US 20120174253A1

(19) United States
(12) Patent Application Publication
Patterson et al.
(10) Pub. No.: US 2012/0174253 A1
(43) Pub. Date: Jul. 5, 2012

(54) GENERATION OF HIGH POLYHYDROXYBUTRATE PRODUCING OILSEEDS

(76) Inventors: **Nii Patterson**, Chelmsford, MA (US); **Jihong Tang**, Brighton, MA (US); **Edgar Benjamin Cahoon**, Lincoln, NE (US); **Jan G. Jaworski**, Town and Country, MO (US); **Wenyu Yang**, Saint Louis, MO (US); **Oliver P. Peoples**, Arlington, MA (US); **Kristi D. Snell**, Belmont, MA (US)

(21) Appl. No.: **13/395,616**

(22) PCT Filed: **Sep. 15, 2010**

(86) PCT No.: **PCT/US10/48962**

§ 371 (c)(1), (2), (4) Date: **Mar. 12, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/242,522, filed on Sep. 15, 2009.

Publication Classification

(51) **Int. Cl.**
A01H 1/02 (2006.01)
A01H 5/00 (2006.01)
C07C 59/147 (2006.01)

(52) **U.S. Cl.** **800/264**; 554/115; 800/298; 800/306; 800/312; 800/314; 800/317.3; 800/320.1; 800/322

(57) **ABSTRACT**

Transgenic oilseed plants, plant material, plant cells, and genetic constructs for synthesis of polyhydroxyalkanoates ("PHA") are provided. In a preferred embodiment, the transgenic oilseed plants synthesize (poly)3-hydroxybutyrate ("PHB") in the seed. Genes utilized include phaA, phaB, phaC, all of which are known in the art. The genes can be introduced in the plant, plant tissue, or plant cell using conventional plant molecular biology techniques.

Figure 1

S line (male sterile) X **M** line (maintainer)
(phaA and phaC)

↓

S line X **R** line (restorer)
(*phaA* and *phaC*) (*phaB*)

↓

Hybrid seeds
(*phaA*/*phaB*/*phaC*)

GENERATION OF HIGH POLYHYDROXYBUTRATE PRODUCING OILSEEDS

FIELD OF THE INVENTION

[0001] The invention is in the field of polymer production in transgenic plants. Methods for generating industrial oilseeds producing high levels of polyhydroxybutyrate (PHB) and industrial oilseeds producing high levels of PHB are described.

BACKGROUND OF THE INVENTION

[0002] Production of polyhydroxyalkanoates (PHAs), a family of naturally occurring renewable and biodegradable plastics, in crops has the potential of providing a renewable source of polymers, chemical intermediates and bio-energy from one crop if plant residues remaining after polymer isolation are converted to liquid fuels and/or energy. PHAs can provide an additional revenue stream that would make bioenergy crops more economically viable.

[0003] PHAs are a natural component of numerous organisms in multiple ecosystems and accumulate in a wide range of bacteria as a granular storage material when the microbes are faced with an unfavorable growth environment, such as a limitation in an essential nutrient (Madison et al., *Microbiol. Mol. Biol. Rev.,* 1999, 63, 21-53; Suriyamongkol et al., *Biotechnol Adv,* 2007, 25, 148-175). The monomer unit composition of these polymers is largely dictated by available carbon source as well as the native biochemical pathways present in the organism. Today PHAs are produced industrially from renewable resources in bacterial fermentations providing an alternative to plastics derived from fossil fuels. PHAs possess properties enabling their use in a variety of applications currently served by petroleum-based plastics and are capable of matching or exceeding the performance characteristics of fossil fuel derived plastics with a broad spectrum of properties that can be obtained by varying the monomer composition of homo- and co-polymers, or by manipulating properties such as molecular weight (Sudesh et al., *Prog. Polym. Sci.,* 2000, 25, 1503-1555; Sudesh et al., *CLEAN—Soil, Air, Water,* 2008, 36, 433-442).

[0004] Industrial production of PHAs in crop plants would provide a low cost, renewable source of plastics. Production of PHAs in plants has been an as yet unsolved goal for plant scientists and has previously been demonstrated in a number of crops unsuitable for industrial production or in industrially useful crops at levels to low to be commercially attractive [for review, see (Suriyamongkol et al., *Biotechnol Adv,* 2007, 25, 148-175); (van Beilen et al., *The Plant Journal,* 2008, 54, 684-701) and references within] including maize (Poirier et al., 2002, Polyhydroxyalkanoate production in transgenic plants, in Biopolymers, Vol 3a, Steinbuchel, A. (ed), Wiley-VHC Verlag GmbH, pgs 401-435), sugarcane (Purnell et al., *Plant Biotechnol. J.,* 2007, 5, 173-184), switchgrass (Somleva et al., *Plant Biotechnol J,* 2008, 6, 663-678), flax (Wrobel et al., *J. Biotechnol.,* 2004, 107, 41-54; Wrobel-Kwiatkowsk et al., *Biotechnol Prog,* 2007, 23, 269-277), cotton (John et al., *Proceedings of the National Academy of Sciences of the United States of America,* 1996, 93, 12768-12773), alfalfa (Saruul et al., *Crop Sci.,* 2002, 42, 919-927), tobacco (Arai et al., *Plant Biotechnol.,* 2001, 18, 289-293; Bohmert et al., *Plant Physiol.,* 2002, 128, 1282-1290; Lossl et al., *Plant Cell Reports,* 2003, 21, 891-899; Lössl et al., *Plant Cell Physiol,* 2005, 46, 1462-1471), potato (Bohmert et al., *Plant Physiol.,* 2002, 128, 1282-1290), and oilseed rape (Valentin et al., *Int. J. Biol. Macromol.,* 1999, 25, 303-306; Slater et al., *Nat. Biotechnol.,* 1999, 17, 1011-1016). Most of the efforts to produce PHAs in plants have focused on production of the homopolymer P3HB or the copolymer poly-3-hydroxybutyrate-co-3-hydroxyvalerate (P3HBV). While there have been some efforts to produce medium chain length PHAs in plants, these studies have yielded barely detectable levels of polymer (Romano et al., *Planta,* 2005, 220, 455-464; Mittendorf et al., *Proceedings of the National Academy of Sciences of the United States of America,* 1998, 95, 13397-13402; Poirier et al., *Plant Physiol.,* 1999, 121, 1359-1366; Matsumoto, *Journal of Polymers and the Environment,* 2006, 14, 369-374; Wang et al., *Chinese Science Bulletin,* 2005, 50, 1113-1120).

[0005] To date, the highest levels of polymer have been obtained when the homopolymer poly-3-hydroxybutyrate (P3HB or PHB) is produced in plastids (Suriyamongkol et al., *Biotechnol Adv,* 2007, 25, 148-175; van Beilen et al., *The Plant Journal,* 2008, 54, 684-701; Bohmert et al., *Molecular Biology and Biotechnology of Plant Organelles,* 2004, 559-585). This is likely due to the high flux of acetyl-CoA, the precursor for PHB in these organelles during fatty acid biosynthesis (Bohmert et al, *Molecular Biology and Biotechnology of Plant Organelles,* 2004, 559-585). Expression of three genes encoding β-ketothiolase, acetoacetyl CoA reductase, and PHA synthase, allows the conversion of acetyl-CoA within the plastid to PHB. Previous work has reported producing levels of PHB in *Brassica napus* up to a maximum of 6.7% of seed weight, a level too low for commercial production

SUMMARY OF THE INVENTION

[0006] Transgenic oilseed plants, plant material, plant cells, and genetic constructs for synthesis of polyhydroxyalkanoates ("PHA") are provided. In a preferred embodiment, the transgenic oilseed plants synthesize (poly)3-hydroxybutyrate ("PHB") in the seed. Host plants, plant tissue, and plant material have been engineered to express genes encoding enzymes in the biosynthetic pathway for PHB production such that polymer precursors in the plastid are polymerized to polymer. Genes utilized include phaA, phaB, phaC, all of which are known in the art. The genes can be introduced in the plant, plant tissue, or plant cell using conventional plant molecular biology techniques.

[0007] It is an object of the invention to provide methods and compositions for producing transgenic oilseeds having commercially viable levels of polyhydroxyalkanoates in the seed, for example greater than 7%, 10%, 15%, or 19% polyhydroxyalkanoate or more of the total dry seed weight.

[0008] It is another object of the invention to provide oilseeds having increased levels of polyhydroxyalkanoate greater than 7%, 10%, 15%, or 19% polyhydroxyalkanoate or more of the total dry seed weight and having impaired germination relative to non-transgenic oilseeds.

[0009] Using a non-traditional screening method to identify transgenic lines than those used in all other reported studies, it has been discovered that very high levels of PHA, for example PHB can be produced in the oilseed but that oilseeds with high levels of PHA fail to germinate or germinate but produce impaired seedlings which do not survive to produce viable fertile plants. The failure to produce viable progeny explains why previous researchers failed to demonstrate that commercial levels of PHA can be produced in transgenic oilseeds. A preferred PHA produced in oilseeds is PHB.

[0010] In another embodiment the transgenes encoding PHA biosynthesis genes are expressed in a seed specific manner such that the PHA accumulates in the seed. In this embodiment the level of PHA accumulated is greater than 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18% and 19% of the dry weight of the seed.

[0011] Methods and compositions for producing hybrid lines are also provided. Hybrid lines can be created by crossing a line containing one or more PHAs, for example PHB genes with a line containing the other gene(s) needed to complete the PHA biosynthetic pathway. Use of lines that possess cytoplasmic male sterility with the appropriate maintainer and restorer lines allows these hybrid lines to be produced efficiently.

[0012] In still another embodiment the oilseeds produced by the disclosed methods produce high levels of PHA and are impaired in their ability to germinate and survive to produce viable plants relative to oilseeds containing little or no PHA, for example less than 7% PHA of the dry weight of the seed. Germination can be impaired by 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or 100% relative to oilseeds with less than 7% PHA. Impaired germination provides a built in mechanism for gene containment reducing the risk of unwanted growth of these oilseeds when a different crop is planted on the production fields.

[0013] Transgenic plants useful for the invention include dicots or monocots. Preferred host plants are oilseed plants, but are not limited to members of the Brassica family including *B. napus, B. rapa, B. carinata* and *B. juncea*. Additional preferred host plants include industrial oilseeds such as *Camelina sativa*, Crambe, jatropha, and castor. Other preferred host plants include *Arabidopsis thaliana, Calendula, Cuphea*, maize, soybean, cottonseed, sunflower, palm, coconut, safflower, peanut, mustards including *Sinapis alba,* and tobacco.

[0014] Other embodiments provide plant material and plant parts of the transgenic plants including seeds, flowers, stems, and leaves. The oilseeds can be used for the extraction of PHA biopolymer or as a source of PHA biopolymer based chemical intermediates. The residual parts of the seed can be used as meal for animal feed or steam and power generation and a source of vegetable oil for industrial oelochemicals or biofuel.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] FIG. **1** is a schematic diagram describing a strategy for creating hybrid seeds using cytoplasmic male sterility.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

[0016] Unless otherwise indicated, the disclosure encompasses all conventional techniques of plant breeding, microbiology, cell biology and recombinant DNA, which are within the skill of the art. See, e.g., Sambrook and Russell, Molecular Cloning: A Laboratory Manual, 3rd edition (2001); Current Protocols In Molecular Biology [(F. M. Ausubel, et al. eds., (1987)]; Plant Breeding Principles and Prospects (Plant Breeding, Vol 1) M. D. Hayward, N. O. Bosemark, I. Romagosa; Chapman & Hall, (1993.); Coligan, Dunn, Ploegh, Speicher and Wingfeld, eds. (1995) Current Protocols in Protein Science (John Wiley & Sons, Inc.); the series Methods in Enzymology (Academic Press, Inc.): PCR 2: A Practical Approach (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)].

[0017] Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Lewin, Genes VII, published by Oxford University Press, 2000; Kendrew et al. (eds.), The Encyclopedia of Molecular Biology, published by Wiley-Interscience., 1999; and Robert A. Meyers (ed.), Molecular Biology and Biotechnology, a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995; Ausubel et al. (1987) Current Protocols in Molecular Biology, Green Publishing; Sambrook and Russell. (2001) Molecular Cloning: A Laboratory Manual 3rd. edition.

[0018] A number of terms used herein are defined and clarified in the following section.

[0019] The term PHB refers to polyhydroxybutyrate and is used interchangeably with the term PHA which refers to polyhydroxyalkanoate.

[0020] The tend PHB also encompasses copolymers of hydroxybutyrate with other hydroxyacid monomers.

[0021] The term "PHA copolymer" refers to a polymer composed of at least two different hydroxyalkanoic acid monomers.

[0022] The term "PHA homopolymer" refers to a polymer that is composed of a single hydroxyalkanoic acid monomer.

[0023] As used herein, a "vector" is a replicon, such as a plasmid, phage, or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. The vectors can be expression vectors.

[0024] As used herein, an "expression vector" is a vector that includes one or more expression control sequences

[0025] As used herein, an "expression control sequence" is a DNA sequence that controls and regulates the transcription and/or translation of another DNA sequence. Control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, a ribosome binding site, and the like. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[0026] As used herein, "operably linked" means incorporated into a genetic construct so that expression control sequences effectively control expression of a coding sequence of interest.

[0027] As used herein, "transformed" and "transfected" encompass the introduction of a nucleic acid into a cell by a number of techniques known in the art.

[0028] "Plasmids" are designated by a lower case "p" preceded and/or followed by capital letters and/or numbers.

[0029] As used herein the term "heterologous" means from another host. The other host can be the same or different species.

[0030] The term "cell" refers to a membrane-bound biological unit capable of replication or division.

[0031] The term "construct" refers to a recombinant genetic molecule including one or more isolated polynucleotide sequences.

[0032] Genetic constructs used for transgene expression in a host organism comprise in the 5'-3' direction, a promoter sequence; a nucleic acid sequence encoding the desired transgene product; and a termination sequence. The open reading frame may be orientated in either a sense or anti-sense direction. The construct may also comprise selectable marker gene (s) and other regulatory elements for expression.

[0033] The term "plant" is used in it broadest sense. It includes, but is not limited to, any species of woody, ornamental or decorative, crop or cereal, fruit or vegetable plant, and photosynthetic green algae (e.g., *Chlamydomonas reinhardtii*). It also refers to a plurality of plant cells that are largely differentiated into a structure that is present at any stage of a plant's development. Such structures include, but are not limited to, a fruit, shoot, stem, leaf, flower petal, etc. The term "plant tissue" includes differentiated and undifferentiated tissues of plants including those present in roots, shoots, leaves, pollen, seeds and tumors, as well as cells in culture (e.g., single cells, protoplasts, embryos, callus, etc.). Plant tissue may be in planta, in organ culture, tissue culture, or cell culture. The term "plant part" as used herein refers to a plant structure, a plant organ, or a plant tissue.

[0034] A non-naturally occurring plant refers to a plant that does not occur in nature without human intervention. Non-naturally occurring plants include transgenic plants and plants produced by non-transgenic means such as plant breeding.

[0035] The term "plant cell" refers to a structural and physiological unit of a plant, comprising a protoplast and a cell wall. The plant cell may be in form of an isolated single cell or a cultured cell, or as a part of higher organized unit such as, for example, a plant tissue, a plant organ, or a whole plant.

[0036] The term "plant cell culture" refers to cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes and embryos at various stages of development.

[0037] The term "plant material" refers to leaves, stems, roots, flowers or flower parts, fruits, pollen, egg cells, zygotes, seeds, cuttings, cell or tissue cultures, or any other part or product of a plant.

[0038] A "plant organ" refers to a distinct and visibly structured and differentiated part of a plant such as a root, stem, leaf, flower bud, or embryo.

[0039] "Plant tissue" refers to a group of plant cells organized into a structural and functional unit. Any tissue of a plant, whether in a plant or in culture, is included. This term includes, but is not limited to, whole plants, plant organs, plant seeds, tissue culture and any groups of plant cells organized into structural and/or functional units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

[0040] "Seed germination" refers to growth of an embryonic plant contained within a seed resulting in the formation and emergence of a seedling.

[0041] "Cotyledon" refers to the embryonic first leaves of a seedling.

[0042] "Early plantlet development" refers to growth of the cotyledon containing seedling to form a plantlet.

II. Transgenic Plants

[0043] Transgenic plants have been developed that produce increased levels of biopolymers such as polyhydroxyalkanoates (PHAs) in seeds. Methods and constructs for engineering plants for seed specific production of PHA, in particular PHB, are described. One embodiment provides transgenic plants for the direct, large scale production of PHAs in crop plants or in energy crops where a plant by-product, such as oil, can be used for production of energy. Proof of concept studies for polyhydroxybutyrate (PHB) synthesis in canola (Valentin et al., *Int. J. Biol. Macromol.,* 1999, 25, 303-306; Houmiel et al., *Planta,* 1999, 209, 547-550; Slater et al., *Nat. Biotechnol.,* 1999, 17, 1011-1016.) have been reported. There have been instances where high level PHB production in the chloroplasts of plants has led to decreases in total plant growth (Bohmert et al., *Molecular Biology and Biotechnology of Plant Organelles,* 2004, 559-585; Bohmert et al., *Planta,* 2000, 211, 841-845) for unidentified reasons. There have been several studies that have attempted to alleviate this problem by inducible expression of enzymes (Bohmert et al., *Plant Physiol.,* 2002, 128, 1282-1290; Lössl et al., *Plant Cell Physiol,* 2005, 46, 1462-1471; Kourtz et al., *Transgenic Res,* 2007, 16, 759-769).

[0044] Transgenic oilseeds comprising at least about 8% dry weight PHA are provided. In one embodiment we provide transgenic oilseeds having at least 10% PHA dry weight and which are impaired in germination and plant survival.

[0045] A. Genetic Constructs for Transformation

[0046] Suitable genetic constructs include expression cassettes for enzymes for production of polyhydroxyalkanoates, in particular from the polyhydroxybutyrate biosynthetic pathway. In one embodiment, the construct contains operatively linked in the 5' to 3' direction, a seed specific promoter that directs transcription of a nucleic acid sequence in the nucleus; a nucleic acid sequence encoding one of the PHB biosynthetic enzymes; and a 3' polyadenylation signal that increases levels of expression of transgenes. In one embodiment, enzymes for formation of polymer precursors are targeted to the plastid using appropriate plastid-targeting signals. In another embodiment, the PHA pathway is expressed directly from the plastid genome using appropriate plastidial promoters and regulatory sequences.

[0047] DNA constructs useful in the methods described herein include transformation vectors capable of introducing transgenes into plants. As used herein, "transgenic" refers to an organism in which a nucleic acid fragment containing a heterologous nucleotide sequence has been introduced. The transgenes in the transgenic organism are preferably stable and inheritable. The heterologous nucleic acid fragment may or may not be integrated into the host genome.

[0048] Several plant transformation vector options are available, including those described in "Gene Transfer to Plants" (Potrykus, et al., eds.) Springer-Verlag Berlin Heidelberg New York (1995); "Transgenic Plants: A Production System for Industrial and Pharmaceutical Proteins" (Owen, et al., eds.) John Wiley & Sons Ltd. England (1996); and "Methods in Plant Molecular Biology: A Laboratory Course Manual" (Maliga, et al. eds.) Cold Spring Laboratory Press, New York (1995). Plant transformation vectors generally include one or more coding sequences of interest under the transcriptional control of 5' and 3' regulatory sequences, including a promoter, a transcription termination and/or polyadenylation signal, and a selectable or screenable marker gene. For the expression of two or more polypeptides from a single transcript, additional RNA processing signals and ribozyme sequences can be engineered into the construct (U.S. Pat. No. 5,519,164). This approach has the advantage of locating multiple transgenes in a single locus, which is advantageous in subsequent plant breeding efforts.

[0049] Engineered minichromosomes can also be used to express one or more genes in plant cells. Cloned telomeric repeats introduced into cells may truncate the distal portion of a chromosome by the formation of a new telomere at the integration site. Using this method, a vector for gene transfer can be prepared by trimming off the arms of a natural plant chromosome and adding an insertion site for large inserts (Yu et al., Proc Natl Acad Sci USA, 2006, 103, 17331-6; Yu et al., *Proc Natl Acad Sci USA,* 2007, 104, 8924-9). The utility of engineered minichromosome platforms has been shown using Cre/lox and FRT/FLP site-specific recombination systems on a maize minichromosome where the ability to undergo recombination was demonstrated (Yu et al., *Proc Natl Acad Sci USA,* 2006, 103, 17331-6; Yu et al., *Proc Natl Acad Sci USA,* 2007, 104, 8924-9). Such technologies could be applied to minichromosomes, for example, to add genes to an engineered plant. Site specific recombination systems have also been demonstrated to be valuable tools for marker gene removal (Kerbach, S. et al., *Theor Appl Genet,* 2005, 111, 1608-1616), gene targeting (Chawla, R. et al., *Plant Biotechnol J,* 2006, 4, 209-218; Choi, S. et al., *Nucleic Acids Res,* 2000, 28, E19; Srivastava, V, & Ow, D W, *Plant Mol Biol,* 2001, 46, 561-566; Lyznik, L A, et al., *Nucleic Acids Res,* 1993, 21, 969-975), and gene conversion (Djukanovic, V, et al., *Plant Biotechnol J,* 2006, 4, 345-357).

[0050] An alternative approach to chromosome engineering in plants involves in vivo assembly of autonomous plant minichromosomes (Carlson et al., *PLoS Genet,* 2007, 3, 1965-74). Plant cells can be transformed with centromeric sequences and screened for plants that have assembled autonomous chromosomes de novo. Useful constructs combine a selectable marker gene with genomic DNA fragments containing centromeric satellite and retroelement sequences and/or other repeats.

[0051] Another approach is Engineered Trait Loci ("ETL") technology (U.S. Pat. No. 6,077,697 to Hadlaczky et al.; US Patent Application 2006/0143732). This system targets DNA to a heterochromatic region of plant chromosomes, such as the pericentric heterochromatin, in the short arm of acrocentric chromosomes. Targeting sequences may include ribosomal DNA (rDNA) or lambda phage DNA. The pericentric rDNA region supports stable insertion, low recombination, and high levels of gene expression. This technology is also useful for stacking of multiple traits in a plant (US Patent Application 2006/0246586, 2010/0186117 and PCT WO 2010/037209).

[0052] Zinc-finger nucleases (ZFNs) are also useful in that they allow double strand DNA cleavage at specific sites in plant chromosomes such that targeted gene insertion or deletion can be performed (Shukla et al., *Nature,* 2009; Townsend et al., *Nature,* 2009).

[0053] For direct expression of transgenes from the plastid genome, a vector to transform the plant plastid chromosome by homologous recombination (as described in U.S. Pat. No. 5,545,818 to McBride et al.) is used in which case it is possible to take advantage of the prokaryotic nature of the plastid genome and insert a number of transgenes as an operon. WO 2010/061186 describes an alternative method for introducing genes into the plastid chromosome using an adapted endogenous cellular process for the transfer of RNAs from the cytoplasm to the plastid where they are incorporated by homologous recombination. This plastid transformation procedure is also suitable for practicing the disclosed compositions and methods.

[0054] A transgene may be constructed to encode a multifunctional enzyme through gene fusion techniques in which the coding sequences of different genes are fused with or without linker sequences to obtain a single gene encoding a single protein with the activities of the individual genes. Transgenes encoding a bifunctional protein containing thiolase and reductase activities (Kourtz, L., K. et al. (2005), Plant Biotechnol. 3: 435-447) and a trifunctional protein having each of the three enzyme activities required for PHB expression in plants (Mullaney and Rehm (2010), Journal of Biotechnology 147: 31-36) have been described. Such synthetic fusion gene/enzyme combinations can be further optimized using molecular evolution technologies.

[0055] A transgene may be constructed to encode a series of enzyme activities separated by intein sequences such that on expression, two or more enzyme activities are expressed from a single promoter as described by Snell in U.S. Pat. No. 7,026,526 to Metabolix, Inc.

[0056] 1. Genes Involved in Polyhydroxyalkanoate Synthesis

[0057] In a preferred embodiment, the products of the transgenes are enzymes and other factors required for production of a biopolymer, such as a polyhydroxyalkanoate (PHA).

[0058] For PHA production, transgenes encode enzymes such as beta-ketothiolase, acetoacetyl-CoA reductase, PHB ("short chain") synthase, PHA ("long chain") synthase, threonine dehydratase, dehydratases such as 3-OH acyl ACP, isomerases such as Δ 3-cis, Δ 2-trans isomerase, propionyl-CoA synthetase, hydroxyacyl-CoA synthetase, hydroxyacyl-CoA transferase, R-3-hydroxyacyl-ACP:CoA transferase, thioesterase, fatty acid synthesis enzymes and fatty acid beta-oxidation enzymes. Useful genes are well known in the art, and are disclosed for example by Snell and Peoples *Metab. Eng.* 4: 29-40 (2002); Bohmert et. al. in *Molecular Biology and Biotechnology of Plant Organelles*. H. Daniell, C. D. Chase Eds., Kluwer Academic Publishers, Netherlands, 2004, pp. 559-585; (Suriyamongkol et al., *Biotechnol Adv,* 2007, 25, 148-175; van Beilen et al., *The Plant Journal,* 2008, 54, 684-701).

[0059] PHA Synthases

[0060] Examples of PHA synthases include a synthase with medium chain length substrate specificity, such as phaC1 from *Pseudomonas oleovorans* (WO 91/000917; Huisman, et al. *J. Biol. Chem.* 266, 2191-2198 (1991)) or *Pseudomonas aeruginosa* (Timm, A. & Steinbuchel, A. *Eur. J. Biochem.* 209: 15-30 (1992)), the synthase from *Alcaligenes eutrophus* with short chain length specificity (Peoples, O. P. & Sinskey, A. J. *J. Biol. Chem.* 264:15298-15303 (1989)), or a two subunit synthase such as the synthase from *Thiocapsa pfennigii* encoded by phaE and phaC (U.S. Pat. No. 6,011,144). Other useful PHA synthase genes have been isolated from, for example, *Alcaligenes latus* (Accession ALU47026), *Burkholderia* sp. (Accession AF153086), *Aeromonas caviae* (Fukui & Doi, *J. Bacteriol.* 179: 4821-30 (1997)), *Acinetobacter* sp. strain RA3849 (Accession L37761), *Rhodospirillum rubrum* (U.S. Pat. No. 5,849,894), *Rhodococcus ruber* (Pieper & Steinbuechel, *FEMS Microbiol. Lett.* 96(1): 73-80 (1992)), and *Nocardia corallina* (Hall et. al., *Can. J. Microbiol.* 44: 687-91 (1998)), *Arthrospira* sp. PCC 8005 (Accessions ZP_07166315 and ZP_07166316), *Cyanothece* sp. PCC 7425 (Accessions ACL46371 and ACL46370) and *Synechocystis* sp. PCC6803 (Accession BAA17430; Hein et al. (1998), Archives of Microbiology 170: 162-170).

[0061] PHA synthases with broad substrate specificity useful for producing copolymers of 3-hydroxybutyrate and longer chain length (from 6 to 14 carbon atoms) hydroxyacids have also been isolated from *Pseudomonas* sp. A33 (Appl. Microbiol. Biotechnol. 42: 901-909 (1995)) and *Pseudomonas* sp. 61-3 (Accession AB014757; Kato, et al. Appl. Microbiol. Biotechnol. 45: 363-370 (1996)).

[0062] A range of PHA synthase genes and genes encoding additional metabolic steps useful in PHA biosynthesis are described by Madison and Huisman. *Microbiology and Molecular biology Reviews* 63:21-53 (1999)) and Suriyamongkol et al. (Suriyamongkol et al., *Biotechnol Adv,* 2007, 25, 148-175).

[0063] Hydratase and Dehydrogenase

[0064] An alpha subunit of beta-oxidation multienzyme complex pertains to a multifunctional enzyme that minimally possesses hydratase and dehydrogenase activities. The subunit may also possess epimerase and A 3-cis, A 2-trans isomerase activities. Examples of alpha subunits of the beta-oxidation multienzyme complex are FadB from *E. coli* (DiRusso, C. C. J. Bacteriol. 1990, 172, 6459-6468), FaoA from *Pseudomonas fragi* (Sato, S., Hayashi, et al. J. Biochem. 1992, 111, 8-15), and the *E. coli* open reading frame 1714 that contains homology to multifunctional α subunits of the β-oxidation complex (Genbank Accession #1788682). A β subunit of the β-oxidation complex refers to a polypeptide capable of forming a multifunctional enzyme complex with its partner α subunit. The β subunit possesses thiolase activity. Examples of β subunits are FadA from *E. coli* (DiRusso, C. C. *J. Bacteriol.* 172: 6459-6468 (1990)), FaoB from *Pseudomonas fragi* (Sato, S., Hayashi, M., Imamura, S., Ozeki, Y., Kawaguchi, A. *J. Biochem.* 111: 8-15 (1992)), and the *E. coli* open reading frame f436 that contains homology to α subunits of the β-oxidation complex (Genbank Accession # AE000322; gene b2342).

[0065] Reductases

[0066] The transgene can encode a reductase. A reductase refers to an enzyme that can reduce β-ketoacyl CoAs to R-3-OH-acyl CoAs, such as the NADH dependent reductase from *Chromatium vinosum* (Liebergesell, M., & Steinbuchel, A. *Eur. J. Biochem.* 209: 135-150 (1992)), the NADPH dependent reductase from *Alcaligenes eutrophus* (Accession J04987, Peoples, O. P. & Sinskey, A. J. *J. Biol. Chem.* 264: 15293-15297 (1989))), the NADPH reductase from *Zoogloea ramigera* (Accession P23238; Peoples, O. P. & Sinskey, A. J. Molecular Microbiology 3: 349-357 (1989)) or the NADPH reductase from *Bacillus megaterium* (U.S. Pat. No. 6,835, 820), *Alcaligenes latus* (Accession ALU47026), *Rhizobium meliloti* (Accession RMU17226), *Paracoccus denitrificans* (Accession D49362), *Burkholderia* sp. (Accession AF153086), *Pseudomonas* sp. strain 61-3 (Accession AB014757), *Acinetobacter* sp. strain RA3849 (Accession L37761), *P. denitrificans*, (Accession P50204), and *Synechocystis* sp. Strain PCC6803 (Taroncher-Oldenburg et al., (2000), Appl. Environ. Microbiol. 66: 4440-4448).

[0067] Thiolases

[0068] The transgene can encode a thiolase. A beta-ketothiolase refers to an enzyme that can catalyze the conversion of acetyl CoA and an acyl CoA to a β-ketoacyl CoA, a reaction that is reversible. An example of such thiolases are PhaA from *Alcaligenes eutropus* (Accession J04987, Peoples, O. P. & Sinskey, A. J. J. Biol. Chem. 264: 15293-15297 (1989)), BktB from *Alcaligenes eutrophus* (Slater et al. *J Bacteriol.* 180(8):1979-87 (1998)) and thiolases from the following *Rhizobium meliloti* (Accession RMU17226), *Z. ramigera* (Accession P07097), *Paracoccus denitrificans* (Accession D49362), *Burkholderia* sp. (Accession AF153086), *Alcaligenes latus* (Accession ALU47026), *Allochromatium vinosum* (Accession P45369), *Thiocystis violacea* (Accession P45363); *Pseudomonas* sp. strain 61-3 (Accession AB014757), *Acinetobacter* sp. strain RA3849 (Accession L37761) and *Synechocystis* sp. Strain PCC6803 (Taroncher-Oldenburg et al., (2000), Appl. Environ. Microbiol. 66: 4440-4448).

[0069] Oxidases

[0070] An acyl CoA oxidase refers to an enzyme capable of converting saturated acyl CoAs to Δ 2 unsaturated acyl CoAs. Examples of acyl CoA oxidases are POX1 from *Saccharomyces cerevisiae* (Dmochowska, et al. Gene, 1990, 88, 247-252) and ACX1 from *Arabidopsis thaliana* (Genbank Accession # AF057044).

[0071] Catalases

[0072] The transgene can also encode a catalase. A catalase refers to an enzyme capable of converting hydrogen peroxide to hydrogen and oxygen. Examples of catalases are KatB from *Pseudomonas aeruginosa* (Brown, et al.): Bacterial. 177: 6536-6544 (1995)) and KatG from *E. coli* (Triggs-Raine, B. L. & Loewen, P. C. *Gene* 52: 121-128 (1987)).

[0073] 2. Promoters

[0074] Plant promoters can be selected to control the expression of the transgene in different plant tissues or organelles for all of which methods are known to those skilled in the art (Gasser & Fraley, *Science* 244:1293-99 (1989)). In one embodiment, promoters are selected from those of eukaryotic or synthetic origin that are known to yield high levels of expression in plant and algae cytosol. In another embodiment, promoters are selected from those of plant or prokaryotic origin that are known to yield high expression in plastids. In certain embodiments the promoters are inducible. Inducible plant promoters are known in the art.

[0075] Suitable constitutive promoters for nuclear-encoded expression include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in U.S. Pat. No. 6,072,050; the core CAMV 355 promoter, (Odell et al. (1985) *Nature* 313:810-812); rice actin (McElroy et al. (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen et al. (1989) *Plant Mol. Biol.* 12:619-632 and Christensen et al. (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last et al. (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten et al. (1984) *EMBO J.* 3:2723-2730); and ALS promoter (U.S. Pat. No. 5,659,026). Other constitutive promoters include, for example, U.S. Pat. Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142.

[0076] "Tissue-preferred" promoters can be used to target a gene expression within a particular tissue such as seed, leaf or root tissue. Tissue-preferred promoters include Yamamoto et al. (1997) *Plant J* 12(2)255-265; Kawamata et al. (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen et al (1997) *Mol. Gen. Genet.* 254(3):337-343; Russell et al. (1997) *Transgenic Res.* 6(2):157-168; Rinehart et al. (1996) *Plant Physiol.* 112 (3):1331-1341; Van Camp et al (1996) *Plant Physiol.* 112(2): 525-535; Canevascini et al. (1996) *Plant Physiol.* 112(2): 513-524; Yamamoto et al. (1994) *Plant Cell Physiol.* 35(5): 773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco et al. (1993) *Plant Mol. Biol.* 23(6):1129-1138; Matsuoka et al. (1993) *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; and Guevara-Garcia et al. (1993) *Plant J.* 4(3):495-505.

[0077] "Seed-preferred" promoters include both "seed-specific" promoters (those promoters active during seed development such as promoters of seed storage proteins) as well as "seed-germinating" promoters (those promoters active during seed germination). See Thompson et al. (1989) *BioEssays* 10:108. Such seed-preferred promoters include, but are not limited to, Cim1 (cytokinin-induced message); cZ19B1 (maize 19 kDa zein); milps (myo-inositol-1-phosphate synthase); and ce1A (cellulose synthase). Gama-zein is a preferred endosperm-specific promoter. Glob-1 is a preferred embryo-specific promoter. For dicots, seed-specific promoters include, but are not limited to, bean β-phaseolin, napin β-conglycinin, soybean lectin, cruciferin, oleosin, the *Lesquerella* hydroxylase promoter, and the like. For monocots, seed-specific promoters include, but are not limited to, maize 15 kDa zein, 22 kDa zein, 27 kDa zein, g-zein, waxy, shrunken 1, shrunken 2, globulin 1, etc. Additional seed specific promoters useful for practicing this invention are described in the Examples disclosed herein.

[0078] Leaf-specific promoters are known in the art. See, for example, Yamamoto et al. (1997) *Plant J.* 12(2):255-265; Kwon et al. (1994) *Plant Physiol.* 105:357-67; Yamamoto et al. (1994) *Plant Cell Physiol.* 35(5):773-778; Gotor et al. (1993) *Plant J.* 3:509-18; Orozco et al. (1993) *Plant Mol. Biol.* 23(6):1129-1138; and Matsuoka et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590.

[0079] Root-preferred promoters are known and may be selected from the many available from the literature or isolated de nova from various compatible species. See, for example, Hire et al. (1992) Plant Mol. Biol. 20(2): 207-218 (soybean root-specific glutamine synthetase gene); Keller and Baumgartner (1991) Plant Cell 3(10):1051-1061 (root-specific control element in the GRP 1.8 gene of French bean); Sanger et al. (1990) Plant Mol. Biol. 14(3):433-443 (root-specific promoter of the mannopine synthase (MAS) gene of *Agrobacterium tumefaciens*); and Miao et al. (1991) Plant Cell 3(1):1 1'-22 (full-length cDNA clone encoding cytosolic glutamine synthetase (GS), which is expressed in roots and root nodules of soybean). See also U.S. Pat. Nos. 5,837,876; 5,750,386; 5,633,363; 5,459,252; 5,401,836; 5,110,732; and 5,023,179.

[0080] Plastid specific promoters include the PrbcL promoter [Allison L. A. et al., EMBO 15: 2802-2809 (1996); Shiina T. et al., Plant Cell 10: 1713-1722 (1998)]; the PpsbA promoter [Agrawal G K, et al., Nucleic Acids Research 29: 1835-1843 (2001)]; the Prrn 16 promoter [Svab Z & Maliga P., Proc. Natl. Acad. Sci. USA 90: 913-917 (1993), Allison L A et al., EMBO 15: 2802-2809 (1996)]; the PaccD promoter (WO97/06250; Hajdukiewicz P T J et al., EMBO J. 16: 4041-4048 (1997)).

[0081] Chemical-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize 1n2-2 promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1 a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena et al. *Proc. Natl. Acad. Sci. USA* 88:10421-10425 (1991) and McNellis et al. *Plant J* 14(2):247-257 (1998)) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz et al. *Mol. Gen. Genet.* 227:229-237 (1991), and U.S. Pat. Nos. 5,814,618 and 5,789,156), herein incorporated by reference in their entirety.

[0082] In one embodiment, coordinated expression of the three transgenes, phaA, phaB, and phaC, necessary for conversion of acetyl-CoA to PHB is controlled by a seed specific promoter, such as the soybean oleosin promoter (Rowley et al., *Biochim Biophys Acta,* 1997, 1345, 1-4) or the promoter from the lesquerlla hydroxylase gene (U.S. Pat. No. 6,437,220 B1). In another embodiment, coordinated expression of the three transgenes, phaA, phaB, and phaC, necessary for conversion of acetyl-CoA to PHB is controlled by a promoter active primarily in the biomass plant, such as the maize chlorophyll A/B binding protein promoter (Sullivan et al., *Mol. Gen. Genet.,* 1989, 215, 431-40). It has been previously shown that plants transformed with multi-gene constructs produced higher levels of polymer than plants obtained from crossing single transgene lines (Valentin et al., *Int. J. Biol. Macromol.,* 1999, 25, 303-306; Bohmert et al., *Planta,* 2000, 211, 841-845).

[0083] In one embodiment, the final molecular weight of the polymer produced is controlled by the choice of promoter for expression of the PHA synthase gene. As described in U.S. Pat. No. 5,811,272, high PHA synthase activity will lower polymer molecular weight and low PHA synthase activity will increase polymer molecular weight. In another embodiment, a strong promoter is used for expression of the genes encoding plastid-targeted monomer producing enzymes while a weaker promoter is used to control expression of synthase.

[0084] 3. Transcription Termination Sequences

[0085] At the extreme 3' end of the transcript of the transgene, a polyadenylation signal can be engineered. A polyadenylation signal refers to any sequence that can result in polyadenylation of the mRNA in the nucleus prior to export of the mRNA to the cytosol, such as the 3' region of nopaline synthase (Bevan, M., Barnes, W. M., Chilton, M. D. Nucleic Acids Res. 1983, 11, 369-385).

[0086] 4. Selectable Markers

[0087] Genetic constructs may encode a selectable marker to enable selection of plastid transformation events. There are many methods that have been described for the selection of transformed plants [for review see (Miki et al., *Journal of Biotechnology,* 2004, 107, 193-232) and references incorporated within]. Selectable marker genes that have been used extensively in plants include the neomycin phosphotransferase gene nptII (U.S. Pat. No. 5,034,322, U.S. Pat. No. 5,530,196), hygromycin resistance gene (U.S. Pat. No. 5,668,298), the bar gene encoding resistance to phosphinothricin (U.S. Pat. No. 5,276,268), the expression of aminoglycoside 3"-adenyltransferase (aadA) to confer spectinomycin resistance (U.S. Pat. No. 5,073,675), the use of inhibition resistant 5-enolpyruvyl-3-phosphoshikimate synthetase (U.S. Pat. No. 4,535,060) and methods for producing glyphosate tolerant plants (U.S. Pat. No. 5,463,175; U.S. Pat. No. 7,045,684). Methods of plant selection that do not use antibiotics or herbicides as a selective agent have been previously described and include expression of glucosamine-6-phosphate deaminase to inactive glucosamine in plant selection medium (U.S. Pat. No. 6,444,878) and a positive/negative system that utilizes D-amino acids (Erikson et al., *Nat Biotechnol,* 2004, 22, 455-8). European Patent Publication No. EP 0 530 129 A1 describes a positive selection system which enables the transformed plants to outgrow the non-transformed lines by expressing a transgene encoding an enzyme that activates an inactive compound added to the growth media. U.S. Pat. No. 5,767,378 describes the use of mannose or xylose for the positive selection of transgenic plants. Methods for positive selection using sorbitol dehydrogenase to convert sorbitol to fructose for plant growth have also been described (WO 2010/102293). Screenable marker genes include the beta-glucuronidase gene (Jefferson et al., 1987, *EMBO J.* 6: 3901-3907; U.S. Pat. No. 5,268,463) and native or modified green fluorescent protein gene (Cubitt et al., 1995, *Trends Biochem. Sci.* 20: 448-455; Pan et al., 1996, *Plant Physiol.* 112: 893-900).

[0088] Transformation events can also be selected through visualization of fluorescent proteins such as the fluorescent proteins from the nonbioluminescent Anthozoa species which include DsRed, a red fluorescent protein from the *Discosoma* genus of coral (Matz et al. (1999), Nat Biotechnol 17: 969-73). An improved version of the DsRed protein has been developed (Bevis and Glick (2002), Nat Biotech 20: 83-87) for reducing aggregation of the protein. Visual selection can also be performed with the yellow fluorescent proteins (YFP) including the variant with accelerated maturation of the signal (Nagai, T. et al. (2002), Nat Biotech 20: 87-90), the blue fluorescent protein, the cyan fluorescent protein, and the green fluorescent protein (Sheen et al. (1995), Plant J 8: 777-84; Davis and Vierstra (1998), Plant Molecular Biology 36: 521-528). A summary of fluorescent proteins can be found in Tzfira et al. (Tzfira et al. (2005), Plant Molecular Biology 57: 503-516) and Verkhusha and Lukyanov (Verkhusha, V. V. and K. A. Lukyanov (2004), Nat Biotech 22: 289-296) whose references are incorporated in entirety. Improved versions of many of the fluorescent proteins have been made for various applications. Use of the improved versions of these proteins or the use of combinations of these proteins for selection of transformants will be obvious to those skilled in the art. It is also practical to simply analyze progeny from transformation events for the presence of the PHB thereby avoiding the use of any selectable marker.

[0089] For plastid transformation constructs, a preferred selectable marker is the spectinomycin-resistant allele of the plastid 16S ribosomal RNA gene (Staub J M, Maliga P, *Plant Cell* 4: 39-45 (1992); Svab Z, Hajdukiewicz P, Maliga P, *Proc. Natl., Acad. Sci. USA* 87: 8526-8530 (1990)). Selectable markers that have since been successfully used in plastid transformation include the bacterial aadA gene that encodes aminoglycoside adenyltransferase (AadA) conferring spectinomycin and streptomycin resistance (Svab et al., *Proc, Natl. Acad. Sci. USA,* 1993, 90, 913-917), nptII that encodes aminoglycoside phosphotransferase for selection on kanamycin (Carrer H, Hockenberry T N, Svab Z, Maliga P., *Mol. Gen. Genet.* 241: 49-56 (1993); Lutz K A, et al., *Plant J.* 37: 906-913 (2004); Lutz K A, et al., *Plant Physiol.* 145: 1201-1210 (2007)), aphA6, another aminoglycoside phosphotransferase (Huang F—C, et al, *Mol. Genet. Genomics* 268: 19-27 (2002)), and chloramphenicol acetyltransferase (Li, W., et al. (2010), Plant Mol Biol, DOI 10.1007/s11103-010-9678-4). Another selection scheme has been reported that uses a chimeric betaine aldehyde dehydrogenase gene (BADH) capable of converting toxic betaine aldehyde to nontoxic glycine betaine (Daniell H, et al., *Curr. Genet.* 39: 109-116 (2001)).

[0090] 5. Plastid Targeting Signals

[0091] Plastid targeting sequences are known in the art and include the chloroplast small subunit of ribulose-1,5-bisphosphate carboxylase (Rubisco) (de Castro Silva Filho et al, *Plant Mal. Biol.* 30:769-780 (1996); Schnell et *J. Biol. Chem.* 266(5):3335-3342 (1991)); 5-(enolpyruvyl)shikimate-3-phosphate synthase (EPSPS) (Archer et al. *J. Bioenerg. Biomemb.* 22(6):789-810 (1990)); tryptophan synthase (Zhao et al. *J. Biol. Chem.* 270(11):6081-6087 (1995)); plastocyanin (Lawrence et al. *J. Biol. Chem.* 272(33):20357-20363 (1997)); chorismate synthase (Schmidt et al. *J. Biol. Chem.* 268(36):27447-27457 (1993)); and the light harvesting chlorophyll a/b binding protein (LHBP) (Lamppa et al. *J. Biol. Chem.* 263:14996-14999 (1988)). See also Von Heijne et al. *Plant Mol. Biol. Rep.* 9:104-126 (1991); Clark et al. *J. Biol. Chem.* 264:17544-17550 (1989); Della-Cioppa et al. *Plant Physiol.* 84:965-968 (1987); Romer et al. *Biochem. Biophys. Res. Commun.* 196:1414-1421 (1993); and Shah et al. *Science* 233:478-481 (1986). Alternative plastid targeting signals have also been described in the following: US 2008/0263728; Miras, S. et al. (2002), J Biol Chem 277(49): 47770-8; Miras, S. et al. (2007), J Biol Chem 282: 29482-29492.

[0092] B. Exemplary Host Plants

[0093] Plants transformed in accordance with the present disclosure may be monocots or dicots. The transformation of suitable agronomic plant hosts using vectors for nuclear transformation or direct plastid transformation can be accomplished with a variety of methods and plant tissues. Representative plants useful in the methods disclosed herein include the Brassica family including *B. napus, B. rapa, B. carinata* and *B. juncea*; industrial oilseeds such as *Camelina sativa*, Crambe, jatropha, castor; *Calendula, Cuphea, Arabidopsis thaliana*; maize; soybean; cottonseed; sunflower; palm; coconut; safflower; peanut; mustards including *Sinapis alba*; sugarcane flax and tobacco, also are useful with the methods disclosed herein. Representative tissues for transformation using these vectors include protoplasts, cells, callus tissue, leaf discs, pollen, and meristems.

[0094] C. Methods of Plant Transformation

[0095] Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway et al. (1986) *Biotechniques* 4:320-334), electroporation (Riggs et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606), *Agrobacterium*-mediated transformation (Townsend et al., U.S. Pat. No. 5,563,055; Zhao et al. WO US98/01268), direct gene transfer (Paszkowski et al. (1984) EMBO J. 3:2717-2722), and ballistic particle acceleration (see, for example, Sanford et al., U.S. Pat. No. 4,945,050; Tomes et al. (1995) *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); and McCabe et al. *Biotechnology* 6:923-926 (1988)). Also see Weissinger et al. *Ann. Rev. Genet.* 22:421-477 (1988); Sanford et al, Particulate Science and Technology 5:27-37 (1987) (onion); Christou et al. *Plant Physiol.* 87:671-674 (1988) (soybean); McCabe et al. (1988) BioTechnology 6:923-926 (soybean); Finer and McMullen *In Vitro Cell Dev. Biol.* 27P:175-182 (1991) (soybean); Singh et al. Theor. Appl. Genet. 96:319-324 (1998) (soybean); Dafta et al. (1990) *Biotechnology* 8:736-740 (rice); Klein et al. *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (1988) (maize); Klein et al. *Biotechnology* 6:559-563 (1988) (maize); Tomes, U.S. Pat. No. 5,240,855; Buising et al., U.S. Pat. Nos. 5,322,783 and 5,324,646; Tomes et al. (1995) in Plant Cell, Tissue, and Organ Culture Fundamental Methods, ed. Gamborg (Springer-Verlag, Berlin) (maize); Klein et al. Plant Physiol. 91:440-444 (1988) (maize); Fromm et al. *Biotechnology* 8:833-839 (1990) (maize); Hooykaas-Van Slogteren et al. *Nature* 311:763-764 (1984); Bowen et al., U.S. Pat. No. 5,736,369 (cereals); Bytebier et al. *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (1987) (Liliaceae); De Wet et al. in The Experimental Manipulation of Ovule Tissues, ed. Chapman et al. (Longman, N.Y.), pp. 197-209 (1985) (pollen); Kaeppler et al. *Plant Cell Reports* 9:415-418 (1990) and Kaeppler et al. *Theor. Appl. Genet.* 84:560-566 (1992) (whisker-mediated transformation); D'Halluin et al. *Plant Cell* 4:1495-1505 (1992) (electroporation); Li et al. *Plant Cell Reports* 12:250-255 (1993) and Christou and Ford Annals of Botany 75:407-413 (1995) (rice); Osjoda et al. *Nature Biotechnology* 14:745-750 (1996) (maize via *Agrobacterium tumefaciens*); all of which are herein incorporated by reference in their entirety. References for protoplast transformation and/or gene gun for Agrisoma technology are described in WO 2010/037209. Methods for transforming plant protoplasts are available including transformation using polyethylene glycol (PEG), electroporation, and calcium phosphate precipitation (see for example Potrykus et al., 1985, Mol. Gen. Genet., 199, 183-188; Potrykus et al., 1985, Plant Molecular Biology Reporter, 3, 117-128), Methods for plant regeneration from protoplasts have also been described [Evans et al., in Handbook of Plant Cell Culture, Vol 1, (Macmillan Publishing Co., New York, 1983); Vasil, 1K in Cell Culture and Somatic Cell Genetics (Academic, Orlando, 1984)].

[0096] Methods for transformation of plastids such as chloroplasts are known in the art. See, for example, Svab et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation may be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase (McBride et al., *Proc. Natl. Acad. Sci. USA,* 1994, 91:7301-7305) or by use of an integrase, such as the phiC31 phage site-specific integrase, to target the gene insertion to a previously inserted phage attachment site (Lutz et al., *Plant J,* 2004, 37, 906-13). Plastid transformation vectors can be designed such that the transgenes are expressed from a promoter sequence that has been inserted with the transgene during the plastid transformation process or, alternatively, from an endogenous plastidial promoter such that an extension of an existing plastidial operon is achieved (Herz et al., *Transgenic Research,* 2005, 14, 969-982). An alternative method for plastid transformation as described in WO 2010/061186 wherein RNA produced in the nucleus of a plant cell can be targeted to the plastid genome can also be used to practice the disclosed invention. Inducible gene expression from the plastid genome using a synthetic riboswitch has also been reported (Verhounig et al. (2010), Proc Natl Acad Sci USA 107: 6204-6209). Methods for designing plastid transformation vectors are described by Lutz et al. (Lutz et al., *Plant Physiol,* 2007, 145, 1201-10).

[0097] Recombinase technologies which are useful for producing the disclosed transgenic plants include the cre-lox, FLP/FRT and Gin systems. Methods by which these technologies can be used for the purpose described herein are described for example in (U.S. Pat. No. 5,527,695; Dale And Ow, 1991*, Proc. Natl. Acad. Sci. USA* 88: 10558-10562; Medberry et al., 1995*, Nucleic Acids Res.* 23: 485-490).

[0098] D. Methods for Reproducing Transgenic Plants

[0099] Following transformation by any one of the methods described above, the following procedures can be used to obtain a transformed plant expressing the transgenes: select the plant cells that have been transformed on a selective medium; regenerate the plant cells that have been transformed to produce differentiated plants; select transformed plants expressing the transgene producing the desired level of desired polypeptide(s) in the desired tissue and cellular location.

[0100] In plastid transformation procedures, further rounds of regeneration of plants from explants of a transformed plant or tissue can be performed to increase the number of transgenic plastids such that the transformed plant reaches a state of homoplasmy (all plastids contain uniform plastomes containing transgene insert).

[0101] The cells that have been transformed may be grown into plants in accordance with conventional techniques. See, for example, McCormick et al, *Plant Cell Reports* 5:81-84 (1986). These plants may then be grown, and either pollinated with the same transformed variety or different varieties, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that constitutive expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure constitutive expression of the desired phenotypic characteristic has been achieved.

[0102] In some scenarios, it may be advantageous to insert a multi-gene pathway into the plant by crossing of lines containing portions of the pathway to produce hybrid plants in which the entire pathway has been reconstructed. This is especially the case when high levels of product in a seed compromises the ability of the seed to germinate or the resulting seedling to survive under normal soil growth conditions. Hybrid lines can be created by crossing a line containing one or more PHB genes with a line containing the other gene(s) needed to complete the PHB biosynthetic pathway. Use of lines that possess cytoplasmic male sterility (Esser, K. et al., 2006, Progress in Botany, Springer Berlin Heidelberg. 67, 31-52) with the appropriate maintainer and restorer lines allows these hybrid lines to be produced efficiently. Cytoplasmic male sterility systems are already available for some Brassicaceae species (Esser, K. et al., 2006, Progress in Botany, Springer Berlin Heidelberg. 67, 31-52). These Brassicaceae species can be used as gene sources to produce cytoplasmic male sterility systems for other oilseeds of interest such as Camelina.

III. Methods for Use

[0103] The disclosed genetic constructs can be used to produce industrial oilseed plants for high levels of PHA production. Specifically, PHA is produced in the seed.

[0104] The transgenic plants can be grown and harvested. The polyhydroxyalkanoate can be isolated from the oilseeds and the remaining plant material can be used as a feedstock for industrial use, preferably for the production of oleochemicals, energy or for use as feed for animals. The polyhydroxyalkanoate harvested from the plants can then be used to produce plastics, rubber material, coating material, and binders for paints, or as a feedstock for producing chemical derivatives such as hydroxyacids, esters, alkenoic acids or amines. PHA also has several medical applications.

[0105] The present invention will be further understood by reference to the following non-limiting examples.

EXAMPLES

Example 1

Design and Construction of Transformation Vectors for Production of PHB in Oilseeds

[0106] Five different vectors for seed specific expression of the PHB pathway were constructed containing different seed specific promoters for production of PHB in oilseeds (Table 1). Vector pMBXS490, a pCAMBIA based plasmid (Centre for Application of Molecular Biology to International Agriculture, Canberra, Australia), contains the following gene expression cassettes: (1) an expression cassette for PHA synthase containing the promoter from the soybean oleosin isoform A gene, a DNA fragment encoding the signal peptide of the small subunit of rubisco from pea (*P. sativum*) and the first 24 amino acids of the mature protein (Cashmore, A. R. 1983, In Genetic Engineering of Plants, pp. 29-38), a DNA fragment encoding a hybrid PHA synthase (PhaC; U.S. Pat. No. 6,316,262) in which the first nine amino acids at the N-terminus of this synthase are derived from the *Pseudomonas oleovorans* phaC1 gene and the remainder of the synthase coding sequence is derived from *Zoogloea ramigera* phaC gene, and the 3' termination sequence from the soybean oleosin isoform A gene; (2) an expression cassette for reductase containing the promoter from the soybean oleosin isoform A gene, a DNA fragment encoding the signal peptide and the first 24 amino acids of the mature protein of the small subunit of rubisco from pea, a DNA fragment encoding a NADPH dependent reductase (PhaB) from *Ralstonia eutropha eutropha* (Peoples, O. & A. Sinskey, 1989, J. Biol. Chem., 264, 15293-15297), and the 3' termination sequence from the soybean oleosin isoform A gene; (3) an expression cassette for thiolase containing the promoter from the soybean glycinin (gy1) gene (Iida et al., 1995, Plant Cell Reports, 14, 539-544), a DNA fragment encoding the signal peptide and the first 24 amino acids of the mature protein of the small subunit of rubisco from pea, the phaA gene encoding a β-ketothiolase (PhaA) from *Ralstonia eutropha* (Peoples, O. & A. Sinskey, 1989, J. Biol. Chem., 264, 15293-15297), and a 3' termination sequence from the soybean glycinin gene; (4) an expression cassette for DsRed, a protein that can be visualized in seeds by placing them in light of the appropriate wavelength, containing the promoter from the cassava mosaic virus (CMV), a DNA fragment encoding a modified red fluorescent protein from *Discosoma* sp. (DsRed) in which eleven amino acids have been added to the C-terminus to increase solubility and/or prevent aggregation of the protein, and a termination sequence from the *Agrobacterium tumefaciens* nopaline synthase gene.

TABLE 1

Summary of transformation vectors containing seed specific promoters

Plasmid	Promoter controlling expression of pha genes	Selectable or visible marker
pMBXS490	Oleosin	DsRed
pMBXS364	LH	DsRed
pMBXS355	LH	bar
pMBXS491	Napin	DsRed
pMBXS492	Glycinin	DsRed

[0107] Promoters are as follows: LH, promoter from the *Lesquerella fendleri* bifunctional oleate 12-hydroxylase: saturate gene (U.S. Pat. No. 6,437,220 B1); Oleosin, promoter from the soybean oleosin isoform A gene (Rowley and Herman, 1997, Biochim. Biophys. Acta 1345, 1-4); Napin, promoter from the *Brassica* napes napin gene (Ellenstrom, M. et al., 1996, Plant Molecular Biology, 32: 1019-1027); Glycinin, promoter from the soybean glycinin (gy1) gene (Iida, A. et al., 1995, Plant Cell Reports, 14:539-544).

[0108] Vectors pMBXS364, pMBXS355, pMBXS491, and pMBXS492 contain the same PHB pathway genes as pMBXS490 with the exception that the expression of these genes is under the control of different promoters as outlined in Table 1. Vector pMBXS355 contains an expression cassette for the bar gene, encoding phosphinothricin acetyltransferase whose expression is under the control of the 35S promoter. Expression of the bar gene allows selection of transformants based on their resistance to bialaphos. All other vectors in Table 1 contain expression cassettes for DsRed allowing the identification of transgenic seeds under the appropriate wavelength of light.

Example 2

Transformation of Camelina

[0109] In preparation for plant transformation experiments, seeds of *Camelina sativa* cultivar Suneson or Celine were sown directly into 4 inch pots filled with soil (Metro mix) in the greenhouse. Growth conditions were maintained at 24° C. during the day and 18° C. during the night. Plants were grown until flowering. Plants with a number of unopened flower buds were used in 'floral dip' transformations.

[0110] *Agrobacterium* strain GV3101 was transformed with the construct of interest using electroporation. A single colony of GV3101 containing the construct of interest was obtained from a freshly streaked plate and was inoculated into 5 mL LB medium. After overnight growth at 28° C., 2 mL of culture was transferred to a 500-mL flask containing 300 mL of LB and incubated overnight at 28° C. Cells were pelleted by centrifugation (6,000 rpm, 20 min), and diluted to an OD600 of ~0.8 with infiltration medium containing 5% sucrose and 0.05% (v/v) Silwet-L77 (Lehle Seeds, Round Rock, Tex., USA). Camelina plants were transformed by "floral dip" using transformation constructs as follows. Pots containing plants at the flowering stage were placed inside a 460 mm height vacuum desiccator (Bel-Art, Pequannock, N.J., USA). Inflorescences were immersed into the *Agrobacterium* inoculum contained in a 500-ml beaker. A vacuum (85 kPa) was applied and held for 5 min. Plants were removed from the desiccator and were covered with plastic bags in the dark for 24 h at room temperature. Plants were removed from the bags and returned to normal growth conditions within the greenhouse for seed formation.

[0111] To identify Camelina seeds expressing DsRed, fully mature seeds were harvested from transformed plants and placed in a desiccator with anhydrous calcium sulfate as desiccant for at least 2 days prior to screening. DsRed expressing seeds were visualized in a darkroom with a green LumaMax LED flashlight (Lab Safety Supply, Inc., Janesville, Wis.) and a pair of KD's Dark Red glasses (Pacific Coast Sunglasses Inc., Santa Maria, Calif.).

[0112] To identify bialaphos resistant seeds, seeds from floral dip transformations were sterilized in 70% ethanol and 10% bleach, and washed in water. Sterilized seeds were placed on germination and selection medium in square Petri dishes. The germination and selection medium contained 10 mg/L bialaphos (Gold BioTechnology, 130178-500) in ½× MS medium, which was made with Murashige & Skoog medium mixture (Caisson Labs, MSP09) at half concentration. The plates were sealed and placed in a growth chamber for germination under a 16-h photoperiod, 3,000 lux light intensity, and temperatures of 23/20° C. at day/night. Seedlings with greenish cotyledons were picked and transferred to soil about six days after initiation of germination.

Example 3

Production of PHB in Seeds of Camelina

[0113] In initial transformation experiments with pMBXS490, 24 DsRed positive seeds were isolated. Four of these seeds were sacrificed to determine their PHB content using a previously described gas chromatography/butanolysis technique performed essentially as previously described (Somleva et al., 2008, *Plant Biotechnol. J.,* 663-678). These four seeds contained 19.9, 12.0, 9.8, and 6.4% dwt PHB in the seed. When other seeds from this transformation were planted in soil, seedlings possessed whitish cotyledons and their growth was severely impaired. Only a few T_1 seeds with low levels of PHB were capable of germination and survival in soil in a greenhouse. These seedlings were still weak and possessed white or variegated cotyledons.

[0114] In transformations of pMBXS355 and pMBXS364, seeds from transformed plants were screened for resistance to bialophos and or visual screening for DsRed, respectively. Despite having the same promoter controlling the expression of the PHB biosynthetic pathway, the maximum PHB production in pMBXS355 (0.54% PHB) was significantly lower than the amount produced by pMBXS364 (3.4%) (Table 2). This is likely due to difficulty in distinguishing between weak pMBXS355 seedlings that produced higher levels of PHB and the non-transformed, bialophos sensitive seedlings.

TABLE 2

Comparison of PHB production in Lines isolated using bialaphos selection or visual screening

Vector	Selectable or Screenable Marker	# of Lines Tested	# of Lines w/ PHB in T2 Seeds	Range of PHB Production (% seed weight)
pMBXS355	Bar[1]	204	5	0.05 to 0.54%
pMBXS364	DsRed[2]	170	85	0.5 to 3.4%

[1]Selection of transformants performed by germination of seeds on tissue culture plates containing 10 mg/L bialophos.
[2]Selection of transformants performed by visual screening for DsRed expression.

[0115] In transformations with pMBX491 and pMBX492 containing the PHB genes under the control of the napin and glycinin promoters, respectively, were healthier than transformants obtained from pMBX490 transformations. For pMBX491, T2 seeds were isolated containing 8% PHB in DsRed seeds picked from the segregating population. These seeds possessed a 75% germination rate and a 60% survival rate under greenhouse conditions in soil. The cotyledons after 11 days were chlorotic and the growth of this line was significantly delayed compared to wild-type. For pMBX492, T2 seeds were isolated containing 6.9% PHB in DsRed seeds picked from the segregating population. These seeds possessed a 75% germination rate and a 70% survival rate under greenhouse conditions in soil. After 11 days, the cotyledons and first true leaves of this transformant were green. The growth of this line was somewhat delayed compared to wild-type but faster than the pMBXS491 line.

[0116] The 19% dwt PHB produced in a single seed obtained from Camelina plants transformed with construct pMBXS490 was an unexpected result and is the highest level of PHB reported in oilseeds to date. Previous studies with *Brassica napus* produced up to 7.7% dwt PHB. These seeds were obtained from transformation of *Brassica napus* using stem segments as the explants and selection of the transformed explants (Fry, J. et al., 1987, 6, 321-325) using glyphosate resistance obtained from expression of a gene encoding 5-enolpyruvylshikimate-3-phosphate synthase. Researchers did not report any germination issues with seeds isolated from the transformed plants [Houmiel et al., 1999, Planta, 209, 547-550; Valentin et al., 1999, Int. J. Biol. Macromol. 25, 303-306].

[0117] The use of DsRed as a visual marker in Camelina enabled the identification of high PHB producing seeds that would not have germinated in a typical seed screening procedure where an antibiotic or herbicide selectable marker, such as glyphosate resistance, is employed to provide resistance to the selection agent during seed germination and seedling development in tissue culture medium.

Example 4

Transformation of *Brassica napus, Brassica carinata*, and *Brassica juncea*

[0118] Transformation of *Brassica carinata*

[0119] *Brassica carinata* can be transformed using a previously described floral dip method (Shiv et al., 2008, Journal of Plant Biochemistry and Biotechnology 17, 1-4). Briefly constructs of interest are transformed into *Agrobacterium* strain GV-3101 and cells are grown in liquid medium. Cells are harvested and resuspended in a transformation medium consisting of ½ MS salts, 5% sucrose, and 0.05% Silwet L-77. *Brassica* carinata plants are grown in a greenhouse until inflorescences develop and approximately 25% of their flowers are opened. Plants are submerged in the prepared *Agrobacterium* solution for approximately 1 minute, and covered for 24 hours. Plants are returned to the greenhouse and allowed to set seed. Transformed seeds are screened by picking DsRed seeds under the appropriate wavelength of light as described above.

[0120] Transformation of *Brassica napus*

[0121] *Brassica* seeds are surface sterilized in 10% commercial bleach (Javex, Colgate-Palmolive) for 30 min with gentle shaking. The seeds are washed three times in sterile distilled water and placed in germination medium comprising Murashige-Skoog (MS) salts and vitamins, 3% (w/v) sucrose and 0.7% (w/v) phytagar, pH 5.8 at a density of 20 per plate and maintained at 24° C. an a 16 h light/8 h dark photoperiod at a light intensity of 60-80 $\mu Em^{-2} s^{-1}$ for 4-5 days.

[0122] Constructs of interest are introduced into *Agrobacterium tumefacians* strain EHA101 (Hood et. al., 1986, J. Bacteriol. 168: 1291-1301) by electroporation. Prior to transformation of cotyledonary petioles, single colonies of strain EHA101 harboring each construct are grown in 5 ml of minimal medium supplemented with appropriate antibiotics for 48 hr at 28° C. One ml of bacterial suspension was pelleted by centrifugation for 1 min in a microfuge. The pellet was resuspended in 1 ml minimal medium.

[0123] For transformation, cotyledons are excised from 4 or in some cases 5 day old seedlings so that they included ~2 mm of petiole at the base. Individual cotyledons with the cut surface of their petioles are immersed in diluted bacterial suspension for 1 s and immediately embedded to a depth of ~2 mm in co-cultivation medium, MS medium with 3% (w/v) sucrose and 0.7% phytagar and enriched with 20 μM benzyladenine. The inoculated cotyledons are plated at a density of 10 per plate and incubated under the same growth conditions for 48 h. After co-cultivation, the cotyledons are transferred to regeneration medium comprising MS medium supplemented with 3% sucrose, 20 μM benzyladenine, 0.7% (w/v) phytagar, pH 5.8, 300 mg/L timentinin and 20 mg/L kanamycin sulfate.

[0124] After 2-3 weeks regenerant shoots obtained are cut and maintained on "shoot elongation" medium (MS medium containing, 3% sucrose, 300 mg/L timentin, 0.7% (w/v) phytagar, 300 mg/L timentinin and 20 mg/L kanamycin sulfate, pH 5.8) in Magenta jars. The elongated shoots are transferred to "rooting" medium comprising MS medium, 3% sucrose, 2 mg/L indole butyric acid, 0.7% phytagar and 500 mg/L carbenicillin. After roots emerge, plantlets are transferred to potting mix (Redi Earth, W.R. Grace and Co.). The plants are maintained in a misting chamber (75% relative humidity) under the same growth conditions. Plants are allowed to self pollinate to produce seeds. Seeds are screened by visualization of DsRed as described above.

[0125] *Brassica napus* can also be transformed using the floral dip procedure described by Shiv et al. (Shiv et al., 2008*, Journal of Plant Biochemistry and Biotechnology* 17, 1-4) as described above for *Brassica* carinata.

[0126] Transformation of *Brassica juncea*

[0127] *Brassica juncea* can be transformed using hypocotyl explants according to the methods described by Barfield and Pua (Barfield and Pua, Plant Cell Reports, 10, 308-314) or Pandian et al. (Pandian, et al., 2006*, Plant Molecular Biology Reporter* 24: 103a-103i) as follows.

[0128] *B. juncea* seeds are sterilized 2 min in 70% (v/v) ethanol and washed for 20 min in 25% commercial bleach (10 g/L hypochlorite). Seeds are rinsed 3× in sterile water. Surface-sterilized seeds are plated on germination medium (1× MS salts, 1×MS vitamins, 30 g/L sucrose, 500 mg/L MES. pH 5.5) and kept in the cold room for 2 days. Seeds are incubated for 4-6 days at 24° C. under low light (20 $\mu m\ m^{-1}s^{-1}$). Hypocotyl segments are excised and rinsed in 50 mL of callus induction medium (1× MS salts, 1× B5 vitamins, 30 g/L sucrose, 500 mg/L MES, 1.0 mg/L 2.4-D, 1.0 mg/L kinetin pH 5.8) for 30 min without agitation. This procedure is repeated but with agitation on orbital shaker (~140 g) for 48 h at 24° C. in low light (10 $\mu m\ m^{-1}s^{-1}$).

[0129] *Agrobacterium* can be prepared as follows: Cells of *Agrobacterium* strain AGL1 (Lazo, G. et al. (1991), *Biotechnology,* 9: 963-967) containing the construct of interest are grown in 5 mL of LB medium with appropriate antibiotic at 28° C. for 2 days. The 5 mL culture is transferred to 250 mL flask with 45 mL of LB and cultured for 4 h at 28° C. Cells is pelleted and resuspended in BM medium (1× MS salts, 1× B5 vitamins, 30 g/L sucrose, 500 mg/L MES, pH 5.8). The optical density at 600 nm is adjusted to 0.2 with BM medium and used for inoculation.

[0130] Explants are cocultivated with *Agrobacterium* for 20 min after which time the *Agrobacterium* suspension is removed. Hypocotyl explants are washed once in callus induction medium after which cocultivation proceeds for 48 h with gentle shaking on orbital shaker. After several washes in CIM, explants are transferred to selective shoot-inducing medium (500 mg/L AgNO2, 0.4 mg/L zeatin riboside, 2.0 mg/L benzylamino purine, 0.01 mg/L GA, 200 mg/L Timentin appropriate selection agent and 8 g/L agar added to basal medium) plates for regeneration at 24° C. Root formation is induced on root-inducing medium (0.5×MS salts, 0.5× B5 vitamins, 10 g/L sucrose, 500 mg/L MES, 0.1 mg/L indole-3-butyric acid, 200 mg/L Timentin, appropriate selection agent and S g/L agar, pH 5.8).

[0131] Plantlets are transferred to are removed from agar, gently washed, and transferred to potting soil in pots. Plants are grown in a humid environment for a week and then transferred to the greenhouse.

Example 5

Production of Hybrid Lines that are not Capable of Germinating

[0132] In previous experiments in *Arabidopsis*, lower levels of PHB were obtained when lines expressing individual PHB genes were crossed to produce a plant containing the entire PHB biosynthetic pathway (Nawrath, C., Y. Poirier, et al., 1994, Proc. Natl. Acad. Sci. USA 91, 12760-12764) than when multi-gene constructs containing the entire PHB biosynthetic pathway were constructed and transformed (Bohmert, K., I. et al., 2000, Planta 211, 841-845;U.S. Pat. No. 6,448,473). This observation led to the subsequent predominant use of multi-gene constructs for PHB production in plants. However, in some scenarios, it may be advantageous to insert a multi-gene pathway into the plant by crossing of lines containing portions of the pathway to produce hybrid plants in which the entire pathway has been reconstructed. This is especially the case when high levels of product in a seed compromises the ability of the seed to germinate or the resulting seedling to survive under normal soil growth conditions. Hybrid lines can be created by crossing a line containing one or more PHB genes with a line containing the other gene(s) needed to complete the PHB biosynthetic pathway. Use of lines that possess cytoplasmic male sterility (Esser, K. et al., 2006, Progress in Botany, Springer Berlin Heidelberg. 67, 31-52) with the appropriate maintainer and restorer lines allows these hybrid lines to be produced efficiently. Cytoplasmic male sterility systems are already available for some Brassicaceae species (Esser, K. et al., 2006, Progress in Botany, Springer Berlin Heidelberg. 67, 31-52). These Brassicaceae species can be used as gene sources to produce cytoplasmic male sterility systems for other oilseeds of interest such as Camelina. Cytoplasmic male sterility has also been reported upon expression of a β-ketothiolase from the chloroplast genome in tobacco (Ruiz, O. N. and H. Daniell, 2005, Plant Physiol. 138, 1232-1246). Male sterility has also been reported upon expression of the faoA gene encoding the α-subunit of the fatty acid β-oxidation complex from *Pseudomonas putida* (U.S. Pat. No. 6,586,658).

[0133] High PHB producing lines that are not capable of germination can be produced using oilseed lines that possess cytoplasmic male sterility (CMS) controlled by an extra-nuclear genome (i.e. mitochondria or chloroplast). The male sterile line is typically maintained by crossing with a maintainer line that is genetically identical except that it possesses normal fertile cytoplasm and is therefore male fertile. Transformation of the maintainer line with one or more genes for the PHB biosynthetic pathway and crossing this modified maintainer line with the original male sterile line will produce a male sterile line possessing a portion of the PHB biosynthetic pathway. In this example, insertion of the phaA and phaC genes into the maintainer line and crossing with the original male cytoplasmic sterile line will form a male sterile line containing the phaA and phaC genes.

[0134] Fertility can be restored to this line using a "restorer line" that carries the appropriate nuclear restorer genes. Alternatively, the restorer line can be transformed with the remaining genes required to complete the PHB biosynthetic pathway and crossed with the previously created male sterile line containing phaA and phaC to produce a hybrid line containing the entire PHB biosynthetic pathway.

[0135] Crosses can be performed in the field by planting multiple rows of the male sterile line, the line that will produce the seed, next to a few rows of the male fertile line. Harvested seed can be used for subsequent plantings or as the PHB containing seed for crushing and extraction. When expression cassettes for the PHB genes in this example are controlled by strong promoters, such as the soybean oleosin promoter, high PHB producing seeds generated in this manner will possess weak seedlings upon germination and will not be able to survive field conditions under normal growth circumstances unless treated with a material that promotes seedling strength/vigor. This adds a level of gene containment.

[0136] Cytoplasmic male sterility systems are already available for some Brassicaceae species (Esser, K., 2006, Progress in Botany, Springer Berlin Heidelberg. 67, 31-52). These Brassicaceae species can be used as gene sources to produce cytoplasmic male sterility systems for other oilseeds of interest such as Camelina. Cytoplasmic male sterility has also been reported upon expression of a β-ketothiolase from the chloroplast genome in tobacco (Ruiz, O. N. and H. Daniell, 2005, Plant Physiol. 138, 1232-1246). Overexpression of β-ketothiolase in Camelina to generate a male sterile line and subsequent crossing with a line expressing phaB and phaC could also be used for hybrid seed production.

[0137] Male sterile lines have also been produced in *Brassica napus* by overexpression of the faoA gene from *Pseudomonas putida* under the control of the phaseolin promoter sequence (U.S. Pat. No. 6,586,658).

[0138] Double haploid technology can be used to speed up the breeding process. In the double haploid technique, immature pollen grains (haploids) are exposed to treatments that result in doubling of the existing genetic material resulting in homozygous, true breeding material in a single generation.

[0139] The references, patents, and patent applications cited throughout are incorporated by reference where permissible in their entireties.

```
Vector: pMBXS490
                                                    (SEQ ID NO: 1)
  1 GGGGATCCGT ACGTAAGTAC GTACTCAAAA TGCCAACAAA TAAAAAAAAA

 51 GTTGCTTTAA TAATGCCAAA ACAAATTAAT AAAACACTTA CAACACCGGA

101 TTTTTTTTAA TTAAAATGTG CCATTTAGGA TAAATAGTTA ATATTTTTAA

151 TAATTATTTA AAAAGCCGTA TCTACTAAAA TGATTTTTAT TTGGTTGAAA

201 ATATTAATAT GTTTAAATCA ACACAATCTA TCAAAATTAA ACTAAAAAAA

251 AAATAAGTGT ACGTGGTTAA CATTAGTACA GTAATATAAG AGGAAAATGA

301 GAAATTAAGA AATTGAAAGC GAGTCTAATT TTTAAATTAT GAACCTGCAT

351 ATATAAAAGG AAAGAAAGAA TCCAGGAAGA AAAGAAATGA AACCATGCAT

401 GGTCCCCTCG TCATCACGAG TTTCTGCCAT TTGCAATAGA AACACTGAAA

451 CACCTTTCTC TTTGTCACTT AATTGAGATG CCGAAGCCAC CTCACACCAT

501 GAACTTCATG AGGTGTAGCA CCCAAGGCTT CCATAGCCAT GCATACTGAA

551 GAATGTCTCA AGCTCAGCAC CCTACTTCTG TGACGTGTCC CTCATTCACC

601 TTCCTCTCTT CCCTATAAAT AACCACGCCT CAGGTTCTCC GCTTCACAAC

651 TCAAACATTC TCTCCATTGG TCCTTAAACA CTCATCAGTC ATCACCGCGG

701 CCGCGGAATT CATGGCTTCT ATGATATCCT CTTCCGCTGT GACAACAGTC

751 AGCCGTGCCT CTAGGGGGCA ATCCGCCGCA GTGGCTCCAT TCGGCGGCCT

801 CAAATCCATG ACTGGATTCC CAGTGAAGAA GGTCAACACT GACATTACTT

851 CCATTACAAG CAATGGTGGA AGAGTAAAGT GCATGCAGGT GTGGCCTCCA
```

-continued

```
 901 ATTGGAAAGA AGAAGTTTGA GACTCTTTCC TATTTGCCAC CATTGACGAG
 951 AGATTCTAGA GTGACTGACG TTGTCATCGT ATCCGCCGCC CGCACCGCGG
1001 TCGGCAAGTT TGGCGGCTCG CTGGCCAAGA TCCCGGCACC GGAACTGGGT
1051 GCCGTGGTCA TCAAGGCCGC GCTGGAGCGC GCCGGCGTCA AGCCGGAGCA
1101 GGTGAGCGAA GTCATCATGG GCCAGGTGCT GACCGCCGGT TCGGGCCAGA
1151 ACCCCGCACG CCAGGCCGCG ATCAAGGCCG GCCTGCCGGC GATGGTGCCG
1201 GCCATGACCA TCAACAAGGT GTGCGGCTCG GGCCTGAAGG CCGTGATGCT
1251 GGCCGCCAAC GCGATCATGG CGGGCGACGC CGAGATCGTG GTGGCCGGCG
1301 GCCAGGAAAA CATGAGCGCC GCCCCGCACG TGCTGCCGGG CTCGCGCGAT
1351 GGTTTCCGCA TGGGCGATGC CAAGCTGGTC GACACCATGA TCGTCGACGG
1401 CCTGTGGGAC GTGTACAACC AGTACCACAT GGGCATCACC GCCGAGAACG
1451 TGGCCAAGGA ATACGGCATC ACACGCGAGG CGCAGGATGA GTTCGCCGTC
1501 GGCTCGCAGA ACAAGGCCGA AGCCGCGCAG AAGGCCGGCA AGTTTGACGA
1551 AGAGATCGTC CCGGTGCTGA TCCCGCAGCG CAAGGGCGAC CCGGTGGCCT
1601 TCAAGACCGA CGAGTTCGTG CGCCAGGGCG CCACGCTGGA CAGCATGTCC
1651 GGCCTCAAGC CCGCCTTCGA CAAGGCCGGC ACGGTGACCG CGGCCAACGC
1701 CTCGGGCCTG AACGACGGCG CCGCCGCGGT GGTGGTGATG TCGGCGGCCA
1751 AGGCCAAGGA ACTGGGCCTG ACCCCGCTGG CCACGATCAA GAGCTATGCC
1801 AACGCCGGTG TCGATCCCAA GGTGATGGGC ATGGGCCCGG TGCCGGCCTC
1851 CAAGCGCGCC CTGTCGCGCG CCGAGTGGAC CCCGCAAGAC CTGGACCTGA
1901 TGGAGATCAA CGAGGCCTTT GCCGCGCAGG CGCTGGCGGT GCACCAGCAG
1951 ATGGGCTGGG ACACCTCCAA GGTCAATGTG AACGGCGGCG CCATCGCCAT
2001 CGGCCACCCG ATCGGCGCGT CGGGCTGCCG TATCCTGGTG ACGCTGCTGC
2051 ACGAGATGAA GCGCCGTGAC GCGAAGAAGG GCCTGGCCTC GCTGTGCATC
2101 GGCGGCGGCA TGGGCGTGGC GCTGGCAGTC GAGCGCAAAT AACTCGAGGC
2151 GGCCGCAGCC CTTTTTGTAT GTGCTACCCC ACTTTTGTCT TTTTGGCAAT
2201 AGTGCTAGCA ACCAATAAAT AATAATAATA ATAATGAATA AGAAAACAAA
2251 GGCTTTAGCT TGCCTTTTGT TCACTGTAAA ATAATAATGT AAGTACTCTC
2301 TATAATGAGT CACGAAACTT TTGCGGGAAT AAAAGGAGAA ATTCCAATGA
2351 GTTTTCTGTC AAATCTTCTT TTGTCTCTCT CTCTCTCTCT TTTTTTTTTT
2401 TCTTTCTTCT GAGCTTCTTG CAAAACAAAA GGCAAACAAT AACGATTGGT
2451 CCAATGATAG TTAGCTTGAT CGATGATATC TTTAGGAAGT GTTGGCAGGA
2501 CAGGACATGA TGTAGAAGAC TAAAATTGAA AGTATTGCAG ACCCAATAGT
2551 TGAAGATTAA CTTTAAGAAT GAAGACGTCT TATCAGGTTC TTCATGACTT
2601 AAGCTTTAAG AGGAGTCCAC CATGGTAGAT CTGACTAGTA GAAGGTAATT
2651 ATCCAAGATG TAGCATCAAG AATCCAATGT TTACGGGAAA AACTATGGAA
2701 GTATTATGTG AGCTCAGCAA GAAGCAGATC AATATGCGGC ACATATGCAA
2751 CCTATGTTCA AAAATGAAGA ATGTACAGAT ACAAGATCCT ATACTGCCAG
2801 AATACGAAGA AGAATACGTA GAAATTAAGA AAGAAGAACC AGGCGAAGAA
```

-continued

```
2851 AAGAATCTTG AAGACGTAAG CACTGACGAC AACACTGAAA AGAAGAAGAT
2901 AAGGTCGGTG ATTGTGAAAG AGACATAGAG GACACATGTA AGGTGGAAAA
2951 TGTAAGGGCG GAAAGTAACC TTATCACAAA GGAATCTTAT CCCCCACTAC
3001 TTATCCTTTT ATATTTTTCC GTGTCATTTT TGCCCTTGAG TTTTCCTATA
3051 TAAGGAACCA AGTTCGGCAT TTGTGAAAAC AAGAAAAAAT TGGTGTAAGC
3101 TATTTTCTTT GAAGTACTGA GGATACAACT TCAGAGAAAT TTGTAAGAAA
3151 GTGGATCGAA ACCATGGCCT CCTCCGAGAA CGTCATCACC GAGTTCATGC
3201 GCTTCAAGGT GCGCATGGAG GGCACCGTGA ACGGCCACGA GTTCGAGATC
3251 GAGGGCGAGG GCGAGGGCCG CCCCTACGAG GGCCACAACA CCGTGAAGCT
3301 GAAGGTGACC AAGGGCGGCC CCCTGCCCTT CGCCTGGGAC ATCCTGTCCC
3351 CCCAGTTCCA GTACGGCTCC AAGGTGTACG TGAAGCACCC CGCCGACATC
3401 CCCGACTACA AGAAGCTGTC CTTCCCCGAG GGCTTCAAGT GGGAGCGCGT
3451 GATGAACTTC GAGGACGGCG GCGTGGCGAC CGTGACCCAG GACTCCTCCC
3501 TGCAGGACGG CTGCTTCATC TACAAGGTGA AGTTCATCGG CGTGAACTTC
3551 CCCTCCGACG GCCCCGTGAT GCAGAAGAAG ACCATGGGCT GGGAGGCCTC
3601 CACCGAGCGC CTGTACCCCC GCGACGGCGT GCTGAAGGGC GAGACCCACA
3651 AGGCCCTGAA GCTGAAGGAC GGCGGCCACT ACCTGGTGGA GTTCAAGTCC
3701 ATCTACATGG CCAAGAAGCC CGTGCAGCTG CCCGGCTACT ACTACGTGGA
3751 CGCCAAGCTG GACATCACCT CCCACAACGA GGACTACACC ATCGTGGAGC
3801 AGTACGAGCG CACCGAGGGC CGCCACCACC TGTTCCTGGT ACCAATGAGC
3851 TCTGTCCAAC AGTCTCAGGG TTAATGTCTA TGTATCTTAA ATAATGTTGT
3901 CGGCGATCGT TCAAACATTT GGCAATAAAG TTTCTTAAGA TTGAATCCTG
3951 TTGCCGGTCT TGCGATGATT ATCATATAAT TTCTGTTGAA TTACGTTAAG
4001 CATGTAATAA TTAACATGTA ATGCATGACG TTATTTATGA GATGGGTTTT
4051 TATGATTAGA GTCCCGCAAT TATACATTTA ATACGCGATA GAAAACAAAA
4101 TATAGCGCGC AAACTAGGAT AAATTATCGC GCGCGGTGTC ATCTATGTTA
4151 CTAGATCGGG AATTAAACTA TCAGTGTTTG ACAGGATATA TTGGCGGGTA
4201 AACCTAAGAG AAAAGAGCGT TTATTAGAAT AACGGATATT TAAAAGGGCG
4251 TGAAAAGGTT TATCCGTTCG TCCATTTGTA TGTGCATGCC AACCACAGGG
4301 TTCCCCTCGG GATCAAAGTA CTTTGATCCA ACCCCTCCGC TGCTATAGTG
4351 CAGTCGGCTT CTGACGTTCA GTGCAGCCGT CTTCTGAAAA CGACATGTCG
4401 CACAAGTCCT AAGTTACGCG ACAGGCTGCC GCCCTGCCCT TTTCCTGGCG
4451 TTTTCTTGTC GCGTGTTTTA GTCGCATAAA GTAGAATACT TGCGACTAGA
4501 ACCGGAGACA TTACGCCATG AACAAGAGCG CCGCCGCTGG CCTGCTGGGC
4551 TATGCCCGCG TCAGCACCGA CGACCAGGAC TTGACCAACC AACGGGCCGA
4601 ACTGCACGCG GCCGGCTGCA CCAAGCTGTT TTCCGAGAAG ATCACCGGCA
4651 CCAGGCGCGA CCGCCCGGAG CTGGCCAGGA TGCTTGACCA CCTACGCCCT
4701 GGCGACGTTG TGACAGTGAC CAGGCTAGAC CGCCTGGCCC GCAGCACCCG
```

-continued

4751 CGACCTACTG GACATTGCCG AGCGCATCCA GGAGGCCGGC GCGGGCCTGC

4801 GTAGCCTGGC AGAGCCGTGG GCCGACACCA CCACGCCGGC CGGCCGCATG

4851 GTGTTGACCG TGTTCGCCGG CATTGCCGAG TTCGAGCGTT CCCTAATCAT

4901 CGACCGCACC CGGAGCGGGC GCGAGGCCGC CAAGGCCCGA GGCGTGAAGT

4951 TTGGCCCCCG CCCTACCCTC ACCCCGGCAC AGATCGCGCA CGCCCGCGAG

5001 CTGATCGACC AGGAAGGCCG CACCGTGAAA GAGGCGGCTG CACTGCTTGG

5051 CGTGCATCGC TCGACCCTGT ACCGCGCACT TGAGCGCAGC GAGGAAGTGA

5101 CGCCCACCGA GGCCAGGCGG CGCGGTGCCT TCCGTGAGGA CGCATTGACC

5151 GAGGCCGACG CCCTGGCGGC CGCCGAGAAT GAACGCCAAG AGGAACAAGC

5201 ATGAAACCGC ACCAGGACGG CCAGGACGAA CCGTTTTTCA TTACCGAAGA

5251 GATCGAGGCG GAGATGATCG CGGCCGGGTA CGTGTTCGAG CCGCCCGCGC

5301 ACGTCTCAAC CGTGCGGCTG CATGAAATCC TGGCCGGTTT GTCTGATGCC

5351 AAGCTGGCGG CCTGGCCGGC CAGCTTGGCC GCTGAAGAAA CCGAGCGCCG

5401 CCGTCTAAAA AGGTGATGTG TATTTGAGTA AAACAGCTTG CGTCATGCGG

5451 TCGCTGCGTA TATGATGCGA TGAGTAAATA AACAAATACG CAAGGGGAAC

5501 GCATGAAGGT TATCGCTGTA CTTAACCAGA AAGGCGGGTC AGGCAAGACG

5551 ACCATCGCAA CCCATCTAGC CCGCGCCCTG CAACTCGCCG GGGCCGATGT

5601 TCTGTTAGTC GATTCCGATC CCCAGGGCAG TGCCCGCGAT TGGGCGGCCG

5651 TGCGGGAAGA TCAACCGCTA ACCGTTGTCG GCATCGACCG CCCGACGATT

5701 GACCGCGACG TGAAGGCCAT CGGCCGGCGC GACTTCGTAG TGATCGACGG

5751 AGCGCCCCAG GCGGCGGACT TGGCTGTGTC CGCGATCAAG GCAGCCGACT

5801 TCGTGCTGAT TCCGGTGCAG CCAAGCCCTT ACGACATATG GGCCACCGCC

5851 GACCTGGTGG AGCTGGTTAA GCAGCGCATT GAGGTCACGG ATGGAAGGCT

5901 ACAAGCGGCC TTTGTCGTGT CGCGGGCGAT CAAAGGCACG CGCATCGGCG

5951 GTGAGGTTGC CGAGGCGCTG GCCGGGTACG AGCTGCCCAT TCTTGAGTCC

6001 CGTATCACGC AGCGCGTGAG CTACCCAGGC ACTGCCGCCG CCGGCACAAC

6051 CGTTCTTGAA TCAGAACCCG AGGGCGACGC TGCCCGCGAG GTCCAGGCGC

6101 TGGCCGCTGA AATTAAATCA AAACTCATTT GAGTTAATGA GGTAAAGAGA

6151 AAATGAGCAA AAGCACAAAC ACGCTAAGTG CCGGCCGTCC GAGCGCACGC

6201 AGCAGCAAGG CTGCAACGTT GGCCAGCCTG GCAGACACGC CAGCCATGAA

6251 GCGGGTCAAC TTTCAGTTGC CGGCGGAGGA TCACACCAAG CTGAAGATGT

6301 ACGCGGTACG CCAAGGCAAG ACCATTACCG AGCTGCTATC TGAATACATC

6351 GCGCAGCTAC CAGAGTAAAT GAGCAAATGA ATAAATGAGT AGATGAATTT

6401 TAGCGGCTAA AGGAGGCGGC ATGGAAAATC AAGAACAACC AGGCACCGAC

6451 GCCGTGGAAT GCCCCATGTG TGGAGGAACG GGCGGTTGGC CAGGCGTAAG

6501 CGGCTGGGTT GTCTGCCGGC CCTGCAATGG CACTGGAACC CCCAAGCCCG

6551 AGGAATCGGC GTGACGGTCG CAAACCATCC GGCCCGGTAC AAATCGGCGC

6601 GGCGCTGGGT GATGACCTGG TGGAGAAGTT GAAGGCCGCG CAGGCCGCCC

6651 AGCGGCAACG CATCGAGGCA GAAGCACGCC CCGGTGAATC GTGGCAAGCG

-continued

6701 GCCGCTGATC GAATCCGCAA AGAATCCCGG CAACCGCCGG CAGCCGGTGC

6751 GCCGTCGATT AGGAAGCCGC CCAAGGGCGA CGAGCAACCA GATTTTTTCG

6801 TTCCGATGCT CTATGACGTG GGCACCCGCG ATAGTCGCAG CATCATGGAC

6851 GTGGCCGTTT TCCGTCTGTC GAAGCGTGAC CGACGAGCTG GCGAGGTGAT

6901 CCGCTACGAG CTTCCAGACG GGCACGTAGA GGTTTCCGCA GGGCCGGCCG

6951 GCATGGCCAG TGTGTGGGAT TACGACCTGG TACTGATGGC GGTTTCCCAT

7001 CTAACCGAAT CCATGAACCG ATACCGGGAA GGGAAGGGAG ACAAGCCCGG

7051 CCGCGTGTTC CGTCCACACG TTGCGGACGT ACTCAAGTTC TGCCGGCGAG

7101 CCGATGGCGG AAAGCAGAAA GACGACCTGG TAGAAACCTG CATTCGGTTA

7151 AACACCACGC ACGTTGCCAT GCAGCGTACG AAGAAGGCCA AGAACGGCCG

7201 CCTGGTGACG GTATCCGAGG GTGAAGCCTT GATTAGCCGC TACAAGATCG

7251 TAAAGAGCGA AACCGGGCGG CCGGAGTACA TCGAGATCGA GCTAGCTGAT

7301 TGGATGTACC GCGAGATCAC AGAAGGCAAG AACCCGGACG TGCTGACGGT

7351 TCACCCCGAT TACTTTTTGA TCGATCCCGG CATCGGCCGT TTTCTCTACC

7401 GCCTGGCACG CCGCGCCGCA GGCAAGGCAG AAGCCAGATG GTTGTTCAAG

7451 ACGATCTACG AACGCAGTGG CAGCGCCGGA GAGTTCAAGA AGTTCTGTTT

7501 CACCGTGCGC AAGCTGATCG GGTCAAATGA CCTGCCGGAG TACGATTTGA

7551 AGGAGGAGGC GGGGCAGGCT GGCCCGATCC TAGTCATGCG CTACCGCAAC

7601 CTGATCGAGG GCGAAGCATC CGCCGGTTCC TAATGTACGG AGCAGATGCT

7651 AGGGCAAATT GCCCTAGCAG GGGAAAAAGG TCGAAAAGGT CTCTTTCCTG

7701 TGGATGTACC GTACATTGGG AACCCAAAGC CGTACATTGG GAACCGGAAC

7751 CCGTACATTG GGAACCCAAA GCCGTACATT GGGAACCGGT CACACATGTA

7801 AGTGACTGAT ATAAAAGAGA AAGAAGGCCA TTTTTCCGCC TAAAACTCTT

7851 TAAAACTTAT TAAAACTCTT AAAACCCGCC TGGCCTGTGC ATAACTGTCT

7901 GGCCAGCGCA CAGCCGAAGA GCTGCAAAAA GCGCCTACCC TTCGGTCGCT

7951 GCGCTCCCTA CGCCCCGCCG CTTCGCGTCG GCCTATCGCG GCCGCTGGCC

8001 GCTCAAAAAT GGCTGGCCTA CGGCCAGGCA ATCTACCAGG GCGCGGACAA

8051 GCCGCGCCGT CGCCACTCGA CCGCCGGCGC CCACATCAAG GCACCCTGCC

8101 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG

8151 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG

8201 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCGCA GCCATGACCC

8251 AGTCACGTAG CGATAGCGGA GTGTATACTG GCTTAACTAT GCGGCATCAG

8301 AGCAGATTGT ACTGAGAGTG CACCATATGC GGTGTGAAAT ACCGCACAGA

8351 TGCGTAAGGA GAAAATACCG CATCAGGCGC TCTTCCGCTT CCTCGCTCAC

8401 TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT

8451 CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG

8501 AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC

8551 GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCCTGACGAG CATCACAAAA

-continued

```
8601 ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC
8651 CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT
8701 GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC
8751 TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC
8801 TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC
8851 CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT
8901 CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA
8951 GGCTGGCCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG
9001 AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA
9051 AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT
9101 GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA
9151 AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA
9201 ACTCACGTTA AGGGATTTTG GTCATGCATT CTAGGTACTA AAACAATTCA
9251 TCCAGTAAAA TATAATATTT TATTTTCTCC CAATCAGGCT TGATCCCCAG
9301 TAAGTCAAAA AATAGCTCGA CATACTGTTC TTCCCCGATA TCCTCCCTGA
9351 TCGACCGGAC GCAGAAGGCA ATGTCATACC ACTTGTCCGC CCTGCCGCTT
9401 CTCCCAAGAT CAATAAAGCC ACTTACTTTG CCATCTTTCA CAAAGATGTT
9451 GCTGTCTCCC AGGTCGCCGT GGGAAAAGAC AAGTTCCTCT TCGGGCTTTT
9501 CCGTCTTTAA AAAATCATAC AGCTCGCGCG GATCTTTAAA TGGAGTGTCT
9551 TCTTCCCAGT TTTCGCAATC CACATCGGCC AGATCGTTAT TCAGTAAGTA
9601 ATCCAATTCG GCTAAGCGGC TGTCTAAGCT ATTCGTATAG GGACAATCCG
9651 ATATGTCGAT GGAGTGAAAG AGCCTGATGC ACTCCGCATA CAGCTCGATA
9701 ATCTTTTCAG GGCTTTGTTC ATCTTCATAC TCTTCCGAGC AAAGGACGCC
9751 ATCGGCCTCA CTCATGAGCA GATTGCTCCA GCCATCATGC CGTTCAAAGT
9801 GCAGGACCTT TGGAACAGGC AGCTTTCCTT CCAGCCATAG CATCATGTCC
9851 TTTTCCCGTT CCACATCATA GGTGGTCCCT TTATACCGGC TGTCCGTCAT
9901 TTTTAAATAT AGGTTTTCAT TTTCTCCCAC CAGCTTATAT ACCTTAGCAG
9951 GAGACATTCC TTCCGTATCT TTTACGCAGC GGTATTTTTC GATCAGTTTT
10001 TTCAATTCCG GTGATATTCT CATTTTAGCC ATTTATTATT TCCTTCCTCT
10051 TTTCTACAGT ATTTAAAGAT ACCCCAAGAA GCTAATTATA ACAAGACGAA
10101 CTCCAATTCA CTGTTCCTTG CATTCTAAAA CCTTAAATAC CAGAAAACAG
10151 CTTTTTCAAA GTTGTTTTCA AAGTTGGCGT ATAACATAGT ATCGACGGAG
10201 CCGATTTTGA AACCGCGGTG ATCACAGGCA GCAACGCTCT GTCATCGTTA
10251 CAATCAACAT GCTACCCTCC GCGAGATCAT CCGTGTTTCA AACCCGGCAG
10301 CTTAGTTGCC GTTCTTCCGA ATAGCATCGG TAACATGAGC AAAGTCTGCC
10351 GCCTTACAAC GGCTCTCCCG CTGACGCCGT CCCGGACTGA TGGGCTGCCT
10401 GTATCGAGTG GTGATTTTGT GCCGAGCTGC CGGTCGGGGA GCTGTTGGCT
10451 GGCTGGTGGC AGGATATATT GTGGTGTAAA CAAATTGACG CTTAGACAAC
10501 TTAATAACAC ATTGCGGACG TTTTTAATGT ACTGAATTAA CGCCGAATTA
```

-continued

10551 ATTCCTAGGC CACCATGTTG GGCCCGGGGC GCGCCGTACG TAGTGTTTAT

10601 CTTTGTTGCT TTTCTGAACA ATTTATTTAC TATGTAAATA TATTATCAAT

10651 GTTTAATCTA TTTTAATTTG CACATGAATT TTCATTTTAT TTTTACTTTA

10701 CAAAACAAAT AAATATATAT GCAAAAAAAT TTACAAACGA TGCACGGGTT

10751 ACAAACTAAT TTCATTAAAT GCTAATGCAG ATTTTGTGAA GTAAAACTCC

10801 AATTATGATG AAAAATACCA CCAACACCAC CTGCGAAACT GTATCCCAAC

10851 TGTCCTTAAT AAAAATGTTA AAAAGTATAT TATTCTCATT TGTCTGTCAT

10901 AATTTATGTA CCCCACTTTA ATTTTTCTGA TGTACTAAAC CGAGGGCAAA

10951 CTGAAACCTG TTCCTCATGC AAAGCCCCTA CTCACCATGT ATCATGTACG

11001 TGTCATCACC CAACAACTCC ACTTTTGCTA TATAACAACA CCCCCGTCAC

11051 ACTCTCCCTC TCTAACACAC ACCCCACTAA CAATTCCTTC ACTTGCAGCA

11101 CTGTTGCATC ATCATCTTCA TTGCAAAACC CTAAACTTCA CCTTCAACCG

11151 CGGCCGCATG GCTTCTATGA TATCCTCTTC CGCTGTGACA ACAGTCAGCC

11201 GTGCCTCTAG GGGGCAATCC GCCGCAGTGG CTCCATTCGG CGGCCTCAAA

11251 TCCATGACTG GATTCCCAGT GAAGAAGGTC AACACTGACA TTACTTCCAT

11301 TACAAGCAAT GGTGGAAGAG TAAAGTGCAT GCAGGTGTGG CCTCCAATTG

11351 GAAAGAAGAA GTTTGAGACT CTTTCCTATT TGCCACCATT GACGAGAGAT

11401 TCTAGAGTGA GTAACAAGAA CAACGATGAG CTGCAGTGGC AATCCTGGTT

11451 CAGCAAGGCG CCCACCACCG AGGCGAACCC GATGGCCACC ATGTTGCAGG

11501 ATATCGGCGT TGCGCTCAAA CCGGAAGCGA TGGAGCAGCT GAAAAACGAT

11551 TATCTGCGTG ACTTCACCGC GTTGTGGCAG GATTTTTTGG CTGGCAAGGC

11601 GCCAGCCGTC AGCGACCGCC GCTTCAGCTC GGCAGCCTGG CAGGGCAATC

11651 CGATGTCGGC CATCATGTCC GCATCTTACC TGCTCAACGC CAAATTCCTC

11701 AGTGCCATGG TGGAGGCGGT GGACACCGCA CCCCAGCAAA AGCAGAAAAT

11751 ACGCTTTGCC GTGCAGCAGG TGATTGATGC CATGTCGCCC GCGAACTTCC

11801 TCGCCACCAA CCCGGAAGCG CAGCAAAAAC TGATTGAAAC CAAGGGCGAG

11851 AGCCTGACGC GTGGCCTGGT CAATATGCTG GGCGATATCA ACAAGGGCCA

11901 TATCTCGCTG TCGGACGAAT CGGCCTTTGA AGTGGGCCGC AACCTGGCCA

11951 TTACCCCGGG CACCGTGATT TACGAAAATC CGCTGTTCCA GCTGATCCAG

12001 TACACGCCGA CCACGCCGAC GGTCAGCCAG CGCCCGCTGT TGATGGTGCC

12051 GCCGTGCATC AACAAGTTCT ACATCCTCGA CCTTCAACCG GAAAATTCGC

12101 TGGTGCGCTA CGCGGTGGAG CAGGGCAACA CCGTGTTCCT GATCTCGTGG

12151 AGCAATCCGG ACAAGTCGCT GGCCGGCACC ACCTGGGACG ACTACGTGGA

12201 GCAGGGCGTG ATCGAAGCGA TCCGCATCGT CCAGGACGTC AGCGGCCAGG

12251 ACAAGCTGAA CATGTTCGGC TTCTGCGTGG GCGGCACCAT CGTTGCCACC

12301 GCACTGGCGG TACTGGCGGC GCGTGGCCAG CACCCGGCGG CCAGCCTGAC

12351 CCTGCTGACC ACCTTCCTCG ACTTCAGCGA CACCGGCGTG CTCGACGTCT

12401 TCGTCGATGA AACCCAGGTC GCGCTGCGTG AACAGCAATT GCGCGATGGC

-continued

12451 GGCCTGATGC CGGGCCGTGA CCTGGCCTCG ACCTTCTCGA GCCTGCGTCC

12501 GAACGACCTG GTATGGAACT ATGTGCAGTC GAACTACCTC AAAGGCAATG

12551 AGCCGGCGGC GTTTGACCTG CTGTTCTGGA ATTCGGACAG CACCAATTTG

12601 CCGGGCCCGA TGTTCTGCTG GTACCTGCGC AACACCTACC TGGAAAACAG

12651 CCTGAAAGTG CCGGGCAAGC TGACGGTGGC CGGCGAAAAG ATCGACCTCG

12701 GCCTGATCGA CGCCCCGGCC TTCATCTACG GTTCGCGCGA AGACCACATC

12751 GTGCCGTGGA TGTCGGCGTA CGGTTCGCTC GACATCCTCA ACCAGGGCAA

12801 GCCGGGCGCC AACCGCTTCG TGCTGGGCGC GTCCGGCCAT ATCGCCGGCG

12851 TGATCAACTC GGTGGCCAAG AACAAGCGCA GCTACTGGAT CAACGACGGT

12901 GGCGCCGCCG ATGCCCAGGC CTGGTTCGAT GGCGCGCAGG AAGTGCCGGG

12951 CAGCTGGTGG CCGCAATGGG CCGGGTTCCT GACCCAGCAT GGCGGCAAGA

13001 AGGTCAAGCC CAAGGCCAAG CCCGGCAACG CCCGCTACAC CGCGATCGAG

13051 GCGGCGCCCG GCCGTTACGT CAAAGCCAAG GGCTGAGCGG CCGCTGAGTA

13101 ATTCTGATAT TAGAGGGAGC ATTAATGTGT TGTTGTGATG TGGTTTATAT

13151 GGGGAAATTA AATAAATGAT GTATGTACCT CTTGCCTATG TAGGTTTGTG

13201 TGTTTTGTTT TGTTGTCTAG CTTTGGTTAT TAAGTAGTAG GGACGTTCGT

13251 TCGTGTCTCA AAAAAAGGGG TACTACCACT CTGTAGTGTA TATGGATGCT

13301 GGAAATCAAT GTGTTTTGTA TTTGTTCACC TCCATTGTTG AATTCAATGT

13351 CAAATGTGTT TTGCGTTGGT TATGTGTAAA ATTACTATCT TTCTCGTCCG

13401 ATGATCAAAG TTTTAAGCAA CAAAACCAAG GGTGAAATTT AAACTGTGCT

13451 TTGTTGAAGA TTCTTTTATC ATATTGAAAA TCAAATTACT AGCAGCAGAT

13501 TTTACCTAGC ATGAAATTTT ATCAACAGTA CAGCACTCAC TAACCAAGTT

13551 CCAAACTAAG ATGCGCCATT AACATCAGCC AATAGGCATT TTCAGCAAGG

13601 CGCGCCCGCG CCGATGTATG TGACAACCCT CGGGATTGTT GATTTATTTC

13651 AAAACTAAGA GTTTTTGTCT TATTGTTCTC GTCTATTTTG GATATCAATC

13701 TTAGTTTTAT ATCTTTTCTA GTTCTCTACG TGTTAAATGT TCAACACACT

13751 AGCAATTTGG CCTGCCAGCG TATGGATTAT GGAACTATCA AGTCTGTGAC

13801 GCGCCGTACG TAGTGTTTAT CTTTGTTGCT TTTCTGAACA ATTTATTTAC

13851 TATGTAAATA TATTATCAAT GTTTAATCTA TTTTAATTTG CACATGAATT

13901 TTCATTTTAT TTTTACTTTA CAAAACAAAT AAATATATAT GCAAAAAAAT

13951 TTACAAACGA TGCACGGGTT ACAAACTAAT TTCATTAAAT GCTAATGCAG

14001 ATTTTGTGAA GTAAAACTCC AATTATGATG AAAAATACCA CCAACACCAC

14051 CTGCGAAACT GTATCCCAAC TGTCCTTAAT AAAAATGTTA AAAAGTATAT

14101 TATTCTCATT TGTCTGTCAT AATTTATGTA CCCCACTTTA ATTTTTCTGA

14151 TGTACTAAAC CGAGGGCAAA CTGAAACCTG TTCCTCATGC AAAGCCCCTA

14201 CTCACCATGT ATCATGTACG TGTCATCACC CAACAACTCC ACTTTTGCTA

14251 TATAACAACA CCCCCGTCAC ACTCTCCCTC TCTAACACAC ACCCCACTAA

14301 CAATTCCTTC ACTTGCAGCA CTGTTGCATC ATCATCTTCA TTGCAAAACC

14351 CTAAACTTCA CCTTCAACCG CGGCCGCATG GCTTCTATGA TATCCTCTTC

-continued

14401 CGCTGTGACA ACAGTCAGCC GTGCCTCTAG GGGGCAATCC GCCGCAGTGG

14451 CTCCATTCGG CGGCCTCAAA TCCATGACTG GATTCCCAGT GAAGAAGGTC

14501 AACACTGACA TTACTTCCAT TACAAGCAAT GGTGGAAGAG TAAAGTGCAT

14551 GCAGGTGTGG CCTCCAATTG GAAAGAAGAA GTTTGAGACT CTTTCCTATT

14601 TGCCACCATT GACGAGAGAT TCTAGAGTGA CTCAGCGCAT TGCGTATGTG

14651 ACCGGCGGCA TGGGTGGTAT CGGAACCGCC ATTTGCCAGC GGCTGGCCAA

14701 GGATGGCTTT CGTGTGGTGG CCGGTTGCGG CCCCAACTCG CCGCGCCGCG

14751 AAAAGTGGCT GGAGCAGCAG AAGGCCCTGG GCTTCGATTT CATTGCCTCG

14801 GAAGGCAATG TGGCTGACTG GGACTCGACC AAGACCGCAT TCGACAAGGT

14851 CAAGTCCGAG GTCGGCGAGG TTGATGTGCT GATCAACAAC GCCGGTATCA

14901 CCCGCGACGT GGTGTTCCGC AAGATGACCC GCGCCGACTG GGATGCGGTG

14951 ATCGACACCA ACCTGACCTC GCTGTTCAAC GTCACCAAGC AGGTGATCGA

15001 CGGCATGGCC GACCGTGGCT GGGGCCGCAT CGTCAACATC TCGTCGGTGA

15051 ACGGGCAGAA GGGCCAGTTC GGCCAGACCA ACTACTCCAC CGCCAAGGCC

15101 GGCCTGCATG GCTTCACCAT GGCACTGGCG CAGGAAGTGG CGACCAAGGG

15151 CGTGACCGTC AACACGGTCT CTCCGGGCTA TATCGCCACC GACATGGTCA

15201 AGGCGATCCG CCAGGACGTG CTCGACAAGA TCGTCGCGAC GATCCCGGTC

15251 AAGCGCCTGG GCCTGCCGGA AGAGATCGCC TCGATCTGCG CCTGGTTGTC

15301 GTCGGAGGAG TCCGGTTTCT CGACCGGCGC CGACTTCTCG CTCAACGGCG

15351 GCCTGCATAT GGGCTGAGCG GCCGCTGAGT AATTCTGATA TTAGAGGGAG

15401 CATTAATGTG TTGTTGTGAT GTGGTTTATA TGGGGAAATT AAATAAATGA

15451 TGTATGTACC TCTTGCCTAT GTAGGTTTGT GTGTTTTGTT TTGTTGTCTA

15501 GCTTTGGTTA TTAAGTAGTA GGGACGTTCG TTCGTGTCTC AAAAAAAGGG

15551 GTACTACCAC TCTGTAGTGT ATATGGATGC TGGAAATCAA TGTGTTTTGT

15601 ATTTGTTCAC CTCCATTGTT GAATTCAATG TCAAATGTGT TTTGCGTTGG

15651 TTATGTGTAA AATTACTATC TTTCTCGTCC GATGATCAAA GTTTTAAGCA

15701 ACAAAACCAA GGGTGAAATT TAAACTGTGC TTTGTTGAAG ATTCTTTTAT

15751 CATATTGAAA ATCAAATTAC TAGCAGCAGA TTTTACCTAG CATGAAATTT

15801 TATCAACAGT ACAGCACTCA CTAACCAAGT TCCAAACTAA GATGCGCCAT

15851 TAACATCAGC CAATAGGCAT TTTCAGCAAG GCGCGTAA

pMBXS364

(SEQ ID NO: 2)

1 CATGCCAACC ACAGGGTTCC CCTCGGGATC AAAGTACTTT GATCCAACCC

51 CTCCGCTGCT ATAGTGCAGT CGGCTTCTGA CGTTCAGTGC AGCCGTCTTC

101 TGAAAACGAC ATGTCGCACA AGTCCTAAGT TACGCGACAG GCTGCCGCCC

151 TGCCCTTTTC CTGGCGTTTT CTTGTCGCGT GTTTTAGTCG CATAAAGTAG

201 AATACTTGCG ACTAGAACCG GAGACATTAC GCCATGAACA AGAGCGCCGC

251 CGCTGGCCTG CTGGGCTATG CCCGCGTCAG CACCGACGAC CAGGACTTGA

301 CCAACCAACG GGCCGAACTG CACGCGGCCG GCTGCACCAA GCTGTTTTCC

-continued

```
 351 GAGAAGATCA CCGGCACCAG GCGCGACCGC CCGGAGCTGG CCAGGATGCT

 401 TGACCACCTA CGCCCTGGCG ACGTTGTGAC AGTGACCAGG CTAGACCGCC

 451 TGGCCCGCAG CACCCGCGAC CTACTGGACA TTGCCGAGCG CATCCAGGAG

 501 GCCGGCGCGG GCCTGCGTAG CCTGGCAGAG CCGTGGGCCG ACACCACCAC

 551 GCCGGCCGGC CGCATGGTGT TGACCGTGTT CGCCGGCATT GCCGAGTTCG

 601 AGCGTTCCCT AATCATCGAC CGCACCCGGA GCGGGCGCGA GGCCGCCAAG

 651 GCCCGAGGCG TGAAGTTTGG CCCCCGCCCT ACCCTCACCC CGGCACAGAT

 701 CGCGCACGCC CGCGAGCTGA TCGACCAGGA AGGCCGCACC GTGAAAGAGG

 751 CGGCTGCACT GCTTGGCGTG CATCGCTCGA CCCTGTACCG CGCACTTGAG

 801 CGCAGCGAGG AAGTGACGCC CACCGAGGCC AGGCGGCGCG GTGCCTTCCG

 851 TGAGGACGCA TTGACCGAGG CCGACGCCCT GGCGGCCGCC GAGAATGAAC

 901 GCCAAGAGGA ACAAGCATGA AACCGCACCA GGACGGCCAG GACGAACCGT

 951 TTTTCATTAC CGAAGAGATC GAGGCGGAGA TGATCGCGGC CGGGTACGTG

1001 TTCGAGCCGC CCGCGCACGT CTCAACCGTG CGGCTGCATG AAATCCTGGC

1051 CGGTTTGTCT GATGCCAAGC TGGCGGCCTG GCCGGCCAGC TTGGCCGCTG

1101 AAGAAACCGA GCGCCGCCGT CTAAAAAGGT GATGTGTATT TGAGTAAAAC

1151 AGCTTGCGTC ATGCGGTCGC TGCGTATATG ATGCGATGAG TAAATAAACA

1201 AATACGCAAG GGGAACGCAT GAAGGTTATC GCTGTACTTA ACCAGAAAGG

1251 CGGGTCAGGC AAGACGACCA TCGCAACCCA TCTAGCCCGC GCCCTGCAAC

1301 TCGCCGGGGC CGATGTTCTG TTAGTCGATT CCGATCCCCA GGGCAGTGCC

1351 CGCGATTGGG CGGCCGTGCG GGAAGATCAA CCGCTAACCG TTGTCGGCAT

1401 CGACCGCCCG ACGATTGACC GCGACGTGAA GGCCATCGGC CGGCGCGACT

1451 TCGTAGTGAT CGACGGAGCG CCCCAGGCGG CGGACTTGGC TGTGTCCGCG

1501 ATCAAGGCAG CCGACTTCGT GCTGATTCCG GTGCAGCCAA GCCCTTACGA

1551 CATATGGGCC ACCGCCGACC TGGTGGAGCT GGTTAAGCAG CGCATTGAGG

1601 TCACGGATGG AAGGCTACAA GCGGCCTTTG TCGTGTCGCG GGCGATCAAA

1651 GGCACGCGCA TCGGCGGTGA GGTTGCCGAG GCGCTGGCCG GGTACGAGCT

1701 GCCCATTCTT GAGTCCCGTA TCACGCAGCG CGTGAGCTAC CCAGGCACTG

1751 CCGCCGCCGG CACAACCGTT CTTGAATCAG AACCCGAGGG CGACGCTGCC

1801 CGCGAGGTCC AGGCGCTGGC CGCTGAAATT AAATCAAAAC TCATTTGAGT

1851 TAATGAGGTA AAGAGAAAAT GAGCAAAAGC ACAAACACGC TAAGTGCCGG

1901 CCGTCCGAGC GCACGCAGCA GCAAGGCTGC AACGTTGGCC AGCCTGGCAG

1951 ACACGCCAGC CATGAAGCGG GTCAACTTTC AGTTGCCGGC GGAGGATCAC

2001 ACCAAGCTGA AGATGTACGC GGTACGCCAA GGCAAGACCA TTACCGAGCT

2051 GCTATCTGAA TACATCGCGC AGCTACCAGA GTAAATGAGC AAATGAATAA

2101 ATGAGTAGAT GAATTTTAGC GGCTAAAGGA GGCGGCATGG AAAATCAAGA

2151 ACAACCAGGC ACCGACGCCG TGGAATGCCC CATGTGTGGA GGAACGGGCG

2201 GTTGGCCAGG CGTAAGCGGC TGGGTTGTCT GCCGGCCCTG CAATGGCACT

2251 GGAACCCCCA AGCCCGAGGA ATCGGCGTGA CGGTCGCAAA CCATCCGGCC
```

-continued

2301 CGGTACAAAT CGGCGCGGCG CTGGGTGATG ACCTGGTGGA GAAGTTGAAG
2351 GCCGCGCAGG CCGCCCAGCG GCAACGCATC GAGGCAGAAG CACGCCCCGG
2401 TGAATCGTGG CACGCGGCCG CTGATCGAAT CCGCAAAGAA TCCCGGCAAC
2451 CGCCGGCAGC CGGTGCGCCG TCGATTAGGA AGCCGCCCAA GGGCGACGAG
2501 CAACCAGATT TTTTCGTTCC GATGCTCTAT GACGTGGGCA CCCGCGTCAG
2551 TCGCAGCATC ATGGACGTGG CCGTTTTCCG TCTGTCGAAG CGTGACCGAC
2601 GAGCTGGCGA GGTGATCCGC TACGAGCTTC CAGACGGGCA CGTAGAGGTT
2651 TCCGCAGGGC CGGCCGGCAT GGCCAGTGTG TGGGATTACG ACCTGGTACT
2701 GATGGCGGTT TCCCATCTAA CCGAATCCAT GAACCGATAC CGGGAAGGGA
2751 AGGGAGACAA GCCCGGCCGC GTGTTCCGTC CACACGTTGC GGACGTACTC
2801 AAGTTCTGCC GGCGAGCCGA TGGCGGAAAG CAGAAAGACG ACCTGGTAGA
2851 AACCTGCATT CGGTTAAACA CCACGCACGT TGCCATGCAG CGTACGAAGA
2901 AGGCCAAGAA CGGCCGCCTG GTGACGGTAT CCGAGGGTGA AGCCTTGATT
2951 AGCCGCTACA AGATCGTAAA GAGCGAAACC GGGCGGCCGG AGTACATCGA
3001 GATCGAGCTA GCTGATTGGA TGTACCGCGA GATCACAGAA GGCAAGAACC
3051 CGGACGTGCT GACGGTTCAC CCCGATTACT TTTTGATCGA TCCCGGCATC
3101 GGCCGTTTTC TCTACCGCCT GGCACGCCGC GCCGCAGGCA AGGCAGAAGC
3151 CAGATGGTTG TTCAAGACGA TCTACGAACG CAGTGGCAGC GCCGGAGAGT
3201 TCAAGAAGTT CTGTTTCACC GTGCGCAAGC TGATCGGGTC AAATGACCTG
3251 CCGGAGTACG ATTTGAAGGA GGAGGCGGGG CAGGCTGGCC CGATCCTAGT
3301 CATGCGCTAC CGCAACCTGA TCGAGGGCGA AGCATCCGCC GGTTCCTAAT
3351 GTACGGAGCA GATGCTAGGG CAAATTGCCC TAGCAGGGGA AAAAGGTCGA
3401 AAAGGTCTCT TTCCTGTGGA TAGCACGTAC ATTGGGAACC CAAAGCCGTA
3451 CATTGGGAAC CGGAACCCGT ACATTGGGAA CCCAAAGCCG TACATTGGGA
3501 ACCGGTCACA CATGTAAGTG ACTGATATAA AAGAGAAAAA AGGCGATTTT
3551 TCCGCCTAAA ACTCTTTAAA ACTTATTAAA ACTCTTAAAA CCCGCCTGGC
3601 CTGTGCATAA CTGTCTGGCC AGCCCACAGC CGAAGAGCTG CAAAAAGCGC
3651 CTACCCTTCG GTCGCTGCGC TCCCTACGCC CCGCCGCTTC GCGTCGGCCT
3701 ATCGCGGCCG CTGGCCGCTC AAAAATGGCT GGCCTACGGC CAGGCAATCT
3751 ACCAGGGCGC GGACAAGCCG CGCCGTCGCC ACTCGACCGC CGGCGCCCAC
3801 ATCAAGGCAC CCTGCCTCGC GCGTTTCGGT GATGACGGTG AAAACCTCTG
3851 ACACATGCAG CTCCCGGAGA CGGTCACAGC TTGTCTGTAA GCGGATGCCG
3901 GGAGCAGACA AGCCCGTCAG GGCGCGTCAG CGGGTGTTGG CGGGTGTCGG
3951 GGCGCAGCCA TGACCCAGTC ACGTAGCGAT AGCGGAGTGT ATACTGGCTT
4001 AACTATGCGG CATCAGAGCA GATTGTACTG AGAGTGCACC ATATGCGGTG
4051 TGAAATACCG CACAGATGCG TAAGGAGAAA ATACCGCATC AGGCGCTCTT
4101 CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA
4151 GCGGTATCAG CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG

-continued

```
4201 GGATAACGCA GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA
4251 ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCCT
4301 GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC
4351 AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT
4401 CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
4451 TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC
4501 GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC
4551 AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA GTCCAACCCG
4601 GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG
4651 CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA
4701 ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG
4751 CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC
4801 CACCGCTGGT AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA
4851 GAAAAAAAGG ATCTCAAGAA GATCCTTTGA TCTTTTCTAC GGGGTCTGAC
4901 GCTCAGTGGA ACGAAAACTC ACGTTAAGGG ATTTTGGTCA TGCATTCTAG
4951 GTACTAAAAC AATTCATCCA GTAAAATATA ATATTTTATT TTCTCCCAAT
5001 CAGGCTTGAT CCCCAGTAAG TCAAAAAATA GCTCGACATA CTGTTCTTCC
5051 CCGATATCCT CCCTGATCGA CCGGACGCAG AAGGCAATGT CATACCACTT
5101 GTCCGCCCTG CCGCTTCTCC CAAGATCAAT AAAGCCACTT ACTTTGCCAT
5151 CTTTCACAAA GATGTTGCTG TCTCCCAGGT CGCCGTGGGA AAAGACAAGT
5201 TCCTCTTCGG GCTTTTCCGT CTTTAAAAAA TCATACAGCT CGCGCGGATC
5251 TTTAAATGGA GTGTCTTCTT CCCAGTTTTC GCAATCCACA TCGGCCAGAT
5301 CGTTATTCAG TAAGTAATCC AATTCGGCTA AGCGGCTGTC TAAGCTATTC
5351 GTATAGGGAC AATCCGATAT GTCGATGGAG TGAAAGAGCC TGATGCACTC
5401 CGCATACAGC TCGATAATCT TTTCAGGGCT TTGTTCATCT TCATACTCTT
5451 CCGAGCAAAG GACGCCATCG GCCTCACTCA TGAGCAGATT GCTCCAGCCA
5501 TCATGCCGTT CAAAGTGCAG GACCTTTGGA ACAGGCAGCT TTCCTTCCAG
5551 CCATAGCATC ATGTCCTTTT CCCGTTCCAC ATCATAGGTG GTCCCTTTAT
5601 ACCGGCTGTC CGTCATTTTT AAATATAGGT TTTCATTTTC TCCCACCAGC
5651 TTATATACCT TAGCAGGAGA CATTCCTTCC GTATCTTTTA CGCAGCGGTA
5701 TTTTTCGATC AGTTTTTTCA ATTCCGGTGA TATTCTCATT TTAGCCATTT
5751 ATTATTTCCT TCCTCTTTTC TACAGTATTT AAAGATACCC CAAGAAGCTA
5801 ATTATAACAA GACGAACTCC AATTCACTGT TCCTTGCATT CTAAAACCTT
5851 AAATACCAGA AAACAGCTTT TTCAAAGTTG TTTTCAAAGT TGGCGTATAA
5901 CATAGTATCG ACGGAGCCGA TTTTGAAACC GCGGTGATCA CAGGCAGCAA
5951 CGCTCTGTCA TCGTTACAAT CAACATGCTA CCCTCCGCGA GATCATCCGT
6001 GTTTCAAACC CGGCAGCTTA GTTGCCGTTC TTCCGAATAG CATCGGTAAC
6051 ATGAGCAAAG TCTGCCGCCT TACAACGGCT CTCCCGCTGA CGCCGTCCCG
6101 GACTGATGGG CTGCCTGTAT CGAGTGGTGA TTTTGTGCCG AGCTGCCGGT
```

-continued

6151 CGGGGAGCTG TTGGCTGGCT GGTGGCAGGA TATATTGTGG TGTAAACAAA

6201 TTGACGCTTA GACAACTTAA TAACACATTG CGGACGTTTT TAATGTACTG

6251 AATTAACGCC GAATTAATTC GGGGGATCTG GATTTTAGTA CTGGATTTTG

6301 GTTTTAGGAA TTAGAAATTT TATTGATAGA AGTATTTTAC AAATACAAAT

6351 ACATACTAAG GGTTTCTTAT ATGCTCAACA CATGAGCGAA ACCCTATAGG

6401 AACCCTAATT CCCTTATCTG GGAACTACTC ACACATTTTT ATGGAGAAAC

6451 TCGAGTTAAC CCTGAGACTG TTGGACAGAG CTCATTGGTA CCAGGAACAG

6501 GTGGTGGCGG CCCTCGGTGC GCTCGTACTG CTCCACGATG GTGTAGTCCT

6551 CGTTGTGGGA GGTGATGTCC AGCTTGGCGT CCACGTAGTA GTAGCCGGGC

6601 AGCTGCACGG GCTTCTTGGC CATGTAGATG GACTTGAACT CCACCAGGTA

6651 GTGGCCGCCG TCCTTCAGCT TCAGGGCCTT GTGGGTCTCG CCCTTCAGCA

6701 CGCCGTCGCG GGGGTACAGG CGCTCGGTGG AGGCCTCCCA GCCCATGGTC

6751 TTCTTCTGCA TCACGGGGCC GTCGGAGGGG AAGTTCACGC CGATGAACTT

6801 CACCTTGTAG ATGAAGCAGC CGTCCTGCAG GGAGGAGTCC TGGGTCACGG

6851 TCGCCACGCC GCCGTCCTCG AAGTTCATCA CGCGCTCCCA CTTGAAGCCC

6901 TCGGGGAAGG ACAGCTTCTT GTAGTCGGGG ATGTCGGCGG GGTGCTTCAC

6951 GTACACCTTG GAGCCGTACT GGAACTGGGG GGACAGGATG TCCCAGGCGA

7001 AGGGCAGGGG GCCGCCCTTG GTCACCTTCA GCTTCACGGT GTTGTGGCCC

7051 TCGTAGGGGC GGCCCTCGCC CTCGCCCTCG ATCTCGAACT CGTGGCCGTT

7101 CACGGTGCCC TCCATGCGCA CCTTGAAGCG CATGAACTCG GTGATGACGT

7151 TCTCGGAGGA GGCCATTTTG GTAGACTCGA GAGAGATAGA TTTGTAGAGA

7201 GAGACTGGTG ATTTCAGCGT GTCCTCTCCA AATGAAATGA ACTTCCTTAT

7251 ATAGAGGAAG GTCTTGCGAA GGATAGTGGG ATTGTGCGTC ATCCCTTACG

7301 TCAGTGGAGA TATCACATCA ATCCACTTGC TTTGAAGACG TGGTTGGAAC

7351 GTCTTCTTTT TCCACGATGC TCCTCGTGGG TGGGGGTCCA TCTTTGGGAC

7401 CACTGTCGGC AGAGGCATCT TGAACGATAG CCTTTCCTTT ATCGCAATGA

7451 TGGCATTTGT AGGTGCCACC TTCCTTTTCT ACTGTCCTTT TGATGAAGTG

7501 ACAGATAGCT GGGCAATGGA ATCCGAGGAG GTTTCCCGAT ATTACCCTTT

7551 GTTGAAAAGT CTCAATAGCC CTTTGGTCTT CTGAGACTGT ATCTTTGATA

7601 TTCTTGGAGT AGACGAGAGT GTCGTGCTCC ACCATGTTAT CACATCAATC

7651 CACTTGCTTT GAAGACGTGG TTGGAACGTC TTCTTTTTCC ACGATGCTCC

7701 TCGTGGGTGG GGGTCCATCT TTGGGACCAC TGTCGGCAGA GGCATCTTGA

7751 ACGATAGCCT TTCCTTTATC GCAATGATGG CATTTGTAGG TGCCACCTTC

7801 CTTTTCTACT GTCCTTTTGA TGAAGTGACA GATAGCTGGG CAATGGAATC

7851 CGAGGAGGTT TCCCGATATT ACCCTTTGTT GAAAAGTCTC AATAGCCCTT

7901 TGGTCTTCTG AGACTGTATC TTTGATATTC TTGGAGTAGA CGAGAGTGTC

7951 GTGCTCCACC ATGTTGGCAA GCTGCTCTAG CCAATACGCA AACCGCCTCT

8001 CCCCGCGCGT TGGCCGATTC ATTAATGCAG CTGGCACGAC AGGTTTCCCG

-continued

8051 ACTGGAAAGC GGGCAGTGAG CGCAACGCAA TTAATGTGAG TTAGCTCACT

8101 CATTAGGCAC CCCAGGCTTT ACACTTTATG CTTCCGGCTC GTATGTTGTG

8151 TGGAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGC TATGACCATG

8201 ATTACGAATT CAGGTACCAT TTAAATCCTG CAGGGTTTAA ACAGTGTTTT

8251 ACTCCTCATA TTAACTTCGG TCATTAGAGG CCACGATTTG ACACATTTTT

8301 ACTCAAAACA AAATGTTTGC ATATCTCTTA TAATTTCAAA TTCAACACAC

8351 AACAAATAAG AGAAAAAACA AATAATATTA ATTTGAGAAT GAACAAAAGG

8401 ACCATATCAT TCATTAACTC TTCTCCATCC ATTTCCATTT CACAGTTCGA

8451 TAGCGAAAAC CGAATAAAAA ACACAGTAAA TTACAAGCAC AACAAATGGT

8501 ACAAGAAAAA CAGTTTTCCC AATGCCATAA TACTCGAACG GCGCGCCTCA

8551 GCCCATATGC AGGCCGCCGT TGAGCGAGAA GTCGGCGCCG GTCGAGAAAC

8601 CGGACTCCTC CGACGACAAC CAGGCGCAGA TCGAGGCGAT CTCTTCCGGC

8651 AGGCCCAGGC GCTTGACCGG GATCGTCGCG ACGATCTTGT CGAGCACGTC

8701 CTGGCGGATC GCCTTGACCA TGTCGGTGGC GATATAGCCC GGAGAGACCG

8751 TGTTGACGGT CACGCCCTTG GTCGCCACTT CCTGCGCCAG TGCCATGGTG

8801 AAGCCATGCA CGCCGGCCTT GGCGGTGGAG TAGTTGGTCT GGCCGAACTG

8851 GCCCTTCTGC CCGTTCACCG ACGAGATGTT GACGATGCGG CCCCAGCCAC

8901 GGTCGGCCAT GCCGTCGATC ACCTGCTTGG TGACGTTGAA CAGCGAGGTC

8951 AGGTTGGTGT CGATCACCGC ATCCCAGTCG GCGCGGGTCA TCTTGCGGAA

9001 CACCACGTCG CGGGTGATAC CGGCGTTGTT GATCAGCACA TCAACCTCGC

9051 CGACCTCGGA CTTGACCTTG TCGAATGCGG TCTTGGTCGA GTCCCAGTCA

9101 GCCACATTGC CTTCCGAGGC AATGAAATCG AAGCCCAGGG CCTTCTGCTG

9151 CTCCAGCCAC TTTTCGCGGC GCGGCGAGTT GGGGCCGCAA CCGGCCACCA

9201 CACGAAAGCC ATCCTTGGCC AGCCGCTGGC AAATGGCGGT TCCGATACCA

9251 CCCATGCCGC CGGTCACATA CGCAATGCGC TGAGTCACTC TAGAATCTCT

9301 CGTCAATGGT GGCAAATAGG AAAGAGTCTC AAACTTCTTC TTTCCAATTG

9351 GAGGCCACAC CTGCATGCAC TTTACTCTTC CACCATTGCT TGTAATGGAA

9401 GTAATGTCAG TGTTGACCTT CTTCACTGGG AATCCAGTCA TGGATTTGAG

9451 GCCGCCGAAT GGAGCCACTG CGGCGGATTG CCCCCTAGAG GCACGGCTGA

9501 CTGTTGTCAC AGCGGAAGAG GATATCATAG AAGCCATTTT ACTAGTAAGA

9551 AGCTGAAAAT ATCAAAAGAA GGAACAGTCA TTAATCTATT GCATGTACTA

9601 GATTTTAGAT ATGAGTGGTC AAAAAAAACT TACGTTAATA ACGATGAAGA

9651 AGACAATGAT CCTCAGCACA ATCTCTCTCT CTCTCTCTTG GCTTCTCTTC

9701 TGGTGAATAG CACGAGAGAG GGTTTAAATG GAAGGCTCGT GGGTCCAAAA

9751 TGGGTGGCGG AGGAAATAGG AGAAGTAGGC AGTGACAAGT AATGTAGTAT

9801 TTAGTATTTG ATGAATGACA CATTTTCATT TCAGCATCAT CACCAACCAT

9851 CCTTTTGTTC CTTTGCTTCA ACTGTCACTT TCAATTGACA AAATTTTTTA

9901 TGTTTTCATG AGAAAACTAA ATTCTTATAA AGATTCATCT TCTTGAGTAT

9951 TATACGTGTA GTTTATGAAC AACACGTGTT GTTCCTATAT TTTTGTTCTG

-continued

10001 TTACCTCTAG AATAAAGTTG TCACCATTTC ATGAGTTCAA TTTTTCTTTA

10051 ATAGCCCCAA AAACAAAAGA TGATTCACAA GAAAGATGCG AATATTTTGC

10101 TATGAATCTT TTCTTAAGAG AAGCAATTAC ATTTTCACAA TAAAATTAGA

10151 TCCACGACTT AACCTAGTTT ATGTTGATTA TTTCTAGTGT TAGTATTAAG

10201 CAAAAATAAA ACTTATGAAT ACGAAGGCCT TTAAAGGAAA CTAAAGAAAG

10251 GACAAGGTAT AAACGTCCTA GAAAGTTCTA GGGTTTAGGC TTAGGGTCTA

10301 AGATATATGC TTTGAGTTTT ATGGCTTAGT AACACATTTT TGTAACACTT

10351 CTTTGTAACA TTTCTTGATA TGTTGGAGAA GTAACTCGTC TGGACAATAG

10401 TTATTTCCAA TATATAGGAA AAACGGCCTA AACAATAGCC GACGGGGACA

10451 AATACATCAT AAACAAAAAA TCCCGGTTAC AAACTTCCTA AAAAGCCATT

10501 CGGTCCACTC CGTTAAGCCT GAACTGTGCC TCCGTTATGC AAAAACGCCG

10551 TTGACCATCC GTAACCTAGT TGACTGACGG ATTATGGATT TAATCCGTTT

10601 TAAGGCCGTT AATAACACCA AAACGACGTC GTTTTGGTGT TTAAATTTTT

10651 TTTAACAACA ATTAAACCAA ACGACGTCGT TTTGGTTTAA TTAAATTTTT

10701 TTATCAAAAA CCCAAGCCCA AGCCCAAAAC TCTTAACAAA AGATAAAGCC

10751 CATCTCTATT TTTTCTAATT AAAACGCACA GCATTATGTT TCTTCTCTAA

10801 CGGATATATT TTCAATCTCA TAAATTGGGG ATTAGGGTTC TTATTTCCCA

10851 ATTCTCAATC TCTCAAAATT CTCCAAAATT CTCTGAAATT GATAATGCCT

10901 TCTTCTTCTT CAAACTCGTT TTTCTCTTTT GACAGTGAGC TTGAAGATGA

10951 TAACCATCGT GGTTTTCCTA AGACCTGTCG ATTTGGATGT CGTGTTGTGA

11001 TCAGAACCTC AAGAACTCCA AAAAACCTAG GTAGATTATT CCATACCTGT

11051 GAGAAAAATT TCAAAAGAGG AGGATTCCAC ACCTGGAAGT GGACTGATGT

11101 GTCTTTAGTA GAAGAAGTAG AGGACATAAA GGCTTACATT CATAACCGTG

11151 AGAAGTGTCA CGATGAAGAA ATGTTATTAT TGAAGGCTCA GATTCGTGGC

11201 TGTGAGAAGA TGATTGAAGG CTTGAAAGGA GAAGCAAAAC GTATGAAGCT

11251 AATTGTTGTT GCCGGAATAG TTGTGTTTGG TTGCTTTTTG TGTCTCTCTA

11301 AGTGATGTAT GAGATGAATG TTTGTGTATG TGATGTTGTT TTGTCTCAAT

11351 AATTAGTCAC TGATGTTGTA TGTAATGTTG TGTTTTGCAT CTCTAATTAG

11401 TTAATAATGA ATGTTGTTCT TATGTAATGT TTGATTTAAT CAATGGCTTT

11451 TGCAAATAAA TCCATAACAG AACNTATTCA ATATTTTCGA AAACATAACA

11501 AAGGTTTCAA AAGAAATTGC ATGTTGATTA GCTGAGTTTT CAAACAAAAT

11551 GCATTACATA GACAGACCCT GCTTCATAAT CCCCAAAACA CAAAAGAGAA

11601 GCATGCTAAT AACCGCAACT AATATCCAAA GACAGCTTCA TAATCCCAAA

11651 ACACAAAAAA AGAAGATTCA TAACCGATCC TTCATGTATT TAAAGAAAAT

11701 CAGACAACAA GCAAAGACTT AATCTTCCTG AGTAACTGAT GAGCTCAAGT

11751 CGACGTTTAA ACAGTGTTTT ACTCCTCATA TTAACTTCGG TCATTAGAGG

11801 CCACGATTTG ACACATTTTT ACTCAAAACA AAATGTTTGC ATATCTCTTA

11851 TAATTTCAAA TTCAACACAC AACAAATAAG AGAAAAAACA AATAATATTA

-continued

```
11901 ATTTGAGAAT GAACAAAAGG ACCATATCAT TCATTAACTC TTCTCCATCC
11951 ATTTCCATTT CACAGTTCGA TAGCGAAAAC CGAATAAAAA ACACAGTAAA
12001 TTACAAGCAC AACAAATGGT ACAAGAAAAA CAGTTTTCCC AATGCCATAA
12051 TACTCGAACT ACGTATTATT TGCGCTCGAC TGCCAGCGCC ACGCCCATGC
12101 CGCCGCCGAT GCACAGCGAG GCCAGGCCCT TCTTCGCGTC ACGGCGCTTC
12151 ATCTCGTGCA GCAGCGTCAC CAGGATACGG CAGCCCGACG CGCCGATCGG
12201 GTGGCCGATG GCGATGGCGC CGCCGTTCAC ATTGACCTTG GAGGTGTCCC
12251 AGCCCATCTG CTGGTGCACC GCCAGCGCCT GCGCGGCAAA GGCCTCGTTG
12301 ATCTCCATCA GGTCCAGGTC TTGCGGGGTC CACTCGGCGC GCGACAGGGC
12351 GCGCTTGGAG GCCGGCACCG GGCCCATGCC CATCACCTTG GGATCGACAC
12401 CGGCGTTGGC ATAGCTCTTG ATCGTGGCCA GCGGGGTCAG GCCCAGTTCC
12451 TTGGCCTTGG CCGCCGACAT CACCACCACC GCGGCGGCGC CGTCGTTCAG
12501 GCCCGAGGCG TTGGCCGCGG TCACCGTGCC GGCCTTGTCG AAGGCGGGCT
12551 TGAGGCCGGA CATGCTGTCC AGCGTGGCGC CCTGGCGCAC GAACTCGTCG
12601 GTCTTGAAGG CCACCGGGTC GCCCTTGCGC TGCGGGATCA GCACCGGGAC
12651 GATCTCTTCG TCAAACTTGC CGGCCTTCTG CGCGGCTTCG GCCTTGTTCT
12701 GCGAGCCGAC GGCGAACTCA TCCTGCGCCT CGCGTGTGAT GCCGTATTCC
12751 TTGGCCACGT TCTCGGCGGT GATGCCCATG TGGTACTGGT TGTACACGTC
12801 CCACAGGCCG TCGACGATCA TGGTGTCGAC CAGCTTGGCA TCGCCCATGC
12851 GGAAACCATC GCGCGAGCCC GGCAGCACGT GCGGGGCGGC GCTCATGTTT
12901 TCCTGGCCGC CGGCCACCAC GATCTCGGCG TCGCCCGCCA TGATCGCGTT
12951 GGCGGCCAGC ATCACGGCCT TCAGGCCCGA GCCGCACACC TTGTTGATGG
13001 TCATGGCCGG CACCATCGCC GGCAGGCCGG CCTTGATCGC GGCCTGGCGT
13051 GCGGGGTTCT GGCCCGAACC GGCGGTCAGC ACCTGGCCCA TGATGACTTC
13101 GCTCACCTGC TCCGGCTTGA CGCCGGCGCG CGCCGGCGCG GCCTTGATGA
13151 CCACGGCACC CAGTTCCGGT GCGGGGTTCT TGGCCAGCGA GCCGCCAAAC
13201 TTGCCGACCG CGGTGCGGGC GGCGGATACG ATGACAACGT CAGTCACTCT
13251 AGAATCTCTC GTCAATGGTG GCAAATAGGA AAGAGTCTCA AACTTCTTCT
13301 TTCCAATTGG AGGCCACACC TGCATGCACT TTACTCTTCC ACCATTGCTT
13351 GTAATGGAAG TAATGTCAGT GTTGACCTTC TTCACTGGGA ATCCAGTCAT
13401 GGATTTGAGG CCGCCGAATG GAGCCACTGC GGCGGATTGC CCCCTAGAGG
13451 CACGGCTGAC TGTTGTCACA GCCGAACAGG ATATCATAGA AGCCATTTTG
13501 GATCCAAGAA GCTGAAAATA TCAAAAGAAG GAACAGTCAT TAATCTATTG
13551 CATGTACTAG ATTTTAGATA TGAGTGGTCA AAAAAAACTT ACGTTAATAA
13601 CGATGAAGAA GACAATGATC CTCAGCACAA TCTCTCTCTC TCTCTCTTGG
13651 CTTCTCTTCT GGTGAATAGC ACGAGAGAGG GTTTAAATGG AAGGCTCGTG
13701 GGTCCAAAAT GGGTGGCGGA GGAAATAGGA GAAGTAGGCA GTGACAAGTA
13751 ATGTAGTATT TAGTATTTGA TGAATGACAC ATTTTCATTT CAGCATCATC
13801 ACCAACCATC CTTTTGTTCC TTTGCTTCAA CTGTCACTTT CAATTGACAA
```

-continued

13851 AATTTTTTAT GTTTTCATGA GAAAACTAAA TTCTTATAAA GATTCATCTT

13901 CTTGAGTATT ATACGTGTAG TTTATGAACA ACACGTGTTG TTCCTATATT

13951 TTTGTTCTGT TACCTCTAGA ATAAAGTTGT CACCATTTCA TGAGTTCAAT

14001 TTTTCTTTAA TAGCCCCAAA AACAAAAGAT GATTCACAAG AAAGATGCGA

14051 ATATTTTGCT ATGAATCTTT TCTTAAGAGA AGCAATTACA TTTTCACAAT

14101 AAAATTAGAT CCACGACTTA ACCTAGTTTA TGTTGATTAT TTCTAGTGTT

14151 AGTATTAAGC ACAAATAAGA CTTATGAATA CGAAGGCCTT TAAAGGAAAC

14201 TAAAGAAAGG ACAAGGTATA AACGTCCTAG AAAGTTCTAG GGTTTAGGCT

14251 TAGGGTCTAA GATATATGCT TTGAGTTTTA TGGCTTAGTA ACACATTTTT

14301 GTAACACTTC TTTGTAACAT TTCTTGATAT GTTGGAGAAG TAACTCGTCT

14351 GGACAATAGT TATTTCCAAT ATATAGGAAA AACTTCCTAA ACAATAGCCG

14401 ACGGGGACAA ATACATCATA AACAAAAGAT CCCGGTTACA AACTTCCTAA

14451 AAAGCCATTC GGTCCACTCC GTTAAGCCTG AACTGTGCCT CCGTTATGCA

14501 AAAACGCCGT TGACCATCCG TAACCTAGTT GACTGACGGA TTATGGATTT

14551 AATCCGTTTT AAGGCCGTTA ATAACACCAA AACGACGTCG TTTTGGTGTT

14601 TTAATTTTTT TTAACAACAA TTAAACCAAA CGACGTCGTT TTGGTTTAAT

14651 TAAATTTTTT TATCAAAAAC CCAAGCCCAA GCCCAAAACT CTTAACAAAA

14701 GATAAAGCCC ATCTCTATTT TTCTCTAATTA AAACGCACAG CATTATGTTT

14751 CTTCTCTAAC GGATATATTT TCAATCTCAT AAATTGGGGA TTAGGGTTCT

14801 TATTTCCCAA TTCTCAATCT CTCAAAATTC TCCAAAATTC TCTGAAATTG

14851 ATAATGCCTT CTTCTTCTTC AAACTCGTTT TTCTCTTTTG ACAGTGAGCT

14901 TGAAGATGAT AACCATCGTG GTTTTCCTAA GACCTGTCGA TTTGGATGTC

14951 GTGTTGTGAT CAGAACCTCA ATAACACCAA AAAACCTAGG TAGATTATTC

15001 CATACCTGTG AGAAAAATTT CAAAAGAGGA GGATTCCACA CCTGGAAGTG

15051 GACTGATGTG TCTTTAGTAG AAGAAGTAGA GGACATAAAG GCTTACATTC

15101 ATAACCGTGA GAAGTGTCAC GATGAAGAAA TGTTATTATT GAAGGCTCAG

15151 ATTCGTGGCT GTGAGAAGAT GATTGAAGGC TTGAAAGGAG AAGCAAAACG

15201 TATGAAGCTA ATTGTTCTTG CCGGAATAGT TGTGTTTGGT TGCTTTTTGT

15251 GTCTCTCTAA GTGATGTATG AGATGAATGT TTGTGTATGT GATGTTGTTT

15301 TGTCTCAATA ATTAGTCACT GATGTTGTAT GTAATGTTGT GTTTTGCATC

15351 TCTAATTAGT TAATAATGAA TGTTGTTCTT ATGTAATGTT TGATTTAATC

15401 AATGGCTTTT GCAAATAAAT CCATAACAGA ACNTATTCAA TATTTTCGAA

15451 AACATAACAA AGGTTTCAAA AGAAATTGCA TTAGCATTAG CTGAGTTTTC

15501 AAACAAAATG CATTACATAG ACAGACCCTG CTTCATAATC CCCAAAACAC

15551 AAAAGAGAAG CATGCTAATA ACCGCAACTA ATATCCAAAG ACAGCTTCAT

15601 AATCCCAAAA CACAAAAAAA GAAGATTCAT AACCGATCCT TCATGTATTT

15651 AAAGAAAATC AGACAACAAG CAAAGACTTA ATCTTCCTGA GTAATGGAAG

15701 AGCTCAACTG CAGGTTTAAA CAGTGTTTTA CTCCTCATAT TAACTTCGGT

-continued

```
15751 CATTAGAGGC CACGATTTGA CACATTTTTA CTCAAAACAA AATGTTTGCA

15801 TATCTCTTAT AATTTCAAAT TCAACACACA ACAAATAAGA GAAAAAACAA

15851 ATAATATTAA TTTGAGAATG AACAAAAGGA CCATATCATT CATTAACTCT

15901 TCTCCATCCA TTTCCATTTC ACAGTTCGAT AGCGAAAACC GAATAAAAAA

15951 CACAGTAAAT TACAAGCACA ACAAATGGTA CAAGAAAAAC AGTTTTCCCA

16001 ATGCCATAAT ACTCGAACGC GATCGCTCAG CCCTTGGCTT TGACGTAACG

16051 GCCGGGCGCC GCCTCGATCG CGGTGTAGCG GGCGTTGCCG GGCTTGGCCT

16101 TGGGCTTGAC CTTCTTGCCG CCATGCTGGG TCAGGAACCC GGCCCATTGC

16151 GGCCACCAGC TGCCCGGCAC TTCCTGCGCG CCATCGAACC AGGCCTGGGC

16201 ATCGGCGGCG CCACCGTCGT TGATCCAGTA GCTGCGCTTG TTCTTGGCCA

16251 CCGAGTTGAT CACGCCGGCG ATATGGCCGG ACGCGCCCAG CACGAAGCGG

16301 TTGGCGCCCG GCTTGCCCTG GTTGAGGATG TCGAGCGAAC CGTACGCCGA

16351 CATCCACGGC ACGATGTGGT CTTCGCGCGA ACCGTAGATG AAGGCCGGGG

16401 CGTCGATCAG GCCGAGGTCG ATCTTTTCGC CGGCCACCGT CAGCTTGCCC

16451 GGCACTTTCA GGCTGTTTTC CAGGTAGGTG TTGCGCAGGT ACCAGCAGAA

16501 CATCGGGCCC GGCAAATTGG TGCTGTCCGA ATTCCAGAAC AGCAGGTCAA

16551 ACGCCGCCGG CTCATTGCCT TTGAGGTAGT TCGACTGCAC ATAGTTCCAT

16601 ACCAGGTCGT TCGGACGCAG GCTCGAGAAG GTCGAGGCCA GGTCACGGCC

16651 CGGCATCAGG CCGCCATCGC GCAATTGCTG TTCACGCAGC GCGACCTGGG

16701 TTTCATCGAC GAAGACGTCG AGCACGCCGG TGTCGCTGAA GTCGAGGAAG

16751 GTGGTCAGCA GGGTCAGGCT GGCCGCCGGG TGCTGGCCAC GCGCCGCCAG

16801 TACCGCCAGT GCGGTGGCAA CGATGGTGCC GCCCACGCAG AAGCCGAACA

16851 TGTTCAGCTT GTCCTGGCCG CTGACGTCCT GGACGATGCG GATCGCTTCG

16901 ATCACGCCCT GCTCCACGTA GTCGTCCCAG GTGGTGCCGG CCAGCGACTT

16951 GTCCGGATTG CTCCACGAGA TCAGGAACAC GGTGTTGCCC TGCTCCACCG

17001 CGTAGCGCAC CAGCGAATTT TCCGGTTGCA GGTCGAGGAT GTAGAACTTG

17051 TTGATGCACG GCGGCACCAT CAACAGCGGG CGCTGGCTGA CCGTCGGCGT

17101 GGTCGGCGTG TACTGGATCA GCTGGAACAG CGGATTTTCG TAAATCACGG

17151 TGCCCGGGGT AATGGCCAGG TTGCGGCCCA CTTCAAAGGC CGATTCGTCC

17201 GACAGCGAGA TATGGCCCTT GTTGATATCG CCCAGCATAT TGACCAGGCC

17251 ACGCGTCAGG CTCTCGCCCT TGGTTTCAAT CAGTTTTTGC TGCGCTTCCG

17301 GGTTGGTGGC GAGGAAGTTC GCGGGCGACA TGGCATCAAT CACCTGCTGC

17351 ACGGCAAAGC GTATTTTCTG CTTTTGCTGG GGTGCGGTGT CCACCGCCTC

17401 CACCATGGCA CTGAGGAATT TGGCGTTGAG CAGGTAAGAT GCGGCATTGA

17451 AGGCCGACAT CGGATTGCCC TGCCAGGCTG CCGAGCTGAA GCGGCGGTCG

17501 CTGACGGCTG GCGCCTTGCC AGCCAAAAAA TCCTGCCACA ACGCGGTGAA

17551 GTCACGCAGA TAATCGTTTT TCAGCTGCTC CATCGCTTCC GGTTTGAGCG

17601 CAACGCCGAT ATCCTGCAAC ATGGTGGCCA TCGGGTTCGC CTCGGTGGTG

17651 GGCGCCTTGC TGAACCAGGA TTGCCACTGC AGCTCATCGT TGTTCTTGTT
```

-continued

17701 ACTCACTCTA GAATCTCTCG TCAATGGTGG CAAATAGGAA AGAGTCTCAA

17751 ACTTCTTCTT TCCAATTGGA GGCCACACCT GCATGCACTT TACTCTTCCA

17801 CCATTGCTTG TAATGGAAGT AATGTCAGTG TTGACCTTCT TCACTGGGAA

17851 TCCAGTCATG GATTTGAGGC CGCCGAATGG AGCCACTGCG GCGGATTGCC

17901 CCCTAGAGGC ACGGCTGACT GTTGTCACAG CGGAAGAGGA TATCATAGAA

17951 GCCATTTTTG TACAAAGAAG CTGAAAATAT CAAAAGAAGG AACAGTCATT

18001 AATCTATTGC ATGTACTAGA TTTTAGATAT GAGTGGTCAA AAAAAACTTA

18051 CGTTAATAAC GATGAAGAAG ACAATGATCC TCAGCACAAT CTCTCTCTCT

18101 CTCTCTTGGC TTCTCTTCTG GTGAATAGCA CGAGAGAGGG TTTAAATGGA

18151 AGGCTCGTGG GTCCAAAATG GGTGGCGGAG GAAATAGGAG AAGTAGGCAG

18201 TGACAAGTAA TGTAGTATTT AGTATTTGAT GAATGACACA TTTTCATTTC

18251 AGCATCATCA CCAACCATCC TTTTGTTCCT TTGCTTCAAC TGTCACTTTC

18301 AATTGACAAA ATTTTTTATG TTTTCATGAG AAAACTAAAT TCTTATAAAG

18351 ATTCATCTTC TTGAGTATTA TACGTGTAGT TTATGAACAA CACGTGTTGT

18401 TCCTATATTT TTGTTCTGTT ACCTCTAGAA TAAAGTTGTC ACCATTTCAT

18451 GAGTTCAATT TTTCTTTAAT AGCCCCAAAA ACAAAAGATG ATTCACAAGA

18501 AAGATGCGAA TATTTTGCTA TGAATCTTTT CTTAAGAGAA GCAATTACAT

18551 TTTCACAATA AAATTAGATC CACGACTTAA CCTAGTTTAT GTTGATTATT

18601 TCTAGTGTTA GTATTAAGCA AAAATAAAAC TTATGAATAC GAAGGCCTTT

18651 AAAGGAAACT AAAGAAAGGA CAAGGTATAA ACGTCCTAGA AAGTTCTAGG

18701 GTTTAGGCTT AGGGTCTAAG ATATATGCTT TGAGTTTTAT GGCTTAGTAA

18751 CACATTTTTG TAACACTTCT TTGTAACATT TCTTGATATG TTGGAGAAGT

18801 AACTCGTCTG GACAATAGTT ATTTCCAATA TATAGGAAAA ACGGCCTAAA

18851 CAATAGCCGA CGGGGACAAA TACATCATAA ACAAAAAATC CCGGTTACAA

18901 ACTTCCTAAA AAGCCATTCG GTCCACTCCG TTAAGCCTGA ACTGTGCCTC

18951 CGTTATGCAA AAACGCCGTT GACCATCCGT AACCTAGTTG ACTGACGGAT

19001 TATGGATTTA ATCCGTTTTA AGGCCGTTAA TAACACCAAA ACGACGTCGT

19051 TTTGGTGTTT TAATTTTTTT TAACAACAAT TAAACCAAAC GACGTCGTTT

19101 TGGTTTAATT AAATTTTTTT ATCAAAAACC CAAGCCCAAG CCCAAAACTC

19151 TTAACAAAAG ATAAAGCCCA TCTCTATTTT TTCTAATTAA AACGCACAGC

19201 ATTATGTTTC TTCTCTAACG GATATATTTT CAATCTCATA AATTGGGGAT

19251 TAGGGTTCTT ATTTCCCAAT TCTCAATCTC TAACACTTCT CCAAAATTCT

19301 CTGAAATTGA TAATGCCTTC TTCTTCTTCA AACTCGTTTT TCTCTTTTGA

19351 CAGTGAGCTT GAAGATGATA ACCATCGTGG TTTTCCTAAG ACCTGTCGAT

19401 TTGGATGTCG TGTTGTGATC AGAACCTCAA GAACTCCAAA AAACCTAGGT

19451 AGATTATTCC ATACCTGTGA GAAAAATTTC AAAAGAGGAG GATTCCACAC

19501 CTGGAAGTGG ACTGATGTGT CTTTAGTAGA AGAAGTAGAG GACATAAAGG

19551 CTTACATTCA TAACCGTGAG AAGTGTCACG ATGAAGAAAT GTTATTATTG

-continued

19601 AAGGCTCAGA TTCGTGGCTG TGAGAAGATG ATTGAAGGCT TGAAAGGAGA

19651 AGCAAAACGT ATGAAGCTAA TTGTTGTTGC CGGAATAGTT GTGTTTGGTT

19701 GCTTTTTGTG TCTCTCTAAG TGATGTATGA GATGAATGTT TGTGTATGTG

19751 ATGTTGTTTT GTCTCAATAA TTAGTCACTG ATGTTGTATG TAATGTTGTG

19801 TTTTGCATCT CTAATTAGTT AATAATGAAT GTTGTTCTTA TGTAATGTTT

19851 GATTTAATCA ATGGCTTTTG CAAATAAATC CATAACAGAA CNTATTCAAT

19901 ATTTTCGAAA ACATAACAAA GGTTTCAAAA GAAATTGCAT TAGCATTAGC

19951 TGAGTTTTCA AACAAAATGC ATTACATAGA CAGACCCTGC TTCATAATCC

20001 CCAAAACACA AAAGAGAAGC ATGCTAATAA CCGCAACTAA TATCCAAAGA

20051 CAGCTTCATA ATCCCAAAAC ACAAAAAAAG AAGATTCATA ACCGATCCTT

20101 CATGTATTTA AAGAAAATCA GACAACAAGC AAAGACTTAA TCTTCCTGAG

20151 TAACTGATGA GCTCAAAAGC TTGGCACTGG CCGTCGTTTT ACAACGTCGT

20201 GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC

20251 CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT

20301 CCCAACAGTT GCGCAGCCTG AATGGCGAAT GCTAGAGCAG CTTGAGCTTG

20351 GATCAGATTG TCGTTTCCCG CCTTCAGTTT AAACTATCAG TGTTTGACAG

20401 GATATATTGG CGGGTAAACC TAAGAGAAAA GAGCGTTTAT TAGAATAACG

20451 GATATTTAAA AGGGCGTGAA AAGGTTTATC CGTTCGTCCA TTTGTATGTG

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 2

<210> SEQ ID NO 1
<211> LENGTH: 15888
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Plasmid Vector

<400> SEQUENCE: 1

ggggatccgt acgtaagtac gtactcaaaa tgccaacaaa taaaaaaaaa gttgctttaa 60

taatgccaaa acaaattaat aaaacactta caacaccgga tttttttaa ttaaaatgtg 120

ccatttagga taaatagtta atatttttaa taattattta aaaagccgta tctactaaaa 180

tgatttttat ttggttgaaa atattaatat gtttaaatca acacaatcta tcaaaattaa 240

actaaaaaaa aaataagtgt acgtggttaa cattagtaca gtaatataag aggaaaatga 300

gaaattaaga aattgaaagc gagtctaatt tttaaattat gaacctgcat atataaaagg 360

aaagaaagaa tccaggaaga aaagaaatga aaccatgcat ggtcccctcg tcatcacgag 420

tttctgccat ttgcaataga aacactgaaa cacctttctc tttgtcactt aattgagatg 480

ccgaagccac ctcacaccat gaacttcatg aggtgtagca cccaaggctt ccatagccat 540

gcatactgaa gaatgtctca agctcagcac cctacttctg tgacgtgtcc ctcattcacc 600

ttcctctctt ccctataaat aaccacgcct caggttctcc gcttcacaac tcaaacattc 660

tctccattgg tccttaaaca ctcatcagtc atcaccgcgg ccgcggaatt catggcttct 720

-continued

```
atgatatcct cttccgctgt gacaacagtc agccgtgcct ctagggggca atccgccgca    780
gtggctccat tcggcggcct caaatccatg actggattcc cagtgaagaa ggtcaacact    840
gacattactt ccattacaag caatggtgga agagtaaagt gcatgcaggt gtggcctcca    900
attggaaaga agaagtttga gactctttcc tatttgccac cattgacgag agattctaga    960
gtgactgacg ttgtcatcgt atccgccgcc cgcaccgcgg tcggcaagtt tggcggctcg   1020
ctggccaaga tcccggcacc ggaactgggt gccgtggtca tcaaggccgc gctggagcgc   1080
gccggcgtca agccggagca ggtgagcgaa gtcatcatgg gccaggtgct gaccgccggt   1140
tcgggccaga accccgcacg ccaggccgcg atcaaggccg gcctgccggc gatggtgccg   1200
gccatgacca tcaacaaggt gtgcggctcg ggcctgaagg ccgtgatgct ggccgccaac   1260
gcgatcatgg cgggcgacgc cgagatcgtg gtggccggcg gccaggaaaa catgagcgcc   1320
gccccgcacg tgctgccggg ctcgcgcgat ggtttccgca tgggcgatgc caagctggtc   1380
gacaccatga tcgtcgacgg cctgtgggac gtgtacaacc agtaccacat gggcatcacc   1440
gccgagaacg tggccaagga atacggcatc acacgcgagg cgcaggatga gttcgccgtc   1500
ggctcgcaga acaaggccga agccgcgcag aaggccggca agtttgacga agagatcgtc   1560
ccggtgctga tcccgcagcg caagggcgac ccggtggcct tcaagaccga cgagttcgtg   1620
cgccagggcg ccacgctgga cagcatgtcc ggcctcaagc ccgccttcga caaggccggc   1680
acggtgaccg cggccaacgc ctcgggcctg aacgacggcg ccgccgcggt ggtggtgatg   1740
tcggcggcca aggccaagga actgggcctg accccgctgg ccacgatcaa gagctatgcc   1800
aacgccggtg tcgatcccaa ggtgatgggc atgggcccgg tgccggcctc caagcgcgcc   1860
ctgtcgcgcg ccgagtggac cccgcaagac ctggacctga tggagatcaa cgaggccttt   1920
gccgcgcagg cgctggcggt gcaccagcag atgggctggg acacctccaa ggtcaatgtg   1980
aacggcggcg ccatcgccat cggccacccg atcggcgcgt cgggctgccg tatcctggtg   2040
acgctgctgc acgagatgaa gcgccgtgac gcgaagaagg gcctggcctc gctgtgcatc   2100
ggcggcggca tgggcgtggc gctggcagtc gagcgcaaat aactcgaggc ggccgcagcc   2160
ctttttgtat gtgctacccc acttttgtct tttggcaat agtgctagca accaataaat   2220
aataataata ataatgaata agaaaacaaa ggctttagct tgccttttgt tcactgtaaa   2280
ataataatgt aagtactctc tataatgagt cacgaaactt ttgcgggaat aaaaggagaa   2340
attccaatga gttttctgtc aaatcttctt ttgtctctct ctctctctct tttttttttt   2400
tctttcttct gagcttcttg caaaacaaaa ggcaaacaat aacgattggt ccaatgatag   2460
ttagcttgat cgatgatatc tttaggaagt gttggcagga caggacatga tgtagaagac   2520
taaaattgaa agtattgcag acccaatagt tgaagattaa ctttaagaat gaagacgtct   2580
tatcaggttc ttcatgactt aagctttaag aggagtccac catggtagat ctgactagta   2640
gaaggtaatt atccaagatg tagcatcaag aatccaatgt ttacgggaaa aactatggaa   2700
gtattatgtg agctcagcaa gaagcagatc aatatgcggc acatatgcaa cctatgttca   2760
aaaatgaaga atgtacagat acaagatcct atactgccag aatacgaaga agaatacgta   2820
gaaattgaaa aagaagaacc aggcgaagaa aagaatcttg aagacgtaag cactgacgac   2880
aacaatgaaa agaagaagat aaggtcggtg attgtgaaag agacatagag gacacatgta   2940
aggtggaaaa tgtaagggcg gaaagtaacc ttatcacaaa ggaatcttat cccccactac   3000
```

-continued

```
ttatcctttt atatttttcc gtgtcatttt tgcccttgag ttttcctata taaggaacca   3060
agttcggcat ttgtgaaaac aagaaaaaat tggtgtaagc tattttcttt gaagtactga   3120
ggatacaact tcagagaaat ttgtaagaaa gtggatcgaa accatggcct cctccgagaa   3180
cgtcatcacc gagttcatgc gcttcaaggt gcgcatggag ggcaccgtga acggccacga   3240
gttcgagatc gagggcgagg gcgagggccg cccctacgag ggccacaaca ccgtgaagct   3300
gaaggtgacc aagggcggcc ccctgccctt cgcctgggac atcctgtccc cccagttcca   3360
gtacggctcc aaggtgtacg tgaagcaccc cgccgacatc cccgactaca agaagctgtc   3420
cttccccgag ggcttcaagt gggagcgcgt gatgaacttc gaggacggcg gcgtggcgac   3480
cgtgacccag gactcctccc tgcaggacgg ctgcttcatc tacaaggtga agttcatcgg   3540
cgtgaacttc ccctccgacg gccccgtgat gcagaagaag accatgggct gggaggcctc   3600
caccgagcgc ctgtaccccc gcgacggcgt gctgaagggc gagacccaca aggccctgaa   3660
gctgaaggac ggcggccact acctggtgga gttcaagtcc atctacatgg ccaagaagcc   3720
cgtgcagctg cccggctact actacgtgga cgccaagctg gacatcacct cccacaacga   3780
ggactacacc atcgtggagc agtacgagcg caccgagggc cgccaccacc tgttcctggt   3840
accaatgagc tctgtccaac agtctcaggg ttaatgtcta tgtatcttaa ataatgttgt   3900
cggcgatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct   3960
tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta   4020
atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta   4080
atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc   4140
atctatgtta ctagatcggg aattaaacta tcagtgtttg acaggatata ttggcgggta   4200
aacctaagag aaaagagcgt ttattagaat aacggatatt taaaagggcg tgaaaaggtt   4260
tatccgttcg tccatttgta tgtgcatgcc aaccacaggg ttcccctcgg gatcaaagta   4320
ctttgatcca acccctccgc tgctatagtg cagtcggctt ctgacgttca gtgcagccgt   4380
cttctgaaaa cgacatgtcg cacaagtcct aagttacgcg acaggctgcc gccctgccct   4440
tttcctggcg ttttcttgtc gcgtgtttta gtcgcataaa gtagaatact tgcgactaga   4500
accggagaca ttacgccatg aacaagagcg ccgccgctgg cctgctgggc tatgcccgcg   4560
tcagcaccga cgaccaggac ttgaccaacc aacgggccga actgcacgcg gccggctgca   4620
ccaagctgtt ttccgagaag atcaccggca ccaggcgcga ccgcccggag ctggccagga   4680
tgcttgacca cctacgccct ggcgacgttg tgacagtgac caggctagac cgcctggccc   4740
gcagcacccg cgacctactg gacattgccg agcgcatcca ggaggccggc gcgggcctgc   4800
gtagcctggc agagccgtgg gccgacacca ccacgccggc cggccgcatg gtgttgaccg   4860
tgttcgccgg cattgccgag ttcgagcgtt ccctaatcat cgaccgcacc cggagcgggc   4920
gcgaggccgc caaggcccga ggcgtgaagt ttggcccccg ccctaccctc accccggcac   4980
agatcgcgca cgcccgcgag ctgatcgacc aggaaggccg caccgtgaaa gaggcggctg   5040
cactgcttgg cgtgcatcgc tcgaccctgt accgcgcact tgagcgcagc gaggaagtga   5100
cgcccaccga ggccaggcgg cgcggtgcct tccgtgagga cgcattgacc gaggccgacg   5160
ccctggcggc cgccgagaat gaacgccaag aggaacaagc atgaaaccgc accaggacgg   5220
ccaggacgaa ccgtttttca ttaccgaaga gatcgaggcg gagatgatcg cggccgggta   5280
```

-continued

cgtgttcgag ccgcccgcgc acgtctcaac cgtgcggctg catgaaatcc tggccggttt 5340
gtctgatgcc aagctggcgg cctggccggc cagcttggcc gctgaagaaa ccgagcgccg 5400
ccgtctaaaa aggtgatgtg tatttgagta aaacagcttg cgtcatgcgg tcgctgcgta 5460
tatgatgcga tgagtaaata aacaaatacg caaggggaac gcatgaaggt tatcgctgta 5520
cttaaccaga aaggcgggtc aggcaagacg accatcgcaa cccatctagc ccgcgccctg 5580
caactcgccg gggccgatgt tctgttagtc gattccgatc cccagggcag tgcccgcgat 5640
tgggcggccg tgcgggaaga tcaaccgcta accgttgtcg gcatcgaccg cccgacgatt 5700
gaccgcgacg tgaaggccat cggccggcgc gacttcgtag tgatcgacgg agcgccccag 5760
gcggcggact tggctgtgtc cgcgatcaag gcagccgact tcgtgctgat tccggtgcag 5820
ccaagccctt acgacatatg ggccaccgcc gacctggtgg agctggttaa gcagcgcatt 5880
gaggtcacgg atggaaggct acaagcggcc tttgtcgtgt cgcgggcgat caaaggcacg 5940
cgcatcggcg gtgaggttgc cgaggcgctg gccgggtacg agctgcccat tcttgagtcc 6000
cgtatcacgc agcgcgtgag ctacccaggc actgccgccg ccggcacaac cgttcttgaa 6060
tcagaacccg agggcgacgc tgcccgcgag gtccaggcgc tggccgctga aattaaatca 6120
aaactcattt gagttaatga ggtaaagaga aaatgagcaa aagcacaaac acgctaagtg 6180
ccggccgtcc gagcgcacgc agcagcaagg ctgcaacgtt ggccagcctg gcagacacgc 6240
cagccatgaa gcgggtcaac tttcagttgc cggcggagga tcacaccaag ctgaagatgt 6300
acgcggtacg ccaaggcaag accattaccg agctgctatc tgaatacatc gcgcagctac 6360
cagagtaaat gagcaaatga ataaatgagt agatgaattt tagcggctaa aggaggcggc 6420
atggaaaatc aagaacaacc aggcaccgac gccgtggaat gccccatgtg tggaggaacg 6480
ggcggttggc caggcgtaag cggctgggtt gtctgccggc cctgcaatgg cactggaacc 6540
cccaagcccg aggaatcggc gtgacggtcg caaaccatcc ggcccggtac aaatcggcgc 6600
ggcgctgggt gatgacctgg tggagaagtt gaaggccgcg caggccgccc agcggcaacg 6660
catcgaggca gaagcacgcc ccggtgaatc gtggcaagcg gccgctgatc gaatccgcaa 6720
agaatcccgg caaccgccgg cagccggtgc gccgtcgatt aggaagccgc ccaagggcga 6780
cgagcaacca gattttttcg ttccgatgct ctatgacgtg ggcacccgcg atagtcgcag 6840
catcatggac gtggccgttt tccgtctgtc gaagcgtgac cgacgagctg gcgaggtgat 6900
ccgctacgag cttccagacg ggcacgtaga ggtttccgca gggccggccg gcatggccag 6960
tgtgtgggat tacgacctgg tactgatggc ggtttcccat ctaaccgaat ccatgaaccg 7020
ataccgggaa gggaagggag acaagcccgg ccgcgtgttc cgtccacacg ttgcggacgt 7080
actcaagttc tgccggcgag ccgatggcgg aaagcagaaa gacgacctgg tagaaacctg 7140
cattcggtta aacaccacgc acgttgccat gcagcgtacg aagaaggcca agaacggccg 7200
cctggtgacg gtatccgagg gtgaagcctt gattagccgc tacaagatcg taaagagcga 7260
aaccgggcgg ccggagtaca tcgagatcga gctagctgat tggatgtacc gcgagatcac 7320
agaaggcaag aacccggacg tgctgacggt tcaccccgat tactttttga tcgatcccgg 7380
catcggccgt tttctctacc gcctggcacg ccgcgccgca ggcaaggcag aagccagatg 7440
gttgttcaag acgatctacg aacgcagtgg cagcgccgga gagttcaaga agttctgttt 7500
caccgtgcgc aagctgatcg ggtcaaatga cctgccggag tacgatttga aggaggaggc 7560

-continued

```
ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac ctgatcgagg gcgaagcatc   7620
cgccggttcc taatgtacgg agcagatgct agggcaaatt gccctagcag gggaaaaagg   7680
tcgaaaaggt ctctttcctg tggatagcac gtacattggg aacccaaagc cgtacattgg   7740
gaaccggaac ccgtacattg ggaacccaaa gccgtacatt gggaaccggt cacacatgta   7800
agtgactgat ataaaagaga aaaaaggcga tttttccgcc taaaactctt taaaacttat   7860
taaaactctt aaaacccgcc tggcctgtgc ataactgtct ggccagcgca cagccgaaga   7920
gctgcaaaaa gcgcctaccc ttcggtcgct gcgctcccta cgccccgccg cttcgcgtcg   7980
gcctatcgcg gccgctggcc gctcaaaaat ggctggccta cggccaggca atctaccagg   8040
gcgcggacaa gccgcgccgt cgccactcga ccgccggcgc ccacatcaag gcaccctgcc   8100
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca   8160
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg   8220
ttggcgggtg tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg   8280
gcttaactat gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat   8340
accgcacaga tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac   8400
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt   8460
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca   8520
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc   8580
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact   8640
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct   8700
gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag   8760
ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca   8820
cgaacccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa   8880
cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc   8940
gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag   9000
aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg   9060
tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca   9120
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc   9180
tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgcatt ctaggtacta   9240
aaacaattca tccagtaaaa tataatattt tattttctcc caatcaggct tgatccccag   9300
taagtcaaaa aatagctcga catactgttc ttccccgata tcctccctga tcgaccggac   9360
gcagaaggca atgtcatacc acttgtccgc cctgccgctt ctcccaagat caataaagcc   9420
acttactttg ccatctttca caaagatgtt gctgtctccc aggtcgccgt gggaaaagac   9480
aagttcctct tcgggctttt ccgtctttaa aaaatcatac agctcgcgcg gatctttaaa   9540
tggagtgtct tcttcccagt ttcgcaatcc cacatcggcc agatcgttat tcagtaagta   9600
atccaattcg gctaagcggc tgtctaagct attcgtatag ggacaatccg atatgtcgat   9660
ggagtgaaag agcctgatgc actccgcata cagctcgata atcttttcag ggctttgttc   9720
atcttcatac tcttccgagc aaaggacgcc atcggcctca ctcatgagca gattgctcca   9780
gccatcatgc cgttcaaagt gcaggacctt tggaacaggc agctttcctt ccagccatag   9840
```

-continued

```
catcatgtcc ttttcccgtt ccacatcata ggtggtccct ttataccggc tgtccgtcat   9900
ttttaaatat aggttttcat tttctcccac cagcttatat accttagcag gagacattcc   9960
ttccgtatct tttacgcagc ggtatttttc gatcagtttt ttcaattccg gtgatattct  10020
cattttagcc atttattatt tccttcctct tttctacagt atttaaagat accccaagaa  10080
gctaattata acaagacgaa ctccaattca ctgttccttg cattctaaaa ccttaaatac  10140
cagaaaacag ctttttcaaa gttgttttca aagttggcgt ataacatagt atcgacggag  10200
ccgattttga aaccgcggtg atcacaggca gcaacgctct gtcatcgtta caatcaacat  10260
gctaccctcc gcgagatcat ccgtgtttca aacccggcag cttagttgcc gttcttccga  10320
atagcatcgg taacatgagc aaagtctgcc gccttacaac ggctctcccg ctgacgccgt  10380
cccggactga tgggctgcct gtatcgagtg gtgattttgt gccgagctgc cggtcgggga  10440
gctgttggct ggctggtggc aggatatatt gtggtgtaaa caaattgacg cttagacaac  10500
ttaataacac attgcggacg tttttaatgt actgaattaa cgccgaatta attcctaggc  10560
caccatgttg ggcccggggc gcgccgtacg tagtgtttat ctttgttgct tttctgaaca  10620
atttatttac tatgtaaata tattatcaat gtttaatcta ttttaatttg cacatgaatt  10680
ttcattttat ttttacttta caaaacaaat aaatatatat gcaaaaaaat ttacaaacga  10740
tgcacgggtt acaaactaat ttcattaaat gctaatgcag attttgtgaa gtaaaactcc  10800
aattatgatg aaaaatacca ccaacaccac ctgcgaaact gtatcccaac tgtccttaat  10860
aaaaatgtta aaaagtatat tattctcatt tgtctgtcat aatttatgta ccccacttta  10920
atttttctga tgtactaaac cgagggcaaa ctgaaacctg ttcctcatgc aaagccccta  10980
ctcaccatgt atcatgtacg tgtcatcacc caacaactcc acttttgcta tataacaaca  11040
cccccgtcac actctccctc tctaacacac accccactaa caattccttc acttgcagca  11100
ctgttgcatc atcatcttca ttgcaaaacc ctaaacttca ccttcaaccg cggccgcatg  11160
gcttctatga tatcctcttc cgctgtgaca acagtcagcc gtgcctctag ggggcaatcc  11220
gccgcagtgg ctccattcgg cggcctcaaa tccatgactg gattcccagt gaagaaggtc  11280
aacactgaca ttacttccat tacaagcaat ggtggaagag taaagtgcat gcaggtgtgg  11340
cctccaattg gaaagaagaa gtttgagact ctttcctatt tgccaccatt gacgagagat  11400
tctagagtga gtaacaagaa caacgatgag ctgcagtggc aatcctggtt cagcaaggcg  11460
cccaccaccg aggcgaaccc gatggccacc atgttgcagg atatcggcgt tgcgctcaaa  11520
ccggaagcga tggagcagct gaaaaacgat tatctgcgtg acttcaccgc gttgtggcag  11580
gatttttgg ctggcaaggc gccagccgtc agcgaccgcc gcttcagctc ggcagcctgg  11640
cagggcaatc cgatgtcggc cttcaatgcc gcatcttacc tgctcaacgc caaattcctc  11700
agtgccatgg tggaggcggt ggacaccgca ccccagcaaa agcagaaaat acgctttgcc  11760
gtgcagcagg tgattgatgc catgtcgccc gcgaacttcc tcgccaccaa cccggaagcg  11820
cagcaaaaac tgattgaaac caagggcgag agcctgacgc gtggcctggt caatatgctg  11880
ggcgatatca acaagggcca tatctcgctg tcggacgaat cggcctttga agtgggccgc  11940
aacctggcca ttaccccggg caccgtgatt tacgaaaatc cgctgttcca gctgatccag  12000
tacacgccga ccacgccgac ggtcagccag cgcccgctgt tgatggtgcc gccgtgcatc  12060
aacaagttct acatcctcga cctgcaaccg gaaaattcgc tggtgcgcta cgcggtggag  12120
```

-continued

```
cagggcaaca ccgtgttcct gatctcgtgg agcaatccgg acaagtcgct ggccggcacc  12180
acctgggacg actacgtgga gcagggcgtg atcgaagcga tccgcatcgt ccaggacgtc  12240
agcggccagg acaagctgaa catgttcggc ttctgcgtgg gcggcaccat cgttgccacc  12300
gcactggcgg tactggcggc gcgtggccag cacccggcgg ccagcctgac cctgctgacc  12360
accttcctcg acttcagcga caccggcgtg ctcgacgtct tcgtcgatga aacccaggtc  12420
gcgctgcgtg aacagcaatt gcgcgatggc ggcctgatgc cgggccgtga cctggcctcg  12480
accttctcga gcctgcgtcc gaacgacctg gtatggaact atgtgcagtc gaactacctc  12540
aaaggcaatg agccggcggc gtttgacctg ctgttctgga attcggacag caccaatttg  12600
ccgggcccga tgttctgctg gtacctgcgc aacacctacc tggaaaacag cctgaaagtg  12660
ccgggcaagc tgacggtggc cggcgaaaag atcgacctcg gcctgatcga cgccccggcc  12720
ttcatctacg gttcgcgcga agaccacatc gtgccgtgga tgtcggcgta cggttcgctc  12780
gacatcctca accagggcaa gccgggcgcc aaccgcttcg tgctgggcgc gtccggccat  12840
atcgccggcg tgatcaactc ggtggccaag aacaagcgca gctactggat caacgacggt  12900
ggcgccgccg atgcccaggc ctggttcgat ggcgcgcagg aagtgccggg cagctggtgg  12960
ccgcaatggg ccgggttcct gacccagcat ggcggcaaga aggtcaagcc caaggccaag  13020
cccggcaacg cccgctacac cgcgatcgag gcggcgcccg gccgttacgt caaagccaag  13080
ggctgagcgg ccgctgagta attctgatat tagagggagc attaatgtgt tgttgtgatg  13140
tggtttatat ggggaaatta aataaatgat gtatgtacct cttgcctatg taggtttgtg  13200
tgttttgttt tgttgtctag ctttggttat taagtagtag ggacgttcgt tcgtgtctca  13260
aaaaaagggg tactaccact ctgtagtgta tatggatgct ggaaatcaat gtgttttgta  13320
tttgttcacc tccattgttg aattcaatgt caaatgtgtt ttgcgttggt tatgtgtaaa  13380
attactatct ttctcgtccg atgatcaaag ttttaagcaa caaaaccaag ggtgaaattt  13440
aaactgtgct ttgttgaaga ttcttttatc atattgaaaa tcaaattact agcagcagat  13500
tttacctagc atgaaatttt atcaacagta cagcactcac taaccaagtt ccaaactaag  13560
atgcgccatt aacatcagcc aataggcatt ttcagcaagg cgcgcccgcg ccgatgtatg  13620
tgacaaccct cgggattgtt gatttatttc aaaactaaga gtttttgtct tattgttctc  13680
gtctattttg gatatcaatc ttagttttat atcttttcta gttctctacg tgttaaatgt  13740
tcaacacact agcaatttgg cctgccagcg tatggattat ggaactatca agtctgtgac  13800
gcgccgtacg tagtgtttat ctttgttgct tttctgaaca atttatttac tatgtaaata  13860
tattatcaat gtttaatcta ttttaatttg cacatgaatt ttcattttat ttttacttta  13920
caaaacaaat aaatatatat gcaaaaaaat ttacaaacga tgcacgggtt acaaactaat  13980
ttcattaaat gctaatgcag attttgtgaa gtaaaactcc aattatgatg aaaaatacca  14040
ccaacaccac ctgcgaaact gtatcccaac tgtccttaat aaaaatgtta aaaagtatat  14100
tattctcatt tgtctgtcat aatttatgta ccccacttta atttttctga tgtactaaac  14160
cgagggcaaa ctgaaacctg ttcctcatgc aaagccccta ctcaccatgt atcatgtacg  14220
tgtcatcacc caacaactcc acttttgcta tataacaaca cccccgtcac actctccctc  14280
tctaacacac accccactaa caattccttc acttgcagca ctgttgcatc atcatcttca  14340
ttgcaaaacc ctaaacttca ccttcaaccg cggccgcatg gcttctatga tatcctcttc  14400
```

-continued

```
cgctgtgaca acagtcagcc gtgcctctag gggcaatcc gccgcagtgg ctccattcgg  14460

cggcctcaaa tccatgactg gattcccagt gaagaaggtc aacactgaca ttacttccat  14520

tacaagcaat ggtggaagag taaagtgcat gcaggtgtgg cctccaattg gaaagaagaa  14580

gtttgagact ctttcctatt tgccaccatt gacgagagat tctagagtga ctcagcgcat  14640

tgcgtatgtg accggcggca tgggtggtat cggaaccgcc atttgccagc ggctggccaa  14700

ggatggcttt cgtgtggtgg ccggttgcgg ccccaactcg ccgcgccgcg aaaagtggct  14760

ggagcagcag aaggccctgg gcttcgattt cattgcctcg gaaggcaatg tggctgactg  14820

ggactcgacc aagaccgcat tcgacaaggt caagtccgag gtcggcgagg ttgatgtgct  14880

gatcaacaac gccggtatca cccgcgacgt ggtgttccgc aagatgaccc gcgccgactg  14940

ggatgcggtg atcgacacca acctgacctc gctgttcaac gtcaccaagc aggtgatcga  15000

cggcatggcc gaccgtggct ggggccgcat cgtcaacatc tcgtcggtga acgggcagaa  15060

gggccagttc ggccagacca actactccac cgccaaggcc ggcctgcatg gcttcaccat  15120

ggcactggcg caggaagtgg cgaccaaggg cgtgaccgtc aacacggtct ctccgggcta  15180

tatcgccacc gacatggtca aggcgatccg ccaggacgtg ctcgacaaga tcgtcgcgac  15240

gatcccggtc aagcgcctgg gcctgccgga agagatcgcc tcgatctgcg cctggttgtc  15300

gtcggaggag tccggtttct cgaccggcgc cgacttctcg ctcaacggcg gcctgcatat  15360

gggctgagcg gccgctgagt aattctgata ttagagggag cattaatgtg ttgttgtgat  15420

gtggtttata tggggaaatt aaataaatga tgtatgtacc tcttgcctat gtaggtttgt  15480

gtgttttgtt ttgttgtcta gctttggtta ttaagtagta gggacgttcg ttcgtgtctc  15540

aaaaaaaggg gtactaccac tctgtagtgt atatggatgc tggaaatcaa tgtgttttgt  15600

atttgttcac ctccattgtt gaattcaatg tcaaatgtgt tttgcgttgg ttatgtgtaa  15660

aattactatc tttctcgtcc gatgatcaaa gttttaagca acaaaaccaa gggtgaaatt  15720

taaactgtgc tttgttgaag attcttttat catattgaaa atcaaattac tagcagcaga  15780

ttttacctag catgaaattt tatcaacagt acagcactca ctaaccaagt tccaaactaa  15840

gatgcgccat taacatcagc caataggcat tttcagcaag gcgcgtaa              15888
```

```
<210> SEQ ID NO 2
<211> LENGTH: 20500
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: SYnthetic Plasmid Vector
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (11474)..(11474)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15433)..(15433)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (19892)..(19892)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 2
```

```
catgccaacc acagggttcc cctcgggatc aaagtacttt gatccaaccc ctccgctgct     60

atagtgcagt cggcttctga cgttcagtgc agccgtcttc tgaaaacgac atgtcgcaca    120
```

-continued

```
agtcctaagt tacgcgacag gctgccgccc tgccctttc ctggcgtttt cttgtcgcgt    180
gttttagtcg cataaagtag aatacttgcg actagaaccg gagacattac gccatgaaca    240
agagcgccgc cgctggcctg ctgggctatg cccgcgtcag caccgacgac caggacttga    300
ccaaccaacg ggccgaactg cacgcggccg gctgcaccaa gctgttttcc gagaagatca    360
ccggcaccag gcgcgaccgc ccggagctgg ccaggatgct tgaccaccta cgccctggcg    420
acgttgtgac agtgaccagg ctagaccgcc tggcccgcag cacccgcgac ctactggaca    480
ttgccgagcg catccaggag gccggcgcgg gcctgcgtag cctggcagag ccgtgggccg    540
acaccaccac gccggccggc cgcatggtgt tgaccgtgtt cgccggcatt gccgagttcg    600
agcgttccct aatcatcgac cgcacccgga gcgggcgcga ggccgccaag gcccgaggcg    660
tgaagtttgg cccccgccct accctcaccc cggcacagat cgcgcacgcc cgcgagctga    720
tcgaccagga aggccgcacc gtgaaagagg cggctgcact gcttggcgtg catcgctcga    780
ccctgtaccg cgcacttgag cgcagcgagg aagtgacgcc caccgaggcc aggcggcgcg    840
gtgccttccg tgaggacgca ttgaccgagg ccgacgccct ggcggccgcc gagaatgaac    900
gccaagagga acaagcatga aaccgcacca ggacggccag gacgaaccgt ttttcattac    960
cgaagagatc gaggcggaga tgatcgcggc cgggtacgtg ttcgagccgc ccgcgcacgt   1020
ctcaaccgtg cggctgcatg aaatcctggc cggtttgtct gatgccaagc tggcggcctg   1080
gccggccagc ttggccgctg aagaaaccga gcgccgccgt ctaaaaaggt gatgtgtatt   1140
tgagtaaaac agcttgcgtc atgcggtcgc tgcgtatatg atgcgatgag taaataaaca   1200
aatacgcaag gggaacgcat gaaggttatc gctgtactta accagaaagg cgggtcaggc   1260
aagacgacca tcgcaaccca tctagcccgc gccctgcaac tcgccggggc cgatgttctg   1320
ttagtcgatt ccgatcccca gggcagtgcc cgcgattggg cggccgtgcg ggaagatcaa   1380
ccgctaaccg ttgtcggcat cgaccgcccg acgattgacc gcgacgtgaa ggccatcggc   1440
cggcgcgact tcgtagtgat cgacggagcg ccccaggcgg cggacttggc tgtgtccgcg   1500
atcaaggcag ccgacttcgt gctgattccg gtgcagccaa gcccttacga catatgggcc   1560
accgccgacc tggtggagct ggttaagcag cgcattgagg tcacggatgg aaggctacaa   1620
gcggcctttg tcgtgtcgcg ggcgatcaaa ggcacgcgca tcggcggtga ggttgccgag   1680
gcgctggccg ggtacgagct gcccattctt gagtcccgta tcacgcagcg cgtgagctac   1740
ccaggcactg ccgccgccgg cacaaccgtt cttgaatcag aacccgaggg cgacgctgcc   1800
cgcgaggtcc aggcgctggc cgctgaaatt aaatcaaaac tcatttgagt taatgaggta   1860
aagagaaaat gagcaaaagc acaaacacgc taagtgccgg ccgtccgagc gcacgcagca   1920
gcaaggctgc aacgttggcc agcctggcag acacgccagc catgaagcgg gtcaactttc   1980
agttgccggc ggaggatcac accaagctga agatgtacgc ggtacgccaa ggcaagacca   2040
ttaccgagct gctatctgaa tacatcgcgc agctaccaga gtaaatgagc aaatgaataa   2100
atgagtagat gaattttagc ggctaaagga ggcggcatgg aaaatcaaga acaaccaggc   2160
accgacgccg tggaatgccc catgtgtgga ggaacgggcg gttggccagg cgtaagcggc   2220
tgggttgtct gccggccctg caatggcact ggaaccccca agcccgagga atcggcgtga   2280
cggtcgcaaa ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga   2340
gaagttgaag gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg   2400
```

-continued

```
tgaatcgtgg caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc   2460
cggtgcgccg tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc   2520
gatgctctat gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg   2580
tctgtcgaag cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca   2640
cgtagaggtt tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact   2700
gatggcggtt tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa   2760
gcccggccgc gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga   2820
tggcggaaag cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt   2880
tgccatgcag cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga   2940
agccttgatt agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga   3000
gatcgagcta gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct   3060
gacggttcac cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct   3120
ggcacgccgc gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg   3180
cagtggcagc gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc   3240
aaatgacctg ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt   3300
catgcgctac cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca   3360
gatgctaggg caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga   3420
tagcacgtac attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa   3480
cccaaagccg tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa   3540
aggcgatttt tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc   3600
ctgtgcataa ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg   3660
gtcgctgcgc tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc   3720
aaaaatggct ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc   3780
actcgaccgc cggcgcccac atcaaggcac cctgcctcgc gcgtttcggt gatgacggtg   3840
aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg   3900
ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggcgcagcca   3960
tgacccagtc acgtagcgat agcggagtgt atactggctt aactatgcgg catcagagca   4020
gattgtactg agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa   4080
ataccgcatc aggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg   4140
gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg   4200
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa   4260
ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg   4320
acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc   4380
tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc   4440
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc   4500
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg   4560
ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc   4620
actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga   4680
```

-continued

```
gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc   4740

tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac   4800

caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg   4860

atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc   4920

acgttaaggg attttggtca tgcattctag gtactaaaac aattcatcca gtaaaatata   4980

atattttatt ttctcccaat caggcttgat ccccagtaag tcaaaaaata gctcgacata   5040

ctgttcttcc ccgatatcct ccctgatcga ccggacgcag aaggcaatgt cataccactt   5100

gtccgccctg ccgcttctcc caagatcaat aaagccactt actttgccat ctttcacaaa   5160

gatgttgctg tctcccaggt cgccgtggga aaagacaagt tcctcttcgg gcttttccgt   5220

ctttaaaaaa tcatacagct cgcgcggatc tttaaatgga gtgtcttctt cccagttttc   5280

gcaatccaca tcggccagat cgttattcag taagtaatcc aattcggcta agcggctgtc   5340

taagctattc gtatagggac aatccgatat gtcgatggag tgaaagagcc tgatgcactc   5400

cgcatacagc tcgataatct tttcagggct ttgttcatct tcatactctt ccgagcaaag   5460

gacgccatcg gcctcactca tgagcagatt gctccagcca tcatgccgtt caaagtgcag   5520

gacctttgga acaggcagct ttccttccag ccatagcatc atgtcctttt cccgttccac   5580

atcataggtg gtccctttat accggctgtc cgtcattttt aaatataggt tttcattttc   5640

tcccaccagc ttatatacct tagcaggaga cattccttcc gtatctttta cgcagcggta   5700

tttttcgatc agtttttca attccggtga tattctcatt ttagccattt attatttcct   5760

tcctcttttc tacagtattt aaagataccc caagaagcta attataacaa gacgaactcc   5820

aattcactgt tccttgcatt ctaaaacctt aaataccaga aaacagcttt ttcaaagttg   5880

ttttcaaagt tggcgtataa catagtatcg acggagccga ttttgaaacc gcggtgatca   5940

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt   6000

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag   6060

tctgccgcct tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat   6120

cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga   6180

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt   6240

taatgtactg aattaacgcc gaattaattc gggggatctg gattttagta ctggattttg   6300

gttttaggaa ttagaaattt tattgataga agtattttac aaatacaaat acatactaag   6360

ggtttcttat atgctcaaca catgagcgaa accctatagg aaccctaatt cccttatctg   6420

ggaactactc acacattatt atggagaaac tcgagttaac cctgagactg ttggacagag   6480

ctcattggta ccaggaacag gtggtggcgg ccctcggtgc gctcgtactg ctccacgatg   6540

gtgtagtcct cgttgtggga ggtgatgtcc agcttggcgt ccacgtagta gtagccgggc   6600

agctgcacgg gcttcttggc catgtagatg gacttgaact ccaccaggta gtggccgccg   6660

tccttcagct tcagggcctt gtgggtctcg cccttcagca cgccgtcgcg ggggtacagg   6720

cgctcggtgg aggcctccca gcccatggtc ttcttctgca tcacggggcc gtcggagggg   6780

aagttcacgc cgatgaactt caccttgtag atgaagcagc cgtcctgcag ggaggagtcc   6840

tgggtcacgg tcgccacgcc gccgtcctcg aagttcatca cgcgctccca cttgaagccc   6900

tcggggaagg acagcttctt gtagtcgggg atgtcggcgg ggtgcttcac gtacaccttg   6960
```

-continued

```
gagccgtact ggaactgggg ggacaggatg tcccaggcga agggcagggg gccgcccttg   7020
gtcaccttca gcttcacggt gttgtggccc tcgtaggggc ggccctcgcc ctcgccctcg   7080
atctcgaact cgtggccgtt cacggtgccc tccatgcgca ccttgaagcg catgaactcg   7140
gtgatgacgt tctcggagga ggccattttg gtagactcga gagagataga tttgtagaga   7200
gagactggtg atttcagcgt gtcctctcca aatgaaatga acttccttat atagaggaag   7260
gtcttgcgaa ggatagtggg attgtgcgtc atcccttacg tcagtggaga tatcacatca   7320
atccacttgc tttgaagacg tggttggaac gtcttctttt tccacgatgc tcctcgtggg   7380
tgggggtcca tctttgggac cactgtcggc agaggcatct tgaacgatag cctttccttt   7440
atcgcaatga tggcatttgt aggtgccacc ttccttttct actgtccttt tgatgaagtg   7500
acagatagct gggcaatgga atccgaggag gtttcccgat attacccttt gttgaaaagt   7560
ctcaatagcc ctttggtctt ctgagactgt atctttgata ttcttggagt agacgagagt   7620
gtcgtgctcc accatgttat cacatcaatc cacttgcttt gaagacgtgg ttggaacgtc   7680
ttcttttcc acgatgctcc tcgtgggtgg gggtccatct ttgggaccac tgtcggcaga   7740
ggcatcttga acgatagcct ttcctttatc gcaatgatgg catttgtagg tgccaccttc   7800
cttttctact gtcctttga tgaagtgaca gatagctggg caatggaatc cgaggaggtt   7860
tcccgatatt accctttgtt gaaaagtctc aatagccctt tggtcttctg agactgtatc   7920
tttgatattc ttggagtaga cgagagtgtc gtgctccacc atgttggcaa gctgctctag   7980
ccaatacgca aaccgcctct ccccgcgcgt tggccgattc attaatgcag ctggcacgac   8040
aggtttcccg actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact   8100
cattaggcac cccaggcttt acactttatg cttccggctc gtatgttgtg tggaattgtg   8160
agcggataac aatttcacac aggaaacagc tatgaccatg attacgaatt caggtaccat   8220
ttaaatcctg cagggtttaa acagtgtttt actcctcata ttaacttcgg tcattagagg   8280
ccacgatttg acacattttt actcaaaaca aaatgtttgc atatctctta taatttcaaa   8340
ttcaacacac aacaaataag agaaaaaaca aataatatta atttgagaat gaacaaaagg   8400
accatatcat tcattaactc ttctccatcc atttccattt cacagttcga tagcgaaaac   8460
cgaataaaaa acacagtaaa ttacaagcac aacaaatggt acaagaaaaa cagttttccc   8520
aatgccataa tactcgaacg gcgcgcctca gcccatatgc aggccgccgt tgagcgagaa   8580
gtcggcgccg gtcgagaaac cggactcctc cgacgacaac caggcgcaga tcgaggcgat   8640
ctcttccggc aggcccaggc gcttgaccgg gatcgtcgcg acgatcttgt cgagcacgtc   8700
ctggcggatc gccttgacca tgtcggtggc gatatagccc ggagagaccg tgttgacggt   8760
cacgcccttg gtcgccactt cctgcgccag tgccatggtg aagccatgca ggccggcctt   8820
ggcggtggag tagttggtct ggccgaactg gcccttctgc ccgttcaccg acgagatgtt   8880
gacgatgcgg ccccagccac ggtcggccat gccgtcgatc acctgcttgg tgacgttgaa   8940
cagcgaggtc aggttggtgt cgatcaccgc atcccagtcg gcgcgggtca tcttgcggaa   9000
caccacgtcg cgggtgatac cggcgttgtt gatcagcaca tcaacctcgc cgacctcgga   9060
cttgaccttg tcgaatgcgg tcttggtcga gtcccagtca gccacattgc cttccgaggc   9120
aatgaaatcg aagcccaggg ccttctgctg ctccagccac ttttcgcggc gcggcgagtt   9180
ggggccgcaa ccggccacca cacgaaagcc atccttggcc agccgctggc aaatggcggt   9240
```

-continued

```
tccgatacca cccatgccgc cggtcacata cgcaatgcgc tgagtcactc tagaatctct   9300
cgtcaatggt ggcaaatagg aaagagtctc aaacttcttc tttccaattg gaggccacac   9360
ctgcatgcac tttactcttc caccattgct tgtaatggaa gtaatgtcag tgttgacctt   9420
cttcactggg aatccagtca tggatttgag gccgccgaat ggagccactg cggcggattg   9480
ccccctagag gcacggctga ctgttgtcac agcggaagag gatatcatag aagccatttt   9540
actagtaaga agctgaaaat atcaaaagaa ggaacagtca ttaatctatt gcatgtacta   9600
gattttagat atgagtggtc aaaaaaaact tacgttaata acgatgaaga agacaatgat   9660
cctcagcaca atctctctct ctctctcttg gcttctcttc tggtgaatag cacgagagag   9720
ggtttaaatg gaaggctcgt gggtccaaaa tgggtggcgg aggaaatagg agaagtaggc   9780
agtgacaagt aatgtagtat ttagtatttg atgaatgaca cattttcatt tcagcatcat   9840
caccaaccat ccttttgttc ctttgcttca actgtcactt tcaattgaca aaattttta   9900
tgttttcatg agaaaactaa attcttataa agattcatct tcttgagtat tatacgtgta   9960
gtttatgaac aacacgtgtt gttcctatat ttttgttctg ttacctctag aataaagttg  10020
tcaccatttc atgagttcaa tttttcttta atagccccaa aaacaaaaga tgattcacaa  10080
gaaagatgcg aatattttgc tatgaatctt ttcttaagag aagcaattac attttcacaa  10140
taaaattaga tccacgactt aacctagttt atgttgatta tttctagtgt tagtattaag  10200
caaaaataaa acttatgaat acgaaggcct ttaaaggaaa ctaaagaaag gacaaggtat  10260
aaacgtccta gaaagttcta gggtttaggc ttagggtcta agatatatgc tttgagtttt  10320
atggcttagt aacacatttt tgtaacactt ctttgtaaca tttcttgata tgttggagaa  10380
gtaactcgtc tggacaatag ttatttccaa tatataggaa aaacggccta aacaatagcc  10440
gacggggaca aatacatcat aaacaaaaaa tcccggttac aaacttccta aaaagccatt  10500
cggtccactc cgttaagcct gaactgtgcc tccgttatgc aaaaacgccg ttgaccatcc  10560
gtaacctagt tgactgacgg attatggatt taatccgttt taaggccgtt aataacacca  10620
aaacgacgtc gttttggtgt tttaattttt tttaacaaca attaaaccaa acgacgtcgt  10680
tttggtttaa ttaaattttt ttatcaaaaa cccaagccca agcccaaaac tcttaacaaa  10740
agataaagcc catctctatt ttttctaatt aaaacgcaca gcattatgtt tcttctctaa  10800
cggatatatt ttcaatctca taaattgggg attagggttc ttatttccca attctcaatc  10860
tctcaaaatt ctccaaaatt ctctgaaatt gataatgcct tcttcttctt caaactcgtt  10920
tttctctttt gacagtgagc ttgaagatga taaccatcgt ggttttccta agacctgtcg  10980
atttggatgt cgtgttgtga tcagaacctc aagaactcca aaaaacctag gtagattatt  11040
ccatacctgt gagaaaaatt tcaaaagagg aggattccac acctggaagt ggactgatgt  11100
gtctttagta gaagaagtag aggacataaa ggcttacatt cataaccgtg agaagtgtca  11160
cgatgaagaa atgttattat tgaaggctca gattcgtggc tgtgagaaga tgattgaagg  11220
cttgaaagga gaagcaaaac gtatgaagct aattgttgtt gccggaatag ttgtgtttgg  11280
ttgctttttg tgtctctcta agtgatgtat gagatgaatg tttgtgtatg tgatgttgtt  11340
ttgtctcaat aattagtcac tgatgttgta tgtaatgttg tgttttgcat ctctaattag  11400
ttaataatga atgttgttct tatgtaatgt ttgatttaat caatggcttt tgcaaataaa  11460
tccataacag aacntattca atattttcga aaacataaca aaggtttcaa aagaaattgc  11520
```

-continued

attagcatta gctgagtttt caaacaaaat gcattacata gacagaccct gcttcataat 11580

ccccaaaaca caaaagagaa gcatgctaat aaccgcaact aatatccaaa gacagcttca 11640

taatcccaaa acacaaaaaa agaagattca taaccgatcc ttcatgtatt taaagaaaat 11700

cagacaacaa gcaaagactt aatcttcctg agtaactgat gagctcaagt cgacgtttaa 11760

acagtgtttt actcctcata ttaacttcgg tcattagagg ccacgatttg acacattttt 11820

actcaaaaca aaatgtttgc atatctctta taatttcaaa ttcaacacac aacaaataag 11880

agaaaaaaca aataatatta atttgagaat gaacaaaagg accatatcat tcattaactc 11940

ttctccatcc atttccattt cacagttcga tagcgaaaac cgaataaaaa acacagtaaa 12000

ttacaagcac aacaaatggt acaagaaaaa cagttttccc aatgccataa tactcgaact 12060

acgtattatt tgcgctcgac tgccagcgcc acgcccatgc cgccgccgat gcacagcgag 12120

gccaggccct tcttcgcgtc acggcgcttc atctcgtgca gcagcgtcac caggatacgg 12180

cagcccgacg cgccgatcgg gtggccgatg gcgatggcgc cgccgttcac attgaccttg 12240

gaggtgtccc agcccatctg ctggtgcacc gccagcgcct gcgcggcaaa ggcctcgttg 12300

atctccatca ggtccaggtc ttgcggggtc cactcggcgc gcgacagggc gcgcttggag 12360

gccggcaccg ggcccatgcc catcaccttg ggatcgacac cggcgttggc atagctcttg 12420

atcgtggcca gcggggtcag gcccagttcc ttggccttgg ccgccgacat caccaccacc 12480

gcggcggcgc cgtcgttcag gcccgaggcg ttggccgcgg tcaccgtgcc ggccttgtcg 12540

aaggcgggct tgaggccgga catgctgtcc agcgtggcgc cctggcgcac gaactcgtcg 12600

gtcttgaagg ccaccgggtc gcccttgcgc tgcgggatca gcaccgggac gatctcttcg 12660

tcaaacttgc cggccttctg cgcggcttcg gccttgttct gcgagccgac ggcgaactca 12720

tcctgcgcct cgcgtgtgat gccgtattcc ttggccacgt tctcggcggt gatgcccatg 12780

tggtactggt tgtacacgtc ccacaggccg tcgacgatca tggtgtcgac cagcttggca 12840

tcgcccatgc ggaaaccatc gcgcgagccc ggcagcacgt gcggggcggc gctcatgttt 12900

tcctggccgc cggccaccac gatctcggcg tcgcccgcca tgatcgcgtt ggcggccagc 12960

atcacggcct tcaggcccga gccgcacacc ttgttgatgg tcatggccgg caccatcgcc 13020

ggcaggccgg ccttgatcgc ggcctggcgt gcggggttct ggcccgaacc ggcggtcagc 13080

acctggccca tgatgacttc gctcacctgc tccggcttga cgccggcgcg ctccagcgcg 13140

gccttgatga ccacggcacc cagttccggt gccgggatct tggccagcga gccgccaaac 13200

ttgccgaccg cggtgcgggc ggcggatacg atgacaacgt cagtcactct agaatctctc 13260

gtcaatggtg gcaaatagga aagagtctca aacttcttct ttccaattgg aggccacacc 13320

tgcatgcact ttactcttcc accattgctt gtaatggaag taatgtcagt gttgaccttc 13380

ttcactggga atccagtcat ggatttgagg ccgccgaatg gagccactgc ggcggattgc 13440

cccctagagg cacggctgac tgttgtcaca gcggaagagg atatcataga agccattttg 13500

gatccaagaa gctgaaaata tcaaaagaag gaacagtcat taatctattg catgtactag 13560

attttagata tgagtggtca aaaaaaactt acgttaataa cgatgaagaa gacaatgatc 13620

ctcagcacaa tctctctctc tctctcttgg cttctcttct ggtgaatagc acgagagagg 13680

gtttaaatgg aaggctcgtg ggtccaaaat gggtggcgga ggaaatagga gaagtaggca 13740

gtgacaagta atgtagtatt tagtatttga tgaatgacac attttcattt cagcatcatc 13800

-continued

```
accaaccatc cttttgttcc tttgcttcaa ctgtcacttt caattgacaa aatttttat  13860
gttttcatga gaaaactaaa ttcttataaa gattcatctt cttgagtatt atacgtgtag  13920
tttatgaaca acacgtgttg ttcctatatt tttgttctgt tacctctaga ataaagttgt  13980
caccatttca tgagttcaat ttttctttaa tagccccaaa aacaaaagat gattcacaag  14040
aaagatgcga atattttgct atgaatcttt tcttaagaga agcaattaca ttttcacaat  14100
aaaattagat ccacgactta acctagttta tgttgattat ttctagtgtt agtattaagc  14160
aaaaataaaa cttatgaata cgaaggcctt taaaggaaac taaagaaagg acaaggtata  14220
aacgtcctag aaagttctag ggtttaggct tagggtctaa gatatatgct ttgagtttta  14280
tggcttagta acacattttt gtaacacttc tttgtaacat ttcttgatat gttggagaag  14340
taactcgtct ggacaatagt tatttccaat atataggaaa aacggcctaa acaatagccg  14400
acggggacaa atacatcata aacaaaaaat cccggttaca aacttcctaa aaagccattc  14460
ggtccactcc gttaagcctg aactgtgcct ccgttatgca aaaacgccgt tgaccatccg  14520
taacctagtt gactgacgga ttatggattt aatccgtttt aaggccgtta ataacaccaa  14580
aacgacgtcg ttttggtgtt ttaatttttt ttaacaacaa ttaaaccaaa cgacgtcgtt  14640
ttggtttaat taaatttttt tatcaaaaac ccaagcccaa gcccaaaact cttaacaaaa  14700
gataaagccc atctctattt tttctaatta aaacgcacag cattatgttt cttctctaac  14760
ggatatattt tcaatctcat aaattgggga ttagggttct tatttcccaa ttctcaatct  14820
ctcaaaattc tccaaaattc tctgaaattg ataatgcctt cttcttcttc aaactcgttt  14880
ttctcttttg acagtgagct tgaagatgat aaccatcgtg gttttcctaa gacctgtcga  14940
tttggatgtc gtgttgtgat cagaacctca agaactccaa aaaacctagg tagattattc  15000
catacctgtg agaaaaattt caaaagagga ggattccaca cctggaagtg gactgatgtg  15060
tctttagtag aagaagtaga ggacataaag gcttacattc ataaccgtga gaagtgtcac  15120
gatgaagaaa tgttattatt gaaggctcag attcgtggct gtgagaagat gattgaaggc  15180
ttgaaaggag aagcaaaacg tatgaagcta attgttgttg ccggaatagt tgtgtttggt  15240
tgctttttgt gtctctctaa gtgatgtatg agatgaatgt ttgtgtatgt gatgttgttt  15300
tgtctcaata attagtcact gatgttgtat gtaatgttgt gttttgcatc tctaattagt  15360
taataatgaa tgttgttctt atgtaatgtt tgatttaatc aatggctttt gcaaataaat  15420
ccataacaga acntattcaa tattttcgaa aacataacaa aggtttcaaa agaaattgca  15480
ttagcattag ctgagttttc aaacaaaatg cattacatag acagaccctg cttcataatc  15540
cccaaaacac aaaagagaag catgctaata accgcaacta atatccaaag acagcttcat  15600
aatcccaaaa cacaaaaaaa gaagattcat aaccgatcct tcatgtattt aaagaaaatc  15660
agacaacaag caaagactta atcttcctga gtaactgatg agctcaactg caggtttaaa  15720
cagtgtttta ctcctcatat taacttcggt cattagaggc cacgatttga cacattttta  15780
ctcaaaacaa aatgtttgca tatctcttat aatttcaaat tcaacacaca acaaataaga  15840
gaaaaaacaa ataatattaa tttgagaatg aacaaaagga ccatatcatt cattaactct  15900
tctccatcca tttccatttc acagttcgat agcgaaaacc gaataaaaaa cacagtaaat  15960
tacaagcaca acaaatggta caagaaaaac agttttccca atgccataat actcgaacgc  16020
gatcgctcag cccttggctt tgacgtaacg gccgggcgcc gcctcgatcg cggtgtagcg  16080
```

-continued

```
ggcgttgccg ggcttggcct tgggcttgac cttcttgccg ccatgctggg tcaggaaccc  16140
ggcccattgc ggccaccagc tgcccggcac ttcctgcgcg ccatcgaacc aggcctgggc  16200
atcggcggcg ccaccgtcgt tgatccagta gctgcgcttg ttcttggcca ccgagttgat  16260
cacgccggcg atatggccgg acgcgcccag cacgaagcgg ttggcgcccg gcttgccctg  16320
gttgaggatg tcgagcgaac cgtacgccga catccacggc acgatgtggt cttcgcgcga  16380
accgtagatg aaggccgggg cgtcgatcag gccgaggtcg atcttttcgc cggccaccgt  16440
cagcttgccc ggcactttca ggctgttttc caggtaggtg ttgcgcaggt accagcagaa  16500
catcgggccc ggcaaattgg tgctgtccga attccagaac agcaggtcaa acgccgccgg  16560
ctcattgcct ttgaggtagt tcgactgcac atagttccat accaggtcgt tcggacgcag  16620
gctcgagaag gtcgaggcca ggtcacggcc cggcatcagg ccgccatcgc gcaattgctg  16680
ttcacgcagc gcgacctggg tttcatcgac gaagacgtcg agcacgccgg tgtcgctgaa  16740
gtcgaggaag gtggtcagca gggtcaggct ggccgccggg tgctggccac gcgccgccag  16800
taccgccagt gcggtggcaa cgatggtgcc gcccacgcag aagccgaaca tgttcagctt  16860
gtcctggccg ctgacgtcct ggacgatgcg gatcgcttcg atcacgccct gctccacgta  16920
gtcgtcccag gtggtgccgg ccagcgactt gtccggattg ctccacgaga tcaggaacac  16980
ggtgttgccc tgctccaccg cgtagcgcac cagcgaattt tccggttgca ggtcgaggat  17040
gtagaacttg ttgatgcacg gcggcaccat caacagcggg cgctggctga ccgtcggcgt  17100
ggtcggcgtg tactggatca gctggaacag cggattttcg taaatcacgg tgcccggggt  17160
aatggccagg ttgcggccca cttcaaaggc cgattcgtcc gacagcgaga tatggccctt  17220
gttgatatcg cccagcatat tgaccaggcc acgcgtcagg ctctcgccct tggtttcaat  17280
cagtttttgc tgcgcttccg ggttggtggc gaggaagttc gcgggcgaca tggcatcaat  17340
cacctgctgc acggcaaagc gtatttttctg cttttgctgg ggtgcggtgt ccaccgcctc  17400
caccatggca ctgaggaatt tggcgttgag caggtaagat gcggcattga aggccgacat  17460
cggattgccc tgccaggctg ccgagctgaa gcggcggtcg ctgacggctg gcgccttgcc  17520
agccaaaaaa tcctgccaca acgcggtgaa gtcacgcaga taatcgtttt tcagctgctc  17580
catcgcttcc ggtttgagcg caacgccgat atcctgcaac atggtggcca tcgggttcgc  17640
ctcggtggtg ggcgccttgc tgaaccagga ttgccactgc agctcatcgt tgttcttgtt  17700
actcactcta gaatctctcg tcaatggtgg caaataggaa agagtctcaa acttcttctt  17760
tccaattgga ggccacacct gcatgcactt tactcttcca ccattgcttg taatggaagt  17820
aatgtcagtg ttgaccttct tcactgggaa tccagtcatg gatttgaggc cgccgaatgg  17880
agccactgcg gcggattgcc ccctagaggc acggctgact gttgtcacag cggaagagga  17940
tatcatagaa gccatttttg tacaaagaag ctgaaaatat caaaagaagg aacagtcatt  18000
aatctattgc atgtactaga ttttagatat gagtggtcaa aaaaaactta cgttaataac  18060
gatgaagaag acaatgatcc tcagcacaat ctctctctct ctctcttggc ttctcttctg  18120
gtgaatagca cgagagaggg tttaaatgga aggctcgtgg gtccaaaatg ggtggcggag  18180
gaaataggag aagtaggcag tgacaagtaa tgtagtattt agtatttgat gaatgacaca  18240
ttttcatttc agcatcatca ccaaccatcc ttttgttcct ttgcttcaac tgtcactttc  18300
aattgacaaa attttttatg ttttcatgag aaaactaaat tcttataaag attcatcttc  18360
```

-continued

```
ttgagtatta tacgtgtagt ttatgaacaa cacgtgttgt tcctatattt ttgttctgtt  18420
acctctagaa taaagttgtc accatttcat gagttcaatt tttctttaat agccccaaaa  18480
acaaaagatg attcacaaga aagatgcgaa tattttgcta tgaatctttt cttaagagaa  18540
gcaattacat tttcacaata aaattagatc cacgacttaa cctagtttat gttgattatt  18600
tctagtgtta gtattaagca aaaataaaac ttatgaatac gaaggccttt aaaggaaact  18660
aaagaaagga caaggtataa acgtcctaga aagttctagg gtttaggctt agggtctaag  18720
atatatgctt tgagttttat ggcttagtaa cacatttttg taacacttct ttgtaacatt  18780
tcttgatatg ttggagaagt aactcgtctg gacaatagtt atttccaata tataggaaaa  18840
acggcctaaa caatagccga cggggacaaa tacatcataa acaaaaaatc ccggttacaa  18900
acttcctaaa aagccattcg gtccactccg ttaagcctga actgtgcctc cgttatgcaa  18960
aaacgccgtt gaccatccgt aacctagttg actgacggat tatggattta atccgtttta  19020
aggccgttaa taacaccaaa acgacgtcgt tttggtgttt taattttttt taacaacaat  19080
taaaccaaac gacgtcgttt tggtttaatt aaattttttt atcaaaaacc caagcccaag  19140
cccaaaactc ttaacaaaag ataaagccca tctctatttt ttctaattaa aacgcacagc  19200
attatgtttc ttctctaacg gatatatttt caatctcata aattggggat tagggttctt  19260
atttcccaat tctcaatctc tcaaaattct ccaaaattct ctgaaattga taatgccttc  19320
ttcttcttca aactcgtttt tctcttttga cagtgagctt gaagatgata accatcgtgg  19380
ttttcctaag acctgtcgat ttggatgtcg tgttgtgatc agaacctcaa gaactccaaa  19440
aaacctaggt agattattcc atacctgtga gaaaaatttc aaaagaggag gattccacac  19500
ctggaagtgg actgatgtgt ctttagtaga agaagtagag gacataaagg cttacattca  19560
taaccgtgag aagtgtcacg atgaagaaat gttattattg aaggctcaga ttcgtggctg  19620
tgagaagatg attgaaggct tgaaaggaga agcaaaacgt atgaagctaa ttgttgttgc  19680
cggaatagtt gtgtttggtt gctttttgtg tctctctaag tgatgtatga gatgaatgtt  19740
tgtgtatgtg atgttgtttt gtctcaataa ttagtcactg atgttgtatg taatgttgtg  19800
ttttgcatct ctaattagtt aataatgaat gttgttctta tgtaatgttt gatttaatca  19860
atggcttttg caaataaatc cataacagaa cntattcaat attttcgaaa acataacaaa  19920
ggtttcaaaa gaaattgcat tagcattagc tgagttttca aacaaaatgc attacataga  19980
cagaccctgc ttcataatcc ccaaaacaca aaagagaagc atgctaataa ccgcaactaa  20040
tatccaaaga cagcttcata atcccaaaac acaaaaaaag aagattcata accgatcctt  20100
catgtattta aagaaaatca gacaacaagc aaagacttaa tcttcctgag taactgatga  20160
gctcaaaagc ttggcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt  20220
tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga  20280
ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat gctagagcag  20340
cttgagcttg gatcagattg tcgtttcccg ccttcagttt aaactatcag tgtttgacag  20400
gatatattgg cgggtaaacc taagagaaaa gagcgtttat tagaataacg gatatttaaa  20460
agggcgtgaa aaggtttatc cgttcgtcca tttgtatgtg                        20500
```

1. An oilseed comprising greater than 7% polyhydroxyalkanoate (PHA) dry weight of the oilseed, wherein germination of the oilseed is impaired relative to an oilseed having less than 7% polyhydroxyalkanoate.

2. The oilseed of claim **1**, wherein the PHA comprises (poly) 3-hydroxybutyrate (PHB).

3. The oilseed of claim **1** comprises greater than 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18% or 19% of the dry weight of the seed.

4. The oilseed of claim **1**, wherein the oilseed is produced by a transgenic plant genetically engineered to produce PHA.

5. The oilseed of claim **4**, wherein the PHA is PHB.

6. The oilseed of claim **4**, wherein the plant transformed to produce the transgenic plant is selected from the group consisting of members of the *Brassica* family: *B. napus, B. rapa, B. carinata* and *B. juncea*; industrial oilseeds: *Camelina sativa*, Crambe, jatropha, castor; *Arabidopsis thaliana; Calendula, Cuphea*; maize; soybean; cottonseed; sunflower; palm; coconut; safflower; peanut; mustards including *Sinapis alba*; and tobacco.

7. The oilseed of claim **1**, wherein germination of the oilseed is impaired by 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or 100% relative to an oilseed comprising less than 7% PHA.

8. A transgenic plant genetically engineered to produce PHA, wherein the transgenic plant produces the oilseed of claim **1**.

9. The transgenic plant of claim **8** wherein the plant transformed to produce the oilseed of claim **1** is selected from the group consisting of members of the *Brassica* family; *B. napus, B. rapa, B. carinata* and *B. juncea*; industrial oilseeds: *Camelina sativa*, Crambe, Jatropha, castor; *Arabidopsis thaliana; Calendula, Cuphea*; maize; soybean; cottonseed; sunflower; palm; coconut; safflower; peanut; mustards: *Sinapis alba*; and tobacco.

10. (canceled)

11. A method for producing a hybrid transgenic plant line comprising crossing a plant line comprising one or more PHB biosynthetic pathway genes with a plant line containing the remaining PHB biosynthetic pathway gene(s) needed to complete the PHB biosynthetic pathway.

12. The method of claim **11** wherein the plant lines comprise cytoplasmic male sterility (CMS) controlled by an extranuclear genome.

13. The method of claim **11** wherein the male sterile line is maintained by crossing with a maintainer line that is genetically identical and comprises normal fertile cytoplasm.

14. The method of claim **13** wherein the maintainer line is transformed with one or more genes for the PHB biosynthetic pathway.

15. The method of claim **14** wherein crossing the transformed maintainer line with the original male sterile line produces a male sterile line possessing a portion of the PHB biosynthetic pathway.

16. The method of claim **15** wherein insertion of the phaA and phaC genes into the maintainer line and crossing with the original male cytoplasmic sterile line forms a male sterile line containing the phaA and phaC genes.

* * * * *