

METHODS AND COMPOSITIONS FOR PRODUCING ETHYLENE FROM RECOMBINANT MICROORGANISMS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No.
5 62/942,895, filed December 03, 2019 which is incorporated herein by reference.

TECHNICAL FIELD

[0002] The present disclosure relates to recombinant microorganisms having
an improved ethylene producing ability, methods of producing the same, and methods of
10 harvesting ethylene from such recombinant organisms. A benefit of the recombinant
microorganisms and the methods disclosed herein can include increased production of
ethylene from microbial cultures. An additional benefit can be the use of carbon dioxide to
produce bio-ethylene useful as a feedstock for the production of plastics, textiles, and
chemical materials, and for use in other applications. Another benefit of the methods and
15 systems disclosed herein can include reduction of excess carbon dioxide from the
environment.

BACKGROUND

[0003] The increased demand for power worldwide has led to an excess of
carbon dioxide from burning fossil fuels such as oil and gas, contributing substantially to
20 what many are calling a global warming crisis. Industry is so desperate to prevent carbon
dioxide from entering the atmosphere that they have resorted to sequestering carbon dioxide
from exhaust streams and the atmosphere. They then store the carbon dioxide in subterranean
environments. However, all current known methods just remove carbon dioxide from the
atmosphere by storing it under ground. They do not actually convert the carbon dioxide back
25 into any other useful material.

[0004] The limited supply of petroleum and its harmful effects on the
environment have prompted developments in renewable sources of fuels and chemicals.
Ethylene is the most widely produced organic compound in the world, useful in a broad
spectrum of industries including plastics, solvents, and textiles. Ethylene is currently
30 produced by steam cracking fossil fuels or dehydrogenating ethane. With millions of metric
tons of ethylene being produced each year, however, more than enough carbon dioxide is
produced by such processes to greatly contribute to the global carbon footprint. Producing

ethylene through renewable methods would accordingly help to meet the huge demand from the energy and chemical industries, while also helping to protect the environment.

[0005] Since ethylene is a potentially renewable feedstock, there has been a great deal of interest in developing technologies to produce ethylene from renewable sources, such as carbon dioxide and biomass. Bio-ethylene is currently produced using ethanol derived from corn or sugar cane. A variety of microbes, including bacteria and fungi, naturally produce ethylene in small amounts. Heterologous expression of an ethylene producing enzyme has been demonstrated in several microbial species, where the hosts have been able to utilize a variety of carbon sources, including lignocellulose and carbon dioxide.

10 [0006] Based on modern history, it is fair to say that excess carbon dioxide in the atmosphere will not be reduced until it becomes profitable to reduce it. There remains a need for improvements in microbial bio-ethylene systems and processes, in order to produce ethylene at a commercial scale. There remains a need to produce hydrocarbons through more efficient renewable technologies. There remains a need to remove excess carbon dioxide
15 from the atmosphere. There remains a need for improved methods to produce ethylene from a renewable feedstock for industrial and commercial applications.

SUMMARY

[0007] Embodiments herein are directed to a recombinant microorganism
20 having an improved ethylene producing ability, wherein the recombinant microorganism expresses at least one ethylene forming enzyme (EFE) protein having an amino acid sequence at least 95% identical to SEQ ID NO: 1 (see attached Appendix) by expressing a non-native EFE expressing nucleotide sequence, wherein an amount of EFE protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism
25 lacking the non-native EFE expressing nucleotide sequence. In an embodiment, the recombinant microorganism also expresses at least one alpha-ketoglutarate permease (AKGP) protein having an amino acid sequence at least 95% identical to SEQ ID NO: 2 (see attached Appendix) by expressing a non-native AKGP expressing nucleotide sequence, wherein an amount of AKGP protein produced by the recombinant microorganism is greater
30 than that produced relative to a control microorganism lacking the non-native AKGP expressing nucleotide sequence. In an embodiment, the amount of EFE protein produced by the recombinant microorganism is from about 5% to about 200% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence.

[0008] In various embodiments, the recombinant microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*,
5 algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell.

[0009] In an embodiment, the non-native EFE expressing nucleotide sequence is inserted into a bacterial vector plasmid, a high copy number bacterial vector plasmid, a bacterial vector plasmid having an inducible promoter, a nucleotide guide of a homologous
10 recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3 (see attached Appendix), and the non-native EFE expressing nucleotide sequence is inserted into a vector plasmid of a *Chlamydomonas* sp. bacterium.

[0010] In an embodiment, a non-native EFE expressing nucleotide sequence and a non-native AKGP expressing nucleotide sequence are inserted into a bacterial vector plasmid, a high copy number bacterial vector plasmid, a bacterial vector plasmid having an inducible promoter, a nucleotide guide of a homologous recombination system, a CRISPR
15 CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4 (see attached Appendix), and the non-native EFE expressing nucleotide sequence and the AKGP expressing nucleotide sequence are inserted into a vector plasmid of an *Escherichia* sp. bacterium.

[0011] In an embodiment, the recombinant microorganism further includes a
25 non-native AKGP expressing nucleotide sequence, wherein the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 5 (see attached Appendix), and the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus* sp.
30 bacterium. In another embodiment, the recombinant microorganism further includes a non-native AKGP expressing nucleotide sequence, wherein the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6 (see attached Appendix), and the non-native EFE expressing nucleotide sequence and non-native AKGP

expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus* sp. bacterium.

[0012] In certain embodiments, the recombinant microorganism expresses at least one phosphoenolpyruvate synthase (PEP) protein having an amino acid sequence at least 95% identical to SEQ ID NO. 15 by expressing a non-native PEP expressing nucleotide sequence, wherein an amount of PEP protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native PEP expressing nucleotide sequence, and wherein an amount of AKG produced by the recombinant microorganism is greater than that produced relative to the control microorganism. In certain such embodiments, the recombinant microorganism includes a microorganism selected from the group consisting of a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*, algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, and a plant cell.

[0013] In certain embodiments, the recombinant microorganism expresses at least one phosphoenolpyruvate synthase (PEP) protein having an amino acid sequence at least 95% identical to SEQ ID NO. 15 by expressing a non-native PEP expressing nucleotide sequence, wherein an amount of PEP protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native PEP expressing nucleotide sequence, and wherein an amount of AKG produced by the recombinant microorganism is greater than that produced relative to the control microorganism.

[0014] In certain embodiments, the recombinant microorganism expresses at least one citrate synthase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 17 by expressing a non-native citrate synthase expressing nucleotide sequence, wherein an amount of citrate synthase protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native citrate synthase expressing nucleotide sequence.

[0015] In certain embodiments, the recombinant microorganism expresses at least one isocitrate dehydrogenase (IDH) protein having an amino acid sequence at least 95% identical to SEQ ID NO. 20 by expressing a non-native IDH expressing nucleotide sequence,

wherein an amount of IDH protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native IDH expressing nucleotide sequence, and wherein an amount of AKG produced by the recombinant microorganism is greater than that produced relative to the control microorganism.

5 **[0016]** In certain embodiments, the recombinant microorganism contains a deletion in a glucose-1-phosphate adenylyltransferase expressing nucleotide sequence, wherein an amount of glucose-1-phosphate adenylyltransferase protein produced by the recombinant microorganism is less than that produced relative to a control microorganism lacking the deletion.

10 **[0017]** In certain embodiments, the recombinant microorganism expresses at least one sucrose permease protein having an amino acid sequence at least 95% identical to SEQ ID NO. 24 by expressing a non-native sucrose permease expressing nucleotide sequence,
wherein an amount of sucrose permease protein produced by the recombinant microorganism
15 is greater than that produced relative to a control microorganism lacking the non-native sucrose permease expressing nucleotide sequence.

[0018] In certain embodiments, the recombinant microorganism expresses at least one sucrose permease protein having an amino acid sequence at least 95% identical to SEQ ID NO. 24 by expressing a non-native sucrose permease expressing nucleotide
20 sequence,
wherein an amount of sucrose permease protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native sucrose permease expressing nucleotide sequence.

[0019] In certain embodiments, the recombinant microorganism expresses at least one sucrose permease protein having an amino acid sequence at least 95% identical to SEQ ID NO. 24 by expressing a non-native sucrose permease expressing nucleotide
25 sequence,
wherein an amount of sucrose permease protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native
30 sucrose permease expressing nucleotide sequence.

[0020] In certain embodiments, the recombinant microorganism expresses at least one protein selected from the group consisting of a sucrose phosphate synthase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 26, a sucrose-6-phosphatase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 28,

a glycogen phosphorylase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 30, and a UTP-glucose-1-phosphate uridylyltransferase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 32, by expressing a non-native nucleotide sequence encoding the at least one protein, wherein an amount of the at least one protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native nucleotide sequence encoding the at least one protein, wherein an amount of sucrose produced by the recombinant microorganism is greater than that produced relative to the control microorganism.

[0021] In certain embodiments, the recombinant microorganism contains at least one deletion in at least one nucleotide sequence, wherein the at least one nucleotide sequence encodes at least one protein selected from an invertase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 34, a glucosylglycerol-phosphate synthase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 36, and a glycogen synthase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 38, wherein an amount of the at least one protein produced by the recombinant microorganism is less than that produced relative to a control microorganism lacking the at least one deletion.

[0022] Embodiments herein are directed to methods of producing a recombinant microorganism having an improved ethylene producing ability. In an embodiment, the method includes producing the recombinant microorganism by inserting a non-native EFE expressing nucleotide sequence or a combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence into a bacterial plasmid of a microorganism, wherein the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3 or SEQ ID NO. 4. In another embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence expresses an amino acid sequence at least 95% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In another embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 7 (See Appendix).

[0023] In various embodied methods, the microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*,

algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell.

5 **[0024]** In an embodiment of methods herein, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3 and the microorganism is a *Chlamydomonas* sp. bacterium. In another embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4 and the microorganism is an *Escherichia* sp. bacterium. In another embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence expresses an amino acid sequence at least 10 95% identical to SEQ ID NO. 5 or SEQ ID NO. 6, and the microorganism is a *Synechococcus* sp. bacterium. In another embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 7 and the microorganism is *Synechococcus* sp. bacterium.

15 **[0025]** Methods of producing ethylene are embodied herein. An embodiment of such a method includes providing a recombinant microorganism having an improved ethylene producing ability, wherein the recombinant microorganism expresses at least one ethylene forming enzyme (EFE) protein having an amino acid sequence at least 95% identical to SEQ ID NO: 1 by expressing a non-native EFE expressing nucleotide sequence, 20 wherein an amount of EFE protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native EFE expressing nucleotide sequence; culturing the recombinant microorganism in a bioreactor culture vessel under conditions sufficient to produce ethylene in the bioreactor culture vessel; and harvesting ethylene from the bioreactor culture vessel.

25 **[0026]** In an embodiment of methods of producing ethylene herein, the recombinant microorganism contains a non-native EFE expressing nucleotide sequence or a combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence inserted into a bacterial plasmid of the microorganism, wherein the non-native EFE expressing nucleotide sequence has a nucleotide sequence at 30 least 95% identical to SEQ ID NO. 3 or SEQ ID NO. 4. In another embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 7. In another embodiment, the combined non-native EFE expressing nucleotide

sequence and non-native AKGP expressing nucleotide sequence expresses an amino acid sequence at least 95% identical to SEQ ID NO. 5 or SEQ ID NO. 6.

[0027] In various embodiments of producing ethylene herein, the recombinant microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*,
5 *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*, algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell.

[0028] An embodiment of a method of producing ethylene further includes
10 increasing an amount of ethylene production by adding at least one activator to a culture containing the recombinant microorganism located within the bioreactor culture vessel. In an embodiment, such a method includes adding CO₂ to a culture atmosphere contained within the bioreactor culture vessel at rate of between about 100 ml/minute and about 500 ml/minute. In an embodiment, such a method further includes decreasing an amount of
15 ethylene production by removing at least one molecular switch from the cell culture containing the recombinant microorganism located within the bioreactor culture vessel. In an embodiment, such a method further includes controlling the amount of ethylene produced from the microbial culture by increasing or decreasing the concentration of at least one nutrient or the amount of at least one stimulus when culturing the recombinant
20 microorganism. In an embodiment, the concentration of at least one nutrient and the amount of at least one stimulus are at a ratio of from about 0.5-1.5 gr./liter to about 0.1 mM in the microbial culture. In an embodiment, such a method further includes removing the amount of ethylene produced from the microbial culture by condensing the ethylene from a gaseous to a liquid state, or wherein the amount of ethylene recovered is from about 0.5 ml
25 to about 10 ml/liter/h.

[0029] Embodiments herein are directed to a recombinant microorganism having an improved alpha-ketoglutarate (AKG) producing ability. In certain embodiments, the recombinant microorganism expresses at least one phosphoenolpyruvate synthase (PEP) protein having an amino acid sequence at least 95% identical to SEQ ID NO. 15 by
30 expressing a non-native PEP expressing nucleotide sequence, and wherein an amount of PEP protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native PEP expressing nucleotide sequence.

[0030] In certain embodiments, the recombinant microorganism expresses at least one isocitrate dehydrogenase (IDH) protein having an amino acid sequence at least 95%

identical to SEQ ID NO. 20 by expressing a non-native IDH expressing nucleotide sequence, and wherein an amount of IDH protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native IDH expressing nucleotide sequence; wherein an amount of AKG produced by the recombinant microorganism is greater than that produced relative to the control microorganism.

[0031] In certain embodiments, the recombinant microorganism expresses at least one citrate synthase protein having an amino acid sequence at least 95% identical to SEQ ID NO. 17 by expressing a non-native citrate synthase expressing nucleotide sequence, wherein an amount of citrate synthase protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native citrate synthase expressing nucleotide sequence.

[0032] In certain embodiments, the recombinant microorganism contains a deletion in a glucose-1-phosphate adenylyltransferase expressing nucleotide sequence, wherein an amount of glucose-1-phosphate adenylyltransferase protein produced by the recombinant microorganism is less than that produced relative to a control microorganism lacking the deletion.

[0033] In certain embodiments, the recombinant microorganism includes a microorganism selected from the group consisting of a Cyanobacteria, a Synechococcus, Synechococcus elongatus, Synechococcus leopoliensis, Synechocystis, Anabaena, a Pseudomonas, Pseudomonas syringae, Pseudomonas savastanoi, Chlamydomonas, Chlamydomonas reinhardtii, Escherichia, Escherichia coli, Geobacteria, algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, and a plant cell.

BRIEF DESCRIPTION OF THE DRAWINGS

[0034] The foregoing summary, as well as the following detailed description of the embodiments, will be better understood when read in conjunction with the attached drawings. For the purpose of illustration, there are shown in the drawings some embodiments, which may be preferable. It should be understood that the embodiments depicted are not limited to the precise details shown. Unless otherwise noted, the drawings are not to scale.

[0035] Figure 1 is a flow chart depicting an embodiment of a method of producing ethylene herein.

[0036] Figure 2 is an illustration of a vector plasmid for expression of an ethylene forming enzyme (EFE) protein according to embodiments herein.

[0037] Figure 3A is a photograph of an SDS-PAGE gel showing expression of an EFE protein according to embodiments herein.

5 [0038] Figure 3B is a photograph of a Western blot showing expression of an EFE protein according to embodiments herein.

[0039] Figure 4A is a graph showing the growth rate of *E. coli* BL 21 PUC19 EFE over time according to embodiments herein.

[0040] Figure 4B is a graph showing ethylene yield over time for an *E. coli* BL 21 PUC19 EFE culture according to embodiments herein.

[0041] Figure 5A is a photograph showing growth of bacterial colonies according to embodiments herein.

[0042] Figure 5B is a photograph showing growth of bacterial colonies according to embodiments herein.

15 [0043] Figure 6 is a photograph of a Southern blot showing the results of a cloning experiment for AKG and sucrose production according to embodiments herein.

[0044] Figure 7A is a photograph of a Southern blot showing the results of a cloning experiment for sucrose production according to embodiments herein.

[0045] Figure 7B is a photograph of a flask bacterial culture according to
20 embodiments herein.

[0046] Figure 8A is a photograph of a Southern blot showing the results of a cloning experiment for ethylene production according to embodiments herein.

[0047] Figure 8B is a photograph of a Southern blot showing the results of a cloning experiment for ethylene production according to embodiments herein.

25 [0048] Figure 9A is a photograph of a Southern blot showing the results of a cloning experiment for AKG production according to embodiments herein.

[0049] Figure 9B is a photograph of a flask bacterial culture according to
embodiments here.

30 DETAILED DESCRIPTION

[0050] Unless otherwise noted, all measurements are in standard metric units.

[0051] Unless otherwise noted, all instances of the words “a,” “an,” or “the” can refer to one or more than one of the word that they modify.

[0052] Unless otherwise noted, the phrase “at least one of” means one or more than one of an object. For example, “at least one nutrient” means one nutrient, more than one nutrient, or any combination thereof.

[0053] Unless otherwise noted, the term “about” refers to $\pm 10\%$ of the non-
5 percentage number that is described, rounded to the nearest whole integer. For example, about 100 ml/minute, would include 90 to 110 ml/minute. Unless otherwise noted, the term “about” refers to $\pm 5\%$ of a percentage number. For example, about 95% would include 90 to 100%. When the term “about” is discussed in terms of a range, then the term refers to the appropriate amount less than the lower limit and more than the upper limit. For example,
10 from about 100 to about 500 ml/minute would include from 90 to 550 ml/minute.

[0054] Unless otherwise noted, measurable properties (height, width, length, ratio etc.) as described herein are understood to be averaged measurements.

[0055] Unless otherwise noted, the terms “provide”, “provided” or “providing” refer to the supply, production, purchase, manufacture, assembly, formation,
15 selection, configuration, conversion, introduction, addition, or incorporation of any element, amount, component, reagent, quantity, measurement, or analysis of any composition of matter, method or system of any embodiment herein.

[0056] Sequence identity is herein defined as a relationship between two or more amino acid (polypeptide or protein) sequences or two or more nucleic acid
20 (polynucleotide) sequences, as determined by comparing the sequences. Usually, sequence identities or similarities are compared over the whole length of the sequences compared. In the art, “identity” also means the degree of sequence relatedness between amino acid or nucleic acid sequences, as the case may be, as determined by the match between strings of such sequences. “Similarity” between two amino acid sequences is determined by comparing
25 the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide. “Identity” and “similarity” can be readily calculated by various methods, known to those skilled in the art. In an embodiment, sequence identity is determined by comparing the whole length of the sequences as identified herein.

[0057] Exemplary methods to determine identity are designed to give the
30 largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Exemplary computer program methods to determine identity and similarity between two sequences include e.g. the BestFit, BLASTP (Protein Basic Local Alignment Search Tool), BLASTN (Nucleotide Basic Local Alignment Search Tool), and FASTA (Altschul, S. F. et al., J. Mol. Biol. 215:403-410 (1990), publicly

available from NCBI and other sources (BLAST.RTM. Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894). A most exemplary algorithm used is EMBOSS (European Molecular Biology Open Software Suite). Exemplary parameters for amino acid sequences comparison using EMBOSS are gap open 10.0, gap extend 0.5, Blosum matrix. Exemplary
5 parameters for nucleic acid sequences comparison using EMBOSS are gap open 10.0, gap extend 0.5, DNA full matrix (DNA identity matrix). In embodiments, it is possible to compare the DNA/ protein sequences among different species to determine the homology of sequences using online data such as Gene bank, KEG, BLAST and Ensemble.

[0058] Optionally, in determining the degree of amino acid similarity, the
10 skilled person may also take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled person. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of
15 amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-
20 tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or his; Asp to glu; Cys to ser or ala; Gln to asn; Glu to asp; Gly to pro; His to asn or gln; Ile to leu or val; Leu to ile or val; Lys to arg; gln or glu; Met to leu or ile; Phe to met, leu or tyr; Ser to thr; Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

[0059] Unless otherwise noted, the term "adapted" or "codon adapted" refers to "codon optimization" of polynucleotides as disclosed herein, the sequence of which may
30 be native or non-native, or may be adapted for expression in other microorganisms. Codon optimization adapts the codon usage for an encoded polypeptide towards the codon bias of the organism in which the polypeptide is to be expressed. Codon optimization generally helps to increase the production level of the encoded polypeptide in the host cell.

[0060] Carbon dioxide emissions resulting from the use of fossil fuels continue to rise on a global scale. Reduction of atmospheric carbon dioxide levels is a key to mitigating or reversing climate change. Carbon capture and storage (CCS) is a prominent technology for removal of industrial carbon dioxide from the atmosphere; it has been
5 estimated that over 20 trillion tons of carbon dioxide captured from refining and other industrial processes can be transported and stored in various types of subterranean environments or storage tanks. Although CCS is a cost effective and affordable way to reduce carbon dioxide emissions compared to other currently available methods, the problem remains that the carbon dioxide is merely being stored underground until it escapes.
10 Therefore, CCS methods do not provide a sustainable solution to reduce excess carbon dioxide in the atmosphere. Also, there is little financial incentive for industries to pump carbon dioxide into subterranean environments, unless they are forced to by environmental regulations, or they are paid to do it as part of their business model. Arguably, global warming is a crisis because it is more lucrative to produce carbon dioxide than to dispose of
15 carbon dioxide.

[0061] There remains a need to remove excess carbon dioxide from the atmosphere in more efficient and sustainable ways. There remains a need for technologies that can harness the over-abundance of carbon dioxide to make useful products, and for other applications that are beneficial to industry and the environment.

[0062] The challenges of the limited supply of petroleum, and the harmful effects of petroleum operations on the environment, have prompted a growing emphasis on maximizing output from existing resources, and in developing renewable sources of fuels and chemicals that can minimize environmental impacts. Since ethylene is a potentially renewable feedstock, there has been a great deal of interest in developing technologies to
25 produce ethylene from renewable sources, such as carbon dioxide and biomass. Ethylene is the most widely produced organic compound in the world, useful in a broad spectrum of industries including plastics, solvents, and textiles. Ethylene is currently produced by steam cracking fossil fuels or dehydrogenating ethane. With millions of metric tons of ethylene being produced each year, however, more than enough carbon dioxide is produced by such
30 processes to greatly contribute to the global carbon footprint. Producing ethylene through renewable methods would accordingly help to meet the huge demand from the energy and chemical industries, while also helping to protect the environment.

[0063] Conventional methods have been developed to produce bio-ethylene using ethanol derived from corn or sugar cane. However, the production of bio-ethylene from

biomass (e.g. corn and sugar cane) is a time-consuming and cost-ineffective process, requiring land, transportation, and digestion of biomass. For example, there are massive inefficiencies associated with the growing and transportation of corn and sugar cane, which by itself causes CO₂ emission. A variety of microbes, including bacteria and fungi, naturally produce ethylene in small amounts. Such microbes make use of an ethylene-forming enzyme (EFE). A type of ethylene pathway, such as is found in *Pseudomonas syringae* and *Penicillium digitatum*, uses alpha-ketoglutarate (AKG) and arginine as substrates in a reaction catalyzed by an ethylene-forming enzyme. Ethylene-forming enzymes provide a promising target, because expression of a single gene can be sufficient for ethylene production. Techniques making use of heterologous expression of an EFE have been demonstrated in several microbial species, where the microbial hosts have been able to utilize a variety of carbon sources in the Calvin cycle, including lignocellulose and carbon dioxide. Plus, recent developments in cost-effective high throughput genetic sequencing technologies have led to an increased understanding of microbial gene expression. However, the currently available technologies do not produce industrially relevant quantities of ethylene through microbial activity. There remains a need for improvements in microbial bio-ethylene production that can produce ethylene at a commercial scale. There remains a need for methods to produce ethylene useful for industrial and other applications using carbon dioxide feedstocks.

20 **[0064]** Embodiments of the present disclosure can provide a benefit not only of removing carbon dioxide from the environment along with the benefit of producing a valuable organic compound capable of being sold commercially. Embodiments of the present disclosure can thus provide a renewable alternative to conventional carbon dioxide storage, by using recombinant microbial technology to convert the carbon dioxide into ethylene as a useful organic compound. One benefit of the embodiments of the present disclosure is that the methods can make it economically profitable for an oil or natural gas company to remove carbon dioxide from the environment. An oil company, or a contractor thereof, could instead of pumping carbon dioxide into a subterranean environment or leaving the sequestered carbon dioxide underground, use the carbon dioxide as a carbon source for a culture of recombinant microorganisms to convert the carbon dioxide to ethylene in a cost-effective way. Also, much the carbon dioxide generated by transportation can be avoided because the process can be practiced on-site or would be expected to consumer more carbon dioxide than it produces.

[0065] The most effective methods for protecting the environment are those methods that people actually use. The more profitable those methods are; the more likely people are to use them. One of the benefits of the methods disclosed herein is the cost-effectiveness of using a bioreactor system. Embodiments of the present disclosure can provide a benefit of engineering a photosynthetic ethylene producing microorganism, by adapting the relevant metabolic signaling pathways to produce ethylene on an industrial scale. Such embodiments can make it profitable to remove carbon dioxide from the atmosphere and to passively generate valuable organic compounds while the microbes do the work – on a scale previously unimaginable.

[0066] What would happen to the global warming crisis if it became more profitable, or just as profitable, to convert carbon dioxide into valuable organic compounds as it did to generate the carbon dioxide in the first place? The presently disclosed methods might transform energy producers from global warming companies to global cooling companies.

[0067] The present disclosure relates to recombinant microorganisms having an improved ethylene producing ability. The present disclosure relates to methods of producing ethylene, including providing a recombinant microorganism having an improved ethylene producing ability according to various embodiments herein. As a general overview of a method disclosed herein, referring to FIG. 1, the method includes providing a recombinant microorganism expressing at least one EFE protein according to embodiments disclosed herein 102; culturing the recombinant microorganism in a bioreactor culture vessel under conditions sufficient to produce ethylene in the bioreactor culture vessel 104; increasing an amount of ethylene production by adding at least one activator to the culture within the bioreactor culture vessel, or adding carbon dioxide to a culture atmosphere within the bioreactor culture vessel 106; decreasing an amount of ethylene production by removing at least one molecular switch from the cell culture 108; controlling an amount of ethylene produced from the microbial culture by increasing or decreasing the concentration of at least one nutrient or the amount of at least one stimulus when culturing the recombinant microorganism 110; and removing an amount of ethylene produced from the microbial culture by condensing the ethylene from a gaseous to a liquid state 112. As an illustration of a vector plasmid for expression of an EFE protein according to embodiments herein, referring to FIG. 2, a non-native EFE expressing nucleotide sequence is inserted into the vector plasmid of a *Chlamydomonas* sp. bacterium. As an illustration of a recombinant microorganism having an improved ethylene producing ability herein, referring to the illustration of an SDS-PAGE gel in FIG. 3A and the illustration of a Western blot in FIG. 3B,

an EFE protein is expressed from a vector plasmid of an *Escherichia* sp. bacterium having a non-native EFE expressing nucleotide sequence inserted into the vector plasmid, as shown by the arrows.

5

Embodiments of Recombinant Microorganisms

- 10 **[0068]** The present disclosure relates to a recombinant microorganism having an improved ethylene producing ability. In such embodiments, the recombinant microorganism expresses at least one ethylene forming enzyme (EFE) protein having an amino acid sequence at least 95% identical to SEQ ID NO: 1 by expressing a non-native EFE expressing nucleotide sequence. In some embodiments, the non-native EFE expressing
- 15 nucleotide sequence encodes an EFE of *Pseudomonas savastanoi*. In an embodiment, the EFE protein has an amino acid sequence at least 80% or at least 90% identical to SEQ ID NO: 1. In an embodiment, the EFE protein has an amino acid sequence at least 98% identical to SEQ ID NO: 1. In various embodiments, an amount of EFE protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism
- 20 lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 5% to about 200% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 50% to about 150%
- 25 or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 75% to about 100% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence.
- 30 **[0069]** In an embodiment, the recombinant microorganism also expresses at least one alpha-ketoglutarate permease (AKGP) protein by expressing a non-native AKGP expressing nucleotide sequence. In an embodiment, the AKGP protein has an amino acid sequence at least 95% identical to SEQ ID NO: 2. In an embodiment, the AKGP protein has an amino acid sequence at least 80% or at least 90% identical to SEQ ID NO: 2. In an

embodiment, the AKGP protein has an amino acid sequence at least 98% identical to SEQ ID NO: 2. In an embodiment, the original sequence for SEQ ID NO: 2 was from AKGP from *Pseudomonas syringe*, but sequence innovation was performed to improve the expression of this sequence in *Synechococcus elongatus*. In various embodiments, an amount of AKGP protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native AKGP expressing nucleotide sequence. In some embodiments, the amount of AKGP protein produced by the recombinant microorganism is from about 5% to about 200% or more greater than that produced relative to the control microorganism lacking the non-native AKGP expressing nucleotide sequence. In some embodiments, the amount of AKGP protein produced by the recombinant microorganism is from about 50% to about 150% or more greater than that produced relative to the control microorganism lacking the non-native AKGP expressing nucleotide sequence. In some embodiments, the amount of AKGP protein produced by the recombinant microorganism is from about 75% to about 100% or more greater than that produced relative to the control microorganism lacking the non-native AKGP expressing nucleotide sequence.

[0070] In various embodiments, the recombinant microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*, algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell. In some embodiments, the recombinant microorganism can include *Saccharomyces cerevisiae*, *Pseudomonas putida*, *Trichoderma viride*, *Trichoderma reesei*, and tobacco.

[0071] In an embodiment, the non-native EFE expressing nucleotide sequence is inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3, and the non-native EFE expressing nucleotide sequence is inserted into a vector plasmid of a *Chlamydomonas* sp. bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in a *Chlamydomonas* sp. bacterium.

[0072] In an embodiment, a non-native EFE expressing nucleotide sequence and a non-native AKGP expressing nucleotide sequence are inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE
5 expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4, and the non-native EFE expressing nucleotide sequence and the AKGP expressing nucleotide sequence are inserted into a vector plasmid of an *Escherichia* sp. bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing
10 nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in an *Escherichia* sp. bacterium, or in an *Escherichia coli* bacterium.

[0073] In an embodiment, the recombinant microorganism further includes a non-native AKGP expressing nucleotide sequence. In some such embodiments, the non-
15 native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 5. In some embodiments, the non-native expressing nucleotide sequence and non-native AKGP expression nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6. In such embodiments, the combined amino acid sequence can
20 include one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a His-TEV sequence; in an embodiment, the His-TEV sequence includes SEQ ID NO. 10. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 90% identical to SEQ ID NO. 5 or
25 SEQ ID NO. 6. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 98% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus* sp. bacterium.
30 In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 95% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 90% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide

sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 98% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Cyanobacteria*. In an embodiment, the non-native EFE
5 expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Synechococcus sp.* bacterium.

Embodiments of Methods of Producing a Recombinant Microorganism

[0074] Embodiments herein are directed to methods of producing a
10 recombinant microorganism having an improved ethylene producing ability. In an embodiment, the method includes producing a recombinant microorganism by inserting a non-native EFE expressing nucleotide sequence into a bacterial plasmid of a microorganism. In some embodiments, the non-native EFE expressing nucleotide sequence encodes an EFE of *Pseudomonas savastanoi*. In an embodiment, the non-native EFE expressing nucleotide
15 sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for
20 expression in a *Chlamydomonas sp.* bacterium. In an embodiment, the *Chlamydomonas sp.* bacterium includes *Chlamydomonas reinhardtii*.

[0075] In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90%
25 identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in an *Escherichia sp.* bacterium. In an embodiment, the *Escherichia sp.* bacterium includes *E. coli*. In an embodiment, the non-native EFE expressing nucleotide sequence
30 includes an N-terminal NdeI cloning site (SEQ ID NO. 8 (See Appendix)). In an embodiment, the non-native EFE expressing nucleotide sequence includes one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a histidine tag at the C-terminal end, followed by a stop codon and a HindIII cloning site (SEQ ID NO. 9 (See Appendix)).

[0076] In an embodiment, the method includes producing a recombinant microorganism by inserting a combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence into a bacterial plasmid of a microorganism. In one such embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence expresses an amino acid sequence at least 95% identical to SEQ ID NO. 5. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expression nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6. In such embodiments, the combined amino acid sequence can include one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a His-TEV sequence; in an embodiment, the His-TEV sequence includes SEQ ID NO. 10 (See Appendix). In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 90% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 98% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus sp.* bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 95% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 90% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 98% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Cyanobacteria*. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Synechococcus sp.* bacterium.

[0077] In various embodied methods, the microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*,

algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell. In some embodiments, the recombinant microorganism can include *Saccharomyces cerevisiae*, *Pseudomonas putida*, *Trichoderma viride*, *Trichoderma reesei*, and tobacco.

5 **[0078]** In embodiments of methods herein, the non-native EFE expressing nucleotide sequence is inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3, and the non-native EFE
10 expressing nucleotide sequence is inserted into a vector plasmid of a *Chlamydomonas* sp. bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence is
15 codon adapted for expression in a *Chlamydomonas* sp. bacterium.

[0079] In an embodiment, a non-native EFE expressing nucleotide sequence and a non-native AKGP expressing nucleotide sequence are inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE
20 expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4, and the non-native EFE expressing nucleotide sequence and the AKGP expressing nucleotide sequence are inserted into a vector plasmid of an *Escherichia* sp. bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing
25 nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in an *Escherichia* sp. bacterium, or in an *Escherichia coli* bacterium.

[0080] In an embodiment, the recombinant microorganism further includes a non-native AKGP expressing nucleotide sequence. In some such embodiments, the non-
30 native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 5. In some embodiments, the non-native expressing nucleotide sequence and non-native AKGP expression nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6. In such embodiments, the combined amino acid sequence can

include one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a His-TEV sequence; in an embodiment, the His-TEV sequence includes SEQ ID NO. 10. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 90% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 98% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus sp.* bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 95% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 90% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 98% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Cyanobacteria*. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Synechococcus sp.* bacterium.

Embodiments of Methods of Producing Ethylene

[0081] Methods of producing ethylene are embodied herein. An embodiment of such a method includes providing a recombinant microorganism having an improved ethylene producing ability. In an embodiment, the recombinant microorganism expresses at least one ethylene forming enzyme (EFE) protein having an amino acid sequence at least 95% identical to SEQ ID NO: 1 by expressing a non-native EFE expressing nucleotide sequence, wherein an amount of EFE protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native EFE expressing nucleotide sequence; culturing the recombinant microorganism in a bioreactor culture vessel under conditions sufficient to produce ethylene in the bioreactor culture vessel; and harvesting ethylene from the bioreactor culture vessel.

[0082] In some embodiments of methods of producing ethylene, the non-native EFE expressing nucleotide sequence encodes an EFE of *Pseudomonas savastanoi*. In an embodiment, the EFE protein has an amino acid sequence at least 80% or at least 90% identical to SEQ ID NO: 1. In an embodiment, the EFE protein has an amino acid sequence at least 98% identical to SEQ ID NO: 1. In various embodiments, an amount of EFE protein produced by the recombinant microorganism is greater than that produced relative to a control microorganism lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 5% to about 200% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 50% to about 150% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence. In some embodiments, the amount of EFE protein produced by the recombinant microorganism is from about 75% to about 100% or more greater than that produced relative to the control microorganism lacking the non-native EFE expressing nucleotide sequence.

[0083] In an embodiment of methods of producing ethylene, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 80% or at least 90% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in a *Chlamydomonas sp.* bacterium. In an embodiment, the *Chlamydomonas sp.* bacterium includes *Chlamydomonas reinhardtii*.

[0084] In an embodiment of methods herein, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in an *Escherichia sp.* bacterium. In an embodiment, the *Escherichia sp.* bacterium includes *E. coli*. In an embodiment, the non-native EFE expressing nucleotide sequence includes an N-terminal NdeI cloning site (SEQ ID NO. 8). In an embodiment, the non-native EFE expressing nucleotide sequence includes one or more purification tags; in an

embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a histidine tag at the C-terminal end, followed by a stop codon and a HindIII cloning site (SEQ ID NO. 9).

[0085] In an embodiment, the method of producing ethylene includes

5 producing a recombinant microorganism by inserting a combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence into a bacterial plasmid of a microorganism. In one such embodiment, the combined non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence expresses an amino acid sequence at least 95% identical to SEQ ID NO. 5. In some

10 embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expression nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6. In such embodiments, the combined amino acid sequence can include one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a His-TEV sequence; in an

15 embodiment, the His-TEV sequence includes SEQ ID NO. 10. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 90% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid

20 sequence at least 98% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus sp.* bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 95% identical to SEQ

25 ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 90% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 98% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE

30 expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Cyanobacteria*. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Synechococcus sp.* bacterium.

[0086] In various embodied methods, the microorganism includes a *Cyanobacteria*, a *Synechococcus*, *Synechococcus elongatus*, *Synechococcus leopoliensis*, *Synechocystis*, *Anabaena*, a *Pseudomonas*, *Pseudomonas syringae*, *Pseudomonas savastanoi*, *Chlamydomonas*, *Chlamydomonas reinhardtii*, *Escherichia*, *Escherichia coli*, *Geobacteria*,
5 algae, microalgae, electrosynthesis bacteria, a photosynthetic microorganism, yeast, filamentous fungi, or a plant cell. In some embodiments, the recombinant microorganism can include *Saccharomyces cerevisiae*, *Pseudomonas putida*, *Trichoderma viride*, *Trichoderma reesei*, and tobacco.

[0087] In embodiments of methods herein, the non-native EFE expressing
10 nucleotide sequence is inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID NO. 3, and the non-native EFE expressing nucleotide sequence is inserted into a vector plasmid of a *Chlamydomonas* sp.
15 bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 3. In an embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in a *Chlamydomonas* sp. bacterium.

20 [0088] In an embodiment, a non-native EFE expressing nucleotide sequence and a non-native AKGP expressing nucleotide sequence are inserted into a bacterial vector plasmid, a nucleotide guide of a homologous recombination system, a CRISPR CAS system, a phage display system, or a combination thereof. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 95% identical to SEQ ID
25 NO. 4, and the non-native EFE expressing nucleotide sequence and the AKGP expressing nucleotide sequence are inserted into a vector plasmid of an *Escherichia* sp. bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 90% identical to SEQ ID NO. 4. In an embodiment, the non-native EFE expressing nucleotide sequence has a nucleotide sequence at least 98% identical to SEQ ID NO. 4. In an
30 embodiment, the non-native EFE expressing nucleotide sequence is codon adapted for expression in an *Escherichia* sp. bacterium, or in an *Escherichia coli* bacterium.

[0089] In an embodiment, the recombinant microorganism further includes a non-native AKGP expressing nucleotide sequence. In some such embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide

sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 5. In some embodiments, the non-native expressing nucleotide sequence and non-native AKGP expression nucleotide sequence express a combined amino acid sequence at least 95% identical to SEQ ID NO. 6. In such embodiments, the combined amino acid sequence can include one or more purification tags; in an embodiment, the purification tag includes a histidine tag. In an embodiment, the purification tag includes a His-TEV sequence; in an embodiment, the His-TEV sequence includes SEQ ID NO. 10. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 90% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In other embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence express a combined amino acid sequence at least 98% identical to SEQ ID NO. 5 or SEQ ID NO. 6. In some embodiments, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence are inserted into a bacterial plasmid of a *Synechococcus sp.* bacterium. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 95% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 90% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and non-native AKGP expressing nucleotide sequence include a nucleotide sequence at least 98% identical to SEQ ID NO. 7. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Cyanobacteria*. In an embodiment, the non-native EFE expressing nucleotide sequence and the non-native AKGP expressing nucleotide sequence are codon adapted for expression in a *Synechococcus sp.* bacterium.

[0090] Embodiments of producing ethylene herein include culturing a recombinant microorganism in a bioreactor culture under conditions sufficient to produce ethylene in the bioreactor culture vessel. A bioreactor culture according to embodied methods can include one or more suitable reagents or growth media for supporting the growth of the recombinant microorganism culture. Such reagents or culture media can include water, one or more carbohydrates, one or more amino acids or amino acid derivatives, one or more buffers, sea water, Luria broth, Luria Bertani broth, BG-11 media, carbon dioxide, light, temperature, electricity, or combinations thereof.

[0091] An embodiment of a method of producing ethylene includes increasing an amount of ethylene production by adding at least one activator to a culture containing the recombinant microorganism located within the bioreactor culture vessel. The addition of such an activator can include increasing a concentration of one or more substrates of the EFE enzyme being expressed by the recombinant microorganism culture. Such a substrate can include alpha-ketoglutarate or arginine, or combinations thereof as well as other sources of carbon such as glycerol and glucose. In other embodiments, adding at least one activator can include adding a molecular switch. In some embodiments, adding at least one activator can include insertion of an inducible promoter upstream of the EFE gene; one such promoter includes an IPTG promoter. In such embodiments, IPTG can be added as a molecular switch to the culture media. In some embodiments, adding at least one activator can include adding one or more nutrients or stimuli to the culture. Such nutrients or stimuli can include one or more carbohydrates, one or more amino acids or amino acid derivatives, one or more EFE substrates, succinate, carbon dioxide, light, temperature, electricity, glycerol, sugars, or combinations thereof. In such embodiments, adding at least one activator to the culture can provide a benefit of controlling the cycles of ethylene production and enhancing the ethylene production rate. In some embodiments, the ethylene produced can be removed from the bioreactor culture vessel as it is produced. In such embodiments, removal of the ethylene can include condensing ethylene produced as a gas into a liquid form for removal from the bioreactor culture vessel.

[0092] In an embodiment, a method of producing ethylene includes adding CO₂ to a culture atmosphere contained within the bioreactor culture vessel at rate of between about 100 ml/minute and about 500 ml/minute. In an embodiment, the method includes adding CO₂ to a culture atmosphere contained within the bioreactor culture vessel at rate of between about 150 ml/minute and about 450 ml/minute. In an embodiment, the method includes adding CO₂ to a culture atmosphere contained within the bioreactor culture vessel at rate of between about 250 ml/minute and about 350 ml/minute. Such embodiments can provide a benefit of enhancing or controlling the rate of ethylene production in the bioreactor culture vessel, as well as providing a benefit of converting CO₂ into a useful product.

[0093] In an embodiment, a method of producing ethylene includes decreasing an amount of ethylene production by removing at least one molecular switch from the microbial culture containing the recombinant microorganism located within the bioreactor culture vessel. In an embodiment, such a method further includes controlling the amount of ethylene produced from the microbial culture by increasing or decreasing the concentration of

at least one nutrient or the amount of at least one stimulus when culturing the recombinant microorganism. In an embodiment, the concentration of at least one nutrient and the amount of at least one stimulus are at a ratio of from about 0.5-1.5 gr./liter to about 0.1 mM in the microbial culture. In an embodiment, such a method further includes removing the amount of ethylene produced from the microbial culture by condensing the ethylene from a gaseous to a liquid state, or wherein the amount of ethylene recovered is from about 0.5 ml to about 10 ml/liter/h. Such embodiments can provide a benefit of controlling the amount of ethylene production by controlling the rate of activity of the Calvin cycle in the microbial culture. For example, it is possible to shift from an expression system to a growth system, where the cells are allowed to grow for 5-7 days and their growth conditions are monitored. When the cells reach an exponential growth condition (meaning that the cells are metabolically active), it is possible to shift from the growth system to an expression system, where the cells are shifted to an ethylene production cycle to produce ethylene for harvesting. This expression system might be maintained for 7 or more days.

EXAMPLES

Example 1. Cloning of Ethylene Forming Enzyme Gene Sequence into *Chlamydomonas reinhardtii* Vector Plasmid

[0094] An EFE (Ethylene Forming Enzyme) protein will be expressed and produced in *Chlamydomonas reinhardtii*. Plasmid pChlamy_4-EFE was generated successfully, to be used in EFE protein expression (Creative Enzymes, Shirley, NY).

[0095] The polynucleotide coding for the *Pseudomonas savastanoi* pv. *Phaseolicola* EFE protein (GenBank: KPB44727.1, SEQ ID NO: 1) was cloned into the pChlamy_4 vector plasmid (ThermoFisher). Other reagents and use of instruments were provided by Creative Biostructure.

[0096] The Ethylene-forming enzyme (EFE) gene sequence from strain *Pseudomonas savastanoi* pv. *Phaseolicola* (GenBank: KPB44727.1) was used for the preparation of EFE recombinant protein. The corresponding nucleotide sequences were codon adapted for expression in *Chlamydomonas reinhardtii* and synthesized (SEQ ID NO: 3). The EFE construct was cloned into the pChlamy_4 vector with the *KpnI* and *PstI* restriction enzyme sites.

[0097] According to the pChlamy_4vector character, EFE with an N-tag was designed and generated. The pChlamy_4 vector contains the ATG initiation codon (vector ATG) for proper initiation of translation at position 497-499, found at the beginning of the *Sh*

ble gene after the removal of Intron-1 Rbc S2. The FMDV 2A peptide gene flanking the Multiple Cloning Site 1 (MCS1) is in frame with the *Sh ble* gene. To use the N-terminal 6× His-V5-TEV tag, the EFE sequence was cloned in-frame after the TEV site, into the KphI/PstI digested pChlamy_4 vector. A TAA (stop codon) was designed for proper translation termination. The resulting sequence chromatogram is shown in FIG. 2; referring to FIG. 2, the EFE protein gene coding sequences are shown in the arrow labeled “EFE-protein”. An open reading frame orientation was confirmed by plasmid validation by nucleotide sequencing.

Example 2: Protein Expression Evaluation of Expression of EFE in *E. coli*

[0098] The polynucleotide coding for the *Pseudomonas savastanoi* pv. Phaseolicola EFE protein (GenBank: KPB44727.1, SEQ ID NO: 1) was cloned into the pET-30a(+) vector plasmid. The corresponding nucleotides sequences were codon adapted for expression in *E. coli* (SEQ ID NO: 4), containing an optional His tag at the C-terminal end followed by a stop codon and HindII site (SEQ ID NO: 9). An NdeI site was used for cloning at the 5-prime end, where the NdeI site contains an ATG start codon (SEQ ID NO: 8). *E. coli* BL21(DE3) competent cells were transformed with the recombinant plasmid. A single colony was inoculated into LB medium containing kanamycin; cultures were incubated in 37 °C at 200 rpm. Once cell density reached to OD=0.6-0.8 at 600 nm, 0.5 mM IPTG was introduced for induction. A pilot expression of EFE (about 44.5 kDa)_BL21(DE3) was conducted. SDS-PAGE (FIG. 3A) and Western blotting (FIG. 3B) were used to monitor the EFE protein expression (GenScript USA, Inc., Piscataway, NJ). Referring to FIG. 3A and FIG. 3B:

[0099] SDS PAGE (left) and Western blot (right, using anti His antibody (GenScript , Cat.No.A00186)) analysis of Pilot expression of EFE in *E. coli* expression in construct pET 30a(+).

Lane M₁: Protein marker
 Lane M₂: Western blot marker
 Lane PC₁: BSA (1 μg)
 Lane PC₂: BSA (2 μg)
 Lane NC: Cell lysate without induction
 Lane 1: Cell lysate with induction for 16 h at 15°C
 Lane 2: Cell lysate with induction for 4 h at 37°C

Lane NC₁: Supernatant of cell lysate without induction

Lane 3: Supernatant of cell lysate with induction for 16 h at 15 °C

Lane 4: Supernatant of cell lysate with induction for 4 h at 37 °C

Lane NC₂: Pellet of cell lysate without induction

5 Lane 5: Pellet of cell lysate with induction for 16 h at 15°C

Lane 6: Pellet of cell lysate with induction for 4 h at 37°C

[0100] The results of SDS-PAGE and Western blots showed that EFE was expressed in *E. coli*. The highest EFE expression conditions were found with induction for 16 h at 15°C that resulted in an expression level of 5 mg/L and a solubility of 30%.

10 **Example 3: Recombinant EFE and AKGP Expressing Nucleotide Sequences Adapted for Expression in *Synechococcus* spp. Bacteria**

[0101] A combined polynucleotide sequence (EFE_-P2A-aKGP, SEQ ID NO: 7) for expressing the EFE protein and for expressing the AKGP protein (SEQ ID NO. 2) was generated, after adaptation of the nucleotide sequence for expression in *Cyanobacteria* species *Synechococcus elongates* and *Synechococcus leopoliensis* (GenScript, Piscataway, NJ). An codon adaptation analysis algorithm was used which adapts a variety of parameters that are critical to the efficiency of gene expression, including but not limited to codon usage bias, GC content, CpG dinucleotide content, mRNA secondary structure, cryptic splicing sites, premature PolyA sites, internal chi sites and ribosomal binding sites, negative CpG islands, RNA instability motif (ARE), repeat sequences (direct repeat, reverse repeat, and Dyad repeat), and restriction sites that may interfere with cloning. A codon usage bias adjustment was performed using the distribution of codon usage frequency along the length of the gene sequence, with a resulting Codon Adaptation Index (CAI) of 0.95. A CAI of 1.0 is considered to be perfect in the desired expression organism, and a CAI of greater than 0.8 is regarded as good, in terms of high gene expression level. The Frequency of Optimal Codons (FOP) was measured as the percentage distribution of favorable codons in computed codon quality groups, with the value of 100 set for the codon with the highest usage frequency for a given amino acid in the desired expression organism. A result of 80% of the codons was found in the highest codon quality group of 91-100, 3% in the second highest quality group of 81-90, and 14% in the third highest quality group of 71-80. A GC content adjustment was performed resulting in an average GC content of 56.46%, with the ideal percentage range of GC content being between 30-70%. One optional HindIII cloning site

(SEQ ID NO: 12) was incorporated at the 5-prime end at position 1 of the sequence; one optional KpnI cloning site (SEQ ID NO: 13) was incorporated at the 3-prime end at position 2524 of the sequence.

[0102] The corresponding combined EFE and AKGP amino acid sequences expressed by SEQ ID NO. 7 have a P2A cleavage sequence (SEQ ID NO: 11) inserted between the EFE and AKGP amino acid sequences. The encoded amino acid sequence having the EFE and AKGP sequence together is shown in SEQ ID NO. 5 (EFE-P2A_pSyn_6 (No His). An optional His-TEV sequence may be included at the N-terminus (SEQ ID NO: 10), resulting in the amino acid sequence of SEQ ID NO. 6 (EFE-P2A-aKGP_pSyn_6).

Example 4: Lab Scale Experimental Procedures.

[0103] 1. Tailored-designed DNA constructions will be generated that encode the critical intermediates of a synthetic bio-ethylene pathway. 2. Carefully selected photosynthetic microorganisms will then be expanded for cloning and gene expression. 3. Genetic and metabolic engineering of microorganisms will then be performed for continuous production of bio-ethylene. 4. Bioengineered microorganisms will then be selected and expanded in a photobioreactor. 5. Bioreactor culture conditions (including CO₂ concentration, light exposure time and wave-length, temperature, pH) will be adapted. 6. Samples will be collected and analyzed by HPLC to measure bio-ethylene synthesis. 7. Bio-ethylene production in genetically engineered microorganisms will be adapted. 8. Ethylene production processes will be scaled up.

Example 5: Engineering the production of AKG in *Cyanobacteria*

[0104] Previous work has shown that deletion of the *glgC* gene leads to an increase in AKG production in *Cyanobacteria*. It has also been shown that over expression of the *ppc* gene (SEQ ID NO. 14, Genbank P74299), which encodes phosphoenolpyruvate synthase (SEQ ID NO. 15), and the *gluA* gene (SEQ ID NO. 16, Genbank Q59977), which encodes citrate synthase (SEQ ID NO. 17), can enhance the production of AKG as a substrate for producing other compounds. Based on this research, two categories of genes were chosen, including genes that are related directly to AKG synthesis and secretion pathways, including *ppc* and *gluA* (overexpression), and genes that are involved in energy storage pathways, including *glgC* (deletion), which plays a critical role in the glycogen synthesis pathway. A construct of *ppc*-p2A-*gluA* (SEQ ID NO. 18) was created for cloning into the pSyn6 plasmid

before integration into *Synechococcus elongatus* and growth of transformed colonies. PCR was performed on pSyn6-PPC-gltA colonies to confirm the expression of the construct in Cyanobacteria; an expected band size for PPC-gltA of 4621 base pairs was observed.

[0105] For the overexpression of AKG in *Cyanobacteria*, a construct of an IDH gene (SEQ ID NO: 19), which encodes isocitrate dehydrogenase (SEQ ID NO. 20), was made by cloning the IDH gene into the pSyn6 plasmid. Successful cloning of the IDH gene into the pSyn6-IDH plasmid was confirmed by growth of bacterial colonies (FIG. 5A) and by gel electrophoresis and DNA analysis (FIG. 6). *Synechococcus elongatus* strain S2434-IDH integrating the IDH construct was confirmed by bacterial culture growth (FIG. 9B) and gel electrophoresis and DNA analysis (FIG. 9A). Cell culture growth was shown to be improved significantly by increasing the bicarbonate concentration in the growth medium by 0.5 g/L or 1.0 g/L. A plasmid for deletion of the *glgC* gene (SEQ ID NO. 21, Genbank CP000100.1), which encodes glucose-1-phosphate adenylyltransferase (SEQ ID NO. 22), in *Cyanobacteria* (*Synechococcus elongatus*), was also made and confirmed.

15 **Example 6: Engineering the production of sucrose in *Cyanobacteria***

[0106] For production of sucrose in *Cyanobacteria*, a construct of a *cscB* gene from *E. coli* (SEQ ID. NO. 23, Genbank P300000), which encodes sucrose permease (SEQ ID NO. 24), was made by cloning of the *cscB* gene into the pSyn6 plasmid. Successful cloning of the *cscB* gene into the pSyn6-cscB construct was confirmed by bacterial colony growth (FIG. 5B) and by gel electrophoresis and DNA analysis (FIG. 6). *Synechococcus elongatus* strain UTEX S2434 (S2434-cscB) integrating *cscB* was confirmed by bacterial culture growth (FIG. 7B) and by gel electrophoresis and DNA analysis (FIG. 7A).

[0107] In addition to the overexpression of the *cscB* gene, and the deletion of the *glgC* gene, other gene targets include overexpression of the *sps* gene (SEQ ID NO. 25, Genbank A0A0H3K0V9), which encodes sucrose phosphate synthase (SEQ ID NO. 26); the *spp* gene (SEQ ID NO. 27, Genbank Q7BII3), which encodes sucrose-6-phosphatase (SEQ ID NO. 28), the *glgP* gene (SEQ ID NO. 29, Genbank Q31RP3), which encodes glycogen phosphorylase (SEQ ID NO. 30), and the *galU* gene (SEQ ID NO. 31, Genbank P0AEP3), which encodes UTP-glucose-1-phosphate uridylyltransferase (SEQ ID NO. 32), to reroute the intermediates to sucrose. In addition, deletion of the *inv* gene (SEQ ID NO. 33, Genbank P74573), which encodes invertase (SEQ ID NO. 34), and the *ggpS* gene (SEQ ID NO. 35, Genbank P74258), which encodes glucosylglycerol-phosphate synthase (SEQ ID NO. 36), will prevent conversion to alternative products; and deletion of the *glgA* gene (SEQ ID NO.

37, Genbank P74521), which encodes glycogen synthase (SEQ ID NO. 38), will eliminate the conversion of substrate to glycogen, which potentially can increase the sucrose yield.

Example 7: Engineering the production of ethylene in *E. coli*

[0108] For engineering production of ethylene in *E. coli*, gene construct pUC-EFE (SEQ ID NO. 39) was made encoding ethylene forming enzyme (EFE) under an IPTG-inducible promoter in a high copy number plasmid, pUC19. Expression of the PUC-EFE plasmid in *E. coli* was confirmed by colony growth on agar media supplemented with ampicillin, IPTG and X-gal, and observance of the expected band size of 2322 base pairs by gel electrophoresis and DNA analysis (FIG. 8A and FIG. 8B). In FIG. 8A, the arrow shows the EFE DNA construct; in FIG. 8B, the arrow shows the DNA element controlling plasmid copy number. DNA sequencing results confirmed the presence of the plasmid. EFE production was confirmed by SDS-PAGE and Western blot analysis. Ethylene expression levels of 5 mg/L and 30% solubility were observed under induction conditions of 16 hours at 15 degrees Celsius.

[0109] For engineering production of ethylene in *E. coli*, a plasmid was constructed for continuous production of EFE in *E. coli*. In all approaches, the EFE expression was under control of the chloroplast psbA promoter. In a first construct EFE-AKGP-psbA (SEQ ID NO. 40), the EFE and AKGP genes were placed under control of the psbA promoter (SEQ ID NO. 41) and the rrnB terminator (SEQ ID NO. 42). In a second construct EFE-psbA (SEQ ID NO. 43), only EFE gene expression was placed under control of the psbA promoter (SEQ ID NO. 41) and the T7 terminator (SEQ ID NO. 44). Both constructs were cloned into a pUC19 plasmid backbone, to take advantage of the high copy number of the plasmid, before expressing the protein in *E. coli* BL21 (DE3), DH5alpha, or MG1655 cell lines. A pUC-psb-EFE plasmid was constructed (SEQ ID NO. 45).

[0110] The effect of growth media, as well as AKG and arginine supplementation, on ethylene production was measured. The results indicated that a maximum ethylene production of 0.037 lb/gallon/month for *E. coli* BL 21 PUC19 EFE was obtained when fermented under the conditions shown in Table 1.

Table 1. Conditions for the production of ethylene for *E. coli* BL 21 PUC19 EFE at 30 °C

Media	MOPS
Glucose	4g/L

IPTG	0.5 mM
Arginine	3 mM
AKG	2 mM
Induction	Induced at the start

- 5 **[0111]** The results of the observed growth rate of *E. coli* BL 21 PUC19 EFE is shown in FIG. 4A. The observed ethylene yield under the conditions shown in Table 1 is shown in FIG. 4B. Gas chromatography analysis of headspace samples confirmed the production of ethylene by the *E. coli* culture.

APPENDIX

SEQ ID NO: 1 –

MIHAPSRWGVFPSLGLCSPDVVWNEHPSLYMDKEETSMTNLQTFELPTEVTGCAADI
5 SLGRALIQAWQKDGIFQIKTDSEQDRKTQEAMAASKQFCKEPLTFKSSCVSDLTYSG
YVASGEEVTAGKPDFPEIFTVCKDLSVGDQRVKAGWPCHGPVPWPNNQYQKSMKT
FMEELGLAGERLLKLTALGFELPINTFTDLTRDGWHHMRVLRFPQTSTLSRGIGAHT
DYGLLVIAAQDDVGGLYIRPPVEGEKRNRLWLPGESSAGMFEHDEPWTFVTPTPGV
WTVFPGDILQFMTGGQLLSTPHKVKLNTRERFACAYFHEPNFEASAYPLFEPSANERI
10 HYGEHFTNMFMRCYPDRITTQRINKENRLAHLEDLKKYSDTRATGS

SEQ ID NO: 2 -

MTESITSNGTLVASDTRRRVWAIVSASSGNLVEWFDYVYSFCSLYFAHIFFPNGNTT
TQLLQTAGVFAAGFLMRPIGGWLFGRIDRRGRKTSMLISVCMFCGSLIIACLPGY
15 DAIGTWAPALLLLARLFQGLSVGGEYGTSAITYMSEIALEGRKGFYASFQYVTLIGGQ
LLAILVVVILQQILTDSQLHEWGWRIPFAMGAALAIVALWLRRQLDETSQKEVRALK
EAGSFKGLWRNRKAFLMVLGFTAGGSLSFYFTTYMQKYL VNTTGMHANVASVIM
TAALFVFMILIQPLIGALS D KIGRRTSMLIFGGMSALCTVPILTALQHVSSPYAAAFALV
MLAMVIVSFYTSISGILKAEMFPAQVRALGVGLSYAVANALFGGSAEYVALSLKSW
20 GSETTFFWYVTIMGALAFIVSLMLHRKGKGIRL

SEQ ID NO. 3 -

ATGATTCACGCCCCGTCGCGCTGGGGCGTGTTTCCCTCGCTGGGCCTGTGCAGCC
CCGACGTGGTGTGGAACGAGCACCCGAGCCTGTACATGGACAAGGAGGAGACGT
25 CGATGACCAACCTGCAGACGTTTCGAGCTGCCGACCGAGGTGACCGGCTGCGCCG
CCGACATCTCCCTGGGCCGGGCGCTGATCCAGGCGTGGCAGAAGGACGGCATCT
TCCAGATCAAGACCGACAGCGAGCAGGACCGGAAGACCCAGGAGGCGATGGCG
GCCTCCAAGCAGTTCTGCAAGGAGCCCCTGACCTTCAAGTCGTCCTGCGTCAGCG
ACCTGACCTACTCGGGCTACGTGGCCTCGGGCGAGGAGGTGACCGCCGGCAAGC
30 CGGACTTTCCGGAGATCTTCACCGTGTGCAAGGACCTGAGCGTGGGCGACCAGC
GGGTCAAGGCGGGCTGGCCCTGCCACGGCCCCGTGCCGTGGCCGAACAACACCT
ACCAGAAGTCCATGAAGACGTTTCATGGAGGAGCTGGGCCTGGCCGGCGAGCGCC
TGCTGAAGCTGACCGCGCTGGGCTTCGAGCTGCCCATCAACACGTTACCGACCT
GACCCGGGACGGCTGGCACCATGCGCGTCCTGCGGTTTCCGCCCCAGACCAG
35 CACGCTGAGCCGCGGCATTGGCGCGCACACGGACTACGGCCTGCTGGTGATTGC
CGCGCAGGACGACGTGGGCGGCCTGTACATTCGCCCCGCCGTGGAGGGCGAGAA
GCGCAACCGGAACCTGGCTGCCCGGCGAGTCCTCGGCGGGCATGTTTCGAGCACGA
CGAGCCCTGGACGTTCTGTGACCCCCACGCCGGGCGTGTGGACGGTGTTCCTGGC
GACATCCTGCAGTTCATGACCGGCGGCCAGCTG
40 CTGTCGACGCCGCACAAGGTGAAGCTGAACACCCGGGAGCGCTTCGCCTGCGCG
TACTTCCACGAGCCGAACCTCGAGGCCTCGGCCTACCCCTGTTTCGAGCCCTCCG
CGAACGAGCGCATCCACTACGGCGAGCACTTCACCAATATGTTTATGCGCTGCTA

CCCCGACCGCATCACCACCCAGCGCATCAACAAGGAGAATCGCCTGGCGCACCT
GGAGGACCTGAAGAAGTACAGCGACACCCGCGCCACCGGCTCG

SEQ ID NO. 4-

5 ATGATACACGCTCCAAGTAGATGGGGAGTATTTCCCTCACTAGGGTTATGCAGCC
CGGACGTTGTGTGGAATGAGCATCCGAGCCTGTACATGGACAAAGAGGAAACCA
GCATGACCAACCTGCAGACCTTTGAACTGCCGACCGAAGTGACCGGTTGCGCGG
CGGACATCAGCCTGGGTCGTGCGCTGATTCAGGCGTGGCAAAAGGATGGTATCT
TCCAGATTAAAACCGACAGCGAGCAGGATCGTAAGACCCAAGAAGCGATGGCG
10 GCGAGCAAGCAATTTTGCAAAGAGCCGCTGACCTTCAAAGCAGCTGCGTTAGC
GACCTGACCTACAGCGGTTATGTGGCGAGCGGCGAGGAAGTTACCGCGGGCAAG
CCGGATTTCCCGGAAATTTTACCGTGTGCAAGGACCTGAGCGTGGGCGATCAGC
GTGTTAAAGCGGGTTGGCCGTGCCATGGTCCGGTTCCGTGGCCGAACAACACCTA
TCAAAGAGCATGAAAACCTTTATGGAGGAACTGGGTCTGGCGGGCGAGCGTCT
15 GCTGAAACTGACCGCGCTGGGTTTTGAACTGCCGATCAACACCTTCACCGACCTG
ACCCGTGATGGCTGGCACCACATGCGTGTGCTGCGTTTCCCGCCGCAGACCAGCA
CCCTGAGCCGTGGTATTGGTGCGCACACCGACTACGGTCTGCTGGTGATTGCGGC
GCAAGACGATGTTGGTGGCCTGTATATCCGTCCGCCGGTGGAGGGCGAAAAGCG
TAACCGTAACTGGCTGCCGGGCGAGAGCAGCGCGGGCATGTTTGAGCACGACGA
20 ACCGTGGACCTTCGTTACCCCGACCCCGGGTGTGTGGACCGTTTTTCCGGGCGAT
ATTCTGCAGTTCATGACCGGTGGCCAACTGCTGAGCACCCCGCACAAGGTTAAAC
TGAACACCCGTGAACGTTTCGCGTGCGCGTACTTTCACGAGCCGAAGTTTCAAGC
GAGCGCGTATCCGCTGTTTCGAGCCGAGCGCGAACGAACGTATCCACTACGGCGA
GCACTTCACCAACATGTTTATGCGTTGCTATCCGGATCGTATCACCACCAACGT
25 ATTAACAAAGAAAACCGTCTGGCGCACCTGGAAGACCTGAAGAAATACAGCGAC
ACCCGTGCGACCGGCAGC

SEQ ID NO. 5-

MIHAPSRWGVFPSLGLCSPDVVWNEHPSLYMDKEETSMTNLQTFELPTEVTGCAADI
30 SLGRALIQAWQKDGIFQIKTDSEQDRKTQEAMAASKQFCKEPLTFKSSCVSDLTYSG
YVASGEEVTAGKPDFPEIFTVCKDLSVG DQRVKAGWPCHGPVPWPNNYQKSMKT
FMEELGLAGERLLKLTA LGFELPINTFTDLTRDGWHHMRVLRFPQTSTLSRGIGAHT
DYGLLVIAAQDDVGGLYIRPPVEGEKRNRLWLP GESSAGMFEHDEPWTFVTPTPGV
WTVFPGDILQFMTGGQLLSTPHKVKLNTRERFACAYFHEPNFEASAYPLFEPSANERI
35 HYGEHFTNMFMRCYPDRITTQRINKENRLAHLEDLKKYSDTRATGSGATNFSLLKQ
AGDVEENPGPMTESITSNGTLVASDTRRRVWAIVSASSGNLVEWFDYFYVSFCSLYF
AHIFFPSGNTTTQLLQTAGVFAAGFLMRPIGGWLFGRIDRRGRKTSMLISVCMCF
GSLIIACLPGYDAIGTWAPALLLLARLFQGLSVGGEYGTSATYMSEIALEGRKGFYAS
FQYVTLIGGQLLAILVVVILQQILTDSQLHEWGWRI PFAMGAALAIVALWLRRLDE
40 TSQKEVRALKEAGSFKGLWRNRKAFLMVLGFTAGGSLSFYTFTTYMQKYLNTTG
MHANVASVIMTAALFVFM LIQPLIGALSDKIGRRTSMLIFGGMSALCTVPILTALQHV
SSPYAAFALVMLAMVIVSFYTSISGILKAEMFPAQVRALGVGLSYAVANALFGGSAE
YVALSLKSWGSETTFWYVTIMGALAFIVSLMLHRKGKGIRL

SEQ ID. NO. 6-

MHHHHHHENLYFQGKLMIHAPSRWGVFPSLGLCSPDVVWNEHPSLYMDKEETSMT
NLQTFELPTEVTGCAADISLGRALIQAQWQKDGIFQIKTDSEQDRKTQEAMAASKQFC
5 KEPLTFKSSCVSDLTYSYGYVASGEEVTAGKPDFPEIFTVCKDLSVGDQQRVKAGWPCH
GPVPWPNNNTYQKSMKTFMEELGLAGERLLKLTALGFELPINTFTDLTRDGWHHMRV
LRFPPQTSTLSRGIGAHTDYGLLVIAAQDDVGGLYIRPPVEGEKRNRNWLPGESSAG
MFEHDEPWTFVTPTPGVWTVFPGDILQFMTGGQLLSTPHKVKLNTRERFACAYFHEP
NFEASAYPLFEPSANERIHYGHEFTNMFMRCCYPDRITTQRINKENRLAHLEDLKKYS
10 DTRATGSGATNFSLLKQAGDVEENPGPMTESITSNGTLVASDTRRRRVWAIVSASSGN
LVEWFDFYVYSFCSLYFAHIFPSPGNNTTQLLQTAGVFAAGFLMRPIGGWLFGRIDR
RGRKTSMLISVCMCFGSLIACLPGYDAIGTWAPALLLLARLFQGLSVGGEYGTSA
TYMSEIALEGRKGFYASFQYVTLIGGQLLAILVVILQQILTDSQLHEWGWRIPFAMG
AALAIVALWLRRLDETSQKEVRALKEAGSFKGLWRNRKAFLMVLGFTAGGSLSFY
15 TFTTYMQKYL VNTTGMHANVASVIMTAALFVFM LIQPLIGALSDKIGRRTSMLIFGG
MSALCTVPILTALQHVSSPYAAAFALVMLAMVIVSFYTSISGILKAEMFPAQVRALGV
GLSYAVANALFGGSAEYVALSLKSWGSETTFFWYVTIMGALAFIVSLMLHRKGKGI
RL

20 SEQ ID NO. 7-

ATGATTCATGCCCCCTCCCGCTGGGGCGTGTTTCCCAGTCTGGGTCTCTGCTCCCC
CGATGTGGTGTGGAACGAACACCCAGCCTGTACATGGATAAGGAAGAGACCAG
TATGACCAATCTGCAAACCTTTGAACTGCCACCGAGGTGACCGGTTGCGCCGCC
GATATTAGCCTCGGTGCGGCCCTGATTCAAGCCTGGCAAAAGGATGGCATCTTCC
25 AAATCAAGACCGATTCCGAACAAGATCGCAAGACCCAAGAGGCCATGGCCGCCA
GCAAACAATTTTGCAAGGAACCCCTGACCTTTAAATCCAGCTGCGTGAGCGATCT
CACCTACAGTGGCTATGTGGCCAGTGGTGAAGAGGTGACCGCCGGCAAGCCCGA
TTTTCCCGAGATTTTACCCTGTGCAAGGATCTGAGTGTGGGTGATCAACGCGTG
AAAGCCGGTTGGCCCTGCCATGGTCCCGTGCCCTGGCCCAACAATACCTATCAAA
30 AATCCATGAAGACCTTTATGGAAGAACTCGGTCTGGCCGGTGAACGCCTGCTCA
AACTGACCGCCCTCGGCTTTGAGCTGCCCATTAAACACCTTTACCGATCTCACCCG
CGATGGTTGGCACCACATGCGCGTGCTGCGCTTTCCTCCCCAAACCAGCACCTG
AGCCGCGGTATTGGTGCCACACCGATTACGGCCTGCTCGTGATTGCCGCCAAG
ATGATGTGGGCGGTCTGTATATTCGCCCTCCCGTGGAAGGCGAGAAACGCAACC
35 GCAATTGGCTCCCCGGCGAAAGTTCCGCCGGCATGTTTGAACACGATGAACCCTG
GACCTTTGTGACGCCCACGCCCGGCGTGTTGACCGTGTTTCCCGGTGATATTCTG
CAATTTATGACCGGCGGTCAACTGCTCTCCACGCCCCACAAAGTGAAGCTCAACA
CCCGCGAACGCTTTGCCTGCGCCTACTTTCACGAACCAATTTTGAGGCCAGTGC
CTATCCCCTGTTTGAACCCTCCGCCAACGAGCGCATTCACTACGGCGAGCACTTT
40 ACCAATATGTTTATGCGCTGCTATCCCGATCGCATTACCACCCAACGCATTAACA
AGGAAAATCGCCTGGCCACCTCGAGGATCTGAAAAAGTATAGTGATACCCGCG
CCACCGGTAGTGGTGCCACCAACTTTAGCCTGCTCAAACAAGCCGGCGATGTGG
AAGAGAACCCCGGTCCCATGACCGAAAGTATTACCAGCAATGGCACCTGGTGG
CCAGTGATACCCGTCGCCGCGTGTTGGGCCATTGTGAGTGCCAGCAGTGGTAACCT

GGTGGAGTGGTTTGATTTTTACGTGTATAGCTTTTGCAGTCTCTACTTTGCCCACA
TTTTCTTTCCAGTGGAATACCACCACCAACTGCTGCAAACCGCCGGCGTGTT
TGCCGCCGGTTTTCTGATGCGCCCCATTGGCGGTTGGCTCTTTGGCCGCATTGCCG
ATCGTCGCGGTCGCAAGACCAGCATGCTGATTAGCGTGTGCATGATGTGCTTTGG
5 CTCCCTGATTATTGCCTGCCTCCCCGGCTATGATGCCATTGGCACCTGGGCCCCC
GCCCTGCTCCTGCTGGCCCGCCTCTTTCAAGGCCTGAGCGTGGGCGGTGAATACG
GCACCAGCGCCACCTATATGAGTGAAATTGCCCTGGAGGGCCGCAAAGGTTTTT
ACGCCAGTTTTCAATATGTGACCCTGATTGGCGGTCAACTGCTCGCCATTCTCGT
GGTGGTGATTCTCCAACAAATTCTGACCGATTCCCAACTGCACGAATGGGGCTGG
10 CGCATTCCCTTTGCCATGGGTGCCGCCCTGGCCATTGTGGCCCTGTGGCTCCGTC
GCCAACTCGATGAAACCAGCCAAAAAGAGGTGCGCGCCCTGAAAGAAGCCGGC
AGTTTTAAAGGTCTCTGGCGCAACCGCAAGGCCTTTCTCATGGTGCTGGGCTTTA
CCGCCGGCGGTAGTCTGTCCTTTTACACCTTTACCACCTACATGAAAAATATCT
CGTGAACACCACCGGCATGCACGCCAATGTGGCCAGCGTGATTATGACCGCCGC
15 CCTGTTTGTGTTTATGCTCATTCAACCCCTGATTGGCGCCCTCAGCGATAAGATTG
GTCGTCGCACCAGTATGCTGATTTTTGGCGGTATGAGTGCCCTCTGCACCGTGCC
CATTCTCACCGCCCTGCAACACGTGTCCAGCCCCTACGCCGCCTTTGCCCTCGTG
ATGCTGGCCATGGTGATTGTGTCCTTTTATACCAGCATTAGTGGCATTCTGAAGG
CCGAAATGTTTCCCGCCCAAGTGCGCGCCCTGGGCGTGGGTCTCAGTTACGCCGT
20 GGCCAATGCCCTGTTTGGCGGTTCCGCCGAATATGTGGCCCTGTCCCTCAAAGC
TGGGGCAGTGAGACCACCTTTTTCTGGTACGTGACCATTATGGGTGCCCTGGCCT
TTATTGTGAGCCTGATGCTCCACCGCAAAGGCAAGGGTATTCGCCTCTAG

SEQ ID NO. 8: - CATATG

25

SEQ ID NO. 9: - CACCACCACCATCATCATTAAATGAAAGCTT

SEQ ID NO. 10: - MHHHHHHENLYFQGKL

30 SEQ ID NO. 11: - GATNFSLLKQAGDVEENPGP

SEQ ID NO. 12: - AAGCTT

SEQ ID NO. 13: - GGTACC

35

SEQ ID NO. 14: -

ATGACTGATTTTTTACGCGATGACATCAGGTTCTCGGTCAAATCCTCGGTGAGG
TAATTGCGGAACAAGAAGGCCAGGAGGTTTATGAACTGGTCGAACAAGCGCGCC
TGACTTCTTTTGATATCGCCAAGGGCAACGCCGAAATGGATAGCCTGGTTCAGGT

TTTCGACGGCATTACTCCAGCCAAGGCAACACCGATTGCTCGCGCATTTTCCCAC
TTCGCTCTGCTGGCTAACCTGGCGGAAGACCTCTACGATGAAGAGCTTCGTGAAC
AGGCTCTCGATGCAGGCGACACCCCTCCGGACAGCACTCTTGATGCCACCTGGCT
GAAACTCAATGAGGGCAATGTTGGCGCAGAAGCTGTGGCCGATGTGCTGCGCAA
5 TGCTGAGGTGGCGCCGGTTCTGACTGCGCACCCAACTGAGACTCGCCGCCGCACT
GTTTTTGATGCGCAAAAGTGGATCACCACCCACATGCGTGAACGCCACGCTTTGC
AGTCTGCGGAGCCTACCGCTCGTACGCAAAGCAAGTTGGATGAGATCGAGAAGA
ACATCCGCCGTGCGATCACCATTTTGTGGCAGACCGCGTTGATTTCGTGTGGCCCG
CCCACGTATCGAGGACGAGATCGAAGTAGGGCTGCGCTACTACAAGCTGAGCCT
10 TTTGGAAGAGATTCCACGTATCAACCGTGATGTGGCTGTTGAGCTTCGTGAGCGT
TTCGGCGAGGGTGTTCTTTGAAGCCCGTGGTCAAGCCAGGTTCTGGATTGGTG
GAGACCACGACGGTAACCCTTATGTCACCGCGGAAACAGTTGAGTATTCCACTC
ACCGCGCTGCGGAAACCGTGCTCAAGTACTATGCACGCCAGCTGCATTCCCTCGA
GCATGAGCTCAGCCTGTGCGACCGCATGAATAAGGTCACCCCGCAGCTGCTTGC
15 GCTGGCAGATGCAGGGCACAACGACGTGCCAAGCCGCGTGGATGAGCCTTATCG
ACGCGCCGTCCATGGCGTTCGCGGACGTATCCTCGCGACGACGGCCGAGCTGAT
CGGCGAGGACGCCGTTGAGGGCGTGTGGTTCAAGGTCTTTACTCCATACGCATCT
CCGGAAGAATTCTTAAACGATGCGTTGACCATTGATCATTCTCTGCGTGAATCCA
AGGACGTTCTCATTGCCGATGATCGTTTGTCTGTGCTGATTTCTGCCATCGAGAG
20 CTTTGGATTCAACCTTTACGCACTGGATCTGCGCCAAACTCCGAAAGCTACGAG
GACGTCCTCACCAGCTTTTCGAACGCGCCCAAGTCACCGCAAACCTACCGCGAG
CTGTCTGAAGCAGAGAACTTGAGGTGCTGCTGAAGGAACTGCGCAGCCCTCGT
CCGCTGATCCCGCACGGTTCAGATGAATACAGCGAGGTCACCGACCGCGAGCTC
GGCATCTTCCGCACCGCGTCGGAGGCTGTTAAGAAATTCGGGCCACGGATGGTG
25 CCTCACTGCATCATCTCCATGGCATCATCGGTACCGATGTGCTCGAGCCGATGG
TGTTGCTCAAGGAATTCGGAATCATCGCAGCCAACGGCGACAACCCACGCGGCA
CCGTCGATGTCATCCCACTGTTTCGAAACCATCGAAGATCTCCAGGCCGGCGCCGG
AATCCTCGACGAACTGTGGAATAATTGATCTCTACCGCAACTACCTCCTGCAGCGC
GACAACGTCCAGGAAGTCATGCTCGGTTACTCCGATTCCAACAAGGATGGCGGA
30 TATTTCTCCGCAAACCTGGGCGCTTTACGACGCGGAACTGCAGCTCGTCAACTAT
GCCGATCAGCCGGGGTCAACGTTTCGCTGTTCCACGGCCGTGGTGGCACCGTCGG
CCGCGGTGGCGGACCTTCCTACGACGCGATTCTTGCCCAGCCCAGGGGGGCTGTC
CAAGGTTCCGTGCGCATCACCGAGCAGGGCGAGATCATCTCCGCTAAGTACGGC
AACCCCGAAACCGCGCGCCGAAACCTCGAAGCCCTGGTCTCAGCCACGCTTGAG
35 GCATCGCTTCTCGACGTCTCCGAACCTACCGATCACCAACGCGCGTACGACATCA
TGAGTGAGATCTCTGAGCTCAGCTTGAAGAAGTACGCCTCCTTGGTGCACGAGG
ATCAAGGCTTCATCGATTACTTCACCCAGTCCACGCCGCTGCAGGAGATTGGATC
CCTCAACATCGGATCCAGGCCTTCCTCACGCAAGCAGACCTCCTCGGTGGAAGAT
TTGCGAGCCATCCCATGGGTGCTCAGCTGGTCAAGTCTCGTGTGCTGCTGCCAG
40 GCTGGTTTTGGTGTGCGGAACCGCATTAGAGCAGTGGATTGGCGAAGGGGAGCAGG
CCACCCAACGCATTGCCGAGCTGCAAACACTCAATGAGTCCTGGCCATTTTTACC
CTCAGTGTTGGATAACATGGCTCAGGTGATGTCCAAGGCAGAGCTGCGTTTGGCA
AAGCTCTACGCAGACCTGATCCCAGATACGGAAGTAGCCGAGCGAGTCTATTCC
GTCATCCGCGAGGAGTACTTCCTGACCAAGAAGATGTTCTGCGTAATCACCGGCT
45 CTGATGATCTGCTTGATGACAACCCACTTCTCGCACGCTCTGTCCAGCGCCGATA
CCCCTACCTGCTTCCACTCAACGTGATCCAGGTAGAGATGATGCGACGCTACCGA

AAAGGCGACCAAAGCGAGCAAGTGTCCCGCAACATTCAGCTGACCATGAACGGT
CTTTCCACTGCGGTGCGCAACTCCGGC

SEQ ID NO. 15 –

5 MTDFLRDDIRFLGQILGEVIAEQEGQEVYELVEQARLTSFDIAKGNAEMDSLQVQFD
GITPAKATPIARAFSHFALLANLAEDLYDEELREQALDAGDTPPDSTLDATWLKLE
GNVGAEAVADVLRNAEVAPVLTAHPTETRRRTVFDAQKWITTHMRERHALQSAEP
TARTQSKLDEIEKNIRRRITILWQTALIRVARPRIEDEIEVGLRYYKLSLLEEIPRINRDV
10 AVELRERFGEVPLKPVVKPGSWIGGDHGNPYVTAETVEYSTHRAAETVLKYYAR
QLHSLEHELSLSDRMNKVTPQLLALADAGHNDVPSRVDEPYRRVHGVGRILATT
AELIGEDAVEGVWFKVFTPYASPEEFLNDALTIDHSLRESKDVLIADDRLSVLISAIES
FGFNLYALDLRQNSESYEDVLTelferaQVTANYRELSEAEKLEVLLKELRSPRPLIP
HGSDEYSEVTDRELGIFRTASEAVKKFGPRMVPHCIISSMASSVTDVLEPMVLLKEFGL
15 IAANGDNPRGTVDVIPLFETIEDLQAGAGILDELWKIDLYRNYLLQRDNVQEVMLGY
SDSNKDGGYFSANWALYDAELQLVELCRSAGVNVRLFHGRGGTVGRGGGPSYDAI
LAQPRGAVQGSVRITEQGEIISAKYGNPETARRNLEALVSATLEASLLDVSELTDHQR
AYDIMSEISELSLKKYASLVHEDQGFIDYFTQSTPLQEIGSLNIGSRPSSRKQTSSVEDL
RAIPWVLSWSQSRVMLPGWFGVGTALQWIGEGEQATQRIAEQLTLNESWPFLPSV
LDNMAQVMSKAELRLAKLYADLIPDTEVAERVYSVIREEYFLTKKMFCVITGSDDL
20 DDNPLLARSVQRRYPYLLPLNVIQVEMMRRYRKGDQSEQVSRNIQLTMNGLSTAVR
NSG

SEQ ID NO. 16 –

ATGTTTGAAAGGGATATCGTGGCTACTGATAACAACAAGGCTGTCTCTGCACTACC
25 CCGGTGGCGAGTTCGAAATGGACATCATCGAGGCTTCTGAGGGTAACAACGGTG
TTGTCTGGGCAAGATGCTGTCTGAGACTGGACTGATCACTTTTGACCCAGGTTA
TGTGAGCACTGGCTCCACCGAGTCGAAGATCACCTACATCGATGGCGATGCGGG
AATCCTGCGTTACCGCGGCTATGACATCGCTGATCTGGCTGAGAATGCCACCTTC
AACGAGGTTTCTTACCTACTTATCAACGGTGAGCTACCAACCCAGATGAGCTTC
30 ACAAGTTTAACGACGAGATTCCGCCACCACACCTTCTGGACGAGGACTTCAAGTC
CCAGTTCAACGTGTTCCACGCGACGCTCACCCAATGGCAACCTTGGCTTCCTCG
GTAAACATTTTGTCTACCTACTACCAGGACCAGCTGAACCCACTCGATGAGGCAC
AGCTTGATAAGGCAACCGTTTCGCCTCATGGCAAAGGTTCCAATGCTGGCTGCGTA
CGCACACCGCGCACGCAAGGGTGCTCCTTACATGTACCCAGACAACCTCCCTCAAT
35 GCGCGTGAGAACTTCTGCGCATGATGTTTCGGTTACCCAACCGAGCCATACGAG
ATCGACCCAATCATGGTCAAGGCTCTGGACAAGCTGCTCATCCTGCACGCTGACC
ACGAGCAGAACTGCTCCACCTCCACCGTTCGTATGATCGGTTCCGCACAGGCCAA
CATGTTTGTCTCCATCGCTGGTGGCATCAACGCTCTGTCCGGCCCACTGCACGGT
GGCGCAAACCGAGGCTGTTCTGGAGATGCTCGAAGACATCAAGAGCAACACGGT
40 GCGACGCAACCGAGTTCATGAACAAGGTCAAGAACAAGGAAGACGGCGTCCG
CCTCATGGGCTTCGGACACCGCGTTTACAAGAAGTACGATCCACGTGCAGCAATC
GTCAAGGAGACCGCACACGAGATCCTCGAGCACCTCGGTGGCGACGATCTTCTG
GATCTGGCAATCAAGCTGGAAGAAATTGCACTGGCTGATGATTACTTCATCTCCC

GCAAGCTCTACCCGAACGTAGACTTCTACACCGGCCTGATCTACCGCGCAATGGG
CTTCCCAACTGACTTCTTCACCGTATTGTTTCGCAATCGGTCGTCTGCCAGGATGG
ATCGCTCACTACCGCGAGCAGCTCGGTGCAGCAGGCAACAAGATCAACCGCCCA
CGCCAGGTCTACACCGGCAACGAATCCCGCAAGTTGGTTCTCGCGAGGAGCGC
5 TAA

SEQ ID NO. 17 –

MFERDIVATDNNKAVLHYPGGEFEMDIIIEASEGNNGVVLGKMLSETGLITFDPGYVS
TGSTESKITYIDGDAGILRYRGYDIADLAENATFNEVSYLLINGELPTDELHKFNDEI
10 RHHTLLDEDFKSQFNVPFPRDAHPMATLASSVNILSTYYQDQLNPLDEAQLDKATVRL
MAKVPMLAAYAHRARKGAPYMYPDNSLNARENFLRMMFGYPTEPYEIDPIMVKAL
DKLLILHADHEQNCSTSTVRMIGSAQANMFVSIAGGINALSGPLHGGANQAVLEMLE
DIKSNHGGDATEFMNKVKNKEDGVRLMGFGHRVYKNYDPRAAIVKETAHEILEHL
GGDDLLDLAIKLEEIALADDYFISRKLYPNVDFYTGILIYRAMGFPTDFFTVLFAIGRLP
15 GWIAHYREQLGAAGNKINRPRQVYTGNESRKLVPREER

SEQ ID NO. 18 –

ATGACTGATTTTTTACGCGATGACATCAGGTTCTCGGTCAAATCCTCGGTGAGG
TAATTGCGGAACAAGAAGGCCAGGAGGTTTATGAACTGGTCGAACAAGCGCGCC
20 TGACTTCTTTTGATATCGCCAAGGGCAACGCCGAAATGGATAGCCTGGTTCAGGT
TTTCGACGGCATTACTCCAGCCAAGGCAACACCGATTGCTCGCGCATTTTCCCAC
TTCGCTCTGCTGGCTAACCTGGCGGAAGACCTCTACGATGAAGAGCTTCGTGAAC
AGGCTCTCGATGCAGGCGACACCCCTCCGGACAGCACTCTTGATGCCACCTGGCT
GAAACTCAATGAGGGCAATGTTGGCGCAGAAGCTGTGGCCGATGTGCTGCGCAA
25 TGCTGAGGTGGCGCCGGTTCTGACTGCGCACCCAACTGAGACTCGCCGCCGCACT
GTTTTTGATGCGCAAAAGTGGATCACACCCACATGCGTGAACGCCACGCTTTCG
AGTCTGCGGAGCCTACCGCTCGTACGCAAAGCAAGTTGGATGAGATCGAGAAGA
ACATCCGCCGTGCGATCACCATTTTTGTGGCAGACCGCGTTGATTCTGTGGCCCG
CCCACGTATCGAGGACGAGATCGAAGTAGGGCTGCGCTACTACAAGCTGAGCCT
30 TTTGGAAGAGATTCCACGTATCAACCGTGATGTGGCTGTTGAGCTTCGTGAGCGT
TTCGGCGAGGGTGTTCTTTGAAGCCCGTGGTCAAGCCAGGTTCTGGATTGGTG
GAGACCACGACGGTAACCCTTATGTACCGCGGAAACAGTTGAGTATTCCACTC
ACCGCGCTGCGGAAACCGTGCTCAAGTACTATGCACGCCAGCTGCATTCCCTCGA
GCATGAGCTCAGCCTGTGCGACCGCATGAATAAGGTCACCCCGCAGCTGCTTGC
35 GCTGGCAGATGCAGGGCACAACGACGTGCCAAGCCGCGTGGATGAGCCTTATCG
ACGCGCCGTCCATGGCGTTCGCGGACGTATCCTCGCGACGACGGCCGAGCTGAT
CGGCGAGGACGCCGTTGAGGGCGTGTGGTTCAAGGTCTTTACTCCATACGCATCT
CCGGAAGAATTCTTAAACGATGCGTTGACCATTGATCATTCTCTGCGTGAATCCA
AGGACGTTCTCATTGCCGATGATCGTTTGTCTGTGCTGATTTCTGCCATCGAGAG
40 CTTTGGATTCAACCTTTACGCACTGGATCTGCGCCAAAACCTCCGAAAGCTACGAG
GACGTCCTCACCGAGCTTTTCGAACGCGCCCAAGTCACCGCAAACCTACCGCGAG
CTGTCTGAAGCAGAGAACTTGAGGTGCTGCTGAAGGAACTGCGCAGCCCTCGT
CCGCTGATCCCGCACGGTTCAGATGAATACAGCGAGGTCACCGACCGCGAGCTC

GGCATCTTCCGCACCGCGTCGGAGGCTGTTAAGAAATTCGGGCCACGGATGGTG
CCTCACTGCATCATCTCCATGGCATCATCGGTACCGATGTGCTCGAGCCGATGG
TGTTGCTCAAGGAATTCGGACTCATCGCAGCCAACGGCGACAACCCACGCGGCA
CCGTCGATGTCATCCCACTGTTTCGAAACCATCGAAGATCTCCAGGCCGGCGCCGG
5 AATCCTCGACGAACTGTGAAAAATTGATCTCTACCGCAACTACCTCCTGCAGCGC
GACAAACGTCCAGGAAGTCATGCTCGGTTACTCCGATTCCAACAAGGATGGCGGA
TATTTCTCCGCAAACTGGGCGCTTTACGACGCGGAACTGCAGCTCGTCGAACTAT
GCCGATCAGCCGGGGTCAACGTTTCGCTGTTCACGGCCGTGGTGGCACCGTCGG
CCGCGGTGGCGGACCTTCCTACGACGCGATTCTTGCCCAGCCCAGGGGGGCTGTC
10 CAAGGTTCCGTGCGCATCACCGAGCAGGGCGAGATCATCTCCGCTAAGTACGGC
AACCCCGAAACCGCGCGCCGAAACCTCGAAGCCCTGGTCTCAGCCACGCTTGAG
GCATCGTTTCTCGACGTCTCCGAACTCACCGATCACCAACGCGCGTACGACATCA
TGAGTGAGATCTCTGAGCTCAGCTTGAAGAAGTACGCCTCCTTGGTGCACGAGG
ATCAAGGCTTCATCGATTACTTCACCCAGTCCACGCCGCTGCAGGAGATTGGATC
15 CCTCAACATCGGATCCAGGCCTTCCTCACGCAAGCAGACCTCCTCGGTGGAAGAT
TTGCGAGCCATCCCATGGGTGCTCAGCTGGTCAAGTCTCGTGTGTCATGCTGCCAG
GCTGGTTTGGTGTGCGGAACCGCATTAGAGCAGTGGATTGGCGAAGGGGAGCAGG
CCACCCAACGCATTGCCGAGCTGCAAACACTCAATGAGTCCTGGCCATTTTTACC
CTCAGTGTTGGATAACATGGCTCAGGTGATGTCCAAGGCAGAGCTGCGTTTGGCA
20 AAGCTCTACGCAGACCTGATCCCAGATACGGAAGTAGCCGAGCGAGTCTATTCC
GTCATCCGCGAGGAGTACTTCCTGACCAAGAAGATGTTCTGCGTAATCACCGGCT
CTGATGATCTGCTTGATGACAACCCACTTCTCGCACGCTCTGTCCAGCGCCGATA
CCCCTACCTGCTTCCACTCAACGTGATCCAGGTAGAGATGATGCGACGCTACCGA
AAAGGCGACCAAAGCGAGCAAGTGTCCCGCAACATTCAGCTGACCATGAACGGT
25 CTTTCCACTGCGGTGCGCAACTCCGGCGCCACCAACTCAACTCCGGCGCCACCAA
CTTTAGCCTGCTCAAACAAGCCGGCGATGTGGAAGAGAACCCCGGTCCCATGTTT
GAAAGGGATATCGTGGCTACTGATAACAACAAGGCTGTCTTGCCTGCACTACCCCGGT
GGCGAGTTCGAAATGGACATCATCGAGGCTTCTGAGGGTAACAACGGTGTTGTC
CTGGGCAAGATGCTGTCTGAGACTGGACTGATCACTTTTGACCCAGGTTATGTGA
30 GCACTGGCTCCACCGAGTCGAAGATCACCTACATCGATGGCGATGCGGGAATCC
TGCGTTACCGCGGCTATGACATCGCTGATCTGGCTGAGAATGCCACCTTCAACGA
GGTTTCTTACCTACTTATCAACGGTGAGCTACCAACCCAGATGAGCTTCAACGA
TTTAACGACGAGATTGCGCACCAACCCCTTCTGGACGAGGACTTCAAGTCCCAGT
TCAACGTGTTCCACGCGACGCTACCCAATGGCAACCTTGGCTTCCTCGGTTAA
35 CATTTTGTCTACCTACTACCAGGACCAGCTGAACCCACTCGATGAGGCACAGCTT
GATAAGGCAACCGTTTCGCTCATGGCAAAGGTTCCAATGCTGGCTGCGTACGCA
CACCGCGCACGCAAGGGTGCTCCTTACATGTACCCAGACAACCTCCCTCAATGCGC
GTGAGAACTTCCTGCGCATGATGTTTCGGTTACCCAACCGAGCCATACGAGATCGA
CCCAATCATGGTCAAGGCTCTGGACAAGCTGCTCATCCTGCACGCTGACCACGAG
40 CAGAACTGCTCCACCTCCACCGTTTCGTATGATCGGTTCCGCACAGGCCAACATGT
TTGTCTCCATCGCTGGTGGCATCAACGCTCTGTCCGGCCCACTGCACGGTGGCGC
AAACCAGGCTGTTCTGGAGATGCTCGAAGACATCAAGAGCAACCACGGTGGCGA
CGCAACCGAGTTCATGAACAAGGTCAAGAACAAGGAAGACGGCGTCCGCTCAT
GGGCTTCGGACACCGCGTTTACAAGAACTACGATCCACGTGCAGCAATCGTCAA
45 GGAGACCGCACACGAGATCCTCGAGCACCTCGGTGGCGACGATCTTCTGGATCT
GGCAATCAAGCTGGAAGAAATTGCACTGGCTGATGATTACTTCATCTCCCGCAAG

CTCTACCCGAACGTAGACTTCTACACCGGCCTGATCTACCGCGCAATGGGCTTCC
CAACTGACTTCTTCACCGTATTGTTTCGCAATCGGTCGTCTGCCAGGATGGATCGC
TCACTACCGCGAGCAGCTCGGTGCAGCAGGCAACAAGATCAACCGCCACGCCA
GGTCTACACCGGCAACGAATCCCGCAAGTTGGTTCCTCGCGAGGAGCGCTAA

5

SEQ ID NO. 19 -

ATGTACGAGAAGATTCAACCCCCTAGCGAAGGCAGCAAAATTCGCTTTGAAGCC
GGCAAGCCGATCGTTCCCGACAACCCGATCATTCCTTCATTTCGTGGTGACGGCG
CTGGCGTTGATATCTGGCCCGCAACTGAGCGCGTTCTCGATGCCGCTGTCGCTAA
10 AGCCTATGGCGGTCAGCGCAAAATCACTTGGTTCAAAGTCTACGCGGGTGATGA
AGCCTGCGACCTCTACGGCACCTACCAATATCTGCCTGAAGATACGCTGACAGCG
ATCCGCGAGTACGGCGTGGCAATCAAAGGCCCGCTGACGACGCCGATCGGTGGT
GGCATTTCGATCGCTGAACGTGGCGCTACGGCAAATCTTCGATCTCTATGCCTGCG
TCCGCCCCCTGTCGCTACTACACCGGCACACCCTCGCCCCACCGCACGCCCGAACA
15 ACTCGATGTGGTGGTCTACCGCGAAAACACCGAGGATATCTACCTCGGCATCGA
ATGGAAGCAAGGTGATCCCACCGGCGATCGCCTGATCAAGCTGCTGAACGAGGA
CTTCATTCCCAACAGCCCCAGCTTGGGTAAAAGCAAATCCGTTTGGATTCCGGC
ATTGGTATTAAGCCGATCAGTAAAACGGGTAGCCAGCGTCTGATTTCGTCTGTCGA
TCGAGCATGCCCTACGCCTCGAAGGCCGCAAGCGACATGTCACCCTTGTCCACAA
20 GGGCAACATCATGAAGTTCACGGAAGGTGCTTTCCGGGACTGGGGCTATGAACT
GGCCACGACTGAGTTCCGAACCGACTGTGTGACTGAACGGGAGAGCTGGATTCT
TGCCAACCAAGAAAGCAAGCCGGATCTCAGCTTGGAAGACAATGCGCGGCTCAT
CGAACCTGGCTACGACGCGATGACGCCCGAAAAGCAGGCAGCAGTGGTGGCTGA
AGTGAAGCTGTGCTCGACAGCATCGGCGCCACCCACGGCAACGGTCAGTGGAA
25 GTCTAAGGTGCTGGTTGACGATCGCATTGCTGACAGCATCTTCCAGCAGATTCAA
ACCCGCCCCGGGTGAATACTCGGTGCTGGCGACGATGAACCTCAATGGCGACTAC
ATCTCTGATGCAGCGGCGGCGGTGGTCGGTGGCCTGGGCATGGCCCCCGGTGCC
AACATTGGCGACGAAGCGGCGATCTTTGAAGCGACCCACGGCGCAGCGCCCAAG
CACGCTGGCCTCGATCGCATTAAACCCGGCTCGGTTCATCCTCTCCGGCGTGATGA
30 TGCTGGAGTACCTAGGCTGGCAAGAGGCTGCTGACTTGATCACCAAGGGCATCA
GCCAAGCGATCGCTAACCGTGAGGTCACCTACGATCTGGCTCGGTTGATGGAAC
CGGCGGTTGATCAACCACTCAAGTGCTCGGAATTTGCCGAAGCCATCGTCAAGC
ATTCGACGATTAG

35 SEQ ID NO. 20 -

MYEKIQPPSEGSKIRFEAGKPIVPDNPIIPFIRGDGTGVDIWPATERVLDAAVAKAYGG
QRKITWFKVYAGDEACDLYGTYQYLPEDTLTAIREYGVAIKGPLTTPIGGGIRSLNVA
LRQIFDLYACVRPCRYTGTSPHRTPEQLDVVVYRENTEDIYLGIEWKQGDPTGDR
LIKLLNEDFIPNSPSLGKKQIRLDSGIGIKPISKTGSQLIRRAIEHALRLEGRKRHVTLV
40 HKGNIMKFTEGAFRDWGYELATTEFRDTCVTERESWILANQESKPDLSLEDNARLIE
PGYDAMTPEKQAAVVAEVKAVLDSIGATHGNGQWWSKVLVDDRIADSIFQQIQRTP
GEYSVLATMNLNGDYISDAAA VVGGLGMAPGANIGDEAAIFEATHGTAPKHAGL

DRINPGSVILSGVMMLEYLGWQEAADLITKGISQAIANREVTYDLARLMEPAVDQPL
KCSEFAEAIVKHFDD

SEQ ID NO. 21 –

5 GTGAAAAACGTGCTGGCGATCATTCTCGGTGGAGGCGCAGGCAGTCGTCTCTATC
CACTAACCAAACAGCGCGCCAAACCAGCGGTCCCCCTGGCGGGCAAATACCGCT
TGATCGATATTCCTGTCAGCAATTGCATCAACGCTGACATCAACAAAATCTATGT
GCTGAC
GCAGTTTAACTCTGCCTCGCTCAACCGCCACCTCAGTCAGACCTACAACCTCTCC
10 AGCGGCTTTGGCAATGGCTTTGTTGAGGTGCTAGCAGCTCAGATTACGCCGGAGA
ACCCCAACTGGTTCCAAGGCACCGCCGATGCGGTTCCGCCAGTATCTCTGGCTAAT
CAAAGAGTGGGATGTGGATGAGTACCTGATCCTGTCGGGGGATCATCTCTACCG
CATGGACTATAGCCAGTTCATTACGCGGCACCGAGACACCAATGCCGACATCAC
ACTCTCGGTCTTGCCGATCGATGAAAAGCGCGCCTCTGATTTTGGCCTGATGAAG
15 CTAGATGGCAGCGGCCGGGTGGTCGAGTTCAGCGAAAAGCCCAAAGGGGATGA
ACTCAGGGCGATGCAAGTCGATACCACGATCCTCGGGCTTGACCCTGTCGCTGCT
GCTGCCCAGCCCTTCATTGCCTCGATGGGCATCTACGTCTTCAAGCGGGATGTTT
TGATCGATTTGCTCAGCCATCATCCCGAGCAAACCGACTTTGGCAAGGAAGTGAT
TCCCGCTGCAGCCACCCGCTACAACACCCAAGCCTTTCTGTTCAACGACTACTGG
20 GAAGACATCGGCACGATCGCCTCATTCTACGAGGCCAATCTGGCGCTGACTCAG
CAACCTAGCCCACCCTTCAGCTTCTACGACGAGCAGGCGCCGATTTACACCCGCG
CTCGCTACCTGCCGCCAACCAAGCTGCTCGATTGCCAGGTGACCCAGTCGATCAT
TGGCGAGGGCTGCATTCTCAAGCAATGCACCGTTCAGAATTCCGTCTTAGGGATT
CGCTCCCGCATTGAGGCCGACTGCGTGATCCAGGACGCCTTGTTGATGGGCGCTG
25 ACTTCTACGAAACCTCGGAGCTACGGCACCAGAATCGGGCCAATGGCAAAGTGC
CGATGGGAATCGGCAGTGGCAGCACCATCCGTGCGGCCATCGTCGACAAAAATG
CCCACATTGGCCAGAACGTTTCAGATCGTCAACAAAGA
CCATGTGGAAGAGGCCGATCGCGAAGATCTGGGCTTTATGATCCGCAGCGGCAT
TGTCGTTGTGGTCAAAGGGGGCGGTTATTCCCGACAACACGGTGATCTAA
30

SEQ ID NO. 22 –

MKNVLAIILGGGAGSRLYPLTKQRAKPAVPLAGKYRLIDIPVSNINADINKIYVL
FNSASLNRHLSQTYNLSSGFGNGFVEVLAAQITPENPNWFQGTADAVRQYLWLIKE
WDVDEYLILSGDHLYRMDYSQFIQRHRDTNADITLSVLPIDEKRASDFGLMKLDGSG
35 RVVEFSEKPKGDELAMQVDTTILGLDPVAAAAQPFIASMGIIYVFKRDVLIDLLSHH
PEQTDFGKEVIPAAATRYNTQAFLFNDYWEDIGTIA SFYEANLALTQQSPPF SFYDE
QAPIYTRARYLPPTKLLDCQVTQSIIGEGCILKQCTVQNSVLGIRSRIEADCVIQDALL
MGADFYETSELRHQNRANGKVPMGIGSGSTIRRAIVDKNAHIGQNVQIVNKDHVEE
ADREDLGFMIRSGIVVVVKGAVIPDNTVI
40

SEQ ID NO. 23 –

ATGGCACTGAATATTCCATTCAGAAATGCGTACTATCGTTTTGCATCCAGTTACT
CATTTCTCTTTTTTATTTCTGGTCGCTGTGGTGGTCGTTATACGCTATTTGGCTGA
AAGGACATCTAGGATTAACAGGGACGGAATTAGGTACACTTTATTCGGTCAACC
AGTTTACCAGCATTCTATTTATGATGTTCTACGGCATCGTTCAGGATAAACTCGGT
5 CTGAAGAAACCGCTCATCTGGTGTATGAGTTTCATTCTGGTCTTGACCGGACCGT
TTATGATTTACGTTTATGAACCGTTACTGCAAAGCAATTTTTCTGTAGGTCTAATT
CTGGGGGCGCTCTTTTTTGGCCTGGGGTATCTGGCGGGATGCGGTTTGCTTGACA
GCTTCACCGAAAAAATGGCGCGAAATTTTCATTTCGAATATGGAACAGCGCGCG
CCTGGGGATCTTTTGGCTATGCTATTGGCGCGTTCTTTGCCGGTATATTTTTTAGT
10 ATCAGTCCCCATATCAACTTCTGGTTGGTCTCGCTATTTGGCGCTGTATTTATGAT
GATCAACATGCGTTTTTAAAGATAAGGATCACCAGTGCATAGCGGCGGATGCGGG
AGGGGTAAAAAAGAGGATTTTATCGCAGTTTCAAGGATCGAAACTTCTGGGTT
TTCGTCATATTTATTGTGGGGACGTGGTCTTTCTATAACATTTTTGATCAACA
CTTTCTGTCTTTTATGCAGGTTTATTCGAATCACACGATGTAGGAACGCGCCTGT
15 ATGGTTATCTCAACTCATTCCAGGTGGTACTCGAAGCGCTGTGCATGGCGATTAT
TCCTTTCTTTGTGAATCGGGTAGGGCCAAAAATGCATTACTTATCGGTGTTGTG
ATTATGGCGTTGCGTATCCTTTCTGCGCGTTGTTTCGTTAACCCTGGATTATTT
ATTAGTGAAGCTGTTACATGCCATTGAGGTTCCACTTTGTGTCATATCCGTCTTCA
AATACAGCGTGGCAAACCTTTGATAAGCGCCTGTCGTCGACGATCTTTCTGATTGG
20 TTTTCAAATTGCCAGTTCGCTTGGGATTGTGCTGCTTTCAACGCCGACTGGGATA
CTCTTTGACCACGCAGGCTACCAGACAGTTTTCTTCGCAATTCGGGTATTGTCTG
CCTGATGTTGCTATTTGGCATTTTCTTCTGAGTAAAAAACGCGAGCAAATAGTT
ATGGAAACGCCTGTACCTTCAGCAATATAG

25 SEQ ID NO> 24 -

MALNIPFRNAYYRFASSYSFLFFISWSLWWSLYAIWLKGHLGLTGTGLTLYSVNQF
TSILFMMFYGIVQDKLGLKKPLIWCMFSILVLTGPFMIYVYEPLLQSNFSVGLILGALF
FGLGYLAGCGLLDSFTEKMARNFHFHEYGTARAWGSFGYAIGAFAFFAGIFFSISPHINFW
LVSLFGAVFMMINMRFKDKDHQCIAADAGGVKKEDFIAVFKDRNFWVFVIFIVGTW
30 SFYNIFDQQLFPVFYAGLFE

SEQ ID NO. 25 -

ATGGTGGCAGCTCAAAATCTCTACATTCTGCACATTCAGACCCATGGTCTGCTGC
GAGGGCAGAACTTGGAACCTGGGGCGAGATGCCGACACCGGGCGGGCAGACCAAG
35 TACGTCTTAGAACTGGCTCAAGCCCAAGCTAAATCCCCACAAGTCCAACAAGTC
GACATCATCACCCGCCAAATCACCGACCCCCGCGTCAGTGTTGGTTACAGTCAGG
CGATCGAACCCTTTGCGCCCAAGGTCTGATTGTCCGTTTGCCTTTTGGCCCCAA
ACGCTACCTCCGTAAAGAGCTGCTTTGGCCCCATCTCTACACCTTTGCGGATGCA
ATTCTCCAATATCTGGCTCAGCAAAAGCGCACCCCGACTTGGAATCAGGCCCACT
40 ATGCTGATGCTGGCCAAGTGGGATCACTGCTGAGTCGCTGGTTGAATGTACCGCT
AATTTTCACAGGGCATTCTCTGGGGCGGATCAAGCTAAAAAAGCTGTTGGAGCA
AGACTGGCCGCTTGAGGAAATTGAAGCGCAATTCAATATTCAACAGCGAATTGA
TGCGGAGGAGATGACGCTCACTCATGCTGACTGGATTGTCGCCAGCACTCAGCA

GGAAGTGGAGGAGCAATACCGCGTTTACGATCGCTACAACCCAGAGCGCAAAC
 TGTCATTCCACCGGGTGTCTGATACCGATCGCTTCAGGTTTCAGCCCTTGGGCGAT
 CGCGGTGTTGTTCTCCAACAGGAACTGAGCCGCTTCTGCGCGACCCAGAAAAAC
 CTCAAATTCTCTGCCTCTGTGCCCCGCACCTCGCAAAAATGTACCGGCGCTGGT
 5 GCGAGCCTTTGGCGAACATCCTTGGCTGCGCAAAAAAGCCAACCTTGTCTTAGTA
 CTGGGCAGCCGCCAAGACATCAACCAGATGGATCGCGGCAGTCGGCAGGTGTTC
 CAAGAGATTTTCCATCTGGTCGATCGCTACGACCTCTACGGCAGCGTCGCCTATC
 CCAAACAGCATCAGGCTGATGATGTGCCGGAGTTCTATCGCCTAGCGGCTCATTC
 CGGCGGGGTATTCTGTCATCCGGCGCTGACCGAACCTTTTGGTTTGACAATTTTG
 10 GAGGCAGGAAGCTGCGGCGTGCCGGTGGTGGCAACCCATGATGGCGGCCCCCAG
 GAAATTCTCAAACACTGTGATTTGCGCACTTTAGTTGATGTCAGCCGACCCGCTA
 ATATCGCGACTGCACTCGCCACCCTGCTGAGCGATCGCGATCTTTGGCAGTGCTA
 TCACCGCAATGGCATTGAAAAAGTTCCCGCCATTACAGCTGGGATCAACATGTC
 AATACCCTGTTTGAGCGCATGGAAACGGTGGCTTTGCCTCGTCGTCGTGTCTGCA
 15 GTTTCGTACGGAGTCGCAACGCTTGATTGATGCCAAACGCCTTGTCGTTAGTGA
 CATCGACAACACACTGTTGGGCGATCGTCAAGGACTCGAGAATTTAATGACCTAT
 CTCGATCAGTATCGCGATCATTTTGCCTTTGGAATTGCCACGGGGCGTCGCCTAG
 ACTCTGCCCCAAGAAGTCTTGAAAGAGTGGGGCGTTTCTTCGCCAAACTTCTGGGT
 GACTTCCGTCGGCAGCGAGATTCATATGGCACCGATGCTGAACCGGATATCAG
 20 CTGGGAAAAGCATATCAATCGCAACTGGAATCCTCAGCGAATTTCGGGCAGTAAT
 GGCACAACTACCTTTTCTTGAAGTGCAGCCGGAAGAGGATCAAACACCTTCAA
 AGTCAGCTTCTTTGTCCGCGATCGCCACGAGACTGTGCTGCGAGAAGTACGGCAA
 CATCTTCGCCGCCATCGCCTGCGGCTGAAGTCAATCTATTCCCATCAGGAGTTTC
 TTGACATTCTGCCGCTAGCTGCCTCGAAAGGGGATGCGATTGCCACCTCTCACT
 25 CCGCTGGCGGATTCTCTTGAGAACATTTTGGTGGCAGGCGATTCTGGTAACGAT
 GAGGAAATGCTCAAGGGCCATAATCTCGGCGTTGTAGTTGGCAATTACTACCG
 GAATTGGAGCCACTGCGCAGCTACGAGCGCGTCTATTTTGCTGAGGGCCACTATG
 CTAATGGCATTCTGGAAGCCTTAAAACACTATCGCTTTTTTGAGGCGATCGCTTA
 A

30

SEQ ID NO. 26 -

MVAAQNLYILHIQTHGLLRGQNLELGRDADTGGQTKYVLELAQAQAKSPQVQQVDI
 ITRQITDPRVSVGYSQAIEPFAPKGRIVRLPFGPKRYLRKELLWPHLYTFADAILQYLA
 QQKRTPTWIQAHYADAGQVGSLLSRWLNPLIFTGHSLGRIKLKLLLEQDWPLEEIE
 35 AQFNIQQRIDAEEMTLTHADWIVASTQQEVEEQYRVYDRYNPERKLVIPPGVDTRF
 RFQPLGDRGVVLQQELSRFLRDPEKPQILCLCRPAPRKNVPALVRAFGHEHPWLKKKA
 NLVLVLGSRQDINQMDRGSRQVFQEIFHLVDRYDLYGSVAYPKQHQADDVPEFYRL
 AAHSGGVFVNPALTEPFGLTILEAGSCGVPVVATHDGGPQEILKHCDFTGLVDVSRP
 ANIATALATLLSDRDLWQCYHRNGIEKVP AHYSWDQHVNTLFERMETVALPRRA
 40 VSFVRSRKRLIDAKRLVVSIDNTLLGDRQGLENLMTYLDQYRDHFAFGIATGRRLD
 SAQEVLEKEWGVPSPNFWVTSVGSEIHYGTDAEPDISWEKHINRNWNPNQIRAVMAQ
 LPFLELQPEEDQTPFKVSFFVRDRHETVLRVQRHLRRHRLRLKSIYSHQEFDLPLA
 ASKGD AIRHLSLRWRIPLNILVAGDSGND EMLKGHN LGVVVGNYSPELEPLRSYE
 RVYFAEGHYANGILEALKHYRFFEAIA

SEQ ID NO. 27 –

ATGCGACAGTTATTGCTAATTTCTGACCTGGACAATACCTGGGTCGGAGATCAAC
AAGCCCTGGAACATTTGCAAGAATATCTAGGCGATCGCCGGGGAAATTTTATTT
5 GGCCTATGCCACGGGGCGTTCCTACCATTCCGCGAGGGAGTTGCAAAAACAGGT
GGGACTCATGGAACCGGACTATTGGCTCACCGCGGTGGGGAGTGAAATTTACCA
TCCAGAAGGCCTGGACCAACATTGGGCTGATTACCTCTCTGAGCATTGGCAACGG
GATATCCTCCAGGCGATCGCCGATGGTTTTGAGGCCTTAAACCCCAATCTCCCT
TGGAACAAAACCCATGGAAAATTAGCTATCATCTCGATCCCCAGGCTTGCCCCAC
10 CGTCATCGACCAATTAACGGAGATGTTGAAGGAAACCGGCATCCCGGTGCAGGT
GATTTTCAGCAGTGGCAAAGATGTGGATTTATTGCCCCAACGGAGTAACAAAGG
TAACGCCACCCAATATCTGCAACAACATTTAGCCATGGAGCCGTCTCAAACCTG
GTGTGTGGGGACTCCGGCAATGATATTGGCTTATTGAAACTTCCGCTCGGGGTG
TCATTGTCCGTAATGCCAGCCGGAATTATTGCACTGGTATGACCAATGGGGGGA
15 TTCTCGTCATTATCGGGCCCAATCGAGCCATGCTGGCGCTATCCTAGAGGCGATC
GCCCATTTGATTTTTTTGAGCTGA

SEQ ID NO. 28 -

MRQLLLISDLNTWVGDQQALEHLQEYLGDRRGNFYLAAYATGRSYHSARELQKQV
20 GLMEPDYWLTA VGSEIYHPEGLDQHWADYLSEHWQRDILQAIADGFEALKPQSPLE
QNPWKISYHLDPA CPTVIDQLTEMLKETGIPVQVIFSSGKDVDLLPQRSNKG NATQ
YLQQHLAMEPSQTLVCGDSGNDIGLFETSARGVIVRNAQPELLHWYDQWGDSRHY
RAQSSHAGAILEAIAHFDFLS

25 SEQ ID NO. 29 –

ATGAGTGATTCCACCGCCCAACTCAGCTACGACCCCAACACGAGCTACCTCGAGC
CCAGTGCGTTGGTCTGTGAGGATGAACGGACTTCTGTGACTCCCGAGACCTTGAA
ACGGGCTTACGAGGCCCATCTCTACTACAGCCAGGGCAAAACCTCAGCGATCGC
CACCTGCGTGATCACTACATGGCACTGGCCTACATGGTCCGCGATCGCCTCCTG
30 CAACGGTGGCTAGCTTCACTGTGACCTATCAACAACAGCACGTCAAAGTGGTCT
GTTACCTGTCCGCTGAGTTTTTTGATGGGTGCGCACCTCGAAAACCTGCCTGATCAA
CCTGCATCTTCACGACCGGCTTCAGCAAGTTTTTGGATGAACTGGGTCTCGATTTT
GAGCAACTGCTAGAGAAAGAGGAAGAACCCGGGCTAGGCAACGGTGGCCTCGG
TCGCTCGCAGCTTGTTTCCTCGACTCCATGGCTACCCTCGACATTCCTGCCGTCG
35 GCTATGGCATTGCTATGAGTTCGGTATCTTCCACCAAGAACTCCACAACGGCTG
GCAGATCGAAATCCCGATAACTGGCTGCGCTTTGGCAACCCTTGGGAGCTAGA
GCGGCGGAACAGGCCGTGGAAATTAAGTTGGGCGGCCACACGGAGGCCTACCA
CGATGCGCGAGGCCGCTACTGCGTCTCTTGATCCCCGATCGCGTCATTGCGGCC
ATCCCCTACGACACCCCCGTACCGGGCTACGACACCAATAACGTCAGCATGTTGC
40 GGCTCTGGAAGGCTGAGGGCACACGGAACCTCAACCTTGAGGCTTTCAACTCAG
GCAACTACGACGATGCGGTTGCCGACAAAATGTCGTCGGAAACGATCTCGAAGG
TGCTCTATCCCAACGACAACACCCCCCAAGGGCGGGAACCTGCGGCTGGAGCAGC

AGTATTTCTTCGTCTCGGCTTCGCTCCAAGACATCATCCGTCGCCACTTGATGAAC
 CACGGTCATCTTGAGCGGCTGCATGAGGCGATCGCAGTCCAGCTTAACGACACC
 CATCCCAGCGTGGCGGTGCCGGAGTTGATGCGCCTCCTGATCGATGAGCATCACC
 TGA CT TGGGACAATGCTTGGACGATTACACAGCGCACCTTCGCCTACACCAACCA
 5 CACGCTGCTACCTGAAGCCTTGGAAACGCTGGCCCGTGGGCATGTTCCAGCGCACT
 TTACCGCGCTTGATGGAGATTATCTACGAAATCAACTGGCGCTTCTTGGCCAATG
 TGCGGGCCTGGTATCCCGGTGACGACACGAGAGCTCGCCGCCTCTCCCTGATTGA
 GGAAGGAGCTGAGCCCCAGGTGCGCATGGCTCACCTCGCCTGCGTGGGCAGTCA
 TGCCATCAACGGTGTGGCAGCCCTGCATACGCAACTGCTCAAGCAAGAAACCCT
 10 GCGAGATTTCTACGAGCTTTGGCCCGAGAAATTCTTCAACATGACCAACGGTGTG
 ACGCCCCGCCGCTGGCTGCTGCAAAGTAATCCTCGCCTAGCCAACCTGATCAGCG
 ATCGCATTGGCAATGACTGGATTCATGATCTCAGGCAACTGCGACGGCTGGAAG
 ACAGCGTGAACGATCGCGAGTTTTTACAGCGCTGGGCAGAGGTCAAGCACCAAA
 ATAAGGTCGATCTGAGCCGCTACATCTACCAGCAGACTCGCATAGAAGTCGATC
 15 CGCACTCTCTCTTTGATGTGCAAGTCAAACGGATTCACGAATACAAACGCCAGCT
 CCTCGCTGTCATGCATATCGTGACGCTCTACAACTGGCTGAAGCACAATCCCCAG
 CTCAACCTGGTGCCGCGCACTTTTATCTTTGCGGGCAAAGCGGCCCGGGTTACT
 ACCGTGCCAAGCAAATCGTCAAACCTGATCAATGCGGTTCGGGAGCATCATCAACC
 ATGATCCCGATGTCCAAGGGCGACTGAAGGTCGTCTTCCTACCTAACTTCAACGT
 20 TTCCTTGGGGCAGCGCATTTATCCAGCTGCCGATTTGTTCGGAGCAAATCTCAACT
 GCAGGGAAAGAAGCGTCCGGCACCGGCAACATGAAGTTCACCATGAATGGCGCG
 CTGACAATCGGAACCTACGATGGTGCCAACATCGAGATCCGCGAGGAAGTCGGC
 CCCGAAAACCTTCTTCCTGTTTGGCCTGCGAGCCGAAGATATCGCCCCGACGCCAAA
 GTCGGGGCTATCGACCTGTGGAGTTCTGGAGCAGCAATGCGGAACCTGCGGGCAG
 25 TCCTCGATCGCTTTAGCAGTGGTCACTTCACACCGGATCAGCCCAACCTCTTCCA
 AGACTTGGTCAGCGATCTGCTGCAGCGGGATGAGTACATGTTGATGGCGGACTA
 TCAGTCCTACATCGACTGCCAGCGCGAAGCTGCTGCTGCCTACCGCGATTCCGAT
 CGCTGGTGGCGGATGTCGCTACTCAACACCGCGAGATCGGGCAAGTTCTCTCCG
 ATCGCACGATCGCTGACTACAGCGAACAGATCTGGGAGGTCAAACCAGTCCCCG
 30 TCAGCCTAAGCACTAGCTTTTAG

SEQ ID NO. 30 -

MSDSTAQLSYDPTTSYLEPSGLVCEDERTSVTPETLKRAYEAHLYYSQGKTSAIATLR
 DHYMALAYMVRDRLLQRWLASLSTYQQQHVKVVCYLSAEFLMGRHLENCLINLHL
 35 HDRVQQVLDELGLDFEQLLEKEEEPGLGNGGLGRLAACFLDSMATLDIPAVGYGIR
 YEFGIFHQELHNGWQIEIPDNWLRFGNPWELERREQAVEIKLGGHTEAYHDARGRY
 CVSWIPDRVIRAIPTYDTPVPGYDTNNVSMRLRLWKAEGTTELNLEAFNSGNYDDAVA
 DKMSSETISKVLYPNDNTPQGRELRLEQQYFFVSASLQDIIRRLMNHGHLERLHEAI
 AVQLNDTHPSVAVPELMRLLIDEHHLTWDNAWTITQRTFAYTNHTLLPEALERWPV
 40 GMFQRTLPRIMEIIEINWRFLANVRAWYPGDDTRARRLSLIEEGAEPQVRMAHLA
 CVGSHAINGVAALHTQLLKQETLRDFYELWPEKFFNMTNGVTPRRWLLQSNPRLAN
 LISDRIGNDWIHDRLRRLLEDSDVNDREFLQRWAEVKHQNKVDLSRYIYQQTRIEVD
 PHSFLFDVQVKRIHEYKRQLLAVMHIVTLYNWLKHNPQLNLVPRTFIFAGKAAPGY
 RAKQIVKLINAVGSIINHDPDVQGRLKVFLPNFNVSLGQRIYPAADLSEQISTAGKE
 45 ASGTGNMKFTMNGALTIGTYDGANIEIREEVGPENFFLFLGLRAEDIARRQSRGYRVE

FWSSNAELRAVLDRFSSGHFTPDQPNLFQDLVSDLLQRDEYMLMADYQSYIDCQRE
AAAAYRDSRWWRMSLLNTARSGKFSSDRTIADYSEQIWEVKPVPVSLSTSF

SEQ ID NO. 31 –

5 ATGGCTGCCATTAATACGAAAGTCAAAAAAGCCGTTATCCCCGTTGCGGGATTA
GGAACCAGGATGTTGCCGGCGACGAAAGCCATCCCGAAAGAGATGCTGCCACTT
GTCGATAAGCCATTAATTCAATACGTCGTGAATGAATGTATTGCGGCTGGCATT
CTGAAATTGTGCTGGTTACACACTCATCTAAAACTCTATTGAAAACCACTTTGA
TACCAGTTTTGAACTGGAAGCAATGCTGGAAAAACGTGTAAAACGTCAACTGCT
10 TGATGAAGTGCAGTCTATTTGTCCACCGCACGTGACTATTATGCAAGTTCGTCAG
GGTCTGGCGAAAGGCCTGGGACACGCGGTATTGTGTGCTCACCCGGTAGTGGGT
GATGAACCGGTAGCTGTTATTTTGCCTGATGTTATTCTGGATGAATATGAATCCG
ATTTGTCACAGGATAACCTGGCAGAGATGATCCGCCGCTTTGATGAAACGGGTC
ATAGCCAGATCATGGTTGAACCGGTTGCTGATGTGACCGCATATGGCGTTGTGGA
15 TTGCAAAGGCGTTGAATTAGCGCCGGGTGAAAGCGTACCGATGGTTGGTGTGGT
AGAAAAACCGAAAGCGGATGTTGCGCCGTCTAATCTCGCTATTGTGGGTCGTTAC
GTACTTAGCGCGGATATTTGGCCGTTGCTGGCAAAAACCCCTCCGGGAGCTGGTG
ATGAAATTCAGCTCACCGACGCAATTGATATGCTGATCGAAAAAGAAACGGTGG
AAGCCTATCATATGAAAGGGAAGAGCCATGACTGCGGTAATAAATTAGGTTACA
20 TGCAGGCCTTCGTTGAATACGGTATTCGTCATAACACCCTTGGCACGGAATTTAA
AGCCTGGCTTGAAGAAGAGATGGGCATTAAGAAGTAA

SEQ ID NO. 32 -

MAAINTKVKKAVIPVAGLGTRMLPATKAIPKEMLPLVDKPLIQYVVNECIAAGITEIV
25 LVTHSSKNSIENHFDTSFELEAMLEKRVKRQLLDEVQSICPPHVTIMQVRQGLAKGL
GHAVLCAHPVVGDEPVAVILPDVILDEYESDLSQDNLAEMIRRFDETGHSSQIMVEPV
ADVTAYGVVDCKGVELAPGESVPMVGVEKPKADVAPSNLAIVGRYVLSADIWPL
LAKTPPGAGDEIQLTDAIDMLIEKETVEAYHMKGKSHDCGNKLGVMQAFVEYGIRH
NTLGTEFKAWLEEEMGIKK

30

SEQ ID NO. 33 –

ATGAAATCCCCCAGGCTCAACAAATCCTAGACCAGGCCCGCCGTTTGTCTACG
AAAAAGCCATGGTCAAAATCAATGGGCAATACGTGGGGACGGTGGCGGCCATTC
CCCAATCGGATCACCATGATTTGAACTATACGGAAGTTTTCATTGCGGACAATGT
35 GCCGGTGATGATCTTCTTGTTACTGCAAAATGAAACGGAAATTGTCCAAAACCTT
TTGGAAATTTGCCTCACCTCCAAAGTAAGGGCTTTCCACCTACGGCATTTTTCC
CACTAGTTTTGTGGAAACGGAAAACCATGAACTCAAGGCAGACTATGGCCAACG
GGCGATCGGTTCGAGTTTGCTCGGTGGATGCGTCCCTCTGGTGGCCTATTTTGGCC
TATTACTACGTGCAAAGAACCGGCAATGAAGCCTGGGCTAGACAAACCCATGTG
40 CAATTGGGGCTACAAAAGTTTTTAAACCTCATTCTCCATCCAGTCTTTCGGGATG
CACCCACTTTGTTTGTGCCCCGACGGGGCCTTTATGATTGACCGCCCCATGGATGT
GTGGGGAGCGCCGTTGGAAATCCAAACCCTGCTCTACGGAGCCCTGAAAAGTGC

GCGGGGTTACTGTTAATCGACCTCAAGGCGAAGGGTTATTGCAGCAATAAAGA
 CCATCCTTTTGACAGCTTCACGATGGAGCAGAGTCATCAATTTAACCTGAGTGTG
 GATTGGCTCAAAAACTCCGCACCTATCTGCTCAAGCATTATTGGATTAATTGCA
 ATATTGTCCAAGCTCTCCGCCGCCGTCCACGGAACAGTACGGTGAAGAAGCCA
 5 GCAACGAACATAATGTCCACACAGAAACCATTCCCAACTGGCTCCAGGATTGGC
 TCGGCGATCGGGGAGGCTATTTAATCGGCAATATCCGCACGGGTGCCCCGATTT
 TCGCTTTTTCTCCCTGGGTAATTGCTTGGGGGCAATTTTCGATGTCACTAGCTTGG
 CCCAGCAACGTTTCCTTTTTCCGTTTGGTATTAAATAATCAGCGGGAGTTATGTGC
 CCAAATGCCCCTGAGGATTTGCCATCCCCCCTCAAAGATGACGATTGGCGCAGT
 10 AAAACCGGCTTTGACCGCAAAAATTTACCCTGGTGCTACCACAACGCCGGCCATT
 GGCCCTGTTTATTTTGGTTTCTGGTGGTGGCGGTGCTCCGCCATAGCTGCCATTCC
 AACTACGGCACGGTGGAGTATGCGGAAATGGGGAACCTAATTCGCAATAACTAT
 GAGGTGCTTTTGCGCCGTTTGCCAAGCATAAATGGGCTGAATATTTTGATGGCC
 CCACGGGCTTTTGGGTCGGGCAACAATCCCGTTCCTACCAAACCTGGACCATTGT
 15 GGGCCTATTGCTAGTACACCATTTACAGAAGTTAACCCCGACGATGCTTTGATG
 TTCGATTTGCCTAGTTTGAAAAGTTTGCATCAAGCGCTGCATTAA

SEQ ID NO. 34 -

MKSPQAQQILDQARRLLYEKAMVKINGQYVGTVAaipQSDHHDlnyTEVFIRDnVP
 20 VMIFLLLQNETeIVQNfLEICLTlQSKGFPTyGIFPTsFvETENhELKADYgQRAIGRV
 CSVDASLWWPILAYYYVQRTGNEAWARQTHVQLGLQKFLNLILHPVFRDAPTLFVP
 DGAFMIDRPMdVWGAPLEIQTLlyGALKsAAGLLlDLKAKGYCSNKDHPfDSFTM
 EQSHQfNLSVDWLKklRtYLLKHYwINCnIVqALRRRPTEqYGEEASNEhNVHTETI
 PNWLQDWLgDRGGYLIGNIRtGRpDfRfSLGNCLGAIFDVTSLAQQRsFFRLVLNN
 25 QRELCAQMPLRICHPPLKDDdWRsKTGFDRKNLPwCYHNAGHWpCLFWFLVVAVL
 RHSCHsNYGTVEYAEMGNLIRNNYEVLlRRLPKHkWAeyFDGPTGFwVGQqSRsY
 QTWTIVGLLLVHHfTEVNPDDAlMFdLPSLkSLHqALH

SEQ ID NO. 35 -

30 ATGAATTCATCCCTTGTGATCCTTTACCACCGTGAGCCCTACGACGAAGTTAGGG
 AAAATGGCAAAACGGTGTATCGAGAGAAAAAGAGTCCCAACGGGATTTTGCCCA
 CCCTCAAAGTTTTTTTGCCGATGCGGAACAGAGCACCTGGGTGCGATGGAAAC
 AGGTTTCGCCGAAGCAAAAGGATGATTTTCAGGCGGATATGTCCATTGAAGGCC
 TTGGCGATCGTTGTACGGTGCGCCGGGTGCCCCTGACGGCGGAGCAGGTAAAAA
 35 ACTTCTATCACATCACTTCCAAGGAAGCCTTTTGGCCCATTCTCCACTCTTTCCCC
 TGGCAGTTCACCTACGATTCTTCTGATTGGGATAATTTTCAGCACATTAACCGCTT
 ATTTGCCGAGGCGGCCTGTGCCGATGCCGATGACAATGCATTGTTTTGGGTCCAC
 GACTATAACCTCTGGTTAGCGCCCCCTTACATTCGTCAGCTCAAGCCCAACGCCA
 AGATTGCCTTTTTCCACCACACCCCTTCCCCAGCGTTGATATTTTCAATATTTTG
 40 CCCTGGCGGGAGGCGATCGTAGAAAGCTTGCTGGCCTGTGATCTCTGTGGTTTTTC
 ATATTCCCCGCTACGTAGAAAATTTTGTGCGCCGTGGCCCGTAGTCTCAAGCCGGT
 GGAAATCACCAGACGGGTTGTGGTAGACCAAGCCTTTACCCCTACGGTACGGC
 CCTGGCGGAACCGGAACCTACCAACCCAGTTGCGTTATGGCGATCGCCTCATTAAC

CTCGATGCGTTTCCCGTGGGCACCAATCCGGCAAATATCCGGGCGATCGTGGCCA
AAGAAAGTGTGCAACAAAAAGTTGCTGAAATTAAACAAGATTTAGGCGGTAAGA
GGCTAATTGTTTCCGCTGGGCGGGTGGATTACGTGAAGGGCACCAAGGAAATGT
TGATGTGCTATGAACGTCTACTGGAGCGTCGCCCCGAATTGCAGGGGGAAATTA
5 GCCTGGTAGTCCCCGTAGCCAAGGCCGCTGAGGGAATGCGTATTTATCGCAACG
CCCAAAACGAAATTGAACGACTGGCAGGGAAAATTAACGGTCGCTTTGCCAAAC
TGTCCTGGACACCAGTGATGCTGTTACCTCTCCTTTAGCCTATGAGGAGCTCATT
GCCCTGTTCTGTGCCGCCGACATTGCCTGGATCACTCCCCTGCGGGATGGGCTAA
ACCTGGTGGCTAAGGAGTATGTGGTGGCTAAAAATGGCGAAGAAGGAGTTCTGA
10 TCCTCTCGGAATTTGCCGGTTGTGCGGTGGAATAACCGATGCGGTGTTGACTAA
CCCCTACGCTTCCAGCCGTATGGACGAATCCATTGACCAGGCCCTGGCCATGGAC
AAAGACGAACAGAAAAACGCATGGGGAGAATGTACGCCGCCATTAAGCGTTA
CGACGTTCAACAATGGGCCAATCACCTACTGCGGGAAGCCTACGCCGATGTGGT
ACTGGGAGAGCCCCCCCCAAATGTAG
15

SEQ ID NO. 36 -

MNSSLVILYHREPYDEVRENGKTVYREKKSPNGILPTLKSFFADAEQSTWVAWKQV
SPKQKDDFQADMSIEGLGDRCTVRRVPLTAEQVKNFYHITSKEAFWPILHSFPWQFT
20 YDSSDWDNFQHINRLF AEACADADDNALFWVHDYNLWLAPLYIRQLKPNAKIAFF
HHTPFPSVDIFNILPWREAIVESLLACDL CGFHIPRYVENFVAVARSLKPVEITRRVVV
DQAFPTYGTALAEPELTTQLRYGDRNLNDAFPVGTNPANIRAIVAKESVQKVAEIK
QDLGGKRLIVSAGRVDYVKGTKEMLMCYERLLERRPELQGEISLVVPVAKAAEGMR
IYRNAQNEIERLAGKINGRFAKL SWTPVMLFTSPLAYEELIALFCAADIAWITPLRDG
25 LNLVAKEYVVAKNGEGLVILSEFAGCAVELPD AVL TNPYASSRMD ESIDQALAMD
KDEQKKRMGRMYAAIKRYDVQQWANHLLREAYADVVLGEPPQM

SEQ ID NO. 37 -

ATGAAGATTTTATTTGTGGCGGCGGAAGTATCCCCCCTAGCAAAGGTAGGTGGC
30 ATGGGGGATGTGGTGGGTTCCTGCCTAAAGTTCTGCATCAGTTGGGCCATGATG
TCCGTGTCTTCATGCCCTACTACGGTTTCATCGGCGACAAGATTGATGTGCCCAA
GGAGCCGGTCTGGAAAGGGGAAGCCATGTTCCAGCAGTTTGCTGTTTACCAGTCC
TATCTACCGGACACCAAAAATTCCTCTCTACTTGTTTCGGCCATCCAGCTTTCTGACTC
CCGAAGGATCTATGGCGGAGATGACGAGGCGTGGCGGTTCACTTTTTTTTCTAAC
35 GGGGCAGCTGAATTTGCCTGGAACCATGGAAGCCGGAATTATCCATTGCCAT
GATTGGCACACTGGCATGATCCCTGTTTGGATGCATCAGTCCCCAGACATCGCCA
CCGTTTTACCATCCATAATCTTGCTTACCAAGGGCCCTGGCGGGGCTTGCTTGA
AACTATGACTTGGTGTCCTTGGTACATGCAGGGAGACAATGTGATGGCGGCGGC
GATTCAATTTGCCAATCGGGTGA CTACCGTTTCTCCCACCTATGCCAACAGATC
40 CAAACCCCGGCCTATGGGGAAAAGCTGGAAGGGTTATTGTCCTACCTGAGTGGT
AATTTAGTCGGTATTCTCAACGGTATTGATACGGAGATT TACAACCCGGCGGAAG
ACCGCTTTATCAGCAATGTTTTCGATGCGGACAGTTTGGACAAGCGGGTGAAAA
ATAAAATTGCCATCCAGGAGGAAACGGGGTTAGAAATTAATCGTAATGCCATGG

TGGTGGGTATAGTGGCTCGCTTGGTGGAACAAAAGGGGATTGATTTGGTGATTCA
GATCCTTGACCGCTTCATGTCCTACACCGATTCCCAGTTAATTATCCTCGGCACTG
GCGATCGCCATTACGAAACCCAACTTTGGCAGATGGCTTCCCGATTTCTTGGGCG
GATGGCGGTGCAATTACTCCACAACGATGCCCTTTCCCGTCGAGTCTATGCCGGG
5 GCGGATGTGTTTTTAATGCCTTCTCGCTTTGAGCCCTGTGGGCTGAGTCAATTGAT
GGCCATGCGTTATGGCTGTATCCCCATTGTGCGGCGGACAGGGGGTTTGGTGAT
ACGGTATCCTTCTACGATCCTATCAATGAAGCCGGCACCGGCTATTGCTTTGACC
GTTATGAACCCCTGGATTGCTTTACGGCCATGGTGCGGGCCTGGGAGGGTTTCCG
TTTCAAGGCAGATTGGCAAAAATTACAGCAACGGGCCATGCGGGCAGACTTTAG
10 TTGGTACCGTTCGCGCGGGGAATATATCAAAGTTTATAAGGGCGTGGTGGGGAA
ACCGGAGGAATTAAGCCCCATGGAAGAGGAAAAAATCGCTGAGTTAACTGCTTC
CTATCGCTAA

15 SEQ ID NO. 38 -

MKILFVAAEVSPILAKVGGMGDVVGSPLPKVLHQLGHDVRFVMPYYGFIGDKIDVPKE
PVWKGEAMFQQFAVYQSYLPDTKIPLYLFGHPAFDSRRIYGGDDEAWRFTFFSNGA
AEFAWNHWKPEIIHCHDWHTGMIPVWMHQSPDIATVFTIHNLAYQGPWRGLLETMT
WCPWYMQGDNVMAAAIQFANRVTTVSPTYAQQIQTPAYGEKLEGLLSYLSGNLVGI
20 LNGIDTEIYNPAEDRFISNVFDADSLDKRVKNKIAIQEETGLEINRNMVVGIVARLV
EQKGIDLVIQILDRFMSYTDSQLIILGTGDRHYETQLWQMASRFPGRMAVQLLHNDAL
LSRRVYAGADVFLMPSRFEPCLSQLMAMRYGCIPIVRRTGGLVDTVSFYDPINEAG
TGYCFDRYEPLDCFTAMVRAWEGFRFKADWQKLQQRAMRADFSWYRSAGEYIKV
YKGVVGKPEELSPMEEEKIAELTASYR

25

SEQ ID NO. 39 -

TGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGC
AACTCGTAGGACAGGTGGTACCTACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGC
30 CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAT
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG
ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCT
GTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGA
35 ACCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC
AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGAAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
CCTTCGAAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
40 GCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTGCTAGCGAA
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT

AAGGGATTTTGGTCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAA
GGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAAA
TTCCAACATGGATGCTGATTATATGGGTATAAATGGGCTCGCGATAATGTCGGG
CAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTG
5 TTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCA
GACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG
TACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCCGGGAAAACAGCATTC
CAGGTATTAGAAGAATATCCTGATTGAGGTGAAAATATTGTTGATGCGCTGGCAG
TGTTCTGCGCCGGTTGCATTGATTCTGTTTGTAAATTGTCCTTTTAACAGCGAT
10 CGCGTATTTTCGCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATG
CGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAG
AAATGCATAAACTTTTGCCATTCTCACCAGGATTCAGTCGTCCTCATGGTGATTTC
TCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTG
GACGAGTCGGAATCGCAGACCGATAACCAGGATCTTGCCATCCTATGGAACCTGCC
15 TCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGAT
AATCCTGATATGAATAAATTGCAGTTTCATTGATGCTCGATGAGTTTTTCTAAGA
ATTAATTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
GGTTCGCGCACATTTCCCCGAAAAGTGCCACCTGAAATTGTAAACGTTAATATT
TTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGC
20 CGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGATAGGGTTGAG
TGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAGAAGTGGACTCCAACGTC
AAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCC
TAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAA
GGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAA
25 GGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGG
TCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCG
CGTCCCATTCGCCAATCCGGATATAGTTCTCCTTTTTCAGCAAAAAACCCCTCAAG
ACCCGTTTAGAGGGCCCCAAGGGGTTATGCTAGTTATTGCTCAGCGGTGGCAGCAG
CCAACCTCAGCTTCCTTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTG
30 GTGGTGCTCGAGTGCGGCCGAAGCTTTCATTAATGATGATGGTGGTGGTGGCTG
CCGGTCGCACGGGTGTCGCTGTATTTCTTCAGGTCTTCCAGGTGCGCCAGACGGT
TTTCTTTGTTAATACGTTGGGTGGTGATACGATCCGGATAGCAACGCATAAACAT
GTTGGTGAAGTGCTCGCCGTAGTGGATACGTTTCGTTTCGCGCTCGGCTCGAACAGC
GGATACGCGCTCGCTTCGAAGTTCGGCTCGTGAAAGTACGCGCACGCGAAACGT
35 TCACGGGTGTTTCAGTTTAACCTTGTGCGGGGTGCTCAGCAGTTGGCCACCGGTCA
TGAAGTGCAGAATATCGCCCGGAAAAACGGTCCACACACCCGGGGTTCGGGGTAA
CGAAGGTCCACGGTTCGTCGTGCTCAAACATGCCCCGCGCTGCTCTCGCCCGGCAG
CCAGTTACGGTTACGCTTTTTCGCCCTCCACCGGCGGACGGATATACAGGCCACCA
ACATCGTCTTGCGCCGCAATCACAGCAGACCGTAGTCGGTGTGCGCACCAATAC
40 CACGGCTCAGGGTGCTGGTCTGCGGCGGGAAACGCAGCACACGCATGTGGTGCC
AGCCATCACGGGTGAGGTGCGGTGAAGGTGTTGATCGGCAGTTCAAAACCCAGCG
CGGTCAGTTTCAGCAGACGCTCGCCCGCCAGACCCAGTTCCTCCATAAAGGTTTT
CATGCTCTTTTGATAGGTGTTGTTTCGGCCACGGAACCGGACCATGGCACGGCCAA
CCCGCTTTAACACGCTGATCGCCACGCTCAGGTCTTGCACACGGTAAAAATTT
45 CCGGGAAATCCGGCTTGCCCGCGGTAACTTCCTCGCCGCTCGCCACATAACCGCT
GTAGGTCAGGTCGCTAACGCAGCTGCTTTTGAAGGTCAGCGGCTCTTTGCAAAT

TGCTTGCTCGCCGCCATCGCTTCTTGGGTCTTACGATCCTGCTCGCTGTCGGTTTT
AATCTGGAAGATAACCATCCTTTTGCCACGCCTGAATCAGCGCACGACCCAGGCTG
ATGTCCGCCGCGCAACCGGTCACTTCGGTCGGCAGTTCAAAGGTCTGCAGGTGG
TCATGCTGGTTTCCTCTTTGTCCATGTACAGGCTCGGATGCTCATTCCACACAACG
5 TCCGGGCTGCATAACCCTAGTGAGGGAAATACTCCCCATCTACTTGGAGCGTGTA
TCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGT
TATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTTCGCGGGATCGAGATC
GATCTCGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGT
GCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGC
10 CACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGG
CCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGC GGCGGCGGT
GCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAG
GGAGAGCGTCGAGATCCCGGACACCATCGAATGGCGCAAAACCTTTTCGCGGTAT
GGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCACT
15 AACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC
GTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCG
GCGATGGCGGAGCTGAATTACATTCCTCAACCGCGTGGCACAACAACCTGGCGGGC
AAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGT
CGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGT
20 GGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAA
TCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATACTATCCGCTGGATGACCAG
GATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGT
CTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGA
CTGGGCGTGGAGCATCTGGTTCGATTGGGTACCAGCAAATCGCGCTGTTAGCG
25 GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATC
TACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCA
TGTCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGC
GATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGA
GTCCGGGCTGCGCGTGGTGCGGACATCTCGGTAGTGGGATACGACGATACCGA
30 AGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCTG
CTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTG
AAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACTGGCG
CCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGG
CACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTA
35 AGTTAGCTCACTCATTAGGCACCGGGATCTCGACCGATGCCCTTGAGAGCCTTCA
ACCCAGTCAGCTCCTTCCGG

SEQ ID NO. 40 –

ATTTAGCGTCTTCTAATCCAGTGTAGACAGTAGTTTTGGCTCCGTTGAGCACTGTA
40 GCCTTGGGCGATCGCTCTAAACATTACATAAATTCACAAAGTTTTCGTTACATAA
AAATAGTGTCTACTTAGCTAAAAATTAAGGGTTTTTTACACCTTTTTGACAGTTAA
TCTCCTAGCCTAAAAAGCAAGAGTTTTTAATAAGACTCTTGCCCTTTACAACCT
CGAAGGAGCGTCAGATCTCATATGCACCACCACCATCACACGAAAACCTGTAC
TTTCAGGGCAAGCTTATGATTCATGCCCCCTCCCGCTGGGGCGTGTTTCCAGTCT
45 GGGTCTCTGCTCCCCCGATGTGGTGTGGAACGAACACCCAGCCTGTACATGGAT

AAGGAAGAGACCAGTATGACCAATCTGCAAACCTTTGAACTGCCACCGAGGTG
ACCGGTTGCGCCGCCGATATTAGCCTCGGTCGCGCCCTGATTCAAGCCTGGCAAA
AGGATGGCATCTTCCAAATCAAGACCGATTCCGAACAAGATCGCAAGACCCAAG
AGGCCATGGCCGCCAGCAAACAATTTTGAAGGAACCCCTGACCTTTAAATCCA
5 GCTGCGTGAGCGATCTCACCTACAGTGGCTATGTGGCCAGTGGTGAAGAGGTGA
CCGCCGGCAAGCCCCGATTTTCCCGAGATTTTACCGTGTGCAAGGATCTGAGTGT
GGGTGATCAACGCGTGAAAGCCGGTTGGCCCTGCCATGGTCCCGTGCCCTGGCCC
AACAATACCTATCAAAAATCCATGAAGACCTTTATGGAAGAACTCGGTCTGGCC
GGTGAACGCCTGCTCAAACCTGACCGCCCTCGGCTTTGAGCTGCCCATTAACACCT
10 TTACCGATCTCACCCGCGATGGTTGGCACCACATGCGCGTGCTGCGCTTTCCTCC
CCAAACCAGCACCCCTGAGCCGCGGTATTGGTGCCACACCGATTACGGCCTGCTC
GTGATTGCCGCCCAAGATGATGTGGGCGGTCTGTATATTCGCCCTCCCGTGGAAG
GCGAGAAACGCAACCGCAATTGGCTCCCCGCGCAAAGTTCCGCCGGCATGTTTG
AACACGATGAACCCTGGACCTTTGTGACGCCCACGCCCGGCGTGTTGGACCGTGTT
15 TCCCGGTGATATTCTGCAATTTATGACCGGCGGTCAACTGCTCTCCACGCCCCAC
AAAGTGAAGCTCAACACCCGCGAACGCTTTGCCTGCGCCTACTTTCACGAACCCA
ATTTTGAGGCCAGTGCCTATCCCCTGTTTGAACCCTCCGCCAACGAGCGCATTCA
CTACGGCGAGCACTTTACCAATATGTTTATGCGCTGCTATCCCGATCGCATTACC
ACCCAACGCATTAACAAGGAAAATCGCCTGGCCACCTCGAGGATCTGAAAAAG
20 TATAGTGATACCCGCGCCACCGGTAGTGGTGCCACCAACTTTAGCCTGCTCAAAC
AAGCCGGCGATGTGGAAGAGAACCCCGGTCCCATGACCGAAAGTATTACCAGCA
ATGGCACCCCTGGTGGCCAGTGATACCCGTCGCCGCGTGTTGGGCCATTGTGAGTGC
CAGCAGTGGTAACCTGGTGGAGTGGTTTGTATTTTACGTGTATAGCTTTTGCAGT
CTCTACTTTGCCACATTTTCTTTCCAGTGGCAATACCACCACCCAACCTGCTGCA
25 AACCGCCGGCGTGTTTGGCGCCGGTTTCTGATGCGCCCCATTGGCGGTTGGCTC
TTTGGCCGCATTGCCGATCGTCGCGGTGCGAAGACCAGCATGCTGATTAGCGTGT
GCATGATGTGCTTTGGCTCCCTGATTATTGCCTGCCTCCCCGGCTATGATGCCATT
GGCACCTGGGCCCCCGCCCTGCTCCTGCTGGCCCGCCTCTTTCAAGGCCTGAGCG
TGGGCGGTGAATACGGCACCGCCACCTATATGAGTGAAATTGCCCTGGAGG
30 GCCGCAAAGGTTTTTACGCCAGTTTTCAATATGTGACCCTGATTGGCGGTCAACT
GCTCGCCATTCTCGTGGTGGTGATTCTCCAACAAATTCTGACCGATTCCCAACTG
CACGAATGGGGCTGGCGCATTCCCTTTGCCATGGGTGCCGCCCTGGCCATTGTGG
CCCTGTGGCTCCGTCGCCAACTCGATGAAACCAGCCAAAAAGAGGTGCGCGCCC
TGAAAGAAGCCGGCAGTTTTAAAGGTCTCTGGCGCAACCGCAAGGCCTTTCTCAT
35 GGTGCTGGGCTTTACCGCCGGCGGTAGTCTGTCTTTTACACCTTTACCACCTACA
TGCAAAAATATCTCGTGAACACCACCGGCATGCACGCCAATGTGGCCAGCGTGA
TTATGACCGCCGCCCTGTTTGTGTTTATGCTCATTCAACCCCTGATTGGCGCCCTC
AGCGATAAGATTGGTCGTCGCACCAAGTATGCTGATTTTGGCGGTATGAGTGCCC
TCTGCACCGTGCCATTCTCACCGCCCTGCAACACGTGTCCAGCCCCTACGCCGC
40 CTTTGCCCTCGTGATGCTGGCCATGGTGATTGTGTCTTTTATACCAGCATTAGTG
GCATTCTGAAGGCCGAAATGTTTCCCGCCCAAGTGCGCGCCCTGGGCGTGGGTCT
CAGTTACGCCGTGGCCAATGCCCTGTTTGGCGGTTCCGCCGAATATGTGGCCCTG
TCCCTCAAAAGCTGGGGCAGTGAGACCACCTTTTTCTGGTACGTGACCATTATGG
GTGCCCTGGCCTTTATTGTGAGCCTGATGCTCCACCGCAAAGGCAAGGGTATTGCG
45 CCTCTAGGGTACCAGGCAAACCCATCCCCAACCCCTGCTGGGCCTGGATAGCAC
CGGTGGTGGTCACCACCACCATCACCCTAGAGTACTGTATGCATCGAGTGCCTG

GCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAAGTCAGAAGTGAAAC
GCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAGTACC
AGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTT

5 SEQ ID NO. 41 –

ATTAGCGTCTTCTAATCCAGTGTAGACAGTAGTTTTGGCTCCGTTGAGCACTGTA
GCCTTGGGCGATCGCTCTAAACATTACATAAATTCACAAAGTTTTCGTTACATAA
AAATAGTGTCTACTTAGCTAAAAATTAAGGGTTTTTTACACCTTTTTGACAGTTAA
TCTCTAGCCTAAAAAGCAAGAGTTTTTAATAAGACTCTTGCCCTTTACAACCT
10 C

SEQ ID NO. 42 –

TGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAAGTCAGAAG
TGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGA
15 ACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTT

SEQ ID NO. 43 –

GAGGTTGTAAAGGGCAAGAGTCTTAGTTAAAAACTCTTGCTTTTTAGGCTAGGAG
ATTAAGTGTCAAAAAGGTGTAAAAAACCTTAATTTTTAGCTAAGTAGACACTAT
20 TTTTATGTAAACGAAAACCTTGTGAATTTATGTAATGTTTAGAGCGATCGCCCAAG
GCTACAGTGCTCAACGGAGCCAAAACCTACTGTCTACACTGGATTAGAAGACGCT
AAATGGTACCTACGATCTCATATGATACACGCTCCAAGTAGATGGGGAGTATTC
CCTCACTAGGGTTATGCAGCCCGGACGTTGTGTGGAATGAGCATCCGAGCCTGTA
CATGGACAAAGAGGAAACCAGCATGACCAACCTGCAGACCTTTGAACTGCCGAC
25 CGAAGTGACCGGTTGCGCGGCGGACATCAGCCTGGGTCTGCGCTGATTCAGGC
GTGGCAAAAGGATGGTATCTTCCAGATTAAAACCGACAGCGAGCAGGATCGTAA
GACCCAAGAAGCGATGGCGGCGAGCAAGCAATTTTGCAAAGAGCCGCTGACCTT
CAAAAGCAGCTGCGTTAGCGACCTGACCTACAGCGGTTATGTGGCGAGCGGCGA
GGAAGTTACCGCGGGCAAGCCGATTTCCCGGAAATTTTTACCGTGTGCAAGGA
30 CCTGAGCGTGGGCGATCAGCGTGTTAAAGCGGGTTGGCCGTGCCATGGTCCGTT
CCGTGGCCGAACAACACCTATCAAAAGAGCATGAAAACCTTTATGGAGGAACTG
GGTCTGGCGGGCGAGCGTCTGCTGAAACTGACCGCGCTGGGTTTTGAACTGCCG
ATCAACACCTTCACCGACCTGACCCGTGATGGCTGGCACCACATGCGTGTGCTGC
GTTTCCCGCCGCAGACCAGCACCTGAGCCGTGGTATTGGTGCGCACACCGACTA
35 CGGTCTGCTGGTGATTGCGGCGCAAGACGATGTTGGTGGCCTGTATATCCGTCCG
CCGGTGGAGGGCGAAAAGCGTAACCGTAACCTGGCTGCCGGGCGAGAGCAGCGC
GGGCATGTTTGAGCACGACGAACCGTGACCTTCGTTACCCCGACCCCGGGTGTG
TGGACCGTTTTTCCGGGCGATATTCTGCAGTTCATGACCGGTGGCCAAGTCTGA
GCACCCCGCACAAAGGTTAACTGAACACCCGTGAACGTTTCGCGTGCGCGTACTT
40 TCACGAGCCGAAGTTTGAAGCGAGCGCGTATCCGCTGTTTCGAGCCGAGCGCGAA
CGAACGTATCCACTACGGCGAGCACTTCACCAACATGTTTATGCGTTGCTATCCG
GATCGTATCACCACCCAACGTATTAACAAAGAAAACCGTCTGGCGCACCTGGAA

GACCTGAAGAAATACAGCGACACCCGTGCGACCGGCAGCCACCACCACCATCAT
CATTAATGAAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCCTGAGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAG
CAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG

5

SEQ ID NO. 44 –

CAAAAAACCCCTCAAGACCCGTTTAGAGGCCCAAGGGGTTATGCTAG

SEQ ID NO. 45 –

10 GAGCGTGTATCATATGAGATCTGACGCTCCTTCGAGGTTGTAAAGGGCAAGAGT
CTTAGTTAAAACTCTTGCTTTTTAGGCTAGGAGATTAAGTGTCAAAAAGGTGTA
AAAAACCCCTAATTTTTAGCTAAGTAGACACTATTTTTATGTAACGAAAACCTTG
TGAATTTATGTAATGTTTAGAGCGATCGCCAAAGGCTACAGTGCTCAACGGAGCC
AAAACCTACTGTCTACACTGGATTAGAAGACGCTAAATGGTACCTACGATCTCATA
15 TGATACACGCTCCAAGTAGATGGGGAGTATTTCCCTCACTAGGGTTATGCAGCCC
GGACGTTGTGTGGAATGAGCATCCGAGCCTGTACATGGACAAAGAGGAAACCAG
CATGACCAACCTGCAGACCTTTGAACTGCCGACCGAAGTGACCGGTTGCGCGGC
GGACATCAGCCTGGGTCGTGCGCTGATTCAGGCGTGGCAAAAGGATGGTATCTT
CCAGATTAAAACCGACAGCGAGCAGGATCGTAAGACCCAAGAAGCGATGGCGG
20 CGAGCAAGCAATTTTGCAAAGAGCCGCTGACCTTCAAAAGCAGCTGCGTTAGCG
ACCTGACCTACAGCGGTTATGTGGCGAGCGGCGAGGAAGTTACCGCGGGCAAGC
CGGATTTCCCGGAAATTTTTACCGTGTGCAAGGACCTGAGCGTGGGCGATCAGCG
TGTTAAAGCGGGTTGGCCGTGCCATGGTCCGGTTCCGTGGCCGAACAACACCTAT
CAAAAGAGCATGAAAACCTTTATGGAGGAACTGGGTCTGGCGGGCGAGCGTCTG
25 CTGAAACTGACCGCGCTGGGTTTTGAACTGCCGATCAACACCTTACCGACCTGA
CCCGTGATGGCTGGCACCACATGCGTGTGCTGCGTTTCCCGCCGACAGCAGCAC
CCTGAGCCGTGGTATTGGTGCGCACACCGACTACGGTCTGCTGGTGATTGCGGCG
CAAGACGATGTTGGTGGCCTGTATATCCGTCCGCGGTGGAGGGCGAAAAGCGT
AACCGTAACTGGCTGCCGGGCGAGAGCAGCGCGGGCATGTTTGAGCACGACGAA
30 CCGTGGACCTTCGTTACCCCGACCCGGGTGTGTGGACCGTTTTTCCGGGCGATA
TTCTGCAGTTCATGACCGGTGGCCAACTGCTGAGCACCCCGCACAAAGGTTAACT
GAACACCCGTGAACGTTTCGCGTGCGCGTACTTTCACGAGCCGAAGTTTCAAGCG
AGCGCGTATCCGCTGTTTCGAGCCGAGCGCAACGAACGTATCCACTACGGCGAG
CACTTCACCAACATGTTTATGCGTTGCTATCCGATCGTATCACCACCAACGTA
35 TTAACAAAGAAAACCGTCTGGCGCACCTGGAAGACCTGAAGAAATACAGCGACA
CCCGTGCGACCGGCAGCCACCACCACCATCATCATTAAATGAAAGCTTGCGGCCG
CACTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAAGCCCGAA
AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG
GGGCCTCTAAACGGGTCTTGAGGGGTTTTTTTGCTGAAAGGAGGAACTATATCCGG
40 ATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG
GTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCG
CTTTCTCCCTTCCTTTCTCGCCACGTTCCCGGCTTTCCCGTCAAGCTCTAAAT
CGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCAAAA

AAC TTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTT
TCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTG
GAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCG
ATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
5 TTAACAAAAATATTAACGTTTACAATTTACAGGTGGCACTTTTCGGGGAAATGTGCG
CGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
ATTAATTCTTAGAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATAT
CAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAA
CTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCG
10 ACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTAT
CAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTT
TATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAA
ATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGA
AATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGG
15 CGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTT
CTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATC
ATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCGGTCAG
CCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCAT
GTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATAACAATCGATAGATTGTGCG
20 ACCTGATTGCCCAGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCC
ATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCA
TAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGACCAA
AATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC
AAAGGATCTTCGCTAGCAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATC
25 CCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG
GATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTTC
CGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTA
GCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT
30 CTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCG
GGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTGGGCTGAACGG
GGGGTTCGTGCACACAGCCCAGCTTGAGCGAACGACCTACACCGAACTGAGAT
ACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCAGAGGGAGAAAGGCGG
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
35 CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC
TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACG
CCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTAGGTACCATTTAGCGTC

40