrogenase HoxS subunit delta having the sequence set forth in SEQ ID NO:13 or a sequence having more than 79% sequence identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:13.

SEQ ID NO:6:

Nucleotide sequence of NAD⁺-reducing hydrogenase HoxS subunit alpha:

ATGATGCCATCTGAGCCGCACGGCGCGGGCATGCCGCCCCACGGGAAGCGGCCGCGGTTC
 CCACCCCCCAGGAGGTGAGCGCGGTGGTGGCCGAGGTGGTCGCGGATGCCGTGGCATCGGT
 GGGCGGCGCACGCACCCGGCTCATGGACATCGTCCAGCTGGCCCAGCAGCGTCTCGGCCAT
 CTCTCCGAAGAGACCATGGCGGCCATTGCCGCGGGCTCGCCATTCGCGCGGTGGAAGTGG
 CGGACATGGTGTCTTCTACGCCTTCTGAACCGCGCGCCCAAGGGCCGCTACCACATCCGC
 CTGTGCGCGAGCCCCATCTCGCTGATGAAGGGCGCCGAGGCGGTGGCTGCCGCCTTCTGCCA
 GATCCTCGGCATCGCCATGGGCGAGACCTCGCAGGATGGCGACTTCACCCTGGAATGGACCA
 ACGACATCGGCATGGCCGACCAGGAGCCGGCCGCCCTCGTCAACGGCACGGTGATGACGCA
 GCTCGCGCCCGGCGATGCGGCCATCATCGTCCGGCGGCTGCGGGCCCATCACGCGCCCAAT
 GCCCTGCCGCTGTTCCCTGGAGCCGGCGTGGCCGGCTCCGGCCTGCCCCATGCCCGGATCC
 GCCCCAGCCTGGTGATGCCGGGACAGCTTCTGTTCCGCGAGGACCACACGACGCCGGGCGC
 CGGCATCAAGGCGGCACTCGCCCTCACCCCGGACGAAGTGGTGCAGAAGGTCTCCGCCGCG
 CGCCTGCGCGGGCGGGGTGGCGCCGGCTTCCACCGGTCTCAAATGGAAGCTCTGCCGCC

AGTCGCCCCGCCACCAACCCGCCATGTGATCTGCAATGCGGACGAGGGCGAGCCCCGGCACCTTC
AAGGATCGCGTGCTGCTCACGCAGGCGCCGCACCTCATGTTGACGGCATGACCATCGCCGG
CTACGCCTTGGGGGCGCGGGAGGGCGTGGTCTATCTGCGCGGCGAGTACGCCTATCTGTGG
GAGCCTCTGCATGCGGTCTGCGCGAGCGCTATGGGCTCGGGCTCGCCGGCGCGAACATCC
5 TGGGACACGCGGGCTTCGACTTCGACATCCGCATCCAGCTGGGCGCCGGCGCCTATATCTGC
GGCGAGGAATCCGCGCTGGTGAATCGCTGGAAGGCAAGCGCGGCTCGCCCCGCGACCGCC
CCCCCTTCCCCACCGTGC GCGGCCATCTCCAGCAGCCCACCGCCGTGGACAATGTGGAGACC
TTCGCCTGCGCCGCCCCGCATCCTGGAGGATGGCGTGGAGGCGTTGCGGGGCATCGGCACGC
CCGAATCCGCCGGCACGAAGCTCCTCTCGGTGTCGGGCGATTGCCCGCGCCCCGGCGTGTAT
10 GAGGTGCCCTTCGGCCTCACGGTGAACGCGCTGCTCGACCTTGTGCGCGCGCCGGACGCCG
CCTTCGTGCAGATGGGTGGGCCGTCCGGCCAATGCGTGGCGCCGAAGGATTACGGCCGCCG
CATCGCCTTCGAGGACCTGCCACCGGCGGCTCGGTGATGGTGTTCGGCCCCGGGGCGCGAC
GTGCTCGCCATGGTGC GCGAGTTGCGGATTTCTTCGCCGGCGAATCCTGCGGCTGGTGCAC
GCCCTGCCGGGTGGGCACCACTTGTCAAGGAAGAGCTGGACAAGCTCCTCGCCAACCGCG
15 CCACCCTCGCCGACATCCGCGCGCTGGAGACCCTGGCCACGACCGTCTCCCGCACCAGCCCG
TGCGGCCTCGGCCAGACGGCGCCCAACCCCATCTTTCCACCATGCGCAACCTGCCGGAAGC
CTATGAGGCGAGGCTGAGGCCCCAAGACTTCTGCCCTGGGCCTCGCTCGACGAGGCGCTG
AAGCCCGCCATCGTCATCCAGGGCCGCGCGCCCGTGCCGGAGGAAGAGGCATGA

20 SEQ ID NO: 7:

Amino acid sequence of NAD⁺-reducing hydrogenase HoxS subunit alpha:

MMPSEPHGAGMPPPREAAAVPTPQEVSAVVAEVDADAVASVGGARTRLM DIVQLAQRLGHLSE
ETMAAIAARLAIPPVEVADMVSFYAFLNRAPKGRYHIRLSRSPISLMKGAEAVAAAFQCILGIAMG
ETSQDGDFTLEWTNDIGMADQEPAALVNGTVM TQLAPGDAAIIVGRLRAHHAPNALPLFPAGAV
25 AGSGLPHARIRPSLVMPGQLLFREDHTTPGAGIKAALALTPDEVVQKVSAARLRGRGGAGFPTGL
KWKLCRQSPATTRHVICNADEGEPTGFKDRVLLTQAPHL MFDGMTIAGYALGAREGVVYLRGEY
AYLWEPLHAVLRERYGLGLAGANILGHAGDFDFDIRIQLGAGAYICGEESALVESLEGKRGSPDRP
PFPTVRGHLQQPTAVDNVETFACAARILEDGVEAFAGIGTPESAGTKLLSVSGDCPRPGVYEV PFG
LTVNALLDLVGAPDAA FVQMGGPSGQCVAPKDYGRRIAFEDLPTGG SVMVFGPGRDVLAMVREF
30 ADFFAGESCGWCTPCRVTLLKEELDKLLANRATLADIRALET LATTVSRTSRCGLGQTAPNPIL
STMRNLPEAYEARLRPEDFLPWASLDEALKPAIVIQGRAPVPEEEA

SEQ ID NO: 8:

Nucleotide sequence of NAD⁺-reducing hydrogenase HoxS subunit beta:

35 ATGAGCCGGGGATCCCCCGATGCCGGGAAAGACCGCACCATGAGCGCCACCGACGGCACCA
CCGCCCCCGCAAGATCGTCATCGATCCGGTGACCCGCGTGGAGGGCCACGGCAAGGTCAC
CATCCGCCTGGATGAAGCCGGCGCGGTGGAGGATGCGCGTTTCCACATCGTGGAGTTCCGC

GGCTTCGAGCGGTTTCATCCAGGGCCGGATGTAAGTGGGAAAGTGCCCTTATCATCCAGCGGCT
 GTGCGGCATCTGCCCCGGTGAGCCACCATCTGGCGGCGGCGAAAGCCATGGACCAGGTGGCG
 GGCGTGAGCCGCTACCGCCACCGCCGAGAACTGCGCCGGCTGATGCATTATGGGCAGG
 TGCTGCAATCCAACGCTTTGCACATCTTCCACCTCGCCTCGCCGACCTCCTGTTTCGGCTTCG
 5 ACGCGCCGGCCGAGCAGCGCAACATCATCGCCGTGCTCCAGCGTTATCCGGAGATCGGCAAA
 TGGGCGATCTTCATCAGGAAGTTCGGCCAGGAGGTCATCAAGGCCACCGGCGGGCGCAAGA
 TCCATCCCACCAGCGCCATTCCCGGCGGGGTCAACCAGAACCTCGCCGTGGAGGACCGCGAC
 GCCCTGCGCGCCAAGGTGGGCGAGATCATCAGCTGGTGATGGCGGCGCTGGACCATCACA
 AGGCCTATGTGGCGGAAAACCGGGCGCTGCATGACAGCTTCGCCGCCTTCCCCTCCGCCTTC
 10 ATGAGCCTCGTGGGGCCGGATGGCGGCATGGACCTTTATGACGGCACCTGCGGGTGATCG
 ATGCCGAGGGCGCCCCCTCATCGAAGGCGCGCCGCCCGCCTCCTACCGCGACCACCTCATC
 GAGGAGGTGCGGCCCTGGAGCTATCTGAAATTCCTCCATCTGCGCGCCTTCGGCCGCGACGA
 TGGCTGGTATCGGGTCGGCCCCCTCGCCAGGTCAATTGCGCCGCGTCCATCGACACGCCCC
 GCGCCGAGGCGGCCCGGCGGGACTTCATGGCCGAGGGCGGCGGCAAGCCGGTGATGCCA
 15 CCCTCGCTTATCACTGGGCGCGGCTCATCGTGCTGGTCCATTGCGCGGAGAAGATCGAACAG
 CTGCTGTTTCGACGACGACCTGCAAGGCTGCGATCTGCGTGCGGAGGGCACCCGGCGCGGGG
 AAGGCGTCGCCTGGATCGAGGCGCCGCGCGGCACCCTCATCCACCATTACGAGGTGGACGA
 GAACGACCAGGTGCGCCGCGCCAACCTCATCGTCTCCACCACCCACAATAACGAGGCCATGA
 ACCGCGCCGTGCGGCAGGTGGCGAAGACGGACCTTTCGGTGCGGAGATCACCGAAGGGCT
 20 GCTGAACCATATCGAGGTGGCCATCCGCGCCTTCGACCCCTGCCTGTCCTGCGCCACCCATG
 CGCTGGGCCAGATGCCGCTGATCGTGACGCTTGAAGATGCCTCCGGCGCAGAGATCGCCCG
 CGGAGTGAAGGAATGA

SEQ ID NO: 9:

25 Amino acid sequence of NAD⁺-reducing hydrogenase HoxS subunit beta:
 MSRGSPDAGKDRTMSATDGTAPRKIVDPVTRVEGHGKVTIRLDEAGAVEDARFHIVEFRGFER
 FIQGRMYWEVPLIIQRLCGICPVSHHLAAAKAMDQVAGVDRVPPTAEKLRRMLHYGQVLQSNAL
 HIFHLASPDLLFGFDAPAEQRNIIAVLQRYPEIGKWAIFIRKFGQEVKATGGRKIHPTSAIPGGVN
 QNLAVEDRDALRAKVGEIISWCMAALDHHKAYVAENRALHDSFAAFPSAFMSLVGPDGGMDLY
 30 DGTLRVIDAEGAPLIEGAPPASYRDHLIEEVRPWSYLKFPHLRAFGRDDGWYRVGPLAQVNCAAS
 IDTPRAEAARRDFMAEGGGKPVHATLAYHWARLIVLVHCAEKIEQLLFDDDLQGCDLRAEGTRRG
 EGVAWIEAPRGTLIHHYVDENDQVRRANLIVSTTHNNEAMNRAVRQVAKTDLSGREITEGLLN
 HIEVAIRAFDPCLSCATHALGQMPLIVTLEDASGAIEIARGVKE

35 SEQ ID NO: 10:

Nucleotide sequence of NAD⁺-reducing hydrogenase HoxS subunit gamma:

ATGAGCGAGACCCCTTCACCTTTACCGTGACGGCATCGCGGTCCCGGCCACCCCGGCCA
GAGCGTCATCGAGGCGTGCGATGCGGCGGGCATCTATATCCCGCGCCTGTGCCACCACCCG
GACCTGCCGCCGCGGGGCCATTGCCGGGTGTGCACCTGCATCATCGACGGGCGGCCGGCCA
GCGCCTGCACCATGCCCCGCCGCCAGGGGCATGGTGGTGGAGAACGAGACGCCCGCTTTGCT
5 GGCGGAGCGGCGCACGCTGATCGAGATGCTGTTGCGGGAAGGCAACCATTCTGCCAGTTCT
GCGAGGCGAGCGGCGATTGCGAATTGCAGGCGCTGGGCTACCTGTTGCGCATGGTGGCCCC
GCCCTTCCCCCATCTGTGGCCGAAGCGGCGCGTGGATGCCAGCCATCCGATATCTATATCG
ACCACAATCGCTGCATCCTGTGCTCGCGCTGCGTGCGCGCCTCGCGCACCTGGACGGCAAG
TCCGTGTTGCGCTTCGAGGGGCGCGGCATCGAGATGCATCTGGCGGTGACCGGCGGGCACC
10 TGGACGACAGCGCCATCGCCGCCGCCGACAGGGCGGTTGAGATGTGCCCGGTGGGCTGCAT
CGTCCTCAAGCGCACCGGCTACCGCACGCCCTATGGCCGGCGGCGCTACGACGCCGCGCCC
ATCGGCTCCGACATCACCGCCCGGCGCGGCGGCGCGCAAGGACTGA

SEQ ID NO: 11:

15 Amino acid sequence of NAD⁺-reducing hydrogenase HoxS subunit gamma:
MSETPFTFTVDGIAVPATPGQSVIEACDAAGIYIPRLCHHPDLPPAGHCRVCTCIIDGRPASACTM
PAARGMVVENETPALLAERRTLIEMLFAEGNHFCQFCEASGDCELQALGYLFGMVAPPPHLWPK
RPVDASHPDYIDHNRCILSRCVRSRTLDGKSVFGFEGRGIEMHLAVTGHHLLDDSAIAAADRA
VEMCPVGCIVLKRTGYRTPYGRRRYDAAPIGSDITARRGGAKD

20

SEQ ID NO: 12:

Nucleotide sequence of NAD⁺-reducing hydrogenase HoxS subunit delta:

ATGGCCAAGCCCAAACTCGCCACCTGCGCGCTGGCCGGCTGCTTCGGCTGCCACATGTCCTT
CCTGGACATGGACGAGCGCATCGTCGAGCTCATCGACCTGGTGGACCTCGACGTCTCGCCCC
25 TCGACGACAAGAAAACTTCACCGGCATGGTGGAAATCGGCCTGGTGGAAAGCGGCTGCGC
CGACGAGCGCCATGTGAAGGTGCTGCGCGAGTTCCGCGAGAAATCCCGCATCCTGGTGGCG
GTGGGCGCCTGCGCCATCACCGGCGGCATCCCGGCATTGCGCAACCTCGCCGGCCTCGACG
AATGCCTGAGGGAAGCCTACCTCACCGGCCCCACGGTGGAAAGCGGCGGGGCTCATTCCCAAC
GACCCGGAGCTGCCGCTGCTGCTGGACAAGGTCTATCCGGTGCAGGACTTCGTGAAGATCGA
30 CCATTTCTGCCCGGCTGCCCCCCTCGCCGACGCCATCTGGGCGGCTCTGAAGGCGCTGC
TGACCGGCACCGAGCCGCATCTGCCCTACCGCTTTTCAAGTACGAATGA

SEQ ID NO: 13:

Amino acid sequence of NAD⁺-reducing hydrogenase HoxS subunit delta:

35 MAKPKLATCALAGCFGCHMSFLDMDERIVELIDLVDLVSPLDDKKNFTGMVEIGLVEGGCADER
HVKVLREFREKSRILVAVGACAITGGIPALRNLAGLDECLREAYLTGPTVEGGGLIPNDPELPLLLD
KVYPVQDFVKIDHFLPGCPPSADAIWAALKALLTGTEPHLPYPLFKYE

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a NiFeSe hydrogenase large subunit having the sequence set forth in SEQ ID NO:15 or a sequence having more than 84% sequence identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:15.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a NiFeSe hydrogenase small subunit having the sequence set forth in SEQ ID NO:17 or a sequence having more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:17.

SEQ ID NO: 14:

Nucleotide sequence of Periplasmic [NiFeSe] hydrogenase large subunit:

TCCAGACCCGGGCAACATTGCTCCATGTGCTGGGCACCCTGGCCGGCCGCTGGCCCCATACC
CTCGCGCTCCAGCCCGGCGGGGTGACCCGAAGCGCCGACCAGCACGACCGCATGCGCCTGC
TCGCGACGCTGAAGGCGGTGCGGGCGGCGCTGGAAGAGACCTTGTTTCGGCGCGCCTTTGGA
AGAGGTGGCGGCCCTGGACGGCGCCGCCCGTGGAGGCCTGGCGCGCCAACGGCCCCGGA
AGGGGATTTCCGCCTGTTCTTGGAGATCGCCGCCGACCTGGAGCTGGACCGGCTCGGCCGC
GCGCACGACCGCTTTCTCTCTTCGGCGCCTACGCCAGGACGAGGGGCGCCTTTATGGCGC
CGGCACCTTCGAGGCCGGGACGGCGGGAGGGCTCGATCCCAACGCCATCACCGAGGACCAC
GCCTTCGCCCCGATGGAGGACCGCGCGGCGCCCCATGCGCCCTTTGACGGCTCCACCTTCCC
CGATGCCGACGACACCGAGGGCTACACCTGGTGCAAGGCGCCGCGCCTTGCCGGCCTGCCC
TTCGAGACCGGCGCCTTCGCCCCGCGAGGTGGTGGCGGGCCATCCGCTCGCCCGGGACCTCG
TGACGCGGGAAGGCGGCACTGTGCGCAGCCGCGTGGTTCGGCCGGCTGCTGGAACCGCGC
GCACCCTGATCGCCATGGAGGGCTGGGTGAAGGAACTGCGGCCCGAAGGGCCCTGGTGCGC
CCAGGGCCACCTGCCCCAGGAAGGCCGCGCCTTCGGCCTCACCGAGGCGGCGCGCGGGGC
GCTCGGCCACTGGATGGTGGTGGAGAAGGGCCGATTGCCCGCTACCAGATCATCGCCCCCA
CCACCTGGAATTCTCCCCCGCGACGGCGCGGGCCTGCCCGGCCCGCTGGAGACGGCCCT
GGTGGGCGCGCCGTGCGGCAGGGAGAGACGACGCCCGTGAGCGTGCAGCACATCGTGCG
CTCCTTCGACCCGTGCATGGTCTGCACTGTGCATTGA

SEQ ID NO: 15:

Amino acid sequence of Periplasmic [NiFeSe] hydrogenase large subunit:

MSAETRRLVVGPFNRVEGDLEVRDLDVQDGRVQQAQFVSSPLFRGFERILEGRDPRDALVIAPRICGI
 CSVSQSHAAALALAGLQGIAPTHDGRIATNLIVAAENVADHLTHFHVFFMPDFARAVYEDRPWFA
 QAARRFKANQGVSVRRALQTRATLLHVLGTLAGRWPHTLALQPGGVTRSADQHDRMRLATLKA
 VRAALEETLFGAPLEEVAALDGAAVEAWRANGPEGDFRLFLEIAADLELDRLGRAHDRFLSFGAY
 5 AQDEGRLYGAGTFEAGTAGGLDPNAITEDHAFARMEDRAAPHAPFDGSTFPDADDTEGYTWCK
 APRLAGLPFETGAFARQVVAGHPLARDLVLTREGGTVRSRVVGRLLETARTLIAMEGWVKELRPEG
 PWCAQGHLPQEGRAFLTEAARGALGHWMVVEKGRIARYQIIAPTTWNFSPRDGAGLPGPLETA
 LVGAPVRQGETTPVSVQHIVRSFDPCMCTVH

10 SEQ ID NO: 16:

Nucleotide sequence of Periplasmic [NiFeSe] hydrogenase small subunit
 ACGGGGGAGGAAGCCCGCGCCATCTTCGACGCCATCCTTGCCGGCGTTATCGTCTCGACGC
 CCTGTGCGTGGAAGGCGCGCTGCTGCGCGGGCCGAACGGCACCGGGCGCTTCCATGTGCTG
 GCGGGCACGGACACCCCCACCATCGACTGGGCGCGGCAGCTCGCCGGCATGGCGCGCCACG
 15 TGGTGGCGGTGGGCACCTGCGCCGCTATGGGGGCGTGACGGCGGCGGGCATCAACCCAC
 CGATGCCTGCGGCCTCCAGTTCGACGGACGCCGGAAGGGTGGGGCGCTGGGGGCGGACTTC
 CGCTCCCGCTCGGGGCTTCCGGTCATCAATGTGGCCGGCTGCCCCACCATCCCAACTGGGT
 GACGGAAACCTGATGCTGCTCGCCTGCGGCCTGCTGGGCGAGGCCGACCTCGACGTCTATG
 GCCGCGCGCTTCTATGCGGACCTGCTGGTGCATCACGGCTGCGCGCAACGAATACTAT
 20 GAATACAAGGCGAGCGCCGAGAAGATGAGCGACCTCGGCTGCATGATGGAGCATCTGGGCT
 GCCTCGGCACCCAGGCCACGCCGACTGCAACACGCGCCTTTGGAATGGCGAGGGCTCGTG
 CACCCGCGGCGGCTATGCCTGCATCAACTGCACGGCGCCGGAATTCGAGGAGCCGGGCCAC
 GCCTTCCTGGAGACGCCCAAGATCGGCGGCATCCCCATCGGCCTGCCCACCGACATGCCCAA
 GGCCTGGTTCATCGCCTTGTCTCCCTCGCCAAGGCGGCGACGCCGAGCGGCTGCGCAAG
 25 AACGCGGTGTCCGACCATGTGGTCACGCCGCCCGCCGTCAAGGACATCAAGCGGCGATGA

SEQ ID NO: 17:

Amino acid sequence of Periplasmic [NiFeSe] hydrogenase small subunit
 MSTPFSVLWLQSGGCGGCTMSLLCAEAPDLATTLDAAGIGFLWHPALSEETGEEARAIFDAILAG
 30 VIVLDALCVEGALLRGPNGTGRFHVLAGTDTPTIDWARQLAGMARHVVAVGTCAAYGGVTAAGI
 NPTDACGLQFDGRRKGGALGADFRSRLPVINVAGCPTHNPWVTETLMLLACGLLGEADLDVY
 GRPRFYADLLVHHGCPRNEYEYKASAEKMSDLGCMMEHLGCLGTQAHADCNTRLWNNEGSC
 TGGYACINCTAPEFEPEGHAFLETPKIGGIPIGLPTDMPKAWFIALSSLAKAATPERLRKNAVSDHV
 VTPPAVKDIKRR

35

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase gamma chain

atpG_1 having the sequence set forth in SEQ ID NO:19 or a sequence having more than 70% identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:19.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit alpha atpA_1 having the sequence set forth in SEQ ID NO:21 or a sequence having more than 78% identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:21.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit b atpF_1 having the sequence set forth in SEQ ID NO:23 or a sequence having more than 62% identity, e.g. more than 70% identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:23.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit c, sodium ion specific atpE_1 having the sequence set forth in SEQ ID NO:25 or a sequence having more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:25.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit a atpB_1 having the sequence set forth in SEQ ID NO:27 or a sequence having more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:27.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase epsilon chain atpC_1 having the sequence set forth in SEQ ID NO:29 or a sequence having more than 71% identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity,

e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:29.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit beta
5 atpD_1 having the sequence set forth in SEQ ID NO:31 or a sequence having more than 84% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:31.

In another embodiment, the variant chemoautotrophic bacterial strain used in
10 the process of the invention comprises a gene encoding an ATP synthase subunit beta atpD_2 having the sequence set forth in SEQ ID NO:33 or a sequence having more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:33.

In another embodiment, the variant chemoautotrophic bacterial strain used in
15 the process of the invention comprises a gene encoding an ATP synthase gamma chain atpG_2 having the sequence set forth in SEQ ID NO:35 or a sequence having more than 86% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:35.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit alpha
20 atpA_2 having the sequence set forth in SEQ ID NO:37 or a sequence having more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:37.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit delta
25 atpH having the sequence set forth in SEQ ID NO:39 or a sequence having more than 85% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:39.
30

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit b
35 atpF_2 having the sequence set forth in SEQ ID NO:41 or a sequence having more than 87% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:41.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit b' atpG_3 having the sequence set forth in SEQ ID NO:43 or a sequence having more than 81% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:43.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding ATP synthase subunit c atpE_2 having the sequence set forth in SEQ ID NO:45 or a sequence having more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:45.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase subunit a atpB_2 having the sequence set forth in SEQ ID NO:47 or a sequence having more than 92% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:47.

In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding an ATP synthase protein I atpI having the sequence set forth in SEQ ID NO:49 or a sequence having more than 60% identity, e.g. more than 70% identity, such as more than 80% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:49.

SEQ ID NO: 18:

Nucleotide sequence of ATP synthase gamma chain atpG_1

GTGACCGAGCGCCTGTCCGACGTCAACGCCCGCATCGCCTCGGTGCGGCAGCTCTCATCGGT
CATCACGGCCATGCGGGGCATTGCGGCGGCGCGGGCGCGGGAGGCGCGGGGTGCGCTCGA
CGGCATCCGCGCCTATGCGCAGACCATCGCCGAGGCCATCGGCCATGTGCTCGCCGTGCTGC
30 CCGAGGAGGCCCGCGCCCGGTCTCTCCGGGCACCGGCATCGGGGCCATGCGGTCTCGCCCT
GTGCGCGGAGCAGGGCTTTGCCGGCGTCTTCAACGAGCGGGTGCTGGACGAGGCCGCCCGG
CTGCTGACCGGCGGGGCGGGGCCGCGCCGAGCTGCTGCTGGTGGGCGACCGGGGCCTGATG
GTGGCCCCGCGAGCGGGGGCTCGATGTCTCTGGTGGTGGCCATGGTGGCCCATGCGGGGCC
AGGCCTCGGCGCTGGCGGACCGCATCAGCGAGGAGCTCTACCGGCGGATCGATGCGGGACG
35 GGTGACGCGGGTGTCGGTGGTGACGCCGAGCCCGCCGCGTCCGCCGCCATCGAGACGGTG
GTGAAAGTGCTGGTGCCGTTCGACTTCGCCCCGCTCCCCCTGGCGCGGGTGGCATCCGCCCC
GCTCATGACCATGCCGCCGCCGCGGCTGCTGGCCCAGCTGTCGGAGGAATATGTGTTGCCCG

AGCTGTGCGAGGCGCTCACCTTGTCTTCGCGGCGGAGAACGAGGCCCGCATGCGGGCCAT
GATCGCCGCGCGCCAATGTGGCCGATACCCTGGAGGGCCTCGTCGGCCGCGCCCGGCAG
ATGCGCCAGGAGGAGATCACCAACGAGATCATCGAGCTGGAAGGCGGCGCCGCGCAGCGCCC
GGCATGCGGATTGA

5

SEQ ID NO: 19:

Amino acid sequence of ATP synthase gamma chain atpG_1

MTERLSDVNARIASVRQLSSVITAMRGIAAARAREARGRLDGIRAYAQTIAEAIGHVLAVLPPEAR
ARSSGHRHRGHAVIALCAEQGFAGVFNervLDEAARLLTGGAGPAELLVGDRLMVARERGLD
10 VSWVPMVAHAGQASALADRISEELYRRIDAGRVRTSVVHAEPASAAIETVVKLVLPDFARF
PLARVASAPLMTMPPRLLAQLSEYVFAELCEALTSFAAENEARMRAMIAARANVADTLEGLVG
RARQMRQEEITNEIELEGGAGSARHAD

SEQ ID NO: 20:

15 Nucleotide sequence of ATP synthase subunit alpha atpA_1

ATGAGCACGGGCGCGCAAGCGAGCGAGGATTGGCTACCCGGAGCCGGGCGGCCCTGGCC
GGGACGCGCCTTTCCAGCAATCCCAATCGGTGGGCCGGTGGAGGAGATGGCCGACGGCA
TCGCCCCGCTCTCCGGCCTGCCGATGTGCGGCTCGACGAGCTTCTACCTTCGAGGGCGGC
CAGACCGGCTATGCCCTCACCTCGATCGCACCGAGATCGCCGTGGTGTCTGGATGACGC
20 CTCCGGCGTGGAGGCGGGCGCCCGGTGTTCCGGCACCGGCGAGGTGGTGAAGGTGCCGGT
GGGGCCGGGGCTGCTGGGCCGCATCGTCGACCCCTCGGCCGCGCCATGGACCGCTCCGAG
CCGGTGGTGGCGAGGCGACCATCCCATCGAGCGGCCGCGCCGGCCATCATGCCCGCG
ACCTGGTCTCGCAGCCGGTTCAGACCGGCACGCTGGTGGTGGATGCGCTGTTCTCCCTCGGC
CGGGGCCAGCGCGAGCTCATCATCGGCGACCGGGCTACCGGCAAGACCGCCATCGCGGTGG
25 ACACCATCATCAGCCAGAAGCATTTCGACATCGTGTGCATCTACGTGGCGGTGGGCCAGCGC
GCCGCCGCGTGGAGCGGGTGGTGGAGGCGGTGCGCGCCACGGGGCGATCGAGCGCTGC
ATCTTCGTGGTTCGCTCGGCCCGCCTCGCCAGGGCTGCAATGGATCGCGCCGTTCCGCCG
CATGACCATGGCGGAATATTTCCGCGACAACGGCCAGCATGCGCTCATCATCATGATGATCT
CACCAAGCATGCGGCCACCCATCGCGAGCTGGCGCTGCTACCCACGAGCCGCCGGGCCGC
30 GAGGCCTATCCCGGCGACATCTTCTATGTGCACGCCCCCCTTCTGGAGCGGGCCCAAGCT
CTCCGCCGAGCTGGGCGGTGGCTCGCTCACGGCCCTGCCATCGCGGAGACGGACGCGGGA
AACCTCTCCGCCTATATCCCCACCAACCTCATCTCCATACCGATGGGCAGATCGTGTGGAT
TCGCGGCTGTTTCGCGCCAACAGCGCCCGCGGTGGATGTGGGCTCTCCGTGAGCCGGG
TGGGCGGCAAGGCGCAGCATCCCGCGCTTCGGGCGGTGTCCGGGCGCATCCGGCTCGATTA
35 TTCCCAGTTCCTGGAGCTGGAATGTTACCCGCTTCGGCGGCATACCGATACCCGCGTGAA
GGCGCAGATCACCCGGGGCGAGCGCATCCGCGCGCTGCTACCCAGCCGCGCTTTTCCACCC
TGCGCCTTCAGGACGAGGTGGCGCTGCTGGCCGCGCTGGCGGAGGGGGTGTTCGACACTTT

GGCCCCGGGGCTGATGGGCGCCGTGCGTGCCCGCATTCCGGCCCAGCTGGATGCGCAGGTG
AAGGACGTGGCCTCGGCCCTCGCCGAGGGCAAGGTGCTGGAGGAGGGCTTGACGCCCCGTC
TCGTGGCGGCCGTGCGGGCCGTGCGGGCGGACGTGGCCGCGACCGCGAAGGCCGGGCCGT
GA

5

SEQ ID NO: 21:

Amino acid sequence of ATP synthase subunit alpha atpA_1

MSTGAQASEDWLTRSRAALAGTRLSQQSQSVGRVEEMADGIARVSGLPDVRLELTFEGGQT
GYALTDRTEIAVVLLDDASGVEAGARVFGTGEVVKVPVGPGLLGRIVDPLGRPMRSEPVVAQA
10 HHPIERPAPAIIRDVLSQPVQTGTLVVDALFSLGRGQRELIIGDRATGKTAIAVDTIISQKHS
DIVCIYVAVGQRAAAVERVVEAVRAHGAIERCIFVVASAAASPGLQWIAPFAGMTMAEYFRDNGQHA
LIIIDDLTKHAATHRELALLTHEPPGREAYPGDIFVYHARLLERAALKSAELGGGSLTALPIAETDAG
NLSAYIPTNLISITDGQIVLDSRLFAANQRPADVGLSVSRVGGKAQHPALRAVSGRIRLDYSQFL
ELEMFTFRGGITDTRVKAQITRGERIRALLTQPRFSTLRLQDEVALLAALAEGVFDTLAPGLMGAV
15 RARIPAQLDAQVKDVASALAEKGVLEEGLHARLVAAVRAVAADVAATAKAGP

SEQ ID NO: 22:

Nucleotide sequence of ATP synthase subunit b atpF_1

ATGCAGATCGACTGGTGGACGCTGGGCCTGCAGACGGTCAACGTCCTCGTTCTCATCTGGCT
20 CCTGAGCCGCTTCCTGTTCAAGCCGGTGGCGCAGGTCATCGCGCAGCGCCGTGCCGAGATCG
AGAAGCTGGTGGAGGATGCGCGCGCCGCCAAGGCCCGCCGAGGCCGAGCGGGACACGG
CGAAGGCGGAGGAGGCGCGCCTTGCCGCCGAGCGCGGCGCCCGCATGGCGGCGGTGCGCA
AGGAGGCGGAGGCGCAGAAGGCGGCATTGCTGGCCGCCGCCAAGACCGAGGCCGAGGCCC
TGCACGCGGCCGCGGAAGCGGCCATCGTCCGGGCGCGGGCGAGCGAGGAGGAAGCCGCCG
25 CCGACCGCGCCAGCGCCTTGCCGTGGACATCGCCGCCAAGCTGCTGGACCGGCTGCCCCGA
CGACGCCCCGGTTCGCGGGCTTCATCGATGGCCTCGCCGAGGGGCTTGAAGCCCTGCCCCGAG
GCGAGCCGGGCGGTGATCGGCGTCGACGGCGCGCCAGTGCGCGTGACGGCCGCGCGCGCC
CTTATGCCGGCGGAGGAGGAGGCCTGCCGCACGCGGCTCTCCAGGCGCTGGGCCGTCCGG
TGACGCTGGCCGTGACCATCGACCCCGCCCTCATCGCCGGCCTGGAGATGGAGACGCCCCAC
30 GCGGTGGTGCGAATTCCTTCAAGGCCGATCTCGACCGCGTCACCGCGGCGCTCACCCATCA
TGGGACCTGA

SEQ ID NO: 23:

Amino acid sequence of ATP synthase subunit b atpF_1

MQIDWWTLGLQTVNVLVLIWLLSRFLFKPVAQVIAQRRAEIEKLVEDARAATAAAEAERDTAKAE
35 EARLAAERGARMMAVAKEAEQAALLAAKTEAEALHAAAEAAIVRARASEEEAAADRASRLAV

DIAAKLLDRLPDDARVAGFIDGLAEGLEALPEASRAVIGVDGAPVRVTAARALMPAEEEEACRTRLS
QALGRPVTLAVTIDPALIAGLEMETPHAVVRNSFKADLDRVTAALTHHGT

SEQ ID NO: 24:

5 Nucleotide sequence of ATP synthase subunit c, sodium ion specific atpE_1
ATGACTGTCGAGATGGTCAGCATCTTCGCGGGCGGCTCGCCGTCTCTTCGGCGCCATCGG
GCCGGCCCTGGGCGAGGGCCGGGCGGTGGCCGCGGCCATGGACGCCATCGCCCGCCAGCC
GGAGGCGGCCGAACCTTGTCGCGCACGCTCTTCGTCGGCCTCGCCATGATCGAGACCATGG
CGATCTACTGCCTGGTGATCGCGCTCCTGGTGCTCTTCGCCAATCCGTTCTGAAGTGA

10

SEQ ID NO: 25:

Amino acid sequence of ATP synthase subunit c, sodium ion specific atpE_1
MTVEMVSIFAAALAVSFGAIGPALGEGRAVAAAMDIAIRQPEAAGTLSRTLFVGLAMIETMAIYCL
VIALLVLFANPFVK

15

SEQ ID NO: 26:

Nucleotide sequence of ATP synthase subunit a atpB_1
ATGGGCTCGCCGCTGATCCTCGAACCCCTGTTCCATATCGGGCCCGTGCCCATCACCGCGCC
GGTGGTGGTCACCTGGCTCATCATGGCCGCCTTCATTGGGCTGGCGCGGCTCATCACCCGGA
20 AGCTTTCCACCGATCCCACCCGGACCCAGGCGGCGGTGGAACGGTGCTGACCGCCATCGAT
TCCCAGATCGCCGACACCATGCAGGCCGATCCCGCGCCTTATCGCGCGCTCATCGGCACCAT
CTTCCTTTATGTGCTGGTGGCCAACTGGTCTCGCTCATCCCGGGCATCGAGCCGCCACGG
CGCATATCGAGACCGATGCGGCGCTCGCTTTCATCGTGTTCCGCCACCATCGGGTTCGGG
TTGAAGACAAGGGGTGTGAAGGGCTATCTCGCCACCTTCGCCGAACCCCTCCTGGGTGATGAT
25 CCCGCTCAATGTGGTGGAGCAGATCACCCGGACCTTCTCGCTCATCGTGCGCCTGTTCCGGCA
ACATCATGAGCGGGGTGTTCTGGTGGCATCATCCTGTCCCTCGCCGGGCTGCTGGTGCCC
ATCCCCCTCATGGCGCTCGATCTCCTGACCGGCGCCGTGCAGGCCTACATCTTCGCGGTGCT
GGCCTGCGTGTTTCATCGGCGCGGCCATTGGCGAGGCGCCGGCAAAGCCCCAATCGAAGGAG
CCAGGGAAAACATCATGA

30

SEQ ID NO: 27:

Amino acid sequence of ATP synthase subunit a atpB_1
MGSPLILEPLFHIGVPITAPVVVTWLIMAAFIGLARLITRKLSTDPTRTQAAVETVLTADSQIADT
MQADPAPYRALIGTIFLYVLVANWSSLIPGIEPPTAHMETDAALAFIVFAATIGFGLKTRGVKGYLAT
35 FAEPSWVMIPLNVEQITRTFSLIVRLFNGIMSGVFVVGIIISLAGLLVPIPLMALDLLTGAVQAYIF
AVLACVFIGAAIGEAPAKPQSKEPGKTS

ACGAGCATGCGCGCTCGCCAACGAGGTGCGCCGGACCATCGAGCATTATCGCGAGCTTCAG
GATGTGATCTCGCTGCTGGGCATGGAGGAATTGGGCACCGAGGATCGCCGCATCGTGGAGC
GGGCGCGCCGGCTCCAGCGCTTCCTCACCCAGCCCTTCACGGTCACCGAGGCCTTCACCGGC
GTGCCCCGGCCGCTCGGTGGCCATCGCCGACACCATCGCCGGCTGCAGGATGATCCTGTCCG
5 GCGCCTGCGACGACTGGCAGGAAAGCGCCCTCTACATGGTGGGCACCATCGACGAGGCCCCG
CCAGAAGGAGGAGGCCGCTCGCGCCAAGGCGGGCAGGGCGCCCCGGCCGGGACGGCAGC
CGAGACGGCGGAGGCCGCCCGTGA

SEQ ID NO: 31:

10 Amino acid sequence of ATP synthase subunit beta atpD_1
MAAADEEAQSAAGPASGRVVAVRGAVIDIAFAQPPLPLDDALLITDGRGGTVLVEVQSHMDRH
TVRAIALQATTGLSRGLEAARVGGPVKVPVGDHVLGRLLDVTGAIGDKGGPLPADVPTRPIHHAP
PSFAAQGGTSDLFRTGIKVIDLLAPLAQGGKAAMFGGAGVGKTVLVMELIHAMVASYKGISVFAG
VGERSREGHEMLLDMTDSGVLDRTVLVYGQMNEPPGARWRVPMTALTIAEYFRDEKHQNVLLLM
15 DNIFRFVQAGAEVSGLLGRPPSRVGYQPTLASEVAALQERITSVGEASVTAIEAVYVPADDFTDPA
VTTIAAHVDSMVLSRAMAAEGMYPVDPISSSSVLLDPLIVGDEHARVANEVRRTIEHYRELQD
VISLLGMEELGTEDRRIVERARRLQRFLTQPFTVTEAFTGVPGRSVAIADTIAGCRMILSGACDDW
QESALYMGVTIDEARQKEEAARAKAGQGAPAGTAAETAEAAP

20 SEQ ID NO: 32:

Nucleotide sequence of ATP synthase subunit beta atpD_2
ATGGCGAACAAGGTCGGACGCATCACCCAGATCATCGGCGCCGTCGTCGACGTGCAGTTCGA
CGGGCATCTGCCGGCGATTCTCAACGCGATCGAGACCACCAACCAGGGCAACCGGCTGGTG
TCGAAGTGGCTCAGCATCTCGGCGAGAACACCGTGCGCTGCATCGCCATGGATGCCACTGAA
25 GGCCTGGTGCCTGGCCAGGAGGTGGCCGACACCGATGCGCCCATCCAGGTGCCCGTGGGCG
CCGCCACCCTCGGCCGCATCATGAACGTGATCGGCGAGCCGGTGGACGAGCTGGGCCCCAT
CGAGGGCGAAGCGCTGCGCGGCATCCATCAGCCGGCCCCCTCCTATGCGGAGCAGGCCACG
GAAGCTGAGATCCTCGTCACCGGCATCAAGGTGGTGGATCTGCTGGCGCCCTATTCCAAGGG
CGGCAAGGTGGGCCTGTTGCGCGGCGCCGGCGTGGGCAAGACCGTGCTCATCATGGAGCTG
30 ATCAACAACGTGGCCAAGGCGCACGGCGGCTATTCCGTGTTGCGCGGCGTGGGTGAGCGCA
CCCGCGAGGGCAACGACCTCTACCACGAGATGATCGAGTCCAACGTGAACAAGGACCCGCAC
GAGAACAAATGGCTCGGCGGCCGGTTCCAAGTGCGCCCTGGTCTATGGCCAGATGAACGAGCC
GCCCCGGCGCCCGCGCCCGTGGCCCTCACCGGCCTACCGTCGCCGAGCATTTCCGCGAC
CAGGGCCAGGACGTGCTGTTCTTCTGTTGACAACATCTTCCGCTTACCCAGGCGGGCTCCGA
35 GGTGTCGGCGCTTCTCGGCCGCATCCCCTCGGCGGTGGGCTACCAGCCGACGCTGGCCACC
GACATGGGCCAGCTGCAGGAGCGCATCACCAACCACCAAGGGCTCCATCACCTCGGTGCA
GGCCATCTACGTGCCGGCGGACGATCTGACCGATCCGGCGCCGGCCGCTCCTTCGCCCCATC

TGGACGCCACCACGGTGCTGTCGCGCTCCATCGCGGAGAAGGGCATCTACCCGGCGGTGGA
TCCGCTGGACTCCACCTCGCGCATGCTGTCTCCCGCCATCCTCGGCGACGAGCACTACAACAC
CGCGCGCCAGGTGCAGCAGACCCTGCAGCGCTACAAGGCGCTCCAGGACATCATCGCCATCC
TGGGCATGGACGAACTCTCCGAAGAGGACAAGCTCACCGTGGCCCCGCGCCCGCAAGATCGA
5 GCGCTTCCTCTCCCAGCCCTTCCACGTGGCCGAGGTGTTACCGGTTGCCCCGCAAGCTGG
TCGACCTCGCCGACACCATCAAGGGCTTCAAGGGCCTGGTGGACGGCAAGTACGACTACCTG
CCCAGCAGGCCTTCTACATGGTGGGCACCATCGAAGAAGCCATCGAGAAGGGCAAGAAGCT
GGCGGCCGAGGCGGCCTGA

10 SEQ ID NO: 33:

Amino acid sequence of ATP synthase subunit beta atpD_2

MANKVGRITQIIGAVVDVQFDGHLPAILNAIETTNQGNRLVLEVAQHLGENTVRCIAMDATEGLV
RGQEVADTDAPIQVPVGAATLGRIMNVIGEPVDELGPIEGEALRGIHQPAPSYAEQATEAEILVTG
IKVVDLLAPYSKGGKVGLFGGAGVGKTVLIMELINNVAKAHGGYSVFAGVGERTRENDLYHEMI
15 ESNVNKDPHENNGSAAGSKCALVYGQMNEPPGARARVALTGLTVAEHFRDQGQDVLFFVDNIFR
FTQAGSEVSALLGRIPSAVGYPQLATDMGQLQERITTTTKGSITSVQAIYVPADDLTDPAAPASF
AHLDTTVLSRSIAEKGIYPAVDPLDSTRMLSPAILGDEHYNTARQVQQLQRYKALQDIAILG
MDELSEEDKLTVARARKIERFLSQPFHVAEVFTGSPGKLVDLADTIKFKGLVDGKYDYLPEQAFY
MVGTIIEEAIEKGKKLAEEAA

20

SEQ ID NO: 34:

Nucleotide sequence of ATP synthase gamma chain atpG_2

ATGGCGAGTCTGAAGGACCTGAGAAACCGCATTGCCTCGGTGAAGGCGACGCAGAAGATCAC
CAAGGCGATGCAGATGGTCGCCGCGGCGAAGCTGCGTCGCGCCAGGCGGCGGCTGAAGC
25 GGCCCGTCCCTATGCGGAACGCATGGAGACGGTGCTCGGAAATCTTGCTCCGGCATGGTGG
TGGGCGCGCAGGCGCCTGTTCTCATGACCGGGACGGGCAAGAGCGACACCCACCTGCTGCT
GGTGTGCACCGGCGAGCGCGCCTGTGCGGCGCCTTCAACTCGTCCATCGTGCGCTTCGCC
GCGAGCGGGCGCAGCTGCTGCTGGCCGAGGGCAAGAAGGTGAAAATCCTGTGCGTGGGCCG
CAAGGGCCACGAGCAGCTGCGCCGCATCTACCCGGACAACATCATCGACGTGGTGGACCTGC
30 GCGCGGTGCGCAACATCGGCTTCAAGGAGGCCGACGCCATCGCCCGCAAGGTGCTGGCCCT
GCTCGATGAAGGCGCATTGACGTCTGCACGCTCTTCTACTCCCACTTCAGGAGCGTGATCGC
CCAGGTGCCGACGGCCAGCAGCTCATTCCGGCCACCTTCGACGAGCGGCCGCGCGTCCGCC
GATGCGCCGGTCTATGAATATGAGCCGGAGGAGGAGGAGATCCTCGCCGAGCTGCTGCCGC
GCAACGTGGCGGTGCAGATCTTCAAGGCCCTCCTCGAGAACCAGGCTTCTTTCTATGGCTCCC
35 AGATGAGCGCCATGGACAACGCCACGCGCAATGCGGGCGAGATGATCAAGAAGCAGACGCT
CACCTACAACCGTACCCGCCAGGCCATGATCACGAAGGAACATCGAGATCATCTCCGGCG
CCGAGGCCGTCTGA

SEQ ID NO: 35:

Amino acid sequence of ATP synthase gamma chain atpG_2

MASLKDLRNRIASVKATQKITKAMQMVAALKLRRQAQAAEAARPYAERMETVLGNLASGMVVGA
5 QAPVLMGTGKSDTHLLLVCTGERGLCGAFNSSIVRFARERAQLLAEGKKVKILCVGRKGHEQL
RRIYPDNIIDVVDLRAVRNIGFKEADAIARKVLALLDEGAFDVCTLFYSHFRSVIAQVPTAQQLIPA
TFDERPAVADAPVYEYEPPEEEILAEELLPRNVAVQIFKALLENQASFYGSQMSAMDNATRNAGEMI
KKQTLTYNRTRQAMITKELIEIISGAEEV

10 SEQ ID NO: 36:

Nucleotide sequence of ATP synthase subunit alpha atpA_2

ATGGACATTCGAGCCGCTGAAATCTCTGCCATCCTGAAAGAGCAGATCCAGAATTTGGCCAG
GAGGCGGAAGTCTCCGAGGTGGGTGAGTTCTGTCCGTGGGTGACGGCATCGCGCGCGTCT
ACGGCCTCGACAACGTCCAGGCGGGCGAGATGGTCGAGTTCGAGAACGGCACGCGCGGCAT
15 GGCGCTGAACCTCGAGCTCGACAATGTGGCATCGTGATCTTCGGTTCGACCGCGAGATCA
AGGAAGGCCAGACCGTCAAGCGGACCGGCGCCATCGTGACGCCCCCGTGGCAAGGGCCT
GCTCGGCCGCGTCGTGGACGCTCTCGGCAACCCGATCGACGGCAAGGGCCCGATCATGTTCA
CCGAGCGTCGCCGGGTGACGTGAAGGCGCCGGGCATCATCCCGCGCAAGTCGGTGACGA
GCCCATGCAGACCGGCCTGAAGGCCATCGATGCGCTCATCCCCATCGGCCGCGGCCAGCGC
20 GAGCTCATCATCGGCGACCGCCAGACCGGCAAGACCGCCGTGGCGCTCGACTCGATCCTGAA
CCAGAAGCCCATCAACCAGGGCGACGACGAGAAGGCCAAGCTCTACTGCGTCTATGTCGCGG
TGGGCCAGAAGCGTTCCACTGTCGCGCAGTTCTGTAAGGTGCTCGAGGAGCACGGCGCGCT
GGAATATTCCATCGTCGTCGCCGCCACCGCCTCGGACGCGGCCCCCATGCAGTTCCTGGCGC
CGTTCACCGGCACCGCCATGGGCGAGTATTTCCGCGACAACGGCATGCACGCCCTCATCATC
25 CATGATGACCTGTCCAAGCAGGCCGTGGCCTACCGCCAGATGTCGCTGCTGCTGCGCCGCCC
GCCGGGCCGCGAGGCCTATCCCGGCGATGTGTTCTACCTGCACTCCCGCCTCTTGAGCGCG
CCGCCAAGCTCAATGACGAGCACGGCGCCGGCTCGCTGACCGCCCTGCCGGTGATCGAGAC
CCAGGCCAACGACGTGTCGGCCTACATCCCGACCAACGTGATCTCCATCACCGACGGTCAGA
TCTTCCTTGAATCCGATCTGTTCTACCAGGGCATCCGCCCCGGCGGTGAACGTGGGCCTGTG
30 GTGTCGCGCGTGGGCTCTTCGGCCAGATCAAGGCGATGAAGCAGGTGGCCGCAAGATCA
AGGGCGAGCTCGCCAGTATCGCGAGCTGGCGGCCTTCGCCAGTTTCGGTTCGGACCTGGA
CGCGGCCACCCAGAAGCTGCTGAACCGCGGCGCCCGCCTCACCGAGCTGCTGAAGCAGAGC
CAGTTCTCGCCCTCAAGGTGGAGGAGCAGGTGGCGGTGATCTATGCCGGCACCAATGGCTA
TCTCGATCCGCTGCCGGTCTCCAAGGTGCGCGAGTTCGAGCAGGGTCTGCTCCTGTGCTGCG
35 GCTCGCAGCATCCGGAGATCCTGGACGCCATCCGCACGTCCAAGGAGCTTTCCAAGGACACC
GCCGAGAAGCTGACGAAGGCCATCGACGCCTTCGCCAAGAGCTTCTCCTGA

SEQ ID NO: 37:

Amino acid sequence of ATP synthase subunit alpha atpA_2

MDIRAAEISAILKEQIQNFGQEAEVSEVGQVLSVGDGIARVYGLDNVQAGEMVEFENGTRGMAL
NLELDNVGIVIFGSDREIKEGQTVKRTGAIVDAPVGKLLGRVVDALGNPIDGKGPIMFTERRRV
5 DVKAPGIIPRKSVHEPMQTGLKAIDALIPIGRGQRELIIGDRQTGKTAVALDSILNQKPINQGDDE
KAKLYCVYVAVGQKRSTVAQFVKVLEEHGALEYSIVVAATASDAAPMQFLAPFTGTAMGEYFRDN
GMHALIIHDDLSKQAVAYRQMSLLRRPPGREAYPGDVFYLSRLLERAAKLNDEHGAGSLTALP
VIETQANDVSAYIPTNVISITDGQIFLESDFYQGIRPAVNVGLSVSRVGSSAQIKAMKQVAGKIK
GELAQYRELAFAAQFGSDLDAATQKLLNRGARLTELKQSQFSPLKVEEQVAVIYAGTNGYLDPLP
10 VSKVREFEQGLLLSLRSQHPEILDAIRTSKELSKDTAEKLTKAIDAFKSF

SEQ ID NO: 38:

Nucleotide sequence of ATP synthase subunit delta atpH

GTGGCGGAAACGATCGTGTCAAGCATGGCGGGACGCTATGCGACCGCGCTGTTTCGAGCTGG
15 CGGACGAAGCCGGTGCCATCGATTCCGTCCAGGCGGATCTTGATCGCCTGTCCGGCCTTCTG
GCCGAGAGCGCGGATCTGGCGCGGCTGGTCAAGAGCCCGGTCTTCACCGCCGAGCAGCAGC
TCGGCGCGATGGCGGCCATTCTCGATCAAGCAGGCATTTCCGGCCTTGCGGGCAAATTCGTG
AAGCTGGTGGCGCAGAACCGCCGCCTGTTTCGCACTGCCGCGCATGATTGCCGAATACGCCGT
CCTGGTGGCCCGGAAGAAGGGCGAGACCTCGGCGAGCGTGACCGTTGCCACCCCCCTGAGC
20 GATGAGCATCTGGCCACGCTCAAGGCGGCCCTGGCTGAAAAGACCGGCAAGGACGTGAAGC
TCGACGTACCGTCGATCCGTCCATCCTCGGTGGTCTCATCGTGAAGCTCGGCTCGCGCATG
GTCGATGCTTCCCTGAAGACCAAACTCAATTCTATCCGGCATGCGATGAAAGAGGTCCGCTGA

SEQ ID NO: 39:

25 Amino acid sequence of ATP synthase subunit delta atpH

MAETIVSGMAGRYATALFELADEAGAIDSVQADLDRLSGLLAESADLARLVKSPVFTAEQQLGAM
AAILDQAGISGLAGKFVKLVQNRRLFALPRMIAEYAVLVARKKGETSASVTVATPLSDEHLATLK
AALAEKTGKDVKLDVTVDPSILGGLIVKLGSRMVDASLTKLNSIRHAMKEVR

30 SEQ ID NO: 40:

Nucleotide sequence of ATP synthase subunit b atpF_2

ATGACCGAAATGGAAGTGGCTGAGCTCTGGGTGCGCCATCGCCTTCCTGGTTTTTCGTAGGCCTC
CTGATCTATGCGGGCGCCACCGCGCCATCGTCTCCGCCCTGGATTCCCGCGGCTCGCGCAT
CGCCTCGGAAGTGGAGGAGGCCGTCGGCTCAAGGAAGAGGCCAGAGCTGGTGGCCGAA
35 TTCAAGCGCAAGCAGCGGAGGCCGAGGCCGAGGCCGAATCCATCGTCACCGGCGCCAAGG
CCGAGGCCGAGCGCCTCGCCGCCGAGGCCAAGGCGAAGATCGAGGATTCGTACCCGCCG
CACCAAGATGGCCGAGGACAAGATCGCCAGGCCGAGCATCAGGCTCTGGCGGACGTGAAG

TCCATCGCCGCCGAGGCGGCGGCCAAGGCGGCGCGAGGTGATCCTCGGCGCCCGAGGCCACCG
GCGCGGTGGCGGAGCGTCTGCTGTCGGGCGCCATCTCCGAGGTCAAGACCAAGCTCAACTG
A

5 SEQ ID NO: 41:

Amino acid sequence of ATP synthase subunit b atpF_2

MTEMELAELWVAIAFLVFVGLLIYAGAHRAIVSALDSRGSRIASELEEARRLKEEAQKLVAEFKRK
QREAEAEASIVTGAKAEERLAAEAKAKIEDFVTRRTKMAEDKIAQAEHQALADVKSIAAEAAA
KAAEVILGAQATGAVAERLLSGAISEVKTCLN

10

SEQ ID NO: 42:

Nucleotide sequence of ATP synthase subunit b' atpG_3

ATGATGATTGCATGGAAGCGGACCTTCGCAGTCGTGACCTTCGGGGCCGCCCTGATGGCCAT
GCCCCGTCGCGGGCGTGGTCGCAGCTGAGACTTCTCCCGCTCCGGCGGCAGTGGCGCAGGCC
15 GATCATGCGGTGCCCCACCGAGGCGGCCGCGCCAGGGCACCGCCGATGCGGCCCATGCCGCCG
CGCCGGGCGAGGCCGCCCATGGTGGCGCGGCCAAGCACGAAACCCATTTCCCGCCCTTCGA
CGGCACCACCTTCGCCTCCAGTTGCTGTGGCTCGCCGTCACCTTCGGCCTGCTTTACTACCT
CATGAGCAAGGTCACGCTGCCGCGCATCGGCCGCATCCTGGAAGAGCGCCACGACCGCATC
GCCGATGATCTGGAGGAAGCCTCCAAGCATCGCGCCGAGAGCGAGGCCGCCAGCGGGCCT
20 ATGAGAAGGCGCTGAGCGAGGCCCCGCGCGAAGGCCCATTCATCGCCGCGGAAACCCGCGA
CCGCTTGCCGCCCCACGCCGACCAACCGCAAGGCGCTGGAGAGCGAGCTCACCGCCAAG
CTGCAGGCGGCCGAGGAGCGCATCGCCACCACCAAGAGCGAAGCCCTCACCCATGTGCGCG
GCATCGCGGTGGACGCCACCCAATCCATCGTCTCCACCCTCATCGGTGTGCGCCCCGCGGCG
GCCGACGTGAAAAAGCGGTGGACGGCGCCCTGTCCCAGCACGGCCAGGCCTGA

25

SEQ ID NO: 43:

Amino acid sequence of ATP synthase subunit b' atpG_3

MMIAWKRTFAVVTFGAALMAMPVAGVVAETSPAPAAVAQADHAVPTEAAGQGTADAHAHAAP
GEAAHGGAAXHETHFPFDGTTFASQLLWLAFTFGLLYLMSKVTLPRIGRILEERHDRIADDEE
30 ASKHRAESEAAQRAYEKALSEARAKAHSIAAETRDRLAAHADTNRKALESELTKLQAAEERIATT
KSEALTHVRGIAVDATQSIVSTLIGVAPAAADVEKAVDGALSQHGQA

SEQ ID NO: 44:

Nucleotide sequence of ATP synthase subunit c atpE_2

ATGGAAGCGGAAGCTGGAAAGTTCATCGGTGCCGGCCTCGCCTGCCTCGGCATGGGTCTCGC
35 TGGCGTCGGCGTCGGTAACATCTTCGGTAACCTCCTCTCCGGCGCCCTGCGCAACCCGTCCG

CTGCCGACGCCAGTTCGCCCCGCGCTTCATCGGCGCCGCCCTCGCGGAAGGTCTCGGCATC
TTCTCGCTGGTCGTTGCGCTCGTCCTGCTGTTGCTGGCCTGA

SEQ ID NO: 45:

- 5 Amino acid sequence of ATP synthase subunit c atpE_2
MEAEAGKFIGAGLACLMGLAGVGVGNIFGNFLSGALRNPSAADGQFARAFIGAALAEGLGIFSL
VVALVLLFVA

SEQ ID NO: 46:

- 10 Nucleotide sequence of ATP synthase subunit a atpB_2
ATGACCGTCGATCCGATCCACCAGTTCGAGATCAAGCGCTACGTGGATCTGCTGAACGTCGG
CGGTGTCCAGTTCTCCTTCACCAACGCAACGGTGTTTCATGATTGGCATCGTCCTGGTGATTTT
CTTCTTCTGACTTTCGCGACACGCGGTGCGACCCCTTGTCGGGGCCGGATGCAGTCGGCGG
CGGAGCTGAGCTACGAGTTCATCGCCAAGATGGTGCGCGACGCGGCCGGCAGCGAGGGAAT
15 GGTGTTCTTTCCCTTCGTCTTCTCGCTCTTCATGTTGCTGCTGGTGGCGAACGTATTGGGGCT
CATCCCTACACCTTCACGGTGACCGCCACCTCATCGTCACCGCCGCCCTGGCGGCGACGG
TGATCCTCACCGTCATCATCTACGGCTTCGTGCGGCACGGCACCCACTTCCTGCACCTGTTTCG
TGCCGTCGGGCGTGCCGGGCTTCCTCCTGCCCTTCCTCGTGCTGATCGAGGTGGTGTCTGTTT
CTGTGCGGCCCCATCAGCCTCTCGCTGCGTCTGTTGCGCAACATGCTGGCGGGCCACATCGC
20 CCTCAAGGTGTTGCGCTTCTTCGTGCTGGGACTGGCCTCGGCCGGCGCGATCGGCTGGTTTCG
GCGCCACCCTGCCCTTCTTCATGATCGTGCGCTCACCGCGCTGGAGCTGCTGGTGGCGGTG
CTGCAGGCCTACGTGTTGCGGGTGCTGACCTCGATCTACCTCAACGACGCCATCCATCCCGGC
CACTGA

25 SEQ ID NO: 47:

- Amino acid sequence of ATP synthase subunit a atpB_2
MTVDPIHQFEIKRYVDLLNVGGVQFSFTNATVFMIGIVLVIFFFLTFATRGRRLVPGRMQSAAELSY
EFIAKMVRDAAGSEGMVFFPFVFSLFMFVLVANVLGLIPYTFTVTAHLIVTAALAATVILTVIYGFV
RHGTHFLHLFVPSGVPGLLPFLVVIEVVSFLSRPISLSLRLFANMLAGHIALKVFAFFVVG LASAGA
30 IGWFGATLPFFMIVALTALELLVAVLQAYVFAVLTSIYLNDAIHPGH

SEQ ID NO: 48:

- Nucleotide sequence of ATP synthase protein I atpI
ATGTCCGAGCCGAATGATCCATCCCGCAGGGACGGTGCGAAGGCGAAAGACGAGACGCAGG
35 ACTCCCGGCCCGGTGAGGCGGATCTTGCTCGGCGCCTCGATGCGCTCGGCACCTCCATCGGT
CAGGTCAAGTCCAGAAGCGGGGAGCCCGCGGCGACGCCGCGCAAGGACACCTCCTCGGCCT
CCGCGCGGGCCCTGGCGTTTCGGCTGGGCGCCGAGTTTGTTTCAGGCGTGCTGGTGGGCTC

GCTCATCGGCTACGGGTTGGATTATGCGTTTTCGCGATTCGCCCTGGGGGCTGATCGCCTTCAC
 GCTGATCGGCTTTGCCGCCGGCGTCCTGAACATGCTGCGCGTGGCGAACAGCGATGCCAAGC
 GCCACAGCGCGGACAGGTGA

5 SEQ ID NO: 49:

Amino acid sequence of ATP synthase protein I atpI

MSEPNDPSRRDGAKAKDETQDSRPGEADLARRLDALGTSIGQVKSRSGEPAATPRKDTSSASG
 AALAFRLGAEFVSGVLVGLSLIGYGLDYAFAISPWGIIAFTLIGFAAGVLNMLRVANSDAKRHSADR

- 10 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a nitrogenase molybdenum-iron protein alpha chain nifD_1 having the sequence set forth in SEQ ID NO:51 or a sequence having more than 60% identity, e.g. more than 70% identity, such as more than 92% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:51.

- 20 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding nitrogenase molybdenum-iron protein alpha chain nifD_2 having the sequence set forth in SEQ ID NO:53 or a sequence having more than 60% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:53.

- 25 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a nitrogenase molybdenum-iron protein beta chain nifK_1 having the sequence set forth in SEQ ID NO:55 or a sequence having more than 87% identity, e.g. more than 90% identity, such as more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:55.

- 30 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a nitrogenase molybdenum-iron protein beta chain nifK_2 having the sequence set forth in SEQ ID NO:57 or a sequence having more than 95% identity, e.g. more than 96% identity, such as more than 97% identity, e.g. more than 98%, such as more than 99% sequence identity to the sequence set forth in SEQ ID NO:57.

- 35 In another embodiment, the variant chemoautotrophic bacterial strain used in the process of the invention comprises a gene encoding a nitrogenase iron protein nifH

having the sequence set forth in SEQ ID NO:59 or a sequence having more than 98.5% sequence identity to the sequence set forth in SEQ ID NO:59.

SEQ ID NO: 50:

5 Nucleotide sequence of Nitrogenase molybdenum-iron protein alpha chain nifD_1
ATGAGTTCGCTCTCCGCCACTATTCAACAGGTCTTCAACGAGCCGGGCTGCGCGAAGAACCA
GAATAAGTCCGAGGCGGAGAAGAAGAAGGGCTGCACCAAGCAGCTGCAACCCGGCGGAGCG
GCCGGCGGCTGCGCGTTCGACGGCGCGAAGATCGCGCTCCAGCCCTTGACCGACGTCGCCC
ACCTGGTGCACGGCCCCATCGCCTGCGAAGGCAATTCTGGGACAATCGTGGCGCCAAGTCC
10 TCCGGCTCGAACATCTGGCGCACCGGCTTACCACGACATCAACGAAACCGACGTGGTGT
CGGCGGCGAGAAGCGTCTGTTCAAGTCCATCAAGGAAATCATCGAGAAGTACGACCCGCCGG
CCGTCTTCGTCTATCAGACCTGCGTCCCCGCCATGATCGGCGACGACATCGACGCGGTGTGC
AAGGCGGCCAGGGAGAAGTTCGAAAGCCGGTGATCCCGATCAATCCCCCGGCTTCGTGG
GGCCGAAGAATCTCGGCAACAAGCTCGCCGGCGAGGCGCTCCTCGACCATGTGATCGGCACC
15 GAGGAGCCCGATTACACGACGGCCTACGACATCAACATCATCGGCGAATACAATCTCTCCGG
CGAGTTGTGGCAGGTGAAGCCGCTGCTGGACGAGCTGGGCATCCGCATCCTCGCTGCATCT
CCGGCGACGGGAAGTACAAGGATGTGGCGTCTCCACCGCGCCAAGGCGGCGATGATGGT
GTGCTCCAAGGCCATGATCAACGTGGCCCGCAAGATGGAGGAGCGCTACGACATCCCCTTCT
TCGAAGGCTCCTTCTACGGCATCGAGGATAGCTCCGATTCCCTGCGCGAGATTGCGCGCATG
20 CTCATCGAGAAGGGCGCCGATCCGGAGCTGATGGACCGCACCGAGGCGCTGATTGAGCGGG
AAGAGAAGAAGGCGTGGGACGCCATCGCCGCCTACAAGCCCCGCTTCAAGGACAAGAAGGT
GCTGCTCATCACGGCGGCGTGAATCCTGGTGGTGGCAGCGCTCCAGGAAGCCGGC
CTCGAACTGGTGGGCACCTCGGTGAAGAAGTCCACCAAGGAGGACAAGGAGCGCATCAAGG
AACTGATGGGCCAGGACGCCACATGATCGACGACATGACGCCCCGCGAAATGTACAAGATG
25 CTGAAGGACGCCAAGGCGGACATCATGCTCTCGGGCGGGCGCTCGCAATTCATCGCGCTCAA
GGCCGCCATGCCCTGGCTCGACATCAACCAGGAGCGCCACCACGCCTATATGGGCTATGTGG
GCATGGTGAAGCTGGTCTGAGGAGATCGACAAGGCGCTCTACAATCCCGTGTGGGAACAGGT
GCGCAAGCCCCGCCCCGTGGGAAAATCCGGAAGACACCTGGCAGGCCCCGTGCGCTCGCCGAA
ATGGAGGCGGAGGCCGCCGCGCTCGCCGCCGATCCGGTGCGCGCGGAAGAGGTGCGCCGG
30 TCCAAGAAGATCTGCAATTGCAAGAGCGTCGACCTCGGAACCATGAGGACGCCATCAAGGC
TCACGCGCTGACCACCGTGGAGGGTGTGCGAGAGCACACCAATGCCTCGGGAGGCTGCGGA
GCCTGCAGCGGGCGGATCGAGGAGATCTTCGAGGCCGTGGGCGTTGTGCGCCGCCCGCCTC
CCGCGGAGGCCGCCCGTCTCCGAGGAGATCGCGCCGATCCGCTCGCTGCGGAGGAAAA
GCGCCGCGCCAAGAAGGCCTGCGGCTGCAAGGAGGTAGCGGTGCGCACCATGAGGATGCC
35 ATCCGCGCCAAGGGTCTGCGAAACATCGCGGAGGTGCGTGCGGCCACCGATGCCAACACCG
GCTGCGGCAATTGCCAGGAGCGGGTGGAGGGCATCCTCGACCGGGTTCTCGCCGAGGCGGC
CTCAGAACTCCAGGCGGCGGAATAG

SEQ ID NO: 51:

Amino acid sequence of Nitrogenase molybdenum-iron protein alpha chain nifD_1

MSSLSATIQQVFNEPGCAKNQNKSEAEKKKGCTKQLQPGGAAGGCAFDGAKIALQPLTDVAHLV
5 HGPIACEGNSWDNRGAKSSGSNIWRTGFTTDINETDVVFGGEKRLFKSIKEIIEKYDPPAVFVYQ
TCVPAMIGDDIDAVCKAAREKFGKPVIPINSPGFVGPKNLGNKLAGEALLDHVIGTEEPDYTTAYD
INIIGEYNLSGELWQVKPLLELGIRILACISGDGKYKDVASSHRAKAAMMVCSKAMINVARKME
ERYDIPFFEGSFYGIEDSSDSLREIARMLIEKGADPELMDRTEALIEREEKKAWDAIAAYKPRFKDK
KVLITGGVKSWSVVAALQEAGLELVGTSVKKSTKEDKERIKELMGQDAHMIDDMTPREMYKML
10 KDAKADIMLSGGRSQFIALKAAMPWLDINQERHHAYMGYVGMVKLVVEIDKALYNPVWEQVRKP
APWENPEDTWQARALAEMEAEEAALADPVRAEVRRSKKICNCKSVDLGTIEDAIKAHALTTVE
GVREHTNASGGCGACSGRIEEIFEAVGVVAAPPPAEAAPSPQEIAPDPLAAEEKRRRAKKACGCKE
VAVGTIEDAIRAKGLRNIAEVRAATDANTGCGNCQERVEGILDRVLAEAASELQAAE

15 SEQ ID NO: 52:

Nucleotide sequence of Nitrogenase molybdenum-iron protein alpha chain nifD_2

ATGAGTGTGCGACAGTCCCAGAGCGTCGCCGAGATCAAGGCGCGCAACAAGGAACATCATCGA
AGAGGTCTCAAGGTCTATCCCGAGAAGACCGCCAAGCGCCGCGCCAAGCACCTGAACGTCC
ACGAAGCCGGCAAGTCCGACTGCGGCGTGAAAGTCCAACATCAAGTCCATCCCGGGCGTGATG
20 ACCATCCGCGGTTGCGCTTATGCCGGCTCCAAGGGTGTGGTGTGGGGTCCCATCAAGGACAT
GATCCACATCTCCACGGCCCCGTGGGCTGCGGCCAGTATAGCTGGGCCGCCCGCCGCAACT
ACTATATCGGCACGACCGGCATCGACACCTTCGTGACGATGCAGTTCACCTCCGACTTCCAGG
AGAAGGACATCGTCTTCGGCGGCGACAAGAAGCTCGCCAAGATCATGGACGAGATCCAGGAG
CTGTTCCCGCTGAACAACGGCATCACCGTTCAGTCCGAGTGCCCCATCGGCCTCATCGGCGA
25 CGACATCGAGGCCGTCTCCAAGCAGAAGTCCAAGGAGTATGAGGGCAAGACCATCGTGCCGG
TGCGCTGCGAGGGCTTCGCGGCGGTGTCCAGTCCCTGGGCCACCACATCGCCAACGACGCC
ATCCGCGATTGGGTGTTGACAAGATCGCGCCGACGCCGAGCCGCGCTTTGAGCCGACCCC
GTACGACGTGCGCATCATCGGCGACTACAATATCGGTGGTGACGCCTGGTCTCCCGTATCCT
CCTGGAGGAGATGGGCCTGCGCGTGATCGCCAGTGGTCCGGCGACGGTTCGCTCGCTGAG
30 CTGGAGGCCACCCCGAAGGCCAAGCTCAACGTGCTGCACTGCTACCGCTCCATGAACATACAT
CTCGCGCCACATGGAAGAGAAGTACGGTATCCCGTGGTGCGAGTACAACCTTCTCGGTCTTC
CAAGATCGCCGAGTCCCTGCGCAAGATCGCCAGCTACTTCGACGACAAGATCAAGGAAGGCG
CGGAGCGCGTCATCGCCAAGTATCAGCCGCTCATGGATGCGGTGATCGCGAAGTATCGTCCC
CGCCTCGAGGGCAAGACCGTGATGCTGTACGTGGGCGGCCTGCGTCCCCGTACGTCATCG
35 GCGCCTACGAGGACCTGGGCATGGAAGTGGTCGGCACGGGCTACGAGTTCGCCCATACGA
CGACTACCAGCGCACCGCCAGCACTACGTCAAGGATGGCACCATCATCTATGACGACGTGA
CCGGCTACGAGTTCGAGAAGTTCGTGAGAAGATCCAGCCGACCTGGTCTGGTTCGGGCATC

AAGGAAAAGTACGTCTTCCAGAAGATGGGCGTGCCGTTCCGCCAGATGCACTCCTGGGACTA
CTCGGGCCCCGTACCACGGCTATGACGGCTTCGCGATCTTCGCGCGCGACATGGACATGGCCA
TCAACAGCCCCGTGTGGAAGATGACCCAGGCTCCGTGGAAGAGCGTCCCCAAGCCGACGATG
CTCGCGGCTGAATGA

5

SEQ ID NO: 53:

Amino acid sequence of Nitrogenase molybdenum-iron protein alpha chain nifD_2

MSVAQSQSVAEIKARNKELIEEV LKVYPEKTAKRRAKHLNVHEAGKSDCGVKSNIKSIPGVMTIR
GCAYAGSKGVVWGPIKDMIHISHGVPVCGQYSWAARRNYYIGTTGIDTFVMTMQFTSDFQEKDIV
10 FGGDKKLAKIMDEIQELFPLNNGITVQSECPIGLIGDDIEAVSKQKSKEYEGKTIVPVRCEGFRGV
SQSLGHHIANDAI RDWVFDKIAPDAEPRFEPTPYDVAIIGDYNIGGDAWSSRILLEEMGLRVIAQ
WSGDGSLAELEATPKAKLNLHCYRSMNYISRHMEEKYGIPWCEYNFFGPSKIAESLRKIASYFD
DKIKEGAERVIAKYQPLMDAVIAKYRPRLEGKTVMLYVGGLRPRHVIGAYEDLGMEVVGTYEFA
HNDDYQRTAQHYVKDGTIIYDDVTGYEFKFEKIQPDLVGSIGIKEKYVFQKMGVPFRQMHSWD
15 YSGPYHGYDGF AIFARDMDMAINSPVWKMTQAPWKSVPKPTMLAAE

SEQ ID NO: 54:

Nucleotide sequence of Nitrogenase molybdenum-iron protein beta chain nifK_1

ATGGCCACCGTTTCCGTCTCCAAGAAGGCCTGCGCGGTCAACCCCTCAAGATGAGCCAGCC
20 GGTGGGCGGCGCGCTCGCCTTCATGGGCGTGCGCAAGGCCATGCCGCTGCTGCACGGCTCG
CAGGGCTGCACCTCCTTCGGCCTGGTGTCTCGTGCGCCACTTCAAGGAAGCCATCCCCAT
GCAGACCACCGCCATGAGCGAGGTGGCGACGGTTCTGGGCGGCCTTGAGAATGTGGAGCAG
GCCATTCTCAACATCTACAATCGCACCAAGCCGGAGATCATCGGCATCTGCTCCACCGGCGTC
ACCGAGACCAAGGGCGATGATGTGACGGCTACATCAAGCTGATCCGGGACAAGTATCCCCA
25 GCTGGCCGACTTCCCGCTGGTCTATGTCTCCACCCCGATTTC AAGGACGCCTTCCAGGACG
GTTGGGAGAAGACCGTGGCGAAGATGGTGGAGGCGCTGGTGAAGCCCGCCGCGACAAGCA
GAAGGACAAGACCCGCGTCAACGTCTGCCCCGGCTGCCACCTCACGCCCCGGCGATCTGGATG
AGATGCGGACCATCTTCGAGGATTCGGGGCTCACACCCTATTTCTGCGCGATCTGGCCGGCT
CGCTGGATGGGCATATCCCCGAGGACTTCTCGCCCACCACCATCGGCGGCATCGGCATCGAT
30 GAGATCGCCACCATGGGCGAGGCGGCCACACCATCTGCATCGGCGCGCAGATGCGCCGGG
CGGGCGAGGCCATGGAGAAGAAGACCGGCATTCCCTTCAAGCTGTTGAGCGCCTGTGCGG
CCTGGAGGCGAACGACGCCTTCATCATGCACCTGTGCGAGATCTCCGGCCGGCCGGTGCCGG
TGAAGTATCGCCGGCAGCGGGGCCAGCTGGTGGATGCCATGCTGGACGGCCACTTCCATCTG
GGCGGTGCGAAGGTGGCCATGGGGGCGGAGCCGGACCTGCTCTACGACGTGGGCTCCTTCC
35 TGCACGAGATGGGCGCCCACATCCTTTCGCGGTCAACCACCCAGTCGCCGGTGCTGGCG
CGCCTGCCTGCCGAGGAGGTGCTTATCGGCGACCTGGAGGATCTGGAGACCCAGGCGAAGG
CGCGCGGATGCGATCTCCTGCTACCCATTCCCATGGGCGCCAGGCGGCGGAGCGCCTCCAC

ATCCCCCTTCTACCGGATCGGCATTCCCATGTTTGACCGGCTGGGGGCGGGGCATCTGTTGTC
GGTGGGCTATCGCGGCACCCGCGACCTCATCTTCCATCTCGCCAACCTTGTGATCGCCGACCA
CGAGGAAAATCACGAGCCGACGCCCCGACACCTGGGCCACCGGCCATGGCGAGCATGCCGCC
GCCCCCACTTCCCATTGA

5

SEQ ID NO: 55:

Amino acid sequence of Nitrogenase molybdenum-iron protein beta chain nifK_1

MATVSVSKKACAVNPLKMSQPVGGALAFMGVRKAMPLHGSQGCTSFGLVLFVRHFKEAIPMQT
TAMSEVATVLGGLENVEQAILNIYNRTKPEIIGICSTGVTETKGDDVDGYIKLIRDKYPQLADFPLV
10 YVSTPDFKDAFQDGWEKTVAKMVEALVKPAADKQKDKTRVNVLPGCHLTPGDLDEMRTIFEDFG
LTPYFLPDLAGSLDGHIPEDFSPTTIGGIGIDEIATMGEEAHTICIGAQMRRAGEAMEKKTGIPFKL
FERLCGLEANDAFIMHLSQISGRPVVPKYRRQRGQLVDAMLDGHFHLGGRKVKAMGAEPDLLYDV
GSFLHEMGAHILSAVTTTQSPVLARLPAEEVLIGDLEDLETQAKARGCDLLTHSHGRQAERLHI
PFYRIGIPMFDRLGAGHLLSVGYRGTRDLIFHLANLVIADHEENHEPTPDTWATGHGEHAAPTS
15 H

SEQ ID NO: 56:

Nucleotide sequence of Nitrogenase molybdenum-iron protein beta chain nifK_2

ATGCCACAAAATGCTGACAATGTGCTCGATCACTTCGAGCTCTCCGTGGTCCCGAATACCAG
20 CAGATGCTGGCCAATAAGAAAAAGATGTTTCGAGAACCCCCGCGATCCGGCCGAAGTCGAGCG
CGTGCGGGAATGGGCGAAGACTCCTGAATACAAGGAGCTGAACTTCGCCCCGCGAGGCGCTC
ACCGTGAATCCGGCCAAGGCTTGTGAGCCGCTGGGCGCGGTGTTGTCGCGCGTTCGGCTTCGA
GAGCACGATCCCCCTTCGTGCACGGCTCGCAGGGTTGCGTCGCGTATTACCGCTCGCACCTCT
CCCCCCTTCAAGGAGCCGTCCTCCTGCGTCTCCTCGTCCATGACCGAGGATGCGGCGGTG
25 TTCGGCGGCCTCAACAACATGATTGACGGCCTCGCCAACACCTACAACATGTACAAGCCGAAG
ATGATCGCCGTCTCCACCACCTGCATGGCGGAAGTCATCGGCGACGATCTGAACGCCTTCATC
AAGACCGCGAAGGAAAAGGGCTCGGTTCCGGCCGAATACGACGTGCCCTTCGCCCACACCCC
GGCGTTTCGTGCGCAGCCATGTCACCGGCTACGACAATGCGCTCAAGGGCATCCTCGAGCACT
TCTGGGACGGCAAGGCCGGCACCGCGCCGAAGCTGGAGCGCGTTCCTCAACGAGAAGATCAA
30 CTTCATCGGCGGCTTCGACGGCTACACCGTCGGCAACACTCGCGAAGTGAAGCGCATCTTCG
AGGCGTTCGGCGCCGATTACACCATCCTCGCCGACAATTCCGAAGTGTTGACACCCCGACC
GACGGCGAGTTCCGCATGTATGACGGCGGCACGACCCTGGAGGACGCGGCGAACGCGGTGC
ACGCCAAGGCCACCATCTCCATGCAGGAATACTGCACGGAGAAGACCCTGCCCATGATCGCC
GGTCATGGCCAGGACGTGGTCGCCCTCAACCACCCCGTGGGCGTGGGCGGCACCGACAAGT
35 TCCTCATGGAGATCGCCCGCCTACCGGCAAGGAGATCCCCGAGGAGCTGACCCGCGAGCG
CGGCCGTCTCGTGGACGCTATCGCGGACTCTTCGCGGCACATCCACGGCAAGAAGTTCGCCA
TCTACGGCGATCCGGATCTGTGCCTGGGCCTCGCCGCGTTCCTGCTGGAGCTGGGCGCCGAG

CCGACCCATGTGCTGGCCACCAACGGCACCAAGAAGTGGGCCGAGAAGGTTGAGGAACTGTT
CGACTCTTCGCCGTTTCGGCGCCAACTGCAAGGTCTATCCCGGCAAGGACCTGTGGCACATGC
GCTCGCTCCTGTTCTGTGGAGCCGGTGGATTTTCATCATCGGCAACACCTACGGCAAGTATCTCG
AGCGCGACACGGGGACCCCGCTGATCCGTATCGGCTTCCCGGTGTTGACCGTCAACCACCAC
5 CACCGCCGTCCGGTGTGGGGCTATCAGGGCGGCATGAACGTCCTGATCAGATCCTCGACAA
GATCTTTGACGAGATCGACCGCAACACCAACGTGCCGGCCAAGACCGACTACTCGTTCGACAT
CATTCGTTGA

SEQ ID NO: 57:

10 Amino acid sequence of Nitrogenase molybdenum-iron protein beta chain nifK_2
MPQNADNVLDHFELFRGPEYQQMLANKKKMFENPRDPAEVERVREWAKTPEYKELNFAREALTV
NPAKACQPLGAVFVAVGFESTIPFVHGSQGCVAAYRSHLSRHFKEPSSCVSSSMTEAAVFGGLN
NMIDGLANTYNMYKPKMIAVSTTCMAEVIGDDLNAFIKTAKEKGSVPAEYDVPFAHTPAFVGSHV
TGYDNALKGILEHFWGKAGTAPKLERPNEKINFIGGFDGYTVGNTREVKRIFEAFGADYITLAD
15 NSEVFDTPDGEFRMYDGGTTLEDAANAVHAKATISMQEYCTEKLPMIAGHGQDVVALNHPVG
VGGTDKFLMEIARLTGKEIPELTRERGRVLDAIADSSAHIHGKKFAIYGDPLCLGLAAFLLELGA
EPHVLATNGTKKWAQKQELFDSSPFGANCKVYPGKDLWHMRSLLFVEPVDFIIGNTYGKYLER
DTGTPLIRIGFPVDRHHHHRRPVWGYQGGMNVLITLIDKIFDEIDRNTNPAKTDYSFDIIR

20 SEQ ID NO: 58:

Nucleotide sequence of Nitrogenase iron protein nifH

GTGGAGTCCGGTGGTCCTGAGCCGGGCGTGGGCTGCGCCGGCCGCGGCGTGATCACCTCCA
TCAACTTCCTGGAGGAGAACGGCGCCTACGAGGACATCGACTATGTGTCCTACGACGTGCTG
GGCGACGTGGTGTGCGGCGGCTTCGCCATGCCATCCGCGAGAACAAGGCGCAGGAAATCT
25 ACATCGTGATGTCCGGCGAGATGATGGCCATGTATGCGGCCAACACATCTCCAAGGGCATC
CTGAAGTATGCCAATTCGGCGGCGTGCGCCTGGGCGGGCTGGTCTGCAACGAGCGCCAGA
CCGACAAGGAGCTGGAGCTGGCGGAGGCTCTGGCGAAGAAGCTCGGCACCGAGCTGATCTA
CTTCGTGCCGCGCGACAACATCGTGCAGCATGCCGAGCTGCGCCGCATGACAGTGATCGAGT
ATGCGCCCGATTCCGCCCAGGCCAGCACTACCGGAACCTGGCCGAGAAGGTGCACGCCAAC
30 AAGGGCAACGGCATCATCCCGACCCCGATCACCATGGACGAGCTGGAAGACATGCTCATGGA
GCACGGCATCATGAAGGCCGTGGACGAGAGCCAGATCGGCAAGACCGCCGCGGAGCTCGCC
GTCTGA

SEQ ID NO: 59:

35 Amino acid sequence of Nitrogenase iron protein nifH
MESGGPEPGVGCAGRGVITSINFLEENGAYEDIDYVSYDVLGDVVCGGFAMPPIRENKAQEIIYVM
SGEMMAMYAANNISKILKYANSGGVRLGGLVCNERQTDKELELAELAKKLGTELIYFVPRDNI

VQHAELRRMTVIEYAPDSAQAQHYRNLAEKVHANKGNGIIPITMDELEDMLMEHGIMKAVDES
QIGKTAAELAV

Genetic modification of VTT-E-193585 and variants thereof, such as strain VTT-E-213595

As described above, in a further main aspect, the invention relates to general methods for genetic modification of bacterial strain VTT-E-193585 or variants thereof and to genetically-modified variants of strain VTT-E-193585. These methods are exemplified in Example 5 herein.

Accordingly, in one aspect, the invention relates to a method for genetic modification of bacterial strain VTT-E-193585 comprising the steps of:

- a) providing bacteria of strain VTT-E-193585 or a genetically-modified or mutated strain generated using bacterial strain VTT-E-193585, such as strain VTT-E-213595,
- b) introducing a nucleic acid construct into said bacteria, wherein said nucleic acid construct comprises:

- i) sequences encoding a selectable marker,
- ii) optionally further sequences to be integrated into the bacterial genome,
- iii) flanking sequences allowing homologous recombination with the bacterial genome, and,

- c) selecting a genetically-modified strain wherein said nucleic acid construct has been integrated into the bacterial genome, on the basis of the selectable marker.

In one embodiment, the nucleic acid construct is a plasmid.

In some embodiments, the selectable marker is a gene providing antibiotic resistance, such as kanamycin or tetracycline resistance and step c) is carried out by growing the bacteria in the presence of antibiotics.

In another embodiment, the selectable marker is a gene encoding a fluorescent protein and step c) is carried out on the basis of fluorescence.

The flanking sequences are typically fully identical to sequences in the bacterial genome to allow for site-specific integration of the nucleic acid construct.

In one embodiment, the method is used to disrupt a gene by insertion. Thus, the flanking sequences are chosen such that upon integration of the nucleic acid construct, an endogenous gene in the bacterial genome has been disrupted and thus inactivated.

In another embodiment, the method is used to insert a gene, such as a heterologous gene or a mutated gene, or multiple copies of a gene into the bacterial genome.

In a further aspect, the invention relates to a variant of bacterial strain VTT-E-193585 comprising a genetic modification wherein said genetic modification comprises

the disruption of a bacterial gene with a selectable marker providing antibiotic resistance, such as kanamycin or tetracycline resistance. Such variants are exemplified herein, for instance in Example 5. The invention also relates to a culture comprising such a variant, and to a process for the production of biomass, said process comprising
5 culturing such variants. The process may have any of the further features described herein above.

Downstream processing

In one embodiment, the process of the invention comprises the further step of
10 harvesting biomass produced during the culture. Biomass can e.g. be harvested by sedimentation (settling based on gravity), filtration, centrifugation or flocculation. Flocculation may require the addition of a flocculation agent. Centrifugation may e.g. be carried out using a continuous flow centrifuge.

In one embodiment, the harvested biomass is subsequently dried. Drying can
15 e.g. be performed using well known methods, including centrifugation, drum drying, evaporation, freeze drying, heating, spray drying, vacuum drying and/or vacuum filtration. The dried biomass may subsequently be used in a product, e.g. a food or feed product or feed or food ingredient.

In another embodiment, the cells of the harvested biomass are lysed. The lysate
20 may in some embodiments be separated into insoluble and soluble fractions, either or both of which may subsequently be concentrated or dried, and subsequently be used in a product, e.g. a food or a feed product.

In one embodiment, biomass is harvested and proteins are isolated from said biomass, resulting in a protein fraction and a fraction comprising non-protein
25 components. Thus, in one embodiment, the process is for the production of protein and comprises a step of culturing strain VTT-E-213595 thereof, followed by a step of harvesting biomass and a further step of isolating proteins from said biomass. In another embodiment, the process is for the production of protein and comprises culturing a variant chemoautotrophic bacterial strain, for example of the genus *Xanthobacter*, in
30 continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide, followed by a step of harvesting biomass and a further step of isolating proteins from said biomass. Depending on the method of protein isolation, the resulting fractions may be more pure or less pure. Thus, the term "protein fraction" means a fraction enriched in proteins.
35 The protein fraction may still comprise significant amounts of other components and also significant amounts of protein may end up in the "fraction comprising non-protein components".

Isolation of proteins may be performed using any suitable method. For example, in one embodiment, proteins are isolated by breaking cells mechanically and separating protein from cell debris through one or more filtration steps, e.g. successive filtration through multiple filters with decreasing pore size. Mechanical breaking may be carried out using any suitable method, e.g. ball milling, sonication, homogenization, high pressure homogenization, mechanical shearing, etc. The resulting filtered protein fraction will be enriched in proteins, but also still contain other smaller components. Protein may optionally be further purified from this fraction using any suitable method.

In another embodiment, a protein fraction is isolated by performing ethanol extraction followed by one or more filtration steps. Such methods are e.g. known from the preparation of soy bean proteins (see e.g. Chapter 5 "Soybean Protein Concentrates" in "Technology of production of edible flours and protein products from soybeans" by Berk FAO Agricultural Services Bulletin No. 97 (1992)). The resulting protein fraction will be enriched in proteins, but also still contain other components. Protein may optionally be further purified from this fraction using any suitable method.

In one embodiment, the process of the invention comprises the further step of hydrolysing the protein fraction obtained from the process of the invention to obtain amino acids and small peptides.

In one embodiment of the process of the invention, the process comprises the further step of producing a food or feed product from said biomass, from said protein fraction or from said fraction comprising non-protein components. Said further step may simply comprise incorporating said biomass, protein fraction or fraction comprising non-protein components in a food or feed product, by adding it during the production of the food or feed product. In other embodiments, further purification or modification of the biomass or fraction thereof is performed during the course of its incorporation into a food or feed product.

In a further aspect, the invention relates to a product, such as biomass, protein, or non-protein components, obtained or obtainable by the process according to the invention.

In one embodiment, the product obtained from the process of the invention comprises more than 40% protein, such as between 40% and 99% protein, e.g. between 40% and 90% protein, such as between 40% and 60% protein. In a particular embodiment, the product comprises between 25% and 75% protein, between 0% and 20% lipid and between 5% and 40% carbohydrates. In a further embodiment, the product comprises between 40% and 60% protein, between 0% and 15% lipid and between 10% and 25% carbohydrate. In an even further embodiment, the product

obtained from the process of the invention comprises between 45% and 55% protein, between 5% and 10% lipid and between 10% and 20% carbohydrates.

As described above, the invention in a further aspect relates to a food or feed product obtained or obtainable by the process according to the invention. When used
5 herein, the terms "food" and "feed" are intended to include not only conventional food and feed products, such as processed foods, but also related products, such as food and feed supplements, e.g. protein bars, powders or shakes, meat replacements, food ingredients, probiotics, prebiotics, nutraceuticals and the like. In certain embodiments, said biomass, said protein fraction or said fraction comprising non-protein components
10 is utilized in the production of a vegetarian or vegan food product.

In further aspects, the invention relates to the production of pharmaceuticals, bioactive compounds, nutraceuticals, antioxidants and/or vitamins using bacteria of strain VTT-E-193585 or variants thereof, such as the variants described herein, for example strain VTT E-213595. Bioactive compounds, nutraceuticals, antioxidants,
15 vitamins may be extracted from the biomass using methods known in the art.

Thus, in further embodiments, the invention relates to the methods for production of biomass described herein comprising a further step of isolating, such as extracting, compounds, such as bioactive compounds, nutraceuticals, antioxidants, vitamins, from said biomass or from the cultivation liquid. In further embodiment, the
20 invention relates to the use of these compounds in a pharmaceutical product or as a food supplement. In one embodiment, the extracted compound is beta-carotene (provitamin A). In another embodiment, the extracted compound is ubiquinone Q10. In another embodiment, the extracted compound is a form of heme iron e.g. cytochrome C. In another embodiment, the extracted compound is vitamin B12.

25 The invention is further illustrated with the following, non-limiting, examples:

EXAMPLES

Example 1. Isolation of bacterial strain capable of chemoautotrophic growth

A sample of 50 mL containing soil and seawater was collected in a sterile falcon tube
30 from the seashore of the Baltic sea in Naantali in Finland. Part of soil sample was mixed with 10 mL of mineral medium in a sterile Erlenmeyer flask. The medium consisted of 1 g/L NH_4OH , 0.23 g/L KH_2PO_4 , 0.29 g/L $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$, 0.005 g/L $\text{NaVO}_3 \cdot \text{H}_2\text{O}$, 0.2 g/L $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.5 g/L $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.01 g/L CaSO_4 , 0.00015 g/L $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$, 0.005 g/L MnSO_4 , 0.0005 g/L $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.0015 g/L H_3BO_3 , 0.001 g/L CoSO_4 ,
35 0.00005 g/L CuSO_4 and 0.0001 g/L NiSO_4 prepared in tap water. The suspension of soil and medium was incubated in a shaking incubator in +30 °C temperature in a sealed steel box that was flushed continuously with a gas mixture: 150 mL/min of N_2 , 18

mL/min of H₂, 3 mL/min of O₂ and 6 mL/min of CO₂. The cultivation was refreshed in seven-day intervals by taking 1 mL of suspension, which was added in sterile conditions to 9 mL of medium in Erlenmeyer flask, and then placed back into the incubation box. After the fourth dilution, there was no noticeable soil left in the suspension. The volume of the cell suspension was increased to 100 mL in order to grow biomass for bioreactor cultivation. The optical density (OD₆₀₀) of the suspension was 1.53 when it was inoculated to 190 mL of mineral medium in 15-vessel 200-mL parallel bioreactor system (Medicel Explorer, Medicel Oy, Finland). The cultivation conditions were 800 rpm agitation, +30°C temperature and the pH was set to 6.8, controlling it with 1 M NaOH.

Gas was fed through a sparger with a gas mixture consisting of 14 mL/min H₂, 3 mL/min O₂ and 6 mL/min CO₂. The head space of the reactor was flushed with 300 mL/min air. Continuous cultivation was fed with mineral medium 6 mL/h and cell suspension was drawn from the reactor via capillary keeping the volume constant at 200 mL. Cell suspension drawn from the reactor was stored at +4°C. A sample was taken from the bioreactor automatically every day, and absorbance at 600 nm was measured to monitor the growth. After 498 hours of bioreactor cultivation, samples were drawn aseptically and suspension was diluted and plated to agar mineral medium plates containing the above minerals and 2% bacteriological agar. Plates were incubated in same conditions as described above for the Erlenmeyer flasks. Colonies were then picked from agar plates and streaked to new agar plates in order to isolate one organism in one colony. This was repeated twice. Single colonies were picked and suspended into 200 µL of medium in a 96-well microtiter plate. The suspension was incubated at +30°C temperature and shaken 625 rpm in an EnzyScreen gas tight box that was flushed continuously with 150 mL/min of N₂, 18 mL/min of H₂, 3 mL/min of O₂ and 6 mL/min of CO₂. The suspension from one well was transferred to an Erlenmeyer flask and supplemented with fresh medium. Volume was increased until there was enough biomass to perform a bioreactor cultivation. The organism was deposited in the VTT culture collection as VTT-E-193585.

16S rRNA sequencing of a sample demonstrated that the sample contained only one organism. The same sample was used for Illumina NextSeq sequencing providing 1x150 bp metagenomic shotgun sequences. Using Unicycler (Wick et al, 2017 PLoS computational biology 13:e1005595), the de novo assembly was made for metagenomic sequences consisting of 101 contigs. The total genome length was 4,846,739 bp and the GC content was 67.9 %. Gene predictions and functional annotations were performed using Prokka (Seemann, 2014 Bioinformatics 30:2068). The genome annotation produced 4,429 genes. Roary pan genomic alignment (Page et al, 2015

- Bioinformatics 31:3691) grouped VTT-E-193585 among *Xanthobacter* species. The strain was therefore identified as a *Xanthobacter* sp., the closest genome being *Xanthobacter tagetidis*. Alignment-based calculation of average nucleotide identity that takes into account only orthologous fragments (OrthoANI) (Lee et al, 2016 Int J Syst Evol Microbiol 66:1100) gave the best match of 80.4% to *Xanthobacter tagetidis* (ATCC 700314; GCF_003667445.1), whereas the proposed species boundary cut-off is 95-96% (see e.g., Chun et al., 2018 Int J Syst Evol Microbiol, 68: 461-466). *Xanthobacter autotrophicus* Py2 gave a match of 79.6%, while the match for *Xanthobacter* sp. 91 was 79.0%. It could thus be concluded that the isolated bacterial strain deposited as VTT-E-193585 belongs to the Phylum: Proteobacteria; to the Class: Alpha Proteobacteria; and to the Order: Rhizobiales. The most probable Family is Xanthobacteraceae, and the Genus *Xanthobacter*. The VTT-E-193585 bacterial strain could not be assigned unequivocally to any known species.
- A search for putative antimicrobial resistance genes was performed. The ABRicate (<https://github.com/tseemann/abricate>) tool was used to search the genome against the Arg-Annot, NCBI, ResFinder, the ecOH, Megares and VFDB databases using blastn or blastp. A threshold of 50 % was set for both identity and coverage, both on nucleotide and protein level. Only two putative antimicrobial resistance genes were identified. These two genes did not contain amino-acid changes linked to antibiotics resistance and thus a resistant phenotype is not expected.

Example 2. Pilot cultivation and analysis of isolated bacterial strain

- The isolated bacterial strain deposited as VTT-E-193585 was cultivated in a conventional 200-liter stirred tank bioreactor (MPF-U, Marubishi Ltd, Japan). Mixing was performed with Rushton-type impellers rotating at 400 rpm. Temperature in the cultivation was maintained at +30°C. pH was maintained at 6.8±0.2 by adding 8 M NaOH or 3.6 M H₃PO₄ by software control. Cultivation medium contained 1 g/L NH₄OH, 0.23 g/L KH₂PO₄, 0.29 g/L Na₂HPO₄ · 2 H₂O, 0.005 g/L NaVO₃ · H₂O, 0.2 g/L FeSO₄ · 7 H₂O, 0.5 g/L MgSO₄ · 7 H₂O, 0.01 g/L CaSO₄, 0.00015 g/L Na₂MoO₄ · 2 H₂O, 0.005 g/L MnSO₄, 0.0005 g/L ZnSO₄ · 7 H₂O, 0.0015 g/L H₃BO₃, 0.001 g/L CoSO₄, 0.00005 g/L CuSO₄ and 0.0001 g/L NiSO₄ prepared in tap water. A mixture containing 1.8-10.5 L/min hydrogen gas, 0.6-2.5 L/min oxygen gas and 1.8-5 L/min carbon dioxide gas was supplied constantly as the main source of energy and carbon. Dissolved oxygen level was maintained at 7.2±0.5 % by adjusting the gas mixture composition. The inoculum for the cultivation was prepared as described in Example 1. Growth was monitored by taking samples manually and analysing the cell density as optical density by measuring absorbance at

600 nm (Ultrospec 2100 pro UV/visible spectrophotometer, Biochrom Ltd., England) and by measuring cell dry weight (CDW) by drying in oven overnight at 105°C. Optical density was also monitored by using an in situ absorbance probe (Trucell 2, Finesse Ltd, USA). A growth curve of the cultivation is presented in Figure 1. The maximum growth rate in batch phase was 0.06 h⁻¹. The maximum cell density was 4.5 g_{CDW}/L at 92 h. After 92 h of cultivation, feed of fresh cultivation medium as described above was started at a dilution rate of 0.01 h⁻¹. During the continued feed, the cell density was on average 2.9 g_{CDW}/L. Cultivation liquid was constantly collected to a cooled (+10°C) tank from which it was fed in 300-liter batches to a continuous centrifugal separator (BTPX-205, Alfa-Laval AB, Sweden). The cell-containing slurry collected from the separator was fed into an atmospheric double drum dryer (Buflovak 6x8 ADDD, Hebel process solutions LLC., USA), heated with 4 bar steam and drums rotating at 3.5 rpm. This resulted in dried cell powder with approximately 96% dry matter content. Analysis results of the dried cell powder are presented in Table 1 for the proximate composition, in Table 2 for the amino acid composition, in Table 3 for the fatty acid composition, and in Table 4 for the vitamin content. Analyses demonstrate that the dried cell powder has high protein content with all the essential amino acids. It also contains more unsaturated than saturated fatty acids and a lot of B-group vitamins. Peptidoglycan content was only 0.002 mg/g_{CDW} and lipopolysaccharide content was 0.01 mg/g_{CDW}. It would be beneficial that these concentrations would be as small as possible. In comparison, in a commercial lactic acid bacteria preparation analysed at the same time, the peptidoglycan content was 0.244 mg/g_{DW} and the lipopolysaccharide content was 0.015 mg/g_{DW}. Cytotoxicity and genotoxicity assays were performed using the supernatant samples of cultivation. No cytotoxicity against HepG2 or HeLa229 human cell lines was observed. No genotoxicity against *Escherichia coli* WP2 *trp*⁻ or CM871 *uvrA recA lexA* strains was observed.

Table 1. Analysis results of dried cell powder of isolated bacterial strain deposited as VTT-E-193585.

Parameter	Method	Unit	Value
Moisture	Drying at 103 °C	g/100g	3.3
Protein	Kjeldahl (N x 6,25)	g/100g	72.2
Fat	Weibull-Stoldt	g/100g	6.0
Saturated fatty acids	Calculation based on Table 3	g/100g	1.8
Mono-unsaturated fatty acids	Calculation based on Table 3	g/100g	3.8
Polyunsaturated fatty acids	Calculation based on Table 3	g/100g	0.4

Omega 3 fatty acids	Calculation based on Table 3	g/100g	< 0.01
Omega 6 fatty acids	Calculation based on Table 3	g/100g	0.4
Dietary fibres	Gravimetric	g/100g	10.5
Ash	Ashing at 550 °C	g/100g	5.8
Glucose	HPLC-ELSD	g/100g	< 0.15
Fructose	HPLC-ELSD	g/100g	< 0.1
Sucrose	HPLC-ELSD	g/100g	< 0.1
Lactose	HPLC-ELSD	g/100g	< 0.25
Maltose	HPLC-ELSD	g/100g	< 0.2
Total sugars	Calculation	g/100g	< 0.8
Carbohydrates	Calculation	g/100g	2.2
Energy	Calculation	kJ/100g	1572
Energy	Calculation	kcal/100g	373

Table 2. Amino acid composition of dried cell powder of isolated bacterial strain deposited as VTT-E-193585.

Parameter	Method	Unit	Value
Lysine	Ion chromatography	%	3.95
Methionine	Ion chromatography	%	1.60
Cystine	Ion chromatography	%	0.39
Aspartic	Ion chromatography	%	6.82
Threonine	Ion chromatography	%	3.47
Serine	Ion chromatography	%	2.75
Glutamic	Ion chromatography	%	8.84
Proline	Ion chromatography	%	3.14
Glycine	Ion chromatography	%	4.40
Alanine	Ion chromatography	%	6.94
Valine	Ion chromatography	%	4.96
Isoleucine	Ion chromatography	%	3.34
Leucine	Ion chromatography	%	6.08
Tyrosine	Ion chromatography	%	2.99
Phenylalanine	Ion chromatography	%	4.58
Histidine	Ion chromatography	%	1.66
Arginine	Ion chromatography	%	4.96
Tryptophan	HPLC	%	1.34

Table 3. Fatty acid composition of dried cell powder of isolated bacterial strain deposited as VTT-E-193585.

Parameter	Method	Unit	Value
C16:0 (Palmitic acid)	GC-MS	%	24.8
C16:1 (Palmitoleic acid)	GC-MS	%	3.0
C18:0 (Stearic acid)	GC-MS	%	4.4
C18:1n9 (Oleic acid)	GC-MS	%	59.9
C18:2n6 (Linoleic acid)	GC-MS	%	6.1
C18:3n3 (alpha-Linolenic acid)	GC-MS	%	0.4

5 **Table 4. Vitamin content of dried cell powder of isolated bacterial strain deposited as VTT-E-193585.**

Parameter	Method	Unit	Value
Vitamin A (RE)	HPLC	IU/100 g	<100
Vitamin E (TE)	HPLC	mg/100 g	0.33
Vitamin D3	HPLC	IU/100 g	<10
Vitamin D2	HPLC	IU/100 g	21.6
Vitamin C	HPLC	mg/100 g	<1
Thiamine chloride Hydrochloride	LC-MS/MS	mg/100 g	0.9
Vitamin B1 (Thiamine)	Calculation	mg/100 g	0.708
Vitamin B2 (Riboflavin)	HPLC	mg/100 g	6.27
Pyridoxine hydrochloride	HPLC	mg/100 g	3.39
Vitamin B6 (Pyridoxine)	Calculation	mg/100 g	2.79
Vitamin B12	LC-MS/MS	µg/100 g	224
Choline chloride	LC-MS/MS	mg/100 g	14.3
Biotin	LC-MS/MS	µg/100 g	15.6
Folic acid	Microbiological	µg/100 g	1270
Niacin (Vitamin B3)	Microbiological	mg/100 g	23.2
Pantothenic acid	Microbiological	mg/100 g	6.53

Example 3. Cultivation of isolated bacterial strain on different nitrogen sources.

The isolated bacterial strain deposited as VTT-E-193585 was cultivated in a 15-vessel parallel bioreactor system at 200 mL volume (Medicel Explorer, Medicel Oy, Finland). Mixing was performed with Rushton-type impellers rotating at 800 rpm. The temperature in the cultivation was maintained at +30°C. pH was maintained at 6.8 by adding 1 M NaOH. The cultivation medium contained 0.23 g/L KH_2PO_4 , 0.29 g/L $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$, 0.005 g/L $\text{NaVO}_3 \cdot \text{H}_2\text{O}$, 0.2 g/L $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.5 g/L $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.01 g/L CaSO_4 , 0.00015 g/L $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$, 0.005 g/L MnSO_4 , 0.0005 g/L $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$, 0.0015 g/L H_3BO_3 , 0.001 g/L CoSO_4 , 0.00005 g/L CuSO_4 and 0.0001 g/L NiSO_4 prepared in tap water. Furthermore, the nitrogen source was varied in the cultivations so that four cultivations contained 18.7 mM NH_4OH , four cultivations contained 9.34 mM urea ($\text{OC}(\text{NH}_2)_2$), four cultivations contained 18.7 mM potassium nitrate (KNO_3), and three cultivations were left without nitrogen source in the medium. A mixture containing 22 mL/min hydrogen gas, 3.2 mL/min air and 6.4 mL/min carbon dioxide gas was supplied constantly as the main source of energy and carbon. Thus, with air, all cultivations were also supplied with nitrogen gas. Growth was monitored by taking samples automatically and analysing the cell density as optical density by measuring absorbance at 600 nm (Ultrospec 2100 pro UV/visible spectrophotometer, Biochrom Ltd., England). Growth curves of the cultivations are presented in Figure 2. Growth on ammonia and urea were comparable. Growth on nitrate or nitrogen gas was clearly slower than on ammonia or urea. Towards the end of the cultivation, the growth on nitrate was better than growth on nitrogen gas as the only source of nitrogen. There was nonetheless growth also in the cultivations in which nitrogen gas was the only source of nitrogen demonstrating that isolated bacterial strain deposited as VTT-E-193585 is capable of nitrogen fixation.

Example 4. Characterization of antibiotic susceptibility

Antibiotic susceptibility of gentamicin, kanamycin, streptomycin, tetracycline, ampicillin, ciprofloxacin, colistin and fosfomycin for the isolated bacterial strain deposited as VTT-E-193585 was analysed according to CLSI M07-A111 standard (Clinical and laboratory standards institute. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 11th ed. CLSI standard M07, 2018) with hand-made microdilution plate for ampicillin, ciprofloxacin and colistin, with VetMIC Lact-1 plate (SVA National Veterinary Institute, Uppsala, Sweden) for gentamicin, kanamycin, streptomycin and tetracycline using broth microdilution method and for fosfomycin using agar dilution method in aerobic conditions at $+35 \pm 2^\circ\text{C}$ for 48 ± 1 hours using cation-adjusted Mueller Hinton Broth -medium (LabM, LAB114, cations Mg^{2+} and Ca^{2+} added separately). *Escherichia coli* ATCC 25922 was used as quality control strain and it was incubated in aerobic conditions, at $+35 \pm 2^\circ\text{C}$ for 18 ± 2 hours. Results of antibiotic susceptibility of

strains are presented in Table 5. The isolation bacterial strain was found to be generally sensitive to antibiotics. For gentamicin, kanamycin, streptomycin and tetracycline minimum inhibitory concentration (MIC) values for VTT-E-193585 were lower or comparable to *E. coli* ATCC 25922, while for ampicillin, ciprofloxacin, colistin and fosfomycin the MIC values were higher in VTT-E-193585.

Table 5. Minimum Inhibitory Concentration (MIC, µg/ml) -values of antibiotics for VTT-E-193585 strain and *Escherichia coli* ATCC 25922

	VTT-E-193585 48 h ± 1 h	<i>E. coli</i> ATCC 25922 18 h ± 2 h
Gentamicin	0.5	0.5
Kanamycin	2	4
Streptomycin	0.5	4
Tetracycline	0.25	1
Ampicillin	16	8
Ciprofloxacin	0.06	0.008
Colistin	4	2
Fosfomycin	32	0.5

Example 5. Construction of *phaC* knockout strains

10 *Media compositions per 1 L of liquid.*

DSM81-LO4 (DSM)

KH ₂ PO ₄	2.3 g
Na ₂ HPO ₄ ·2H ₂ O	2.9 g
Na ₂ SO ₄	5.45 g
(NH ₄) ₂ SO ₄ *	1.19 g
MgSO ₄ ·7H ₂ O	0.5 g
CaSO ₄ ·2H ₂ O	11.7 mg
MnSO ₄ ·H ₂ O	4.4 mg
NaVO ₃	5 mg
NaHCO ₃	0.5 g
Ammonium iron (III) citrate	5 mg
ZnSO ₄ ·7H ₂ O	0.5 mg
H ₃ BO ₃	1.5 mg
CoCl ₂ ·6H ₂ O	1 mg
CuCl ₂ ·2H ₂ O	50 µg
NiCl ₂ ·6H ₂ O	0.1 mg

Na ₂ MoO ₄ ·2H ₂ O	0.15 mg
Riboflavin	0.5 mg
Thiamine-HCl·2 H ₂ O	2.5 mg
Nicotinic acid	2.5 mg
Pyridoxine-HCl	2.5 mg
Ca-pantothenate	2.5 mg
Biotin	5 µg
Folic acid	10 µg
Vitamin B ₁₂	50 µg

*altered for GC-MS analysis

Lysogeny Broth (LB)

Yeast Extract	5g
Tryptone	10 g
NaCl	10 g
Agar (for plates)	15 g

5 **Super optimal broth (SOB)**

Tryptone	20 g
Yeast extract	5 g
NaCl	0.58 g
KCl	0.18 g
MgCl ₂	0.95 g
MgSO ₄	1.20 g

Super optimal broth with ethanol (SOBE)

5.8 mL of ethanol / 1 L of SOB

10 **Tryptic Soy Agar (TSA)**

BBLTM TrypticaseTM Soy Agar (BD) 40 g

Cultivation methods

All *SoF1* cultures in heterotrophic growth conditions were cultivated with 220 rpm shaking at 30°C in 10 mL volume of super optimal broth supplemented with 100 mM of ethanol (SOBE). All *SoF1* cultures in autotrophic conditions were cultivated in DSM81-LO4 media (DSM) with 136 rpm shaking and at the same temperature and volume as in heterotrophic growth conditions. The gas composition in autotrophic conditions was the following: 44% CO₂, 26% N₂, 22% H₂, 7% O₂, and 1% other gases.

20 *Escherichia coli* strains

Used *E. coli* strains and their relevant characteristic are summarized in Table 6. All strains were grown on lysogeny broth (LB). Antibiotics were used in the following concentrations: 100 µg/mL of ampicillin (AMP), 50 µg/mL of kanamycin (KAN) and 10 µg/mL of tetracycline (TET).

5 **Table 6. *E. coli* strains used**

<i>E. coli</i> strain	Relevant characteristics
TOP10	Electrocompetent used for plasmid construction
S17-1 (Simon et al. 1983 Nat Biotechnol 1:784-791)	Chemically competent, chromosomally integrated RP4 derivative
JM109(DE3)	Nalidixic acid resistance

Plasmid construction

The sequencing of the bacterial genome of strain VTT-E-193585 described in Example 1 identified genes *phaC1* (SEQ ID NO:60, encoding the protein set forth in SEQ ID NO:62) and *phaC2* (SEQ ID NO:61, encoding the protein set forth in SEQ ID NO:63) with homology to *phaC* genes found from other *Xanthobacter spp.*, encoding for polyhydroxyalkanoate (PHA) synthases.

Two plasmids were constructed to target deletion of *phaC1* and *phaC2* genes in the genome of *SoF1* (Table 7). The flanking 1000 bps left (LHA) and right (RHA) homology arms of *phaC1* and *phaC2* were amplified from the genomic DNA of *SoF1* with oligos (8 oligos). Both plasmids were constructed from pUC57 with Gibson assembly. Kanamycin resistance gene (*kan*), tetracycline resistance gene (*tet*) and the mobilization region (*mob*) sequences were the same as used in plasmids described in Van den Bergh et al. 1993 J Bacteriol 175:6097-6104.

20 **Table 7. Summary of plasmids constructed**

Plasmid	Backbone	Deletion cassette	<i>phaC</i> target
pSF001	pUC57m (<i>amp</i> , <i>mob</i>)	LHA- <i>kan</i> -RHA	<i>phaC1</i>
pSF002	pUC57m (<i>amp</i> , <i>mob</i>)	LHA- <i>tet</i> -RHA	<i>phaC2</i>

Knockout strain construction

Plasmids were transferred to *SoF1* with conjugation or electroporation. Antibiotic concentrations used for selection of modified *SoF1* strains were 20 µg/mL of KAN and 10 µg/mL of TET.

For conjugation, liquid culture (LC) of *SoF1* was grown in autotrophic conditions as described above for two to three days to reach an OD of 0.7-1. Overnight (O/N) LC of *E. coli* S17-1, with and without plasmid, and JM109(DE3) were grown at 37°C with shaking 220 rpm. New LCs were inoculated from O/N cultures next day and grown to exponential phase (OD 0.3-0.6). *E. coli* cells were centrifuged 5900 rpm 30 s, washed, and resuspended to 1 volume of 0.9% NaCl. *E. coli* and *SoF1* cells were mixed with OD ratio 1:15 and DSM media was added up to 1 mL. S17-1 without plasmid mixed with *SoF1* was used as a negative control and JM109(DE3) with S17-1 containing plasmid was used as a positive control. Mixtures were vacuum filtered through a 0.22 µm GV Durapore® membrane filter (MilliporeSigma, US). Filters were placed on prewarmed TSA plates, cells facing away from the agar, and incubated O/N in autotrophic conditions. Next day, filters were washed with 1 mL of 0.9% NaCl, vortexed and centrifuged 4000 rpm for 1 min and, after removing the filters, cells were resuspended. Cells were plated with serial dilutions to selective plates and grown in appropriate conditions. Conjugated JM109(DE3) control was plated on TSA plates containing 40 µg/mL of nalidixic acid for selection and incubated at 37°C O/N. Conjugated *SoF1* cultures were plated on DSM plates containing KAN and incubated in autotrophic conditions approximately for one week. After colonies appeared on *SoF1* plates, they were reapplied to fresh DSM KAN plates and incubated one more week in autotrophic conditions. Single colonies from these plates were tapped on TSA plates and grown in 30°C in heterotrophic conditions. If growth on TSA plates resulted in *E. coli* growth, colonies were reapplied again to selective DSM plates.

For electroporation, LC of *SoF1* was cultivated in autotrophic conditions for two to three days to reach an OD of 0.7-1.5. Cells were transformed to Falcon tubes and chilled on ice for 15-30 min. Cells were centrifuged at 4°C 4000 rpm for 5-10 min, supernatant was discarded, and the pellet was resuspended to 1 volume of ice-cold double distilled H₂O. Centrifugation was repeated and supernatant was discarded. Washing was repeated with 1 volume of ice-cold 10% glycerol. Cells were resuspended to ice-cold 10% glycerol to reach concentration of around $2 \cdot 10^{10}$ cells/mL. Cells were used immediately for electrotransformation. 40 µL of cells were mixed with 1 µL of plasmid and incubated for 10 min on ice. Cells were transformed to an electroporation cuvette on ice. The cuvette was subjected to a single electric pulse of 2.5 kV with 25 µF capacitance and 400 or 600 Ω resistance. 1 mL of pre-warmed (30°C) SOBE was added immediately and solution was transferred to a Falcon tube and incubated O/N at autotrophic conditions. After incubation, cells were plated on TSA selection plates with multiple dilutions and incubated in heterotrophic conditions.

After transformation with conjugation or electroporation, successful transformants were screened with colony PCR with addition of final concentration of 3% dimethyl sulfoxide to the standard PCR mixture. Constructed *SoF1* strains are summarized in Table 8.

Table 8. Constructed *SoF1* variant strains.

Strain	Genotype
<i>SoF1-2.0</i>	<i>phaC1::kan</i>
<i>SoF1-3.0</i>	<i>phaC2::tet</i>
<i>SoF1-4.0</i>	<i>phaC1::kan</i> and <i>phaC2::tet</i>

5

PHB content analysis

PHB contents of the wild type (WT) strain *SoF1* and the knockout strain *SoF1-2.0* were analysed with gas chromatography-mass spectrometry (GC-MS). Both strains were grown in autotrophic conditions as described above in 5 mL cultivation volume in varying nitrogen concentrations (18.0, 13.5, 9.0 and 4.5 mM of nitrogen) and in heterotrophic conditions. Cultivations were inoculated from cultures grown in autotrophic conditions to gain a starting OD of 0.1. Analysis was done at earliest after one week from inoculation. 1-3 mL of each sample was centrifuged down, and the pellet was stored at -20°C. Pellets were thawed, washed twice with double distilled H₂O and lyophilized for 24-48 h. 10 mg of each sample was subjected to methanolysis by heating at 100°C for 140 min in a solution containing 1 mL chloroform, 150 µL sulfuric acid, 20 µL internal standard (3-hydroxybutyric acid), and 830 µL methanol. 3-hydroxybutyric acid was treated similarly as a reference sample. After samples were cooled to room temperature, water-soluble particles were removed with 0.5 mL of water. Gas chromatography system (7890, Agilent) and HP-FFAP column (19091F-102, Agilent) were used to analyse the chloroform phase.

Results

According to GC-MS, almost no PHB was produced in the *SoF1-2.0* strain whereas in the WT *SoF1* the PHB dry content was 15-30 % when grown in autotrophic conditions (Table 9). Almost no PHB was produced in *SoF1* or *SoF1-2.0* when grown in heterotrophic conditions

Table 9. PHB contents of WT *SoF1* and *SoF1-2.0* grown in autotrophic conditions in different nitrogen concentrations in autotrophic conditions and in heterotrophic conditions (SOBE) determined via GC-MS.

Nitrogen concentration (mM)	<i>SoF1</i> (mg of PHB/g of dry cell mass)	<i>SoF1-2.0</i> (mg of PHB/g of dry cell mass)
4.5	312.9	0.3
9.0	230.0	1.3
13.5	251.9	1.3
18.0	204.0	1.1
Not known, grown in SOBE	0.1	0.2

Autotrophic growth curves of *SoF1* and *SoF1-2.0* are presented in Figure 3. *SoF1-2.0* has a slightly lower growth rate compared to *SoF1*.

phaC1 DNA sequence (SEQ ID NO:60):

```

5  ATGTCCGCAGCCGAGGAAACGTCCACGCACGCTGAATTGAGGCTCCCGCAGGATGGGGTGG
   AGCATGACGTGCGGGCCGCGAGGCCGCGGTGGACCGGTCCGCCGGGGAGCCGTGCGGGAA
   CGCAGGCGTCCGCCGCGCCCGCGAGGCGGCGCGGTCTTCCGCTCCTGTTTCCGCTGCTGCC
   GGAGAAACGCAGCCACAGGACGATACCCCGCCGCAGGACCCGTCTTCCCTGTGCGACGGTCC
   AGGCCTGTGCGCGCCCGTGACCCAGCCCGGCGCGGCGTCCGGCGACTTCGGCGGGCCCCGAG
10  GCCATGGGCGACGCCTTCATGCCGCCTGTGCCGAAGAGCCCATGGAGGGGATGGCCCCCG
   CGCCGGCAATCTCCGCGGCGCCCGCCTCGGTGCCGTCTGCCGGCTGGCCGGAGGATGCGCC
   TTCGGCACTGCTCGATGCGGTGGATGCGTCCGGGGACCTGCCGGCTGCCGCGGAGGGCGCG
   GCCACGGCGCCGAGCCCATTTTCCGCGAGCTGCCCATGACGGCGGCGCCCGCCGCACCTG
   CCATCGCGAACATTCTCGAGCCGGTGGCCGAGGCTCTGTCCGCGCTCGGGGCGGTGGCCGT
15  CTCCCGGCCCCAGGTCCAGCGCAATTCCGCGCGGCGCCCGAACACCCCCCATGCCAGG
   ATGGCGCCCCCTCAGGCGGCGCCAGAACCGGCCCTGCCGTTCCGCCAAACCTGAAGCTGC
   CAAACCTGAAGCTGCCAAACCTGAAGCCGCCAAGCCTGAAGCTGCCAAGCCTGAAGTTGCCA
   AGCCGGACGCCGCGCGCCGGACGCCGCGAAATCTTCGGGCAAGGCGGAGCGTCCGTCCGG
   CGCCGGGGACGGCTCTGGAACGTGCGGCGTGAACATGGAGGCCTTCTCCCGCAACCTCGCC
20  CGTCTGGTGGAAGAGGGCGGCAAGGCCATGGCCGCCTATCTGAGCCCACGGGAGCAGGGCA
   AGACCGACGATCTCGCCGATGACATCGCCGATGCCATGAAGACCGTGGGCCAGGTGGTGGA
   ATACTGGGTGCGCCGATCCCCAGCGCACGGTGGAGGCGCAGTCCCGCCTCATGGGCGGCTAT
   CTCTCGGTCTGGGCCAACACCCTGAAGCGCCTGGCAGGCGAGGAGGCCACGCCGGTGGCCG
   CCCCCGACCCGAAGGATGCCCGCTTCAAGGATGCCGGCTGGAACGACAGCCCCATGTTTCGAT
25  GCCCTCAAGCAGGCCTATCTGGTCACTCGGACTGGGCGCAGAACATGGTGGACGAGGCCAA
   GGGCCTCGATCCCCACACCAAGCACAAGGCGGAGTTCCTGGTGCGGCAGATCGCCAACGCCA
   TCTCGCCCTCCAATTCTCCTCACCAATCCCGAGCTCATCCGCGAGACCCTGCACTCGTCGG
   GCGAGAATCTGGTGAAGGGCATGCAGAACCTACCGCGGATCTGATGGCGGGGACAGGGCAC
   GCTGAAGATCCGCCAGACGGATCTTCCGCCTTCGAGGTGGGCCGCAACCTCGCCACCACGC

```

CGGGCAAGGTGATCTTCGAAAACGAGCTGATGCAGCTCATCCAGTACGAGCCGACCACCGAG
ACGGTGAAGAAGACGCCGGTGCTCATCGTCCCGCCCTGGATCAACAAGTTCTACATCCTCGA
CCTGACGGCGGAAAAATCCTTGATCAAGTGGCTGGTGAGCCAGGGAAGTACGGTCTTCACCA
TCTCCTGGGTCAATCCCGACGGGCGCCTCGCCGCCAAGGGCTTCGACGATTACATGCGCGAC
5 GGCATCATGGCCGCCCTCGACGCGGTGGCGGTGGCCTCCGGCGAGCGGCGGGCCCATGCG
GTGGGCTATTGCGTGGGCGGCACGCTGCTGGCCACCACGCTCGCTTACATGGCCGCCACCG
GGGATGACCGCATCGCCAGCGCCACCTTCCTCACCACCCAGATCGACTTCACCCATGCGGGC
GATCTGAAGGTGTTGCTGGACGAAAGCCAGCTCGCCACCATCGAGCGCAAGATGAAGGAGAT
GGGGTATCTGGAAGGCTCGAAGATGGCCTCCGCCCTCAACATGCTGCGCTCCAACGACCTGA
10 TCTGGCCCTATGTGGTGAACAACATACATGAAGGGCAAGGCGCCGTTTCCGTTTCGACCTGCTGT
TCTGGAATTCGATTCCACCCGCATGCCGGCGGCGAACCATTCTATTATCTGCGCAACTGCT
ACCTACCAACAATATCGCCCGGGGTCTGGCGGAGCTGGCCGGCCTCAAGATCGATGTCACC
AAGGTGTCGATCCAGTCTATTGCTGGCGACGCGGGAAGACCACATCGCCCCGGCCAACTC
GGTCTATATCGGGGCAAATCTCCTGTCCGGCCCCGGTGCGCTACGTGCTGGCGGGCTCGGGCC
15 ACATCGCCGGGGTGGTGAACCCGCGGCCAAGATGAAGTATCAGTACTGGGCCGACGGCCCC
GGTGGGGCCGAGCTACGAGGCCTGGCTGGCCGGGGCGCAGGAGCACAAGGGCTCCTGGTG
GCCGGACTGGTTCAACTGGTTCTCCTTCAACCATCCCGAGGAGGTGCCGGCGCGCGCCATCG
GCGGCGGCCCGCTCGCCCCCATCGAGGACGCACCCGGGCGCTATGTGAAGGAGCGGTGCTA
G

20 *phaC2* DNA sequence (SEQ ID NO:61)

ATGGAAGCGCGAAAGATGCCCCGTTTACCCCCCTCCTCCGCCACCATCCTGCCCCCTGCCCGTC
TCGGCCGCTCCCCCTTCGACGGCGCCCGCGCGTCCCTGCCGGCGACGGCCTCCAGTCCAC
CAACAAGGCCAGCCCCCTCAGCCTTCCCCGCCGCTGGGCGCGTTCTTCGCCCTGCCCGGCC
TGCCCGCCTTCTGTGCCCGGACGAGGAGGATTTGAGGAGGGTTGCGGGCCGGCCGCCTT
25 CCATGCGGTGGACCGGGCGGCGGCGGCCTTGGTGGCCCGCACCACCCAGGGGCTCTCGCCC
GCCGCCCTACCCCTCGTTACATGGATTGGGCGATGCATCTGGCAGCGGCGCCGGGCAAGCA
GGCGGAGCTGGCGGTGAAGGCCACGCGCAAGGCGGCGCGCTTCTGGGCCTATGTGCTCGCC
TCCACCCTCGACCGCACCCAGGCGCCCTGCATCGCGCCCTGGTGGGCGACGAACGGTTTTC
CGCCCCGGCCTGGCAGGACTGGCCCTACCGCTTCTGGTACCAGGCCTTCTGCTCAATCAGC
30 AATGGTGGCACAACGCCACTCATGGCGTGCCGGGCGTGGCGCCGCACAATCAGGACGTGGT
GGCCTTCGCCGCCCGGCAGGTCTGGACATGTTCTCCCCCTCCAACCTCGCCCCCTACCAATCC
CGAGGTGGTGAAGAAGGCGCGCCAGACGCTAGGGGGCAATTTCTGTCAGGGCGCGCGCAAC
TTCATGGAGGACCAGTCGCGCAAAACCACCGGCCGCCCGCCCGTGGGCGCGGAGGCGTTCA
CTCCCGGCAAGGAGGTGGCCATACCCCCGGCGAGGTGATCTACCGCAACCATCTCATCGAG
35 CTGATCCAATACCGGGCGACCACCCCTGACGTGCATGCGGAGCCGATCCTCATCGTCCCCG
CTGGATCATGAAATATTACATCCTCGACCTGTGCCCCGATAACTCCCTGATCCGCTACCTGGT

GGACAAGGGGCACACGGTCTTCTGCATCTCCTGGCGCAATGTGAACGCGGAGGACCGCGATC
TCGGCTTCGAGGACTATCGCAAGATGGGCATCATGGCGGCCCTCGACGCGGTGAACGCGGT
GGTGCCGAACCAAGGTGCATGCGGTGGGTTATTGCCTGGGCGGCACGCTGCTCTCCATCG
CCGCCCGCCCATGGCGCGGGTGGTCGACGACCGGCTGGGCTCCGTACCCTGTTGCGCCGC
5 CCAGACCGACTTCACCGAGCCCGGCGAACTGCAGCTCTTCGTGGACCCGAGCGAGCTTTATG
CCCTGGAAAGCCTCATGTGGGACCAGGGCTATCTCGGCGCCCGGCAGATGGCCGCGCGCTT
CGAGATGCTGCGCTCCAACGACCTCGTCTGGTCGCGCATGGTGCGCGACTATCTCATGGGCG
AGCGCGCACCCATGAACGACCTCATGGCCTGGAACGCAGATGCCACCCGCATGCCCTATCGC
ATGCATTCCCAGTATCTGCGCAACCTGTTCTCGACAACGAGCTGGCCGTGGGCCGTTACATG
10 GTGGAGGGCCGGCCGGTGTGCTGCGAGAACATCCGCGTCCCGCTGTTCTGTGGTGCGCACGG
AGCGGGACCACGTGGCGCCGTGGAAGTCGGTCTACAAGATCCACCAGCTCACCGACACGGA
CGTGACCTTCGTCTCGCCTCCGGCGGGCACAATGCGGGCATCGTCTCCGAGCCCGGCCACA
AGCACCGGCACTATCGCATCCACGACACCAAATTGGGCGAGATGCATGTGAGCCCGGAGGAA
TGGATGGAGGCGAACCGGTGCGCAGGACGGGTCTGGTGGCCGGCCTGGGAGGCATGGCTC
15 GCCGGCCAGTCTCCGGCCGCATCGGCCTGCCGCTCTGGGCGCGCCCGGCTACGAGGTGC
TGGGCCCCGGCGCCCGGCACCTATGTGATGCAAAGGTGA

phaC1 amino acid sequence (SEQ ID NO:62):

MSAAEETSTHAELRLPQDGVHDVAAAEAAVDRSAGEPSGTQASAAPAEAPSSAPVSAAAGET
QPQDDTPPDSSLSDGPLSPVTPQGAASGDFGGPEAMGD AFMPPVPEEPMEGMAPAPAISA
20 APASVPSAGWPEDAPSALLDAVDASGDLPA AEGAATAPEPIFRELPMTAAPAAPAIANILEPVAE
ALSALGAVAVSRPQVQREFRAAPEHPPMPRMAPPQA APEPAPAVPPKPEAAKPEAAKPEAAKPEA
AKPEVAKPDAAAPDAKSSGKAERPSGAGDGS GTSGVNMEAFSRNLARLVEEGGKMAAYLSP
REQKTDLDLADDIADAMKTVGQVVEYVWADPQRTVEAQSRLMGGYLSVWANTLKR LAGEEATP
VAAPDPKDARFKDAGWNDSPMFDALKQAYLVTS DWAQNMVDEAKGLDPHTKHKA EFLVRQIAN
25 AISPSNFVLTNPELIRETLHSSGENLVKGMQNL TADLMAGQGTLKIRQTDLSAFEVGRNLATTPGK
VIFENELMQLIQYEPTTETVKKTPVLIVPPWINKFY ILDLTAEKSLIKWLVSQGLTVFTISWVNP DG
RLAAKGFDYMRDGIMAALDAVAVASGERRAHAVGY CVGGTLLATTLAYMAATGDDRIASATFL
TTQIDFTHAGDLKFVDESQ LATERKMKEMGYLE GSKMASAFNMLRSNDLIWPYVNNYMKGK
APFPFDLLFWNSDSTRMPAANH SYLRNCYL TNNIARGLAELAGLKIDVT KVSIPVYSLATREDHI
30 APANSVYIGANLLSGPVRYVLAGSGHIAGVVNPPAKMKYQYWADGPVGPSYEAWLAGAQEHKG
SWWPDWFNWFNFHPEEVPARAIGGGRLAPIEDAPGRYVKERS

phaC2 amino acid sequence (SEQ ID NO:63):

MEARKMPVSPSSATILPLVSAAPPSTAPAASLPATASSSTNKASPSAFPAAWARSFALPGLPAFL
CPDEEDFEEGSGPAAFHAVDRAAAALVARTTQGLSPAALTAYMDWAMHLAAAPGKQAE LAVKA
35 TRKAARFWAYVLASTLDRTQAPCIAPLVGDERFSAPAWQDWPYRWFYQAFLLNQWVHNATH

GVPGVAPHNQDVVAFARQVLDMFSPSNSPLTNPEVVKKARQTLGANFVQGARNFMEDQSRKT
TGRPPVGAEAFTPGKEVAITPGEVIYRNHLIELIQYRATTPDVHAEPIVIPAWIMKYYILDLSPDNS
LIRYLVDKGHTVFCISWRNVNAEDRDLGFEDYRKMGIMAALDAVNAVVPNQKVHAVGYCLGGTL
LSIAAAAMARVVDDRLGSVTLFAAQTDFTPEGELQLFVDPSELYALESMLWDQGGLGARQMAGA
5 FEMLRSNDLVWSRMVRDYL MGERAPMNDLMAWNADATRM PYRMHSQYLRNLFLDNELAVGRY
MVEGRPVSLQNI RVPLFVVGTERD HVAPWKS VYKIHQLTDTDTFVLASGGHNAGIVSEPGHKH
RHRYIHDTKL GEMHVSPEEWMEANRSQDGSWWPAWEAWLAGQSSGRIGLPPLGAPGYEVLGP
APGTYVMQR

CLAIMS

1. A variant of bacterial strain VTT-E-193585 comprising a genetic modification that reduces the bacterial production of polyhydroxyalkanoic acid (PHA) as compared to strain VTT-E-193585.
5
2. The variant according to claim 1, wherein the genetic modification reduces bacterial PHA synthase activity as compared to strain VTT-E-193585, preferably wherein PHA synthase activity has been reduced to less than 10%, such as less than 5%, for example less than 2%.
10
3. The variant according to claim 1 or 2, wherein the variant comprises a genetic modification reducing the expression level of *phaC1* and/or the activity of the phaC1 enzyme.
15
4. The variant according to any one of the preceding claims, wherein the variant comprises a genetic modification reducing the expression level of *phaC2* and/or the activity of the phaC2 enzyme.
- 20 5. The variant according to any one of the preceding claims, wherein the genetic modification is a gene disruption.
6. The variant according to any one of the preceding claims, wherein the variant comprises gene disruptions of both *phaC1* and *phaC2*.
25
7. The variant according to any one of the preceding claims, wherein the variant is the bacterial strain deposited under number VTT-E-213595.
8. The variant according to any one of the preceding claims, wherein the variant has retained the ability to grow using hydrogen gas as energy source and carbon dioxide as the only carbon source.
30
9. A variant *Xanthobacter* strain comprising a gene disruption of one or more genes encoding a PHA synthase.
35
10. A culture comprising the variant bacteria according to any one of the preceding claims.

11. A process for the production of biomass, said process comprising culturing the variant bacteria of any one of claims 1 to 9.
- 5 12. The process according to claim 11, comprising culturing the variant bacteria in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide.
- 10 13. A process for the production of biomass, said process comprising culturing a variant chemoautotrophic bacterial strain in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide and wherein said variant chemoautotrophic strain comprises a gene disruption of one or more genes encoding a PHA synthase.
- 15 14. The process according to claim 12 or 13, wherein dissolved oxygen in the culture is maintained between 5% and 10%.
- 15 15. The process according to any one of claims 12 to 14, wherein ammonium, urea, nitrate and/or nitrogen gas is used as nitrogen source.
- 20 16. The process according to any one of claims 12 to 15, wherein pH in the culture is maintained between 5.5 and 8.0, e.g. between 6.5 and 7.0, such as at 6.8.
17. The process according to any one of claims 12 to 16, wherein said culture is grown at a temperature between 25°C and 40°C, e.g. between 28°C and 32°C, such as at 30°C.
- 25 18. The process according to any one of claims 12 to 17, comprising the further step of harvesting biomass produced during the culture, optionally comprising a further step of drying the biomass.
- 30 19. A process for the production of protein, comprising performing the process according to claim 18 and a further step of isolating protein from said biomass, wherein the process results in a protein fraction and a fraction comprising non-protein components.

20. The process according to claim 18 or 19, comprising the further step of producing a food or feed product from said biomass, from said protein fraction or from said fraction comprising non-protein components.
- 5 21. A product, such as biomass, protein, or non-protein components obtained or obtainable by the process according to any one of claims 11 to 20.
22. A food or feed product obtained or obtainable by the process according to claim 20.
- 10 23. A method for genetic modification of bacterial strain VTT-E-193585 comprising the steps of:
- a) providing bacteria of strain VTT-E-193585 or a genetically-modified or mutated strain generated using bacterial strain VTT-E-193585, such as VTT-E-213595,
 - b) introducing a nucleic acid construct into said bacteria, wherein said nucleic acid
 - 15 construct comprises:
 - i) sequences encoding a selectable marker,
 - ii) optionally further sequences to be integrated into the bacterial genome,
 - iii) flanking sequences allowing homologous recombination with the bacterial genome, and,
 - 20 c) selecting a genetically-modified strain on the basis of the selectable marker.
24. The method according to claim 23, wherein the selectable marker is a gene providing antibiotic resistance, such as kanamycin or tetracycline resistance.
- 25 25. A variant of bacterial strain VTT-E-193585 comprising a genetic modification wherein said genetic modification comprises the disruption of a bacterial gene with a selectable marker providing antibiotic resistance, such as kanamycin or tetracycline resistance.
- 30 26. A culture comprising the variant according to claim 25.
27. A process for the production of biomass, said process comprising culturing the variant bacteria of claim 25.
- 35 28. The process according to claim 27, comprising culturing the variant bacteria in continuous culture with hydrogen as energy source and an inorganic carbon source, wherein the inorganic carbon source comprises carbon dioxide.

29. The process according to claim 28, further comprising the features specified in any one of claims 14 to 18.
- 5 30. A process for the production of protein, comprising performing the process according to claim 27, 28 or 29 and a further step of isolating protein from said biomass, wherein the process results in a protein fraction and a fraction comprising non-protein components.
- 10 31. The process according to claim 30, comprising the further step of producing a food or feed product from said biomass, from said protein fraction or from said fraction comprising non-protein components.
- 15 32. A product, such as biomass, protein, or non-protein components obtained or obtainable by the process according to any one of claims 27 to 31.
33. A food or feed product obtained or obtainable by the process according to claim 31.