

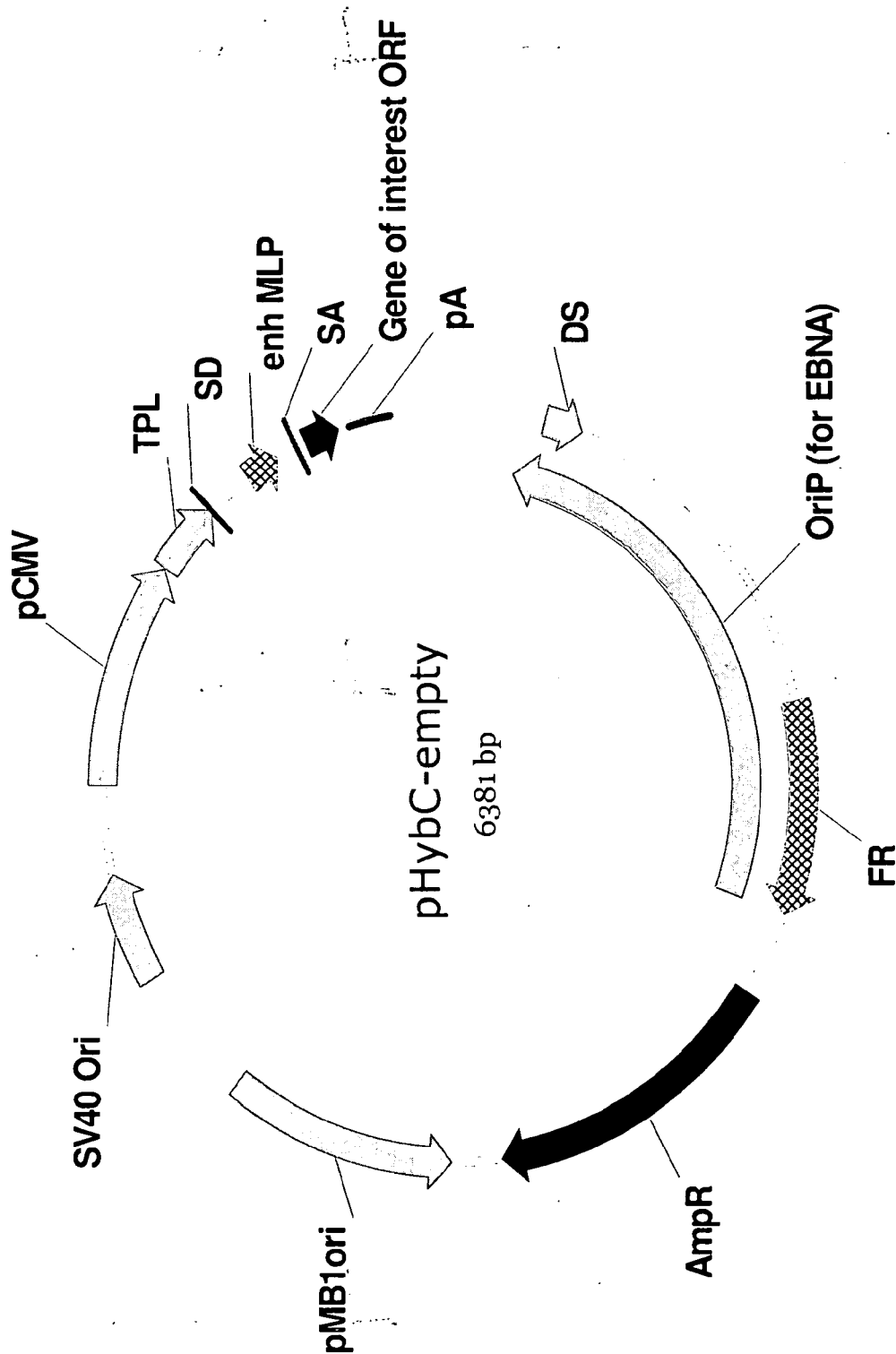


US 20130344537A1

(19) **United States**(12) **Patent Application Publication**
HSIEH(10) **Pub. No.: US 2013/0344537 A1**(43) **Pub. Date: Dec. 26, 2013**(54) **MAMMALIAN EXPRESSION VECTORS AND
USES THEREOF****Publication Classification**(71) Applicant: **Chung-Ming HSIEH**, Newton, MA
(US)(51) **Int. Cl.**
C12P 21/00 (2006.01)
C07K 16/24 (2006.01)(72) Inventor: **Chung-Ming HSIEH**, Newton, MA
(US)(52) **U.S. Cl.**
CPC **C12P 21/00** (2013.01); **C07K 16/241**
(2013.01)(21) Appl. No.: **13/886,341**USPC **435/69.6**; 435/320.1; 435/326; 435/69.1;
435/369; 435/365(22) Filed: **May 3, 2013****Related U.S. Application Data**(57) **ABSTRACT**(60) Continuation of application No. 13/368,618, filed on
Feb. 8, 2012, now Pat. No. 8,455,219, which is a divi-
sion of application No. 12/354,568, filed on Jan. 15,
2009, now Pat. No. 8,187,836.(60) Provisional application No. 61/021,282, filed on Jan.
15, 2008, provisional application No. 61/104,546,
filed on Oct. 10, 2008.

The present invention features nucleic acids for recombinant protein expression in mammalian cell culture. The episomal vectors of the invention promote high protein production in mammalian cells expressing the SV40 T Ag or Epstein-Barr virus nuclear antigen (e.g., COS7 or HEK293-6E cells). The methods and systems are useful, for example, in pharmaceutical drug development and cloning, especially for the production of antibodies.

Figure 1



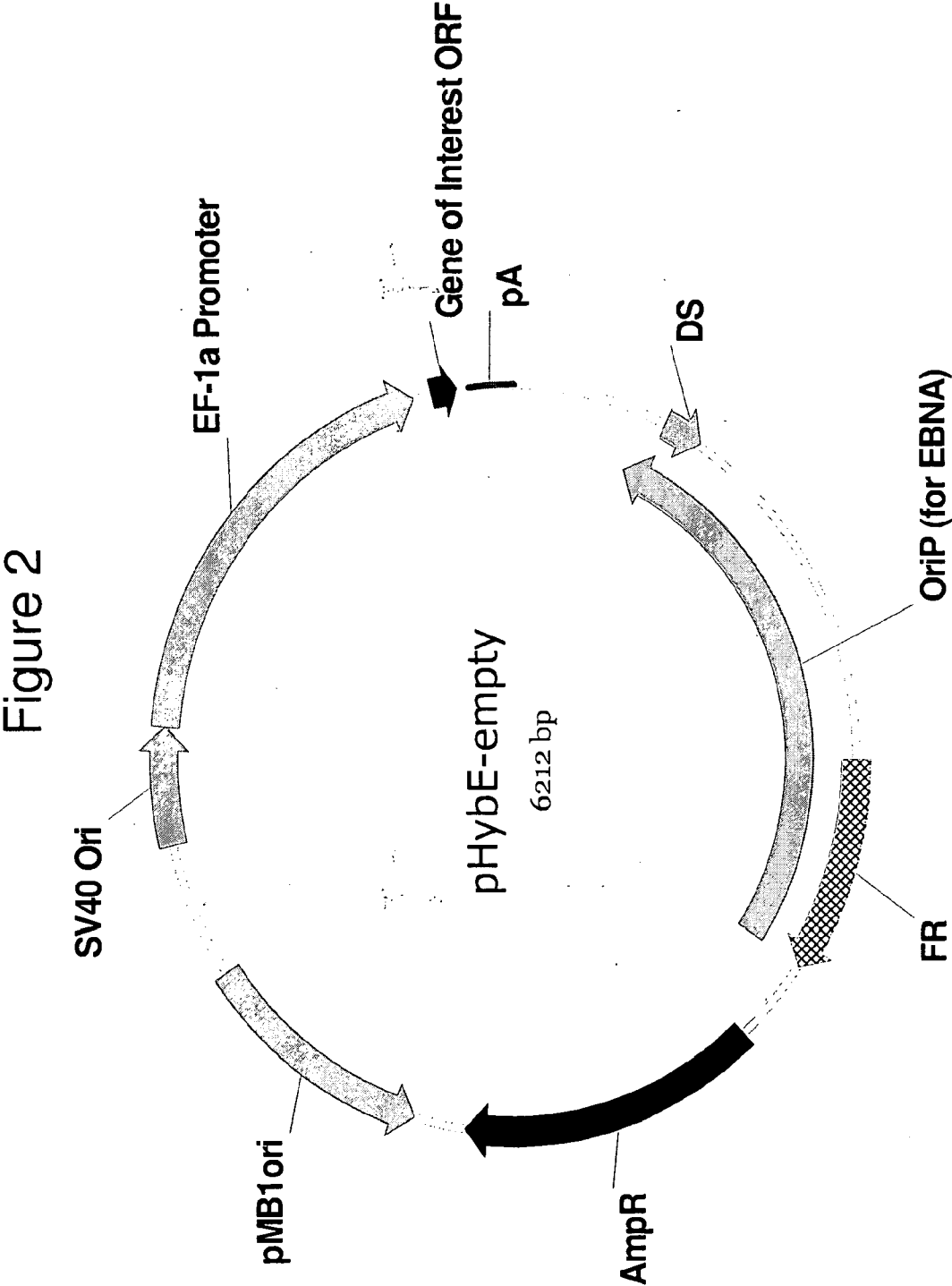


Figure 3

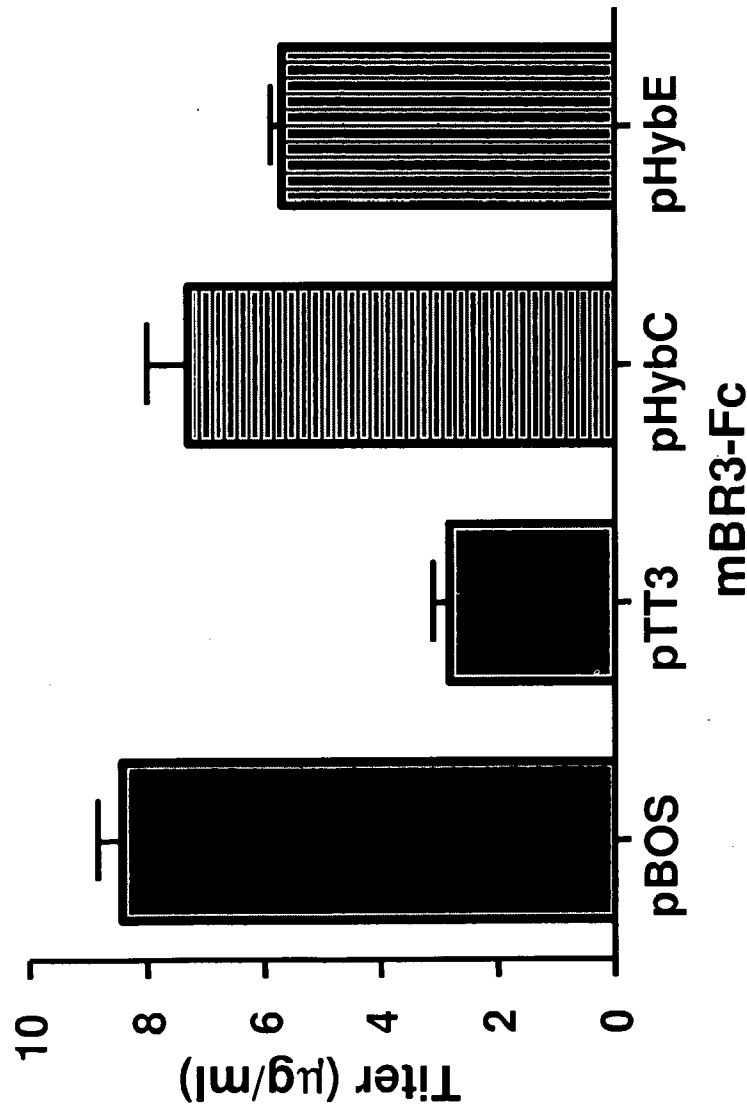


Figure 4

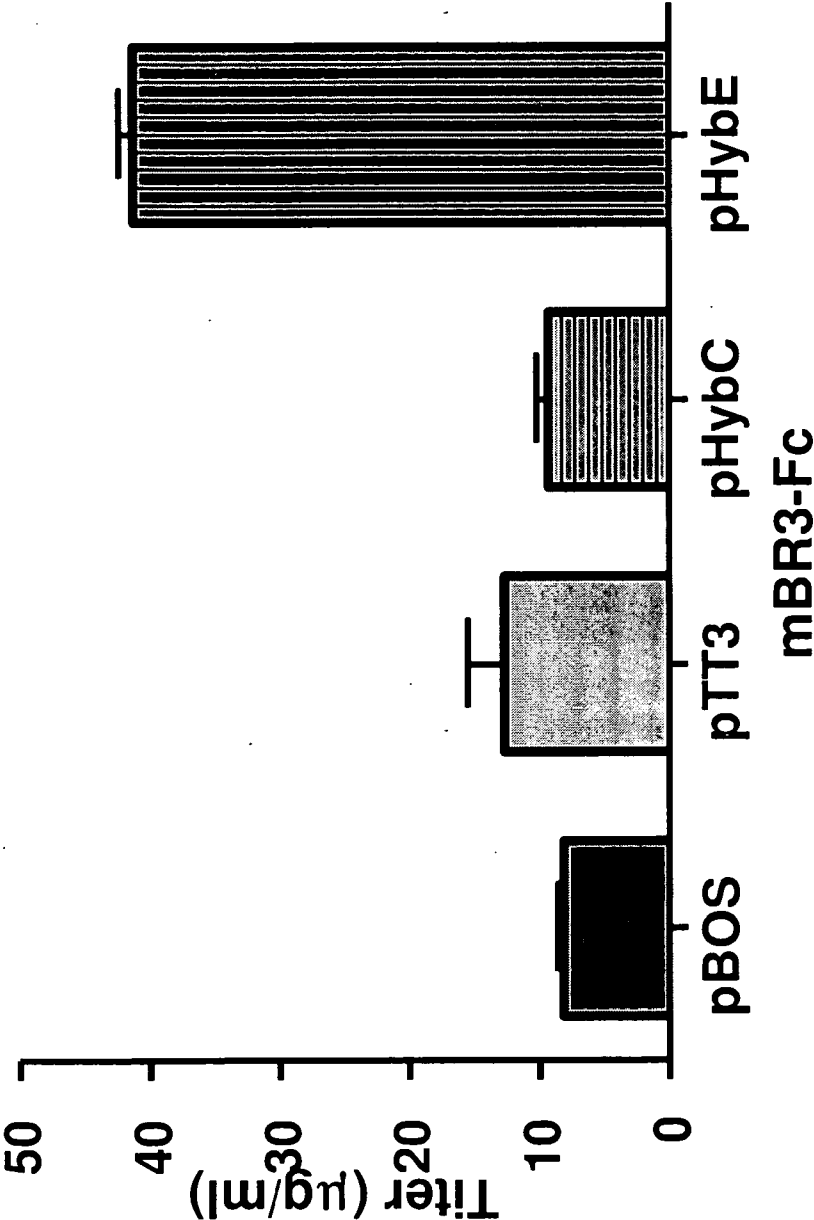


Figure 5

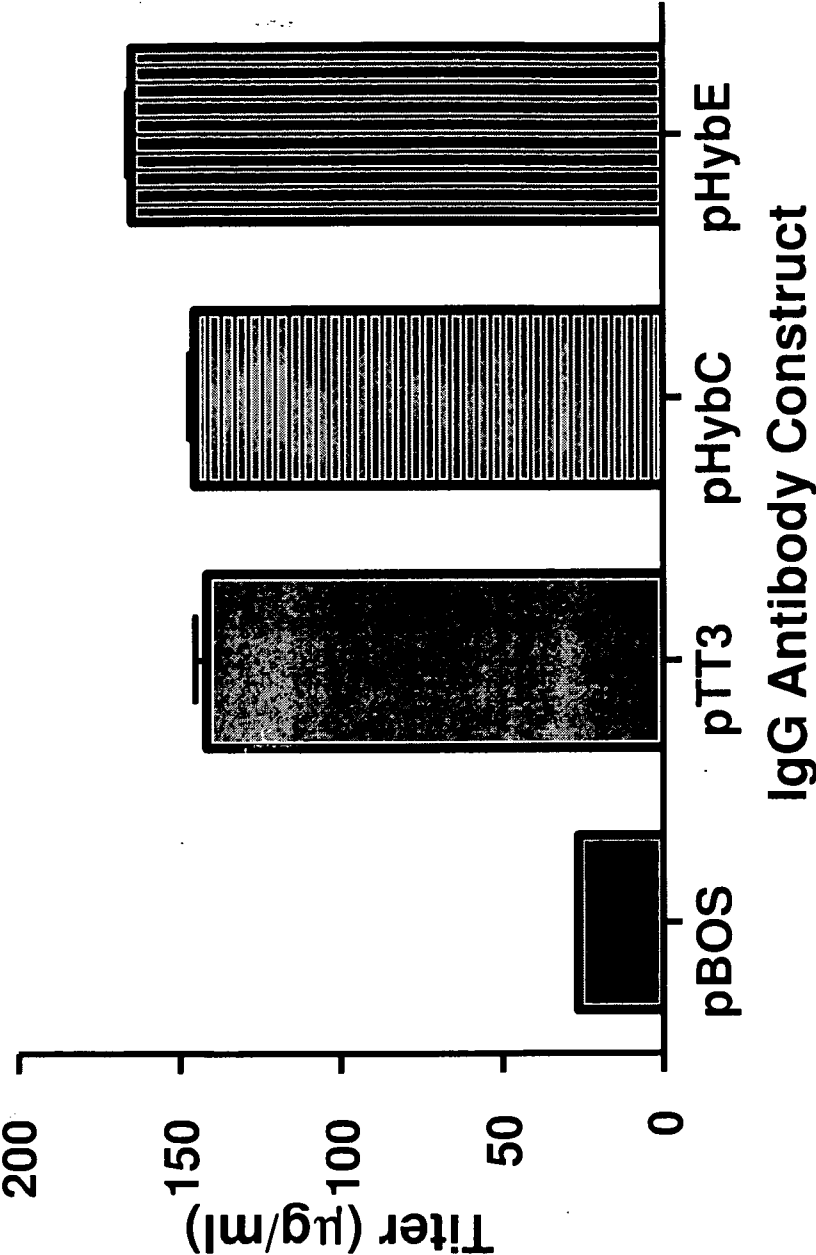


Figure 6

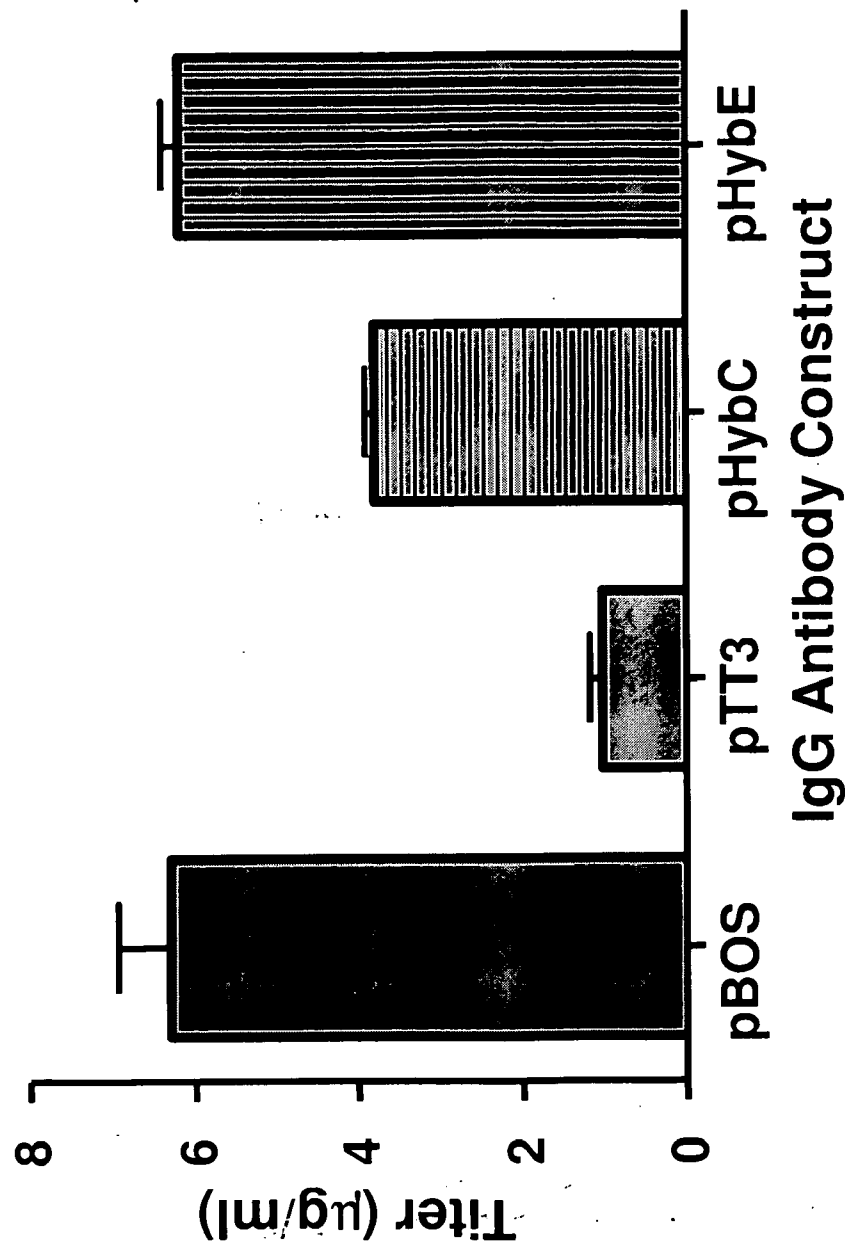


Figure 7

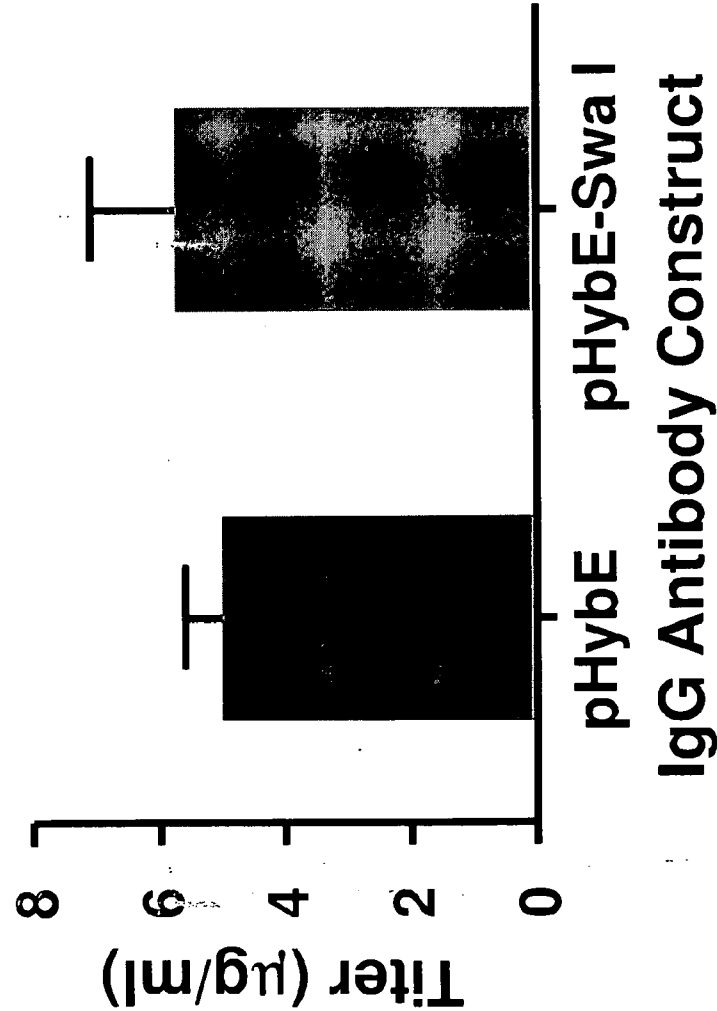


Figure 8

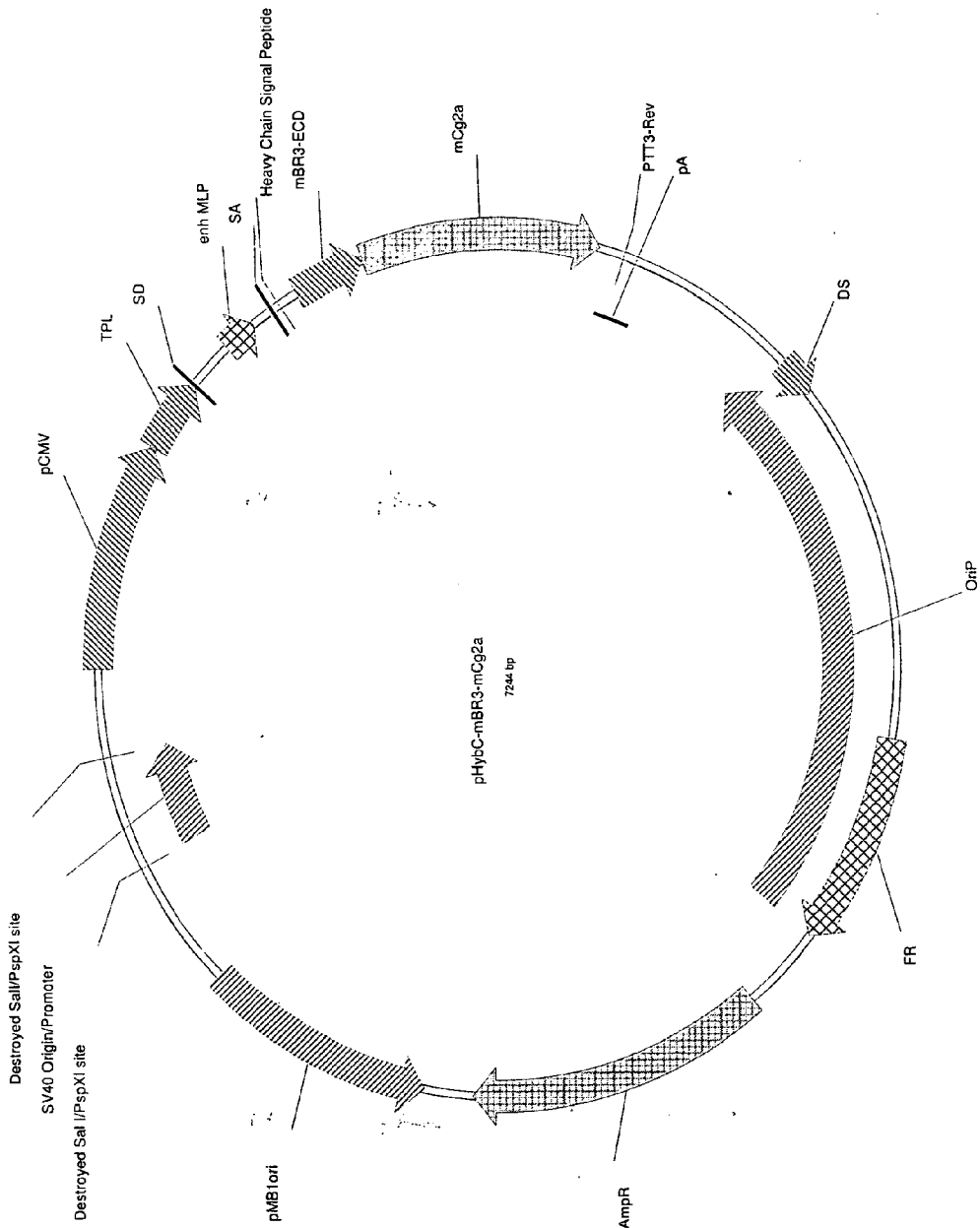


Figure 9

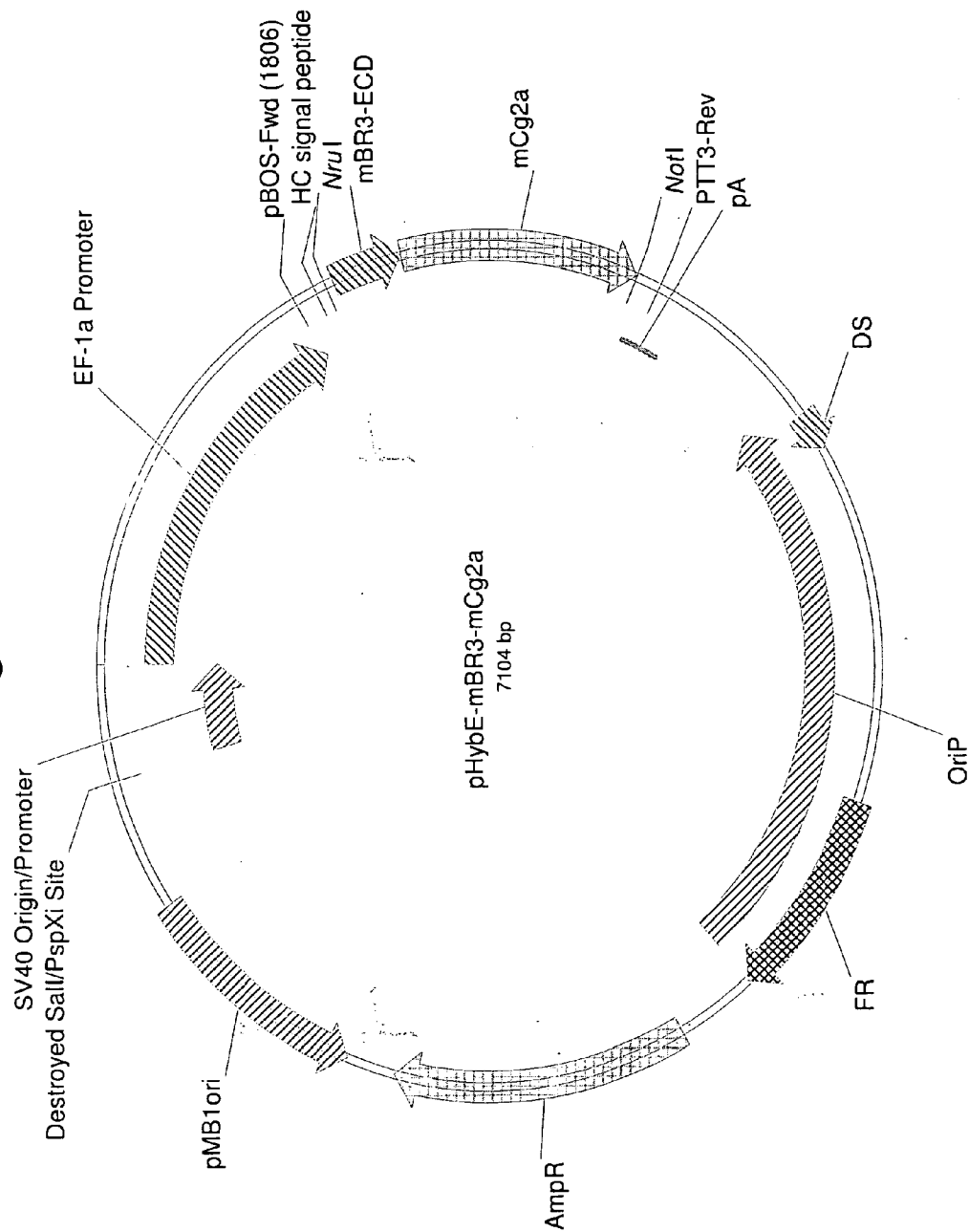


Figure 10

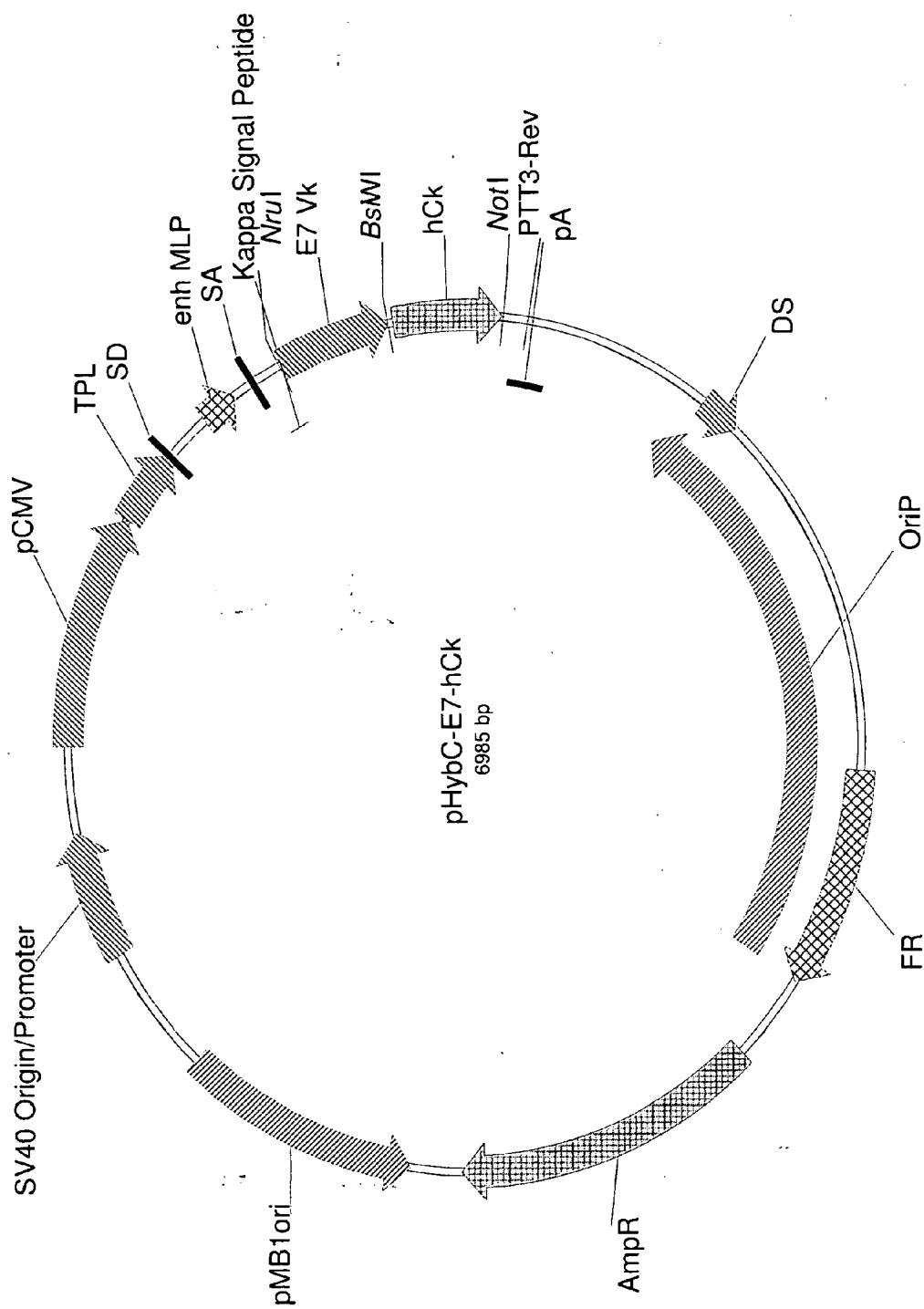


Figure 11

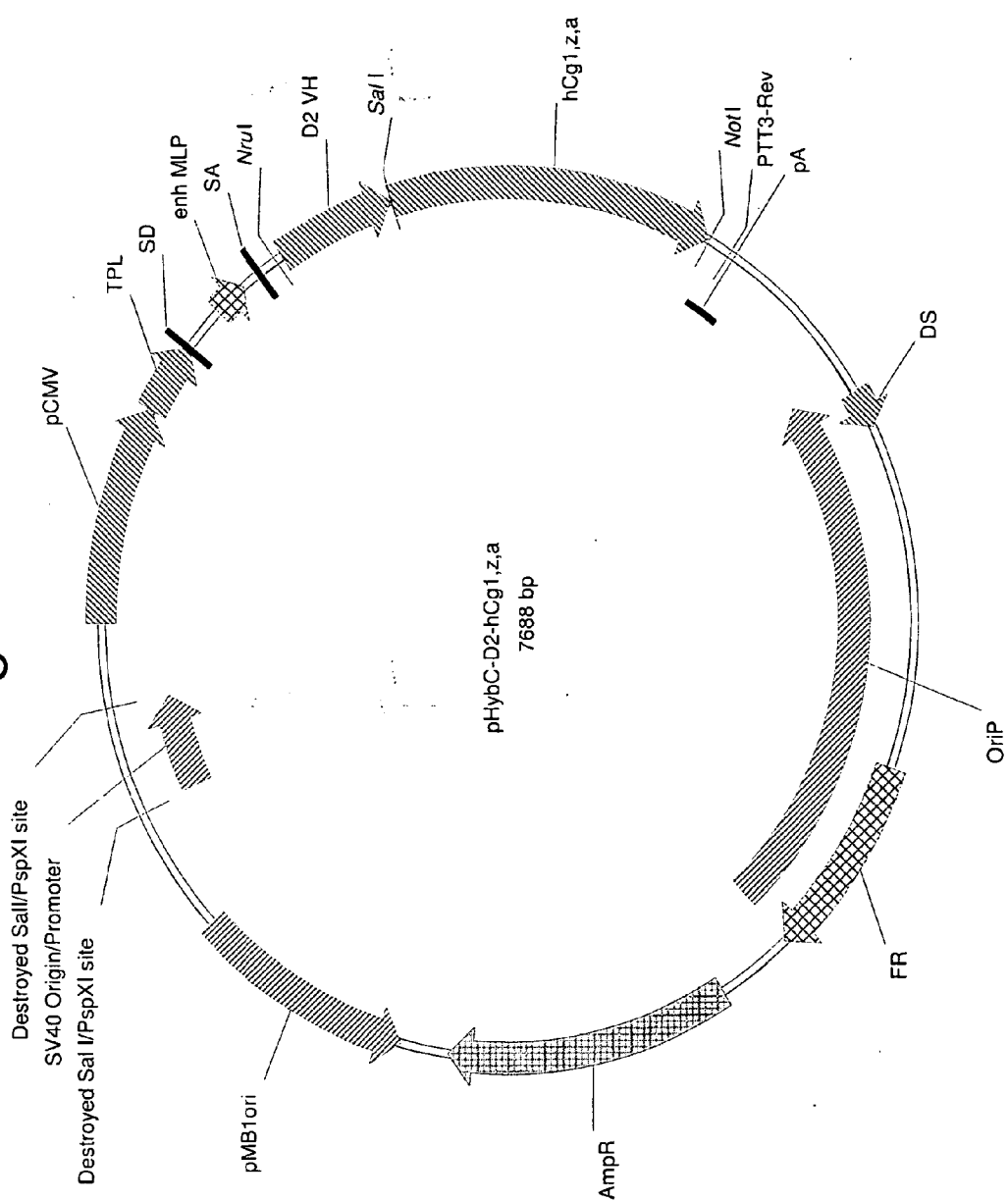


Figure 12

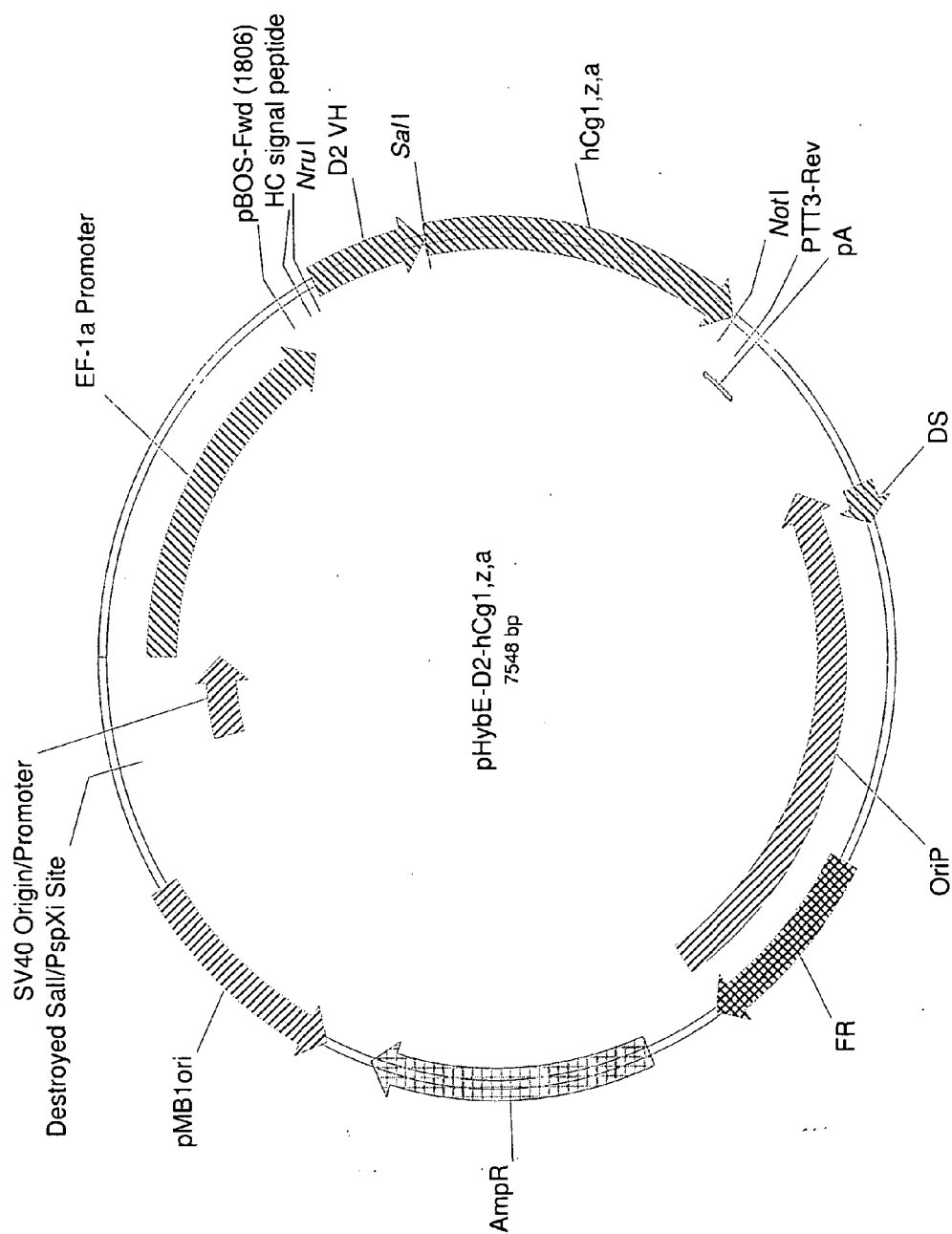


Figure 13

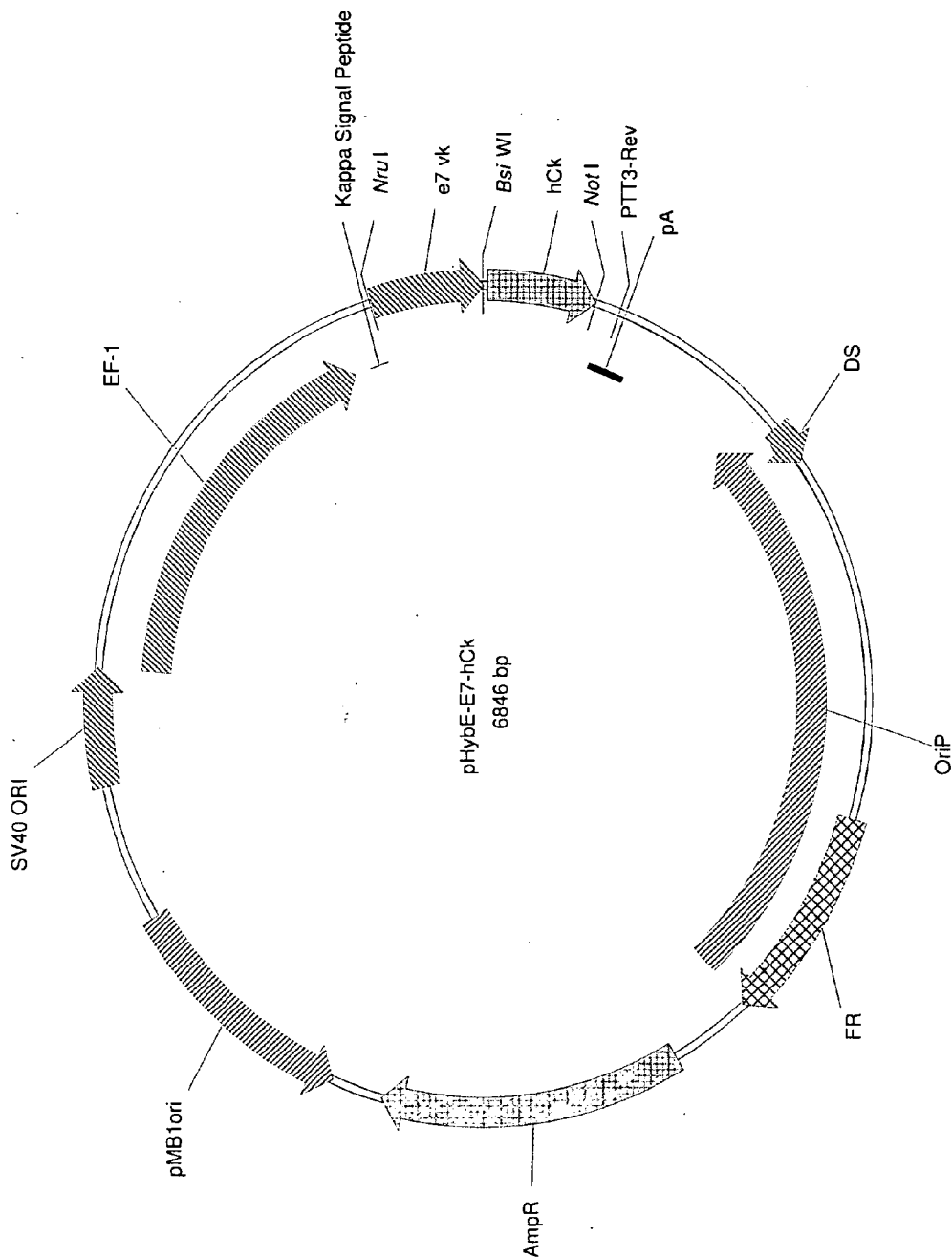


Figure 14

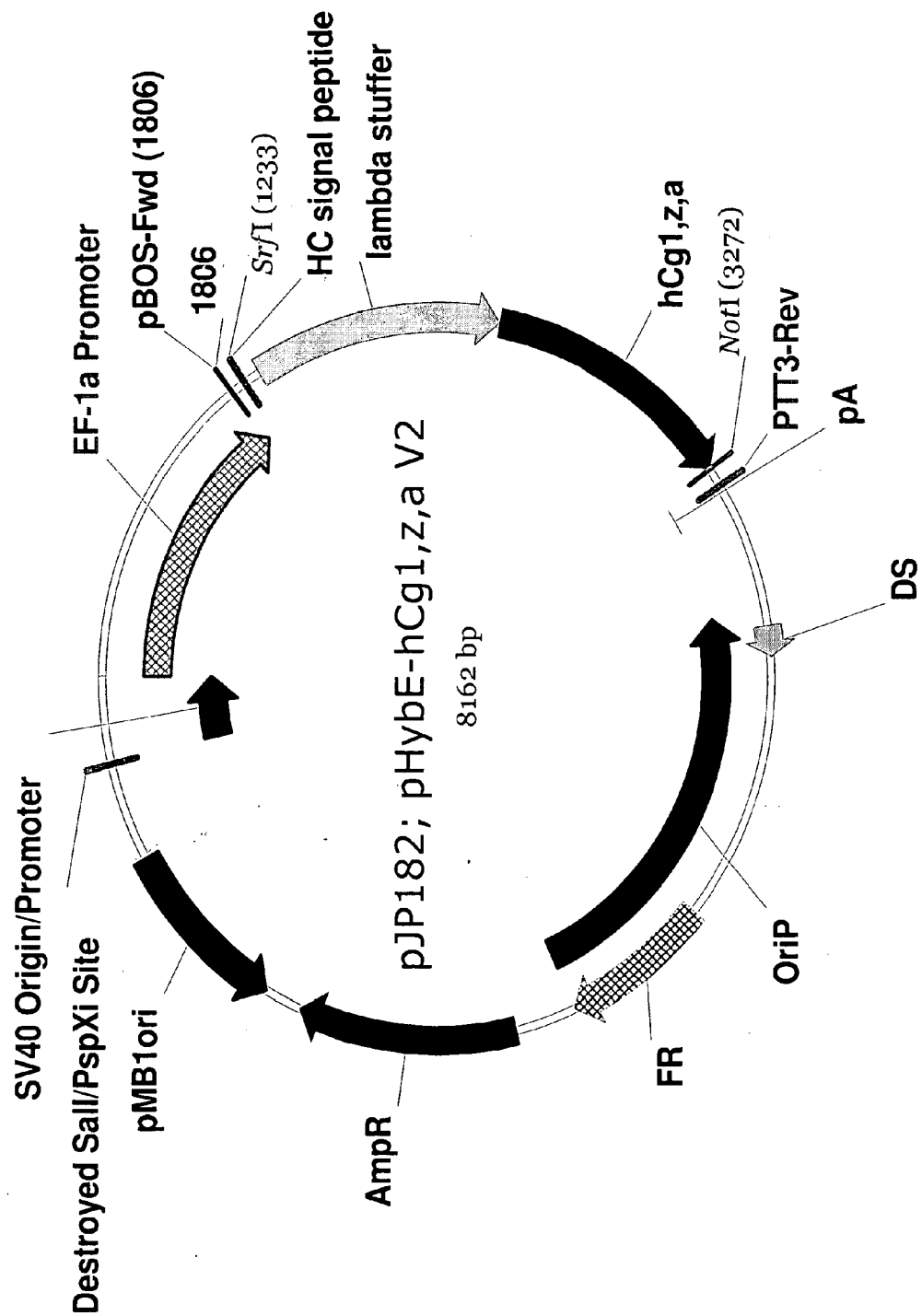


Figure 15

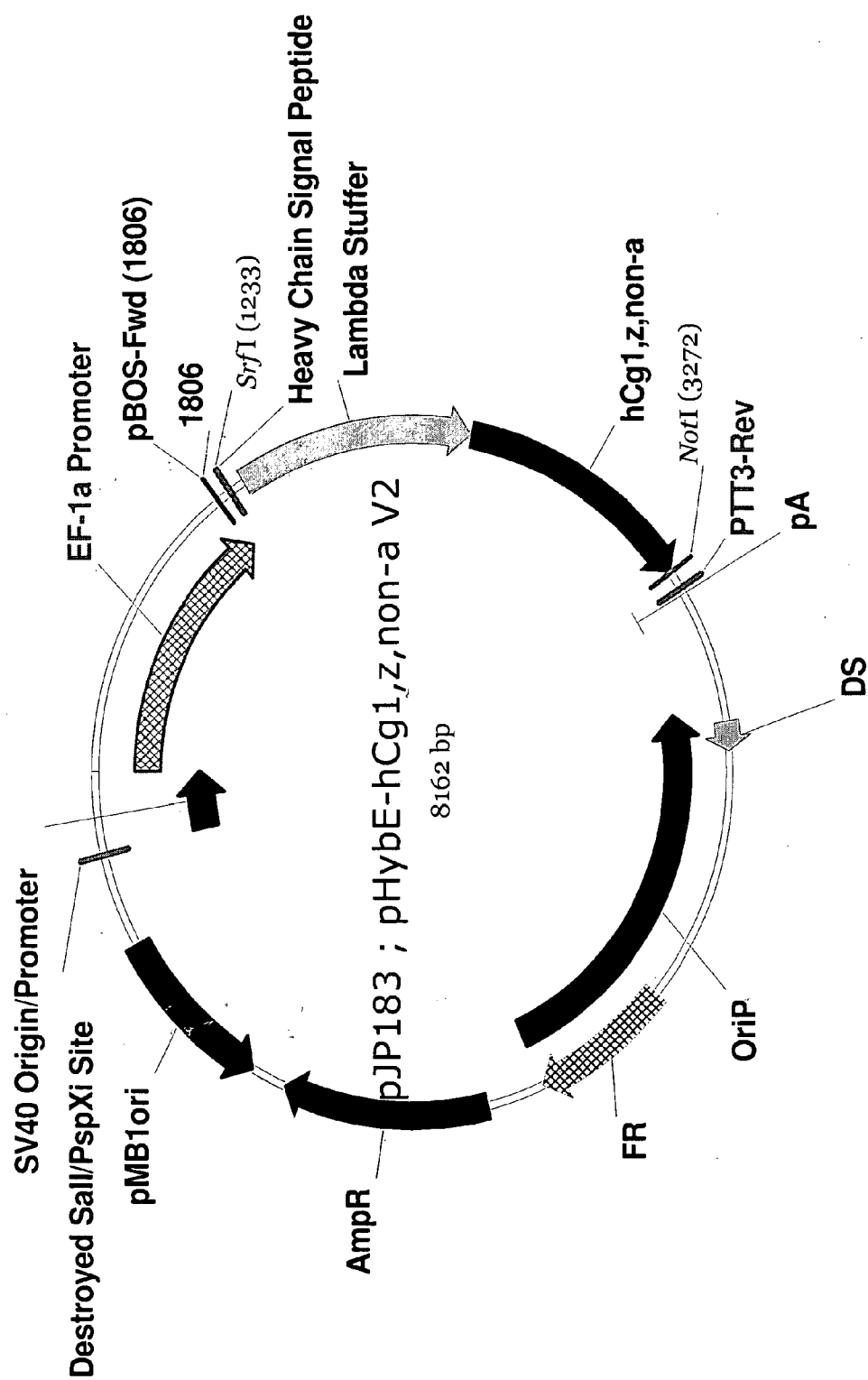


Figure 16

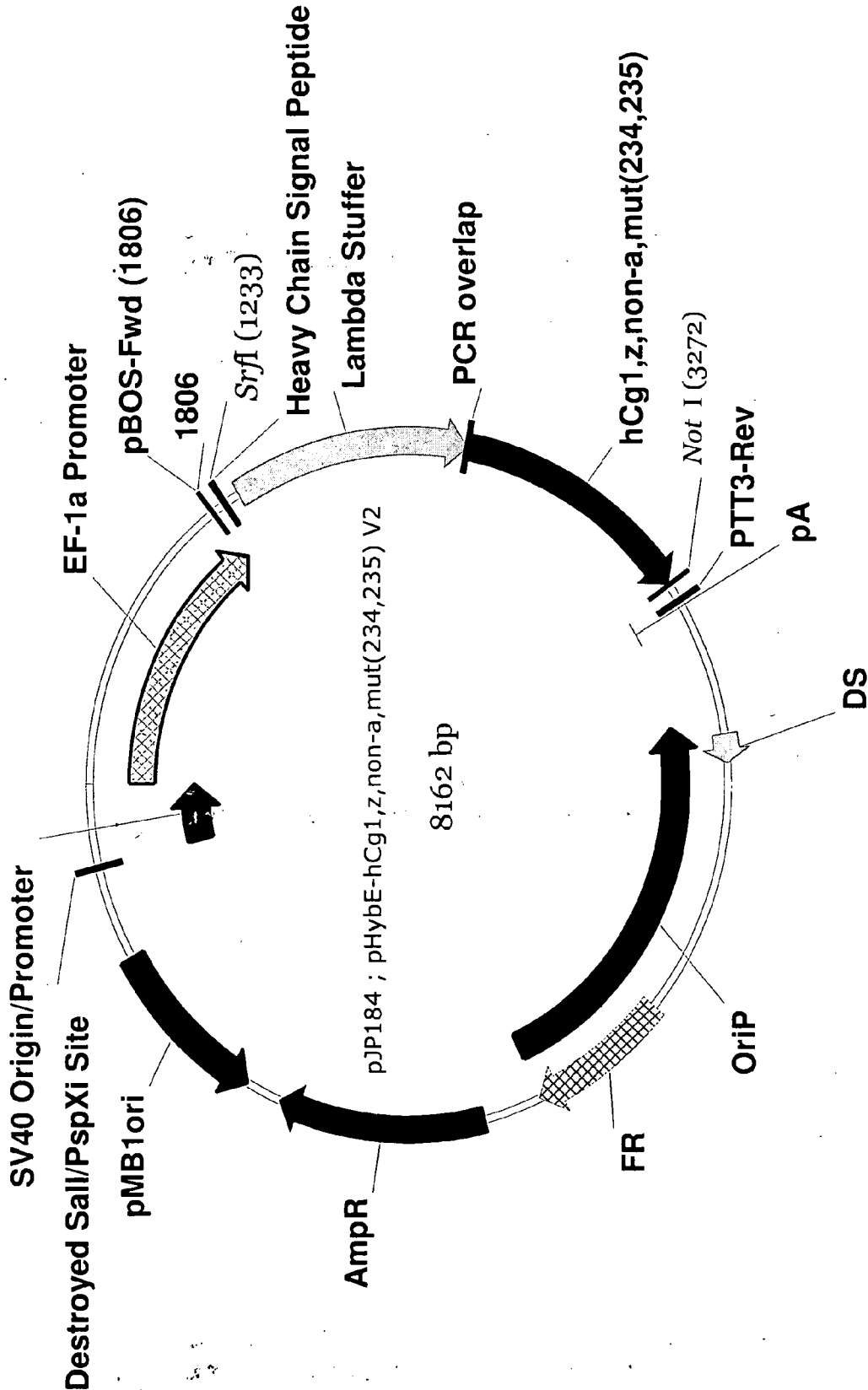


Figure 17

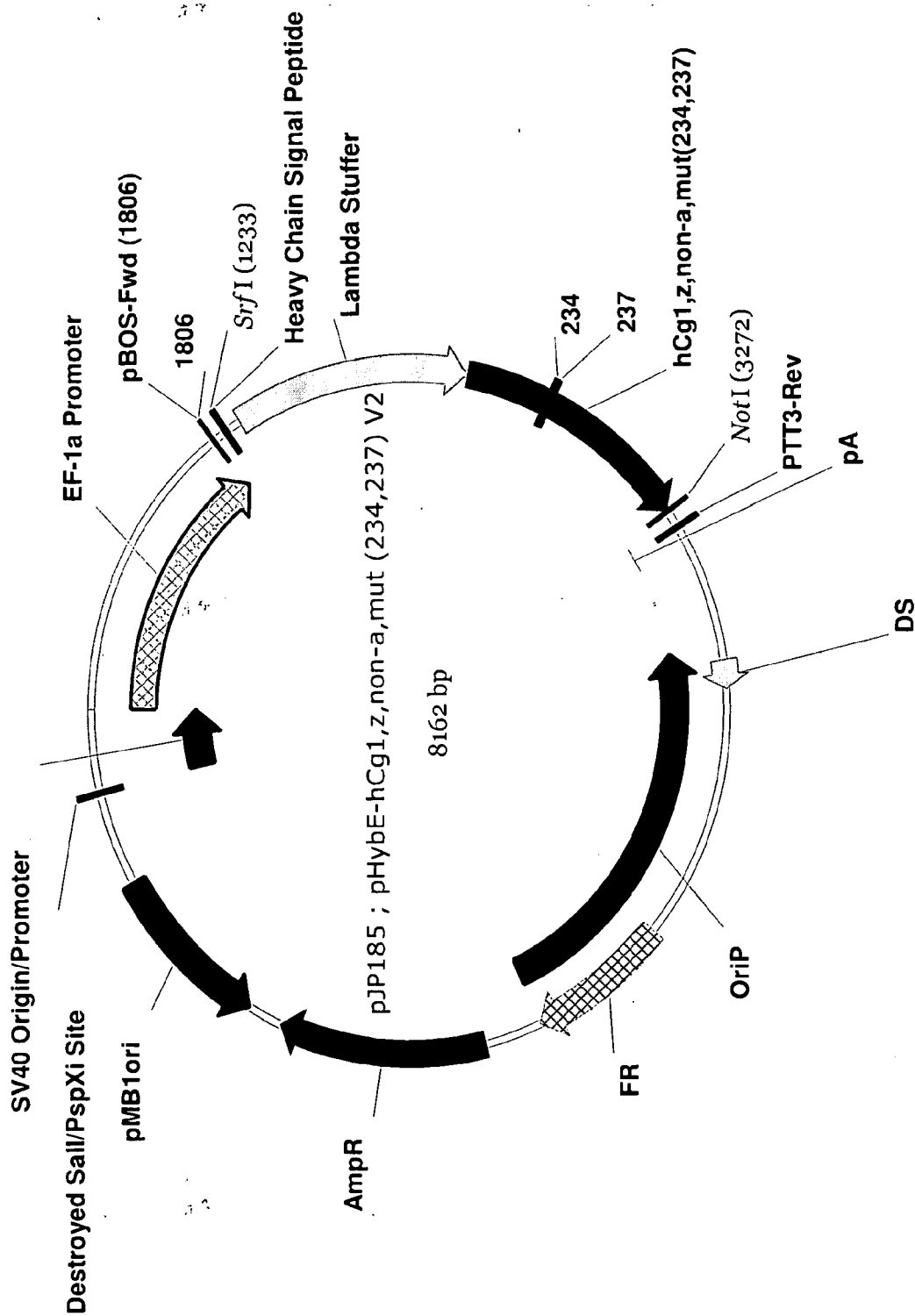


Figure 18

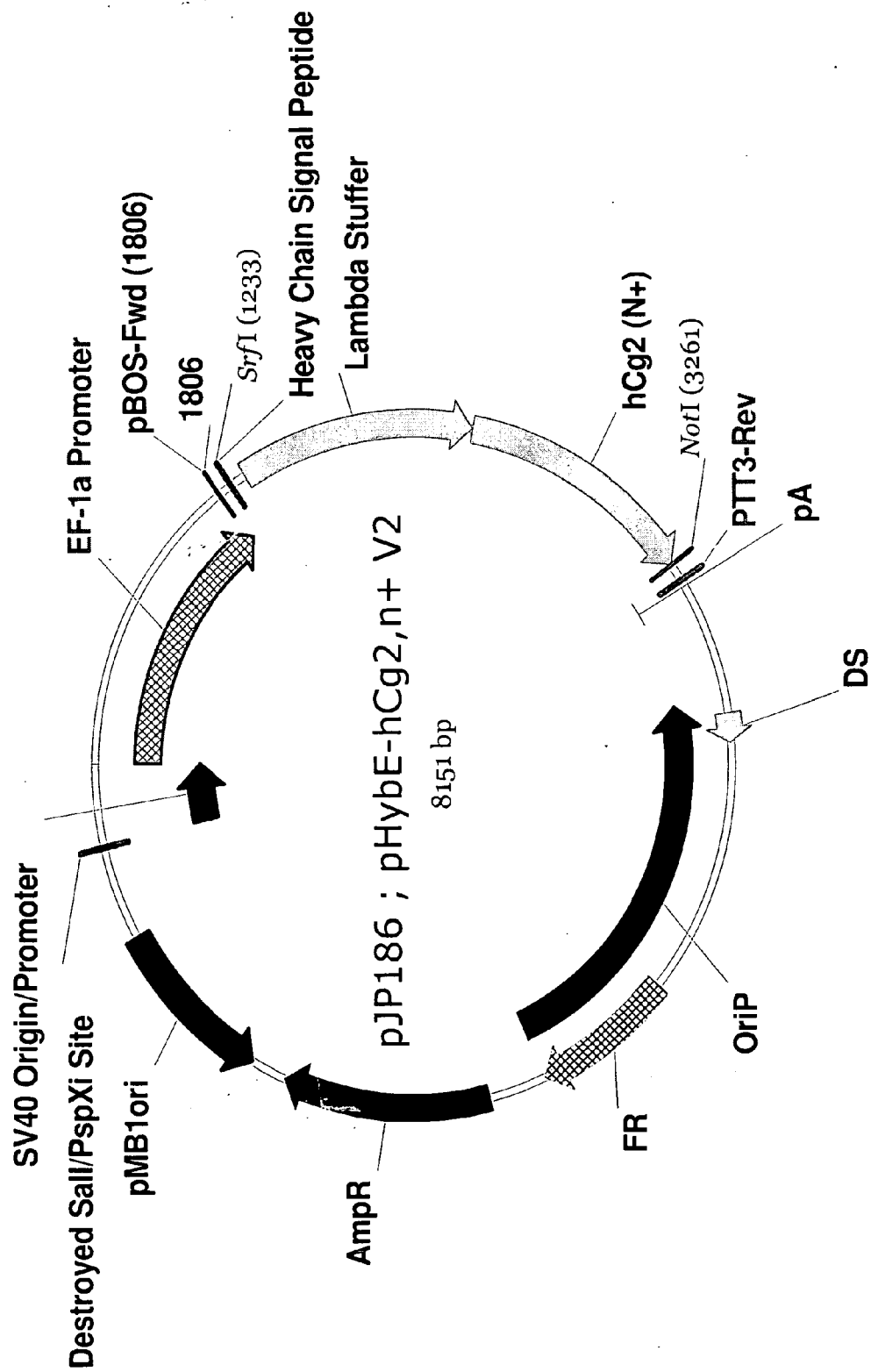


Figure 19

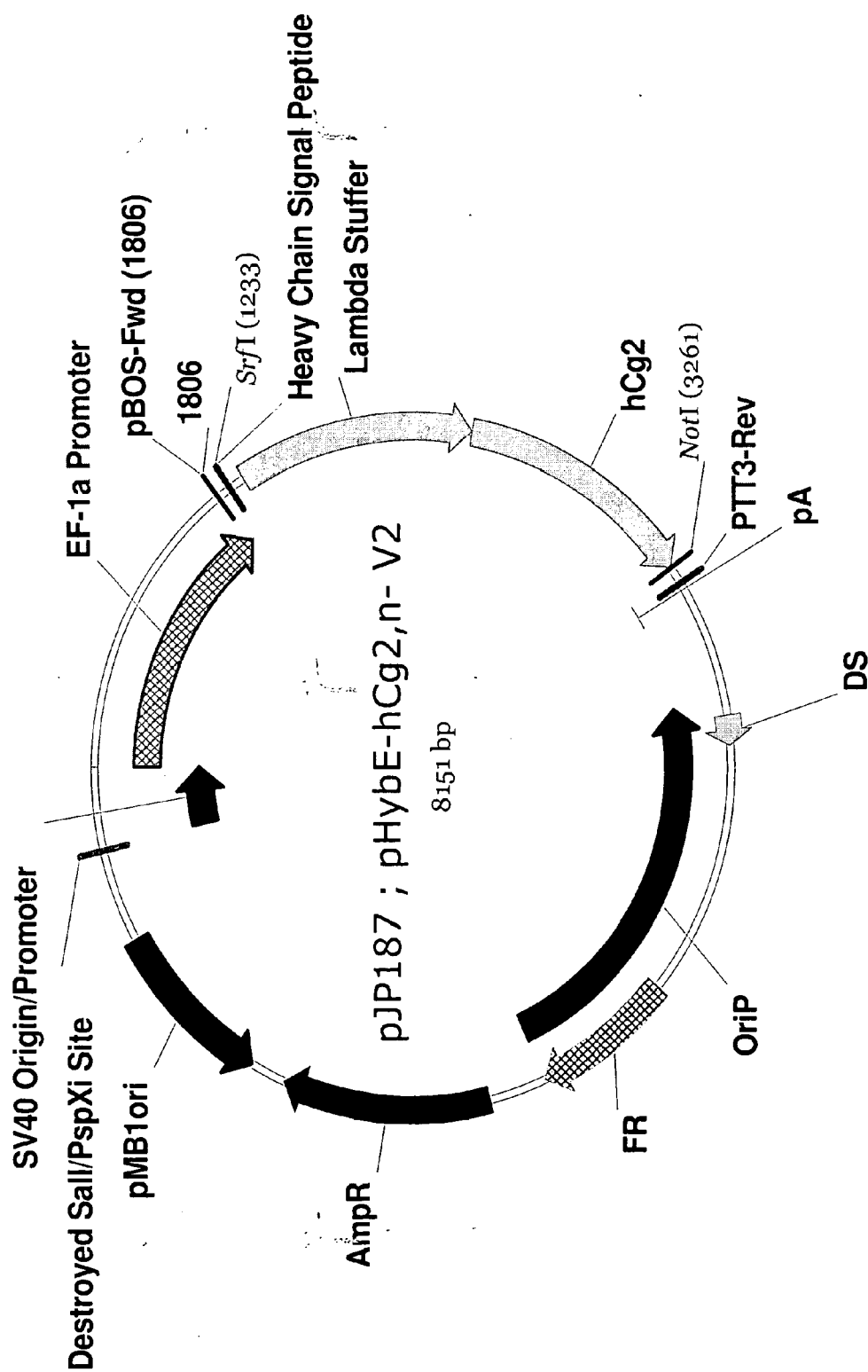


Figure 20

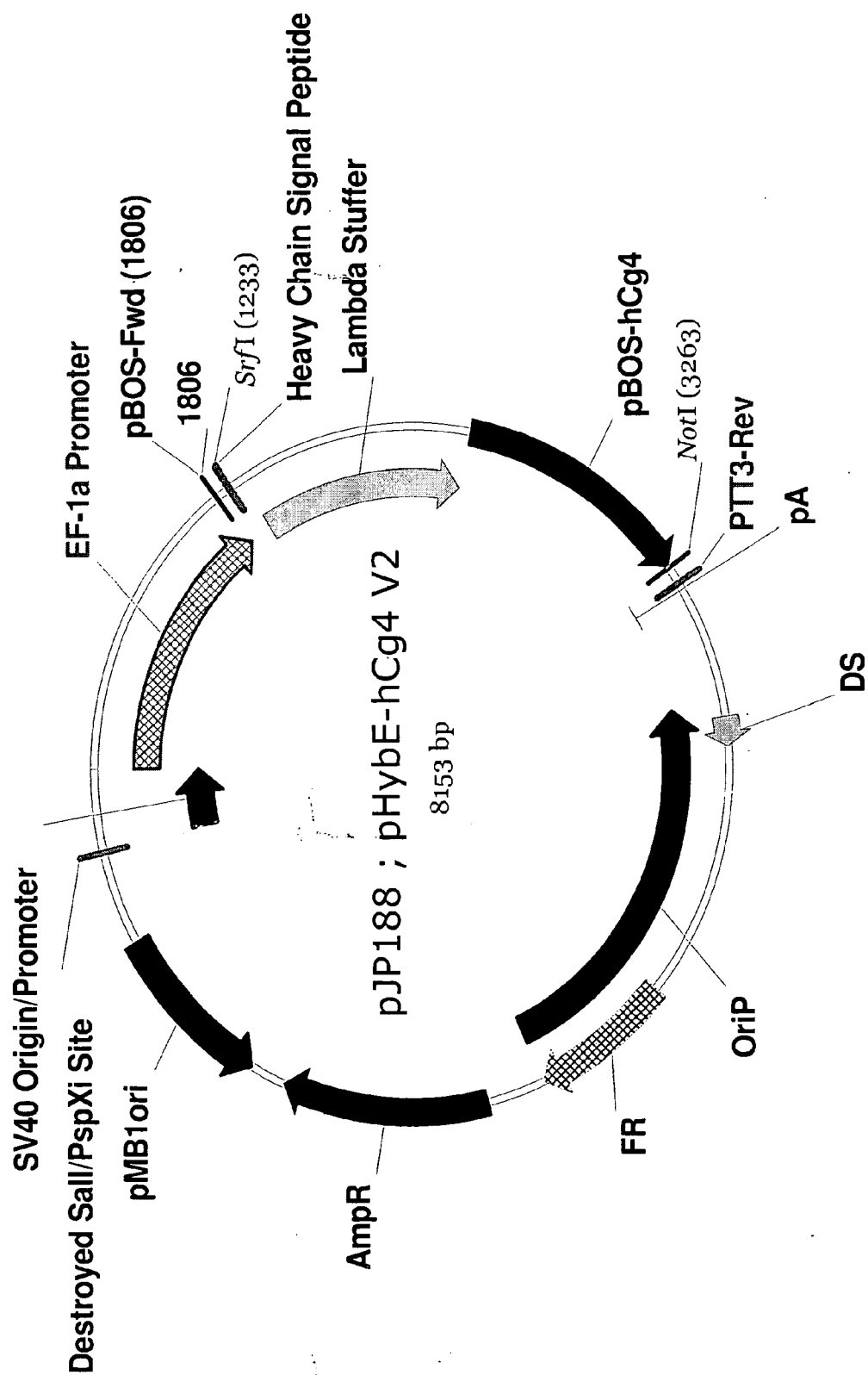


Figure 21

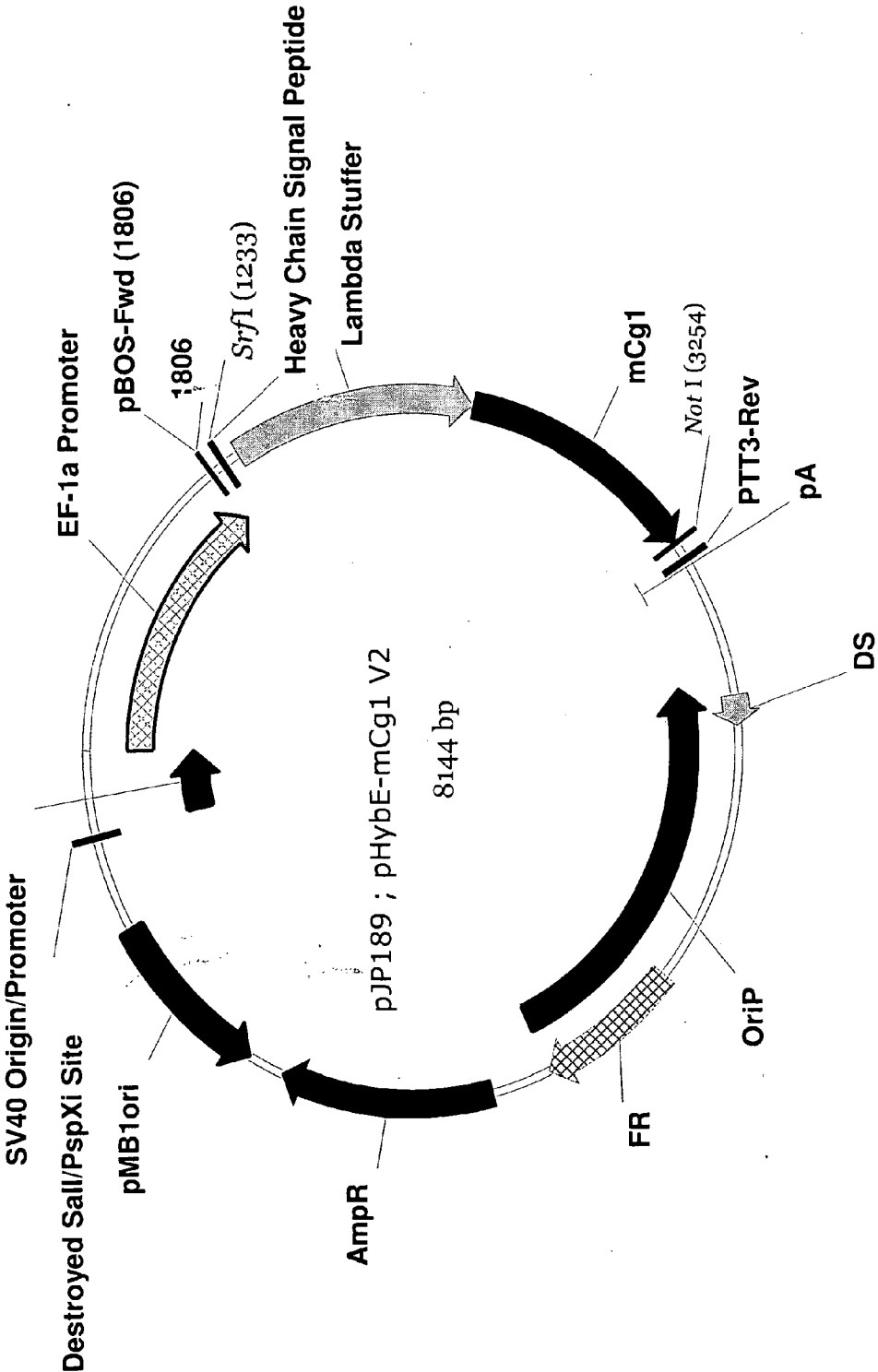


Figure 22

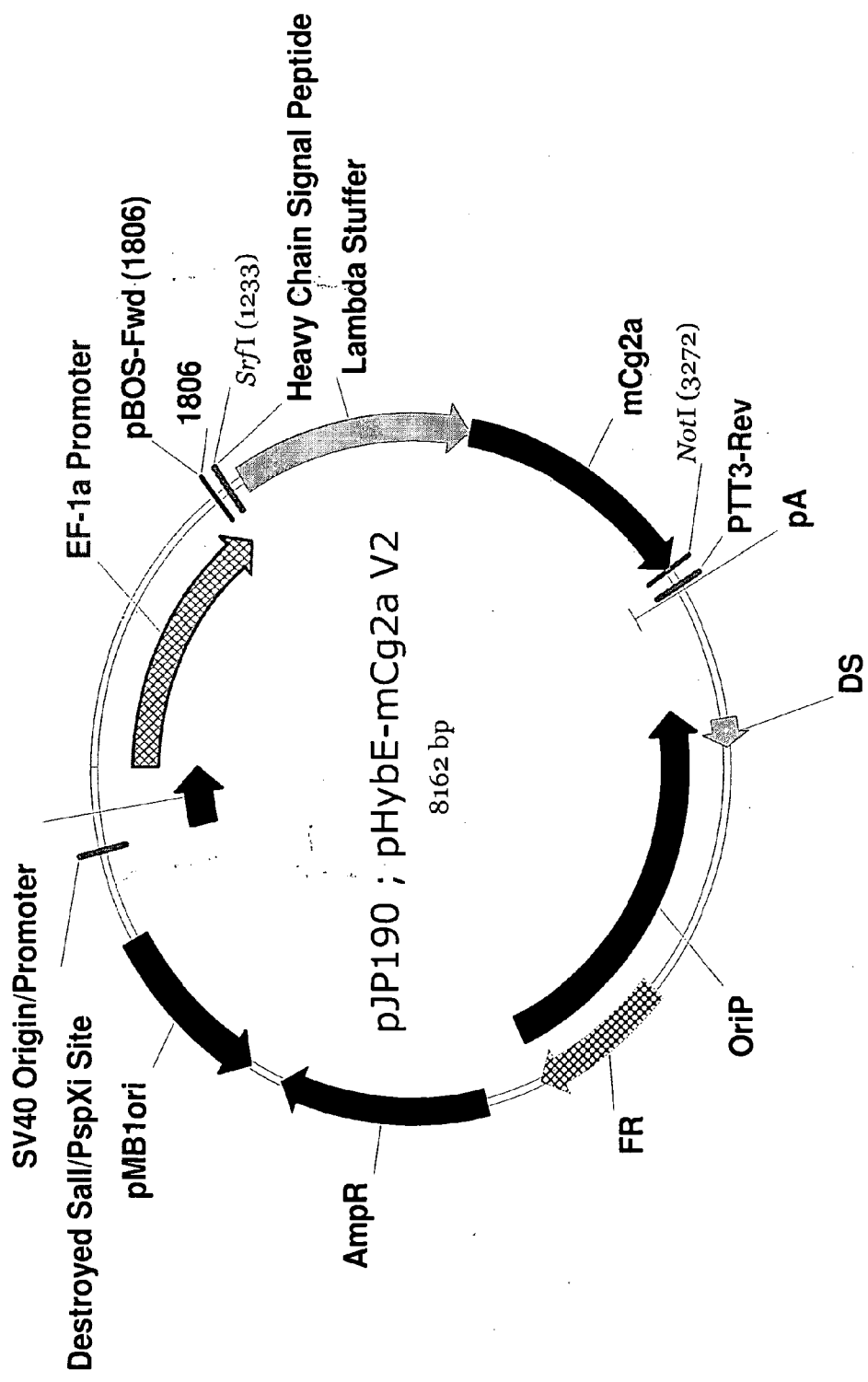


Figure 23

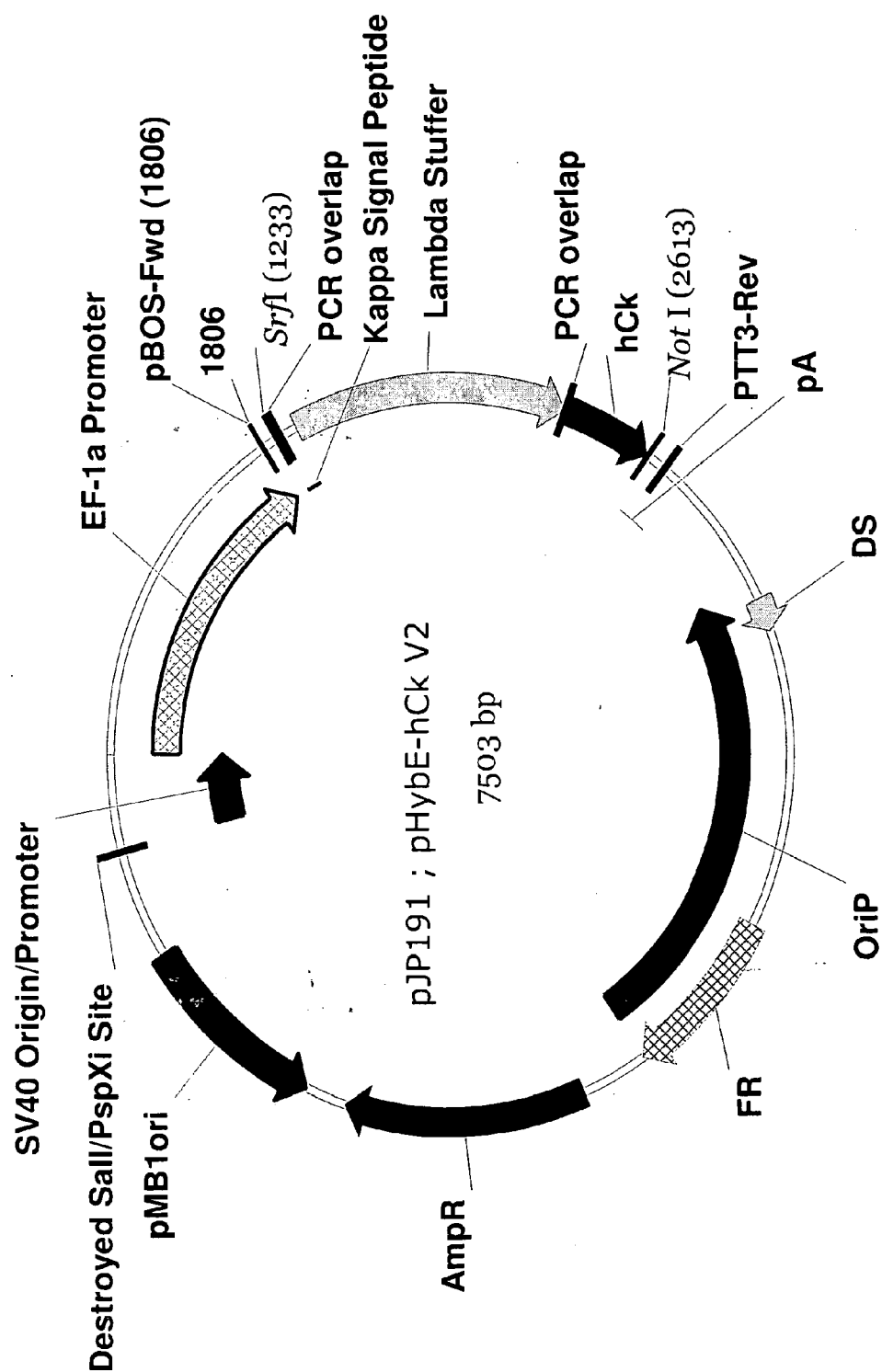


Figure 24

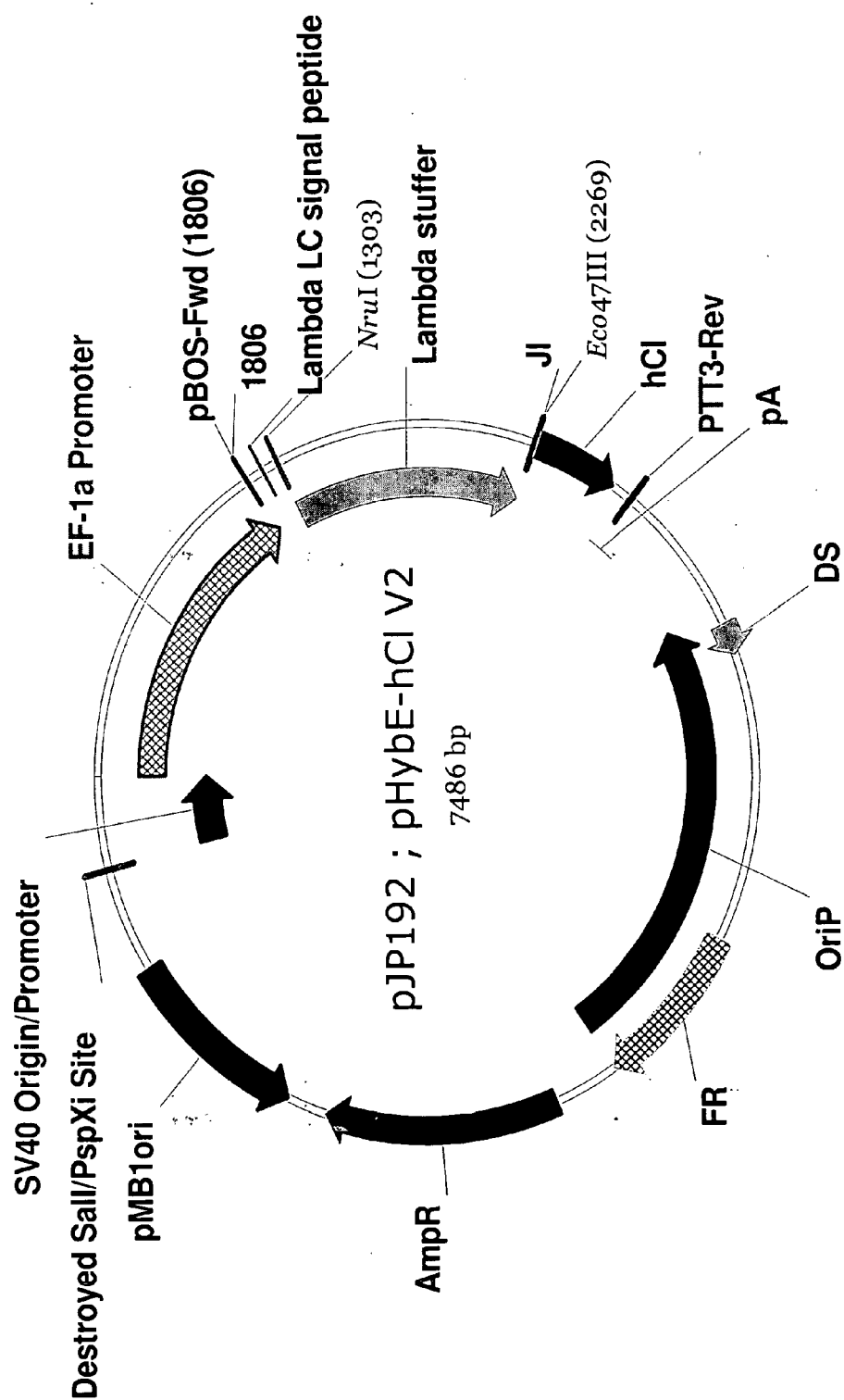
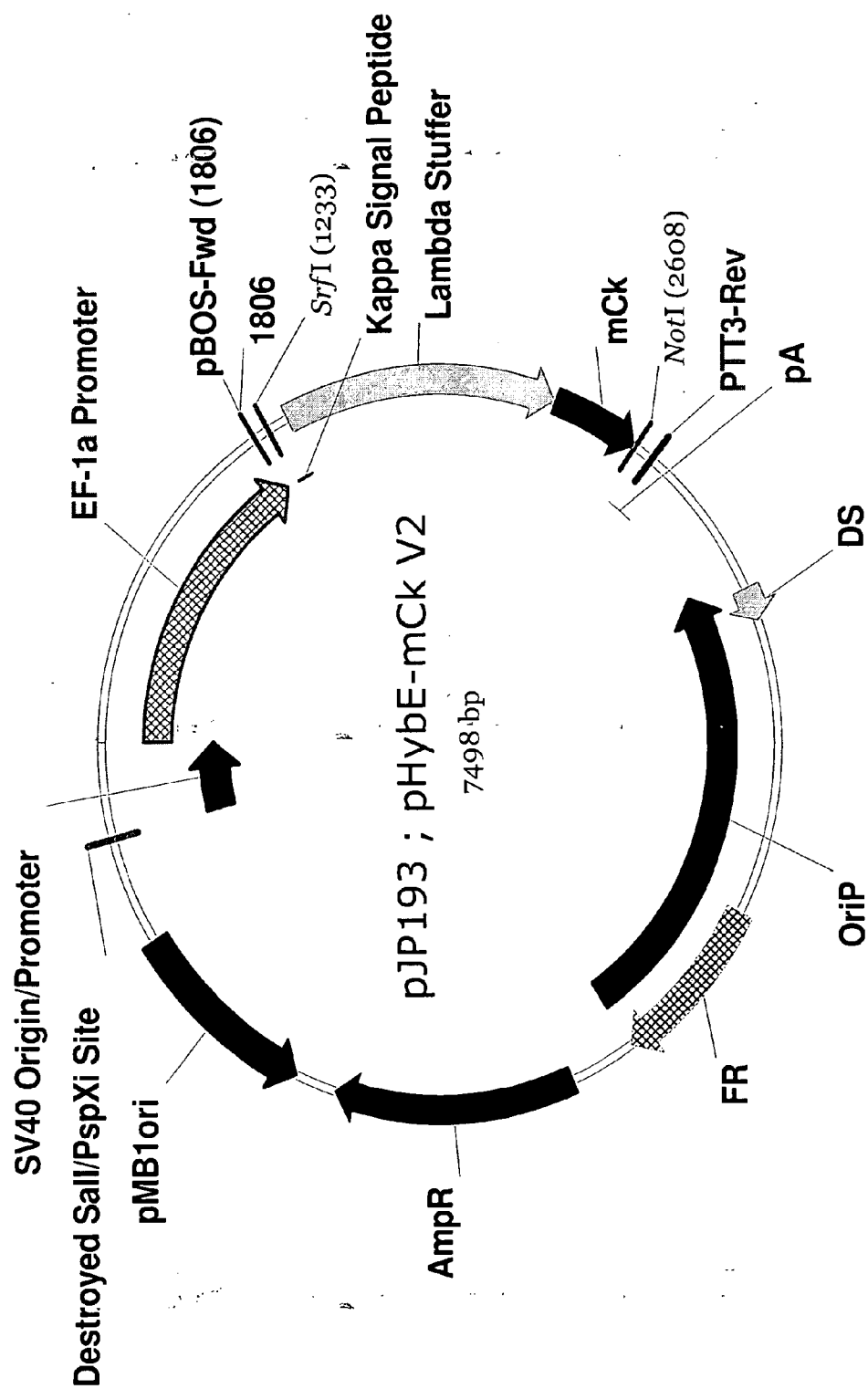


Figure 25



MAMMALIAN EXPRESSION VECTORS AND USES THEREOF

RELATED APPLICATIONS

[0001] This application is a continuation application of U.S. patent application Ser. No. 13/368,618, filed on Feb. 8, 2012, which is a divisional of U.S. patent application Ser. No. 12/354,568, filed on Jan. 15, 2009 (now U.S. Pat. No. 8,187,836), which claims priority to U.S. provisional application Ser. No. 61/021,282, filed on Jan. 15, 2008, and to U.S. provisional application Ser. No. 61/104,546, filed on Oct. 10, 2008, the contents of each of which are hereby incorporated in their entirety.

BACKGROUND OF THE INVENTION

[0002] Stable production of proteins, including biologics, can be accomplished by transfecting host cells with vectors containing DNA that encodes the protein. Maintenance of the vector in the cell line can be achieved through a variety of means, including extrachromosomal replication through episomal origins of replication. Episomal vectors contain an origin of replication that promotes replication of the vector when the sequence is bound by a replication initiation factor. Episomal vectors have several advantages over vectors that require insertion into the host genome. For example, episomal vectors decrease phenotypic changes in the cell that may result from integration of a vector into the host genome. Episomal vectors may also be isolated from the transfected cells using standard DNA extraction protocols.

[0003] With the evolving importance of therapeutic proteins, i.e., biologics, efforts must be made to optimize protein production, while improving efficiency of the overall production process. Thus, improvements in efficiency must be weighed against the protein production capacity of the vector. There is a need for better expression systems that provide efficient cloning options, as well as high levels of the desired protein product. It would be advantageous to decrease the number of cloning steps involved in the production of biologics, especially antibodies, to improve time requirements and minimize cost. It would also be advantageous to provide vectors that provide adequate protein production for both small and large scale cell cultures. The present invention overcomes the limitations of conventional vectors, by providing additional advantages that will be apparent from the detailed description below.

SUMMARY OF THE INVENTION

[0004] Recombinant proteins may be produced by mammalian cell transient transfection, especially during the pharmaceutical drug discovery process. A variety of host cells may be used to express proteins, including mammalian cells such as COS and human embryonic kidney (HEK) cells. Episomal vectors rely on both an origin of replication and a trans-acting replication initiation factor that binds the origin. Replication initiation factors, such as Epstein-Barr virus nuclear antigen (EBNA) that binds the OriP of the Epstein-Barr virus, may be cloned into the episomal vector, or, alternatively, may be expressed by the host cell into which the vector is transfected. Thus, episomal vectors may be specific to certain cell lines that express the trans-acting factor required to activate replication through the origin of replication.

[0005] The present invention eliminates the need for different episomal vector backbones for recombinant protein expression. The present invention provides episomal vectors comprising at least two different episomal origins of replication, which allow the same vector to be used in different cell types for protein expression. Different origins of replication allow the vector to be used in different types of mammalian cells that provide the necessary trans-acting replication factors and allow the vector to replicate. By eliminating the need to re-clone the gene of interest for protein production, the instant invention improves efficiency and reduces costs associated with multiple vectors, while at the same time maintaining protein production levels. A surprising aspect of the invention is that the addition of nucleotides to the vector, i.e., a second origin of replication, does not negatively impact the vector's ability to produce protein at the desired levels.

[0006] In a preferred embodiment, the vectors of the invention comprise antibody heavy or light chain constant regions. Thus, an antibody light or heavy chain variable region may be cloned into the vector upstream of the light or heavy chain constant region, respectively, further improving the efficiency of the expression system. The episomal vectors promote high protein production in mammalian cells expressing the SV40 T Ag or Epstein-Barr virus nuclear antigen (e.g., COS7 or HEK293-6E cells).

[0007] The present invention provides an optimal combination of elements for protein yield, production efficiency, and reduced cost, which are all important elements for protein production, especially in the pharmaceutical industry and the production of biologic proteins, such as antibodies. Other features and advantages of the invention are described in the detailed description and claims below.

[0008] In one aspect, the invention provides an expression vector comprising: a) an OriP origin of replication derived from Epstein-Barr virus (EBV); (b) an SV40 origin of replication; (c)

[0009] an insertion site for inserting a gene of interest; and (d) a nucleic acid sequence encoding an antibody heavy or light chain constant region, operably linked to the insertion site. In an embodiment, the gene of interest is an antibody heavy or light chain variable region, for example, a murine, a humanized, a chimeric or a human antibody heavy or light chain variable region. In a particular embodiment, the antibody heavy chain variable region is the heavy chain variable region of an antibody selected from the group consisting of adalimumab, ABT-325, and ABT-874. In another particular embodiment, the antibody light chain variable region is the light chain variable region of an antibody selected from the group consisting of adalimumab, ABT-325, and ABT-874. The antibody heavy chain constant region is murine, humanized, chimeric or human, for example, and may be an antibody heavy constant region is selected from the group consisting of gamma 1, z, a; gamma 1, z, non-a; gamma 2, n+; gamma 2, n-; and gamma 4. The gamma 1, z, non-a antibody heavy chain constant region may further comprise an alanine mutation at position 234 of the heavy chain constant region. In another embodiment, the gamma 1, z, non-a antibody heavy chain constant region may further comprise an alanine mutation at either position 235 or 237 of the antibody heavy chain constant region.

[0010] In an embodiment, the antibody light chain constant region is a human kappa isotype or a human lambda isotype. In an embodiment, the antibody heavy chain constant region is a murine gamma 1 isotype or a murine gamma 2a isotype.

In another embodiment, the antibody light chain constant region is a murine kappa isotype. In an embodiment, the antibody heavy chain constant region is an Fc domain. In an embodiment, the heavy or light chain antibody variable region is 5' to the insertion site.

[0011] In an embodiment, the expression vector further comprises a promoter operably linked to the insertion site, wherein the promoter is either an EF-1 α promoter or a cytomegalovirus (CMV) promoter.

[0012] In an embodiment, the expression vector further comprises a selectable marker, such as an ampicillin resistance gene.

[0013] In an embodiment, the CMV promoter comprises a nucleic acid sequence that is at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to nucleotides 1 to 608 of SEQ ID NO: 1. In a particular embodiment, the CMV promoter comprises nucleotides 1 to 608 of SEQ ID NO: 1.

[0014] In an embodiment the EF-1 α promoter is human. In an embodiment, the EF-1 α promoter comprises a nucleic acid sequence that is at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to nucleotides 76 to 1267 of SEQ ID NO: 2. In a particular embodiment, the EF-1 α promoter comprises nucleotides 76 to 1267 of SEQ ID NO: 2.

[0015] In an embodiment, the OriP origin of replication comprises a nucleic acid sequence that is at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to nucleotides 1795 to 3545 of SEQ ID NO: 1.

[0016] In an embodiment, the SV40 origin of replication comprises a nucleic acid sequence that is at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to nucleotides 5834 to 6140 of SEQ ID NO: 1. In a particular embodiment, the SV40 origin of replication comprises nucleotides 5834 to 6140 of SEQ ID NO: 1.

[0017] Exemplary expression vector of the invention comprise a nucleic acid sequence that is at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to a sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32. In particular embodiments, the expression vector comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32.

[0018] Expression vectors of the invention are also provided in FIGS. 1, 2, and 14-25. Additional vectors of the invention are described in FIGS. 8-13.

[0019] In another aspect, the invention provides a mammalian host cell comprising the vector of the invention. The mammalian host may be a COS cell, such as a COS 7 cell, or a human embryonic kidney (HEK) cell, such as a HEK-293 cell.

[0020] In another aspect, the invention provides a kit comprising a vector of the invention.

[0021] In another aspect, the invention provides a method of producing a recombinant protein comprising introducing an expression vector of the invention into a mammalian host cell, culturing the mammalian host cell under suitable conditions so as to express the protein, and recovering the protein.

[0022] In another aspect, the invention provides an expression vector comprising a nucleic acid sequence encoding a signal peptide. In one embodiment, the gene of interest is operably linked to a nucleic acid encoding a signal peptide.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] The foregoing and other objects, features and advantages of the present invention, as well as the invention itself, will be more fully understood from the following description of preferred embodiments when read together with the accompanying drawings, in which:

[0024] FIG. 1 shows a map of the empty pHyb-C vector. Features include a SV40 eukaryotic origin of replication, a cytomegalovirus eukaryotic expression promoter (pCMV), Tripartite leader sequence (TPL), a splice donor site (SD), an Adenovirus major late enhancer element (enh MLP), a splice acceptor site (SA), an open reading frame (ORF) region for a gene of interest followed by a poly A signal (pA), a dyad symmetry element (DS), an Epstein Barr virus-derived eukaryotic origin of replication (OriP), a repeat region (FR), an ampicillin resistance marker (AmpR) and a bacterial origin of replication (pMB1ori).

[0025] FIG. 2 shows a map of the empty pHyb-E vector. Features include a SV-40 eukaryotic origin of replication, an EF-1 α eukaryotic promoter, an open reading frame (ORF) region for a gene of interest followed by a poly A signal (pA), a dyad symmetry element (DS), an Epstein Barr virus-derived eukaryotic origin of replication (OriP), a repeat region (FR), an ampicillin resistance marker (AmpR) and a bacterial origin of replication (pMB1ori).

[0026] FIG. 3 shows recombinant Fc fusion protein titers produced by COS cells transfected via electroporation with pBOS, pTT3, pHybC and pHybE vectors.

[0027] FIG. 4 shows recombinant Fc fusion protein titers produced by HEK-293-6E cells transfected using PEI with pBOS, pTT3, pHybC and pHybE vectors.

[0028] FIG. 5 shows antibody titers produced by HEK-293-6E transfected using PEI with pBOS, pTT3, pHybC and pHybE vectors constructed to express an IgG antibody.

[0029] FIG. 6 shows antibody titers produced by COS transfection via electroporation with pBOS, pTT3, pHybC and pHybE vectors constructed to express an IgG antibody.

[0030] FIG. 7 shows antibody titers produced by COS transfection via electroporation with pHyb-E-Swa I (v1) or pHyb-E (v2) vector constructs expressing an IgG antibody.

[0031] FIG. 8 shows a map of the pHybC-mBR3-mCg2a vector (also referred to as "pHybC-mBR3-Fc").

[0032] FIG. 9 shows a map of the pHybE-mBR3-mCg2a vector (also referred to as "pHybE-mBR3-Fc").

[0033] FIG. 10 shows a map of the pHybC-E7-hCk vector (also referred to as "pHybC-E7").

[0034] FIG. 11 shows a map of the pHybC-D2-hCg1,z,a vector (also referred to as "pHybC-D2").

[0035] FIG. 12 shows a map of the pHybE-D2-hCg1,z,a vector (also referred to as "pHybE-D2").

[0036] FIG. 13 shows a map of the pHybE-E7-hCk vector (also referred to as "pHybE-E7").

[0037] FIG. 14 shows a map of pHybE-hCg1,z,a V2 (also referred to as "pJP182").

[0038] FIG. 15 shows a map of pHybE-hCg1,z,non-a V2 (also referred to as "pJP183").

[0039] FIG. 16 shows a map of pHybE-hCg1,z,non-a,mu (234,235) V2 (also referred to as "pJP184").

[0040] FIG. 17 shows a map of pHybE-hCg1,z,non-a,mu (234,237) V2 (also referred to as “pJP185”).

[0041] FIG. 18 shows a map of pHybE-hCg2,n+V2 (also referred to as “pJP186”).

[0042] FIG. 19 shows a map of pHybE-hCg2,n- V2 (also referred to as “pJP187”).

[0043] FIG. 20 shows a map of pHybE-hCg4 V2 (also referred to as “pJP188”).

[0044] FIG. 21 shows a map of pHybE-mCg1 V2 (also referred to as “pJP189”).

[0045] FIG. 22 shows a map of pHybE-mCg2a V2 (also referred to as “pJP190”).

[0046] FIG. 23 shows a map of pHybE-hCk V2 (also referred to as “pJP191”).

[0047] FIG. 24 shows a map of pHybE-hCl V2 (also referred to as “pJP192”).

[0048] FIG. 25 shows a map of pHybE-mCk V2 (also referred to as “pJP193”).

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

[0049] In order that the present invention may be more readily understood, certain terms are first defined herein.

[0050] The term “nucleic acid” or “nucleic acid molecule,” as used herein, is intended to include DNA, RNA, mRNA, cDNA, genomic DNA, and analogs thereof. A nucleic acid molecule may be single-stranded or double-stranded, but preferably is double-stranded DNA. A nucleic acid may be isolated, or integrated into another nucleic acid molecule, e.g., an expression vector or the chromosome of an eukaryotic host cell.

[0051] An “isolated” nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term “isolated” includes nucleic acid molecules that are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an “isolated” nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. Moreover, an “isolated” nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

[0052] The terms “recombinant vector” or “vector,” used interchangeably herein, refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid,” which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Alternatively, a vector can be linear. Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. In a preferred embodi-

ment, the vectors of the invention are episomal mammalian vectors. The term “construct,” as used herein, also refers to a vector.

[0053] Certain vectors are capable of directing the expression of genes to which they are operatively linked. An “expression vector” or “recombinant expression vector” is a nucleic acid molecule encoding a gene that is expressed in a host cell, and, furthermore, contains the necessary elements to control expression of the gene. Typically, an expression vector comprises a transcription promoter, a gene of interest, and a transcription terminator. Gene expression is usually placed under the control of a promoter, and such a gene is said to be “operably linked to” the promoter. Similarly, a regulatory element and a core promoter are operably linked if the regulatory element modulates the activity of the core promoter. In one embodiment, the expression vector of the invention comprises more than one origin of replication, thus not limiting the vector to one cell type.

[0054] As used herein, the term “episomally replicating vector” or “episomal vector” refers to a vector that is typically and very preferably not integrated into the genome of the host cell, but exists in parallel. An episomally replicating vector, as used herein, is replicated during the cell cycle and in the course of this replication the vector copies are distributed statistically in the resulting cells depending on the number of the copies present before and after cell division. Preferably, the episomally replicating vector may take place in the nucleus of the host cell, and preferably replicates during S-phase of the cell cycle. Moreover, the episomally replicating vector is replicated at least once, i.e., one or multiple times, in the nucleus of the host cell during S-phase of the cell cycle. In a very preferred embodiment, the episomally replicating vector is replicated once in the nucleus of the host cell during S-phase of the cell cycle.

[0055] As used herein, the terms “origin of replication sequences” or “origin of replication,” used interchangeably herein, refer to sequences which, when present in a vector, initiate replication. An origin of replication may be recognized by a replication initiation factor or, alternatively, by a DNA helicase.

[0056] As used herein, “recombination” refers to a process by which nucleic acid material, e.g., DNA, is exchanged between two nucleic acid molecules, for example, in a microorganism. As used herein, “homologous recombination” refers to a process by which nucleic acid material is exchanged between two nucleic acid molecules through regions or segments of sequence homology, or preferably, sequence identity (e.g., a high degree of sequence identity). In exemplary embodiments, the nucleic acid material is located on a chromosome or an episome of the microorganism. In another exemplary embodiment, the nucleic acid material is located extrachromasomally, for example, on a plasmid. Recombination can occur between linear and/or circular DNA molecules.

[0057] As used herein, the term “gene of interest” refers to an exogenous DNA sequence that is added to the vector of the invention. The gene of interest, for example, may comprise a coding sequence that can be either spaced by introns or that is a cDNA encoding the open reading frame. The “gene of interest” as used herein, refers to the DNA sequence that is added to the vector of the invention for eventual protein expression. The region of the vector to which the gene of interest is cloned is referred to herein as an “insertion site.” Preferably, the gene of interest comprises a portion of the

antibody or fusion protein that is expressed using a vector of the invention. For example, the heavy chain variable region of the antibody adalimumab, i.e., the gene of interest, is cloned into the vector of the invention that comprises a heavy chain constant region.

[0058] In one embodiment of the invention, the vector comprises an antibody light or heavy chain constant region that is 3' to the insertion site for the gene of interest and is operably linked thereto. Thus, in one embodiment, the gene of interest is a variable region of a light or heavy chain of an antibody that is operably linked to the antibody light or heavy chain constant region encoded in the vector of the invention.

[0059] A nucleotide sequence is "operably linked" when placed into a functional relationship with another nucleotide sequence. For example, DNA encoding a signal peptide is operably linked to DNA encoding a protein or polypeptide if, when expressed, the sequences encode the signal peptide in frame with the protein or polypeptide. Likewise, a promoter or enhancer is operably linked to a nucleotide sequence encoding a protein or polypeptide if expression of the protein or polypeptide is promoted or enhanced. In one embodiment, nucleotide sequences that are operably linked are contiguous (e.g., in the case of a signal sequences). Alternatively, nucleotide sequences that are operably linked can be non-contiguous (e.g., in the case of enhancers). In one embodiment, the nucleic acid sequence encoding an antibody light or heavy chain constant region is operably linked to the gene of interest, e.g., a heavy or light chain variable region.

[0060] The term "promoter" includes any nucleic acid sequence sufficient to direct transcription in a eukaryotic cell, including inducible promoters, repressible promoters and constitutive promoters. Typically, a promoter includes elements that are sufficient to render promoter-dependent gene expression controllable in a cell type-specific, tissue-specific or temporal-specific manner, or inducible by external signals or agents. Such elements can be located in the 5' or 3' or intron sequence regions of a particular gene. Ordinarily, gene expression will be constitutive, although regulatable promoters can be employed in the present invention if desired. Gene expression can also be controlled by transcription-regulation using heat, light, or metals, such as by the use of metallothionein genes or heat shock genes.

[0061] "Upstream" and "downstream" are terms used to describe the relative orientation between two elements present in a nucleotide sequence or vector. An element that is "upstream" of another is located in a position closer to the 5' end of the sequence (i.e., closer to the end of the molecule that has a phosphate group attached to the 5' carbon of the ribose or deoxyribose backbone if the molecule is linear) than the other element. An element is said to be "downstream" when it is located in a position closer to the 3' end of the sequence (i.e., the end of the molecule that has a hydroxyl group attached to the 3' carbon of the ribose or deoxyribose backbone in the linear molecule) when compared to the other element.

[0062] As used herein, the term "stuffer sequence" refers to a nucleic acid sequence, preferably in a vector, which is flanked by restriction enzyme sites at both the 5' and 3' ends. The stuffer sequence is located in a vector at the insertion site for the nucleic acid encoding the gene of interest. During the cloning process, the stuffer sequence is digested away from the vector using the appropriate restriction enzymes, and the nucleic acid encoding the gene of interest is ligated or homologously recombined into the vector at the former position of the stuffer sequence. Preferably, the stuffer sequence is

large enough to provide sufficient distance between the 5' and 3' restriction enzyme sites so that the restriction enzyme can efficiently cut the vector. In addition, it is preferred that the length of the stuffer sequence is different than the size of the nucleic acid encoding the gene of interest, e.g., a stuffer sequence of about 300 base pairs or less or about 400 base pairs or more may be used for a nucleic acid encoding the gene of interest that is about 350 base pairs. In another embodiment, the stuffer sequence is about 1 kb in size.

[0063] The term "recombinant host cell" (or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

[0064] The term "antibody" as referred to herein includes whole antibodies and any antigen binding fragment (i.e., "antigen-binding portion") or single chains thereof. An "antibody" refers to a glycoprotein comprising at least two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds, or an antigen binding portion thereof. Each heavy chain is comprised of a heavy chain variable region (abbreviated herein as V_H) and a heavy chain constant region. The heavy chain constant region is comprised of three domains, CH1, CH2 and CH3. Each light chain is comprised of a light chain variable region (abbreviated herein as V_L) and a light chain constant region. The light chain constant region is comprised of one domain, CL. The V_H and V_L regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDR), interspersed with regions that are more conserved, termed framework regions (FR). Each V_H and V_L is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The six CDRs of a V_H and V_L combination form an antigen binding site. In the case of an antibody composed of two H chains and two L chains, the antibody may contain two identical antigen binding sites, two different antigen binding sites that bind the same antigen, or two antigen binding sites that bind different antigens. The constant regions of the antibodies may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (e.g., effector cells) and the first component (C1q) of the classical complement system.

[0065] The term "antigen-binding portion" of an antibody (or simply "antibody portion"), as used herein, refers to one or more fragments of an antibody that retain the ability to specifically bind to an antigen (e.g., IL-1 α , IL-1 β). The antigen-binding function of an antibody can be performed by fragments of a full-length antibody. Examples of binding fragments encompassed within the term "antigen-binding portion" of an antibody include (i) a Fab fragment, a monovalent fragment consisting of the V_L , V_H , CL and CH1 domains; (ii) a F(ab')₂ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the V_H and C_{H1} domains; (iv) a Fv fragment consisting of the V_L and V_H domains of a single arm of an antibody, (v) a dAb fragment (Ward et al, (1989)

Nature 341:544-546), which consists of a V_H or V_L domain; and (vi) an isolated complementarity determining region (CDR). Furthermore, although the two domains of the Fv fragment, VL and V_H , are coded for by separate genes, they can be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent molecules (known as single chain Fv (scFv); see e.g., Bird et al. (1988) Science 242:423-426; and Huston et al. (1988) Proc. Natl. Acad. Sci. USA 85:5879-5883). Such single chain antibodies are also intended to be encompassed within the term "antigen-binding portion" of an antibody. These antibody fragments are obtained using conventional techniques known to those with skill in the art, and the fragments are screened for utility in the same manner as are intact antibodies. In one embodiment if the invention, the antibody fragment is selected from the group consisting of a Fab, an Fd, an Fd', a single chain Fv (scFv), an scFv_α, and a domain antibody (dAb).

[0066] Still further, an antibody or antigen-binding portion thereof may be part of a larger immunoadhesion molecules, formed by covalent or noncovalent association of the antibody or antibody portion with one or more other proteins or peptides. Examples of such immunoadhesion molecules include use of the streptavidin core region to make a tetrameric scFv molecule (Kipriyanov et al. (1995) *Human Antibodies and Hybridomas* 6:93-101) and use of a cysteine residue, a marker peptide and a C-terminal polyhistidine tag to make bivalent and biotinylated scFv molecules (Kipriyanov et al. (1994) *Mol. Immunol.* 31:1047-1058). Antibody portions, such as Fc, Fab and F(ab')₂ fragments, can be prepared from whole antibodies using conventional techniques, such as papain or pepsin digestion, respectively, of whole antibodies. Moreover, antibodies, antibody portions and immunoadhesion molecules can be obtained using standard recombinant DNA techniques.

[0067] The term "domain" refers to a folded protein structure that retains its tertiary structure independently of the rest of the protein. Generally, domains are responsible for discrete functional properties of proteins, and in many cases may be added, removed or transferred to other proteins without loss of function of the remainder of the protein and/or of the domain. By single antibody variable domain is meant a folded polypeptide domain comprising sequences characteristic of antibody variable domains. It therefore includes complete antibody variable domains and modified variable domains, for example, in which one or more loops have been replaced by sequences that are not characteristic of antibody variable domains, or antibody variable domains that have been truncated or comprise N- or C-terminal extensions, as well as folded fragments of variable domains that retain at least in part the binding activity and specificity of the full-length domain.

[0068] Variable domains of the invention may be combined to form a group of domains; for example, complementary domains may be combined, such as VL domains being combined with VH domains. Non-complementary domains may also be combined, e.g., VH domain and a second VH domain. Domains may be combined in a number of ways, involving linkage of the domains by covalent or non-covalent means.

[0069] A "dAb" or "domain antibody" refers to a single antibody variable domain (V_H or V_L) polypeptide that specifically binds antigen. In one embodiment, the vector of the invention is used to express a dAb.

[0070] The phrase "recombinant antibody" refers to antibodies that are prepared, expressed, created or isolated by recombinant means, such as antibodies expressed using a recombinant expression vector transfected into a host cell, antibodies isolated from a recombinant, combinatorial antibody library, antibodies isolated from an animal (e.g., a mouse) that is transgenic for human immunoglobulin genes (see, e.g., Taylor et al. (1992) Nucl. Acids Res. 20:6287-6295) or antibodies prepared, expressed, created or isolated by any other means that involves splicing of particular immunoglobulin gene sequences (such as human immunoglobulin gene sequences) to other DNA sequences. Examples of recombinant antibodies include chimeric, CDR-grafted and humanized antibodies.

[0071] The term "human antibody" refers to antibodies having variable and constant regions corresponding to, or derived from, human germline immunoglobulin sequences as described by, for example, Kabat et al. (See Kabat, et al. (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242). The human antibodies of the invention, however, may include amino acid residues not encoded by human germline immunoglobulin sequences (e.g., mutations introduced by random or site-specific mutagenesis in vitro or by somatic mutation in vivo), for example in the CDRs and in particular CDR3.

[0072] Recombinant human antibodies of the invention have variable regions, and may also include constant regions, derived from human germline immunoglobulin sequences (See Kabat et al. (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242). In certain embodiments, however, such recombinant human antibodies are subjected to in vitro mutagenesis (or, when an animal transgenic for human Ig sequences is used, in vivo somatic mutagenesis) and thus the amino acid sequences of the VH and VL regions of the recombinant antibodies are sequences that, while derived from and related to human germline VH and VL sequences, may not naturally exist within the human antibody germline repertoire in vivo. In certain embodiments, however, such recombinant antibodies are the result of selective mutagenesis or backmutation or both.

[0073] The term "backmutation" refers to a process in which some or all of the somatically mutated amino acids of a human antibody are replaced with the corresponding germline residues from a homologous germline antibody sequence. The heavy and light chain sequences of a human antibody of the invention are aligned separately with the germline sequences in the VBASE database to identify the sequences with the highest homology. Differences in the human antibody of the invention are returned to the germline sequence by mutating defined nucleotide positions encoding such different amino acid. The role of each amino acid thus identified as candidate for backmutation should be investigated for a direct or indirect role in antigen binding and any amino acid found after mutation to affect any desirable characteristic of the human antibody should not be included in the final human antibody. To minimize the number of amino acids subject to backmutation those amino acid positions found to be different from the closest germline sequence but identical to the corresponding amino acid in a second germline sequence can remain, provided that the second germline sequence is identical and colinear to the sequence of the human antibody of the invention for at least 10, preferably 12

amino acids, on both sides of the amino acid in question. Backmutation may occur at any stage of antibody optimization.

[0074] The term “chimeric antibody” refers to antibodies that comprise heavy and light chain variable region sequences from one species and constant region sequences from another species, such as antibodies having murine heavy and light chain variable regions linked to human constant regions.

[0075] The term “CDR-grafted antibody” refers to antibodies that comprise heavy and light chain variable region sequences from one species but in which the sequences of one or more of the CDR regions of VH and/or VL are replaced with CDR sequences of another species, such as antibodies having murine heavy and light chain variable regions in which one or more of the murine CDRs (e.g., CDR3) has been replaced with human CDR sequences.

[0076] The term “humanized antibody” refers to antibodies that comprise heavy and light chain variable region sequences from a non-human species (e.g., a mouse) but in which at least a portion of the VH and/or VL sequence has been altered to be more “human-like”, i.e., more similar to human germline variable sequences. One type of humanized antibody is a CDR-grafted antibody, in which human CDR sequences are introduced into non-human VH and VL sequences to replace the corresponding nonhuman CDR sequences.

[0077] As used herein, the terms “linked,” “fused” or “fusion” are used interchangeably. These terms refer to the joining together of two more elements or components, by whatever means including chemical conjugation or recombinant means. An “in-frame fusion” or “operably linked” refers to the joining of two or more open reading frames (ORFs) to form a continuous longer ORF, in a manner that maintains the correct reading frame of the original ORFs. Thus, the resulting recombinant fusion protein is a single protein containing two or more segments that correspond to polypeptides encoded by the original ORFs (which segments are not normally so joined in nature). Although the reading frame is thus made continuous throughout the fused segments, the segments may be physically or spatially separated by, for example, an in-frame linker sequence.

[0078] As used herein, the term “Fc region” includes amino acid sequences derived from the constant region of an antibody heavy chain. In some embodiments, an Fc region includes a polypeptide comprising the constant region of an antibody excluding the first constant region immunoglobulin domain.

[0079] An Fc region may be a functionally equivalent analog of an Fc region. A functionally equivalent analog of an Fc region may be a variant Fc region, comprising one or more amino acid modifications to a wild-type or naturally existing Fc region. In some embodiments, variant Fc regions possess at least 50% homology with a naturally existing Fc region, with about 80% to 99% being preferred, including at least about 85% homology, at least about 90% homology, at least about 95% homology, at least about 96% homology, at least about 97% homology, at least 98% homology, or at least about 99% homology. Functionally equivalent analogs of an Fc region may comprise one or more amino acid residues added to or deleted from the N- or C-termini of the protein, preferably no more than 30, most preferably no more than 10. Functionally equivalent analogs of an Fc region include Fc regions operably linked to a fusion partner.

[0080] The terms “Fc fusion” or “Fc fusion protein”, as used herein, include a protein wherein one or more proteins,

polypeptides or small molecules is operably linked to an Fc region or derivative thereof. The term “Fc fusion” as used herein is intended to be synonymous with terms such as “Ig fusion”, “Ig chimera”, and “receptor globulin” (sometimes with dashes) as used in the prior art (Chamow et al., 1996, *Trends Biotechnol* 14:52-60; Ashkenazi et al., 1997, *Curr Opin Immunol* 9:195-200). An Fc fusion combines one or more Fc regions, or variant(s) thereof, of an immunoglobulin with a fusion partner, which in general can be any protein, polypeptide, peptide, or small molecule. In some embodiments, the role of the non-Fc part of an Fc fusion, i.e., the fusion partner, may be to mediate target binding, and thus it can be functionally analogous to the variable regions of an antibody.

[0081] A variety of linkers may be used in the present invention to covalently link Fc polypeptides to a fusion or conjugate partner or to generate an Fc fusion. As used herein, the terms “linker”, “linker sequence”, “spacer”, “tethering sequence” or equivalents thereof refer to a molecule or group of molecules (such as a monomer or polymer) that connects two molecules and can serve to place the two molecules in a preferred configuration. A number of strategies may be used to covalently link molecules together. These include, but are not limited to, polypeptide linkages between N- and C-termini of proteins or protein domains, linkage via disulfide bonds, and linkage via chemical cross-linking reagents.

II. Vectors of the Invention

[0082] The invention provides episomal vectors for expressing proteins in mammalian host cells. The vectors of the invention are based on the inclusion of two episomal origins of replication that allow the vector to be used in any cell line containing trans-acting replication initiation factors to either of the origins of replication. While the vector may also contain the replication initiation factor that binds the origin of replication, in a preferred embodiment the trans-acting replication factor is provided by the host cell. In addition, in one embodiment, the vectors of the invention provide efficient and effective means for production of antibodies and Fc fusion proteins, as the vectors contain heavy or light chain constant regions operably linked to a gene of interest. Examples of vectors of the invention are described in FIGS. 1, 2, and 8 to 25. In addition, sequences of exemplary vectors are provided in SEQ ID NOs: 1 to 32. FIGS. 1 and 2 (and corresponding SEQ ID NOs: 1 and 2) describe the “open” vector, i.e., the vector of the invention that does not contain antibody heavy or light chain constant regions and a gene of interest. FIGS. 8-25 provides maps of vectors of the invention which also comprise various murine or human constant regions, with sites for cloning a gene of interest.

[0083] The vector of the invention comprises at least two distinct origins of replication, e.g., OriP origin of replication derived from Epstein-Barr virus (EBV) and an SV40 origin of replication. The origin of replication may be derived from a DNA virus, more preferably from a DNA virus that allows for episomal replication, including origins of replication derived from, for example, Epstein-Barr virus, Herpes simplex virus, Herpesvirus Saimiri, Murine Gammaherpesvirus 68, Human Cytomegalovirus, Mouse Cytomegalovirus, Pseudorabiesvirus, Simian Virus 40, Polyoma virus, human BK virus, Bovine Papilloma virus, and Adeno-associated virus.

[0084] In one embodiment, the origin of replication is from Epstein-Barr virus, e.g., oriP, or functional parts thereof (examples of Epstein-Barr functional origins are described in

Aiyar et al. (1998) *EMBO Journal*, 17:6394). The Epstein-Barr virus origin of replication (OriP) is composed of 2 main elements and multiple cis-acting elements that facilitate DNA synthesis by the cell and a viral maintenance element. The first of the two main elements contains a family of repeats (FR), which comprise the EBNA binding sites (shown in FIGS. 1 and 2). EBNA is the replication initiation factor that initiates replication of the vector via OriP (see Genbank accession number V01555 (gi:94734074) for EBNA sequence). The second element contained in OriP contains a so called dyad symmetry (DS) and its function is to serve as an origin recognition element. Generally, the DS and FR elements are spaced by several base pairs, typically 1000 bp. The relative orientation of OriP, and in particular of DS and FR, can be altered without affecting OriP function. The orientation of OriP, and in particular of DS and FR, relative to the other elements positioned on the expression vectors of the invention, can be altered without affecting OriP function. In a preferred embodiment of the invention, wherein the origin of replication is an Epstein-Barr virus origin of replication (OriP), and wherein the OriP comprises a family of repeats (FR) and a dyad symmetry (DS), the consecutive order is such that the DS element is between the gene of interest and the FR element. In one embodiment, the vector of the invention comprises an OriP (Epstein-Barr virus) origin of replication comprising nucleotides 1795 to 3545 of SEQ ID NO: 1, or sequences 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0085] In another embodiment, the vector comprises an SV40 origin of replication. The SV40 (Simian Virus 40) origin of replication (described, for example, in FIGS. 1 and 2 as "SV40 Ori") requires a single viral protein, the large T-antigen, for initiation of replication of the vector via this origin. The SV40 origin of replication may be used in episomal vectors to replicate and maintain said vector (see Calos (1996) *Trends Genetics* 12: 462; Harrison et al. (1994) *J Virol* 68:1913; Cooper et al. (1997) *PNAS* 94:6450; and Ascenzi et al. (1997) *Cancer Lett* 118:135). In one embodiment, the vector of the invention comprises an SV40 origin of replication comprising nucleotides 5834 to 6140 of SEQ ID NO: 1, or sequences 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0086] Functional variants of origins of replication describe herein are also encompassed in the meaning of origin of replication according to the present application.

[0087] In addition to the episomal origins of replication, the vector of the invention may also have an origin of replication for replicating the vector in bacteria. An example, as shown in FIGS. 1 and 2 and not meant to be limiting, is the pMB1 ori, which functions in *E. coli*.

[0088] The vector of the invention may also include a selectable marker. The selection marker may facilitate the cloning and amplification of the vector sequences in prokaryotic and eukaryotic organisms. In certain embodiments, the selection marker will confer resistance to a compound or class of compounds, such as an antibiotic. An exemplary selection marker that can be used with the nucleic acid molecules and expression systems of the present invention is one that confers resistance to puromycin. Alternatively, selection markers may be used that confer resistance to hygromycin, gpt, neomycin, zeocin, ouabain, blasticidin, kanamycin, geneticin, gentamicin, ampicillin, tetracycline, streptomycin, spectino-

mycin, nalidixic acid, rifampicin, chloramphenicol, zeocin or bleomycin, or markers such as DHFR, hisD, trpB, or glutamine synthetase.

[0089] Also included in the vector of the invention are regulatory elements that are necessary for transcription and translation of the gene of interest (as well as the selectable marker), into proteins. The transcriptional regulatory elements normally comprise a promoter 5' of the gene sequence to be expressed, transcriptional initiation and termination sites, and polyadenylation signal sequence. The term "transcriptional initiation site" refers to the nucleic acid in the construct corresponding to the first nucleic acid incorporated into the primary transcript, i.e., the mRNA precursor; the transcriptional initiation site may overlap with the promoter sequences. The term "transcriptional termination site" refers to a nucleotide sequence normally represented at the 3' end of a gene of interest or the stretch of sequences to be transcribed, that causes RNA polymerase to terminate transcription. The polyadenylation signal sequence, or poly-A addition signal provides the signal for the cleavage at a specific site at the 3' end of eukaryotic mRNA and the post-transcriptional addition in the nucleus of a sequence of about 100-200 adenine nucleotides (polyA tail) to the cleaved 3' end. The polyadenylation signal sequence includes the sequence AATAAA located at about 10-30 nucleotides upstream from the site of cleavage, plus a downstream sequence.

[0090] A regulatory element that may be included in the vector of the invention is a promoter. The promoter can be constitutive or inducible. An enhancer (i.e., a cis-acting DNA element that acts on a promoter to increase transcription) may be necessary to function in conjunction with the promoter to increase the level of expression obtained with a promoter alone, and may be included as a transcriptional regulatory element. Often, the polynucleotide segment containing the promoter will include the enhancer sequences as well (e.g., CMV IE P/E; SV40 P/E; MPSV P/E). Splice signals may be included where necessary to obtain spliced transcripts. To produce a secreted polypeptide, the selected sequence will generally include a signal sequence encoding a leader peptide that directs the newly synthesized polypeptide to and through the ER membrane where the polypeptide can be routed for secretion. The leader peptide is often but not universally at the amino terminus of a secreted protein and is cleaved off by signal peptidases after the protein crosses the ER membrane. The selected sequence will generally, but not necessarily, include its own signal sequence. Where the native signal sequence is absent, a heterologous signal sequence can be fused to the selected sequence. Numerous signal sequences are known in the art and available from sequence databases such as GenBank and EMBL. Translational regulatory elements include a translational initiation site (AUG), stop codon and poly A signal for each individual polypeptide to be expressed. An internal ribosome entry site (IRES) is included in some constructs.

[0091] Promoters for use in the present invention include viral, mammalian and yeast promoters, e.g., murine beta globin promoter, ubiquitin promoter, polyoma promoter, mammalian cytomegalovirus (CMV) promoter, yeast alcohol oxidase, phosphoglycerokinase promoter, lactose inducible promoters, galactosidase promoter, adeno-associated viral promoter, poxvirus promoter, retroviral promoters, rous sarcoma virus promoter, adenovirus promoters, SV40 promoter, hydroxymethylglutaryl coenzyme A promoter, thymidine kinase promoter, H5R poxvirus promoters, adenovirus type 2

MPC late promoter, alpha-antitrypsin promoter, fox IX promoter, immunoglobulin promoter, CFTR surfactant promoter, albumin promoter and transferrin promoter. A promoter selected for use with nucleic acids and expression vectors of the invention can provide for (1) high levels of expression, e.g., in driving expression of the gene of interest, or (2) decreased levels of expression (after weakening by modification), e.g., in driving expression of the selectable marker gene. Preferably, the promoter driving the gene of interest is a strong promoter, e.g., ubiquitin, CMV, EF-1 α and SR alpha promoters, to increase expression and promote correct splicing of the product of interest.

[0092] In one embodiment, the vector of the invention includes a CMV promoter to drive expression of the gene of interest. Use of the CMV promoter is described in U.S. Pat. Nos. 5,385,839 and 5,849,522, incorporated by reference herein. In one embodiment, the CMV promoter used in the vector of the invention is operably linked to the gene of interest and nucleotides 1 to 608 of SEQ ID NO: 1. Also included in the scope of the invention are CMV promoter sequences that are 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to nucleotides 1 to 608 of SEQ ID NO: 1.

[0093] Another promoter that may be used in the vector of the invention is a promoter from elongation-factor-1 α (EF-1 α), e.g., human EF-1 α . The sequence for the human EF-1 α promoter can be found at GenBank Accession No. NM_001402 (gi:83367078). In one embodiment, the vector of the invention comprises nucleotides 76 to 1267 of SEQ ID NO: 2. Also included in the scope of the invention are EF-1 α promoter sequences that are 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to nucleotides 1 to 608 of SEQ ID NO: 1.

[0094] In one embodiment, the vector comprises a SmaI restriction site for cloning purposes.

[0095] Typically, genes (e.g., selectable markers and GOIs) are sandwiched between a promoter and a polyadenylation site. The poly A sequence used can be from the gene of interest (i.e., the native poly A sequence can be used) or a heterologous poly A sequence can be used (i.e., from a gene different from the GOI), e.g., BGH polyA and SV40 polyA. An mRNA is transcribed from the promoters and stabilized by the polyadenylation signals located 3' to the coding regions. Poly A signals are well-known in the art, and can be selected based on suitability for use with the vectors and host cells employed in the present invention. Examples of poly A signals that can be used include human BGH poly A, SV40 poly A, human beta actin polyA, rabbit beta globin polyA, and immunoglobulin kappa polyA.

[0096] The vector of the invention includes a gene of interest, which the vector as a means for expressing in cell culture. The gene of interest may encode a functional nucleic acid molecule (e.g., an RNA, such as an antisense RNA molecule) or, more typically, encodes a peptide, polypeptide or protein for which increased production is desired. Vectors of the invention can have a gene of interest, inserted at an insertion site such that the gene of interest is operably linked to a regulatory nucleic acid sequence that allows expression of the gene of interest. In one embodiment, the vectors of the invention can be used to express essentially any gene of interest, particularly genes encoding recombinant proteins having therapeutically useful activity or other commercially relevant applications.

[0097] Non-limiting examples of genes of interest include hormones, chemokines, cytokines, lymphokines, antibodies,

receptors, adhesion molecules, and enzymes. A non-exhaustive list of desired products includes, e.g., human growth hormone, bovine growth hormone, parathyroid hormone, thyroid stimulating hormone, follicle stimulating hormone growth, luteinizing hormone; hormone releasing factor; lipoproteins; alpha-1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; calcitonin; glucagon; molecules such as renin; clotting factors such as factor VIIIC, factor IX, tissue factor, and von Willebrands factor; anti-clotting factors such as Protein C, atrial natriuretic factor, lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; RANTES (regulated on activation normally T-cell expressed and secreted); human macrophage inflammatory protein (MIP-1-alpha); a serum albumin such as human serum albumin; mullerian-inhibiting substance; relaxin A- or B-chain; prorelaxin; mouse gonadotropin-associated peptide; DNase; inhibin; activin; receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as bone-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), growth factors including vascular endothelial growth factor (VEGF), nerve growth factor such as NGF-.beta.; platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF, bFGF, FGF-4, FGF-5, FGF-6; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta, including TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, or TGF- β 5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-1), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), e.g., M-CSF, GM-CSF, and G-CSF; interleukins (ILs), e.g., IL-1 to IL-33; superoxide dismutase; T-cell receptors; surface membrane proteins, e.g., HER2; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; receptors for growth factors, cytokines, chemokines, and lymphokines; regulatory proteins; antibodies; chimeric proteins such as immunoadhesins and fragments of any of the above-listed polypeptides. Examples of bacterial polypeptides or proteins include, e.g., alkaline phosphatase and .beta.-lactamase.

[0098] In one aspect of the invention, the vector comprises an antibody heavy or light chain region that is operably linked to the insertion site. Examples of vectors comprising two episomal origins of replication and a light or heavy chain constant region of an antibody, can be found in SEQ ID NOs: 3-32.

[0099] One embodiment of the invention includes vectors that can be used to express a complete antibody, i.e., a variable region linked to the constant region for either the heavy or light chain. Thus, the gene of interest may encode an antibody heavy chain or light chain variable region, which can be of any antibody type, e.g., murine, chimeric, humanized, and human. A gene of interest encoding a heavy chain or light chain variable region may include the full length variable region, or alternatively, may encode only a fragment of the heavy chain or light chain, e.g., the antigen binding portion region. In one embodiment, the gene of interest encodes a murine or human antibody variable region. In such an

instance, the constant region may be matched to the species of the variable region (SEQ ID NOs: 3-8, 27 and 28 encode murine constant regions, while SEQ ID NOs: 9-26 and 29-32 encode human constant regions).

[0100] In one embodiment, the vector of the invention includes a nucleic acid sequence encoding an antibody heavy chain constant region having certain isotype and/or allotype characteristics. The heavy chain constant region may, for example, be a gamma isotype (IgG), such as gamma 1, gamma 2, gamma 3, or gamma 4. In one embodiment, the heavy chain gamma 1 constant region is a certain allotype, including, but not limited to, allotypes z, a and z, non-a. The z, a, allotype is also known as G1m17 and G1m1 allotypes, and corresponds to IGHG1 with Lys at position 214 (within CH1), Asp at 356 (CH3), and Leu at 358 (CH3) (numbering according to the EU number system). The z, non-a allotype, also known as G1m17, and nG1m1 allotypes, corresponds to IGHG1 with Lys at position 214 (within CH1), Glu at 356 (CH3), and Met at 358 (CH3) (numbering according to the EU number system).

[0101] In another embodiment, the heavy chain gamma 2 constant region (hcG2) is a certain allotype, including, but not limited to, n- or n+. The n+ allotype of hcG2, also known as G2m (n) or G2m (23), corresponds to IGHG2 with Thr at position 189 in CH1 and Met at position 282 (numbering according to the EU number system). The n- allotype of hcG2, also known as G2m (n-), corresponds to IGHG2 with Pro at position 189 in CH1 and Val at position 282 (numbering according to the EU number system). Additional details of the n+ and n- allotypes are described in Hougs et al. (2001) *Immunogenetics* 52:242 and Brusco et al. (1995) *Immunogenetics* 42:414.

[0102] In other embodiments, the heavy chain constant region may be an IgM, IgA (IgA1 or IgA2), IgD, or IgE isotype.

[0103] In one embodiment, the heavy chain constant region may have the following human isotype and allotype characteristics: gamma 1, z, a; gamma 1, z, non-a; gamma 2, n+; gamma 2, n-; or gamma 4. In one embodiment, the isotype/allotype gamma 1, z, non-a may include a mutation at position 234 of the heavy chain constant region. In a further embodiment, the isotype/allotype gamma 1, z, non-a may include mutations at position 234 and 235 or 234 and 237 of the heavy chain constant region. Examples of such vectors are provided in FIGS. 8 to 25.

[0104] In another example, the light chain constant region encoded in the vector of the invention may comprise a kappa isotype or lambda isotype.

[0105] The constant regions encoded by the vector of the invention are not limited to human, but may instead include murine or other species of constant regions. In one embodiment, the expression vector of the invention comprises a nucleic acid encoding a heavy chain constant region that is either a murine gamma 1 isotype or a murine gamma 2a isotype, or a light chain constant region that is a murine kappa isotype.

[0106] Two vectors of the invention, pHybC and pHybE, are empty vectors in that these vectors do not contain constant regions, and may be used for cloning genes of interest. Descriptions of pHybC and pHybE are provided below, and maps of these vectors can be found in FIGS. 1 and 2.

pHybC

[0107] The pHybC vector (empty) contains two viral origins of replication, such that the vector may be replicated in

different cell lines. pHybC contains the following elements: an SV40 origin of replication ("SV40 Ori"), which allows for vector plasmid replication in cells expressing the large T antigen protein of SV40 (e.g., a COS7 cell); a CMV promoter ("pCMV") operably linked to the insertion site for a gene of interest; a Tripartite leader sequence (TPL); a splice donor site (SD); an Adenovirus major late enhancer element (enh MLP); a splice acceptor site (SA); an open reading frame (ORF) region for a gene of interest followed by a poly A signal (pA); a dyad symmetry element (DS); an Epstein Barr virus-derived eukaryotic origin of replication (OriP), which permits replication of the vector plasmid in cells expressing the viral EBNA-1 protein (e.g., HEK-293-6E cells); a repeat region (FR); an ampicillin resistance marker (AmpR); and a bacterial origin of replication (pMB1ori). The pHybC vector utilizes the pCMV promoter, one of the strongest promoter elements available. A vector map of pHybC (empty) is described in FIG. 1. The nucleic acid sequence of the pHybC vector is set forth in SEQ ID NO:1.

pHybE

[0108] The pHybE vector (empty) contains two origins of replication, such that the vector may be replicated in different cell lines. pHybE contains the following elements: an SV40 origin of replication ("SV40 Ori"), which allows for vector plasmid replication in cells expressing the large T antigen protein of SV40 (e.g., a COS7 cell); an EF-1a eukaryotic promoter operably linked to the insertion site for a gene of interest; an open reading frame (ORF) region for a gene of interest followed by a poly A signal (pA); a dyad symmetry element (DS); an Epstein Barr virus-derived eukaryotic origin of replication (OriP); a repeat region (FR); an ampicillin resistance marker (AmpR); and a bacterial origin of replication (pMB1ori). A vector map of pHybE (empty) is described in FIG. 2. pHybE is distinguished from pHybC in that it pHybE contains an EF-1a promoter operably linked to the insertion site for the gene of interest, while pHybC contains a CMV promoter. The nucleic acid sequence of the pHybE vector is set forth in SEQ ID NO:2.

[0109] The below-mentioned vectors are based on either pHybE or pHybC, and additionally contain immunoglobulin heavy or light chain constant regions. As with pHybE and pHybC, the following vectors have cloning sites that may be used for the insertion of a gene of interest, e.g., a coding sequence of an immunoglobulin variable region, or an antigen binding portion thereof. In each instance, the cloning site for the gene of interest is adjacent to the coding sequence of a constant region contained within the vector. Thus, the vectors below may be used to express antibody light or heavy chains containing a particular constant region and a particular variable region. As with pHybC and pHybE, each of the below-mentioned vectors of the invention contain multiple origins of replication, such that the antibody light or heavy chain may be expressed in different cell lines using the same vector. Descriptions of additional vectors of the invention are described below (see also vector maps provided in FIGS. 8 to 25). It should be noted that pHyb vectors described as version 1 (V1) have an additional Sma I site upstream of the Srf I restriction site, whereas pHyb vectors described as version 2 (V2) do not have the additional Sma I site.

Vectors of the Invention Comprising Murine Constant Regions pHybC-mCg2a Vector pHybC-mCg2a is based on the pHybC vector (thus contains all of the elements described above for pHybC). This vector also comprises the murine immunoglobulin coding sequence for the gamma 2a heavy

chain constant region. Thus, in one embodiment, the pHybC-mCg2a vector may be used to express an antibody heavy chain comprising an immunoglobulin heavy chain variable region (or portion thereof) and a murine gamma 2 heavy chain constant region. Alternatively, pHybC-mCg2 may be used to express a gene of interest fused to a gamma 2 heavy chain constant region, e.g. an Fc fusion protein. FIG. 8 shows a map of the pHybC-mBR3-mCg2a which comprises the coding sequence for the extracellular domain (ECD) of the murine BR3 protein as the gene of interest. The nucleic acid sequence of pHybC-mBR3-mCg2a is set forth in SEQ ID NO:27.

pHybE-mCk

[0110] Vector pHybE-mCk is based on the pHybE vector (thus contains all of the elements described above for pHybE). pHybE-mCk also comprises the murine immunoglobulin coding sequence for the kappa light chain constant region. Thus, in one embodiment, the pHybE-mCk vector may be used to express an antibody light chain comprising an immunoglobulin light chain variable region and a murine kappa light chain constant region. Alternatively, pHybE-mCk may be used to express a gene of interest fused to a murine kappa light chain constant region. A vector map of pHybE-mCk V2 is provided in FIG. 25. The nucleic acid sequence of pHybE-mCk V1 is set forth in SEQ ID NO:3 and the nucleic acid sequence of pHybE-mCk V2 is set forth in SEQ ID NO:4.

pHybE-mCg1

[0111] pHybE-mCg1 is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the murine immunoglobulin coding sequence for the gamma 1 heavy chain constant region. Thus, in one embodiment, the pHybE-mCg1 vector may be used to express an antibody heavy chain comprising an immunoglobulin heavy chain variable region and a murine gamma 1 heavy chain constant region. Alternatively, pHybE-mCg1 may be used to express a gene of interest fused to a murine gamma 1 heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-mCg1 V2 is provided in FIG. 21. The nucleic acid sequence of pHybE-mCg1 V1 is set forth in SEQ ID NO:5 and the nucleic acid sequence of pHybE-mCg1 V2 is set forth in SEQ ID NO:6.

pHybE-mCg2a

[0112] pHybE-mCg2a is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the murine immunoglobulin coding sequence for the gamma 2a heavy chain constant region. Thus, in one embodiment, the pHybE-mCg2a vector may be used to express an antibody heavy chain comprising an immunoglobulin heavy chain variable region and a murine gamma 2 heavy chain constant region. Alternatively, pHybE-mCg2a may be used to express a gene of interest fused to a gamma 2 heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-mCg2a V2 is provided in FIG. 22. The nucleic acid sequence of pHybE-mCg2a V1 is set forth as SEQ ID NO:7 and the nucleic acid sequence of pHybE-mCg2a V2 is set forth in SEQ ID NO:8. As an example of one embodiment of how the pHybE-mCg2a may be used, FIG. 9 shows a map of pHybE-mBR3-mCg2a. The vector described in FIG. 9 contains the coding sequence for the extracellular domain (ECD) of the murine BR3 protein. The nucleic acid sequence of pHybE-mBR3-mCg2a is set forth in SEQ ID NO:28.

Vectors of the Invention Comprising Human Constant Regions

[0113] pHybC-E7-hCk

[0114] pHybC-E7-hCk is based on the pHybC vector (thus contains all of the elements described above for pHybC). This vector also comprises the human immunoglobulin coding sequence for the kappa light chain constant region. In addition, pHybC-E7-hCk contains the coding sequence of the light chain variable region of adalimumab (also referred to as "E7"). A vector map of pHybC-E7-hCk is provided in FIG. 10, and the nucleic acid sequence of pHybC-E7-hCk is set forth in SEQ ID NO:29.

pHybC-D2-hCg1,z,a

[0115] pHybC-D2-hCg1,z,a is based on the pHybC vector (thus contains all of the elements described above for pHybC). This vector also comprises the coding sequence for the gamma 1,z,a heavy chain constant region. In addition, pHybC-D2-hCg1,z,a contains the coding sequence of the heavy chain variable region of adalimumab (also referred to as "D2"). A vector map of pHybC-D2-hCg1,z,a is provided in FIG. 11. The nucleic acid sequence of pHybC-D2-hCg1,z,a is set forth in SEQ ID NO:30.

pHybE-hCk

[0116] pHybE-hCk is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the human immunoglobulin coding sequence for the kappa light chain constant region. Thus, for example, the pHybE-hCk vector may be used to express an antibody light chain comprising an immunoglobulin variable light chain region and a human kappa light chain constant region. Alternatively, pHybE-hCk may be used to express a gene of interest fused to a kappa light chain constant region. A vector map of pHybE-hCk V2 is provided in FIG. 23. The nucleic acid sequence of pHybE-hCk V1 is set forth in SEQ ID NO:9 and the nucleic acid sequence of pHybE-hCk V2 is set forth in SEQ ID NO:10. A vector map of pHybE-E7-hCk is also provided in FIG. 13. In addition to all of the elements of the pHybE-hCk vector described above, pHybE-E7-hCk contains the coding sequence of the light chain variable region of adalimumab (also referred to as "E7"). The nucleic acid sequence of pHybE-E7-hCk is set forth in SEQ ID NO:32.

pHybE-hCl

[0117] pHybE-hCl is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the human immunoglobulin coding sequence for the lambda light chain constant region. Thus, in one embodiment, the pHybE-hCl vector may be used to express an antibody light chain comprising an immunoglobulin variable light chain region and a human lambda light chain constant region. Alternatively, pHybE-hCl may be used to express a gene of interest fused to a lambda light chain constant region. A vector map of pHybE-hCl V2 is provided in FIG. 24. The nucleic acid sequence of pHybE-hCl V1 is set forth in SEQ ID NO:11 and the nucleic acid sequence of pHybE-hCl V2 is set forth in SEQ ID NO:12.

pHybE-hCg1,z,a

[0118] pHybE-hCg1,z,a is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises human immunoglobulin coding sequence for the gamma 1,z,a heavy chain constant region. Thus, in one embodiment, the pHybE-hCg1,z,a vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human

gamma 1,z,a heavy chain constant region. Alternatively, pHybE-hCg1,z,a may be used to express a gene of interest fused to a gamma 1,z,a heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg1,z,a is provided in FIG. 14. The nucleic acid sequence of pHybE-hCg1,z,a V1 is set forth in SEQ ID NO:13 and the nucleic acid sequence of pHybE-hCg1,z,a V2 is set forth in SEQ ID NO:14. A vector map for pHybE-D2-hCg1,z,a is provided in FIG. 12. In addition to the elements of pHybE-hCg1,z,a described above, pHybE-D2-hCg1,z,a contains the coding sequence of the heavy chain variable region of adalimumab (also referred to as "D2"). The nucleic acid sequence of pHybE-D2-hCg1,z,a is set forth in SEQ ID NO:31.

pHybE-hCg1,z,non-a

[0119] pHybE-hCg1,z,non-a is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises human immunoglobulin coding sequence for the gamma 1,z,non-a heavy chain constant region. Thus, in one embodiment, the pHybE-hCg1,z,non-a vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma 1,z,non-a heavy chain constant region. Alternatively, pHybE-hCg1,z,non-a may be used to express a gene of interest fused to a gamma 1,z,non-a heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg1,z,non-a V2 is provided in FIG. 15. The nucleic acid sequence of pHybE-hCg1,z,non-a V1 is set forth in SEQ ID NO:15 and the nucleic acid sequence of pHybE-hCg1,z,non-a V2 is set forth in SEQ ID NO:16.

pHybE-hCg1,z,non-a,mu(234,235)

[0120] pHybE-hCg1,z,non-a,mu(234,235) is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises human immunoglobulin coding sequence for the gamma 1,z,non-a,mu(234,235) heavy chain constant region. Thus, in one embodiment, the pHybE-hCg1,z,non-a,mu(234,235) vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma 1,z,non-a,mu(234,235) heavy chain constant region. Alternatively, pHybE-hCg1,z,non-a,mu(234,235) may be used to express a gene of interest fused to a gamma 1,z,non-a,mu(234,235) heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg1,z,non-a,mu(234,235) V2 is provided in FIG. 16. The nucleic acid sequence of pHybE-hCg1,z,non-a,mu(234,235) V1 is set forth in SEQ ID NO:17 and the nucleic acid sequence of pHybE-hCg1,z,non-a,mu(234,235) V2 is set forth in SEQ ID NO:18.

pHybE-hCg1,z,non-a,mu(234,237)

[0121] pHybE-hCg1,z,non-a,mu(234,237) is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises human immunoglobulin coding sequence for the gamma 1,z,non-a,mu(234,237) heavy chain constant region. Thus, in one embodiment, the pHybE-hCg1,z,non-a,mu(234,237) vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma 1,z,non-a,mu(234,237) heavy chain constant region. Alternatively, pHybE-hCg1,z,non-a,mu(234,237) may be used to express a gene of interest fused to a gamma 1,z,non-a,mu(234,237) heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg1,z,non-a,mu(234,237) V2 is provided in FIG. 17. The nucleic acid sequence of pHybE-hCg1,z,non-a,mu(234,237) V1 is set forth in SEQ ID

NO:19 and the nucleic acid sequence of pHybE-hCg1,z,non-a,mu(234,237) V2 is set forth in SEQ ID NO:20.

pHybE-hCg2,n-

[0122] pHybE-hCg2,n- is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the human immunoglobulin coding sequence for the gamma 2,n- heavy chain constant region. Thus, in one embodiment, the pHybE-hCg2,n- vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma 2,n- heavy chain constant region. Alternatively, pHybE-hCg2,n- may be used to express a gene of interest fused to a gamma 2,n- heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg2,n- V2 is provided in FIG. 19. The nucleic acid sequence of pHybE-hCg2,n- V1 is set forth in SEQ ID NO:21 and the nucleic acid sequence of pHybE-hCg2,n- V2 is set forth in SEQ ID NO:22.

pHybE-hCg2,n+

[0123] pHybE-hCg2,n+ is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the human immunoglobulin coding sequence for the gamma 2,n+ heavy chain constant region. Thus, in one embodiment, the pHybE-hCg2,n+ vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma 2,n+ heavy chain constant region. Alternatively, pHybE-hCg2,n+ may be used to express a gene of interest fused to a gamma 2,n+ heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg2,n+ is provided in FIG. 18. The nucleic acid sequence of pHybE-hCg2,n+ V1 is set forth in SEQ ID NO:23 and the nucleic acid sequence of pHybE-hCg2,n+ V2 is set forth in SEQ ID NO:24.

pHybE-hCg4

[0124] pHybE-hCg4 is based on the pHybE vector (thus contains all of the elements described above for pHybE). This vector also comprises the human immunoglobulin coding sequence for the gamma4 heavy chain constant region. Thus, in one embodiment, the pHybE-hCg4 vector may be used to express an antibody heavy chain comprising an immunoglobulin variable heavy chain region and a human gamma4 heavy chain constant region. Alternatively, pHybE-hCg4 may be used to express a gene of interest fused to a gamma4 heavy chain constant region, e.g. an Fc fusion protein. A vector map of pHybE-hCg4 is provided in FIG. 20. The nucleic acid sequence of pHybE-hCg4 V1 is set forth in SEQ ID NO:25 and the nucleic acid sequence of pHybE-hCg4 V2 is set forth in SEQ ID NO:26.

[0125] Sequences of the vectors of the invention are provided in SEQ ID NOs: 1-32. In one embodiment, the vector of the invention comprises a sequence set forth in any one of SEQ ID NOs: 1-32 or sequences that are at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical thereto.

[0126] The invention can be used in the production of human and/or humanized antibodies that immunospecifically recognize specific cellular targets, e.g., any of the aforementioned proteins, the human EGF receptor, the her-2/neu antigen, the CEA antigen, Prostate Specific Membrane Antigen (PSMA), CDS, CD11a, CD18, NGF, CD20, CD45, CD52, Ep-cam, other cancer cell surface molecules, TNF-alpha, TGF-b1, VEGF, other cytokines, alpha 4 beta 7 integrin, IgEs, viral proteins (for example, cytomegalovirus). Examples of antibodies that can be produced using the compositions and

methods of the invention include, but are not limited to, an anti-TNF α antibody, an anti-IL-12 antibody, an anti-IL-18 antibody, and an anti-EPO receptor (EPO-R) antibody. In one embodiment, the anti-TNF α antibody is a fully human anti-TNF α antibody, e.g., adalimumab/D2E7 (see U.S. Pat. No. 6,090,382, incorporated by reference herein; Humira®; Abbott Laboratories). In one embodiment, the anti-IL-12 antibody is a fully human, anti-IL-12 antibody, e.g., ABT-874 (Abbott Laboratories; see U.S. Pat. No. 6,914,128, incorporated by reference herein). In one embodiment, the anti-IL-18 antibody is a fully human IL-18 antibody (e.g., ABT-325), e.g. see also antibodies described in US20050147610 A1, incorporated by reference herein. In one embodiment, the anti-EPO/R (also referred to as ABT-007) antibody is a fully human antibody, like that described in US Patent Publication No. US 20060018902 A1, hereby incorporated by reference.

[0127] In addition, the constant regions encoded in the vector may also be used to operably link a constant region, e.g., an Fc domain, to a protein to form a fusion protein, e.g., an Fc-fusion protein. Thus, another example of the type of protein that may be produced using the methods and compositions of the invention include fusion proteins. Examples of such fusion proteins include proteins expressed as a fusion with a portion of an immunoglobulin molecule, proteins expressed as fusion proteins with a zipper moiety, and novel polyfunctional proteins such as a fusion proteins of a cytokine and a growth factor (i.e., GM-CSF and IL-3, MGF and IL-3). WO 93/08207 and WO 96/40918 describe the preparation of various soluble oligomeric forms of a molecule referred to as CD40L, including an immunoglobulin fusion protein and a zipper fusion protein, respectively; the techniques discussed therein are applicable to other proteins. Another fusion protein is a recombinant TNFR:Fc, also known as entanercept. Entanercept (or Enbrel®; Amgen/Wyeth) is a dimer of two molecules of the extracellular portion of the p75 TNF alpha receptor, each molecule consisting of a 235 amino acid TNFR-derived polypeptide that is fused to a 232 amino acid Fc portion of human IgG1. In fact, any molecule can be expressed as a fusion protein including, but not limited to, the extracellular domain of a cellular receptor molecule, an enzyme, a hormone, a cytokine, a portion of an immunoglobulin molecule, a zipper domain, and an epitope.

[0128] Techniques for determining nucleic acid and amino acid “sequence identity” also are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. In general, “identity” refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their “percent identity.” The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff, *Atlas of Protein Sequences and Structure*, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Bio-medical Research Foundation, Washington, D.C., USA, and

normalized by Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). An exemplary implementation of this algorithm to determine percent identity of a sequence is provided by the Genetics Computer Group (Madison, Wis.) in the “BestFit” utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, Wis.). A preferred method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, Calif.). From this suite of packages the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated the “Match” value reflects “sequence identity.” Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

[0129] Two nucleic acid fragments are considered to “selectively hybridize” as described herein. The degree of sequence identity between two nucleic acid molecules affects the efficiency and strength of hybridization events between such molecules. A partially identical nucleic acid sequence will at least partially inhibit a completely identical sequence from hybridizing to a target molecule. Inhibition of hybridization of the completely identical sequence can be assessed using hybridization assays that are well known in the art (e.g., Southern blot, Northern blot, solution hybridization, or the like, see Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, New York; or Ausubel et al. (Eds.), *Current Protocols In Molecular Biology*, John Wiley & Sons, Inc., New York (1997)). Such assays can be conducted using varying degrees of selectivity, for example, using conditions varying from low to high stringency. If conditions of low stringency are employed, the absence of non-specific binding can be assessed using a secondary probe that lacks even a partial degree of sequence identity (for example, a probe having less than about 30% sequence identity with the target molecule), such that, in the absence of non-specific binding events, the secondary probe will not hybridize to the target.

[0130] When utilizing a hybridization-based detection system, a nucleic acid probe is chosen that is complementary to a target nucleic acid sequence, and then by selection of appropriate conditions the probe and the target sequence “selectively hybridize,” or bind, to each other to form a hybrid molecule. A nucleic acid molecule that is capable of hybridizing selectively to a target sequence under “moderately stringent” typically hybridizes under conditions that allow detection of a target nucleic acid sequence of at least about 10-14 nucleotides in length having at least approximately 70% sequence identity with the sequence of the selected nucleic acid probe. Stringent hybridization conditions typically allow detection of target nucleic acid sequences of at least about 10-14 nucleotides in length having a sequence identity of greater than about 90-95% with the sequence of the selected nucleic acid probe. Hybridization conditions useful for probe/target hybridization where the probe and target have a specific degree of sequence identity, can be determined as is known in the art (see, for example, *Nucleic Acid Hybridization: A Practical Approach*, editors B. D. Hames and S. J. Higgins, (1985) Oxford; Washington, D.C.; IRL Press).

[0131] With respect to stringency conditions for hybridization, it is well known in the art that numerous equivalent conditions can be employed to establish a particular stringency by varying, for example, the following factors: the length and nature of probe and target sequences, base composition of the various sequences, concentrations of salts and other hybridization solution components, the presence or absence of blocking agents or detergents in the hybridization solutions (e.g., formamide, dextran sulfate, and polyethylene glycol, and sodium dodecyl sulphate), hybridization reaction temperature and time parameters, as well as, varying wash conditions. The selection of a particular set of hybridization conditions is selected following standard methods in the art (see, for example, see Sambrook, et al., *supra* or Ausubel et al., *supra*). A first polynucleotide is “derived from” second polynucleotide if it has the same or substantially the same base pair sequence as a region of the second polynucleotide, its cDNA, complements thereof, or if it displays sequence identity as described above. A first polypeptide is “derived from” a second polypeptide if it is (i) encoded by a first polynucleotide derived from a second polynucleotide, or (ii) displays sequence identity to the second polypeptides as described above.

[0132] The invention also provides a kit containing one or more vectors of the invention in a suitable vessel such as a vial. The expression vectors can contain at least one cloning site for insertion of a selected sequence of interest, or can have a specific gene of interest already present in the vector. The vector can be provided in a dehydrated or lyophilized form, or in an aqueous solution. The kit can include a buffer for reconstituting the dehydrated polynucleotide. Other reagents can be included in the kit, e.g., reaction buffers, positive and negative control vectors for comparison. Generally, the kit will also include instructions for use of the reagents therein.

III. Uses of Vectors of Invention

[0133] The invention includes methods of expressing proteins using the vectors described herein. Thus, the invention includes a method of producing a recombinant protein comprising introducing the expression vector of the invention into a mammalian host cell, culturing the mammalian host cell under suitable conditions so as to express the protein, and recovering the protein. An advantage of the vector of the invention is that it provides high protein production using mammalian cell culture systems.

[0134] Any cell type capable of gene expression via a nucleic acid or expression vector of the present invention can be used in the present invention as a host cell. The term “host cells” refers to cells that have been transformed with a vector constructed using recombinant DNA techniques.

[0135] Those having ordinary skill in the art can select a particular host cell line that is best suited for expressing the GOI and selectable marker gene via a vector of the present invention. Cells that can be employed in this invention include mammalian cells and cell lines and cell cultures derived therefrom. Mammalian cells, e.g., germ cells or somatic cells, can be derived from mammals, such as mice, rats, or other rodents, or from primates, such as humans or monkeys. It shall be understood that primary cell cultures or immortalized cells can be employed in carrying out the techniques of this invention.

[0136] In particular embodiments, the cell type is mammalian in origin including, but not limited to Chinese hamster ovary (CHO) (e.g., DG44 and DUXB11; Urlaub et al., *Som.*

Cell Molec. Genet. 12:555, 1986; Haynes et al., *Nuc. Acid. Res.* 11:687-706, 1983; Lau et al., *Mol. Cell. Biol.* 4:1469-1475, 1984; *Methods in Enzymology*, 1991, vol. 185, pp 537-566. Academic Press, Inc., San Diego, Calif.), Chinese hamster fibroblast (e.g., R1610), human cervical carcinoma (e.g., HELA), monkey kidney line (e.g., CV1 and COS), murine fibroblast (e.g., BALBc/3T3), murine myeloma (P3. times.63-Ag3.653; NS0; SP2/O), hamster kidney line (e.g., HAK), murine L cell (e.g., L-929), human lymphocyte (e.g., RAM, human kidney (e.g., 293 and 293T). Host cell lines are typically commercially available (e.g., from BD Biosciences, Lexington, Ky.; Promega, Madison, Wis.; Life Technologies, Gaithersburg, Md.) or from the American Type Culture Collection (ATCC, Manassas, Va.).

[0137] In a preferred embodiment, the host cell used in the invention provides in trans the replication initiation factor corresponding to at least one origin of replication included in the vector of the invention. For example, if the vector comprises two origins of replication corresponding the SV40 origin and the OriP origin, any cell line, preferably mammalian, that expresses either the large T-antigen or the EBNA protein can be used. In one embodiment, the vector is transformed into a COS cell or a human embryonic kidney (HEK) cell. For example, COST cells are derived from CV-1 simian cells transformed by an origin-defective mutant of SV40 (Sigma-Aldrich). EBNA may be provided, for example, by using the HEK-293-6E cell.

[0138] Cell lines that have stably integrated replication initiation factors within the genome have the advantage of stable long-term expression of the replication initiation factor and durable support of replication and maintenance of the origin of replication containing plasmids. Examples of commercially available cell lines expressing EBNA-1 are ATCC: 293HEK-EBNA1 and CV1-EBNA1. Specific cell lines over expressing at least one replication initiation factor, preferably the EBNA1 protein or the SV40 large T-antigen, can be generated by transfection and selection of stable cell clones.

[0139] Nucleic acids and expression vectors can be introduced or transformed into an appropriate host cell by various techniques well known in the art (see, e.g., Ridgway, 1973, *Vectors: Mammalian Expression Vectors*, Chapter 24.2, pp. 470-472, Rodriguez and Denhardt eds., Butterworths, Boston, Mass.; Graham et al., 1973, *Virology* 52:456; Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, New York; Davis et al., 1986, *Basic Methods in Molecular Biology*, Elsevier; and Chu et al., 1981, *Gene* 13:197). The terms “transformation” and “transfection”, and their grammatical variations, are used interchangeably herein and refer to the uptake of foreign DNA by a cell by any means practicable. A cell has been “transformed” when an exogenous nucleic acid has been introduced inside the cell membrane. The uptake of the nucleic acid results in a stable transfectant, regardless of the means by which the uptake is accomplished, which may include transfection (including electroporation), protoplast fusion, calcium phosphate precipitation, cell fusion with enveloped DNA, microinjection, and infection with intact virus. Even transient expression at higher than normal levels is useful for functional studies or for the production and recovery of proteins of interest. Transformed cells are grown under conditions appropriate for the production of the protein of interest (e.g., antibody heavy and/or light chains in one embodiment), and assays are performed to identify the encoded polypeptide of interest. Exemplary assay techniques for identifying and

quantifying gene products include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), or fluorescence-activated cell sorter analysis (FACS), immunohistochemistry, and the like.

[0140] Cells used in the present invention can be cultured according to standard cell culture techniques, e.g., they can be fixed to a solid surface or grown in suspension in appropriate nutrient media.

[0141] Also encompassed by the invention is a mammalian host cell comprising the vectors described herein.

[0142] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology and the like, which are within the skill of the art. Such techniques are explained fully in the literature. See, e.g., *Molecular Cloning: A Laboratory Manual*, (J. Sambrook et al., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989); *Current Protocols in Molecular Biology* (F. Ausubel et al., eds., 1987 updated); *Essential Molecular Biology* (T. Brown ed., IRL Press 1991); *Gene Expression Technology* (Goeddel ed., Academic Press 1991); *Methods for Cloning and Analysis of Eukaryotic Genes* (A. Bothwell et al. eds., Bartlett Publ. 1990); *Gene Transfer and Expression* (M. Kriegler, Stockton Press 1990); *Recombinant DNA Methodology* (R. Wu et al. eds., Academic Press 1989); *PCR: A Practical Approach* (M. McPherson et al., IRL Press at Oxford University Press 1991); *Oligonucleotide Synthesis* (M. Gait ed., 1984); *Cell Culture for Biochemists* (R. Adams ed., Elsevier Science Publishers 1990); *Gene Transfer Vectors for Mammalian Cells* (J. Miller & M. Calos eds., 1987); *Mammalian Cell Biotechnology* (M. Butler ed., 1991); *Animal Cell Culture* (J. Pollard et al. eds., Humana Press 1990); *Culture of Animal Cells*, 2nd ed. (R. Freshney et al. eds., Alan R. Liss 1987); *Flow Cytometry and Sorting* (M. Melamed et al. eds., Wiley-Liss 1990); the series *Methods in Enzymology* (Academic Press, Inc.); and *Animal Cell Culture* (R. Freshney ed., IRL Press 1987); and Wirth M. and Hauser H. (1993) *Genetic Engineering of Animal Cells*, In: *Biotechnology Vol. 2* Puhler A (ed.) VCH, Weinheim 663-744.

EXEMPLIFICATION

[0143] The following examples illustrate an innovative solution to eliminate the need to construct separate vectors for different mammalian host cells, e.g., COS7 and HEK-293-6E cells. The following examples also provide vectors containing nucleic acids encoding constant regions of antibodies, for use in the expression of complete light or heavy chains of an antibody or in the expression of Fc fusion proteins.

[0144] Two new vector backbones, termed pHyb-C and pHyb-E, were constructed by combining selected features from various other vectors, i.e., the pBOS and pTT3 vectors (see U.S. Provisional Appln. No. 60/878,165, International Appln. No. PCT/US2007/26482, filed on Dec. 28, 2007 entitled "DUAL-SPECIFIC IL-1A/IL-1b ANTIBODIES" and U.S. Ser. No. 12/006,068, all of which are hereby incorporated by reference herein). Control vector pBOS contains the EF-1 α promoter operably linked to the insertion site for the gene of interest, and carries the SV40 replication origin. Control vector pTT3 contains the CMV promoter operably linked to the insertion site for the gene of interest, and an EBNA replication origin (OriP).

[0145] The vectors of the invention were tested by evaluating protein expression of both a mouse BR3-Fc fusion and a human antibody (adalimumab) in both COS7 and HEK-293-

6E cells. The successful protein expression in COS7 and HEK-293-6E cells demonstrates a unifying vector system for recombinant expression in both cell types.

Example 1

Construction of Vectors pHybC and pHybE

[0146] FIGS. 1 and 2 provide maps of the new vectors, which each contain two origins of replication. FIGS. 1 and 2 represent "empty" versions of the vectors, i.e., do not contain the nucleic acid of the gene of interest or the antibody constant regions (described in more detail below in Example 4). pHybC contains the CMV promoter operably linked to the insertion site for the gene of interest, while pHybE contains the EF-1 α promoter.

[0147] For pHybC-mBR3-Fc construction ("mBR3" refers to the murine version of the third BLYS receptor, and as used herein refers specifically to the coding sequence for the extracellular domain (ECD) portion of the mBR3 protein), the SV40 origin of replication region from the pEF-BOS vector was PCR amplified with primers that introduced PspX I restriction sites at both 5' and 3' ends of the amplified DNA fragment. This insertion fragment was then digested by PspX I. A pTT3-mBR3-Fc construct, having a Sal I restriction site upstream of the CMV promoter, was digested with Sal I. Then the Psp X I-digested insertion fragment was ligated into the Sal I site of pTT3-mBR3-Fc to create the pHybC-mBR3-Fc vector.

[0148] The pHybE-mBR3-Fc construct was created by first amplifying by PCR a 5'-end PspX I modified DNA fragment containing the SV40 origin of replication region through the mBR3 extracellular domain. This product was then digested at 5' by PspX I and 3' by Bsp68 I, which has a site in the leader sequence upstream of the mBR3 extracellular domain sequence. This digested fragment was subsequently subcloned into a Sal I and Bsp68 I-digested pTT3-mBR3-Fc construct to produce the pHybE-mBR3-Fc construct.

[0149] Maps of pHybC-mBR3-Fc and pHybE-mBR3-Fc, which each express the receptor-Fc fusion protein mBR3-Fc, can be found in FIGS. 8 and 9.

[0150] The pHybC-E7 vector expressing the light chain protein of D2E7 antibody (adalimumab) was similarly constructed as the pHybC-mBR3-Fc, i.e. by ligating the same PspX I digested SV40 Ori region that was isolated and digested during the creation of pHybC-mBR3-Fc (described above) into a previously constructed pTT3-E7 vector predigested by Sal I.

[0151] For pHybE-E7 vector construction, an insert fragment was generated by digestion of a pre-existing pBOS-E7 vector with Hind III and BsiW I restriction enzymes. This insert fragment was then ligated into a pHybC-E7 vector predigested with the same enzymes to generate pHybE-E7 for the expression of the D2E7 light chain protein.

[0152] For pHybC-D2 and pHybE-D2 vector construction, an insert fragment consisting of the heavy chain variable and constant coding regions of the D2E7 antibody (Adalimumab) (i.e. the D2 heavy chain coding sequence) was generated by digesting a pre-existing pTT3-D2 vector with Bsp68 I and Not I restriction enzymes. This insert fragment was ligated into pHybC-mBR3-Fc and pHybE-mBR3-Fc vectors predigested with the same enzymes to generate pHybC-D2 and pHybE-D2, respectively, for the expression of the heavy chain protein of D2E7 antibody (Adalimumab).

Example 2

Comparison of Protein Yield

[0153] To determine whether the increase in vector size with the addition of two origins of replication impacted protein production by the vectors, the pHyb-E and pHyb-C vectors described above were compared to control vectors pBOS and pTT3, which each only contained one origin of replication. To compare expression from pBOS, pTT3, pHyb-C and pHyb-E, a mouse BAFF receptor-human Fc fusion protein construct (mBR3-Fc) was subcloned into the four vector backbones and prepared in parallel by endo-free DNA prep kit.

[0154] The four vectors containing the mBR3-Fc sequence were electroporated into COS cells or transfected into HEK-293-6E cells (protocols described below). The cells were incubated for a period of five or seven days. Media samples were taken and the concentration of the mBR3-Fc secreted protein in the media was measured. Titers were determined by IgG ELISA and adjusted by difference in molecular weight between IgG protein standard and the mBR3-Fc protein from the conditioned media after 5 days for COS cells and 7 days for HEK-293-6E cells. The titer adjustment is required to prevent overestimation of mBR3-Fc protein titer due to the use of a much larger human IgG protein as standards in the ELISA.

293 Transfection

[0155] The 293 transient transfection procedure used in the experiment was a modification of the methods published in Durocher et al. (2002); *Nucleic Acids Research* 30(2):E9 and Pham et al. (2005); *Biotechnology Bioengineering* 90(3): 332-44. Reagents that were used in the transfection included:

[0156] HEK 293-6E cells (human embryonic kidney cell line stably expressing EBNA1; obtained from National Research Council Canada) cultured in disposable Erlenmeyer flasks in a humidified incubator set at 130 rpm, 37° C. and 5% CO₂.

[0157] Culture medium: FreeStyle 293 Expression Medium (Invitrogen 12338-018) plus 25 µg/mL Geneticin (G418) (Invitrogen 10131-027) and 0.1% Pluronic F-68 (Invitrogen 24040-032).

[0158] Transfection medium: FreeStyle 293 Expression Medium plus 10 mM HEPES (Invitrogen 15630-080).

[0159] Polyethylenimine (PEI) stock: 1 mg/mL sterile stock solution, pH 7.0, prepared with linear 25 kDa PEI (Polysciences) and stored at less than -15° C.

[0160] Tryptone Feed Medium: 5% w/v sterile stock of Tryptone N1 (Organotechnie, 19554) in FreeStyle 293 Expression Medium.

[0161] Cell preparation for transfection: Approximately 2-4 hours prior to transfection, HEK 293-6E cells were harvested by centrifugation and resuspended in culture medium at a cell density of approximately 1 million viable cells per mL. For each transfection, 40 mL of the cell suspension was transferred into a disposable 250-mL Erlenmeyer flask and incubated for 2-4 hours.

[0162] Transfection: The transfection medium and PEI stock were prewarmed to room temperature (RT). For each transfection, 25 µg of plasmid DNA and 50 µg of polyethylenimine (PEI) were combined in 5 mL of transfection medium and incubated for 15-20 minutes at RT to allow the DNA:PEI complexes to form. For the BR3-Ig transfections,

25 µg of BR3-Ig plasmid was used per transfection. Each 5-mL DNA:PEI complex mixture was added to a 40-mL culture prepared previously and returned to the humidified incubator set at 130 rpm, 37° C. and 5% CO₂. After 20-28 hours, 5 mL of Tryptone Feed Medium was added to each transfection and the cultures were continued for six days.

COS7 Cell Transfection

[0163] Two COS7 150 mm plates per construct were transfected using standard electroporation conditions as follows. For COS7 transfection experiments, COS cells were cultured in DMEM+10% FBS+1× glutamine. Cells from one confluent T-150 flask were used for electroporation. The cells were trypsinized, and spun down in media plus serum to inactivate serum. Cells were then washed in 1×PBS.

[0164] For each T-150, the pellet was resuspended in 0.8 mL electroporation buffer. The COS electroporation buffer included 20 mM Hepes (or P3 buffer), 137 mM NaCl, 5 mM KCl, 0.7 mM Na₂HPO₄, and 6 mM Dextrose. The electroporation buffer was adjusted to a pH of 7.0 and filter sterilized. Sixty micrograms of DNA (30 µg of each heavy and light chain plasmid DNA or 60 µg DNA in the case of an Fc fusion protein) was used for each electroporation. 0.8 mL of cell/buffer/DNA was mixed to each cuvette. (0.4 cm cuvette—Biorad). In addition, one cuvette was set up with buffer only to use as a blank. Cuvettes were put on ice. Cells were electroporated at 250V and 950 µF for 15 to 25 milliseconds. Cuvettes were then returned to ice. The contents of 2 cuvettes were transferred into one 50 mL conical containing 20 mL Hybridoma SFM. A 10 mL pipette was used to break up clumps and transfer to two 150 mm tissue culture dishes, each containing another 20 mL media. Total media volume in each dish was then 30 mL. The dishes were then incubated at 37° C., 5% CO₂ for three days.

[0165] The COS cell conditioned media (supernatant) was collected into 50 mL conical tubes and spun down. Following the spin, the supernatant was filtered using 2 micron (µm) filter. A sample was removed for ELISA analysis. Supernatants were collected after 5 days and analyzed in a standard IgG ELISA to determine their respective protein yields.

[0166] pBOS, pTT3, pHybC and pHybE versions of vectors were tested separately in the mBR3 and adalimumab (D2E7) experiments.

Protein Testing

[0167] The mBR3-Fc fusion protein concentrations in culture supernatants were tested 5 days (for COS7 cells) or 7 days post-transfection (for 293-6E cells) using ELISA and/or Poros A.

Results

[0168] Data showing protein expression levels from the control and experimental transfections are shown in FIG. 3 (COS cells) and FIG. 4 (HEK-293 cells). The data in FIG. 3 shows that pHybC and pHybE were both effective at producing the fusion protein in COS cells, where both vectors expressed higher levels than control vector pTT3. The data presented in FIG. 4 shows that the expression levels from HEK cells transfected with the pHyb-E exceeded the expression seen with the other three vectors, while pHyb-C protein production levels were comparable with the controls. Thus,

both pHyb-C and pHyb-E were able to express the mBr3-Fc fusion protein as well as, if not better than, control vectors pTT3 and pBOS.

Example 3

Comparison of Protein Yield That Requires Co-transfection of Two DNA Constructs

[0169] A human IgG1/k monoclonal antibody to TNF α (adalimumab)/D2E7 was subcloned into the four vector backbones and prepared in parallel by endo-free DNA prep kit.

[0170] The four vectors containing sequences for expression of adalimumab were electroporated into COS cells; HEK-293-6E cells were transfected using poly(ethyleneimine) (PEI).

[0171] The 293 transient transfection procedure used was the same as that described in Example 3, except for the adalimumab transfections, in which 10 μ g of the D2E7 heavy chain (referred to as "D2") plasmid and 15 μ g of the D2E7 light chain (referred to as "E7") plasmid were used per transfection.

[0172] The COS7 transfection experiments were performed as described above, except 30 μ g of each heavy and light chain vector was used per plate transfection.

[0173] The adalimumab antibody concentrations in culture supernatants were tested 7 days post-transfection using ELISA and/or Poros A. Titers were determined by IgG ELISA from the conditioned media after 5 days for COS7 cells and 7 days for HEK-293-6E cells.

[0174] Data showing protein expression levels from the control and experimental transfections are shown in FIG. 5 (HEK-293 cells) and FIG. 6 (COS cells). Data in FIG. 5 shows that both pHybC and pHybE backbone vectors were able to produce more adalimumab than control vector pBOS, and comparable (pHybC) or greater (pHybE) quantities than control vector pTT3 (Durocher, Y. et al. *Nucleic Acids Res.* 30:E9 (2002)). Similarly, the data in FIG. 6 shows that both pHybC and pHybE backbone vectors were able to produce more protein than control vector pTT3 and comparable levels to control vector pBOS.

Example 4

Construction of the pHyb-E Antibody Constant Region Vector

[0175] To facilitate the creation of vectors that could be used for antibody production using the new pHyb-E vector backbone, a panel of twelve different heavy and light chain vectors was generated (overview provided in Tables 2 and 3). Twelve master template pHybE vectors that allow for both human and mouse IgG expression were constructed.

[0176] To create the vectors described in FIGS. 14-25, a 6123 bp Srf I/Not I fragment was isolated from pHybE-stuffer-hCg1,z,a (pJP167) and ligated with Srf I/Not I restriction fragments from the pBOS vectors consisting of the signal peptide coding region, lambda stuffer, and constant region coding region. To create the Srf I/Not I restriction fragments, Srf I/Not I restriction digests were performed, in order to generate insertion fragments consisting of the signal peptide coding region, lambda stuffer, and constant region coding region (for constant region sequences, see Table 1). These fragments were derived from pBOS master templates that had been constructed into the pEF-BOS plasmid DNA (see

Mizushima, S. and Nagata, S. *Nucleic Acids Res.* 18:5322 (1990); also described in U.S. Provisional Application No. 60/878,165, International Application No. PCT/US2007/026482, filed on Dec. 28, 2007 entitled "DUAL-SPECIFIC IL-1A/IL-1b ANTIBODIES") and U.S. Ser. No. 12/006,068, incorporated by reference herein). The insertion fragment for the pHybE-hC1 construct was first modified by overlapping PCR to create an AfeI restriction site at the 3' end of the J region to facilitate cloning into this vector. All inserts were ligated into a previously sequence validated pHybE construct predigested with Srf I and Not I to generate the following vectors.

[0177] The new constant region-containing vectors were then sequence-verified for mouse and human antibody constant regions (see SEQ ID NOS: 3-32).

[0178] The vectors described in Tables 2 and 3 all have a ~1-kb 'stuffer' sequence (of λ phage DNA) that can be swapped out by the variable region sequences. These new master vectors also contain a new Sma I restriction site directly upstream of the Srf I site. This novel Sma I site is useful for transferring the antibody open reading frame from pHyb-E to other expression vectors that also utilize a Sma I site for cloning purposes, such as CHO expression vectors. In addition to the flexibility of alternative cloning sites, these vectors are also backward compatible with existing pBOS, pTT3, and CHO vectors.

[0179] As seen in FIG. 7, preliminary transfection data in COST cells showed that this additional Sma I site (v1 vectors) had no significant effect on the levels of adalimumab expression when compared with the constructs without the additional Sma I site (v2 vectors).

TABLE 1

Constant region sequences	
constant region	location of sequence
mCk	2285 to 2605 of SEQ ID NO: 3
mCyl	2277 to 3251 of SEQ ID NO: 5
mCyl2a	2277 to 3269 of SEQ ID NO: 7
hCk	2287 to 2610 of SEQ ID NO: 9
hC λ	2269 to 2588 of SEQ ID NO: 11
hCyl1, z, a	2277 to 3269 of SEQ ID NO: 13
hCyl1, z, non-a	2277 to 3269 of SEQ ID NO: 15
hCyl1, z, non-a, mut(234, 235)	2277 to 3269 of SEQ ID NO: 17
hCyl1, z, non-a, mut(234, 237)	2277 to 3269 of SEQ ID NO: 19
hCyl2 (n-)	2277 to 3257 of SEQ ID NO: 21
hCyl2 (n+)	2277 to 3257 of SEQ ID NO: 23
hCyl4	2277 to 3260 of SEQ ID NO: 25

TABLE 2

Exemplary Master Set of pHybE Vectors Made for Human and Mouse IgG Expression		
	Heavy Chain Vectors	Light Chain Vectors
Human	pHybE-, hCg1, z, a	pHybE-hCk
	pHybE-, hCg1, z, non-a	pHybE-hC1
	pHybE-, hCg1, z, non-a, (mut 234, 235)	
	pHybE-, hCg1, z, non-a, (mut 234, 237)	
	pHybE-, hCg2, n+	
Mouse	pHybE-, hCg2, n-	
	pHybE-, hCg4	
	pHybE-mCg1	pHybE-mCk
	pHybE-mCg2a	

Summary:

[0180] The preceding experiments described in Examples 1-4 show that the pHyb-C and pHyb-E vectors are functional in more than one cell line while provide ample protein expression that often exceeded the expression levels seen with the original pBOS and pTT3 vectors. This heightened expression was particularly pronounced when the pHyb-E vector was used to express the low yielding mBR3-Fc fusion protein in HEK-293-6E cells. As shown by this data, the pHyb-C and pHyb-E vectors represent a significant advancement in vector technology over previously used vectors.

TABLE 3

Overview of vectors of invention	
pHyb vectors described as version 1 have an additional Swa I site upstream of the Srf I restriction site.	
pHyb vectors described as version 2, do not have additional Swa I site.	

SEQ ID NO	DESCRIPTION OF NUCLEIC ACID
1	pHybC-empty
2	pHybE-empty
3	pJP180; pHybE-mCk V1
4	pJP193; pHybE-mCk V2
5	pJP176; pHybE-mCg1 V1
6	pJP189; pHybE-mCg1 V2
7	pJP177; pHybE-mCg2a V1
8	pJP190; pHybE-mCg2a V2
9	pJP178; pHybE-hCk V1
10	pJP191; pHybE-hCk V2
11	pJP179; pHybE-hCl V1
12	pJP192; pHybE-hCl V2
13	pJP170; pHybE-hCg1, z, a V1
14	pJP182; pHybE-hCg1, z, a V2
15	pJP171; pHybE-hCg1, z, non-a V1
16	pJP183; pHybE-hCg1, z, non-a V2
17	pJP172; pHybE-hCg1, z, non-a, mut(234, 235) V1
18	pJP184; pHybE-hCg1, z, non-a, mut(234, 235) V2
19	pJP173; pHybE-hCg1, z, non-a, mut (234, 237) V1
20	pJP185; pHybE-hCg1, z, non-a, mut (234, 237) V2

TABLE 3-continued

Overview of vectors of invention	
pHyb vectors described as version 1 have an additional Swa I site upstream of the Srf I restriction site.	
pHyb vectors described as version 2, do not have additional Swa I site.	
SEQ ID NO	DESCRIPTION OF NUCLEIC ACID
21	pJP174; pHybE-hCg2, n- V1
22	pJP187; pHybE-hCg2, n- V2
23	pJP181; pHybE-hCg2, n+ V1
24	pJP186; pHybE-hCg2, n+ V2
25	pJP175; pHybE-hCg4 V1
26	pJP188; pHybE-hCg4 V2
27	pHybC-mBR3-mCg2a
28	pHybE-mBR3-mCg2a
29	pHybC-E7-hCk
30	pHybC-D2-hCg1, z, a
31	pHybE-D2-hCg1, z, a
32	pHybE-E7-hCk

EQUIVALENTS

[0181] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

INCORPORATION BY REFERENCE

[0182] The contents of all cited references (including literature references, patents, patent applications, and websites) that maybe cited throughout this application are hereby expressly incorporated by reference in their entirety for any purpose, as are the references cited therein. The practice of the present invention will employ, unless otherwise indicated, conventional techniques of immunology, molecular biology, cell biology, and drug manufacturing and delivery, which are well known in the art. These techniques include, but are not limited to, techniques described in the following publications:

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 32

<210> SEQ ID NO 1
 <211> LENGTH: 6381
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic construct: pHybC-empty
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (1143) .. (1242)
 <223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 1

```

ttgacattga ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag      60
cccatatatg gagttccgcg ttacataact tacggtaaat gggccgcctg gctgaccgcc      120
caacgacccc cgccattga cgtaataat gacgtatgtt cccatagtaa cgccaatagg      180
gactttccat tgacgtcaat ggggtggagta ttacggtaa actgcccact tggcagtaca      240
tcaagtgtat catatgccaa gtccgccccc tattgacgtc aatgacggta aatggccgcg      300
ctggcattat gccagtaga tgaccttacg ggactttcct acttggcagt acatctacgt      360

```

-continued

attagtcac	gctattacca	tggatgatgcg	gttttggcag	tacaccaatg	ggcgtggata	420
gcggtttgac	tcacggggat	ttccaagtct	ccacccatt	gacgtcaatg	ggagtttgtt	480
ttggcacc	aatcaacggg	actttccaaa	atgtcgtaat	aaccccgccc	cgttgacgca	540
aatggg	aggcgtgtac	ggtgggaggt	ctatataagc	agagctcgtt	tagtgaaccg	600
tcagatcctc	actctcttcc	gcacgcgtgt	ctgcgagggc	cagctgttgg	gctcgcgggt	660
gaggacaaac	tcttcgcggg	ctttccagta	ctcttgatc	ggaacccgt	cggcctccga	720
acggtactcc	gccaccgagg	gacctgagcg	agtcgcgcatc	gaccggatcg	gaaaacctct	780
cgagaaaggc	gtctaaccag	tcacagtcgc	aaggtaggct	gagcaccgtg	gcgggcggca	840
gcgggtggcg	gtcgggggtg	tttctggcgg	aggtgctgct	gatgatgtaa	ttaaagtagg	900
cggctctgag	acggcggtatg	gtcaggtga	gggtggtggcag	gcttgagatc	cagctgttgg	960
ggtagtact	ccctctcaaa	agcgggcatt	actctgcgc	taagattgtc	agtttccaaa	1020
aacgaggagg	atttgatatt	cacctggccc	gatctggcca	tacacttgag	tgacaatgac	1080
atccactttg	cctttctctc	cacaggtgtc	cactcccagg	tccaagtgtg	ggcgccacca	1140
tgnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nntgagcggc	cgctcgaggc	1260
cggcaaggcc	ggatcccccg	acctcgacct	ctggctaata	aaggaaattt	attttcattg	1320
caatagtgtg	ttggaatttt	ttgtgtctct	cactcggaag	gacatatggg	agggcaaatac	1380
atttggtcga	gatccctcgg	agatctctag	ctagaggatc	gatccccgcc	ccggacgaac	1440
taaacctgac	tacgacatct	ctgccccctc	ttcgcggggc	agtgcagtga	atcccttcag	1500
ttggttggtg	caacttgcca	actgggccct	gttccacatg	tgacacgggg	ggggaccaaa	1560
cacaaagggg	ttctctgact	gtagttgaca	tccttataaa	tggatgtgca	catttgccaa	1620
cactgagtgg	ctttcatcct	ggagcagact	ttgcagtctg	tggactgcaa	cacaacattg	1680
cctttatgtg	taactcttgg	ctgaagctct	tacaccaatg	ctggggggaca	tgtacctccc	1740
agggggccag	gaagactacg	ggaggctaca	ccaacgtcaa	tcagaggggc	ctgtgtagct	1800
accgataaag	ggacctcaa	gagggcatta	gcaatagtgt	ttataaggcc	cccttggttaa	1860
ccctaaccgg	gtagcatatg	cttcccgggt	agtagtatat	actatccaga	ctaaccctaa	1920
ttcaatagca	tatgttacc	aacgggaagc	atatgctatc	gaattagggt	tagtaaaagg	1980
gtcctaagga	acagcgatat	ctcccacccc	atgagctgtc	acggttttat	ttacatgggg	2040
tcaggattcc	acgagggtag	tgaaccattt	tagtcacaag	ggcagtggtc	gaagatcaag	2100
gagcggggcag	tgaactctcc	tgaatcttcg	cctgcttctt	cattctcctt	cgttttagcta	2160
atagaataac	tgctgagttg	tgaacagtaa	gggtgatgtg	aggtgctcga	aaacaagggt	2220
tcaggtgacg	cccccagaat	aaaatttggg	cgggggggtc	agtgggtggca	ttgtgctatg	2280
acaccaatat	aaccctcaca	aacccttggg	gcaataaata	ctagtgtagg	aatgaaacat	2340
tctgaatatc	tttaacaata	gaaatccatg	gggtggggac	aagccgtaaa	gactggatgt	2400
ccatctcaca	cgaatttatg	gctatgggca	acacataatc	ctagtgcaat	atgatactgg	2460
ggttattaag	atgtgtccca	ggcagggacc	aagacaggtg	aaccatgttg	ttacactcta	2520
tttgtaacaa	ggggaagag	agtggacgcc	gacagcagcg	gactccactg	gttgtctcta	2580
acacccccga	aaattaacg	gggctccacg	ccaatggggc	ccataaacaa	agacaagtgg	2640
ccactctttt	ttttgaaatt	gtggagtggtg	ggcacgcgtc	agccccaca	cgcgcctctg	2700

-continued

cggttttga	ctgtaaaata	agggtgtaat	aacttggtg	attgtaacc	cgctaaccac	2760
tgcggtcaaa	ccacttgccc	acaaaaccac	taatggcacc	cggggaata	cctgcataag	2820
tagtgggcg	ggccaagata	ggggcgcat	tgctgcgat	tggaggacaa	attacacaca	2880
cttgccctg	agcgccaagc	acagggttgt	tggctctcat	attcacgagg	tcgctgagag	2940
cacgggtggc	taatgttgcc	atgggtagca	tatactaccc	aaatatctgg	atagcatatg	3000
ctatccta	ctatatctgg	gtagcatagg	ctatccta	ctatatctgg	gtagcatatg	3060
ctatccta	ctatatctgg	gtagtatatg	ctatccta	ttatatctgg	gtagcatagg	3120
ctatccta	ctatatctgg	gtagcatatg	ctatccta	ctatatctgg	gtagtatatg	3180
ctatccta	ctgtatccg	gtagcatatg	ctatccta	agagattagg	gtagtatatg	3240
ctatccta	ttatatctgg	gtagcatata	ctacccaa	atctggatag	catatgctat	3300
cctaatact	atctgggtag	catatgctat	cctaatact	atctgggtag	cataggctat	3360
cctaatact	atctgggtag	catatgctat	cctaatact	atctgggtag	tatatgctat	3420
cctaatttt	atctgggtag	cataggctat	cctaatact	atctgggtag	catatgctat	3480
cctaatact	atctgggtag	tatatgctat	cctaatact	atcgggtag	catatgctat	3540
cctcatgata	agctgtcaaa	catgagaatt	ttcttgaaga	cgaaggggc	tcgtgatag	3600
cctattttt	taggttaaat	tcagtataat	aatggtttct	tagacgtcag	gtggcacttt	3660
tcggggaat	gtgcgcggaa	cccctatttg	ttatttttct	taaatacatt	caaataatgta	3720
tcgctcatg	agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	3780
gagtattcaa	catttccgtg	tcgcccttat	tcctttttt	gcggcatttt	gccttcctgt	3840
ttttgctcac	ccagaaacgc	tgggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	3900
agtgggttac	atcgaaactg	atctcaacag	cggttaagatc	cttgagagtt	ttcgccccga	3960
agaacgtttt	ccaatgatga	gcacttttaa	agttctgcta	tgtggcgcg	tattatcccg	4020
tgttgacgcc	gggcaagagc	aactcggtcg	cgcatacac	tattctcaga	atgacttgg	4080
tgagtactca	ccagtcacag	aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	4140
cagtgtctgc	ataaccatga	gtgataaac	tgcggccaac	ttacttctga	caacgatcgg	4200
aggaccgaag	gagctaacgc	cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	4260
tcgttgggaa	ccggagctga	atgaagccat	acaaaacgac	gagcgtgaca	ccacgatgcc	4320
tcgagcaatg	gcaacaacgt	tcgcgaaact	attaactggc	gaactactta	ctctagcttc	4380
ccggcaacaa	ttaatagact	ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgctc	4440
ggcccttcg	gctggctggt	ttattgctga	taaatctgga	gccggtgagc	gtgggtctcg	4500
cggtatcatt	gcagcactgg	ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	4560
gacggggagt	caggcaacta	tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	4620
actgattaag	cattggtaac	tgtcagacca	agtttactca	tatatacttt	agattgattt	4680
aaaacttcat	ttttaattta	aaaggatcta	ggtgaagatc	ctttttgata	atctcatgac	4740
caaaatccct	taacgtgagt	tttcgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	4800
aggatcttct	tgagatcctt	tttttctgcy	cgtaactctgc	tgcttgcaaa	caaaaaaac	4860
accgctacca	gcggtggttt	gtttgccgga	tcaagagcta	ccaactcttt	ttccgaaggt	4920
aactggcttc	agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagg	4980

-continued

ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tctgtttacc	5040
agtggtctgt gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt	5100
accggataag ggcgagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttgga	5160
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct	5220
tcccgaaagg agaaaggcgg acaggtatcc ggtaagcggc agggctcgaa caggagagcg	5280
cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca	5340
cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa	5400
cggcagcaac ggggcctttt tacggttcct ggcccttttg tggccttttg ctcacatgtt	5460
ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga	5520
taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcgggaaga	5580
gcgcccaata cgcaaacgcg ctctcccgcg gcgttgggcg attcattaat gcagctggca	5640
cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct	5700
cactcattag gcaccccagg ctttacactt tatgcttcg gctcgatgt tgtgtggaat	5760
tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg ccaagctcta	5820
gctagaggtc gagtccctcc ccagcaggca gaagtatgca aagcatgcat ctcaattagt	5880
cagcaaccat agtcccgcgc ctaactccgc ccaccccgc cctaactccg cccagttccg	5940
cccattctcc gcccctggcg tgactaattt tttttattta tgcagaggcc gaggcgcct	6000
cggcctctga gctattccag aagtagtgag gaggtttttt tggaggccta ggcttttgca	6060
aaaagctttg caaagatgga taaagtttta aacagagagg aatctttgca gctaattggc	6120
cttctaggtc ttgaaaggag ctgcaccaat tctcatgttt gacagcttat catcgagat	6180
ccgggcaacg ttgttgccat tgctgcaggc gcagaactgg taggtatgga agatctatac	6240
attgaatcaa tattggcaat tagccatatt agtcattggt tatatagcat aaatcaatat	6300
tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc	6360
atgtccaata tgaccgcat g	6381

<210> SEQ ID NO 2
 <211> LENGTH: 6212
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic construct: pHybE-empty
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (1320)..(1391)
 <223> OTHER INFORMATION: n is a, c, g, or t
 <400> SEQUENCE: 2

agctttgcaa agatggataa agttttaaac agagaggaat ctttgcagct aatggacctt	60
ctaggctctt aaaggagtgg gaattggctc cgggtgcccg cagtgggcag agcgcacatc	120
gcccacagtc cccgagaagt tggggggagg ggtcgccaat tgaaccggtg cctagagaag	180
gtggcgccgg gtaaaactgg aaagtgatgt cgtgtactgg ctccgccttt tccccagggg	240
tgggggagaa ccgtatataa gtgcagtagt cgccgtgaac gttctttttc gcaacggggt	300
tgccgcgaga acacaggtaa gtgccgtgtg tggttcccgc gggcctggcc tctttacggg	360
ttatggccct tgcgtgcctt gaattacttc cacctggctg cagtacgtga ttcttgatcc	420
cgagcttcgg gttggaagtg ggtgggagag ttcgaggcct tgcgcttaag gagccccttc	480

-continued

gcctcgtgct tgagttgagg cctggcctgg gcgctggggc cgcgcgtgc gaatctggtg	540
gcaccttcgc gcctgtctcg ctgctttcga taagtctcta gccatttaaa atttttgatg	600
acctgctgcg acgctttttt tctggcaaga tagtcttcta aatgcgggcc aagatctgca	660
cactggtatt tcggtttttg gggccgcggg cggcgacggg gcccgctcgt cccagcgcac	720
atgttcggcg aggcggggcc tgcgagcgcg gccaccgaga atcgacggg ggtagtctca	780
agctggcgg cctgctctgg tgcctggcct cgcgcgcgg tgtatcgccc cgccctgggc	840
ggcaaggctg gcccggtcgg caccagtgc gtgagcgga agatggcgc tccccggccc	900
tgctgcagg agctcaaaat ggaggacgc gcgctcgga gagcgggcg gtgagtcacc	960
cacacaaagg aaaagggcct ttccgtctc agccgtcgt tcatgtgact ccacggagta	1020
ccgggcgcg tccaggcacc tcgattagt ctcgagcttt tggagtacgt cgtctttagg	1080
ttggggggag gggttttatg cgatggagtt tccccacact gagggggtg agactgaagt	1140
taggccagct tggcacttga tgtaattct cttggaattt gccctttttg agtttggatc	1200
ttggttcatt ctcaagcctc agacagtgg tcaaagtctt tttcttccat ttcagtgctc	1260
gtgaggaatt ctctagagat cctcgcact cgagatccat tgtccccgg cgaccatgn	1320
nnnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn	1380
nnnnnnnnnn ntgagcgccc gctcgaggc ggcaaggcgg gatccccga cctcgacctc	1440
tggtcaataa aggaaattta ttttcattgc aatagtgtgt tggaaatttt tgtgtctctc	1500
actcggaagg acatatggga gggcaaatca tttggtcgag atccctcgga gatctctagc	1560
tagaggatcg atccccgccc cggacgaact aaacctgact acgacatctc tgcctctctc	1620
tcgcggggca gtgcattgaa tcccttcagt tgggttggtac aacttgccaa ctgggcccctg	1680
ttccacatgt gacacggggg gggaccaaac acaaaggggt tctctgactg tagttgacat	1740
ccttataaat ggatgtgcac atttgccaac actgagtggc tttcatcctg gacgagactt	1800
tgagctctgt ggactgcaac acaacattgc ctttatgtgt aactcttggc tgaagctctt	1860
acaccaatgc tgggggacat gtacctcca ggggcccagg aagactacgg gaggtacac	1920
caacgtcaat cagagggggc tgtgtagcta ccgataagcg gacctcaag agggcattag	1980
caatagtgtt tataaggccc ccttggtaac cctaaacggg tagcatatgc tccccgggta	2040
gtagtatata ctatccagac taaccctaata tcaatagcat atgttaccca acgggaagca	2100
tatgctatcg aattagggtt agtaaaagg tccaaaggaa cagcgatctc tccccccca	2160
tgagctgtca cggttttatt tacatgggtt caggattcca cgagggtagt gaaccatttt	2220
agtcacaagg gcagtggctg aagatcaagg agcgggcagt gaactctcct gaatcttcgc	2280
ctgctctctc attctcctc gtttagctaa tagaataact gctgagttgt gaacagtaag	2340
gtgtatgtga ggtgctcgaa aacaagggtt caggtgacgc cccagaata aaatttgga	2400
ggggggttca gtggtggcat tgtgctatga caccaatata accctcaca accccttggg	2460
caataaatac tagtgtagga atgaacatt ctgaatatct ttaacaatag aaatccatgg	2520
ggtggggaca agccgtaaag actggatgc catctcacac gaatttatgg ctatgggcaa	2580
cacataatcc tagtgcaata tgatactggg gttattaaga tgtgtcccag gcagggacca	2640
agacagtgta accatgttgt tacactctat ttgtaacaag gggaaagaga gtggacgccc	2700
acagcagcgg actccactgg ttgtctctaa ccccccgaa aattaaacgg ggtcccacgc	2760

-continued

caatggggcc cataaacaac gacaagtggc cactcttttt tttgaaattg tggagtgggg	2820
gcacgcgtca gccccacac gccgcctgc ggttttgac tgtaaaataa ggggtgaata	2880
acttgctga ttgtaacccc gctaaccact gcggtcaaac cacttgccca caaaaccact	2940
aatggcacc cggggaatac ctgcataagt aggtgggagg gccagatag gggcgcgatt	3000
gctgcgatct ggaggacaaa ttacacacac ttgcgcctga gcgccaagca cagggttggt	3060
ggctctcata ttcacgaggt cgctgagagc acggtgggct aatgttgcca tgggtagcat	3120
atactaccca aatatctgga tagcatatgc taccctaacc tatactctggg tagcataggc	3180
taccctaacc tatactctggg tagcatatgc taccctaacc tatactctggg tagtatatgc	3240
taccctaacc tatactctggg tagcataggc taccctaacc tatactctggg tagcatatgc	3300
taccctaacc tatactctggg tagtatatgc taccctaacc tggatccggg tagcatatgc	3360
taccctaacc gagattaggg tagtatatgc taccctaacc tatactctggg tagcatatac	3420
tacccaata tctggatagc atatgctatc ctaacttata tctgggtagc atatgctatc	3480
ctaacttata tctgggtagc ataggctatc ctaacttata tctgggtagc atatgctatc	3540
ctaacttata tctgggtagc atatgctatc ctaacttata tctgggtagc ataggctatc	3600
ctaacttata tctgggtagc atatgctatc ctaacttata tctgggtagc atatgctatc	3660
ctaacttgta tccgggtagc atatgctatc ctcatgataa gctgtcaaac atgagaattt	3720
tcttgaagac gaaaggcct cgtgatacgc ctatttttat aggttaatgt catgataata	3780
atggtttctt agacgtcagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt	3840
ttatttttct aaatacatc aaatatgtat ccgctcatga gacaataacc ctgataaatg	3900
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgccttatt	3960
cccttttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta	4020
aaagatgctg aagatcagtt ggggtgcacg gtgggttaca tcgaactgga tctcaacagc	4080
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cactttttaa	4140
gttctgctat gtggcgcggt attatcccgt gttgacgcgc ggcaagagca actcggtcgc	4200
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt	4260
acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataaacct	4320
gcggccaact tacttctgac aacgatcgga ggaccgaagg agtaaccgc ttttttgac	4380
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata	4440
ccaaacgagc agcgtgacac cacgatgcct gcagcaatgg caacaacgtt gcgcaacta	4500
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg	4560
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat	4620
aaatctggag cgggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt	4680
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcactat ggatgaacga	4740
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa	4800
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag	4860
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttcac	4920
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc	4980
gtaactctgt gcttgcaaac aaaaaacca ccgctaccag cggtggtttg tttgccgat	5040
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat	5100

-continued

```

actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct 5160
acatacctcg ctctgctaata cctgttacca gtggtgctg cagtggcga taagtcgtgt 5220
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg 5280
gggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaact gagataccta 5340
cagcgtgagc tatgagaaa cgcacgcctt cccgaaggga gaaaggcgga caggatatccg 5400
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg 5460
tatctttata gtctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc 5520
tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg 5580
gccttttctt ggccttttgc tcacatgttc tttcctgcgt tateccctga ttctgtggat 5640
aacgctatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc 5700
agcaggtcag tgagcgagga agcgggaagag cgcccaatac gcaaaccgcc tctcccgcg 5760
cgttgccga ttcatatgt cagctggcac gacagggttc ccgactggaa agcgggcagt 5820
gagcgcaacg caattaatgt gagttagctc actcattagg caccgcaggc ttacacttt 5880
atgcttcctg ctctgtatgt gtgtggaatt gtgagcgat aacaatttca cacaggaaac 5940
agctatgacc atgattacgc caagctctag ctagaggctg agtccctccc cagcaggcag 6000
aagtatgcaa agcatgcac tcaattagtc agcaaccata gtcccgcccc taactccgcc 6060
catcccgccc ctaactccgc ccagttccgc ccattctcgc ccccatggct gactaatttt 6120
ttttatttat gcagaggcgc aggcgcgcct ggccctctgag ctattccaga agtagtgagg 6180
aggctttttt ggaggcctag gcttttgcaa aa 6212

```

<210> SEQ ID NO 3

<211> LENGTH: 7498

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP180 ; pHybE-mCk V1

<400> SEQUENCE: 3

```

agtgggaatt ggctccgggt cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtgct gcaattgaac cgggtccctag agaagggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctcgc cctttttccc gaggggtggg gagaaaccgt 180
tataagtga gtagtcgcgc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct 300
gccttgaaat acttccacct ggctgcagta cgtgattctt gatcccgagc ttcggggttg 360
aagtgggtgg gagagttcga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgct gggggccgcg cgtgcgaatc tggtggcacc ttcgcgcctg 480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gcggggcgcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctgggtgct ggccctgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggctggcccc 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc 840

```

-continued

aaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcacccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcatg	tgactccacg	gagtaccggg	cgcggtccag	960
gcacctcgat	tagttctcga	gcttttggag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggg	tcattctcaa	1140
gcctcagaca	gtgggtcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgagc	atttaaatgc	cggggcgcac	catggacatg	cgcgtgcccc	1260
cccagctgct	gggcctgctg	ctgctgtggg	tccccggctc	gcgatgcgca	tggatagccg	1320
aaagggatgc	tgaatttgag	aacgaaaagc	tgcgccggga	ggttgaagaa	ctgcggcagg	1380
ccagcgaggc	agatctccag	ccaggaacta	ttgagtacga	acgccatcga	cttacgcgtg	1440
cgcaggccga	cgcacaggaa	ctgaagaatg	ccagagactc	cgtcgaagtg	gtggaaaccg	1500
cattctgtac	tttcgtgctg	tcgcggatcg	caggtgaaat	tgccagtatt	ctcgacgggc	1560
tccccctgtc	ggtgcagcgg	cgttttccgg	aactggaaaa	ccgacatggt	gatttcctga	1620
aacgggatat	catcaaaagc	atgaacaaaag	cagccgcgct	ggatgaactg	ataccggggg	1680
tgctgagtga	atatatcgaa	cagtcagggt	aacaggctgc	ggcattttgt	ccgcgcgggg	1740
cttcgctcac	tgttcaggcc	ggagccacag	accgccgttg	aatgggcgga	tgctaattac	1800
tatctcccga	aagaatccgc	ataccaggaa	gggcgctggg	aaacactgcc	ctttcagcgg	1860
gccatcatga	atgcgatggg	cagcgactac	atccgtgagg	tgaatgtggg	gaagtctgcc	1920
cgtgtcgggt	attccaaaat	gctgctgggt	gtttatgcct	actttataga	gcataagcag	1980
cgcacacccc	ttatctgggt	gccgacggat	ggtgatgccg	agaactttat	gaaaacccac	2040
gttgagccga	ctattcgtga	tattccgtcg	ctgctggcgc	tggccccgtg	gtatggcaaa	2100
aagcacccgg	ataacacgct	caccatgaag	cgtttcacta	atgggcgtgg	cttctggtgc	2160
ctgggcggta	aagcgcaaaa	aaactaccgt	gaaaagtcgg	tggatgtggc	gggttatgat	2220
gaacttgctg	cttttgatga	tgatattgaa	caggaaggct	ctccgacgtt	cctgggtgac	2280
aagcgtgat	gctgcaccaa	ctgtatccat	cttcccacca	tccagtgagc	agttaacatc	2340
tggaggtgcc	tcagtcgtgt	gottcttgaa	caacttctac	cccaaagaca	tcaatgtcaa	2400
gtggaagatt	gatggcagtg	aacgacaaaa	tggcgtcctg	aacagttgga	ctgatcagga	2460
cagcaaaagc	agcacctaca	gcatgagcag	caccctcacg	ttgaccaagg	acgagtatga	2520
acgacataac	agctatacct	gtgaggccac	tcacaagaca	tcaacttcac	ccattgtcaa	2580
gagcttcaac	aggaatgagt	gttaagcggc	cgtcagaggc	cggcaaggcc	ggatcccccg	2640
acctcgacct	ctggctaata	aaggaaaatt	attttcattg	caatagtgtg	ttggaatttt	2700
ttgtgtctct	cactcggaag	gacatatggg	agggcaaatc	atttggtcga	gatccctcgg	2760
agatctctag	ctagaggatc	gatccccgcc	cggacgaac	taaacctgac	tacgacatct	2820
ctgccccctc	ttcgcggggc	agtgcattga	atcccttcag	ttggttggtg	caacttgcca	2880
actgggccct	gttccacatg	tgacacgggg	ggggacccaa	cacaaagggg	ttctctgact	2940
gtagttgaca	tccttataaa	tggatgtgca	catttgccaa	cactgagtgg	ctttcatcct	3000
ggagcagact	ttgcagtctg	tggactgcaa	cacaacattg	cctttatgtg	taactcttgg	3060
ctgaagctct	tacaccaatg	ctgggggaca	tgtacctccc	aggggcccag	gaagactacg	3120
ggaggctaca	ccaacgtcaa	tcagaggggc	ctgtgtagct	accgataagc	ggaccctcaa	3180

-continued

gagggcatta gcaatagtgt ttataaggcc cccttggttaa ccctaaacgg gtagcatatg	3240
cttcccgggt agtagtatat actatccaga ctaaccctaa ttcaatagca tatgttacct	3300
aacgggaagc atatgctatc gaattaggggt tagtaaaagg gtccaaagga acagcgatat	3360
ctcccacccc atgagctgtc acgggttttat ttacatgggg tcaggattcc acgagggtag	3420
tgaaccattt tagtcacaag ggcagtggtc gaagatcaag gagcgggcag tgaactctcc	3480
tgaatcttcg cctgcttctt cattctcctt cgttttagcta atagaataac tctgagttg	3540
tgaacagtaa ggtgtatgtg aggtgctcga aaacaagggt tcagggtgacg ccccagaat	3600
aaaatttga cgggggggtc agtggtggca ttgtgctatg acaccaatat aacctcaca	3660
aacctcttg gcaataaata ctagtgtagg aatgaaacat tctgaatctc tttaacaata	3720
gaaatccatg ggggtgggac aagccgtaaa gactggatgt ccacttcaca cgaatttatg	3780
gctatgggca acacataatc ctagtgcaat atgatactgg gggtattaag atgtgtccca	3840
ggcagggacc aagacagggt aacctgttg ttacactcta tttgtaacaa ggggaaagag	3900
agtggagccg gacagcagcg gactccactg gttgtctcta acaccccgga aaattaaacg	3960
gggctccacg ccaatggggc ccataaacaa agacaagtgg ccactctttt ttttgaatt	4020
gtggagtggg ggcacgcgtc agccccaca cgccgcctg cggttttga ctgtaaaata	4080
aggggtgaat aacttggtg attgtaaccc cgctaaccac tgcgggtcaaa ccacttgccc	4140
acaaaaccac taatggcacc ccggggaata cctgcataag taggtgggcg ggccaagata	4200
ggggcgcgat tgctgcgacg tggaggacaa attacacaca cttgcgcctg agcgccaagc	4260
acagggttgt tggctctcat attcacgagg tcgctgagag cacgggtggc taatgttgcc	4320
atgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	4380
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	4440
gtagtatatg ctatcctaata ttatatctgg gtagcatagg ctatcctaata ctatatctgg	4500
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ctgtatccg	4560
gtagcatatg ctatcctaata agagattagg gtagtatatg ctatcctaata ttatatctgg	4620
gtagcatata ctacccaata atctggatag catatgctat cctaactctat atctgggtag	4680
catatgctat cctaactctat atctgggtag cataggctat cctaactctat atctgggtag	4740
catatgctat cctaactctat atctgggtag tatatgctat cctaatttat atctgggtag	4800
cataggctat cctaactctat atctgggtag catatgctat cctaactctat atctgggtag	4860
tatatgctat cctaactctat atccgggtag catatgctat cctcatgata agctgtcaaa	4920
catgagaatt ttcttgaaga cgaaagggcc tcgtgatacg cctattttta taggttaatg	4980
tcagtataat aatggtttct tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa	5040
ccccatttg ttatttttcc taaatacatt caaatatgta tccgctcatg agacaataac	5100
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gattattcaa catttccgtg	5160
tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc	5220
tgggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaaactg	5280
atctcaacag cggtgaagtc cttgagagtt ttcccccga agaacgtttt ccaatgatga	5340
gcacttttaa agttctgcta tgtggcgcgg tattatcccg tgttgacgcc gggcaagagc	5400
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag	5460

-continued

```

aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtctgcc ataaccatga 5520
gtgataaacac tgcggccaac ttactttctga caacgatcgg aggaccgaag gagctaaccg 5580
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga 5640
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgcagcaatg gcaacaacgt 5700
tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 5760
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttcg gctggctggg 5820
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 5880
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 5940
tggatgaacg aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac 6000
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 6060
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 6120
ttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 6180
tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt 6240
gtttgcgga tcaagagcta ccaactcttt ttccgaaggc aactggcttc agcagagcgc 6300
agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 6360
tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg 6420
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg 6480
cgggctgaac ggggggttcg tgcacacagc ccagcttggg gcgaacgacc tacaccgaac 6540
tgagatacct acagcgtgag ctatgagaaa gcgccacgct tccgaaggg agaaaggcgg 6600
acaggtatcc ggtaacgggc agggctcgaa caggagagcg cagcaggag cttccagggg 6660
gaaacgctg gtatctttat agtcctgtcg gggttcgcca cctctgactt gagcgtcgat 6720
ttttgtgatg ctgcctcagg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 6780
tacggttcct ggcccttttg tggccttttg ctcatatgtt ctttctgctg ttatccctg 6840
attctgtgga taaccgtatt accgccttg agtgagctga taccgctcgc cgcagccgaa 6900
cgaccgagcg cagcagtgca gtgagcagg aagcggaaga gcgcccaata cgcaaaccgc 6960
ctctccccgc gcgttgccg attcattaat gcagctggca cgacaggttt cccgactgga 7020
aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcacccag 7080
ctttacactt tatgcttcg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc 7140
acacaggaaa cagctatgac catgattacg ccaagctcta gctagaggtc ggtccctcc 7200
ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccat agtcccgcc 7260
ctaactccgc ccatccgcc ctaactccg ccagttccg ccattctcc gcccctggc 7320
tgactaat tttttat tgcagaggcc gaggccgct cggcctctga gctattccag 7380
aagtagtgag gaggtctttt tggaggccta ggcttttgca aaaagctttg caaagatgga 7440
taaagt ttaaacagagagg aatctttgca gctaattgac cttctaggtc ttgaaagg 7498

```

<210> SEQ ID NO 4

<211> LENGTH: 7498

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP193 ; pHybE-mCk V2

<400> SEQUENCE: 4

-continued

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgccc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct	300
gccttgaaat acttccacct ggctgcagta cgtgattctt gateccgagc ttcgggttg	360
aagtgggtgg gagagtcca ggccttcgca ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tggtggaacc ttcgcgcctg	480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcggggcggc acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggctcgca gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgctt ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggctggccc	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc	840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag tcaaccacac aaaggaaaag	900
ggcctttccg tcctcagccg tcgcttcagt tgactccacg gactaccggg cgcgctccag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactcgatg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttct tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgaga tccattgtgc ccgggcgcac catggacatg cgcgtgccg	1260
cccagctgct gggcctgctg ctgctgtggt tccccggctc gcgatgcgca tggtatgccg	1320
aaagggatgc tgaatttag aacgaaaagc tgcgcgggga ggttgaagaa ctgcggcagg	1380
ccagcgaggc agatctccag ccaggaaacta ttgagtacga acgccatcga cttacgcgtg	1440
cgcagggcca cgcacaggaa ctgaagaatg ccagagactc cgtgaagtg gtggaaaccg	1500
cattctgtac tttcgtgctg tcgctggatc caggtgaaat tgccagtatt ctgcacgggc	1560
tccccctgtc ggtgcagcgg cgttttccgg aactggaaaa ccgacatgtt gatttcctga	1620
aacgggatat catcaagcc atgaacaaag cagccgcgct ggatgaactg ataccggggt	1680
tgctgagtga atatatcgaa cagtcagggt aacaggctgc ggcatcttgt ccgcgcggg	1740
cttcgctcac tgttcaggcc ggagccacag accgcgctt aatggggcga tgetaattac	1800
tatctccga aagaatccgc ataccaggaa gggcgctggg aaacactgcc ctttcagcgg	1860
gccatcatga atgcgatgg cagcgactac atccgtgagg tgaatgtggt gaagtctgcc	1920
cgtgtcgggt attccaaaat gctgctgggt gtttatgcct actttataga gcataagcag	1980
cgaacacccc ttatctggtt gccgaaggat ggtgatgcg agaactttat gaaaaccac	2040
gttgagccga ctattcgtga tattccgtcg ctgctggcgc tggcccgtg gtatggcaaa	2100
aagcaccggg ataacacgct caccatgaag cgtttcacta atgggcgtgg cttctggtgc	2160
ctgggcggta aagcggcaaa aaactaccgt gaaaagtcgg tggatgtggc gggttatgat	2220
gaacttgctg cttttgatga tgatattgaa caggaaggct ctccgacgtt cctgggtgac	2280

-continued

aagcgtgat gctgcacaa ctgtatccat cttcccacca tccagtgagc agttaacatc	2340
tggagggtgcc tcagtcgtgt gcttcttgaa caacttctac cccaaagaca tcaatgtcaa	2400
gtggaagatt gatggcagtg aacgacaaaa tggcgtctctg aacagttgga ctgatcagga	2460
cagcaaagac agcacctaca gcatgagcag caccctcacg ttgaccaagg acgagtatga	2520
acgacataac agctatacct gtgaggccac tcacaagaca tcaacttcac ccattgtcaa	2580
gagcttcaac aggaatgagt gttaagcggc cgctcgaggc cggcaaggcc ggatcccccg	2640
acctcgacct ctggctaata aaggaaattt attttcattg caatagtgtg ttggaatttt	2700
ttgtgtctct cactcggaag gacatatggg agggcaaatc atttggtcga gatccctcgg	2760
agatctctag ctaggagatc gatccccgcc cgggacgaac taaacctgac tacgacatct	2820
ctgccccctt tcgcgggggc agtgcatgta atcccttcag ttggttggtg caacttgcca	2880
actgggccct gttccacatg tgacacgggg ggggacaaaa cacaaggggg ttctctgact	2940
gtagttgaca tccttataaa tggatgtgca catttgccaa cactgagtgg ctttcatcct	3000
ggagcagact ttgcagtctg tggactgcaa cacaacattg cctttatgtg taactcttgg	3060
ctgaagctct tacaccaatg ctggggggaca tgtacctccc agggggcccag gaagactacg	3120
ggaggctaca ccaacgtcaa tcagaggggc ctgtgtagct accgataagc ggaccctcaa	3180
gagggcatta gcaatagtgt ttataaggcc cccttgtaa ccctaaacgg gtagcatatg	3240
cttcccgggt agtagtatat actatccaga ctaaccctaa ttcaatagca tatgttacct	3300
aacgggaagc atatgctatc gaattagggt tagtaaaagg gtcctaagga acagcgatat	3360
ctccaccccc atgagctgtc acgggtttat ttacatgggg tcaggattcc acgagggtag	3420
tgaaccattt tagtcacaag ggcagtggct gaagatcaag gagcgggcag tgaactctcc	3480
tgaatcttcg cctgcttctt cattctcctt cgttttagcta atagaataac tgctgagttg	3540
tgaacagtaa ggtgtatgtg aggtgctcga aaacaagggt tcaggtgacg cccccagaat	3600
aaaaatttga cgggggggtc agtgggtggc ttgtgctatg acaccaatat aacctcaca	3660
aacctcttgg gcaataaata ctagtgtagg aatgaaacat tctgaatatc tttaacaata	3720
gaaatccatg ggggtgggac aagccgtaaa gactggatgt ccacttcaca cgaatttatg	3780
gctatgggca acacataatc ctagtgcaat atgatactgg ggttattaag atgtgtccca	3840
ggcagggacc aagacagggt aacctgttg ttacactcta tttgtaacaa ggggaaagag	3900
agtggaacgc gacagcagcg gactccactg gttgtctcta acacccccga aaattaaacg	3960
gggctccacg ccaatggggc ccataaacia agacaagtgg ccactctttt ttttgaaatt	4020
gtggagtggg ggcacgcgtc agccccaca gcgcgcctg cggtttttga ctgtaaaata	4080
aggggtgaat aacttggtg attgtaaccc cgctaaccac tgcggtcaaa ccacttgccc	4140
acaaaaccac taatggcacc ccggggaata cctgcataag taggtgggcg ggccaagata	4200
ggggcgcgat tgctgcgacg tggaggacaa attacacaca cttgcgcctg agcgccaagc	4260
acagggttgt tggctctcat attcacgagg tcgctgagag cacgggtggc taatgttgcc	4320
atgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	4380
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	4440
gtagtatatg ctatcctaata ctatatctgg gtagcatagg ctatcctaata ctatatctgg	4500
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ctgtatccgg	4560
gtagcatatg ctatcctaata agagattagg gtagtatatg ctatcctaata ttatatctgg	4620

-continued

gtagcatata ctacccaaat atctggatag catatgctat cctaacttat atctgggtag	4680
catatgctat cctaacttat atctgggtag cataggctat cctaacttat atctgggtag	4740
catatgctat cctaacttat atctgggtag tatatgctat cctaatttat atctgggtag	4800
cataggctat cctaacttat atctgggtag catatgctat cctaacttat atctgggtag	4860
tatatgctat cctaacttgt atccgggtag catatgctat cctcatgata agctgtcaaa	4920
catgagaatt ttcttgaaga cgaaagggcc tcgtgatacg cctattttta taggttaatg	4980
tcagtataat aatggtttct tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa	5040
ccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac	5100
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catctccgtg	5160
tcgccttat tccctttttt gcggcatttt gccttccgtg ttttgctcac ccagaaacgc	5220
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg	5280
atctcaacag cggtaaagtc cttgagagtt ttcgccccga agaacgtttt ccaatgatga	5340
gcacttttaa agttctgcta tgtggcgcgg tattatcccg tgttgacgcc gggcaagagc	5400
aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag	5460
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtctgc ataaccatga	5520
gtgataaacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg	5580
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga	5640
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgcagcaatg gcaacaacgt	5700
tgcgcaaaact attaaactgc gaactactta ctctagcttc ccggcaacaa ttaatagact	5760
ggatggaggc ggataaagt gcaggaccac ttctgcgctc ggcccttcg gctggctggt	5820
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg	5880
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta	5940
tggatgaacg aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac	6000
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta	6060
aaaggatcta ggtgaagtc ctttttgata atctcatgac caaaatccct taacgtgagt	6120
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt	6180
ttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt	6240
gtttgccgga tcaagagcta ccaactcttt ttccgaaggc aactggcttc agcagagcgc	6300
agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg	6360
tagcaccgcc tacatacctc gctctgctaa tctgtttacc agtggctgct gccagtggcg	6420
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag ggcagcggg	6480
cgggctgaac ggggggttgc tgcacacagc ccagcttggg gcgaacgacc tacaccgaac	6540
tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg	6600
acaggatcc ggtaagcgcg agggctcgaa caggagagcg cagcaggag cttccagggg	6660
gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat	6720
ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt	6780
tacgggtcct ggccctttgc tggccttttg ctacatggt ctttctgctg ttatccctg	6840
attctgtgga taacgttatt accgcctttg agtgagctga taccgctcgc cgcagccgaa	6900

-continued

cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc	6960
ctctccccgc gcgttgccg attcattaat gcagctggca cgacaggttt cccgactgga	7020
aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg	7080
ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc	7140
acacaggaaa cagctatgac catgattacg ccaagcteta gctagaggtc gagtccctcc	7200
ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccat agtcccgcgc	7260
ctaactccgc ccatccgcgc cctaactccg cccagttccg cccattctcc gccccatggc	7320
tgactaatTT tttttattta tgcagaggcc gaggcgcct cggcctctga gctattccag	7380
aagtagtgag gaggcctttt tggaggccta ggcttttgca aaaagcttg caaagatgga	7440
taaagtttta aacagagagg aatctttgca gctaattggac cttctaggtc ttgaaagg	7498

<210> SEQ ID NO 5

<211> LENGTH: 8144

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP176 ; pHybE-mCgl V1

<400> SEQUENCE: 5

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggaggggctg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggtccg cctttttccc gaggggtggg gagaaccgta	180
tataagtgca gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgct	300
gccttgaaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360
aagtgggtgg gagagttcga ggccttgccc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcgcgcg cgtgcgaatc tggtggaacc ttcgcgcctg	480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatggt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctgggtgct ggccctgcgc cgcctgtgat cgcgccgcgc tgggcggcaa ggctggcccg	780
gtcggcacca gttgcgtgag cggaaaagatg gccgcttccc ggccctgctg caggagctc	840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcaaccacac aaaggaaaag	900
ggcctttccg tcctcagccg tcgcttcagt tgactccacg gactaccggg cgcctgccag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggt tcattctcaa	1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgagc atttaaatgc cggggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaagg gtccagtgcg catggtatgc cgaaagggat	1320
gctgaaattg agaacgaaaa gctgcgcggg gaggttgaag aactgcggca ggccagcgag	1380
gcagatctcc agccaggaaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440

-continued

gacgcacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tccgtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttccct gaaacgggat	1620
atcatcaaa ccatgaacaa agcagccgcg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgc ggcttcgctc	1740
actgttcagg ccggagccac agaccgccgt tgaatgggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aagggcgcgt ggaaactcgc ccctttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggatgatgc cgagaacttt atgaaaacc acgttgagcc	2040
gactattcgt gatattccgt cgctgctggc gctggccccg tggatatgca aaaagcaccc	2100
ggataaacag ctcaccatga agcgtttcac taatgggcgt ggcttctggg gctggggcgg	2160
taaagcggca aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgcta	2280
aaacgacacc cccatctgtc tatccactgg cccctggatc tgctgcccaa actaactcca	2340
tggtagacct gggatgcctg gtcaaggcct atttccctga gccagtgaac gtgacctgga	2400
actctggatc cctgtccagc ggtgtgcaca ccttccagc tgtcctgcag tctgacctct	2460
acactctgag cagctcagtg actgtccctc ccagcacctg gccagcgag accgtcacct	2520
gcaacgttgc ccacccggcc agcagcacca aggtggacaa gaaaattgtg cccagggatt	2580
gtggttgtaa gccttgcata tgtacagtcc cagaagtatc atctgtcttc atcttcccc	2640
caaagcccaa ggatgtgtc accattactc tgactcctaa ggtcacgtgt gttgtggtag	2700
acatcagcaa ggatgatccc gaggtccagt tcagctggtt ttagatgatg gtggaggtgc	2760
acacagctca gacgcaaccc cgggaggagc agttcaacag cactttccgc tcagtcagtg	2820
aacttcccat catgcaccag gactggctca atggcaagg gttcaaatgc agggcaaca	2880
gtgcagcttt cctgtccccc atcgagaaaa ccactctcaa aaccaaaggc agaccgaagg	2940
ctccacaggt gtacaccatt ccacctccca aggagcagat ggccaaggat aaagtcagtc	3000
tgacctgcat gataacagac ttcttccctg aagacattac tgtggagtgg cagtggaatg	3060
ggcagccagc ggagaactac aagaacactc agcccatcat ggacacagat ggctcttact	3120
tcgtctacag caagctcaat gtgcagaaga gcaactggga ggcaggaaat actttcacct	3180
gctctgtgtt acatgagggc ctgcacaacc accatactga gaagagcctc tcccactctc	3240
ctggtaaata agcggccgct cgaggccggc aaggccggat ccccgacct cgacctctgg	3300
ctaataaagg aaattttatt tcattgcaat agtgtgttgg aattttttgt gtctctcact	3360
cggagggaca tatgggaggg caaatcattt ggtcgagatc cctcgagat ctctagctag	3420
aggatcgatc cccgccccgg acgaactaaa cctgactacg acatctctgc cccttcttcg	3480
cggggcagtg catgtaatcc cttcagttgg ttggtacaac tgccaactg ggccctgttc	3540
cacatgtgac acgggggggg accaaacaca aaggggttct ctgactgtag ttgacatcct	3600
tataaatgga tgtgcacatt tgccaacact gagtggcttt catcctggag cagactttgc	3660
agtctgtgga ctgcaacaca acattgcctt tatgtgtaac tcttggtgga agctcttaca	3720

-continued

ccaatgctgg	gggacatgta	cctcccagg	gcccaggaag	actacgggag	gctacaccaa	3780
cgtaaatcag	aggggcctgt	gtagctaccg	ataagcggac	cctcaagagg	gcattagcaa	3840
tagtgtttat	aaggccccct	tgtaaacct	aaacgggtag	catatgcttc	ccgggtagta	3900
gtatatacta	tccagactaa	ccctaattca	atagcatatg	ttaccaacg	ggaagcatat	3960
gctatcgaat	tagggtagt	aaaagggtcc	taaggaacag	cgatatctcc	caccccatga	4020
gctgtcacgg	ttttatttac	atggggtcag	gattccacga	gggtagtgaa	ccattttagt	4080
cacaagggca	gtggctgaag	atcaaggagc	gggcagtga	ctctcctgaa	tcttcgcctg	4140
cttcttcatt	ctccttcgtt	tagctaatag	aataactgct	gagttgtgaa	cagtaagggt	4200
tatgtgaggt	gctcgaaaac	aaggtttcag	gtgacgcccc	cagaataaaa	tttgacggg	4260
gggttcagt	gtggcattgt	gctatgacac	caatataacc	ctcaciaaac	ccttgggcaa	4320
taaaactag	tgtaggaatg	aaacattctg	aatatcttta	acaatagaaa	tccatggggt	4380
ggggacaagc	cgtaagact	ggatgtccat	ctcacacgaa	tttatggcta	tgggcaacac	4440
ataatcctag	tgcaatatga	tactgggggt	attaagatgt	gtcccaggca	gggaccaaga	4500
cagggtgaacc	atgtttgtac	actctatttg	taacaagggg	aaagagagtg	gacgccgaca	4560
gcagcggact	ccactgggtg	tctctaacac	ccccgaaaat	taaacggggc	tccacgccaa	4620
tggggcccat	aaacaaagac	aagtggccac	tctttttttt	gaaattgtgg	agtgggggca	4680
cgcgtcagcc	cccacacgcc	gccctgcggg	tttggactgt	aaaataaggg	tgtataaact	4740
tgggtgattg	taaccccgct	aaccactgcg	gtcaaaccac	ttgccacaaa	aaccactaat	4800
ggcaccccg	ggaatacctg	cataagtagg	tgggcggggc	aagatagggg	cgcgattgct	4860
gcgatctgga	ggacaaatta	cacacacttg	cgctgagcgg	ccaagcacag	ggttgttggt	4920
cctcatattc	acgaggtcgc	tgagagcacg	gtgggcta	gttgccatgg	gtagcatata	4980
ctacccaaat	atctggatag	catatgctat	cctaactctat	atctgggtag	cataggctat	5040
cctaactctat	atctgggtag	catatgctat	cctaactctat	atctgggtag	tatatgctat	5100
cctaatttat	atctgggtag	cataggctat	cctaactctat	atctgggtag	catatgctat	5160
cctaactctat	atctgggtag	tatatgctat	cctaactctgt	atccgggtag	catatgctat	5220
cctaataagag	attagggtag	tatatgctat	cctaatttat	atctgggtag	catatactac	5280
ccaaatatct	ggatagcata	tgctatccta	atctatatct	gggtagcata	tgctatccta	5340
atctatatct	gggtagcata	ggctatccta	atctatatct	gggtagcata	tgctatccta	5400
atctatatct	gggtagcata	tgctatccta	atctatatct	gggtagcata	tgctatccta	5460
atctatatct	gggtagcata	tgctatccta	atctatatct	gggtagcata	tgctatccta	5520
atctgtatcc	gggtagcata	tgctatccta	atgataagct	gtcaaaccatg	agaattttct	5580
tgaagacgaa	agggcctcgt	gatacgccta	tttttatagg	ttaatgtcat	gataataatg	5640
gtttcttaga	cgtcagggtg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	5700
tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	5760
caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	5820
ttttttgcgg	cattttgcct	tcctgttttt	gctcaccag	aaacgctggg	gaaagtaaaa	5880
gatgtgaag	atcagttggg	tgacagagtg	ggttacatcg	aactggatct	caacagcggg	5940
aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaaggtt	6000
ctgctatgtg	gcgcggtatt	atcccgtgtt	gacgccgggc	aagagcaact	cggtcgcgcg	6060

-continued

```

atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg 6120
gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg 6180
gccaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgcacaaac 6240
atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca 6300
aacgacgagc gtgacaccac gatgcctgca gcaatggcaa caacgttgcg caaactatta 6360
actggcgaac tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat 6420
aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa 6480
tctggagccg gtgagcgtgg gtctcgcggc atcattgcag cactggggcc agatggtaag 6540
ccctcccgtg tcgtagttag ctacacgacg gggagtcagg caactatgga tgaacgaaat 6600
agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt 6660
tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag gatctagggtg 6720
aagatccttt ttgataatct catgacaaaa atcccttaac gtgagttttc gttccactga 6780
gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta 6840
atctgtctgt tgcaaaaaaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa 6900
gagctaccaa ctctttttcc gaaggttaact ggcttcagca gagcgcagat accaaatact 6960
gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca 7020
tacctcgctc tgctaactct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt 7080
accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg 7140
ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag 7200
cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa aggcggacag gtatccggta 7260
agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa cgctcggtat 7320
ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg 7380
tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc 7440
ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac 7500
cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc 7560
gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgct 7620
tgcccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag 7680
cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccagggtttt acactttatg 7740
cttccggctc gtatgttgtg tggaaattgtg agcggataac aatttcacac aggaaacagc 7800
tatgaccatg attacgcaa gctctagcta gaggtcgagt ccctcccag caggcagaag 7860
tatgcaaagc atgcacttca attagtcagc aaccatagtc ccgcccctaa ctcgccccat 7920
cccgccccta actccgcccc gttccgcccc ttctccgccc catggctgac taattttttt 7980
tatttatgca gaggcgaggg ccgcctcggc ctctgagcta ttccagaagt agtgaggagg 8040
cttttttggg ggccctaggct ttgcaaaaa gctttgcaa gatggataaa gttttaaaca 8100
gagaggaatc tttgcagcta atggacctc taggtcttga aagg 8144

```

<210> SEQ ID NO 6

<211> LENGTH: 8144

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

-continued

<223> OTHER INFORMATION: Synthetic construct: pJP189 ; pHybE-mCg1 V2

<400> SEQUENCE: 6

```

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtcgccga      60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa      120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta      180
tataagtga gtagtcgcg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca      240
ggtaagtgcc gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgct      300
gccttgaaat acttccacct ggctgcagta cgtgattctt gateccgagc ttcgggttg      360
aagtgggtgg gagagtcca ggccttcgc ttaaggagcc ccttcgcctc gtgcttgagt      420
tgaggcctgg cctgggcgct ggggcgcgcg cgtgcgaatc tggtggaacc ttcgcgcctg      480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct      540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcgg      600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg      660
gggcctcgca gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc      720
tctggtgctt ggctcgcgc gcctgtgtat cgcctcgccc tgggcggcaa ggctggccc      780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc      840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag tcaaccacac aaaggaaaag      900
ggcctttccg tcctcagccg tcgcttcagt tgactccaag gactaccggg cgccttcag      960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggagggggt      1020
ttatgcgatg gaggttcccc aactcgatg ggtggagact gaagttaggc cagcttgcca      1080
cttgatgtaa ttctccttgg aatttgcctt ttttgagttt ggatcttggg tcattctcaa      1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgtcgtgag gaattctcta      1200
gagatccctc gacctcgaga tccattgtgc ccgggcgcca ccatggagtt tgggctgagc      1260
tggtcttttc ttgtcgcgat ttaaaaggt gtccagtgcg catggtatgc cgaaagggat      1320
gctgaaattg agaacgaaaa gctgcgcggg gaggttgaag aactgcggca ggcagcgag      1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc      1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaaac cgcattctgt      1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg      1560
tcgggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttccct gaaacgggat      1620
atcatcaaa ccatgaacaa agcagccgcg ctggatgaac tgataccggg gttgctgagt      1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggttcgctc      1740
actgttcagg ccggagccac agaccgccgt tgaatgggcg gatgctaatt actatctccc      1800
gaaagaatcc gcataccagg aagggcgctg ggaaacactg cccttcagc gggccatcat      1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg      1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac      1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc      2040
gactattcgt gatattcctg cgtgctggc gctggccccg tggatggca aaaagcaccg      2100
ggataaacac ctcaccatga agcgtttcac taatgggctg ggcttctggt gcctgggcgg      2160
taaaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc      2220

```

-continued

tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctggggtg acaagcgcta	2280
aaacgacacc cccatctgtc tatccactgg cccctggatc tgctgcccaa actaactcca	2340
tggtgaccct gggatgectg gtcaagggtc atttccctga gccagtga ca gtgacctgga	2400
actctggatc cctgtccagc ggtgtgcaca ccttcccagc tgtcctgcag tctgacctct	2460
acactctgag cagctcagtg actgtccctc ccagcacctg gccagcgag accgtcacct	2520
gcaacgttgc ccacccggcc agcagcacca aggtggacaa gaaaattgtg cccagggtatt	2580
gtggttgtaa gccttgcata tgtacagtcc cagaagtatc atctgtcttc atcttcccc	2640
caaagcccaa ggatgtgtc accattactc tgactcctaa ggtcacgtgt gttgtggtag	2700
acatcagcaa ggatgatccc gaggtccagt tcagctggtt tgtagatgat gtggagggtc	2760
acacagctca gacgcaaccc cgggaggagc agttcaacag cactttccgc tcagtcagtg	2820
aacttcccat catgcaccag gactggctca atggcaagga gttcaaatgc aggggtcaaca	2880
gtgcagcttt ccctgcccc atcgagaaaa ccacttccaa aaccaaaggc agaccgaagg	2940
ctccacaggt gtacaccatt ccacctccca aggagcagat ggccaaggat aaagtcagtc	3000
tgacctgcat gataacagac ttcttccctg aagacattac tgtggagtgg cagtggaatg	3060
ggcagccagc ggagaactac aagaacactc agcccatcat ggacacagat ggctcttact	3120
tcgtctacag caagctcaat gtgcagaaga gcaactggga ggacaggaaat actttcacct	3180
gctctgtgtt acatgagggc ctgcacaacc accatactga gaagagcctc tccactctc	3240
ctggtaaata agcggccgct cgaggccggc aaggccggat ccccccacct cgacctctgg	3300
ctaataaagg aaattttatt tcattgcaat agtgtgttgg aattttttgt gtctctcact	3360
cggagggaca tatgggaggg caaatcattt ggtcgagatc cctcgagat ctctagctag	3420
aggatcgatc cccgccccgg acgaactaaa cctgactacg acatctctgc ccttctctcg	3480
cggggcagtg catgtaatcc cttcagttgg ttggtacaac ttgccaaactg ggccctgttc	3540
cacatgtgac acgggggggg accaaacaca aaggggttct ctgactgtag ttgacatcct	3600
tataaatgga tgtgcacatt tgccaacact gagggtgttt catcctggag cagactttgc	3660
agtctgtgga ctgcaacaca acattgcctt tatgtgtaac tcttggctga agctcttaca	3720
ccaatgctgg gggacatgta cctcccaggg gccaggaag actacgggag gctacaccaa	3780
cgtaaatcag aggggcctgt gtagctaccg ataagcggac cctcaaggag gcattagcaa	3840
tagtgtttat aaggccccct tgttaaccct aaacgggtag catatgcttc ccgggtagta	3900
gtatatacta tccagactaa ccctaattca atagcatatg ttaccaaacg ggaagcatat	3960
gctatcgaat tagggtagt aaaaggggtc taaggaacag cgatatctcc caccatga	4020
gctgtcacgg ttttatttac atgggggtcag gattccacga gggtagtgaa ccattttagt	4080
cacaagggca gtggctgaag atcaaggagc gggcagtgaa ctctcctgaa tcttcgctg	4140
cttcttcatt ctccctcggt tagctaatag aataactgct gagttgtgaa cagtaagggtg	4200
tatgtgaggt gctcgaaaac aaggtttcag gtgacgcccc cagaataaaa tttggacggg	4260
gggttcagtg gtggcattgt gctatgacac caatataacc ctcacaaacc ccttgggcaa	4320
taaatactag tgtaggaatg aaacattctg aatatcttta acaatagaaa tccatgggggt	4380
ggggacaagc cgtaaagact ggatgtccat ctcacacgaa tttatggcta tgggcaaac	4440
ataatcctag tgcaatatga tactgggggtt attaagatgt gtcccaggca gggaccaaga	4500

-continued

caggtgaacc	atgttggttac	actctatattg	taacaagggg	aaagagagtg	gacgccgaca	4560
gcagcggact	ccactgggtg	tctctaacac	ccccgaaaat	taaacggggc	tccacgccaa	4620
tggggcccat	aaacaagac	aagtggccac	tctttttttt	gaaattgttg	agtgggggca	4680
cgcgtcagcc	cccacacgcc	gccctgcggt	tttgactgt	aaaataagg	tgtaataact	4740
tggtgattg	taaccccgct	aaccactgcg	gtcaaaccac	tgcccacaa	aaccactaat	4800
ggcaccgccg	ggaataacct	cataagtagg	tggcggggcc	aagatagggg	cgcgattgct	4860
gcgatctgga	ggacaaatta	cacacacttg	cgctgagcg	ccaagcacag	ggttggtggt	4920
cctcatattc	acgaggtcgc	tgagagcagc	gtgggctaata	gttgccatgg	gtagcatata	4980
ctacccaaat	atctggatag	catatgctat	cctaactctat	atctgggtag	cataggctat	5040
cctaactctat	atctgggtag	catatgctat	cctaactctat	atctgggtag	tatatgctat	5100
cctaatttat	atctgggtag	cataggctat	cctaactctat	atctgggtag	catatgctat	5160
cctaactctat	atctgggtag	tatatgctat	cctaactctgt	atccgggtag	catatgctat	5220
cctaatagag	attagggtag	tatatgctat	cctaatttat	atctgggtag	catatactac	5280
ccaaatatct	ggatagcata	tgctatccta	atctatatct	gggtagcata	tgctatccta	5340
atctatatct	gggtagcata	ggctatccta	atctatatct	gggtagcata	tgctatccta	5400
atctatatct	gggtagcata	tgctatccta	atctatatct	gggtagcata	ggctatccta	5460
atctatatct	gggtagcata	tgctatccta	atctatatct	gggtagcata	tgctatccta	5520
atctgtatcc	gggtagcata	tgctatccta	atgataagct	gtcaaactg	agaattttct	5580
tgaagacgaa	agggcctcgt	gatacgcccta	tttttatagg	ttaatgtcat	gataaataatg	5640
gtttcttaga	cgtcagggtg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	5700
tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	5760
caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	5820
ttttttcgcg	cattttgctt	tcctgttttt	gctcaccag	aaacgctggt	gaaagtaaaa	5880
gatgctgaag	atcagttggg	tgacagagt	ggttacatcg	aactggatct	caacagcggg	5940
aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaaagtt	6000
ctgctatgtg	gcgcggtatt	atcccggtgt	gacgccgggc	aagagcaact	cggtcgccgc	6060
atacactatt	ctcagaatga	cttggttag	tactcaccag	tcacagaaaa	gcatcttacg	6120
gatggcatga	cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactgcg	6180
gccaaacttac	ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgacaaac	6240
atgggggatc	atgtaactcg	ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	6300
aacgacgagc	gtgacaccac	gatgcctgca	gcaatggcaa	caacgttgcg	caaactatta	6360
actggcggaac	tacttactct	agcttcccgg	caacaattaa	tagactggat	ggaggcggat	6420
aaagttgcag	gaccacttct	gcgctcggcc	cttcgggctg	gctgggttat	tgctgataaa	6480
tctggagccg	gtgagcgtgg	gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	6540
ccctcccgtg	tcgtagttat	ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	6600
agacagatcg	ctgagatagg	tgccctcactg	attaagcatt	ggtaactgtc	agaccaagtt	6660
tactcatata	tactttagat	tgatttaaaa	cttcattttt	aatttaaaag	gatctagggtg	6720
aagatccttt	ttgataatct	catgacaaaa	atcccttaac	gtgagttttc	gttccactga	6780
gcgtcagacc	ccgtagaaaa	gatcaaagga	tcttcttgag	atcctttttt	tctgcgcgta	6840

-continued

```

atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa 6900
gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact 6960
gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca 7020
tacctcgctc tgctaactct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt 7080
accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg 7140
ggttcgtgca cacagcccag ctgggagcga acgacctaca ccgaactgag atacctacag 7200
cgtgagctat gagaaagcgc cgcgttccc gaaggagaa aggcggacag gtatccggtg 7260
agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa cgctggtat 7320
ctttatagtc ctgctgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg 7380
tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc 7440
ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac 7500
cgtattaccg cttttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc 7560
gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgct 7620
tgcccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc ggcagtgag 7680
cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccaggtctt acactttatg 7740
cttcggctc gtatgttgtg tggaattgtg agcggataac aatttcacac aggaacagc 7800
tatgaccatg attacgcaa gctctagcta gaggtcagtg cctccccag caggcagaag 7860
tatgcaaagc atgcattcca attagtcagc aaccatagtc ccgccctaa ctccgccat 7920
cccgcccta actccgcca gttccgcca ttctcgccc catggctgac taattttttt 7980
tatttatgca gaggcgagg ccgcctggc ctctgagcta ttccagaagt agtgaggagg 8040
cttttttga gccctaggt tttgcaaaaa gctttgcaa gatggataaa gttttaaaca 8100
gagaggaatc tttgcagcta atggacctc taggtcttga aagg 8144

```

<210> SEQ ID NO 7

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP177 ; pHybE-mCg2a V1

<400> SEQUENCE: 7

```

agtgggaatt ggctccgggt cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtgct gcaattgaac cgggtccctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gagggtgagg gagaaccgta 180
tataagtga gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg 360
aagtgggtgg gagagttcga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgct ggggcccgcc cgtgcgaatc tggtggcacc ttcgcgcctg 480
tctcgctgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660

```

-continued

gggcctgcga	gcgcggccac	cgagaatcgg	acgggggtag	tctcaagctg	gccggcctgc	720
tctggtgcct	ggcctcgcgc	cgccgtgtat	cgccccgccc	tgggcggcaa	ggctggcccg	780
gtcggcacca	gttgctgtag	cggaaagatg	gccgcttccc	ggccctgctg	cagggagctc	840
aaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcctg	tgactccacg	gagtaccggg	cgcgctccag	960
gcacctcgat	tagttctcga	gcttttgag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgccct	ttttgagttt	ggatcttggg	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tcattttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgagc	atttaaatgc	cgggcgcca	ccatggagtt	tgggctgagc	1260
tggctttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggtagtc	cgaagggat	1320
gctgaaattg	agaacgaaaa	gctgcgccgg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaac	tattgagtac	gaacgccatc	gacttacgcg	tcgcagggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tggtggaac	cgcattctgt	1500
actttctgtc	tgctcgggat	cgcaggtgaa	attgccagta	ttctcgacgg	gctccccttg	1560
tcggtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttccct	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagcccg	ctggatgaac	tgataccggg	gttgcctgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgccg	ggcttcgctc	1740
actgttcagg	ccggagccac	agaccgccgt	tgaatgggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaacctg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggtgaatgtg	gtgaagtctg	cccggtcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccagcg	atgggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggccccg	tggtaggca	aaaagcaccg	2100
ggataaacag	ctcaccatga	agcgtttcac	taatgggctg	ggcttctggt	gcctgggctg	2160
taaagcgcca	aaaaactacc	gtgaaaagtc	ggtggatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttctgggtg	acaagcgcta	2280
aaacaacagc	cccatcggtc	tatccactgg	cccctgtgtg	tggagatata	actggctcct	2340
cggtgactct	aggatgcctg	gtcaagggtt	atttccctga	gccagtgaac	ttgacctgga	2400
actctggatc	cctgtccagt	ggtgtgcaca	ccttcccagc	tgtcctgcag	tctgacctct	2460
acacctcag	cagctcagtg	actgtaacct	cgagcacctg	gccagccag	tccatcacct	2520
gcaatgtggc	ccaccggcca	agcagcacca	aggtggacaa	gaaaattgag	cccagagggc	2580
ccacaatcaa	gcctgtcct	ccatgcaaat	gccagcacc	taacctcttg	ggtggaccat	2640
ccgtcttcat	cttccctcca	aagatcaagg	atgtactcat	gatctccctg	agcccatag	2700
tcacatgtgt	ggtgggtgat	gtgagcgagg	atgaccagaa	tgtccagatc	agctggtttg	2760
tgaacaacgt	ggaagtacac	acagctcaga	cacaaaccca	tagagaggat	tacaacagta	2820
ctctccgggt	ggtcagtgcc	ctccccatcc	agcaccagga	ctggatgagt	ggcaaggagt	2880
tcaaatgcaa	ggtcaacaac	aaagacctcc	cagcgcccat	cgagagaacc	atctcaaac	2940
ccaaagggtc	agtaagagct	ccacaggtat	atgtcttgcc	tccaccagaa	gaagagatga	3000

-continued

ctaagaaaca ggtcactctg acctgcatgg tcacagactt catgcctgaa gacatttacg	3060
tggagtggaac caacaacggg aaaacagagc taaactacaa gaacactgaa ccagtccctgg	3120
actctgatgg ttcttacttc atgtacagca agctgagagt ggaaaagaag aactgggtgg	3180
aaagaaatag ctactcctgt tcagtgggtc acgaggggtc gcacaatcac cacacgacta	3240
agagcttctc ccggactccg ggtaaaataag cggccgctcg aggcgggcaa ggcgggatcc	3300
cccgacctcg acctctgggt aataaaggaa atttattttc attgcaatag tgtgttgga	3360
ttttttgtgt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctagctagag gatcgatccc cgccccggac gaactaaacc tgactacgac	3480
atctctgccc ctctctcgcg gggcagtgc tgaatccct tcagttgggt ggtacaactt	3540
gccaaactggg ccctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600
gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tcttgagca gactttgcag tctgtggact gcaacacaa attgccttta tgtgtaactc	3720
ttggctgaag ctcttacacc aatgtgggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggagcg tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa ggcccccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtagtaa aagggtccta aggaacagcg	4020
atatctccca ccccatgagc tgtcacgggt ttatttacat ggggtcagga ttccacgagg	4080
gtagtgaacc atttttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctcctgaatc ttcgcctgct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaaggtgta tgtgaggtgc tcgaaaacaa ggtttcaggt gacgccccca	4260
gaataaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaaaccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catggggtgg ggacaagccg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctggggttat taagatgtgt	4500
cccaggcagg gaccaagaca ggtgaacat gttgttacac tctatttga acaaggggaa	4560
agagagtga cgcgcacagc agcggactcc actggttgtc tctaacaccc ccgaaaatta	4620
aacggggctc cagccaatg gggcccataa acaaagacaa gtggccactc tttttttga	4680
aattgtggag tgggggcacg cgtcagcccc cacacgccgc cctgcggttt tggactgtaa	4740
aataaggggt taataacttg gctgattgta accccgctaa ccactgcggt caaaccactt	4800
gcccacaaaa ccactaatgg cccccgggg aatacctgca taagtaggtg ggcgggcaa	4860
gataggggag cgattctgc gatctggagg acaaattaca cacacttgcg cctgagcgcc	4920
aagcacaggg ttgttggtcc tcatattcac gaggtcgtg agagcacggt gggctaattg	4980
tgccatgggt agcatatact acccaaatat ctggatagca tatgctatcc taatctatat	5040
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat	5100
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat	5160
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat	5220
ccgggtagca tatgctatcc taatagagat tagggtagta tatgctatcc taatttatat	5280

-continued

ctgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	5340
gtagcatatg ctatcctaata ctatatctgg gtagcatagg ctatcctaata ctatatctgg	5400
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ttatatctgg	5460
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	5520
gtagtatatg ctatcctaata ctgtatccgg gtagcatatg ctatcctaata gataagctgt	5580
caaacatgag aattttcttg aagacgaaag ggctcgtga tacgcctatt tttataggtt	5640
aatgtcatga taataatggt ttcttagacg tcagggtggca cttttcgggg aaatgtgcgc	5700
ggaaccccta ttgttttatt ttcttaaata cattcaaata tgtatccgct catgagacaa	5760
taacctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc	5820
cgtgtcgccc ttattccctt ttttcgggca ttttccttc ctgtttttgc taccacagaa	5880
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cagcagtggtg ttacatcgaa	5940
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg	6000
atgagcactt ttaaagtctt gctatgtggc gcggtattat cccgtgttga cgccgggcaa	6060
gagcaactcg gtgcgcgat acactattct cagaatgact tggttgagta ctcaccagtc	6120
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc	6180
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta	6240
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag	6300
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgctgcagc aatggcaaca	6360
acgttgcgca aactattaac tggcgaaacta ctactctag cttcccgga acaattaata	6420
gactggatgg aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc	6480
tggtttattg ctgataaate tggagccgtg gagcgtgggt ctcgcggtat cattgcagca	6540
ctggggccag atggtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca	6600
actatggatg aacgaaatag acagatcgt gagatagggt cctcactgat taagcattgg	6660
taactgtcag accaagttaa ctcatatata ctttagattg atttaaaact tcatttttaa	6720
tttaaaagga tctagtgaa gatcctttt gataatctca tgacaaaat cccttaacgt	6780
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	6840
cctttttttc tgccgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcgtg	6900
gtttgtttgc cggatcaaga gctaccaact ctttttcga aggttaactgg cttcagcaga	6960
gcgcagatag caaataactgt tcttctagt tagccgtagt taggccacca cttcaagaac	7020
tctgtagcac cgctacata cctcgtctg ctaatcctgt taccagtggc tgctgccagt	7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag	7140
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	7200
gaactgagat acctacacg tgagctatga gaaagcgcca cgcttccga agggagaaag	7260
gcggacaggt atccggaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca	7320
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt	7380
cgatttttgt gatgctcgtc agggggggcg agcctatgga aaaacgccag caacgcggcc	7440
tttttacggt tcttgccctt ttgctggcct tttgctcaca tgttctttcc tgcggtatcc	7500
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	7560
cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgcgc aatcgcgaaa	7620

-continued

```

ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 7680
tggaagcgg gcagtgcgc caacgcaatt aatgtgagtt agctcactca ttaggcaccc 7740
caggctttac actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa 7800
ttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc 7860
ctccccagca ggcagaagta tgcaaacgat gcattctcaat tagtcagcaa ccatagtccc 7920
gcccctaact ccgcccattcc cgcacctaac tccgcccagt tccgcccatt ctccgcccga 7980
tggtgcacta atttttttta tttatgcaga ggcgaggcc gcctcgccct ctgagctatt 8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaga 8100
tgataaagt tttaaacaga gaggaatctt tgcagctaag ggaccttcta ggtcttgaag 8160
gg 8162

```

<210> SEQ ID NO 8

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP190 ; pHybE-mCg2a V2

<400> SEQUENCE: 8

```

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtgc gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggttgggg gagaaccgta 180
tataagtga gtagtcgcc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt ccgcggggcc tggcctcttt acgggttatg gcccttgct 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcggttggt 360
aagtgggtgg gagagtcca ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tgggtggcac ttcgcgcctg 480
tctcgctgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacctg gtatttcggt 600
tttttgggcc gcggggcgcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctggtgctt ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggctggccc 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc 840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcacccacac aaaggaaaag 900
ggcctttccg tcctcagccg tcgcttcagt tgactccacg gactaccggg cgcgctccag 960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt 1020
ttatgcgatg gattttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca 1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa 1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgcgtgag gaattctcta 1200
gagatccctc gacctcgaga tccattgtgc ccgggcgcca ccatggagtt tgggtgagc 1260
tggttttttc ttgtcgcgat tttaaaaggt gtccagtgcg catggtatgc cgaagggat 1320
gctgaaattg agaacgaaaa gctgcgcggg gaggttgaag aactgcggca ggcagcgag 1380

```

-continued

gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttctgtc tgtcgcggat cgcagggtgaa attgccagta ttctcgacgg gctccccctg	1560
tccgtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaag ccatgaacaa agcagccgcg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggcttcgctc	1740
actgttcagg ccggagccac agaccgcctg tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc	2040
gactattcgt gatattccgt cgctgctggc gctggccccg tggatatgga aaaagcaccg	2100
ggataacacg ctccacctga agcgtttcac taatggcggt ggcttctggt gcctggcgcg	2160
taaagcggca aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgcta	2280
aaacaacagc cccatcggtc tatccactgg cccctgtgtg tggagataca actggctcct	2340
cggtgactct aggatgcctg gtcaagggtt atttccctga gccagtgaac ttgacctgga	2400
actctggatc cctgtccagt ggtgtgcaca ccttcccagc tgtcctgcag tctgacctct	2460
acacctcag cagctcagtg actgtaacct cgagcacctg gccagccag tccatcacct	2520
gcaatgtggc ccaccggca agcagcacca aggtggacaa gaaaattgag ccagagggc	2580
ccacaatcaa gcctgtcct ccatgcaaat gccagcacc taacctcttg ggtggaccat	2640
ccgtcttcat ctccctcca aagatcaagg atgtactcat gatctcctg agcccatag	2700
tcacatgtgt ggtgggtgat gtgagcgagg atgaccaga tgtccagatc agctggtttg	2760
tgaacaacgt ggaagtacac acagctcaga cacaaccca tagagaggat tacaacagta	2820
ctctccgggt ggtcagtgcc ctccccatcc agcaccagga ctggatgagt ggcaaggagt	2880
tcaaatgcaa ggtcaacaac aaagacctcc cagcgcccat cgagagaacc atctcaaac	2940
ccaaagggtc agtaagagct ccacaggat atgtcttgcc tccaccagaa gaagagatga	3000
ctaagaaaca ggtaactctg acctgcatgg tcacagactt catgcctgaa gacatttacg	3060
tggagtggac caacaacggg aaacacagagc taaactacaa gaacactgaa ccagtccctg	3120
actctgatgg ttcttacttc atgtacagca agctgagagt ggaagaag aactgggtgg	3180
aaagaaatag ctactcctgt tcagtgggtc acgagggtct gcacaatcac cacacgacta	3240
agagcttctc ccgactccg ggtaaataag cggccgctcg aggcggcaa ggccggatcc	3300
cccgaacctg acctctggct aataaaggaa atttattttc attgcaatag tgtgttgga	3360
ttttttgtgt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctagttagag gatcgatccc cgcgccggac gaactaaacc tgactacgac	3480
atctctgccc cttcttcgcg gggcagtgca tgtaatccct tcagttggtt ggtacaactt	3540
gccaaactgg ccctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600
gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tcctggagca gactttgcag tctgtggact gcaacacaac attgccttta tgtgtaactc	3720

-continued

ttggctgaag ctcttacacc aatgctgggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggaggc tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa gggcccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtagtaa aagggtccta aggaacagcg	4020
atatctccca ccccatgagc tgtcacggtt ttatttacat ggggtcagga ttccacgagg	4080
gtagtgaacc atttttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctcctgaatc ttcgcctgct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaaggtgta tgtgagggtc tcgaaaacaa ggtttcaggt gacgccccca	4260
gaataaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaaaccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catgggggtg ggacaagccg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctggggttat taagatgtgt	4500
cccaggcagg gaccaagaca ggtgaacat gttgttacac tctatttga acaaggggaa	4560
agagagtgga cgcgcacagc agcggactcc actggttgc tctaacaccc ccgaaaatta	4620
aacggggctc cagccaatg gggcccataa acaaagacaa gtggccactc tttttttga	4680
aattgtggag tgggggcacg cgtcagcccc cacacgccgc cctgcggttt tggactgtaa	4740
aataagggtg taataacttg gctgattgta accccgctaa ccaactgcgg caaaccactt	4800
gcccacaaaa ccactaatgg cccccgggg aatactgca taagtaggtg ggcgggccaa	4860
gataggggag cgattctgc gatctggagg acaaattaca cacacttgcg cctgagcgcc	4920
aagcacaggg ttgttggtcc tcatattcac gaggtcgtg agagcacggg gggctaattgt	4980
tgccatgggt agcatatact acccaaatat ctggatagca tatgctatcc taatctatat	5040
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat	5100
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat	5160
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat	5220
ccgggtagca tatgctatcc taatagagat tagggtagta tatgctatcc taatttatat	5280
ctgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	5340
gtagcatatg ctatcctaata ctatatctgg gtagcatagg ctatcctaata ctatatctgg	5400
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ttatatctgg	5460
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	5520
gtagtatatg ctatcctaata ctgtatccgg gtagcatatg ctatcctcat gataagctgt	5580
caaacatgag aattttcttg aagacgaaag ggcctcgtga tacgcctatt tttatagggt	5640
aatgtcatga taataatggt ttcttagacg tcagggtggca cttttcgggg aaatgtgcgc	5700
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa	5760
taacctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc	5820
cgtgtcgcgc ttattccctt ttttgccgca ttttgccctc ctgtttttgc tcaccagaa	5880
acgtcgtgta aagtaaaaga tgctgaagat cagttgggtg cacgagtggt ttacatcgaa	5940
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg	6000

-continued

atgagcactt	ttaaagttct	gctatgtggc	gcggtattat	cccggttga	cgccgggcaa	6060
gagcaactcg	gtcgccgc	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tgcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6180
atgagtata	acactgcgcg	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	6240
accgcttttt	tgacacaacat	gggggatcat	gtaactcgcc	ttgatcggtg	ggaaccggag	6300
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgcagc	aatggcaaca	6360
acgttgcgca	aactattaac	tgcggaacta	cttactctag	cttcccgga	acaattaata	6420
gactggatgg	aggcgataa	agttgcagga	ccacttctgc	gctcgccct	tccggctggc	6480
tggtttattg	ctgataaatc	tgagccggt	gagcgtgggt	ctcgcggtat	cattgcagca	6540
ctggggccag	atggtaagcc	ctcccgatc	gtagtattct	acacgacggg	gagtcaggca	6600
actatggatg	aacgaaatag	acagatcgct	gagatagggt	cctcactgat	taagcattgg	6660
taactgtcag	accaagttta	ctcatatata	ctttagattg	atttaaaact	tcatttttaa	6720
tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacccaaat	cccttaacgt	6780
gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	6840
cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	accagcgggtg	6900
gtttgtttgc	cggatcaaga	gtacccaact	ctttttccga	aggtaactgg	cttcagcaga	6960
gcgcagatac	caaataactgt	tcttctagt	tagccgtagt	taggccacca	cttcaagaac	7020
tctgtagcac	cgctacata	cctcgctctg	ctaactcctgt	taccagtggc	tgctgccagt	7080
ggcgataagt	cgtgtcttac	cgggttgac	tcaagacgat	agttaccgga	taaggcgag	7140
cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tgagcggaac	gacctacacc	7200
gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgcttcccgga	aggagagaaag	7260
gcggacaggt	atccggttaag	cggcagggtc	ggaacaggag	agcgacagag	ggagcttcca	7320
gggggaaacg	cctgggtatct	ttatagtcct	gtcgggtttc	gccacctctg	acttgagcgt	7380
cgatttttgt	gatgctcgct	agggggggcg	agcctatgga	aaaacgccag	caacgcggcc	7440
tttttacggt	tcctggcctt	ttgctggcct	ttgtctcaca	tgttctttcc	tgcgttatcc	7500
cctgattctg	tgataaacg	tattaccgcc	tttgagttag	ctgataccgc	tcgccgcagc	7560
cgaacgaccg	agcgacgcga	gtcagttagc	gaggaagcgg	aagagcggcc	aatacgcaaa	7620
ccgcctctcc	ccgcgcgttg	gccgattcat	taatgcagct	ggcacgacag	gtttcccgac	7680
tggaagcgg	gcagttagcg	caacgcaatt	aatgtgagtt	agctcactca	ttaggcaccc	7740
caggctttac	actttatgct	tccggtcgt	atgttgtgtg	gaattgtgag	cggataacaa	7800
tttcacacag	gaaacagcta	tgaccatgat	tacgccaagc	tctagctaga	ggctcgagtcc	7860
ctcccagca	ggcagaagta	tgcaaacgat	gcactctaat	tagtcagcaa	ccatagtccc	7920
gcccctaact	cgccccatcc	cgcccctaac	tccgcccagt	tccgcccatt	ctcgcgccca	7980
tggtgacta	atTTTTTTta	tttatgcaga	ggccgaggcc	gcctcgccct	ctgagctatt	8040
ccagaagtag	tgaggaggct	tttttgagg	cctaggcttt	tgcaaaaagc	tttgcaaaaga	8100
tgataaaagt	tttaaacaga	gaggaatctt	tgagctaat	ggaccttcta	ggtcttgaaa	8160
gg						8162

<210> SEQ ID NO 9

<211> LENGTH: 7503

-continued

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic construct: pJP178 ; pHybE-hCk V1

<400> SEQUENCE: 9
agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga      60
gaagttgggg ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa      120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta      180
tataagtgca gtagtcgcg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca      240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctg      300
gccttgaaat acttccacct ggctgcagta cgtgattctt gateccgagc ttcgggttgg      360
aagtgggtgg gagagtctga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt      420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tggtggcacc ttcgcgcctg      480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct      540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt      600
ttttggggcc gcggggcggc acggggcccg tgcgtcccag cgcacatggt cggcgaggcg      660
gggctcgcca gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc      720
tctgggtgct ggctcgcgc gcctgtgtat cgcgccgccc tgggcggcaa ggttgccccg      780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg cagggagctc      840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag tcaaccacac aaaggaaaag      900
ggcctttccg tcctcagccg tcgcttcagt tgactccaag gagtaccggg cgcgcgtccag      960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggagggggt      1020
ttatgcgatg gagtttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca      1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa      1140
gcctcagaca gtggttcaaa gttttttctt tccatttcag gtgtcgtgag gaattctcta      1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcac catggacatg cgcgtgcccg      1260
cccagctgct gggcctgctg ctgctgtggt tccccggctc gcgatgcgca tggtatgccg      1320
aaagggatgc tgaatttgag aacgaaaagc tgcgccggga ggttgaagaa ctgcggcagg      1380
ccagcgaggc agatctccag ccaggaaacta ttgagtacga acgccatcga cttacgcgtg      1440
cgcaggccga cgcacaggaa ctgaagaatg ccagagactc cgtgaagtg gtggaaccg      1500
cattctgtac tttcgtgctg tcgcggatcg caggtgaaat tgccagtatt ctgcacgggc      1560
tccccctgtc ggtgcagcgg cgttttccgg aactggaaaa ccgacatggt gatttcctga      1620
aacgggatat catcaaagcc atgaacaaag cagccgcgct ggatgaactg ataccggggg      1680
tgctgagtga atatatcgaa cagtcagggt aacaggctgc ggcatthtgt ccgcgccggg      1740
cttcgctcac tgttcaggcc ggagccacag accgccgttg aatgggcgga tgctaattac      1800
tatctcccca aagaatccgc ataccaggaa gggcgctggg aaacactgcc ctttcagcgg      1860
gccatcatga atgcgatggg cagcgactac atccgtgagg tgaatgtggt gaagtctgcc      1920
cgtgtcgggt attccaaaat gctgctgggt gtttatgcct actttataga gcataagcag      1980
cgcaaccccc ttatctggtt gccgacggat ggtgatgccg agaactttat gaaaaccac      2040
gttgagccga ctatctgtga tattccgtcg ctgctggcgc tggccccgtg gtatggcaaa      2100

```

-continued

aagcacccggg	ataaacacgct	caccatgaag	cgtttcacta	atgggcgtgg	cttctggtgc	2160
ctgggcggta	aagcgcaaaa	aaactaccgt	gaaaagtcgg	tggatgtggc	gggttatgat	2220
gaacttgctg	cttttgatga	tgatatggaa	caggaagget	ctccgacgtt	cctgggtgac	2280
aagctacgta	cggtggctgc	accatctgtc	ttcatcttcc	cgccatctga	tgagcagttg	2340
aaatctggaa	ctgcctctgt	tgtgtgcctg	ctgaataact	tctatcccag	agaggccaaa	2400
gtacagtggg	aggtggataa	cgccctccaa	tcgggtaact	cccaggagag	tgacacagag	2460
caggacagca	aggacagcac	ctacagcctc	agcagcagcc	tgacgctgag	caaagcagac	2520
tacgagaaac	acaaagtcta	cgctctcgaa	gtcaccctac	agggcctgag	ctcgcccgtc	2580
acaaagagct	tcaacagggg	agagtgttga	gcggccgctc	gaggccggca	aggccggatc	2640
ccccgacctc	gacctctggc	taataaagga	aatttatctt	cattgcaata	gtgtgttgga	2700
atcttttctg	tctctcactc	ggaaggacat	atgggagggc	aaatcatttg	gtcgagatcc	2760
ctcgagatc	tctagctaga	ggatcgatcc	ccgccccgga	cgaactaaac	ctgactacga	2820
catctctgcc	ccttctctgc	ggggcagtcg	atgtaatccc	ttcagttggt	tggtacaact	2880
tgccaactgg	gcctgtttcc	acatgtgaca	cgggggggga	ccaaacacaa	aggggttctc	2940
tgactgtagt	tgacatcctt	ataaatggat	gtgcacattt	gccaacactg	agtggctttc	3000
atcctggagc	agactttgca	gtctgtggac	tgcaacacaa	cattgccttt	atgtgtaact	3060
cttggtgaa	gctcttacac	caatgctggg	ggacatgtac	ctcccagggg	cccaggaaga	3120
ctacgggagg	ctacaccaac	gtcaatcaga	ggggcctgtg	tagctaccga	taagcggacc	3180
ctcaagaggg	cattagcaat	agtgtttata	aggccccctt	gttaacccta	aacgggtagc	3240
atatgcttcc	cgggtagtag	tataactat	ccagactaac	cctaattcaa	tagcatatgt	3300
tacccaacgg	gaagcatatg	ctatcgaatt	agggttagta	aaagggtcct	aaggaacagc	3360
gatattctcc	accccatgag	ctgtcacggt	tttatcttca	tggggtcagg	attccacgag	3420
ggtagtgaac	catttttagtc	acaagggcag	tggtctgaaga	tcaaggagcg	ggcagtgaac	3480
tctcctgaat	cttcgcctgc	ttcttcatcc	tccttcgttt	agctaataga	ataactgctg	3540
agttgtgaac	agtaagggtg	atgtgaggtg	ctcgaaaaca	aggtttcagg	tgacgcccc	3600
agaataaaat	ttggacgggg	ggttcagtg	tggtcatttg	ctatgacacc	aatataaccc	3660
tcacaaaccc	cttgggcaat	aaatactagt	gtaggaatga	aacattctga	atatctttaa	3720
caatagaaat	ccatgggggtg	gggacaagcc	gtaaagactg	gatgtccatc	tcacacgaat	3780
ttatggctat	gggcaacaca	taatcctagt	gcaatatgat	actgggggtta	ttaagatgtg	3840
tcccaggcag	ggaccaagac	aggtgaacca	tggtgttaca	ctctatttgt	aacaagggga	3900
aagagagtgg	acgccgacag	cagcggactc	cactgggtgt	ctctaacacc	ccgaaaaatt	3960
aaacggggct	ccacgccaat	ggggcccata	aacaaagaca	agtggccact	cttttttttg	4020
aaattgtgga	gtggggggc	gcgtcagccc	ccacacgccc	ccctgcggtt	ttggactgta	4080
aaataagggt	gtaataactt	ggctgattgt	aaccccgcta	accactcgcg	tcaaaccact	4140
tgcccacaaa	accactaatg	gcaccccggg	gaatacctgc	ataagtaggt	gggcggggcca	4200
agataggggc	gcgattgctg	cgatctggag	gacaaattac	acacacttgc	gcctgagcgc	4260
caagcacagg	gttgttggtc	ctcatattca	cgaggctcgt	gagagcacgg	tgggcctaag	4320
ttgccatggg	tagcatatac	tacccaaata	tctggatagc	atatgctatc	ctaactctata	4380
tctgggtagc	ataggctatc	ctaactctata	tctgggtagc	atatgctatc	ctaactctata	4440

-continued

tctgggtagt	atatgctatc	ctaatttata	tctgggtagc	ataggctatc	ctaattctata	4500
tctgggtagc	atatgctatc	ctaattctata	tctgggtagt	atatgctatc	ctaattctgta	4560
tccgggtagc	atatgctatc	ctaattagaga	ttagggtagt	atatgctatc	ctaatttata	4620
tctgggtagc	atatactacc	caaatatctg	gatagcatat	gctatcctaa	tctatatctg	4680
ggtagcatat	gctatcctaa	tctatatctg	ggtagcatag	gctatcctaa	tctatatctg	4740
ggtagcatat	gctatcctaa	tctatatctg	ggtagtatat	gctatcctaa	tttatatctg	4800
ggtagcatag	gctatcctaa	tctatatctg	ggtagcatat	gctatcctaa	tctatatctg	4860
ggtagtatat	gctatcctaa	tctgtatccg	ggtagcatat	gctatcctca	tgataagctg	4920
tcaaactga	gaattttctt	gaagacgaaa	gggcctcgtg	atagcctat	ttttataggt	4980
taatgtcatg	ataataatgg	tttcttagac	gtcaggtggc	acttttcggg	gaaatgtgcg	5040
cggaaaccct	atttgtttat	ttttctaaat	acattcaaat	atgtatccgc	tcatgagaca	5100
ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	agtatgagta	ttcaacattt	5160
cctgtgcgcc	cttattccct	tttttgccgc	attttgccct	cctgtttttg	ctcaccaga	5220
aacgctgggtg	aaagtaaaag	atgctgaaga	tcagttgggt	gcacgagtgg	gttacatcga	5280
actggatctc	aacagcggta	agatccttga	gagttttcgc	cccgaagaac	gttttccaat	5340
gatgagcact	tttaaagtgc	tgctatgtgg	cgcggtatta	tcccgtgttg	acgcgggca	5400
agagcaactc	ggtcgcgcga	tacactatc	tcagaatgac	ttggttgagt	actcaccagt	5460
cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	ttatgcagtg	ctgccataac	5520
catgagtgat	aacactgcgg	ccaacttact	tctgacaacg	atcggaggac	cgaaggagct	5580
aaccgccttt	ttgcacaaca	tgggggatca	tgtaactcgc	cttgatcgtt	gggaaccgga	5640
gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	atgcctgcag	caatggcaac	5700
aacgttgccg	aaactattaa	ctggcgaaact	acttactcta	gcttcccgcc	aacaattaat	5760
agactggatg	gaggcggata	aagttgcagg	accacttctg	cgctcggccc	ttccggtctg	5820
ctggtttatt	gctgataaat	ctggagccgg	tgagcgtggg	tctcgcggta	tcattgcagc	5880
actggggcca	gatggtaagc	cctcccgat	cgtagttatc	tacacgacgg	ggagtcaggc	5940
aactatggat	gaacgaaata	gacagatcgc	tgagataggt	gcctcactga	ttaagcattg	6000
gtaactgtca	gaccaagttt	actcatatat	actttagatt	gatttaaaac	ttcattttta	6060
atttaaaagg	atctaggtga	agatcctttt	tgataatctc	atgacaaaaa	tcccttaacg	6120
tgagttttcg	ttccactgag	cgtcagaccc	cgtagaaaag	atcaaaggat	cttcttgaga	6180
tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	aaaccaccgc	taccagcggt	6240
ggtttgtttg	cggatcaag	agctaccaac	tctttttccg	aaggtaactg	gcttcagcag	6300
agcgcagata	ccaaatactg	ttcttctagt	gtagccgtag	ttaggccacc	acttcaagaa	6360
ctctgtagca	ccgctacat	acctcgtct	gctaactctg	ttaccagtgg	ctgctgccag	6420
tgccgataag	tcgtgtctta	ccgggttgga	ctcaagacga	tagttaccgg	ataaggcgca	6480
gcggtcgggc	tgaacggggg	gttcgtgcac	acagcccagc	ttggagcgaa	cgacctacac	6540
cgaactgaga	tacctacagc	gtgagctatg	agaaagcgcc	acgcttcccg	aaggagagaaa	6600
ggcggacagg	tatccggtaa	gcggcagggt	cggaaacagga	gagcgacga	gggagcttcc	6660
agggggaac	gcctggtatc	tttatagtc	tgtcgggttt	cgccacctct	gacttgagcg	6720

-continued

tcgatttttg	tgatgctcgt	cagggggg	gagcctatgg	aaaaacgcca	gcaacgcggc	6780
ctttttacgg	ttcctggcct	tttgctggcc	ttttgctcac	atgttctttc	ctgcgttata	6840
ccttgattct	gtggataacc	gtattaccgc	ctttgagtga	gctgataacc	ctgcgcgcag	6900
ccgaacgacc	gagcgcagcg	agtcagttag	cgaggaagcg	gaagagcgcc	caatacgcaa	6960
accgcctctc	cccgcgcggt	ggccgattca	ttaatgcagc	tggcacgaca	ggtttcccga	7020
ctggaaagcg	ggcagttagc	gcaacgcaat	taatgtgagt	tagctcactc	attaggcacc	7080
ccaggccttta	cacttttatgc	ttccggctcg	tatgttgtgt	ggaattgtga	gcggataaca	7140
atttcacaca	ggaaacagct	atgaccatga	ttacgccaa	ctctagctag	aggtcgagtc	7200
cctccccagc	aggcagaagt	atgcaaagca	tgcattctca	ttagtcagca	accatagtcc	7260
cgccccctaac	tccgccccatc	ccgccccctaa	ctccgccccag	ttccgccccat	tctccgcccc	7320
atggctgact	aatttttttt	atttatgcag	aggccgaggc	cgccctcgcc	tctgagctat	7380
tccagaagta	gtgaggagcg	ttttttggag	gcctaggctt	ttgcaaaaag	ctttgcaaag	7440
atggataaag	ttttaaacag	agaggaatct	ttgcagctaa	tggaccttct	aggtcttgaa	7500
agg						7503

<210> SEQ ID NO 10

<211> LENGTH: 7503

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP191 ; pHybE-hCk V2

<400> SEQUENCE: 10

agtgggaatt	ggctccggtg	cccgtcagtg	ggcagagcgc	acatcgccca	cagtcctccga	60
gaagtgtggg	ggagggtcgc	gcaattgaac	cgggtgcctag	agaaggtggc	gcggggtaaa	120
ctgggaaagt	gatgtcgtgt	actggctccg	cctttttccc	gaggggtggg	gagaaccgta	180
tataagtgca	gtagtgcgcg	tgaacgttct	ttttcgcaac	gggtttgccc	ccagaacaca	240
ggtaagtgcc	gtgtgtggtt	cccgcggggc	tggcctcttt	acgggttatg	gcccttgctg	300
gccttgaaat	acttccacct	ggctgcagta	cgtgattctt	gatcccgagc	ttcgggttg	360
aagtgggtgg	gagagttcga	ggccttgccg	ttaaggagcc	ccttcgcctc	gtgcttgagt	420
tgaggcctgg	cctgggcgct	ggggcgccgc	cgtgcgaatc	tgggtggcacc	ttcgcgcctg	480
tctcgctgct	ttcgataagt	ctctagccat	ttaaaatttt	tgatgacctg	ctgcgacgct	540
ttttttctgg	caagatagtc	ttgtaaatgc	gggccaagat	ctgcacactg	gtatttcggt	600
ttttggggcc	gcgggcgccg	acggggcccg	tgcgtcccag	cgcacatggt	cgcgagggcg	660
gggcctgcga	gcgcggccac	cgagaatcgg	acgggggtag	tctcaagctg	gccggcctgc	720
tctggtgect	ggcctgcgcg	cgccgtgtat	cgcggcgccc	tgggcggcaa	ggctggcccg	780
gtcggcacca	gttgctgtag	cggaaagatg	gccgcttccc	ggcctgctg	cagggagctc	840
aaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcctg	tgaactccacg	gagtaccggg	cgccgtccag	960
gcacctcgat	tagttctcga	gcttttgag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	gggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgccct	ttttgagttt	ggatcttggt	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200

-continued

gagatccctc gacctcgaga tccattgtgc ccggggcgac catggacatg cgcgtgcccc	1260
cccagctgct gggcctgctg ctgctgtggt tccccggctc gcgatgcgca tggatatgcc	1320
aaagggatgc tgaattgag aacgaaaagc tgcgcggga ggttgaagaa ctgcggcagg	1380
ccagcgaggc agatctccag ccaggaacta ttgagtacga acgccatcga cttacgcgtg	1440
cgcaggccga cgcacaggaa ctgaagaatg ccagagactc cgctgaagtg gtggaaccg	1500
cattctgtac tttcgtgctg tcgcggtatc caggtgaaat tgccagtatt ctgcacgggc	1560
tccccctgtc ggtgcagcgg cgttttccgg aactggaaaa ccgacatgtt gatttctcta	1620
aacgggatat catcaagcc atgaacaaag cagccgcgct ggatgaactg ataccggggt	1680
tgtgtagtga atatatcgaa cagtccaggtt aacaggctgc ggcattttgt ccgcgcggg	1740
cttcgctcac tgttcaggcc ggagccacag accgccgttg aatggcgga tgctaattac	1800
tatctccca aagaatccgc ataccaggaa gggcgctggg aaacactgcc ctttcagcgg	1860
gccatcatga atgcgatggg cagcgactac atccgtgagg tgaatgtggt gaagtctgcc	1920
cgtgtcgggt attccaaat gctgtgggt gtttatgcct actttataga gcataagcag	1980
cgcaacaccc ttatctggtt gccgacggat ggtgatgccg agaactttat gaaaaccac	2040
ggtgagccga ctattcgtga tattccgtcg ctgctggcgc tggccccgtg gtatggcaaa	2100
aagcaccggg ataacacgct caccatgaag cgtttcacta atgggcgtgg cttctggtgc	2160
ctgggcggta aagcggcaaa aaactaccgt gaaaagtcgg tggatgtggc gggttatgat	2220
gaacttgctg cttttgatga tgatattgaa caggaaggct ctccgacgtt cctgggtgac	2280
aagctacgta cgggtggtgc accatctgtc ttcattctcc ccgatctga tgagcagttg	2340
aaatctggaa ctgcctctgt tgtgtgcctg ctgaataact tctatcccag agaggccaaa	2400
gtacagtgga aggtggataa cgccctccaa tcgggtaact ccaggagag tgtcacagag	2460
caggacagca aggacagcac ctacagctc agcagcacc tgacgctgag caaagcagac	2520
tacgagaaac acaaagtcta cgcctgcgaa gtcacccatc agggcctgag ctgcgccgtc	2580
acaaagagct tcaacagggg agagtgttga gcggccgctc gaggccggca aggccggatc	2640
ccccgacctc gacctctggc taataaagga aatttatatt cattgcaata gtgtgttga	2700
attttttgtg tctctcactc ggaaggacat atgggagggc aaatcatttg gtcgagatcc	2760
ctcggagatc tctagctaga ggatcgatcc ccgccccgga cgaactaaac ctgactacga	2820
catctctgcc ccttctctgc ggggcagtgc atgtaatccc ttcagttggt tggtaacct	2880
tgccaaactg gccctgttcc acatgtgaca cgggggggga ccaaacacaa aggggttctc	2940
tgactgtagt tgacatcctt ataaatggat gtgcacattt gccaacactg agtggtcttc	3000
atcctggagc agactttgca gtctgtggac tgcaacacaa cattgccttt atgtgtaact	3060
cttggctgaa gctcttacac caatgctggg ggacatgtac ctcccagggg ccaggaaga	3120
ctacgggagg ctacaccaac gtcaatcaga ggggcctgtg tagctaccga taagcggacc	3180
ctcaagaggg cattagcaat agtgtttata agggcccctt gttaacccta aacgggtagc	3240
atatgcttcc cgggtagtag tatatactat ccagactaac cctaattcaa tagcatatgt	3300
tacccaacgg gaagcatatg ctatcgaatt agggtttagta aaagggtcct aaggaacagc	3360
gatatctccc acccatgag ctgtcacggt tttatttaca tggggtcagg attccacgag	3420
ggtagtgaac cattttatgc acaagggcag tggctgaaga tcaaggagcg ggcagtgaac	3480

-continued

tctcctgaat	cttcgcctgc	ttcttcattc	tccttcggtt	agctaataga	ataactgctg	3540
agttgtgaac	agtaaggtgt	atgtgaggtg	ctcgaaaaca	aggtttcagg	tgacgcccc	3600
agaataaaat	ttggacgggg	ggttcagtg	tggtcattgt	ctatgacacc	aatataaccc	3660
tcacaaaccc	cttgggcaat	aaatactagt	gtaggaatga	aacattctga	atatctttaa	3720
caatagaaat	ccatgggggt	gggacaagcc	gtaaagactg	gatgtccatc	tcacacgaat	3780
ttatggctat	gggcaacaca	taatcctagt	gcaatatgat	actgggggtta	ttaagatgtg	3840
ttccaggcag	ggaccaagac	aggtgaacca	tgttggttaca	ctctatttgt	aacaagggga	3900
aagagagtgg	acgccgacag	cagcggactc	cactgggtgt	ctctaacacc	cccgaaaatt	3960
aaacggggct	ccacgccaat	ggggcccata	aacaagaca	agtggccact	cttttttttg	4020
aaattgtgga	gtgggggcac	gcgtcagccc	ccacacgccg	ccctgcgggt	ttggactgta	4080
aaataaggg	gtaataactt	ggctgattgt	aaccccgcta	accactgcgg	tcaaaccact	4140
tgccacaaa	accactaatg	gcacccggg	gaatacctgc	ataagtaggt	ggcggggcca	4200
agataggggc	gcgattgctg	cgatctggag	gacaaattac	acacacttgc	gcctgagcgc	4260
caagcacagg	gttggtggtc	ctcatattca	cgaggtcgct	gagagcacgg	tggtgtaagt	4320
ttgccatggg	tagcatatac	tacccaaata	tctggatagc	atatgctatc	ctaacttata	4380
tctgggtagc	ataggctatc	ctaacttata	tctgggtagc	atatgctatc	ctaacttata	4440
tctgggtagt	atatgctatc	ctaacttata	tctgggtagc	atatgctatc	ctaacttata	4500
tctgggtagc	atatgctatc	ctaacttata	tctgggtagt	atatgctatc	ctaacttata	4560
tcgggtagc	atatgctatc	ctaacttata	ttagggtagt	atatgctatc	ctaacttata	4620
tctgggtagc	atatactacc	caaatactctg	gtagcatat	gctatcctaa	tctatatctg	4680
gtagcatat	gctatcctaa	tctatatctg	gtagcatag	gctatcctaa	tctatatctg	4740
gtagcatat	gctatcctaa	tctatatctg	gtagtatat	gctatcctaa	tttatatctg	4800
gtagcatag	gctatcctaa	tctatatctg	gtagcatat	gctatcctaa	tctatatctg	4860
gtagtatat	gctatcctaa	tctgtatccg	gtagcatat	gctatcctca	tgataagctg	4920
tcaaactga	gaattttctt	gaagacgaaa	gggcctcgtg	atacgctat	ttttataggt	4980
taatgtcatg	ataataatgg	tttcttagac	gtcaggtggc	acttttcggg	gaaatgtgcg	5040
cggaaccctt	atttgtttat	ttttctaaat	acattcaaat	atgtatccgc	tcagagaca	5100
ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	agtatgagta	ttcaacattt	5160
ccgtgtgcgc	cttatccctt	tttttgcggc	attttgcctt	cctgtttttg	ctcaccaga	5220
aacgctgggt	aaagtaaaag	atgctgaaga	tcagttgggt	gcacgagtgg	gttacatcga	5280
actggatctc	aacagcggtg	agatccttga	gagttttcgc	cccgaagaac	gttttccaat	5340
gatgagcaact	tttaaagttc	tgctatgtgg	cgcgggtatta	tcccgtgttg	acgccgggca	5400
agagcaactc	ggtcgccgca	tacactatct	tcagaatgac	ttggttgagt	actcaccagt	5460
cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	ttatgcagtg	ctgccataac	5520
catgagtgat	aacactgcgg	ccaacttact	tctgacaacg	atcggaggac	cgaaggagct	5580
aaccgctttt	ttgcacaaca	tgggggatca	tgtaactcgc	cttgatcggt	gggaaccgga	5640
gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	atgcctgcag	caatggcaac	5700
aacgttgccg	aaactattaa	ctggcggaact	acttactcta	gcttcccggc	aacaattaat	5760
agactggatg	gagggcgata	aagttgcagg	accacttctg	cgctcgcccc	ttccggctgg	5820

-continued

```

ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggtg tcattgcagc 5880
actggggcca gatggtaagc cctcccgat cgtagttatc tacacgacgg ggagtcaggc 5940
aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 6000
gtaactgtca gaccaagttt actcatatat acttttagatt gatttaaaac ttcattttta 6060
atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg 6120
tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga 6180
tccttttttt ctgcgcgtaa tctgtctgct gcaaacaaaa aaaccaccgc taccagcggg 6240
ggtttggttg ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 6300
agcgcagata ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa 6360
ctctgtagca ccgcctacat acctcgtctc gctaatctg ttaccagtgg ctgctgccag 6420
tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca 6480
gcggtcgggc tgaacggggg gtctgtgcac acagcccagc ttggagcgaa cgacctacac 6540
cgaaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaa 6600
ggcgacagc tatccggtaa gcggcagggt cggaacagga gagcgacga gggagcttcc 6660
agggggaaac gcctggtatc tttatagtcc tgcggggtt cgccacctct gacttgagcg 6720
tcgatttttg tgatgctcgt caggggggag gagcctatgg aaaaacgcca gcaacgcggc 6780
ctttttacg ttccctggcct tttgttgcc tttgtctcac atgttctttc ctgcgttatc 6840
ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccgag 6900
ccgaacgacc gagcgcagcg agtcagttag cgaggaagcg gaagagcgcc caatacgcaa 6960
accgcctctc ccgcgcggtt ggccgattca ttaatgcagc tggcacgaca ggtttcccga 7020
ctggaaaagc ggagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc 7080
ccaggcttta cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca 7140
atttcacaca ggaaacagct atgacctga ttacgccaag ctctagctag aggtcgagtc 7200
cctccccagc aggcagaagt atgcaaaagc tgcatctcaa ttagtcagca accatagtcc 7260
cgcccctaac tccgccatc ccgccctaa ctccgccag ttccgcccat tctccgccc 7320
atggctgact aatttttttt atttatgcag aggcagaggc cgctcggcc tctgagctat 7380
tccagaagta gtgaggaggc ttttttgag gcctaggett ttgcaaaaag ctttgcaaag 7440
atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct aggtcttgaa 7500
agg 7503

```

```

<210> SEQ ID NO 11
<211> LENGTH: 7481
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic construct: pJP179 ; pHybE-hCl V1

<400> SEQUENCE: 11

```

```

agtgggaatt ggctccgggt ccgcgcagtg ggagagcgcc acatcgccca cagtccccga 60
gaagttgggg ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgctgtgt actggctccg cttttttccc gaggggtggg gagaaccgta 180
tataagtgca gtagtcgccc tgaacgttct ttttcgcaac gggtttgccc ccagaacaca 240

```

-continued

ggtaagtgcc	gtgtgtggtt	cccgcgggcc	tggcctcttt	acgggttatg	gcccttgct	300
gccttgaatt	acttccacct	ggctgcagta	cgtgattctt	gateccgagc	tccgggttgg	360
aagtgggtgg	gagagtccga	ggccttgccg	ttaaggagcc	ccttcgcctc	gtgcttgagt	420
tgaggcctgg	cctgggcgct	ggggcgccg	cgtgcgaatc	tgttggcacc	tccgcgcctg	480
tctcgctgct	tccgataagt	ctctagccat	ttaaaatttt	tgtgacctg	ctgcgacgct	540
ttttttctgg	caagatagtc	ttgtaaatgc	gggccaagat	ctgcacactg	gtatttcggg	600
ttttggggcc	gcgggcggcg	acggggcccg	tgcgtcccag	cgcacatggt	cggcgaggcg	660
gggctcgca	gcgcggccac	cgagaatcgg	acgggggttag	tctcaagctg	gccggcctgc	720
tctgggtgct	ggcctcgccg	cgccgtgtat	cgccccgccc	tgggcggcaa	ggctggcccg	780
gtcggcacca	gttgcgtgag	cggaaagatg	gccgcttccc	ggccttgctg	cagggagctc	840
aaaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcctg	tgaactccacg	gagtaccggg	cgcgctccag	960
gcacctcgat	tagttctcga	gcttttgag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggg	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgagc	atttaaatgc	ccgggcgcca	ccatgacttg	gacccactc	1260
ctcttctca	ccctcctcct	ccactgcaca	ggaagcttat	cgcgaaaggg	atgctgaaat	1320
tgagaacgaa	aagctgcgcc	gggaggttga	agaactgcgg	caggccagcg	aggcagatct	1380
ccagccagga	actattgagt	acgaacgcca	tcgacttacg	cgtgcgcagg	ccgacgcaca	1440
ggaactgaag	aatgccagag	actccgctga	agtgggtgaa	accgcattct	gtactttcgt	1500
gctgtcgccg	atcgcagggt	aaattgccag	tattctcgac	gggctcccc	tgtcggtgca	1560
gcggcgtttt	ccggaactgg	aaaacgcaca	tgttgatttc	ctgaaacggg	atatcatcaa	1620
agccatgaac	aaagcagccg	cgctggatga	actgataccg	gggttgctga	gtgaatatat	1680
cgaacagtca	ggtaacagg	ctgcggcatt	ttgtccgcgc	cgggcttcgc	tcactgttca	1740
ggccggagcc	acagaccgcc	gttgaatggg	cggatgctaa	ttactatctc	ccgaaagaat	1800
ccgcatacca	ggaagggcgc	tgggaaacac	tgccctttca	gcgggccatc	atgaatgcga	1860
tgggcagcga	ctacatccgt	gaggtgaatg	tggggaagtc	tgcccggtgc	ggttattcca	1920
aaatgctgct	gggtgtttat	gcctacttta	tagagcataa	gcagcgcaac	acccttatct	1980
ggttgccgac	ggatgggtgat	gccgagaact	ttatgaaaac	ccacgttgag	ccgactattc	2040
gtgatattcc	gtcgctgctg	gcgctggccc	cgtgggtatg	caaaaagcac	cgggataaca	2100
cgtcaccat	gaagcgtttc	actaatgggc	gtggcttctg	gtgcctgggc	ggtaaagcgg	2160
caaaaaacta	ccgtgaaaag	tcggtggatg	tggcggttga	tgatgaactt	gctgcttttg	2220
atgatgatat	tgaacaggaa	ggctctccga	cgttcctggg	tgacaagcgt	taaccaagg	2280
ctgccccctc	ggtcactctg	ttcccgccct	cctctgagga	gcttcaagcc	aacaaggcca	2340
cactggtgtg	tctcataagt	gacttctacc	cgggagccgt	gacagtggcc	tggaaggcag	2400
atagcagccc	cgtcaaggcg	ggagtggaga	ccaccacacc	ctccaaacaa	agcaacaaca	2460
agtacgcggc	cagcagctac	ctgagcctga	cgctgagca	gtggaagtcc	cacagaagct	2520
acagctgcc	ggtcacgcat	gaaggagca	ccgtggagaa	gacagtggcc	cctacagaat	2580

-continued

gttcatgagc ggccgctcga ggccggcaag gccggatccc ccgacctcga cctctggcta	2640
ataaaggaaa ttatttttca ttgcaatagt gtgttggaat tttttgtgtc tctcactcgg	2700
aaggacatat gggagggcaa atcatttgggt cgagatccct cggagatctc tagctagagg	2760
atcgatcccc gccccggacg aactaaacct gactacgaca tctctgcccc ttcttcgagg	2820
ggcagtgcat gtaatccctt cagttgggtg gtacaacttg ccaactgggc cctgttccac	2880
atgtgacacg gggggggacc aaacacaaag gggttctctg actgtagttg acatccttat	2940
aaatggatgt gcacatttgc caaactgag tggttttcat cctggagcag actttgcagt	3000
ctgtggactg caacacaaca ttgcctttat gtgtaactct tggtgaagc tcttacacca	3060
atgctggggg acatgtacct ccagggggcc caggaagact acgggaggct acaccaacgt	3120
caatcagagg ggctgtgta gctacgata agcggaccct caagagggca ttagcaatag	3180
tgtttataag gcccccttgt taaccctaaa cgggtagcat atgcttcccg ggtagtagta	3240
tatactatcc agactaacc taattcaata gcataatgta ccaacggga agcatatgct	3300
atcgaattag ggttagtaaa agggtoctaa ggaacagcga tatctccac ccctagagct	3360
gtcacggttt tatttacatg gggtcaggat tccacgagg tagtgaacca ttttagtcac	3420
aagggcagtg gctgaagatc aaggagcggg cagtgaactc tectgaatct tcgctgctt	3480
cttcattctc ctctgttag ctaatagaat aactgctgag ttgtgaacag taagggtgat	3540
gtgagtgct cgaacaacg gtttcagtg acgccccag aataaaattt ggacgggggg	3600
ttcagtggtg gcattgtgct atgacaccaa tataaccctc acaaacccct tgggcaataa	3660
atactagtgt aggaatgaaa cattctgaat atctttaaca atagaaatcc atggggtggg	3720
gacaagccgt aaagactgga tgtccatctc acacgaattt atggctatgg gcaacacata	3780
atcctagtgc aatatgatac tggggttatt aagatgtgtc ccaggcaggg accaagacag	3840
gtgaaccatg ttgttacct ctatttgtaa caaggggaaa gagagtggac gccgacagca	3900
gcggactcca ctggttgtct ctaacacccc cgaataatga acggggctcc acgccaatgg	3960
ggcccataaa caaagacaag tggccactct tttttttgaa attgtggagt gggggcacgc	4020
gtcagcccc acacgcgcgc ctgcggtttt ggactgtaaa ataagggtgt aataacttgg	4080
ctgattgtaa ccccgctaac cactgcggtc aaaccacttg ccacaaaaac cactaatggc	4140
accccgggga atacctgcat aagtaggtgg gcgggccaag ataggggcgc gattgctgcg	4200
atctggagga caaattacac acacttgcgc ctgagcgcca agcacagggt tgttggtcct	4260
catattcacg aggtcgctga gagcacggtg ggctaattgt gccatgggta gcatatacta	4320
cccaaatatc tggatagcat atgctatcct aatctatatc tgggtagcat aggctatcct	4380
aatctatatc tgggtagcat atgctatcct aatctatatc tgggtagtat atgctatcct	4440
aatttatatc tgggtagcat aggctatcct aatctatatc tgggtagcat atgctatcct	4500
aatctatatc tgggtagtat atgctatcct aatctgtatc cgggtagcat atgctatcct	4560
aatagagatt agggtagtat atgctatcct aatttatatc tgggtagcat atactacca	4620
aatatctgga tagcatatgc tatcctaata tatatctggg tagcatatgc tatcctaata	4680
tatatctggg tagcatagga tatcctaata tatatctggg tagcatatgc tatcctaata	4740
tatatctggg tagtatatgc tatcctaatt tatatctggg tagcatagga tatcctaata	4800
tatatctggg tagcatatgc tatcctaata tatatctggg tagtatatgc tatcctaata	4860

-continued

tgtatccggg tagcatatgc tatcctcatg ataagctgtc aaacatgaga attttcttga	4920
agacgaaagg gcctcgtgat acgcctattt ttataggtta atgtcatgat aataatgggt	4980
tcttagacgt cagggtggcac ttttcgggga aatgtgcgcg gaacccttat ttgtttattt	5040
ttctaataac attcaaatat gtatccgctc atgagacaat aacctgata aatgcttcaa	5100
taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt	5160
tttgcggcat tttgccttcc tgtttttgct caccagaaa cgctgggtgaa agtaaaagat	5220
gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag	5280
atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg	5340
ctatgtggcg cggtattatc ccgtgttgac gccgggcaag agcaactcgg tcgccgcata	5400
cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat	5460
ggcatgacag taagagaatt atgcagtgtc gccataacca tgagtgataa cactgcggcc	5520
aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg	5580
ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac	5640
gacgagcgtg acaccacgat gcctgcagca atggcaacaa cggtgcgcaa actattaact	5700
ggcgaactac ttactctage ttcccggcaa caattaatag actggatgga ggcggataaa	5760
gttgacggac cacttctcgc ctccggccctt ccggtcggct ggtttattgc tgataaatct	5820
ggagccgggtg agcgtgggtc tcgcgggtatc attgcagcac tggggccaga tggtaagccc	5880
tcccgatcgc tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga	5940
cagatcgcgtg agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac	6000
tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat ctagggtgaag	6060
atcctttttg ataatctcat gacccaaatc ccttaacgtg agttttcgtt ccaactgagcg	6120
tcagaccccg tagaaaaatg caaaggatct tottgagatc ctttttttct gcgcgtaatc	6180
tgctgcttgc aaacaaaaaa accacgccta ccagcgggtg tttgtttgcc ggatcaagag	6240
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt	6300
cttctagtgt agccgtagt tggccaccac ttcaagaact ctgtagcacc gcctacatac	6360
ctcgcctcgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc	6420
gggttggaact caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt	6480
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt	6540
gagctatgag aaagcggcac gcttcccgaa gggagaaaagg cggacaggta tccggtaagc	6600
ggcagggctg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggatatct	6660
tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca	6720
ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacgggt cctggccttt	6780
tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt	6840
attaccgcct ttgagtgagc tgataccgct cgcgcagacc gaacgaccga gcgcagcgag	6900
tcagtgagcg aggaagcggga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg	6960
ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaagcggg cagtgagcgc	7020
aacgcaatta atgtgagtta gctcactcat taggcacccc aggttttaca ctttatgctt	7080
ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat	7140
gaccatgatt acgccaagct ctagctagag gtcgagtcct tcccagcag gcagaagtat	7200

-continued

gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc	7260
gcccctaact cgcgccagtt cgcgccatc tccgccccat ggctgactaa ttttttttat	7320
ttatgcagag gccgaggccg cctcggcctc tgagctatc cagaagtagt gaggaggctt	7380
ttttggaggc ctaggctttt gcaaaaagct ttgcaaatg ggataaagtt ttaaacagag	7440
aggaatcttt gcagctaatt gaccttctag gtcttgaaag g	7481

<210> SEQ ID NO 12

<211> LENGTH: 7486

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP192 ; pHybE-hCl V2

<400> SEQUENCE: 12

agtggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtcgccga	60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctt	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg	360
aagtgggtgg gagagtcca ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcgcgag cgtgcgaatc tgggtggcacc ttcgcgcctg	480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgcct ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggctggcccg	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc	840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tccctcagcg tcgcttcctg tgactccacg gactaccggg cgccttcag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttctt tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gaactcgaga tccattgtgc cggggcgcca ccatgacttg gacccactc	1260
ctcttctca ccctcctcct ccaactgcaca ggaagcttat cgcgaaaggg atgctgaaat	1320
tgagaacgaa aagctgcgcc gggagggtga agaactgcgg caggccagcg aggcagatct	1380
ccagccagga actattgagt acgaacgcca tcgacttacg cgtgcgcagg ccgacgcaca	1440
ggaactgaag aatgccagag actccgctga agtgggtgaa accgcattct gtactttcgt	1500
gctgtcggcg atcgcagggt aaattgccag tattctcgac gggctcccc tgcgggtgca	1560
gcggcggttt ccggaactgg aaaacgcaca tgttgatttc ctgaaacggg atatcatcaa	1620
agccatgaac aaagcagccg cgctggatga actgataccg gggttgctga gtgaatatat	1680

-continued

cgaacagtca	ggttaacagg	ctgcggcatt	ttgtccgcgc	cgggcttcgc	tcactgttca	1740
ggcggagcc	acagaccgcc	gttgaatggg	cggatgctaa	ttactatctc	cgaagaat	1800
ccgcatacca	ggaagggcgc	tgggaacac	tgcctttca	gcgggccatc	atgaatgcga	1860
tgggcagcga	ctacatccgt	gaggtgaatg	tgggaagtc	tgccctgtgc	ggttattcca	1920
aaatgctgct	gggtgtttat	gcctacttta	tagagcataa	gcagcgcaac	acccttatct	1980
ggttgccgac	ggatggatg	gccgagaact	ttatgaaaac	ccacgttgag	ccgactattc	2040
gtgatatcc	gtcgtgctg	gcgctggccc	cgtggatgg	caaaaagcac	cgggataaca	2100
cgtcaccat	gaagcgtttc	actaatgggc	gtggttctg	gtgctgggc	ggtaaagcgg	2160
caaaaaacta	ccgtgaaaag	tcggtggatg	tggcgggtta	tgtgaactt	gctgcttttg	2220
atgatgat	tgaacaggaa	ggctctccga	cgttctggg	tgacaagcgc	taggtcaacc	2280
caaggctgcc	ccctcggtca	ctctgttccc	gccctcctct	gaggagcttc	aagccaacaa	2340
ggccacactg	gtgtgtctca	taagtgaact	ctaccggga	gccgtgacag	tggcctggaa	2400
ggcagatagc	agccccgtca	aggcgggagt	ggagaccacc	acaccctcca	aacaaagcaa	2460
caacaagtac	gcggccagca	gctacctgag	cctgacgcct	gagcagtgga	agtccacacg	2520
aagctacagc	tgccaggctca	cgcataaagg	gagcaccgtg	gagaagacag	tggcccctac	2580
agaatgttca	tgagcggccg	ctcaggccg	gcaaggccgg	atccccgcac	ctcgacctct	2640
ggctaataaa	ggaaatttat	tttcattgca	atagtgtgtt	ggaatttttt	gtgtctctca	2700
ctcgaaggga	catatgggag	ggcaaatcat	ttggtcgaga	tccctcggag	atctctagct	2760
agaggatcga	tcccccccc	ggacgaacta	aacctgacta	cgacatctct	gccccctctt	2820
cgcggggcag	tgcattgta	cccttcagtt	ggttggtaca	acttgccaac	tgggccctgt	2880
tccacatgtg	acacgggggg	ggaccaaaca	caaaggggtt	ctctgactgt	agttgacatc	2940
cttataaatg	gatgtgcaca	tttgccaaca	ctgagtggct	ttcatcctgg	agcagacttt	3000
gcagtctgtg	gactgaaca	caacattgcc	tttatgtgta	actcttggtc	gaagctctta	3060
caccaatgct	gggggacatg	tacctccag	ggggccagga	agactacggg	aggtacacac	3120
aacgtcaatc	agaggggctc	gtgtagctac	cgataagcgg	accctcaaga	gggcattagc	3180
aatagtgttt	ataagggccc	cttgtaaac	ctaaacgggt	agcatatgct	tcccgggtag	3240
tagtatatac	tatccagact	aaccctaatt	caatagcata	tgttacccaa	cgggaagcat	3300
atgctatcga	attaggggta	gtaaaagggt	cctaaggga	agcgatatct	cccaccccat	3360
gagctgtcac	ggttttattt	acatggggtc	aggattccac	gagggtagtg	aaccatttta	3420
gtcacaaggg	cagtggctga	agatcaaggga	gcgggcagtg	aactctctcg	aatcttcgcc	3480
tgttcttcca	ttctccttcg	tttagctaat	agaataactg	ctgagtgtgt	aacagtaagg	3540
tgtatgtgag	gtgctcgaaa	acaagggttc	aggtgacgcc	cccagaataa	aatttggaag	3600
gggggttcag	tgggtgcatt	gtgctatgac	accaatataa	ccctcacaaa	ccccttgggc	3660
aataaatact	agtgtaggaa	tgaaacattc	tgaatatctt	taacaataga	aatccatggg	3720
gtggggacaa	gccgtaaaga	ctggatgtcc	atctcacacg	aatttatggc	tatgggcaac	3780
acataatcct	agtgaatat	gatactgggg	ttattaagat	gtgtcccagg	cagggaccaa	3840
gacaggtgaa	ccatgttgtt	acactctatt	tgtaacaagg	ggaagagag	tggacgccga	3900
cagcagcgga	ctccactggt	tgtctctaac	accccgaaa	attaaacggg	gctccacgcc	3960
aatggggccc	ataaacaag	acaagtggcc	actctttttt	ttgaaattgt	ggagtggggg	4020

-continued

cacgcgtcag cccccacacg ccgccctgcg gttttggact gtaaaataag ggtgtaataa	4080
cttggctgat tgtaaccccg ctaaccactg cgggtcaaacc acttgccac aaaaccacta	4140
atggcaccce gggaataacc tgcataagta ggtgggcggg ccaagatagg ggcgcgattg	4200
ctgcgatctg gaggacaaat tacacacact tgcgctgag cgccaagcac aggggttgtt	4260
gtcctcatat tcacgaggtc gctgagagca cgggtgggcta atgttgccat gggtagcata	4320
tactacccaa atatctggat agcatatgct atcctaactc atatctgggt agcataggct	4380
atcctaactc atatctgggt agcatatgct atcctaactc atatctgggt agtatatgct	4440
atcctaactt atatctgggt agcataggct atcctaactc atatctgggt agcatatgct	4500
atcctaactc atatctgggt agtatatgct atcctaactc gtatccgggt agcatatgct	4560
atcctaatag agattagggt agtatatgct atcctaactt atatctgggt agcatatact	4620
acccaaatat ctggatagca tatgctatcc taatctatat ctgggtagca tatgctatcc	4680
taatctatat ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc	4740
taatctatat ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc	4800
taatctatat ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc	4860
taatctgtat cgggtagca tatgctatcc tcatgataag ctgtcaaaca tgagaatatt	4920
cttgaagacg aaaggccctc gtgatacgcc tatttttata ggtaaatgtc atgataataa	4980
tggtttctta gacgtcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt	5040
tatttttcta aatacattca aatatgtatc cgctcatgag acaataaccc tgataaatgc	5100
ttcaataata ttgaaaaagg aagagtatga gtattcaaca ttcccggtgc gcccttattc	5160
ccttttttgc ggcattttgc ctctctgttt ttgctcacc agaaacgctg gtgaaagtaa	5220
aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat ctcaacagcg	5280
gtaagatcct tgagagtttt cgcgccgaag aacgttttcc aatgatgagc acttttaag	5340
ttctgctatg tggcgcggtt ttatcccggt ttgacgcgg gcaagagcaa ctcggtcgcc	5400
gcatacacta ttctcagaat gacttggttg agtactcacc agtcacagaa aagcatctta	5460
cggatggcat gacagtaaga gaattatgca gtgctgccat aacctagagt gataaacctg	5520
cggccaactt acttctgaca acgatcggag gaccgaagga gctaaccgct tttttgcaca	5580
acatggggga tcatgttaact cgccttgatc gttgggaacc ggagctgaat gaagccatac	5640
caaacgacga gcgtgacacc acgatgcctg cagcaatggc aacaacgttg cgaaactat	5700
taactggcga actacttact ctagcttccc ggcaacaatt aatagactgg atggaggcgg	5760
ataaagttgc aggaccactt ctgcgctcgg cccttcgggc tggctggttt attgctgata	5820
aatctggagc cggtagcgtt ggggtctcgc gtatcattgc agcactgggg ccagatggta	5880
agccctcccg tatcgtagtt atctacacga cggggagtca ggcaactatg gatgaacgaa	5940
atagacagat cgctgagata ggtgcctcac tgattaagca ttggttaactg tcagaccaag	6000
tttactcata tatacttttag attgatttaa aacttcattt ttaatttaaa aggatctagg	6060
tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact	6120
gagcgtcaga ccccgtagaa aagatcaaaag gatcttcttg agatcctttt tttctgcgcg	6180
taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtgggttgt ttgccggatc	6240
aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgag ataccaataa	6300

-continued

ctgttcttct agttagccg tagttagcc accacttcaa gaactctgta gcaccgccta	6360
catacctcgc tctgctaata ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc	6420
ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg	6480
ggggttcgtg cacacagccc agcttgagc gaacgaccta caccgaactg agatacctac	6540
agcgtgagct atgagaaagc gccacgcttc ccgaaggag aaaggcggac aggtatccgg	6600
taagcggcag ggtcgaaca ggagagcgca cgaggagct tccaggggga aacgcctggt	6660
atctttatag tctgtcggg ttctgccacc tctgactga gcgtcgattt ttgtgatgct	6720
cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggcttttcta cggttcctgg	6780
ccttttgctg gccttttctg cacatgttct ttctgcgtt atccctgat tctgtggata	6840
accgtattac cgcttttag tgagctgata ccgctcgccg cagccgaacg accgagcgca	6900
gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaacgcct cccccgcgc	6960
gttggccgat tcattaatgc agctggcagc acaggtttcc cgactggaaa gcgggcagtg	7020
agcgcaacgc aattaatgtg agttagctca ctcataggc accccaggct ttacacttta	7080
tgcttcggc tcgtatgttg tgtggaattg tgagcggata acaatttcac acaggaaaca	7140
gctatgacca tgattacgc aagctctagc tagaggtcga gtccctcccc agcaggcaga	7200
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgccct aactccgcc	7260
atcccgcccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt	7320
tttatttatg cagagggcga ggccgctcg gcctctgagc tattccagaa gtagtgagga	7380
ggcttttttg gaggcctagg cttttgcaaa aagctttgca aagatggata aagttttaa	7440
cagagaggaa tctttgcagc taatggacct tctaggtctt gaaagg	7486

<210> SEQ ID NO 13

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP170 ; pHybE-hCgl,z,a V1

<400> SEQUENCE: 13

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagtgtggg ggagggtcgc gcaattgaac cgggtcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctcgc cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgcgc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgct	300
gccttgaaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360
aagtgggtgg gagagttcga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcgcgcg cgtgcgaatc tgggtggacc ttcgcgcctg	480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gggggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgctt ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggtggccc	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggcctgctg caggagctc	840

-continued

aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcacccacac aaaggaaaag	900
ggcctttccg tcctcagccg tcgcttcagtg tgactccacg gactaccggg cgccgtccag	960
gcacctcgat tagttctega gcttttggag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc acactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaaggt gtccagtgcg catggtatgc cgaagggat	1320
gctgaaattg agaacgaaaa gctgcgcggg gaggttgaag aactgcggca ggcagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaaac cgcattctgt	1500
actttcgtgc tgctcgggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagccgcg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggcttcgctc	1740
actgttcagg ccggagccac agaccgcgt tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaacc acgttgagcc	2040
gactattcgt gatattccgt cgctgctggc gctggccccg tggatatgga aaaagcaccg	2100
ggataacacg ctcaccatga agcgtttcac taatgggctt ggcttctggg gcctggcg	2160
taaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgttttggat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgct	2280
cgaccaaggg cccatcggtc ttccccctgg caccctctc caagagcacc tctgggggca	2340
cagcgccctt gggctgcctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc cctgaccagc ggcgtgcaca ccttccccgc tgtcctacag tctcaggac	2460
tctactcctt cagcagcggt gtgaccgtgc cctccagcag cttgggcacc cagacctaca	2520
tctgcaacgt gaatcacaag ccagcaaca ccaagggtga caagaaagt gagcccaat	2580
cttgtgacaa aactcacaca tgcccaccgt gccagcacc tgaactcctg gggggaccgt	2640
cagtcttctt cttccccca aaaccaagg acacctcat gatctcccg acccctgagg	2700
tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagttc aactggtacg	2760
tggacggcgt ggaggtgcat aatgccaa gaagcccgcg ggaggagcag tacaacagca	2820
cgtaccgtgt ggtcagcgtc ctcacgtcc tgcaccagga ctggctgaat ggcaaggagt	2880
acaagtgc aa ggtctccaac aaagccctcc cagcccccat cgagaaaacc atctccaaag	2940
ccaaagggca gccccgagaa ccacagggtg acacctgcc cccatcccg gatgagctga	3000
ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatccagc gacatcgccg	3060
tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct cccgtgctgg	3120

-continued

actccgacgg	ctccttcttc	ctctacagca	agctcaccgt	ggacaagagc	aggtggcagc	3180
aggggaacgt	cttctcatgc	tccgtgatgc	atgaggetct	gcacaaccac	tacacgcaga	3240
agagcctctc	cctgtctccg	ggtaaagtag	cggccgctcg	aggccggcaa	ggccggatcc	3300
cccgaacctg	acctctggct	aataaaggaa	atttattttc	attgcaatag	tgtgttgga	3360
ttttttgtgt	ctctcactcg	gaaggacata	tgggagggca	aatcatttgg	tgcagatccc	3420
tccgagatct	ctagctagag	gatcgatccc	cgcgccggac	gaactaaacc	tgactacgac	3480
atctctgccc	cttcttcgcg	gggcagtgca	tgtaatccct	tcagttgggt	ggtacaactt	3540
gccaaactggg	ccctgttcca	catgtgacac	ggggggggac	caaacacaaa	ggggttctct	3600
gactgtagtt	gacatcctta	taaatggatg	tgacacattg	ccaacactga	gtggctttca	3660
tcttgagca	gactttgcag	tctgtggact	gcaacacaac	attgccttta	tgtgtaactc	3720
ttggctgaag	ctcttacacc	aatgtggggg	gacatgtacc	tcccaggggc	ccaggaagac	3780
tacgggaggg	tacaccaacg	tcaatcagag	gggcctgtgt	agctaccgat	aagcggaccc	3840
tcaagagggc	attagcaata	gtgtttataa	ggcccccttg	ttaaccctaa	acgggtagca	3900
tatgcttccc	gggtagtagt	atatactatc	cagactaacc	ctaattcaat	agcatatgtt	3960
acccaacggg	aagcatatgc	tatcgaatta	gggttagtaa	aagggtccta	aggaacagcg	4020
atatctccca	ccccatgagc	tgtcacgggt	ttattttacat	ggggtcagga	ttccacgagg	4080
gtagtgaacc	attttagtca	caagggcagt	ggctgaagat	caaggagcgg	gcagtgaact	4140
ctctgaatc	tctgcctgct	tcttcattct	ccttcgttta	gctaatagaa	taactgctga	4200
gttgtgaaca	gtaagggtga	tgtgaggtgc	tcgaaaacaa	ggtttcaggt	gacgccccca	4260
gaataaaaatt	tggacggggg	gttcagtggg	ggcattgtgc	tatgacacca	atataaccct	4320
cacaaacccc	tggggcaata	aatactagt	taggaatgaa	acattctgaa	tatctttaac	4380
aatagaaatc	catgggggtg	ggacaagccg	taaagactgg	atgtccatct	cacacgaatt	4440
tatggctatg	ggcaacacat	aatcctagt	caatatgata	ctgggggttat	taagatgtgt	4500
cccaggcagg	gaccaagaca	ggtgaacat	gttggttacac	tctatttgta	acaaggggaa	4560
agagagtggg	cgcgcacagc	agcggactcc	actggttgtc	tctaacaccc	ccgaaaatta	4620
aacggggctc	cacgccaatg	gggcccataa	acaaagacaa	gtggccactc	ttttttttga	4680
aattgtggag	tgggggcacg	cgtcagcccc	cacacgccgc	cctgcggttt	tggactgtaa	4740
aataaggggtg	taataacttg	gctgattgta	accccgctaa	ccactgcggg	caaaccactt	4800
gcccacaaaa	ccactaatgg	caccccgggg	aatacctgca	taagtaggtg	ggcgggccaa	4860
gatagggggc	cgattgctgc	gatctggagg	acaaattaca	cacacttgcg	cctgagcgcc	4920
aagcacaggg	ttgttggtcc	tcatattcac	gaggtcgctg	agagcacggg	gggctaattg	4980
tgccatgggt	agcatatact	acccaaatat	ctggatagca	tatgctatcc	taatctatat	5040
ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	5100
ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	taggctatcc	taatctatat	5160
ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatctgtat	5220
ccgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatcctaata	ctatatctgg	5340
gtagcatatg	ctatcctaata	ctatatctgg	gtagcatagg	ctatcctaata	ctatatctgg	5400
gtagcatatg	ctatcctaata	ctatatctgg	gtagtatatg	ctatcctaata	ttatatctgg	5460

-continued

gtagcatagg ctatccta	ctatatctgg gtagcatatg ctatccta	ctatatctgg	5520
gtagtatatg ctatccta	ctgtatccgg gtagcatatg ctatccta	gataagctgt	5580
caaacatgag aatctcttg	aagacgaaag ggctctgga tacgcctatt	tttataggtt	5640
aatgtcatga taataatgtt	ttcttagacg tcaggtggca cttctcgagg	aaatgtgcgc	5700
ggaacccta tttgtttatt	tttctaaata cattcaaata tgtatccgct	catgagacaa	5760
taacctgat aaatgcttca	ataatattga aaaaggaaga gtatgagtat	tcaacatttc	5820
cgtgtcgccc ttattccctt	ttttcgaggc ttttgccttc ctgtttttgc	tcaccagaa	5880
acgctggtga aagtaaaaga	tgctgaagat cagttgggtg cagcagtggtg	ttacatcgaa	5940
ctggatctca acagcggtaa	gatccttgag agttttcgcc ccgaagaacg	ttttccaatg	6000
atgagcactt ttaaagtctt	gctatgtggc gcggtattat cccgtgttga	cgccgggcaa	6060
gagcaactcg gtcgcgcgat	acactattct cagaatgact tggttgagta	ctcaccagtc	6120
acagaaaagc atcttacgga	tggcatgaca gtaagagaat tatgcagtgc	tgccataacc	6180
atgagtata acactgcgcg	caacttactt ctgacaacga tcggaggacc	gaaggagcta	6240
accgcttttt tgcacaacat	gggggatcat gtaactcgcc ttgatcgttg	ggaaccggag	6300
ctgaatgaag ccataccaaa	cgacgagcgt gacaccacga tgctgcagc	aatggcaaca	6360
acgttgcgca aactattaac	tggcgaacta cttactctag cttcccgga	acaattaata	6420
gactggatgg aggcgataa	agttgcagga ccacttctgc gctcgccct	tccggctggc	6480
tggttttatt ctgataaatc	tggagccggt gagcgtgggt ctcgcggtat	cattgcagca	6540
ctggggccag atggtaagcc	ctcccgatc gtagttatct acacgacggg	gagtcaggca	6600
actatggatg aacgaaatag	acagatcgct gagataggtg cctcactgat	taagcattgg	6660
taactgtcag accaagttta	ctcatatata ctttagattg atttaaaact	tcatttttaa	6720
tttaaaagga tctaggtgaa	gatccttttt gataatctca tgacaaaat	cccttaacgt	6780
gagttttcgt tccactgagc	gtcagacccc gtagaaaaga tcaaaggatc	ttcttgagat	6840
cctttttttc tgcgcgtaat	ctgctgcttg caaacaaaaa aaccaccgct	accagcggtg	6900
gtttgtttgc cggatcaaga	gctaccaact ctttttcga aggttaactgg	cttcagcaga	6960
gcgcagatac caaatactgt	tcttctagtg tagccgtagt taggccacca	cttcaagaac	7020
tctgtagcac cgctacata	cctcgctctg ctaatcctgt taccagtggc	tgtgccagt	7080
ggcgataagt cgtgtcttac	cgggttgac tcaagacgat agttaccgga	taaggcgag	7140
cggtcgggct gaacggggg	ttcgtgcaca cagcccagct tggagcgaac	gacctacac	7200
gaactgagat acctacacg	tgagctatga gaaagcgcca cgcttccga	agggagaaag	7260
gcggacaggt atccggaag	cggcagggtc ggaacaggag agcgcacgag	ggagcttcca	7320
gggggaaacg cctggtatct	ttatagtcct gtcgggtttc gccacctctg	acttgagcgt	7380
cgatttttgt gatgctcgtc	aggggggcgg agcctatgga aaaacgccag	caacgcggcc	7440
tttttacggt tcctggcctt	ttgctggcct tttgctcaca tgttctttcc	tgcggtatcc	7500
cctgattctg tgataaacc	tattaccgcc tttgagttag ctgataccgc	tcgccgcagc	7560
cgaacgaccg agcgcagcga	gtcagttagc gaggaagcgg aagagcgccc	aatacgcaaa	7620
ccgcctctcc ccgcgcgtg	gccgattcat taatgcagct ggcacgacag	gtttcccgac	7680
tggaagcgg gcagttagcg	caacgcaatt aatgtgagtt agctcactca	ttaggcaccc	7740

-continued

caggctttac actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa	7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtc	7860
ctccccagca ggcagaagta tgcaaaagcat gcattctcaat tagtcagcaa ccatagtccc	7920
gcccctaact cgcgccatcc cgcgccctaac tccgccagc tccgccatt ctcgccccca	7980
tggctgacta atttttttta tttatgcaga ggcgaggcc gcctcgccct ctgagctatt	8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaga	8100
tgataaagt ttttaacaga gaggaatctt tgcagctaata ggaccttcta ggtcttga	8160
gg	8162

<210> SEQ ID NO 14

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP182 ; pHybE-hCg1,z,a V2

<400> SEQUENCE: 14

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagtgtggg ggagggtcg gcaattgaac cgggtcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgcc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggt cccgcgggc tggcctcttt acgggttatg gcccttgct	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360
aagtgggtgg gagagtgcga ggccttgcc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgtg ggggcccgc cgtgcgaatc tgggtggcacc ttcgcgcctg	480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gggggcgccg acggggccc tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgctt ggcctgcgc gcgcgtgtat cgcgccgcc tgggcggcaa ggtggcccg	780
gtcggcacca gttgcgtgag cggaaagatg gcccttccc gccctgctg caggagctc	840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tccctagccg tcgcttcattg tgactccacg gaggaccgg cgcgctccag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggagggtt	1020
ttatgcgatg gattttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggtcttgggt tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttct tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgaga tccattgtgc cggggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaagg gtccagtgcg catggtatgc cgaagggat	1320
gctgaaattg agaacgaaa gctgcgcgg gaggttgaag aactgcgga gccagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttctgtc tgtcgcggat cgcagggtga attgccagta ttctcgacgg gctccccctg	1560

-continued

tcggtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttccct	gaaacgggat	1620
atcatcaaag	ccatgaacaa	agcagcccg	ctggatgaac	tgataccggg	gttgctgagt	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgcgc	ggcttcgctc	1740
actgttcagg	ccggagccac	agaccgccgt	tgaatgggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaaacactg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggtgaatgtg	gtgaagtctg	cccggtgcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccgacgg	atggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggccccg	tggtatggca	aaaagcaccg	2100
ggataaacacg	ctcaccatga	agcgtttcac	taatgggcgt	ggcttctggt	gcctgggcgg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggtggatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttcctgggtg	acaagcgcg	2280
cgaccaaggg	cccatcggtc	ttccccctgg	caccctctc	caagagcacc	tctgggggca	2340
cagcggccct	gggctgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	cctgaccagc	ggcgtgcaca	ccttccccgc	tgtcctacag	tctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtgc	cctccagcag	cttgggcacc	cagacctaca	2520
tctgcaacgt	gaatcacaa	cccagcaaca	ccaagggtga	caagaaagtt	gagcccaa	2580
cttgtgacaa	aactcacaca	tgccccaccg	gcccagcacc	tgaactcctg	gggggaccgt	2640
cagtcttctc	cttcccccca	aaacccaagg	acacctcat	gatctcccgg	accctgagg	2700
tcacatgcgt	ggtggtggac	gtgagccacg	aagaccctga	ggtcaagttc	aactggtacg	2760
tggaaggcgt	ggaggtgcat	aatgccaaga	caaagcccg	ggaggagcag	tacaacagca	2820
cgtaccgtgt	ggtcagcgtc	ctcaccgtcc	tgaccagga	ctggctgaat	ggcaaggagt	2880
acaagtgcaa	ggtctccaac	aaagccctcc	cagcccccat	cgagaaaacc	atctccaaag	2940
ccaaagggca	gccccgagaa	ccacagggtg	acacctctgc	cccatcccgg	gatgagctga	3000
ccaagaacca	ggtcagcctg	acctgcctgg	tcaaaggctt	ctatcccagc	gacatcgccg	3060
tggaagtggga	gagcaatggg	cagccggaga	acaactacaa	gaccacgcct	cccggtgctg	3120
actccgacgg	ctccttcttc	ctctacagca	agctcaccgt	ggacaagagc	aggtggcagc	3180
aggggaacgt	cttctcatgc	tccgtgatgc	atgaggctct	gcacaaccac	tacacgcaga	3240
agagcctctc	cctgtctccg	ggtaaatgag	cggccgctcg	aggccggcaa	ggccgcatcc	3300
cccgacctcg	acctctggct	aataaaggaa	atttattttc	attgcaatag	tgtgttgga	3360
ttttttgtgt	ctctcactcg	gaaggacata	tgggagggca	aatcatttgg	tcgagatccc	3420
tcggagatct	ctagctagag	gatcgatccc	cgccccggac	gaactaaacc	tgactacgac	3480
atctctgccc	cttcttcg	gggcagtgca	tgtaatccct	tcagttggtt	ggtacaactt	3540
gccaaactgg	ccctgttcca	catgtgacac	ggggggggac	caaacacaaa	ggggttctct	3600
gactgtagtt	gacatcctta	taaatggatg	tgacatttgg	ccaacactga	gtggctttca	3660
tcttgagca	gactttgcag	tctgtggact	gcaacacaa	attgccttta	tgtgtaactc	3720
ttggctgaag	ctcttacacc	aatgtgggg	gacatgtacc	tcccaggggc	ccaggaagac	3780
tacgggaggg	tacaccaacg	tcaatcagag	gggcctgtgt	agctaccgat	aagcggaccc	3840

-continued

tcaagagggc	attagcaata	gtgtttataa	ggcccccttg	ttaaccctaa	acgggtagca	3900
tatgcttccc	gggtagtagt	atatactatc	cagactaacc	ctaattcaat	agcatatgtt	3960
acccaacggg	aagcatatgc	tatcgaatta	gggttagtaa	aagggtccta	aggaacagcg	4020
atatctccca	ccccatgagc	tgtcacgggt	ttatttacat	ggggtcagga	ttccacgagg	4080
gtagtgaacc	attttagtca	caagggcagt	ggctgaagat	caaggagcgg	gcagtgaact	4140
ctctgaatc	ttcgctgct	tcttcattct	ccttcgttta	gctaatagaa	taactgctga	4200
gttgtgaaca	gtaagggtga	tgtaggtg	tcgaaaacaa	ggtttcaggt	gacgccccca	4260
gaataaaatt	tggacggggg	gttcagtggt	ggcattgtgc	tatgacacca	atataaccct	4320
cacaaacccc	ttgggcaata	aatactagt	taggaatgaa	acattctgaa	tatctttaac	4380
aatagaaatc	catgggggtg	ggacaagccg	taaagactgg	atgtccatct	cacacgaatt	4440
tatggctatg	ggcaacacat	aatcctagt	caatatgata	ctgggggttat	taagatgtgt	4500
cccagcgagg	gaccaagaca	ggtgaaccat	gttgttacac	tctatttgta	acaaggggaa	4560
agagagtgga	cgcgcacagc	agcggactcc	actggttgtc	tctaaccacc	ccgaaaatta	4620
aacggggctc	cacgccaatg	gggcccataa	acaagacaa	gtggccactc	ttttttttga	4680
aattgtggag	tgggggcacg	cgtcagcccc	cacacgccgc	cctgcggttt	tggactgtaa	4740
aataaggggtg	taataacttg	gctgattgta	accccgctaa	ccactgcggt	caaaccactt	4800
gcccacaaaa	ccactaatgg	caccccgggg	aatacctgca	taagtaggtg	ggcgggccaa	4860
gataggggag	cgattgctgc	gatctggagg	acaaattaca	cacacttgcg	cctgagcgcc	4920
aagcacaggg	ttgttggtcc	tcatattcac	gaggctcgctg	agagcacggt	gggctaattgt	4980
tgccatgggt	agcatatact	acccaaatat	ctggatagca	tatgctatcc	taatctatat	5040
ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	5100
ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	taggctatcc	taatctatat	5160
ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatctgtat	5220
ccgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatccta	ctatatctgg	5340
gtagcatatg	ctatccta	ctatatctgg	gtagcatagg	ctatccta	ctatatctgg	5400
gtagcatatg	ctatccta	ctatatctgg	gtagtatatg	ctatccta	ttatatctgg	5460
gtagcatagg	ctatccta	ctatatctgg	gtagcatatg	ctatccta	ctatatctgg	5520
gtagtatatg	ctatccta	ctgtatccgg	gtagcatatg	ctatcctcat	gataagctgt	5580
caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttataggtt	5640
aatgtcatga	taataatggt	ttcttagacg	tcagggtggca	cttttcgggg	aaatgtgcgc	5700
ggaaccccta	tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5760
taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5820
cgtgtcgccc	ttattccctt	ttttcgggca	ttttgccttc	ctgtttttgc	tcaccagaa	5880
acgctgggtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggtg	ttacatcgaa	5940
ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	6000
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtgttga	cgccgggcaa	6060
gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6180

-continued

```

atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 6240
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcggtt ggaaccggag 6300
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgctgcagc aatggcaaca 6360
acgttgcgca aactattaac tggcgaacta cttactctag cttcccgga acaattaata 6420
gactggatgg aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc 6480
tggtttattg ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca 6540
ctggggccag atggtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca 6600
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 6660
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 6720
tttaaaagga tctaggtgaa gatccttttt gataatctca tgacaaaat cccttaacgt 6780
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 6840
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcgggtg 6900
gtttgtttgc cggatcaaga gctaccaact ctttttcga aggtaactgg cttcagcaga 6960
gcgcagatac caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac 7020
tctgtagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgcag 7140
cggtcgggct gaacgggggg ttcgtgcaca cagcccgct tggagcgaac gacctacac 7200
gaactgagat acctacacg tgagctatga gaaagcgcca cgcttccga agggagaaag 7260
gcggacaggt atccggaag cggcagggtc ggaacaggag agcgacgag ggagcttcca 7320
gggggaaacg cctggtatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt 7380
cgatttttgt gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc 7440
tttttaacgt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc 7500
cctgattctg tggataaacc tattaccgcc tttgagttag ctgataccgc tcgccgcagc 7560
cgaaacgacc agcgcagcga gtcagtgagc gaggaagcgg aagagcgcgc aatacgcaaa 7620
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 7680
tgaaagcgg gcagtgagc caacgcaatt aatgtgagtt agctcactca ttaggcaccc 7740
caggctttac actttatgct tccggctcgt atgttggtg gaattgtgag cggataaaca 7800
ttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc 7860
ctcccagca ggcagaagta tgcaaagcat gcctctcaat tagtcagcaa ccatagtccc 7920
gcccctaact ccgcccattc cgcccctaac tccgcccagt tccgcccatt ctcgcccaca 7980
tggtgacta atttttttta tttatgcaga ggccgaggcc gcctcgccct ctgagctatt 8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaga 8100
tgataaagt tttaaacaga gaggaatctt tgcagctaag ggaccttcta ggtcttgaaa 8160
gg

```

<210> SEQ ID NO 15

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP171 ; pHybE-hCg1,z,
non-a V1

-continued

<400> SEQUENCE: 15

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtcgccga	60
gaagttgggg ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggtcccg cctttttccc gaggggtggg gagaaccgta	180
tataagtgca gtagtcgccc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctt	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg	360
aagtgggtgg gagagttcga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcgcgag cgtgcgaatc tgggggacc ttcgcgcctg	480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcgggcgccg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgcct ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggctggcccg	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc	840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tccctcagccg tcgcttcagt tgactccacg gaggaccggg cgccttcag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttctt tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaagg gtccagtgcg catggtatgc cgaaagggat	1320
gctgaaattg agaacgaaaa gctgcgccgg gaggttgaag aactgcggca ggcagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgag tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaac cgcattctgt	1500
actttcgtgc tgtcgcggat cgcagggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaag ccatgaacaa agcagcccg cgggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtcgcgccg ggettcgctc	1740
actgttcagg ccggagccac agaccgccgt tgaatgggag gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc	2040
gactattcgt gatattccgt cgtcgtggc gctggccccg tggatatggc aaaagcaccg	2100
ggataaacag ctcacatga agcgtttcac taatgggcgt ggcttctggg gcctgggcgg	2160
taaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220

-continued

tgcttttgat gatgatattg aacaggaagg ctctccgacg ttcttgggtg acaagcgcgt	2280
cgaccaaggg cccatcggtc ttccccctgg caccctcttc caagagcacc tctgggggca	2340
cagcggccct gggctgcctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc cctgaccagc ggcgtgcaca ccttccccgc tgtctacag tctcaggac	2460
tctactccct cagcagcgtg gtgacgtgc cctccagcag cttgggcacc cagacctaca	2520
tctgcaacgt gaatcacaag ccagcaaca ccaaggtgga caagaaagtt gagcccaat	2580
cttgtgacaa aactcacaca tgcccaccgt gccagcacc tgaactcctg gggggaccgt	2640
cagtcttctt cttccccca aaacccaagg acacctctat gatctcccg acccctgagg	2700
tcacatgcgt ggtggtggac gtgagccagc aagacctga ggtcaagttc aactggtacg	2760
tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag tacaacagca	2820
cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat ggcaaggagt	2880
acaagtgcga ggtctccaac aaagccctcc cagcccccat cgagaaaacc atctccaaag	2940
ccaaagggca gccccagaa ccacagggtt acacctgcc cccatccgc gaggagatga	3000
ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatccagc gacatcgccg	3060
tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgctt cccgtgctgg	3120
actccgacgg ctctctcttc ctctacagca agctcaccgt ggacaagagc aggtggcagc	3180
aggggaacgt cttctcatgc tccgtgatgc atgaggtctt gcacaaccac tacacgcaga	3240
agagcctctc cctgtctccg ggtaaatgag cggccgctcg aggcggcaa ggcggatcc	3300
cccgaacctg acctctggct aataaaggaa atttattttt attgcaatag tgtgttgaa	3360
tttttttgtt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctagctagag gatcgatccc cggcccgac gaactaaacc tgactacgac	3480
atctctgccc cttcttcgcg gggcagtgca tgtaatccct tcagttggtt ggtacaactt	3540
gccaaactgg ccctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600
gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tcttgagca gactttgcag tctgtggact gcaacacaa attgccttta tgtgtaactc	3720
ttggctgaag ctcttacacc aatgctgggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggaggg tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa ggcccccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtagtaa aagggtccta aggaacagcg	4020
atatctccca cccatgagc tgtcacggtt ttatttacat ggggtcagga ttccacgagg	4080
gtagtgaacc attttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctctgaatc ttgcctgct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaaggtgta tgtgaggtgc tcgaaaacaa ggtttcaggt gacgccccca	4260
gaataaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaacccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catggggtgg ggacaagccg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctggggttat taagatgtgt	4500
cccagcagg gaccaagaca ggtgaaccat gttgttacac tctatttga acaaggggaa	4560

-continued

agagagtgga	cgccgacagc	agcggactcc	actggttgtc	tctaacaccc	ccgaaaatta	4620
aacggggctc	cacgccaatg	gggcccataa	acaaagacaa	gtggccactc	ttttttttga	4680
aattgtggag	tgggggcacg	cgtcagcccc	cacacgcgcg	cctgcggttt	tggactgtaa	4740
aataagggtg	taataacttg	gctgattgta	accccgctaa	ccactgcggg	caaaccactt	4800
gccacaaaa	ccactaatgg	caccccgggg	aatacctgca	taagtaggtg	ggcgggccaa	4860
gataggggag	cgattgctgc	gatctggagg	acaaattaca	cacacttgcg	cctgagcgcc	4920
aagcacaggg	ttgttggtcc	tcataattac	gaggtcgctg	agagcacggg	gggctaattg	4980
tgccatgggt	agcatatact	acccaaatat	ctggatagca	tatgctatcc	taatctatat	5040
ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	5100
ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	taggctatcc	taatctatat	5160
ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatctgtat	5220
ccgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatcctaata	ctatatctgg	5340
gtagcatatg	ctatcctaata	ctatatctgg	gtagcatagg	ctatcctaata	ctatatctgg	5400
gtagcatatg	ctatcctaata	ctatatctgg	gtagtatatg	ctatcctaata	ttatatctgg	5460
gtagcatagg	ctatcctaata	ctatatctgg	gtagcatatg	ctatcctaata	ctatatctgg	5520
gtagtatatg	ctatcctaata	ctgtatccgg	gtagcatatg	ctatcctaata	gataagctgt	5580
caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	5640
aatgtcatga	taataatggt	ttcttagacg	tcagggtggca	cttttcgggg	aaatgtgcgc	5700
ggaaccccta	tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5760
taacctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5820
cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaaa	5880
acgctgggtg	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggt	ttacatcgaa	5940
ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	6000
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtgttga	cgccgggcaa	6060
gagcaactcg	gtcgccgcat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	tgcataaacc	6180
atgagtgata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	6240
accgcttttt	tgacacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	6300
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgcagc	aatggcaaca	6360
acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccggca	acaattaata	6420
gactggatgg	agggcgataa	agttgcagga	ccacttctgc	gctcgccctt	tccggctggc	6480
tggtttattg	ctgataaatc	tggagccggt	gagcgtgggt	ctcgcggtat	cattgcagca	6540
ctggggccag	atggtaagcc	ctcccgtatc	gtagtatatc	acacgacggg	gagtcaggca	6600
actatggatg	aacgaaaatg	acagatcgct	gagatagggt	cctcactgat	taagcattgg	6660
taactgtcag	accaagttta	ctcatatata	cttttagattg	atttaaaact	tcatttttaa	6720
tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacaaaaat	cccttaacgt	6780
gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	6840

-continued

cctttttttc tgcgcgtaat ctgctgcttg caaacaaaa aaccaccgct accagcggtg	6900
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga	6960
gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca cttcaagaac	7020
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag	7140
cggtcgggct gaacggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	7200
gaactgagat acctacacgc tgagctatga gaaagcgcca cgcttccga agggagaaa	7260
gcggacaggt atccggttaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca	7320
gggggaaacg cctggatatc ttatagtctc gtcgggttcc gccacctctg acttgagcgt	7380
cgatttttgt gatgctcgtc agggggggcg agcctatgga aaaacgccag caacgcggcc	7440
tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc	7500
cctgattctg tgataaacg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	7560
cgaacgaccg agcgcagcga gtcagttagc gaggaagcgg aagagcgccc aatacgcaaa	7620
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac	7680
tgaaagcgg gcagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc	7740
caggctttac actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa	7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc	7860
ctcccagca gccagaagta tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc	7920
gcccctaact ccgcccctcc cgccttaac tccgccagt tccgccatt ctccgccc	7980
tggtgacta atttttttta tttatgcaga ggccgaggcc gcctcgccct ctgagctatt	8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaga	8100
tgataaaagt tttaaacaga gaggaatctt tgcagctaag ggaccttcta ggtcttga	8160
gg	8162

<210> SEQ ID NO 16

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP183 ; pHybE-hCgl,z, non-a V2

<400> SEQUENCE: 16

agtgggaatt ggctccgggt cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaagggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggtccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgcc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360
aagtgggtgg gagagtcca ggccttcgc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgt ggggcgcgc cgtgcgaatc tgggtggcacc ttcgcgcctg	480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600

-continued

ttttggggcc gcgggcgggc acggggcccc tgcgtcccag cgcacatgtt cggcgaggcg	660
gggctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgect ggectcgcg cgccgtgtat cgcgccgcc tggcgggcaa ggctggcccc	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc	840
aaaaatggagg acgcggcgct cgggagagcg ggcgggtgag taccacac aaaggaaaag	900
ggcctttccg tcctcagcgg tcgcttcag tgactccacg gactaccggg cgccgtccag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggagggggt	1020
ttatgcgatg gaggttcccc acactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgcctt ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgctgtgag gaattctcta	1200
gagatccctc gacctcgaga tccattgtgc ccgggcgcca ccatggagtt tgggtgagc	1260
tggctttttc ttgtcgcgat tttaaaaggt gtccagtcg catggtatgc cgaaaggat	1320
gctgaaattg agaacgaaaa gctgcgcgg gaggttgaag aactgcggca ggccagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacggg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaac cgcatctgt	1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagcccg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggcttcgctc	1740
actgttcagg ccggagccac agaccgccgt tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaacc acgttgagcc	2040
gactattcgt gatattccgt cgtgctggc gctggccccg tggatggca aaaagcaccg	2100
ggataaacag ctcaccatga agcgtttcac taatgggctg ggcttctggt gcctggcg	2160
taaagcgga aaaaactacc gtgaaaatc ggtggatgtg gcgggttatg atgaacttc	2220
tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgct	2280
cgaccaaggg cccatcggtc tccccctgg caccctctc caagagcacc tctgggggca	2340
cagcgccctt gggctgcctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc cctgaccagc ggcgtgcaca ccttccccgc tgtcctacag tctcaggac	2460
tctactccct cagcagcgtg gtgaccgtgc cctccagcag cttgggcacc cagacctaca	2520
tctgcaacgt gaatcacaag ccagcaaca ccaagggtga caagaaagt gagcccaat	2580
cttgtagaca aactcacaca tgcccaccgt gccagcacc tgaactcctg gggggaccgt	2640
cagtcttctt cttcccccca aaacccaagg acaccctcat gatctcccg acccctgagg	2700
tcacatgcgt ggtggtggac gtgagccagc aagaccctga ggtcaagttc aactggtacg	2760
tggacggcgt ggaggtgcat aatgccaa caaagcccg ggaggagcag tacaacagca	2820
cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat ggcaaggagt	2880
acaagtgcga ggtctccaac aaagccctcc cagcccccat cgagaaaacc atctccaaag	2940

-continued

ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgc gaggagatga	3000
ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc gacatcgccg	3060
tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct cccgtgctgg	3120
actccgacgg ctctctcttc ctctacagca agctcaccgt ggacaagagc aggtggcagc	3180
aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac tacacgcaga	3240
agagcctctc cctgtctccg ggtaaatgag cggccgctcg aggccggcaa ggccggatcc	3300
cccgaacctc acctctggct aataaaggaa atttattttc attgcaatag tgtgttgga	3360
ttttttgtgt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctagttagag gatcgatccc cgccccggac gaactaaacc tgactacgac	3480
atctctgccc cttcttcgcg gggcagtgca tgtaatccct tcagttgggt ggtacaactt	3540
gccaaactgg cctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600
gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tcttgagca gactttgag tctgtggact gcaacacaac attgccttta tgtgtaactc	3720
ttggctgaag ctcttacacc aatgctggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggaggc tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa ggcccccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtttagta aagggtccta aggaacagcg	4020
atatctccca ccccatgagc tgtcacgggt ttatttacat ggggtcagga ttccacgagg	4080
gtagtgaacc attttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctoctgaatc ttgcctgct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaagggtga tgtgagggtc tcgaaaacaa ggtttcagggt gacgccccca	4260
gaataaaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaaaccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catggggtgg ggacaagccg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctgggggttat taagatgtgt	4500
cccaggcagg gaccaagaca ggtgaaccat gttgttacac tctatttgta acaaggggaa	4560
agagagtgga cgcgcacagc agcggactcc actggttgtc tctaacaccc ccgaaaatta	4620
aacggggctc cagccaatg gggcccataa acaaagacaa gtggccactc ttttttttga	4680
aattgtggag tgggggcacg cgtcagcccc cacacgccgc cctgcggttt tggactgtaa	4740
aataagggtg taataacttg gctgattgta accccgctaa ccactgcggc caaaccactt	4800
gcccacaaaa ccactaatgg cccccgggg aatacctgca taagtagggt ggccggccaa	4860
gataggggag cgattgctgc gatctggagg acaaattaca cacacttgcg cctgagcgcc	4920
aagcacaggg ttgttggtcc tcatattcac gaggtcgtg agagcacggg gggctaatgt	4980
tgccatgggt agcatatact acccaaatat ctggatagca tatgctatcc taatctatat	5040
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat	5100
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat	5160
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat	5220

-continued

cgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatcctaata	ctatatctgg	5340
gtagcatatg	ctatcctaata	ctatatctgg	gtagcatagg	ctatcctaata	ctatatctgg	5400
gtagcatatg	ctatcctaata	ctatatctgg	gtagtatatg	ctatcctaata	ttatatctgg	5460
gtagcatagg	ctatcctaata	ctatatctgg	gtagcatatg	ctatcctaata	ctatatctgg	5520
gtagtatatg	ctatcctaata	ctgtatccgg	gtagcatatg	ctatcctaata	gataagctgt	5580
caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttataggtt	5640
aatgtcatga	taataatggt	ttcttagacg	tcaggtagga	cttttcgggg	aaatgtgcgc	5700
ggaacccta	ttgttttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5760
taacctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5820
cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaa	5880
acgtcgtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggt	ttacatcgaa	5940
ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	cgaagaacg	ttttccaatg	6000
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtgttga	cgccgggcaa	6060
gagcaactcg	gtcgccgcat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6180
atgagtata	acactcgggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	6240
accgcttttt	tgcaacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	6300
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgcagc	aatggcaaca	6360
acgttcgca	aactattaac	tggcgaacta	cttactctag	cttcccgga	acaattaata	6420
gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcgccct	tcgggctggc	6480
tggtttattg	ctgataaata	tggagccggt	gagcgtgggt	ctcgcggtat	cattgcagca	6540
ctggggccag	atggtaagcc	ctcccgatc	gtagttatct	acacgacggg	gagtcaggca	6600
actatggatg	aacgaaatag	acagatcgct	gagatagggt	cctcactgat	taagcattgg	6660
taactgtcag	accaagttta	ctcatatata	cttttagattg	atttaaaact	tcatttttaa	6720
tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacaaaaat	cccttaacgt	6780
gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	6840
cctttttttc	tgccgctaata	ctgctgcttg	caaacaaaaa	aaccaccgct	accagcgggtg	6900
gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	cttcagcaga	6960
gcgcagatac	caaatactgt	tcttctagt	tagccgtagt	taggccacca	cttcaagaac	7020
tctgtagcac	cgctacata	cctcgctctg	ctaactcctgt	taccagtggc	tgctgccagt	7080
ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	taaggcgag	7140
cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggagcgaac	gacctacacc	7200
gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgcttccga	aggagagaaag	7260
gcggacaggt	atccggtaag	cggcagggtc	ggaacaggag	agcgcacgag	ggagcttcca	7320
gggggaaacg	cctggtatct	ttatagtctt	gtcgggtttc	gccacctctg	acttgagcgt	7380
cgatttttgt	gatgctcgct	agggggggcg	agcctatgga	aaaacgccag	caacgcggcc	7440
tttttacggt	tcctggcctt	ttgctggcct	tttgcacaca	tgttctttcc	tgcggtatcc	7500
cctgattctg	tgataaacg	tattaccgcc	tttgagttag	ctgataccgc	tcgcccgcagc	7560

-continued

```

cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa 7620
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 7680
tggaaagcgg gcagtgcgag caacgcaatt aatgtgagtt agctcactca ttaggcaccc 7740
caggctttac actttatgct tccggtcgt atgttggtgtg gaattgtgag cggataacaa 7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc 7860
ctccccagca ggcagaagta tgcaaacgat gcattctcaat tagtcagcaa ccatagtccc 7920
gcccctaact ccgccatcc cgcccctaac tccgccagc tccgccatt ctccgcccc 7980
tggtgacta atttttttta tttatgcaga ggccgaggcc gcctcgccct ctgagctatt 8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaaga 8100
tggataaagt tttaaacaga gaggaatctt tgcagctaata ggaccttcta ggtcttgaaa 8160
gg 8162

```

```

<210> SEQ ID NO 17
<211> LENGTH: 8162
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic construct: pJP172 ; pHybE-hCgl,z,
non-a,mu(234,235) V1

```

```

<400> SEQUENCE: 17

```

```

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtgcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta 180
tataagtga gtagtcgccc tgaacgttct ttttcgcaac gggtttgccc ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg 360
aagtgggtgg gagagttcga ggccttgccc ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggccttg cctgggcgtg ggggcccgcg cgtgcgaatc tgggtggacc ttcgcgcctg 480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcgg 600
ttttggggcc gcggggcggc acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggctcgca gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctggtgctt ggctcgcgc cgcctgtgat cgcgccgccc tgggcggcaa ggttgcccc 780
gtcgccacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagactc 840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcacccacac aaaggaaaag 900
ggcctttccg tcctcagccg tcgcttcagt tgactccacg gactaccggg cgcgctccag 960
gcacctcgat tagttctcga gcttttgagg tacgtcgtct ttaggttggg gggaggggtt 1020
ttatgcgatg gagtttcccc acactgagtg ggtggagact gaagttaggc cagcttgcca 1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa 1140
gcctcagaca gtggttcaaa gtttttttct tccatttcag gtgtcgtgag gaattctcta 1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcca ccatggagtt tgggctgagc 1260
tggttttttc ttgtcgcgat tttaaaaggt gtccagtcgc catggtatgc cgaaagggat 1320

```

-continued

gctgaaattg agaacgaaaa gctgcgccgg gaggttgaag aactgcggca ggccagcgag	1380
gcagatctcc agccaggaac tattgagtag gaacgccatc gacttacgcg tgcgcaggcc	1440
gagcgacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttctgtc tgtcgcggat cgcagggtga attgccagta ttctcgacgg gctccccctg	1560
tccgtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagcccg cttgatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtcgcgcg cggttcgctc	1740
actgttcagg ccggagccac agaccgcctg tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc	2040
gactattcgt gatattccgt cgtctgtggc gctggccccg tggatatgga aaaagcaccg	2100
ggataaacag ctcaccatga agcgtttcac taatggcggt ggcttctggt gcctggcg	2160
taaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgct	2280
cgaccaaggg cccatcggtc ttccccctgg caccctctc caagagcacc tctgggggca	2340
cagcgccctt gggctgctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc cctgaccage ggctgcaca ccttccccgc tgcctacag tctcaggac	2460
tctactccct cagcagcgtg gtgaccgtgc cctccagcag cttgggcacc cagacctaca	2520
tctgcaacgt gaatcacaag cccagcaaca ccaaggtgga caagaaagt gagcccaa	2580
cttgtgacaa aactcacaca tgcccaccgt gccagcacc tgaagccg cggggaccgt	2640
cagtcttctt cttccccca aaacccaagg acaccctcat gatctcccg acccctgagg	2700
tcacatgcgt ggtggtggac gtgagccagc aagaccctga ggtcaagttc aactggtacg	2760
tggacggcgt ggaggtgcat aatgccaa gaagcccg cgaggagcag tacaacagca	2820
cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat ggcaaggagt	2880
acaagtgcga ggtctccaac aaagccctcc cagcccccat cgagaaaacc atctccaaag	2940
ccaaagggca gccccagaa ccacagggtg acaccctgcc cccatcccg gaggatga	3000
ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatccagc gacatcgccg	3060
tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct cccgtgctgg	3120
actccgacgg ctctcttctt ctctacagca agctcaccgt ggacaagagc aggtggcagc	3180
aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac tacacgcaga	3240
agagcctctc cctgtctccg ggtaaatgag cggccgctcg aggcgggcaa ggccggatcc	3300
cccagctcgc acctctggct aataaaggaa atttattttt attgcaatag tgtgttgga	3360
ttttttgtgt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctatgtagag gatcgatccc cggcccgac gaactaaacc tgactacgac	3480
atctctgccc cttcttcgcg gggcagtgca tgtaatccct tcagttggtt ggtacaactt	3540
gccaactggg ccctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600

-continued

gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tectggagca gactttgcag tctgtggact gcaacacaaac attgccttta tgtgtaactc	3720
ttggctgaag ctcttacacc aatgctgggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggaggc tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa gggccccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtagtaa aagggtccta aggaacagcg	4020
atatctccca ccccatgagc tgtcacgggt ttatttacct ggggtcagga ttccacgagg	4080
gtagtgaacc atttttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctctgaatc ttgcctgtct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaaggtgta tgtgagggtc tcgaaaacaa ggtttcaggt gacgccccca	4260
gaataaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaaaccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catgggggtg ggacaagccg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctgggggtat taagatgtgt	4500
cccaggcagg gaccaagaca ggtgaacat gttgttacac tctatttga acaaggggaa	4560
agagagtgga cgcgcacagc agcggactcc actggttgc tctaacaccc cggaaaatta	4620
aacggggctc cagccaatg gggcccataa acaaagacaa gtggccactc tttttttga	4680
aattgtggag tgggggcacg cgtcagcccc cacacgccgc cctgcggttt tggactgtaa	4740
aataaggggtg taataacttg gctgattgta accccgctaa ccactgcggt caaaccactt	4800
gcccacaaaa ccactaatg caccocgggg aatacctgca taagtaggtg ggcgggcca	4860
gataggggag cgattctgc gatctggagg acaaattaca cacacttgcg cctgagcgcc	4920
aagcacaggg ttgttggtcc tcatattcac gaggtcgtg agagcacggg gggctaattg	4980
tgccatgggt agcatatact acccaaatat ctggatagca tatgctatcc taatctatat	5040
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat	5100
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat	5160
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat	5220
ccgggtagca tatgctatcc taatagagat tagggtagta tatgctatcc taatttatat	5280
ctgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	5340
gtagcatatg ctatcctaata ctatatctgg gtagcatagg ctatcctaata ctatatctgg	5400
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ttatatctgg	5460
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	5520
gtagtatatg ctatcctaata ctgtatccgg gtagcatatg ctatcctcat gataagctgt	5580
caaacatgag aattttcttg aagacgaaag ggccctgtga tacgcctatt tttatagggt	5640
aatgtcatga taataatggt ttcttagacg tcagggtggca cttttcgggg aaatgtgcgc	5700
ggaaccacct tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa	5760
taacctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc	5820
cgtgtgcgcc ttattccctt ttttgccgca ttttgccctc ctgtttttgc tccccagaa	5880
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggtg ttacatcgaa	5940

-continued

ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg	6000
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtgttga cgccgggcaa	6060
gagcaactcg gtgcgcgat acactattct cagaatgact tggttgagta ctcaccagtc	6120
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc	6180
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta	6240
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag	6300
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgctgcagc aatggcaaca	6360
acgttgcgca aactattaac tggcgaacta cttactctag cttcccgga acaattaata	6420
gactggatgg aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc	6480
tggtttattg ctgataaate tggagccggt gagcgtgggt ctgcggtat cattgcagca	6540
ctggggccag atggtaagcc ctcccgatc gtagtatat acacgacggg gagtcaggca	6600
actatggatg aacgaaatag acagatcgct gagatagtg cctcactgat taagcattgg	6660
taactgtcag accaagtta ctcatatata ctttagattg atttaaaact tcatttttaa	6720
tttaaaagga tctagtgaa gatccttttt gataatctca tgacccaaat cccttaacgt	6780
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	6840
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaa aaccaccgct accagcgtg	6900
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggttaactgg cttcagcaga	6960
gcgcagatac caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac	7020
tctgtagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag	7140
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	7200
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttccga agggagaaag	7260
gcggacaggt atccggtta cggcagggtc ggaacaggag agcgcacgag ggagcttcca	7320
gggggaaacg cctggtatct ttatagtcct gtccgggttc gccacctctg acttgagcgt	7380
cgatttttgt gatgctcgct agggggggcg agcctatgga aaaacgccag caacgcggcc	7440
tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc	7500
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	7560
cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgcgc aatacgcaaa	7620
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac	7680
tggaaagcgg gcagtgcagc caacgcaatt aatgtgagtt agctcactca ttaggcaccc	7740
caggctttac actttatgct tccggtcgt atgttgtgtg gaattgtgag cggataacaa	7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc	7860
ctccccagca ggcagaagta tgcaaagcat gcactcctaat tagtcagcaa ccatagtccc	7920
gcccctaact ccgcccatcc cgcccctaac tccgccagc tccgccatt ctcgcccaca	7980
tggctgacta atttttttta tttatgcaga ggccgaggcc gcctcgccct ctgagctatt	8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaaga	8100
tgataaaagt tttaaacaga gaggaatctt tgcagctaata ggaccttcta ggtcttga	8160
gg	8162

-continued

<210> SEQ ID NO 18
<211> LENGTH: 8162
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic construct: pJP184 ; pHybE-hCgl,z,
non-a,mu(234,235) V2

<400> SEQUENCE: 18

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggaggggtcg gcaattgaac cgggtcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgcc tgaacgttct ttttcgcaac gggtttgcg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360
aagtgggtgg gagagtgcga ggccttgcc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctcg cctgggcgtg ggggcgcgc cgtgcgaatc tgggtggcacc ttcgcgcctg	480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcgg	600
ttttggggcc gggggcgccg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgctt ggcttcgcgc cgccgtgtat cgcgccgccc tgggcggcaa ggcgtggccc	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagctc	840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tcctcagcgg tcgcttcctg tgactccacg gactaccggg cgcgcgccg	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggtccttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gtttttttct tcattttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgaga tcattgtgc cggggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaagg gtccagtgcg catggtatgc cgaagggat	1320
gctgaaattg agaacgaaa gctgcgccg gaggttgaag aactgcggca ggcagcgag	1380
gcagatctcc agccaggaac tattgagtag gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgtttttc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagcccg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggttcgctc	1740
actgttcagg ccggagccac agaccgcgt tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aagggcgctg ggaaacactg ccctttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980

-continued

ccttatctgg	ttgccgacgg	atgggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggccccg	tggtatggca	aaaagcaccc	2100
ggataacacg	ctcaccatga	agcgtttcac	taatgggctg	ggcttctggt	gcctgggctg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	gggtgatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttcctgggtg	acaagcgctg	2280
cgaccaaggg	cccatcggtc	ttccccctgg	caccctcttc	caagagcacc	tctgggggca	2340
cagcggccct	gggctgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	cctgaccacg	ggcgtgcaca	ccttccccgc	tgctctacag	tctcaggac	2460
tctactccct	cagcagcgtg	gtgacgtg	cctccagcag	cttgggcacc	cagacctaca	2520
tctgcaacgt	gaatcacaag	cccagcaaca	ccaagggtga	caagaaagtt	gagcccaa	2580
cttgtgacaa	aactcacaca	tgccccaccg	gcccagcacc	tgaagccg	gggggaccgt	2640
cagtcttctc	cttcccccca	aaacccaagg	acacctcat	gatctcccg	accctgagg	2700
tcacatgcgt	gggtgtggac	gtgagccacg	aagaccctga	ggccaagt	aactggtacg	2760
tgagcggcgt	ggaggtgcat	aatgccaaga	caaagccg	ggaggagcag	tacaacagca	2820
cgtaccgtgt	ggtcagcgtc	ctcacgtcc	tgaccagga	ctggctgaat	ggcaaggagt	2880
acaagtgcaa	ggctctcaac	aaagccctcc	cagcccccat	cgagaaaacc	atctccaaag	2940
ccaaagggca	gccccgagaa	ccacaggtgt	acacctgcc	cccatccgc	gaggagatga	3000
ccaagaacca	ggtcagcctg	acctgcctgg	tcaaaggctt	ctatcccagc	gacatcgccg	3060
tgagtgggga	gagcaatggg	cagccggaga	acaactacaa	gaccacgcct	cccgtgctgg	3120
actccgacgg	ctccttcttc	ctctacagca	agctcacctg	ggacaagagc	aggtggcagc	3180
aggggaacgt	cttctcatgc	tccgtgatgc	atgaggctct	gcacaaccac	tacacgcaga	3240
agagcctctc	cctgtctccg	ggtaaatgag	cggccgctcg	aggccggcaa	ggccggatcc	3300
cccgacctcg	acctctgggt	aataaaggaa	atttattttc	attgcaatag	tgtgttgga	3360
ttttttgtgt	ctctcactcg	gaaggacata	tgggagggca	aatcatttgg	tcgagatccc	3420
tcggagatct	ctagctagag	gatcgatccc	cgcgccggac	gaactaaacc	tgactacgac	3480
atctctgccc	cttcttcg	gggcagtgca	tgtaatccct	tcagttggtt	ggtacaactt	3540
gccaaactgg	ccctgttcca	catgtgacac	ggggggggac	caaacacaaa	ggggttctct	3600
gactgtagtt	gacatcctta	taaattggatg	tgacatttgg	ccaacactga	gtggctttca	3660
tcttgagca	gactttgcag	tctgtggact	gcaacacaa	attgccttta	tgtgtaactc	3720
ttgggtgaag	ctcttacacc	aatgtgggg	gacatgtacc	tcccaggggc	ccaggaagac	3780
tacgggaggg	tacaccaacg	tcaatcagag	gggcctgtgt	agctaccgat	aagcgaccc	3840
tcaagagggc	attagcaata	gtgtttataa	ggcccccttg	ttaaccctaa	acgggtagca	3900
tatgcttccc	gggtagtagt	atatactatc	cagactaacc	ctaattcaat	agcatatgtt	3960
acccaacggg	aagcatatgc	tatcgaatta	gggttagtaa	aagggtccta	aggaacagcg	4020
atatctcccc	ccccatgagc	tgacacggtt	ttatttacat	ggggtcagga	ttccacgagg	4080
gtagtgaacc	attttagtca	caagggcagt	ggctgaagat	caaggagcgg	gcagtgaact	4140
ctcctgaatc	ttcgctgct	tcttcattct	ccttcgttta	gctaatagaa	taactgctga	4200
gttgtgaaca	gtaagggtga	tgtgaggtgc	tcgaaaacaa	ggtttcaggt	gacgccccca	4260
gaataaaatt	tgagcggggg	gttcagtggg	ggcattgtgc	tatgacacca	atataaccct	4320

-continued

cacaaacccc	ttgggcaata	aatactagtg	taggaatgaa	acattctgaa	tatctttaac	4380
aatagaaatc	catggggtgg	ggacaagccg	taaagactgg	atgtccatct	cacacgaatt	4440
tatggctatg	ggcaacacat	aatcctagtg	caatatgata	ctgggggttat	taagatgtgt	4500
cccaggcagg	gaccaagaca	ggtgaaccat	gttggttacac	tctatttgta	acaaggggaa	4560
agagagtgga	cgccgacagc	agcggactcc	actggttgtc	tctaacaccc	ccgaaaatta	4620
aacggggctc	cacgccaatg	gggcccataa	acaagacaa	gtggccactc	ttttttttga	4680
aattgtggag	tgggggcacg	cgtcagcccc	cacacgccgc	cctgcggttt	tggactgtaa	4740
aataaggggtg	taataacttg	gctgattgta	accccgctaa	ccactgcggt	caaaccactt	4800
gcccacaaaa	ccactaatgg	caccccgggg	aatacctgca	taagtaggtg	ggcgggccaa	4860
gataggggag	cgattgctgc	gatctggagg	acaaattaca	cacacttgcg	cctgagcgcc	4920
aagcacaggg	ttgttggtcc	tcatattcac	gaggtcgctg	agagcacggt	gggctaattg	4980
tgccatgggt	agcatatact	acccaaatat	ctggatagca	tatgctatcc	taatctatat	5040
ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	5100
ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	taggctatcc	taatctatat	5160
ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatctgtat	5220
ccgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatcctaata	ctatatctgg	5340
gtagcatatg	ctatcctaata	ctatatctgg	gtagcatagg	ctatcctaata	ctatatctgg	5400
gtagcatatg	ctatcctaata	ctatatctgg	gtagtatatg	ctatcctaata	ttatatctgg	5460
gtagcatagg	ctatcctaata	ctatatctgg	gtagcatatg	ctatcctaata	ctatatctgg	5520
gtagtatatg	ctatcctaata	ctgtatccgg	gtagcatatg	ctatcctcat	gataagctgt	5580
caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	5640
aatgtcatga	taataatggt	ttcttagacg	tcaggtaggca	cttttcgggg	aaatgtgcgc	5700
ggaaccccta	ttgttttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5760
taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5820
cgtgtcgcgc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaa	5880
acgctgggtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	5940
ctggatctca	acagcggtaa	gaccccttag	agttttcgcc	ccgaagaacg	ttttccaatg	6000
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtgttga	cgcggggcaa	6060
gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6180
atgagtata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	6240
accgcttttt	tgcacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	6300
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgcagc	aatggcaaca	6360
acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccgcca	acaattaata	6420
gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcggccct	tccggctggc	6480
tggtttattg	ctgataaaat	tggagccggt	gagcgtgggt	ctcgcggtat	cattgcagca	6540
ctggggccag	atggtaagcc	ctcccgatc	gtagttatct	acacgacggg	gagtcaggca	6600

-continued

actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcatttg	6660
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa	6720
tttaaaagga tctaggtgaa gatccttttt gataatctca tgacaaaaat cccttaacgt	6780
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	6840
cctttttttc tgcgcgtaat ctgctgcttg caaacaacaa aaccaccgct accagcggtg	6900
gtttgtttgc cggatcaaga gctaccaact ctttttcga aggttaactgg cttcagcaga	6960
gcgcagatac caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac	7020
tctgtagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgcag	7140
cggtcgggct gaacggggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	7200
gaactgagat acctacacg tgagctatga gaaagcgcca cgcttccga agggagaaag	7260
gcggacaggt atccggaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca	7320
gggggaaacg cctggatatc ttatagtcct gtcgggttcc gccacctctg acttgagcgt	7380
cgatttttgt gatgctcgtc agggggggcg agcctatgga aaaacgccag caacgcggcc	7440
tttttaacgt tcctggcctt ttgctggcct tttgtcaca tgttctttcc tgcgttatcc	7500
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc	7560
cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa	7620
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac	7680
tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc	7740
caggctttac actttatgct tccggctcgt atgttggttg gaattgtgag cggataacaa	7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc	7860
ctccccagca ggcagaagta tgcaaagcat gcctctcaat tagtcagcaa ccatagtccc	7920
gcccctaact ccgcccatcc cgcccctaac tccgccagc tccgccatt ctcgcccaca	7980
tggtgacta atttttttta tttatgcaga ggcgaggcc gcctcgccct ctgagctatt	8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaga	8100
tggataaagt tttaaacaga gaggaatctt tgcagctaat ggaccttcta ggtcttga	8160
gg	8162

<210> SEQ ID NO 19

<211> LENGTH: 8162

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP173 ; pHybE-hCg1,z,
non-a,mu (234,237) V1

<400> SEQUENCE: 19

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggagggtcgc gcaattgaac cgggtcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtgcg gtagtcgcgc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg	360

-continued

aagtgggtgg	gagagttcga	ggccttgccg	ttaaggagcc	ccttcgcctc	gtgcttgagt	420
tgaggcctgg	cctgggcgct	ggggccgcgc	cgtgcgaatc	tggtggcacc	ttcgcgcctg	480
tctcgtcgtc	ttcgataagt	ctctagccat	ttaaaatttt	tgatgacctg	ctcgacgct	540
ttttttctgg	caagatagtc	ttgtaaatgc	gggccaagat	ctgcacactg	gtatttcggt	600
ttttggggcc	gcgggcgccg	acggggcccg	tgcgtcccag	cgcacatgtt	cggcgaggcg	660
gggcctgcga	gcgcgccac	cgagaatcgg	acgggggtag	tctcaagctg	gccggcctgc	720
tctggtgcct	ggcctcgcgc	cgcctgtgat	cgcctccccc	tgggcgga	ggctggcccg	780
gtcggcacca	gttgcgtgag	cggaaagatg	gccgcttccc	ggcctcgtc	caggagactc	840
aaaaatggagg	acgcggcgct	cgggagagcg	ggcggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcagt	tgactccacg	gagtaaccgg	cgcctccag	960
gcacctcgat	tagttctcga	gcttttgag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgccct	tttgagtgtt	ggatcttgg	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgagc	atttaaatgc	cggggcgcca	ccatggagtt	tgggctgagc	1260
tggctttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggtatgc	cgaagggat	1320
gctgaaattg	agaacgaaaa	gctgcgcggg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaac	tattgagtac	gaacgccatc	gacttacgcg	tgcgcaggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tggtggaaac	cgcattctgt	1500
actttcgtgc	tgctcgggat	cgcaggtgaa	attgccagta	ttctcgacgg	gtccccctg	1560
tcggtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttctc	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagcccg	ctggatgaac	tgataccggg	gttgcgtgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgcgc	ggcttcgctc	1740
actgttcagg	cggagccac	agaccgcgt	tgaatggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aaggcgctg	ggaaacactg	cccttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggtgaatgtg	gtgaagtctg	cccgtgtcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgaacac	1980
ccttatctgg	ttgccgagcg	atggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgtcgtcggc	gctggccccg	tggtatggca	aaaagcaccg	2100
ggataaacag	ctccatga	agcgtttcac	taatggcggt	ggcttctggt	gcctggcgcg	2160
taaagcgga	aaaaactacc	gtgaaaatgc	ggtggatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttcctgggtg	acaagcgct	2280
cgaccaaggg	cccatcggtc	ttccccctgg	caccctcctc	caagagcacc	tctgggggca	2340
cagcggccct	gggctgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	cctgaccagc	ggcgtgcaca	ccttccccgc	tgctctacag	tcctcaggac	2460
tctactccct	cagcagcgtg	gtgacgtgc	cctccagcag	cttgggcacc	cagacctaca	2520
tctgcaacgt	gaatcacaag	cccagcaaca	ccaagggtga	caagaaagtt	gagcccaa	2580
cttgtagaaa	aactcacaca	tgcccccggt	gcccagcacc	tgaagccctg	ggggcaccgt	2640
cagtcttctc	cttcccccca	aaacccaagg	acaccctcat	gatctcccg	accctgagg	2700

-continued

tcacatgcgt ggtgggtggac gtgagccacg aagaccctga ggtcaagttc aactggtagc	2760
tggacggcgt ggaggtgcat aatgccaaaga caaagccgcg ggaggagcag tacaacagca	2820
cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat ggcaaggagt	2880
acaagtgc aa ggtctccaac aaagccctcc cagcccccat cgagaaaaacc atctccaaag	2940
ccaaagggca gccccgagaa ccacagggtg acaccctgcc cccatcccgc gaggagatga	3000
ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc gacatcgccg	3060
tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct cccgtgctgg	3120
actccgacgg ctctctcttc ctctacagca agctcacctg ggacaagagc aggtggcagc	3180
aggggaacgt ctctcatgc tccgtgatgc atgaggtctt gcacaaccac tacacgcaga	3240
agagcctctc cctgtctccg ggtaaatgag cggccgctcg aggccggcaa ggcggatcc	3300
cccgacctcg acctctggct aataaaggaa atttattttc attgcaatag tgtgttgaa	3360
ttttttgtgt ctctcactcg gaaggacata tgggagggca aatcatttgg tcgagatccc	3420
tcggagatct ctagctagag gatcgatccc cggcccgac gaactaaacc tgactacgac	3480
atctctgccc ctctctcgcg gggcagtgca tgtaatccct tcagttggtt ggtacaactt	3540
gccaaactgg ccctgttcca catgtgacac ggggggggac caaacacaaa ggggttctct	3600
gactgtagtt gacatcctta taaatggatg tgcacatttg ccaacactga gtggctttca	3660
tcttgagca gactttgcag tctgtggact gcaacacaa attgccttta tgtgtaactc	3720
ttggctgaag ctcttacacc aatgtgggg gacatgtacc tcccaggggc ccaggaagac	3780
tacgggaggg tacaccaacg tcaatcagag gggcctgtgt agctaccgat aagcggaccc	3840
tcaagagggc attagcaata gtgtttataa ggcccccttg ttaaccctaa acgggtagca	3900
tatgcttccc gggtagtagt atatactatc cagactaacc ctaattcaat agcatatgtt	3960
acccaacggg aagcatatgc tatcgaatta gggtagtaa aagggtccta aggaacagcg	4020
atatctccca cccatgagc tgtcacggtt ttatttacat ggggtcagga ttccacgagg	4080
gtagtgaacc attttagtca caagggcagt ggctgaagat caaggagcgg gcagtgaact	4140
ctctgaatc ttgcctgct tcttcattct ccttcgttta gctaatagaa taactgctga	4200
gttgtgaaca gtaagggtga tgtgaggtgc tcgaaaacaa ggtttcaggt gacgccccca	4260
gaataaaatt tggacggggg gttcagtggt ggcattgtgc tatgacacca atataaccct	4320
cacaaacccc ttgggcaata aatactagt taggaatgaa acattctgaa tatctttaac	4380
aatagaaatc catgggggtg ggacaagcgg taaagactgg atgtccatct cacacgaatt	4440
tatggctatg ggcaacacat aatcctagt caatatgata ctgggggttat taagatgtgt	4500
cccaggcagg gaccaagaca ggtgaacat gttgttacac tctatttga acaaggggaa	4560
agagagtgga cggcgacgc agcggactcc actgggtgtc tctaacaccc ccgaaaatta	4620
aacggggctc cacgccaatg gggcccataa acaaaagaaa gtggccactc tttttttga	4680
aattgtggag tgggggcacg cgtcagcccc cacacgccgc cctgcggttt tggactgtaa	4740
aataaggggt taataacttg gctgattgta accccgctaa cactgcgggt caaaccactt	4800
gcccacaaaa ccactaatgg cccccgggg aatacctgca taagtaggtg ggcgggcca	4860
gataggggag cgattgctgc gatctggagg acaaattaca cacacttgcg cctgagcgcc	4920
aagcacaggg ttgttggtcc tcatattcac gaggtcgtg agagcacggt gggctaattg	4980

-continued

tgccatgggt agcatatact acccaaatat ctggatagca tatgctatcc taatctatat	5040
ctgggtagca taggctatcc taatctatat ctgggtagca tatgctatcc taatctatat	5100
ctgggtagta tatgctatcc taatttatat ctgggtagca taggctatcc taatctatat	5160
ctgggtagca tatgctatcc taatctatat ctgggtagta tatgctatcc taatctgtat	5220
ccgggtagca tatgctatcc taatagagat taggtagta tatgctatcc taatttatat	5280
ctgggtagca tatactaccc aaatatctgg atagcatatg ctatcctaata ctatatctgg	5340
gtagcatatg ctatcctaata ctatatctgg gtagcatagg ctatcctaata ctatatctgg	5400
gtagcatatg ctatcctaata ctatatctgg gtagtatatg ctatcctaata ttatatctgg	5460
gtagcatagg ctatcctaata ctatatctgg gtagcatatg ctatcctaata ctatatctgg	5520
gtagtatatg ctatcctaata ctgtatccgg gtagcatatg ctatcctaata gataagctgt	5580
caaacatgag aattttcttg aagacgaaag ggccctctga tacgcctatt tttataggtt	5640
aatgtcatga taataatggt ttcttagacg tcagggtggca cttttcgggg aaatgtgcgc	5700
ggaaccacct tttgtttatt ttctaaata cattcaaata tgtatccgct catgagacaa	5760
taaccctgat aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc	5820
cggtctgccc ttattccctt ttttgccgca ttttgccctc ctgtttttgc tcaaccagaa	5880
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cagcagtggtt ttacatcgaa	5940
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg	6000
atgagcactt ttaaagtctt gctatgtggc gcggtattat cccgtgttga cgcggggcaa	6060
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc	6120
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc	6180
atgagtata acactgccc caacttactt ctgacaacga tcggaggacc gaaggagcta	6240
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag	6300
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgcagc aatggcaaca	6360
acgttgccga aactattaac tggcgaacta ctactctag cttcccggca acaattaata	6420
gactggatgg aggcgataa agttgcagga ccacttctgc gctcgccct tccggctggc	6480
tggtttattg ctgataaatc tggagccggt gaggctgggt ctcgcggtat cattgcagca	6540
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gtagcaggca	6600
actatggatg aacgaaatag acagatcgct gagatagggt cctcactgat taagcattgg	6660
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa	6720
tttaaaagga tctaggtgaa gatccttttt gataatctca tgacaaaaat cccttaacgt	6780
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	6840
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg	6900
gtttgtttgc cggatcaaga gctaccaact ctttttcga aggttaactgg cttcagcaga	6960
gcgcagatag caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac	7020
tctgtagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	7080
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgcag	7140
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacac	7200
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttccga agggagaaag	7260
gcggacaggt atccggtgaa cggcagggtc ggaacaggag agcgcacgag ggagcttcca	7320

-continued

```

gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 7380
cgatTTTTgt gatgctcgtc aggggggscg agcctatgga aaaacgccag caacgcggcc 7440
tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc 7500
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc 7560
cgaacgaccg agcgcagcga gtcagttagc gaggaagcgg aagagcgcgc aatacgcaaa 7620
ccgctctccc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 7680
tggaagcgg gcagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc 7740
caggctttac actttatgct tcgggctcgt atgttggttg gaattgtgag cggataacaa 7800
tttcacacag gaaacagcta tgaccatgat tacgccaagc tctagctaga ggtcgagtcc 7860
ctccccagca ggcagaagta tgcaaagcat gcactcctaat tagtcagcaa ccatagtccc 7920
gcccctaact ccgcccctcc cgccttaac tcggccagc tcggccatt ctccgcccc 7980
tggtgacta atttttttta tttatgcaga ggcgagggcc gcctcgccct ctgagctatt 8040
ccagaagtag tgaggaggct tttttggagg cctaggcttt tgcaaaaagc tttgcaaaaga 8100
tgataaagt tttaaacaga gaggaatctt tgcagctaata ggaccttcta ggtcttgaag 8160
gg 8162

```

```

<210> SEQ ID NO 20
<211> LENGTH: 8162
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic construct: pJP185 ; pHybE-hCgl,z,
non-a,mu (234,237) V2

```

```

<400> SEQUENCE: 20

```

```

agtgggaatt ggctccgggt cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgctcgtg actggtccg cctttttccc gaggggtggg gagaaccgta 180
tataagtga gtagtcgcc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgct 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg 360
aagtgggtgg gagagttcga ggccttgcc ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgct ggggcgcgcg cgtgcgaatc tggtggcacc ttcgcgcctg 480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
tttttctcgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctggtgcct ggctcgcgc cgcctgtat cgcgccgccc tgggcggcaa ggttgcccc 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggcctgctg caggagagtc 840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tccccacac aaaggaaaag 900
ggcctttccg tcctcagccg tcgcttcagt tgactccacg gactaccggg cgccttcag 960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt 1020
ttatgcgatg gagtttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca 1080

```

-continued

cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggt	tcattctcaa	1140
gcctcagaca	gtgggtcaaa	gtttttttct	tcattttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgaga	tccattgtgc	cggggcgcca	ccatggagtt	tgggctgagc	1260
tggctttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggtatgc	cgaagggat	1320
gctgaaattg	agaacgaaaa	gctgcgccgg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaac	tattgagtac	gaacgccatc	gacttacgcg	tgcgcaggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tggtggaaac	cgcattctgt	1500
actttcgtgc	tgctcgggat	cgcaggtgaa	attgccagta	ttctcgacgg	gctccccctg	1560
tcggtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttctt	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagccgcg	ctggatgaac	tgataccggg	gttgcgtgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgccg	ggcttcgctc	1740
actgttcagg	cggagccac	agaccgccgt	tgaatgggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaaacactg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggatgaatgtg	gtgaagtctg	cccgtgtcgg	1920
ttattccaaa	atgctcgtgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccagcgg	atggtgatgc	cgagaacttt	atgaaaacc	acgttgagcc	2040
gactattcgt	gatattccgt	cgtcgtcggc	gctggccccg	tggatatggca	aaaagcaccg	2100
ggataaacag	ctcaccatga	agcgtttcac	taatggggcg	ggcttctggt	gcctggggcg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggatgatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttctggggtg	acaagcgcgt	2280
cgaccaaggg	cccacgtggt	ttccccctgg	caccctcttc	caagagcacc	tctgggggca	2340
cagcggccct	gggtgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	cctgaccagc	ggcgtgcaca	ccttccccgc	tgtcctacag	tcctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtgc	cctccagcag	cttgggcacc	cagacctaca	2520
tctgcaacgt	gaatcacaag	cccagcaaca	ccaagggtga	caagaaagtt	gagcccaa	2580
cttgtagaaa	aactcacaca	tgcccaccgt	gccagcacc	tgaagccctg	ggggcaccgt	2640
cagtcttctt	cttcccccca	aaacccaagg	acaccctcat	gatctcccgg	acccctgagg	2700
tcacatgcgt	ggtggtggac	gtgagccacg	aagaccctga	ggtcaagttc	aactggtacg	2760
tggacggcgt	ggaggtgcat	aatgccaaaga	caaagccgcg	ggaggagcag	tacaacagca	2820
cgtaccgtgt	ggtcagcgtc	ctcaccgtcc	tgaccaggga	ctggctgaat	ggcaaggagt	2880
acaagtgcga	ggtctccaac	aaagccctcc	cagcccccat	cagaaaaacc	atctccaaag	2940
ccaaagggca	gccccgagaa	ccacaggtgt	acaccctgcc	cccatcccgc	gaggagatga	3000
ccaagaacca	ggtcagcctg	acctgcctgg	tcaaaggctt	ctatcccagc	gacatcgccg	3060
tggagtggga	gagcaatggg	cagccggaga	acaactacaa	gaccacgcct	cccgtgctgg	3120
actccgacgg	ctccttcttc	ctctacagca	agctcacctg	ggacaagagc	aggtggcagc	3180
aggggaacgt	cttctcatgc	tccgtgatgc	atgaggctct	gcacaaccac	tacacgcaga	3240
agagcctctc	cctgtctccg	ggtaaatgag	cggccgctcg	aggccggcaa	ggccggatcc	3300
cccgaacctg	acctctgggt	aataaaggaa	atttattttc	attgcaatag	tgtgttgga	3360

-continued

ttttttgtgt	ctctcactcg	gaaggacata	tgaggaggca	aatcatttgg	tcgagatccc	3420
tcggagatct	ctagctagag	gatcgatccc	cgccccggac	gaactaaacc	tgactacgac	3480
atctctgccc	cttcttcgcg	gggcagtgca	tgtaatccct	tcagttgggt	ggtacaactt	3540
gccaaactggg	ccctgttcca	catgtgacac	ggggggggac	caaacacaaa	ggggttctct	3600
gactgtagtt	gacatcctta	taaattggatg	tgacacattg	ccaacactga	gtggctttca	3660
tcctggagca	gactttgcag	tctgtggact	gcaacacaac	attgccttta	tgtgtaactc	3720
ttggctgaag	ctcttacacc	aatgctgggg	gacatgtacc	tcccaggggc	ccaggaagac	3780
tacgggaggc	tacaccaacg	tcaatcagag	gggcctgtgt	agctaccgat	aagcggaccc	3840
tcaagagggc	attagcaata	gtgtttataa	ggcccccttg	ttaaccctaa	acgggtagca	3900
tatgcttccc	gggtagtagt	atatactatc	cagactaacc	ctaattcaat	agcatatgtt	3960
acccaacggg	aagcatatgc	tatcgaatta	gggttagtaa	aagggtccta	aggaacagcg	4020
atatctccca	ccccatgagc	tgtaacggtt	ttatttacat	ggggtcagga	ttccacgagg	4080
gtagtgaacc	attttagtca	caagggcagt	ggctgaagat	caaggagcgg	gcagtgaact	4140
ctcctgaatc	ttcgctgct	tcttcattct	ccttcgttta	gctaatagaa	taactgctga	4200
gttgtgaaca	gtaagggtga	tgtgaggtgc	tcgaaaacaa	ggtttcaggt	gacgccccca	4260
gaataaaatt	tggaacgggg	gttcagtggt	ggcattgtgc	tatgacacca	atataaccct	4320
cacaaacccc	ttgggcaata	aatactagt	taggaatgaa	acattctgaa	tatctttaac	4380
aatagaaatc	catgggggtg	ggacaagccg	taaagactgg	atgtccatct	cacacgaatt	4440
tatggctatg	ggcaacacat	aatcctagt	caatatgata	ctgggggttat	taagatgtgt	4500
cccaggcagg	gaccaagaca	ggtgaaccat	gttggttacac	tctatttgta	acaaggggaa	4560
agagagtggg	cgcgcacagc	agcggactcc	actggttgtc	tctaacaccc	ccgaaaatta	4620
aacggggctc	cacgccaatg	gggcccataa	acaaagacaa	gtggccactc	ttttttttga	4680
aattgtggag	tgggggcacg	cgtcagcccc	cacacgcgcg	cctgcggttt	tggaactgtaa	4740
aataagggtg	taataacttg	gctgattgta	accccgctaa	ccactgcggg	caaaccactt	4800
gcccacaaaa	ccactaatgg	caccccgggg	aatacctgca	taagttagtg	ggcgggccaa	4860
gataggggcg	cgattgctgc	gatctggagg	acaaattaca	cacacttgcg	cctgagcgcc	4920
aagcacaggg	ttgttggtcc	tcataattac	gaggtcgtcg	agagcacggg	gggctaatgt	4980
tgccatgggt	agcatatact	acccaaatat	ctggatagca	tatgctatcc	taatctatat	5040
ctgggtagca	taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	5100
ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	taggctatcc	taatctatat	5160
ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatctgtat	5220
ccgggtagca	tatgctatcc	taatagagat	tagggtagta	tatgctatcc	taatttatat	5280
ctgggtagca	tatactaccc	aaatatctgg	atagcatatg	ctatcctaata	ctatatctgg	5340
gtagcatatg	ctatcctaata	ctatatctgg	gtagcatagg	ctatcctaata	ctatatctgg	5400
gtagcatatg	ctatcctaata	ctatatctgg	gtagtatatg	ctatcctaata	ttatatctgg	5460
gtagcatagg	ctatcctaata	ctatatctgg	gtagcatatg	ctatcctaata	ctatatctgg	5520
gtagtatatg	ctatcctaata	ctgtatccgg	gtagcatatg	ctatcctcat	gataagctgt	5580
caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	5640
aatgtcatga	taataatggt	ttcttagacg	tcagggtggca	cttttcgggg	aaatgtgcgc	5700

-continued

ggaaccccta	ttgttttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	5760
taacctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	5820
cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaa	5880
acgtgtgtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	5940
ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	6000
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtgttga	cgccgggcaa	6060
gagcaactcg	gtcgccgcat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	6120
acagaaaagc	atcttacgga	tggtcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	6180
atgagtata	acactcggc	caacttactt	ctgacaacga	tggaggacc	gaaggagcta	6240
accgcttttt	tgacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	6300
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgcagc	aatggcaaca	6360
acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccggca	acaattaata	6420
gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcgccctt	tccggctggc	6480
tggtttattg	ctgataaate	tgtagccggt	gagcgtgggt	ctcgcggtat	cattgcagca	6540
ctggggccag	atggtaagcc	ctcccgtatc	gtagttatct	acacgacggg	gagtcaggca	6600
actatggatg	aacgaaatag	acagatcgct	gagataggtg	cctcactgat	taagcattgg	6660
taactgtcag	accaagttta	ctcatatata	ctttagattg	atttaaaact	tcatttttaa	6720
tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacccaaat	cccttaacgt	6780
gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	6840
cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	accagcggtg	6900
gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	cttcagcaga	6960
gcgcagatac	caaatactgt	tcttctagt	tagccgtagt	taggccacca	cttcaagaac	7020
tctgtagcac	cgctacata	cctcgctctg	ctaactcctg	taccagtggc	tgctgccagt	7080
ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	taaggcgcag	7140
cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggagcgaac	gacctacacc	7200
gaactgagat	acctacacgc	tgagctatga	gaaagcgcca	cgcttcccga	aggagaaaa	7260
gcggacaggt	atccggtaa	cggcagggtc	ggaacaggag	agcgcacgag	ggagcttcca	7320
gggggaaaac	cctggtatct	ttatagtcct	gtcgggtttc	gccacctctg	acttgagcgt	7380
cgatttttgt	gatgctcgct	agggggggcg	agcctatgga	aaaacgccag	caacgcggcc	7440
tttttacggt	tcctggcctt	ttgctggcct	tttgctcaca	tgttctttcc	tgcggtatcc	7500
cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataccgc	tcgccgcagc	7560
cgaacgaccg	agcgcagcga	gtcagttagc	gaggaagcgg	aagagcgcgc	aatacgcaaa	7620
ccgctctctc	ccgcgcgttg	gccgattcat	taatgcagct	ggcacgacag	gtttcccagc	7680
tgaaagcgg	gcagttagcg	caacgcaatt	aatgtgagtt	agctcactca	ttaggcaccc	7740
caggctttac	actttatgct	tccggctcgt	atgttggtg	gaattgtgag	cggataacaa	7800
tttcacacag	gaaacagcta	tgaccatgat	tacgccaaagc	tctagctaga	ggtcgagtc	7860
ctccccagca	ggcagaagta	tgcaaagcat	gcattctcaat	tagtcagcaa	ccatagtcct	7920
gcccctaact	ccgcccattc	cgcccctaac	tccgcccagt	tccgcccatt	ctccgcccga	7980

-continued

tggtgacta atttttttta tttatgcaga ggccgaggcc gcctcgccct ctgagctatt	8040
ccagaagtag tgaggaggct tttttggagg cctaggtctt tgcaaaaagc tttgcaaaga	8100
tgataaagt tttaaacaga gaggaatctt tgcagctaata ggaccttcta ggtcttgaaa	8160
gg	8162

<210> SEQ ID NO 21
 <211> LENGTH: 8151
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic construct: pJP174 ; pHybE-hCg2,n- V1
 <400> SEQUENCE: 21

agtggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga	60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta	180
tataagtga gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctt	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg	360
aagtgggtgg gagagtcca ggccttcgc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tgggtggacc ttcgcgcctg	480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcggggcgcg acggggcccg tgcgtcccag cgcacatggt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggttag tctcaagctg gccggcctgc	720
tctgggtgct ggctcgcgc cgcctgttat cgcgccgccc tgggcggcaa ggcctggccc	780
gtcggcacca gttgcgtgag cggaaagatg gcccttccc ggccctgctg caggagagtc	840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tccccacac aaaggaaaag	900
ggcctttccg tctcagccg tcgcttcag tgactccacg gactaccggg cgccttcag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgcct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttctt tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcca ccatggagtt tgggctgagc	1260
tggtcttttc ttgtcgcgat tttaaaagg gtccagtgc catggtagc cgaagggat	1320
gctgaaattg agaacgaaaa gctgcgccgg gaggttgaag aactgcggca ggccagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tgggtgaaac cgcattctgt	1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagcccg ctggatgaac tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcg ggttcgctc	1740
actgttcagg ccggagccac agaccgcgt tgaatggcg gatgctaatt actatctccc	1800

-continued

gaaagaatcc gcataccagg aagggcgctg ggaaacactg ccctttcagc gggccatcat	1860
gaatgcgatg ggacgcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgcaacac	1980
ccttatctgg ttgccgacgg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc	2040
gactattcgt gatattccgt cgctgctggc gctggccccg tggtaggca aaaagcaccg	2100
ggataaacag ctcacatga agcgtttcac taatgggctg ggcttctggg gectgggagg	2160
taaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgcttttgat gatgatattg aacaggaagg ctctccgacg ttctgggtg acaagcgcgt	2280
cgaccaaggg cccatcggtc tccccctgg cgccctgctc tagaagcacc tccgagagca	2340
cagcgccctt gggctgcctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc tctgaccagc ggcgtgcaca ccttcccagc tgcctgcag tctcaggac	2460
tctactccct cagcagcgtg gtgaccgtgc cctccagcaa cttcggcacc cagacctaca	2520
catgcaacgt agatcacaag cccagcaaca ccaaggtgga caagacagtt gagcgcaa	2580
gttgtgtcga gtgcccaccg tgcccagcac cacctgtggc aggaccgtca gtcttctct	2640
tcccccaaa acccaaggac accctcatga tctcccgac cctgagggtc acgtgcgtgg	2700
tggtaggagt gagccacgaa gaccccgagg tccagttcaa ctggtacgtg gacggcgtgg	2760
aggtagcataa tgccaagaca aagccacggg aggagcagtt caacagcagc ttccgtgtgg	2820
tcagcgtcct caccgttgtg caccaggact ggctgaacgg caaggagtac aagtgcagg	2880
tctcaacaa aggcctccca gccccctcg agaaaacat ctccaaaacc aaagggcagc	2940
cccgagaacc acaggtgtac accctgcccc catcccgga ggagatgacc aagaaccagg	3000
tcagcgtgac ctgctgggtc aaaggcttct accccagcga catcgccgtg gagtgggaga	3060
gcaatgggca gccggagaa aactacaaga ccacacctcc catgctggac tccgacggct	3120
ccttcttct ctacagcaag ctcaccgtgg acaagagcag gtggcagcag gggaaactct	3180
tctcatgctc cgtgatgcat gaggtcttgc acaaccacta cagcagaag agcctctccc	3240
tgctccggg taaatgacgc ggccgctcga ggccggcaag gccggatccc ccgacctcga	3300
cctctggcta ataaaggaaa tttattttca ttgcaatagt gtgttggaat tttttgtgtc	3360
tctcactcgg aaggacatat gggagggcaa atcatttggt cgagatccct cggagatctc	3420
tagctagagg atcgatcccc gccccggacg aactaaacct gactacgaca tctctgcccc	3480
ttcttcggg ggcatgcat gtaatccctt cagttggttg gtacaacttg ccaactgggc	3540
cctgttccac atgtgacagc gggggggacc aaacacaaag gggttctctg actgtagttg	3600
acatccttat aaatggatgt gcacatttgc caacactgag tggttttcat cctggagcag	3660
actttgcagt ctgtggactg caacacaaca ttgcctttat gtgtaactct tggctgaagc	3720
tcttacacca atgctggggg acatgtacct cccaggggac cagggaagact acgggaggct	3780
acaccaacgt caatcagagg ggcctgtgta gctaccgata agcggacctc caagagggca	3840
ttagcaatag tgtttataag gcccccttgt taaccctaaa cgggtagcat atgcttccc	3900
ggtagtagta tatactatcc agactaacc taattcaata gcatatgtta cccaacggga	3960
agcatatgct atcgaaatag ggttagtaaa agggctcctaa ggaacagcga tatctccac	4020
cccatgagct gtcacggttt tatttacatg gggtcaggat tccacaggg tagtgaacca	4080

-continued

ttttagtcac aagggcagtg gctgaagatc aaggagcggg cagtgaactc tctgaaatct	4140
tcgctgctt cttcattctc ctctgttttag ctaatagaat aactgctgag ttgtgaacag	4200
taaggtgtat gtgaggtgct cgaatacaag gtttcaggtg acgccccag aataaaattt	4260
ggacgggggg ttcagtggtg gcattgtgct atgacaccaa tataaccctc acaaacccct	4320
tgggcaataa atactagtgt aggaatgaaa cattctgaat atctttaaca atagaaatcc	4380
atggggtggg gacaagccgt aaagactgga tgtccatctc acacgaattt atggctatgg	4440
gcaacacata atcctagtgc aatatgatac tggggttatt aagatgtgtc ccaggcaggg	4500
accaagacag gtgaaccatg ttgttacact ctatttgtaa caaggggaaa gagagtggac	4560
gccgacagca gcggactcca ctggtgtgct ctaacacccc cgaaaattaa acggggctcc	4620
acgccaatgg ggccataaa caaagacaag tggccactct ttttttgaa attgtggagt	4680
gggggcacgc gtcagcccc acacgccgcc ctgcggtttt ggactgtaaa ataagggtgt	4740
aataacttgg ctgattgtaa ccccgtaac cactgcggtc aaaccacttg ccacaaaaac	4800
cactaatggc accccgggga atacctgcat aagtaggtgg gcgggccaag ataggggcgc	4860
gattgtgcgc atctggagga caaattacac acacttgcgc ctgagcgcca agcacagggt	4920
tgttggtcct catattcacg aggtcgtga gagcacggtg ggctaagtgt gccatgggta	4980
gcataacta cccaaatata tggatagcat atgctatcct aatctatata tgggtagcat	5040
aggctatcct aatctatata tgggtagcat atgctatcct aatctatata tgggtagtat	5100
atgctatcct aatttatata tgggtagcat aggctatcct aatctatata tgggtagcat	5160
atgctatcct aatctatata tgggtagtat atgctatcct aatctgtata cgggtagcat	5220
atgctatcct aatagagatt agggtagtat atgctatcct aatttatata tgggtagcat	5280
atactacca aatatctgga tagcatatgc taccctaata tatatctggg tagcatatgc	5340
taccctaata tatatctggg tagcataggc taccctaata tatatctggg tagcatatgc	5400
taccctaata tatatctggg tagtatatgc taccctaatt tatatctggg tagcataggc	5460
taccctaata tatatctggg tagcatatgc taccctaata tatatctggg tagtatatgc	5520
taccctaata tgtatcggg tagcatatgc taccctcatg ataagctgtc aaacatgaga	5580
atcttcttga agacgaagg gcctcgtgat acgcctatct ttataggtta atgtcatgat	5640
aataatgggt tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaaccctat	5700
ttgtttatct ttctaaatac attcaaatat gtatccgctc atgagacaat aacctgata	5760
aatgcttcaa taatatgtaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct	5820
tattcccttt tttgcggcat tttgccttcc tgtttttgct caccagaaa cgctggtgaa	5880
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa	5940
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt	6000
taaggttctg ctatgtggcg cggtattatc ccgtgttgac gccgggcaag agcaactcgg	6060
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca	6120
tcttacggat ggcatgacag taagagaatt atgcagtgtc gccataacca tgagtgataa	6180
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt	6240
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc	6300
cataccaaac gacgagcgtg acaccacgat gcctgcagca atggcaacaa cgttgcgcaa	6360
actattaact ggcaactac ttactctagc ttcccgcaa caattaatag actggatgga	6420

-continued

```

ggcggataaa gttgcaggac cactctcgcg ctcgccctt cggctggct ggtttattgc 6480
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 6540
tggttaagccc tcccgtatcg tagttatcta cagcagggg agtcaggcaa ctatggatga 6600
acgaaataga cagatcgctg agataggtgc ctactgatt aagcattggt aactgtcaga 6660
ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat ttaaaggat 6720
ctaggtgaag atcctttttg ataatctcat gacaaaaac ccttaacgtg agttttcggt 6780
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 6840
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgc 6900
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 6960
aaatactggt cttctagtgt agccgtagt agggccaccac ttcaagaact ctgtagcacc 7020
gcctacatac ctgcgtctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 7080
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 7140
aacggggggg tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 7200
cctacagcgt gagctatgag aaagcgcac gcttcccgaa gggagaaaagg cggacaggta 7260
tccggtaagc ggcagggctg gaacaggaga gcgcacgagg gagcttcag ggggaaacgc 7320
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttctg 7380
atgctcgtca ggggggcgga gcctatgaa aaacgccagc aacgcggcct ttttacggtt 7440
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 7500
ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga 7560
gcgcagcagc tcagtgcgag aggaagcgga agagcgccca atacgcaaac cgcctctccc 7620
cgcgcggttg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 7680
cagtgcgagc aacgcaatta atgtgagtta gctcactcat taggcacccc aggetttaca 7740
ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg 7800
aaacagctat gaccatgatt acgccaagct ctactagag gtcgagtcct tccccagcag 7860
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 7920
cgcccatccc gcccctaact ccgcccgtt ccgcccattc tccgcccctat ggctgactaa 7980
ttttttttat ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt 8040
gaggaggctt ttttgagggc ctaggctttt gcaaaaagct ttgcaaagat ggataaagtt 8100
ttaaacagag aggaatcttt gcagctaag gaccttctag gtcttgaaag g 8151

```

<210> SEQ ID NO 22

<211> LENGTH: 8151

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP187 ; pHybE-hCg2,n- V2

<400> SEQUENCE: 22

```

agtgggaatt ggctccggtg ccgctcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggaggggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgctgtgt actggctccg cttttttccc gaggggtggg gagaaccgta 180
tataagtgca gtagtcgccc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240

```

-continued

ggtaagtgcc	gtgtgtggtt	cccgcgggcc	tggcctcttt	acgggttatg	gccttgcgt	300
gccttgaatt	acttccacct	ggctgcagta	cgtgattctt	gateccgagc	tccgggttg	360
aagtgggtgg	gagagtccga	ggccttgcgc	ttaaggagcc	ccttcgcctc	gtgcttgagt	420
tgaggcctgg	cctgggcgct	ggggcgccg	cgtgcgaatc	tgttggcacc	tccgcgcctg	480
tctcgctgct	tccgataagt	ctctagccat	ttaaaatttt	tgatgacctg	ctgcgacgct	540
ttttttctgg	caagatagtc	ttgtaaatgc	gggccaagat	ctgcacactg	gtatttcggg	600
ttttggggcc	gcgggcggcg	acggggcccg	tgcgtcccag	cgcacatggt	cggcgaggcg	660
gggcctgcga	gcgcggccac	cgagaatcgg	acgggggttag	tctcaagctg	gcgcgcctgc	720
tctgggtgct	ggcctcgcgc	cgccgtgtat	cgcgccgccc	tgggcggcaa	ggctggcccg	780
gtcggcacca	gttgcgtgag	cggaaagatg	gccgcttccc	ggccttgcctg	caggagagctc	840
aaaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcctg	tgaactccacg	gagtagccgg	cgcgcgtccag	960
gcacctcgat	tagttctcga	gcttttgag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggg	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgaga	tccattgtgc	cggggcgcca	ccatggagtt	tgggctgagc	1260
tggctttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggatgac	cgaaggagat	1320
gctgaaattg	agaacgaaaa	gctgcgcggg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaac	tattgagtac	gaacgccatc	gacttacgcg	tgcgcaggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tgttggaaac	cgcattctgt	1500
actttcgtgc	tgtcgcggat	cgcagggtgaa	attgccagta	ttctcgacgg	gtccccctg	1560
tccgtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttcc	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagccgcg	ctggatgaac	tgatccggg	gttgcgtgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgcgg	ggcttcgctc	1740
actgttcagg	cggagccac	agaccgcctg	tgaatggcgg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaaacactg	cccttccagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggatgaatgtg	gtgaagtctg	cccgtgtcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccgagcg	atggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgtgctggc	gctggccccg	tgttatggca	aaaagcaccg	2100
ggataaacacg	ctcaccatga	agcgtttcac	taatgggcgt	ggcttctggg	gcctgggcgg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggatgatgtg	gcgggttatg	atgaacttgc	2220
tgttttgat	gatgatattg	aacagggaag	ctctccgacg	ttctgggtg	acaagcgcgt	2280
cgaccaaggg	cccacggctc	ttccccctgg	cgcctgctc	tagaagcacc	tccgagagca	2340
cagcggccct	gggtgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	tctgaccagc	ggcgtgcaca	ccttccagc	tgtcctgcag	tcctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtgc	cctccagcaa	cttcggcacc	cagacctaca	2520
catgcaacgt	agatcacaag	cccagcaaca	ccaagggtgga	caagacagtt	gagcgcaaat	2580

-continued

gttggtgcga	gtgcccacg	tgcccagcac	cacctgtggc	aggaccgtca	gtcttctct	2640
ttccccaaa	acccaaggac	accctcatga	tctcccgac	ccctgaggtc	acgtgcgtgg	2700
tggtggacgt	gagccacgaa	gaccccgagg	tccagttcaa	ctggtacgtg	gacggcgtgg	2760
aggtgcataa	tgccaagaca	aagccacggg	aggagcagtt	caacagcacg	ttcgtgtgg	2820
tcagcgtcct	caccgttgtg	caccaggact	ggctgaacgg	caaggagtac	aagtgcgaag	2880
ttccaacaa	aggcctccca	gcccccatcg	agaaaacat	ctccaaaacc	aaagggcagc	2940
cccgagaacc	acagggtgtac	accctgcccc	catcccgga	ggagatgacc	aagaaccagg	3000
tcagcctgac	ctgcctggtc	aaaggcttct	accccagcga	catcgccgtg	gagtgggaga	3060
gcaatgggca	gccggagaac	aactacaaga	ccacacctcc	catgctggac	tccgacggct	3120
ccttcttct	ctacagcaag	ctcacgtgg	acaagagcag	gtggcagcag	gggaacgtct	3180
ttctcatgctc	cgtgatgcat	gaggctctgc	acaaccacta	cacgcagaag	agcctctccc	3240
tgtctccggg	taaatgacgc	ggccgctcga	ggccggcaag	gccggatccc	ccgacctcga	3300
cctctggcta	ataaaggaaa	tttattttca	ttgcaatagt	gtgttggaat	ttttgtgtc	3360
ttcactcgg	aaggacatat	gggagggcaa	atcatttgg	cgagatccct	cggagatctc	3420
tagctagagg	atcgatcccc	gccccggacg	aactaaacct	gactacgaca	tctctgcccc	3480
ttcttcgcg	ggcagtgcat	gtaatccct	cagttggtg	gtacaacttg	ccaactgggc	3540
cctgttccac	atgtgacacg	gggggggacc	aaacacaaag	gggttctctg	actgtagtgt	3600
acatccttat	aaatggatgt	gcacatttgc	caacactgag	tggttttcat	cctggagcag	3660
actttgcagt	ctgtggactg	caacacacaa	ttgcctttat	gtgtaactct	tggtgaagc	3720
tcttacacca	atgctggggg	acatgtacct	cccaggggcc	caggaagact	acgggaggct	3780
acaccaacgt	caatcagagg	ggcctgtgta	gctaccgata	agcggaccct	caagagggca	3840
ttagcaatag	tgtttataag	gcccccttgc	taacctaaa	cgggtagcat	atgcttccc	3900
ggtagtagta	tatactatcc	agactaacc	taattcaata	gcataatgta	cccaacggga	3960
agcatatgct	atcgaattag	ggttagtaaa	agggctcctaa	ggaacagcga	tatctccac	4020
cccatgagct	gtcacggttt	tatttacatg	gggtcaggat	tccacgagg	tagtgaacca	4080
ttttagtcac	aagggcagtg	gctgaagac	aaggagcggg	cagtgaactc	tctgaatct	4140
tcgctgctt	cttcattctc	cttcgtttag	ctaatagaat	aactgctgag	ttgtgaacag	4200
taaggtgtat	gtgaggtgct	cgaaaacaag	gtttcagtg	acgccccag	aataaaattt	4260
ggacggggg	ttcagtggtg	gcattgtgct	atgacaccaa	tataaccctc	acaaaccct	4320
tgggcaataa	atactagtgt	aggaatgaaa	cattctgaat	atctttaaca	atagaaatcc	4380
atggggtggg	gacaaaccgt	aaagactgga	tgtccatctc	acacgaattt	atggctatgg	4440
gcaacacata	atcctagtgc	aatatgatac	tggggttatt	aagatgtgtc	ccaggcagg	4500
accaagacag	gtgaaccatg	ttgttacact	ctatttgtaa	caaggggaaa	gagagtggac	4560
gccgacagca	gcggaactcca	ctggttgtct	ctaacacccc	cgaaaattaa	acggggctcc	4620
acgccaatgg	ggcccataaa	caaagacaag	tggccactct	ttttttgaa	attgtggagt	4680
gggggcacgc	gtcagccccc	acacgcgcgc	ctgcggtttt	ggactgtaaa	ataaggtgt	4740
aataacttgg	ctgattgtaa	ccccgctaac	cactgcggtc	aaaccacttg	cccacaaaac	4800
cactaatggc	accccgggga	atacctgcat	aagtaggtgg	gcgggccaag	ataggggcgc	4860

-continued

gattgctgcg atctggagga caaattacac acacttgccg ctgagcgcca agcacagggt	4920
tgttggctcct catattcacg aggtcgctga gagcacgggt ggctaagtgt gccatgggta	4980
gcataacta cccaaatata tggatagcat atgctatcct aatctatata tgggtagcat	5040
aggctatcct aatctatata tgggtagcat atgctatcct aatctatata tgggtagtat	5100
atgctatcct aatttatata tgggtagcat aggtatcct aatctatata tgggtagcat	5160
atgctatcct aatctatata tgggtagtat atgctatcct aatctgtata cgggtagcat	5220
atgctatcct aatagagatt agggtagtat atgctatcct aatttatata tgggtagcat	5280
atactacca aatatctgga tagcatatgc taccctaata tatatctggg tagcatatgc	5340
taccctaata tatatctggg tagcataggg taccctaata tatatctggg tagcatatgc	5400
taccctaata tatatctggg tagtatatgc taccctaatt tatatctggg tagcataggg	5460
taccctaata tatatctggg tagcatatgc taccctaata tatatctggg tagtatatgc	5520
taccctaata tgtatccggg tagcatatgc taccctcatg ataagctgtc aaacatgaga	5580
atcttcttga agacgaagg gcctcgtagt acgcctatct ttatagggtta atgtcatgat	5640
aataatgggt tcttagacgt cagggtggcag ttttcgggga aatgtgcgcg gaacccttat	5700
ttgtttatct ttctaataac attcaaatat gtatccgctc atgagacaat aacctgata	5760
aatgcttcaa taatatgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct	5820
tattcccttt ttgcgccat ttgcccctcc tgtttttgct caccagaaa cgtgggtgaa	5880
agtaaaagat gctgaagatc agttgggtgc acgagtggtg tacatcgaac tggatctcaa	5940
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt	6000
taaaattctg ctatgtggcg cggtattatc ccgtgttgac gccgggcaag agcaactcgg	6060
tcgcccata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca	6120
tcttacggat ggcatgacag taagagaatt atgcagtgtc gccataacca tgagtataaa	6180
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt	6240
gcacaacatg ggggatcatg taactgcctc tgatcggttg gaaccggagc tgaatgaagc	6300
cataccaaac gacgagcgtg acaccacgat gcctgcagca atggcaacaa cgttgccgaa	6360
actattaact ggccgaactac ttactctagc ttcccggcaa caattaatag actggatgga	6420
ggcgataaaa gttgcaggac cacttctgcg ctccggccct ccggctggct ggtttattgc	6480
tgataaatct ggagccgggt agcgtgggtc tcgcggtatc attgcagcac tggggccaga	6540
tggtaaagccc tcccgtatcg tagttatcta cagcagggg agtcaggcaa ctatggatga	6600
acgaaataga cagatcgctg agataggtgc ctactgatt aagcattggg aactgtcaga	6660
ccaagtctac tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat	6720
ctaggatgaag atcctttttg ataactctat gaccaaaatc ccttaacgtg agttttcgtt	6780
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct	6840
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc	6900
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc	6960
aaatactggt cttctagtgt agccgtagtt agggccaccac ttaagaact ctgtagcacc	7020
gcctacatac ctgcgtctgc taatcctgtt accagtggct gctgccagtg gcgataagtc	7080
gtgtcttacc ggggttgact caagacgata gttaccggat aaggcgcagc ggtcgggctg	7140
aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata	7200

-continued

```

cctacagcgt gagctatgag aaagcgccac gcttcccga gggagaaagg cggacaggta 7260
tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttcag ggggaaacgc 7320
ctggtatctt tatagtctcg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 7380
atgctcgtca gggggcgga gcctatgga aaacgccagc aacgcggcct ttttacgggt 7440
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 7500
ggataaccgt attaccgctt ttgagtgcgc tgataccgct cgccgcagcc gaacgaccga 7560
gcgcagcgag tcagtgcgc aggaagcgga agagcgccca atacgcaaac cgctctccc 7620
cgcggttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 7680
cagtgcgcgc aacgcaatta atgtgagta gctcactcat taggcacccc aggccttaca 7740
ctttatgctt ccggctcgta tgttgtgtg aattgtgagc ggataacaat ttcacacagg 7800
aaacagctat gaccatgatt acgccaagct ctagctagag gtcgagtcct tccccagcag 7860
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 7920
cgcccatccc gccctaact ccgcccagtt ccgcccattc tccgcccctt ggctgactaa 7980
ttttttttat ttatgcagag gccgagggcg cctcggcctc tgagctattc cagaagtagt 8040
gaggaggctt ttttgagggc ctaggctttt gcaaaaagct ttgcaaagat ggataaagtt 8100
ttaaacagag aggaatcttt gcagctaag gaccttctag gtcttgaaag g 8151

```

<210> SEQ ID NO 23

<211> LENGTH: 8151

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP181 ; pHybE-hCg2,n+ V1

<400> SEQUENCE: 23

```

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaagggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta 180
tataagtga gtagtcgccc tgaacgttct ttttcgcaac gggtttgccc ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgctg 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg 360
aagtgggtgg gagagtcca ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgtt ggggcccggc cgtgcgaatc tggtggaacc ttcgcgcctg 480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
ttttttcttg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gcgggcgggc acgggggccc tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctgggtgct ggctcgcgc cgccgtgtat cgcgccgccc tgggcggcaa ggctggccc 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc 840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcaccacac aaaggaaaag 900
ggcctttccg tctcagccg tcgcttcagt tgactccacg gactaccggg cgccgtccag 960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt 1020

```

-continued

ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080
cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggt	tcattctcaa	1140
gcctcagaca	gtggttcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgagc	atttaaatgc	cggggcgcca	ccatggagtt	tgggctgagc	1260
tggctttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggtatgc	cgaagggat	1320
gctgaaattg	agaacgaaaa	gctgcgcggg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaac	tattgagtac	gaacgccatc	gacttacgcg	tgcgcaggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tgggtgaaac	cgcattctgt	1500
actttctgtc	tgtcgcggat	cgcaggtgaa	attgccagta	ttctcgacgg	gtccccctg	1560
tcggtgcagc	ggcgttttcc	ggaactgaa	aaccgacatg	tgatttcct	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagcccg	ctggatgaac	tgataccggg	gttgcctgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgcgg	ggcttcgctc	1740
actgttcagg	ccggagccac	agaccgcgt	tgaatggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaacactg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggtgaatgtg	gtgaagtctg	cccgtgtcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
cettatctgg	ttgccagcgg	atgggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggcccg	tggatatgga	aaaagcaccg	2100
ggataaacag	ctcaccatga	agcgtttcac	taatgggct	ggcttctggt	gcctgggcgg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggtggatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttctgggtg	acaagcgcgt	2280
cgaccaaggg	cccatcggtc	ttccccctgg	cgcctgctc	tagaagcacc	tccgagagca	2340
cagcggccct	gggctgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	tctgaccage	ggcgtgcaca	ccttcccagc	tgtcctgcag	tcctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtga	cctccagcaa	cttcggcacc	cagacctaca	2520
catgcaacgt	agatcacaag	cccagcaaca	ccaagggtgga	caagacagtt	gagcgcaaat	2580
gttgtgtcga	gtgcccaccg	tgcccagcac	cacctgtggc	aggaccgtca	gtcttcctct	2640
tcccccaaaa	acccaaggac	accctcatga	tctcccgac	ccctgaggtc	acgtgcgtgg	2700
tgggtggcgt	gagccacgaa	gaccccgagg	tccagttcaa	ctggtacgtg	gacggcatgg	2760
aggtgcataa	tgccaagaca	aagccacggg	aggagcagtt	caacagcacg	ttcgtgtgg	2820
tcagcgtcct	caccgttgtg	caccaggact	ggctgaacgg	caaggagtac	aagtgcagg	2880
tctccaacaa	aggcctccca	gcccccatcg	agaaaaccat	ctccaaaacc	aaagggcagc	2940
cccgagaacc	acaggtgtac	accctgcccc	catcccgga	ggagatgacc	agaaccagg	3000
tcagcctgac	ctgcctggtc	aaaggcttct	accccagcga	catgcctgtg	gagtgggaga	3060
gcaatgggca	gccgggagac	aactacaaga	ccacacctcc	catgctggac	tccgacggct	3120
ccttctctct	ctacagcaag	ctcacgtgg	acaagagcag	gtggcagcag	gggaacgtct	3180
tctcatgctc	cgtgatgcat	gaggctctgc	acaaccacta	cacgcagaag	agcctctccc	3240
tgtctccggg	taaatgacgc	ggccgctcga	ggccggcaag	gccgatccc	ccgacctcga	3300
cctctggcta	ataaaggaaa	tttattttca	ttgcaatagt	gtgttgaat	ttttgtgtc	3360

-continued

tctcactcgg aaggacatat gggagggcaa atcatttggg cgagatccct cggagatctc	3420
tagctagagg atcgatcccc gcccgggacg aactaaacct gactacgaca tctctgcccc	3480
ttcttcgchg ggcagtgcac gtaatccctt cagttggttg gtacaacttg ccaactgggc	3540
cctgttccac atgtgacacg gggggggacc aaacacaaag gggttctctg actgtagttg	3600
acatccctat aaatggatgt gcacatttgc caacactgag tggctttcat cctggagcag	3660
actttgcagt ctgtggactg caacacacaa ttgcctttat gtgtaactct tggctgaagc	3720
tcttacacca atgctggggg acatgtacct ccagggggcc caggaagact acgggaggct	3780
acaccaacgt caatcagagg ggcctgtgta gctaccgata agcggaccct caagagggca	3840
ttagcaatag tgtttataag gccccttgtg taaccctaaa cgggtagcat atgcttcccg	3900
ggtagtagta tatactatcc agactaacc taattcaata gcatatgtta cccaacggga	3960
agcatatgct atcgaattag ggttagtaaa agggctcctaa ggaacagcga tatctccac	4020
cccatgagct gtcacgggtt tatttacatg gggtcaggat tccacgagg tagtgaacca	4080
ttttagtcac aagggcagtg gctgaagatc aaggagcggg cagtgaactc tcctgaatct	4140
tcgcctgctt cttcattctc ctctgtttag ctaatagaat aactgctgag ttgtgaacag	4200
taaggtgat gtgaggtgct cgaatacaag gtttcagggtg acgccccag aataaaattt	4260
ggacgggggg ttcagtgggt gcattgtgct atgacaccaa tataaccctc acaaacccct	4320
tgggcaataa atactagtgt aggaatgaaa cattctgaat atctttaaca atagaaatcc	4380
atggggtggg gacaagccgt aaagactgga tgtccatctc acacgaattt atggctatgg	4440
gcaacacata atcctagtgc aatatgatac tggggttatt aagatgtgtc ccaggcaggg	4500
accaagacag gtgaaccatg ttgttacct ctatttgtaa caaggggaaa gagagtggac	4560
gcccagacga ggcgactcca ctggttgtct ctaacacccc cgaaaattaa acggggctcc	4620
acgccaatgg ggcacataaa caaagacaag tggccactct tttttttaa attgtggagt	4680
gggggcacgc gtcagcccc acacgccgcc ctgcgggttt ggactgtaaa ataaggtgt	4740
aataacttgg ctgattgtaa ccccgcctaac cactgcggtc aaaccacttg cccacaaaac	4800
cactaatggc accccgggga atacctgcat aagtaggtgg gcgggccaag atagggcgcg	4860
gattgtctgc atctggagga caaattacac acacttgcgc ctgagcgcca agcacagggt	4920
tgttggctct catattcacg aggtcgctga gagcacggtg ggctaattgt gccatgggta	4980
gcatatacta cccaaatata tggatagcat atgctatcct aatctatata tgggtagcat	5040
aggctatcct aatctatata tgggtagcat atgctatcct aatctatata tgggtagtat	5100
atgctatcct aatttatata tgggtagcat aggctatcct aatctatata tgggtagcat	5160
atgctatcct aatctatata tgggtagtat atgctatcct aatctgtata cgggtagcat	5220
atgctatcct aatagagatt agggtagtat atgctatcct aatttatata tgggtagcat	5280
atactacca aatatctgga tagcatatgc taccctaata tatatctggg tagcatatgc	5340
taccctaata tatatctggg tagcataggg taccctaata tatatctggg tagcatatgc	5400
taccctaata tatatctggg tagtatatgc taccctaatt tatatctggg tagcataggg	5460
taccctaata tatatctggg tagcatatgc taccctaata tatatctggg tagtatatgc	5520
taccctaata tgtatccggg tagcatatgc taccctcatg ataagctgtc aaacatgaga	5580
atcttcttga agacgaagg gcctcgtgat acgcctatct ttatagggtta atgtcatgat	5640

-continued

aataatgggtt	tcttagacgt	caggtggcac	ttttcgggga	aatgtgcgcg	gaacccctat	5700
ttgtttat	ttctaaat	attcaaat	gtatccgctc	atgagacaat	aacctgata	5760
aatgcttcaa	taatattgaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	5820
tattcccttt	tttgcggcat	tttgccttcc	tgtttttgct	caccagaaa	cgctggtgaa	5880
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	5940
cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	tgagcacttt	6000
taaagttctg	ctatgtggcg	cggtattatc	cgtgttgac	gccgggcaag	agcaactcgg	6060
tcgccgcata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	6120
tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	tgagtataaa	6180
cactgcggcc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	ccgctttttt	6240
gcacaacatg	gggatcatg	taactcgctt	tgatcggttg	gaaccggagc	tgaatgaagc	6300
catacctaac	gacgagcgtg	acaccacgat	gcctgcagca	atggcaacaa	cgttgcgcaa	6360
actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	actggatgga	6420
ggcggataaa	gttgcaggac	cacttctgcg	ctcgcccttt	ccggctgggt	ggtttattgc	6480
tgataaatct	ggagccgggt	agcgtgggtc	tcgcggtatc	attgcagcac	tggggccaga	6540
tggttaagccc	tcccgatctg	tagttatcta	cacgacgggg	agtcaggcaa	ctatggatga	6600
acgaaataga	cagatcgctg	agataggtgc	ctcactgatt	aagcattggg	aactgtcaga	6660
ccaagtctac	tcatatatac	tttagattga	tttaaaactt	catttttaat	ttaaaaggat	6720
ctaggtgaag	atcctttttg	ataatctcat	gacaaaaatc	ccttaacgtg	agttttcggt	6780
ccactgagcg	tcagaccccg	tagaaaagat	caaaggatct	tcttgagatc	ctttttttct	6840
gcgcgtaatc	tgcgtcttgc	aaacaaaaaa	accaccgcta	ccagcgggtg	ttgttttgcc	6900
ggatcaagag	ctaccaactc	tttttcgcaa	ggtaactggc	ttcagcagag	cgcagatacc	6960
aaatactggt	cttctagtgt	agccgtagtt	agggccaccac	ttcaagaact	ctgtagcacc	7020
gcctacatac	ctcgctctgc	taatcctgtt	accagtgggt	gctgccagtg	gcgataagtc	7080
gtgtcttacc	gggttggtg	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	7140
aacgggggggt	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	7200
cctacagcgt	gagctatgag	aaagcgcac	gcttcccga	gggagaaaag	cggacaggta	7260
tccggtaagc	ggcagggctg	gaacaggaga	gcgcacgagg	gagcttcag	ggggaacgc	7320
ctggtatctt	tatagtcctg	tcgggtttcg	ccacctctga	cttgagcgtc	gattttttgtg	7380
atgctcgtca	ggggggcgga	gcctatggaa	aaacgccagc	aacgcggcct	ttttacgggt	7440
cctggccttt	tgcctgcctt	ttgctcacat	gttctttcct	gcgttatccc	ctgattctgt	7500
ggataaccgt	attaccgcct	ttgagtgagc	tgataccgct	cgcgcagacc	gaacgaccga	7560
gcgcagcgag	tcagtgcgag	aggaagcgga	agagcgccca	atacgcaaac	cgcctctccc	7620
cgcgcgttgg	ccgattcatt	aatgcagctg	gcacgacagg	tttcccgact	ggaaagcggg	7680
cagtgcgcgc	aacgcaatta	atgtgagtta	gctcactcat	taggcacccc	aggctttaca	7740
ctttatgctt	ccggctcgta	tggtgtgtgg	aattgtgagc	ggataacaat	ttcacacagg	7800
aaacagctat	gacctgatt	acgccaagct	ctagctagag	gtcaggtccc	tccccagcag	7860
gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	catagtcccg	cccctaactc	7920
cgcccatccc	gcccctaact	ccgcccagtt	ccgcccattc	tccgcccct	ggctgactaa	7980

-continued

```

ttttttttat ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt 8040
gaggaggcctt ttttgaggc ctaggctttt gcaaaaagct ttgcaaagat ggataaagtt 8100
ttaaacagag aggaatcttt gcagctaatt gaccttctag gtcttgaaag g 8151

```

<210> SEQ ID NO 24

<211> LENGTH: 8151

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP186 ; pHybE-hCg2,n+ V2

<400> SEQUENCE: 24

```

agtgggaatt ggctccggcg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggaggggctg gcaattgaac cgggtcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gagggtgagg gagaaccgta 180
tataagtga gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctg 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttg 360
aagtgggtgg gagagtgcga ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tggtggcacc ttcgcgcctg 480
tctcgtgct ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gggggcgccg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctgggtcct ggctcgcgc gcgcgtgtat cgcgccgccc tgggcggcaa ggcgggcccg 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg cagggaagctc 840
aaaaatggag acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag 900
ggcctttccg tcctcagccg tcgcttcctg tgactccacg gactaccggg cgcgcgtccg 960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt 1020
ttatgcgatg gattttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca 1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa 1140
gcctcagaca gtggttcaaa gtttttttct tcattttcag gtgtcgtgag gaattctcta 1200
gagatccctc gacctcgaga tcattgtgc cggggcgcca ccatggagtt tgggctgagc 1260
tggttttttc ttgtcgcgat tttaaaagg gtccagtgcg catggtatgc cgaaagggat 1320
gctgaaattg agaacgaaa gctgcgcggg gaggttgaag aactgcgga gccagcgag 1380
gcagatctcc agccaggaac tattgagtag gaacgccatc gacttacgcg tgcgcaggcc 1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaac cgcattctgt 1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg 1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat 1620
atcatcaaag ccatgaacaa agcagcccggt ctggatgaac tgataccggg gttgctgagt 1680
gaatatatcg aacagtcagg ttaacaggct cgggcatttt gtcgcgcgg ggcttcgctc 1740
actgttcagg ccggagccac agaccgcgt tgaatggcg gatgctaatt actatctccc 1800

```

-continued

gaaagaatcc	gcataaccagg	aagggcgctg	ggaaacactg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggtgaatgtg	gtgaagtctg	cccggtgcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccgacgg	atggtgatgc	cgagaacttt	atgaaaaccc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggccccg	tggtatggca	aaaagcaccg	2100
ggataaacag	ctcaccatga	agcgtttcac	taatgggcgt	ggcttctggg	gcctgggcgg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggtggatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttctgggtg	acaagcgct	2280
cgaccaaggg	cccatcggtc	ttccccctgg	cgccctgctc	tagaagcacc	tccgagagca	2340
cagcggccct	gggctgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	tctgaccagc	ggcgtgcaca	ccttcccagc	tgctctacag	tctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtaa	cctccagcaa	cttcggcacc	cagacctaca	2520
catgcaacgt	agatcacaa	cccagcaaca	ccaagggtga	caagacagtt	gagcgcaa	2580
ggtgtgtcga	gtgcccaccg	tgcccagcac	cacctgtggc	aggaccgtca	gtcttcctct	2640
tcccccaaaa	acccaaggac	accctcatga	tctcccgac	ccctgaggtc	acgtgcgtgg	2700
tggtggacgt	gagccacgaa	gaccccgagg	tccagttcaa	ctggtacgtg	gacggcatgg	2760
agggtgataa	tgccaagaca	aagccacggg	aggagcagtt	caacagcacg	ttccgtgtgg	2820
tcagcgtcct	caccgttgtg	caccaggact	ggctgaacgg	caaggagtac	aagtgcaagg	2880
tctccaacaa	aggcctccca	gcccccatcg	agaaaacat	ctccaaaacc	aaagggcagc	2940
cccgagaacc	acaggtgtac	accctgcccc	catcccggga	ggagatgacc	aagaaccagg	3000
tcagcctgac	ctgcctggtc	aaaggcttct	accccagcga	catcgccgtg	gagtgggaga	3060
gcaatgggca	gccggagaa	aactacaaga	ccacacctcc	catgctggac	tccgacggct	3120
ccttcttctc	ctacagcaag	ctcacctggg	acaagagcag	gtggcagcag	gggaacgtct	3180
tctcatgctc	cgtgatgcat	gaggctctgc	acaaccacta	cacgcagaag	agcctctccc	3240
tgctctcggg	taaatgacgc	ggcgcctcga	ggcgggcaag	gccggatccc	ccgacctcga	3300
cctctggcta	ataaaggaaa	tttattttca	ttgcaatagt	gtgttggaat	tttttgtgtc	3360
tctcactcgg	aaggacatat	gggagggcaa	atcatttggg	cgagatccct	cggagatctc	3420
tagctagagg	atcgatcccc	gccccggacg	aactaaacct	gactacgaca	tctctgcccc	3480
ttcttcggcg	ggcagtgcat	gtaatccctt	cagttggttg	gtacaacttg	ccaactgggc	3540
cctgttccac	atgtgacacg	gggggggacc	aaacacaaag	gggttctctg	actgtagtgtg	3600
acatccttat	aaatggatgt	gcacatttgc	caacactgag	tggttttcat	cctggagcag	3660
actttgcagt	ctgtggactg	caacacaaca	ttgcctttat	gtgtaactct	tggtgaagc	3720
tcttacacca	atgctggggg	acatgtacct	cccagggggc	cagggaagact	acgggaggct	3780
acaccaacgt	caatcagagg	ggcctgtgta	gctaccgata	agcggaccct	caagagggca	3840
ttagcaatag	tgtttataag	gcccccttgt	taaccctaaa	cgggtagcat	atgcttcccg	3900
ggtagtagta	tatactatcc	agactaacc	taattcaata	gcatatgtta	cccaacggga	3960
agcatatgct	atcgaattag	ggtagtaaa	agggtcctaa	ggaacagcga	tatctccac	4020
cccatagagct	gtcacggttt	tatttacatg	gggtcaggat	tccacgaggg	tagtgaacca	4080
tttagtcac	aagggcagtg	gctgaagatc	aaggagcggg	cagtgaactc	tctgaatct	4140

-continued

tcgectgctt cttcattctc cttcgtttag ctaatagaat aactgctgag ttgtgaacag	4200
taagggtgat gtgaggtgct cgaaaacaag gtttcaggtg acgccccag aataaaattt	4260
ggacgggggg ttcagtgggt gcattgtgct atgacaccaa tataaccctc acaaacccct	4320
tgggcaataa atactagtgt aggaatgaaa cattctgaat atctttaaca atagaaatcc	4380
atgggggtgg gacaagccgt aaagactgga tgtccatctc acacgaattt atggctatgg	4440
gcaacacata atcctagtgc aatatgatac tggggttatt aagatgtgtc ccaggcaggg	4500
accaagacag gtgaaccatg ttgttacact ctatttgtaa caaggggaaa gagagtggac	4560
gccgacagca gcggactcca ctggttgtct ctaacacccc cgaaaattaa acggggctcc	4620
acgccaatgg ggcccataaa caaagacaag tggccactct tttttttgaa attgtggagt	4680
gggggcacgc gtcagcccc acacgccgc ctgcggtttt ggactgtaaa ataagggtgt	4740
aataacttgg ctgattgtaa ccccgctaac cactgcggtc aaaccacttg cccacaaaac	4800
cactaatggc acccggggga atacctgcat aagtaggtgg gcgggccaag ataggggcgc	4860
gattgtgcgc atctggagga caaattacac aacttgcgc ctgagcgcca agcacagggt	4920
tgttggtcct catattcacg aggtcgctga gagcacggtg ggctaattgt gccatgggta	4980
gcataacta cccaaatate tggatagcat atgctatcct aatctatate tgggtagcat	5040
aggctatcct aatctatate tgggtagcat atgctatcct aatctatate tgggtagcat	5100
atgctatcct aatttatate tgggtagcat aggctatcct aatctatate tgggtagcat	5160
atgctatcct aatctatate tgggtagcat atgctatcct aatctatate tgggtagcat	5220
atgctatcct aatagagatt agggtagtat atgctatcct aatttatate tgggtagcat	5280
atactaccca aatatctgga tagcatatgc taccctaate tatatctggg tagcatatgc	5340
taccctaate tatatctggg tagcataggc taccctaate tatatctggg tagcatatgc	5400
taccctaate tatatctggg tagtatatgc taccctaatt tatatctggg tagcataggc	5460
taccctaate tatatctggg tagcatatgc taccctaate tatatctggg tagtatatgc	5520
taccctaate tgtatccggg tagcatatgc taccctcatg ataagctgtc aaacatgaga	5580
atcttcttga agacgaaagg gcctcgtgat acgcctatct ttatagggtta atgtcatgat	5640
aataatgggt tcttagacgt cagggtggc acgttgggga aatgtgcgcg gaaccctat	5700
ttgtttatct ttctaaatgc attcaaatat gtatccgctc atgagacaat aacctgata	5760
aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct	5820
tattcccttt ttgcgcat tttgccttcc tgtttttgct caccagaaa cgctggtgaa	5880
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa	5940
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt	6000
taaaagtctg ctatgtggcg cgggtattac ccgtgttgac gccgggcaag agcaactcgg	6060
tcgccgcata cactattctc agaatgactt gggtgagtac tcaccagtca cagaaaagca	6120
tcttacggat ggcatgacag taagagaatt atgcagtgtc gccataacca tgagtgataa	6180
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt	6240
gcacaacatg ggggatcatg taactgcct tgatcgttgg gaaccggagc tgaatgaagc	6300
cataccaaac gacgagcgtg acaccacgat gcctgcagca atggcaacaa cgttgcgcaa	6360
actattaact ggcaactac ttactctagc ttcccgcaa caattaatag actggatgga	6420

-continued

```

ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggtggct ggtttattgc 6480
tgataaatct ggagccgggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 6540
tggtaaagccc tcccgatcgc tagttatcta cacgacgggg agtcaggcaa ctatggatga 6600
acgaaataga cagatcgctg agatagggtc ctcactgatt aagcattggt aactgtcaga 6660
ccaagtttac tcatatatac tttagattga tttaaaactt cttttttaat ttaaaaggat 6720
ctagggtgaag atcctttttg ataatctcat gaccaaaac ccttaacgtg agttttcgtt 6780
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 6840
gcgcgtaate tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgc 6900
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 6960
aaatactggt cttctagtgt agccgtagt agggccaccac ttcaagaact ctgtagcacc 7020
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 7080
gtgtcttacc ggggttgact caagacgata gttaccggat aaggcgcagc ggtcgggctg 7140
aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 7200
cctacagcgt gagctatgag aaagcgcac gcttcccgaa gggagaaaagg cggacaggta 7260
tccggtaaag gccagggctg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 7320
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 7380
atgctcgtca gggggggcga gcctatggaa aaacgccagc aacgcggcct ttttacgggt 7440
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 7500
ggataaccgt attaccgctt ttgagtgcgc tgataccgct cgcgcagacc gaacgaccga 7560
gcgcagcgag tcagtgcgc aggaagcggga agagcgccca atacgcaaac cgctctccc 7620
cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 7680
cagtgcgcgc aacgcaatta atgtgagtta gctcactcat taggcacccc aggcctttaca 7740
ctttatgctt ccggctcgta tggtgtgtg aattgtgagc ggataacaat ttcacacagg 7800
aaacagctat gaccatgatt acgccaagct ctactagag gtcgagtcct tcccagcag 7860
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 7920
cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgcccctat ggtgactaa 7980
ttttttttat ttatgcagag gccgagggcg cctcggcctc tgagctattc cagaagtagt 8040
gaggaggcct ttttgaggcg ctaggctttt gcaaaaagct ttgcaaagat ggataaagtt 8100
ttaaacagag aggaatcttt gcagctaattg gaccttctag gtcttgaaag g 8151

```

<210> SEQ ID NO 25

<211> LENGTH: 8153

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP175 ; pHyE-hCg4 V1

<400> SEQUENCE: 25

```

agtgggaatt ggctccgggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggaggggtcg gcaattgaac cgggtcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta 180
tataagtga gtagtcgcg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggccc tggcctcttt acgggttatg gcccttgct 300

```

-continued

gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg	360
aagtgggtgg gagagtccga ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcctg ggggcgcgcg cgtgcgaatc tggtggcacc ttcgcgcctg	480
tctcgtctgt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcgggcgccg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg	660
gggcctgcga gcgcgccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc	720
tctggtgcct ggctcgcgc cgcctgtgat cgcctccccc tggcgggcaa ggctggcccg	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc	840
aaaaatggagg acgcggcctg cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tctcagccg tcgcttcagt tgactccacg gactaccggg cgcctccag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gaggttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttctt tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gacctcgagc atttaaatgc ccgggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaagggt gtccagtgcg catggtatgc cgaagggat	1320
gctgaaattg agaacgaaaa gctgcgcggg gaggttgaag aactgcggca ggcagcgag	1380
gcagatctcc agccaggaac tattgagtac gaacgccatc gacttacgcg tgcgcaggcc	1440
gacgcacagg aactgaagaa tgccagagac tccgctgaag tggtggaac cgcattctgt	1500
actttcgtgc tgtcgcggat cgcaggtgaa attgccagta ttctcgacgg gctccccctg	1560
tcggtgcagc ggcgttttcc ggaactggaa aaccgacatg ttgatttctt gaaacgggat	1620
atcatcaaa ccatgaacaa agcagcccg ctaggatgaa tgataccggg gttgctgagt	1680
gaatatatcg aacagtcagg ttaacaggct gcggcatttt gtccgcgcgg ggttcgctc	1740
actgttcagg ccggagccac agaccgcgt tgaatggcg gatgctaatt actatctccc	1800
gaaagaatcc gcataccagg aaggcgctg ggaaacactg ccttttcagc gggccatcat	1860
gaatgcgatg ggcagcgact acatccgtga ggtgaatgtg gtgaagtctg cccgtgtcgg	1920
ttattccaaa atgctgctgg gtgtttatgc ctactttata gagcataagc agcgaacac	1980
ccttatctgg ttgccagcg atggtgatgc cgagaacttt atgaaaaccc acgttgagcc	2040
gactattcgt gatattccgt cgtgctggc gctggccccg tggatatgga aaaagcaccg	2100
ggataaacag ctcaccatga agcgtttcac taatgggcgt ggcttctggg gcctggcg	2160
taaagcgga aaaaactacc gtgaaaagtc ggtggatgtg gcgggttatg atgaacttgc	2220
tgttttgat gatgatattg aacaggaagg ctctccgacg ttcttgggtg acaagcgct	2280
cgaccaaggg cccatccgtc tccccctgg cgcctgctc caggagcacc tccgagagca	2340
cagcgccctt gggtgcctg gtcaaggact acttccccga accggtgacg gtgtcgtgga	2400
actcaggcgc cctgaccagc ggcgtgcaca ccttccccgc tgtcctacag tctcaggac	2460
tctactccct cagcagcgtg gtgacgtgc cctccagcag cttgggcacg aagacctaca	2520
cctgcaatgt agatcacaag ccagcaaca ccaagggtga caagagagtt gagtccaaat	2580

-continued

acggtccgcc	atgcccatca	tgccagcac	ctgaattcct	ggggggacca	tcagtcttcc	2640
tggtccccc	aaaacccaag	gacacctca	tgatctcccg	gacctctgag	gtcacgtgcg	2700
tggtggtgga	cgtgagccag	gaagacccc	aggctcagtt	caactggtac	gtggatggcg	2760
tgagggtgca	taatgccaa	acaaagccg	gggaggagca	gttcaacagc	acgtaccgtg	2820
tggtcagcgt	cctcaccgtc	ctgcaccagg	actggctgaa	cggcaaggag	tacaagtgca	2880
aggctctcaa	caaaggcctc	ccgtctctca	tcgagaaaac	catctccaaa	gccaaagggc	2940
agccccgaga	gccacagggt	tacacctgc	ccccatcca	ggaggagatg	accaagaacc	3000
aggtcagcct	gacctgcctg	gtcaaaggct	tctaccccag	cgacatcgcc	gtggagtggg	3060
agagcaatgg	gcagccggag	aacaactaca	agaccacgcc	tccgtgctg	gactccgacg	3120
gctcctctct	cctctacagc	aggctaaccg	tggacaagag	caggtggcag	gaggggaatg	3180
tcttctcatg	ctccgtgatg	catgaggctc	tgcacaacca	ctacacgcag	aagagcctct	3240
ccctgtctct	gggtaaatga	gcggccgctc	gaggccggca	aggccggatc	ccccgacctc	3300
gacctctggc	taataaagga	aatttatctt	cattgcaata	gtgtgttgga	attttttgtg	3360
tctctcactc	ggaagacat	atgggagggc	aatcatttg	gtcgagatcc	ctcgagatc	3420
tctagctaga	ggatcgatcc	ccgccccgga	cgaactaaac	ctgactacga	catctctgcc	3480
ccttctctgc	ggggcagtgc	atgtaatccc	ttcagttggg	tggtaacaact	tgccaactgg	3540
gcctgttcc	acatgtgaca	cgggggggga	ccaaacacaa	aggggttctc	tgactgtagt	3600
tgacatcctt	ataaatggat	gtgcacattt	gccaaactg	agtggctttc	atcctggagc	3660
agactttgca	gtctgtggac	tgcaacacaa	cattgccttt	atgtgtaact	cttggctgaa	3720
gctcttacac	caatgctggg	ggacatgtac	ctcccagggg	cccaggaaga	ctacgggagg	3780
ctacaccaac	gtcaatcaga	ggggcctgtg	tagctaccga	taagcggacc	ctcaagaggg	3840
cattagcaat	agtgtttata	aggccccctt	gttaacccta	aacgggtagc	atatgcttcc	3900
cgggtagtag	tatatactat	ccagactaac	cctaattcaa	tagcatatgt	tacccaacgg	3960
gaagcatatg	ctatcgat	agggttagta	aaagggtcct	aaggacagc	gatatctccc	4020
accccatgag	ctgtcacggg	tttatcttca	tggggtcagg	attccacgag	ggtagtgaa	4080
catttttagtc	acaagggcag	tggctgaaga	tcaaggagcg	ggcagtgaac	tctcctgaat	4140
cttcgcctgc	ttcttcatc	tccttcgttt	agctaataga	ataactgctg	agttgtgaac	4200
agtaagggtg	atgtgagggt	ctcgaaaaca	aggtttcagg	tgacgcccc	agaataaaat	4260
ttggacgggg	gggtcagtg	tggcattgtg	ctatgacacc	aatataaccc	tcacaaaccc	4320
cttgggcaat	aaatactagt	gtaggatga	aacattctga	atatctttaa	caatagaaat	4380
ccatgggggtg	gggacaagcc	gtaaagactg	gatgtccatc	tcacacgaat	ttatggctat	4440
gggcaacaca	taatcctagt	gcaatatgat	actgggggta	ttaagatgtg	tcccaggcag	4500
ggaccaagac	agggtgaacca	tggtgttaca	ctctatttgt	aacaagggga	aagagagtgg	4560
acgccgacag	cagcggactc	cactgggtgt	ctctaaccac	cccgaataat	aaacggggct	4620
ccacgccaat	ggggcccata	aacaaagaca	agtggccact	cttttttttg	aaattgtgga	4680
gtggggggc	gcgtcagccc	ccacacgcg	ccctgcgggt	tggactgta	aaataagggt	4740
gtaataaact	ggctgattgt	aaccccgcta	accactgcgg	tcaaacact	tgcccacaaa	4800
accactaatg	gcaccccg	gaatacctgc	ataagtaggt	gggcgggcca	agataggggc	4860
gcgattgctg	cgatctggag	gacaaattac	acacacttgc	gcctgagcgc	caagcacagg	4920

-continued

gttggtggtc ctcataattca cgaggtcgct gagagcacgg tgggctaattg ttgccatggg	4980
tagcatatac tacccaaata tctggatagc atatgctatc ctaatctata tctgggtagc	5040
ataggctatc ctaatctata tctgggtagc atatgctatc ctaatctata tctgggtagt	5100
atatgctatc ctaatttata tctgggtagc ataggctatc ctaatctata tctgggtagc	5160
atatgctatc ctaatctata tctgggtagt atatgctatc ctaatctgta tccgggtagc	5220
atatgctatc ctaatagaga ttagggtagt atatgctatc ctaatttata tctgggtagc	5280
atatactacc caaatatctg gatagcatat gctatcctaa tctatatctg ggtagcatat	5340
gctatcctaa tctatatctg ggtagcatag gctatcctaa tctatatctg ggtagcatat	5400
gctatcctaa tctatatctg ggtagtatat gctatcctaa tttatatctg ggtagcatag	5460
gctatcctaa tctatatctg ggtagcatat gctatcctaa tctatatctg ggtagtatat	5520
gctatcctaa tctgtatccg ggtagcatat gctatcctca tgataagctg tcaaacaatga	5580
gaattttctt gaagacgaaa gggcctcgctg atacgcctat ttttataggt taatgtcatg	5640
ataataatgg tttcttagac gtcaggtggc acttttcggg gaaatgtgcg cggaaccct	5700
atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga	5760
taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt cctgtgcgcc	5820
cttattccct tttttcggcg attttgcctt cctgtttttg ctcaccaga aacgctggcg	5880
aaagtaaaag atgctgaaga tcagttgggt gcacgagtggt gttacatcga actggatctc	5940
aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact	6000
tttaaagttc tgetatgtgg cgcggtatta tcccgtgttg acgcccggca agagcaactc	6060
ggtcgccgca tacactatctc tcagaatgac ttggttgagt actcaccagt cacagaaaag	6120
catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat	6180
aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt	6240
ttgcacaaca tgggggatca tgtaactcgc cttgatcggt gggaaccgga gctgaatgaa	6300
gccataccaa acgacgagcg tgacaccacg atgcctgcag caatggcaac aacgttgccg	6360
aaactattaa ctggcgaaact acttactcta gcttcccggc aacaattaat agactggatg	6420
gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt	6480
gctgataaat ctggagccgg tgagcgtggg tctcgggta tcattgcagc actggggcca	6540
gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat	6600
gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca	6660
gaccaagttt actcatatat acttttagatt gattttaaac ttcattttta attttaaagg	6720
atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg	6780
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tctttttttt	6840
ctgcgcgtaa tctgtgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg	6900
ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	6960
ccaaatactg tcttcttagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca	7020
ccgctacat acctcgctct gctaactctg ttaccagtggt ctgctgccag tggcgataag	7080
tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	7140
tgaacggggg gtctgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	7200

-continued

tacctacagc	gtgagctatg	agaaagcgcc	acgcttcccg	aaggagagaaa	ggcggacag	7260
tatccggtaa	gcggcaggg	cggaacagga	gagcgacga	gggagcttcc	aggggaaac	7320
gcctggtatc	tttatagtcc	tgctgggttt	cgccacctct	gacttgagcg	tcgatttttg	7380
tgatgctcgt	cagggggg	gagcctatgg	aaaaacgcca	gcaacgcggc	ctttttacgg	7440
ttcctggcct	tttgcggcc	ttttgctcac	atgttctttc	ctgcgttacc	cctgattct	7500
gtggataacc	gtattaccgc	ctttgagtga	gctgataccg	ctcgccgag	cgaacgacc	7560
gagcgagcg	agtcagtga	cgaggaagcg	gaagagcgcc	caatacgcaa	accgcctctc	7620
cccgcgctt	ggcgcattca	ttaatgcagc	tggcacgaca	ggtttccga	ctggaaagcg	7680
ggcagtgagc	gcaacgcaat	taatgtgagt	tagctcactc	attaggcacc	ccaggcttta	7740
cactttatgc	ttccggctcg	tatgttgtgt	ggaattgtga	gcggataaca	atttcacaca	7800
ggaaacagct	atgaccatga	ttacgccaa	ctctagctag	aggtcgagtc	cctccccagc	7860
aggcagaagt	atgcaaa	gcacatctca	ttagtcagca	accatagtcc	cgccccctaac	7920
ttccgccatc	ccgccctcaa	ctccgccag	ttccgccat	tctccgccc	atggctgact	7980
aatttttttt	atttatgcag	aggccgaggg	cgccctcgcc	tctgagctat	tccagaagta	8040
gtgaggaggc	ttttttggag	gcctaggctt	ttgcaaaaag	ctttgcaaag	atggataaag	8100
ttttaaacag	agaggaatct	ttgcagctaa	tggaccttct	aggtcttgaa	agg	8153

<210> SEQ ID NO 26

<211> LENGTH: 8153

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pJP188 ; pHybE-hCg4 V2

<400> SEQUENCE: 26

agtgggaatt	ggctccgggt	cccgtcagtg	ggcagagcgc	acatcgccca	cagtccccca	60
gaagttgggg	ggaggggtcg	gcaattgaac	cggtgcctag	agaaggtggc	gcggggtaaa	120
ctgggaaagt	gatgtcgtgt	actggtccg	cctttttccc	gagggtgggg	gagaaccgta	180
tataagtga	gtagtcgcc	tgaacgttct	ttttcgcaac	gggtttgccg	ccagaacaca	240
ggtaagtgcc	gtgtgtggtt	cccgccggcc	tggcctcttt	acgggttatg	gcccttgct	300
gccttgaatt	acttccacct	ggctgcagta	cgtgattctt	gatcccgagc	ttcgggttgg	360
aagtgggtgg	gagagtccga	ggccttgcc	ttaaggagcc	ccttcgcctc	gtgcttgagt	420
tgaggcctgg	cctgggcgct	ggggccgcgg	cgtgcgaatc	tggtaggcacc	ttcgcgcctg	480
tctcgtgct	ttcgataagt	ctctagccat	ttaaaatttt	tgatgacctg	ctgcgacgct	540
ttttttctgg	caagatagtc	ttgtaaatgc	gggccaagat	ctgcacactg	gtatttcggt	600
ttttggggcc	gcgggcggcg	acggggcccg	tgcgtcccag	cgcacatggt	cggcgaggcg	660
gggcctgcga	gcgcggccac	cgagaatcgg	acgggggtag	tctcaagctg	gccggcctgc	720
tctggtgcct	ggcctcgcc	cgccgtgtat	cgccccgccc	tggcgggcaa	ggctggcccc	780
gtcggcacca	gttgcgtgag	cgaaagatg	gccgcttccc	ggccctgctg	caggagagctc	840
aaaatggagg	acgcggcgct	cgggagagcg	ggcgggtgag	tcaccacac	aaaggaaaag	900
ggcctttccg	tcctcagccg	tcgcttcagt	tgactccacg	gagtaccggg	cgccgtccag	960
gcacctcgat	tagttctcga	gcttttggag	tacgtcgtct	ttaggttggg	gggaggggtt	1020
ttatgcgatg	gagtttcccc	acactgagtg	ggtggagact	gaagttaggc	cagcttgcca	1080

-continued

cttgatgtaa	ttctccttgg	aatttgcctt	ttttgagttt	ggatcttggg	tcattctcaa	1140
gcctcagaca	gtgggtcaaa	gtttttttct	tccatttcag	gtgtcgtgag	gaattctcta	1200
gagatccctc	gacctcgaga	tccattgtgc	cggggcgcca	ccatggagtt	tgggctgagc	1260
tggttttttc	ttgtcgcgat	tttaaaaggt	gtccagtgcg	catggtatgc	cgaagggat	1320
gctgaaattg	agaacgaaaa	gctgcgccgg	gaggttgaag	aactgcggca	ggccagcgag	1380
gcagatctcc	agccaggaa	tattgagtac	gaacgccatc	gacttacgcg	tgcgcaggcc	1440
gacgcacagg	aactgaagaa	tgccagagac	tccgctgaag	tggtggaaac	cgcattctgt	1500
actttcgtgc	tgctcgggat	cgcagggtga	attgccagta	ttctcgacgg	gctccccctg	1560
tcggtgcagc	ggcgttttcc	ggaactggaa	aaccgacatg	ttgatttcc	gaaacgggat	1620
atcatcaaa	ccatgaacaa	agcagccgcg	ctggatgaac	tgataccggg	gttgcgtgag	1680
gaatatatcg	aacagtcagg	ttaacaggct	gcggcatttt	gtccgcgccg	ggcttcgctc	1740
actgttcagg	cggagccac	agaccgccgt	tgaatgggcg	gatgctaatt	actatctccc	1800
gaaagaatcc	gcataccagg	aagggcgctg	ggaaacactg	ccctttcagc	gggccatcat	1860
gaatgcgatg	ggcagcgact	acatccgtga	ggatgaatgtg	gtgaagtctg	cccgtgtcgg	1920
ttattccaaa	atgctgctgg	gtgtttatgc	ctactttata	gagcataagc	agcgcaacac	1980
ccttatctgg	ttgccagcgg	atggtgatgc	cgagaacttt	atgaaaacc	acgttgagcc	2040
gactattcgt	gatattccgt	cgctgctggc	gctggccccg	tggtatggca	aaaagcaccg	2100
ggataaacag	ctcaccatga	agcgtttcac	taatggggcg	ggcttctggg	gcctggggcg	2160
taaagcggca	aaaaactacc	gtgaaaagtc	ggatgatgtg	gcgggttatg	atgaacttgc	2220
tgcttttgat	gatgatattg	aacaggaagg	ctctccgacg	ttctgggtg	acaagcgcgt	2280
cgaccaaggg	cccatccgtc	ttccccctgg	cgccctgctc	caggagcacc	tccgagagca	2340
cagcggccct	gggtgcctg	gtcaaggact	acttccccga	accggtgacg	gtgtcgtgga	2400
actcaggcgc	cctgaccagc	ggcgtgcaca	ccttccccgc	tgctctacag	tctcaggac	2460
tctactccct	cagcagcgtg	gtgaccgtgc	cctccagcag	cttgggcacg	aagacctaca	2520
cctgcaatgt	agatcacaag	cccagcaaca	ccaagggtga	caagagagtt	gagtccaaat	2580
acgggtccgc	atgcccatca	tgcccagcac	ctgaattcct	ggggggacca	tcagtcttcc	2640
tgttcccccc	aaaaccaca	gacacctca	tgatctcccg	gacccctgag	gtcacgtgcg	2700
tggtggtgga	cgtgagccag	gaagaccccg	aggtccagtt	caactggtac	gtggatggcg	2760
tgagggtgca	taatgccaa	acaaagccgc	gggaggagca	gttcaacagc	acgtaccgtg	2820
tggtcagcgt	cctcaccgtc	ctgcaccagg	actggctgaa	cggcaaggag	tacaagtgca	2880
aggtctccaa	caaaggcctc	ccgtctctca	tcgagaaaac	catctccaaa	gccaaagggc	2940
agccccgaga	gccacagggt	tacacctgc	ccccatccca	ggaggagatg	accaagaacc	3000
aggtcagcct	gacctgcctg	gtcaaaaggct	tctaccccag	cgacatcgcc	gtggagtggg	3060
agagcaatgg	gcagccggag	aacaactaca	agaccacgcc	tcccgctcg	gactccgacg	3120
gctccttctt	cctctacagc	aggctaaccg	tggacaagag	caggtggcag	gaggggaatg	3180
tcttctcatg	ctccgtgatg	catgaggctc	tgcacaacca	ctacacgcag	aagagcctct	3240
ccctgtctct	gggtaaatga	gcggccgctc	gaggccggca	aggccggatc	ccccgacctc	3300
gacctctggc	taataaagga	aatttatctt	cattgcaata	gtgtgttgga	attttttgg	3360

-continued

tctctcactc	ggaaggacat	atgggagggc	aatcatttg	gtcgagatcc	ctcggagatc	3420
tctagctaga	ggatcgatcc	cgcgcccgga	cgaactaaac	ctgactacga	catctctgcc	3480
ccttcttcgc	ggggcagtg	atgtaatccc	ttcagttgg	tgggtacaact	tgccaactgg	3540
gccctgttcc	acatgtgaca	cgggggggga	ccaaacacaa	aggggttctc	tgactgtagt	3600
tgacatcctt	ataaatggat	gtgcacattt	gccaaactg	agtggctttc	atcctggagc	3660
agactttgca	gtctgtggac	tgcaacacaa	cattgccttt	atgtgtaact	cttggctgaa	3720
gctcttacac	caatgctggg	ggacatgtac	ctcccagggg	cccaggaaga	ctacgggagg	3780
ctacaccaac	gtcaatcaga	ggggcctgtg	tagctaccga	taagcggacc	ctcaagaggg	3840
cattagcaat	agtggtttata	aggccccctt	gttaacccta	aacgggtagc	atatgcttcc	3900
cgggtagtag	tatatactat	ccagactaac	cctaattcaa	tagcatatgt	tacccaacgg	3960
gaagcatatg	ctatcgaatt	agggttagta	aaagggtcct	aaggaacagc	gatatctccc	4020
accccatgag	ctgtcacggt	tttatattaca	tggggtcagg	attccacgag	ggtagtgaac	4080
catttttagtc	acaagggcag	tggctgaaga	tcaaggagcg	ggcagtgaac	tctcctgaat	4140
cttcgcctgc	ttcttcattc	tccttcgttt	agctaataga	ataactgctg	agttgtgaac	4200
agtaagggtg	atgtgaggtg	ctcgaaaaca	aggtttcagg	tgacgcccc	agaataaaat	4260
ttggacgggg	ggttcagtg	tggcattgtg	ctatgacacc	aatataaccc	tcacaaaccc	4320
cttgggcaat	aaatactagt	gtaggaatga	aacattctga	atatctttaa	caatagaaat	4380
ccatgggggtg	gggacaagcc	gtaaagactg	gatgtccatc	tcacacgaat	ttatggctat	4440
gggcaacaca	taatcctagt	gcaatatgat	actgggggta	ttaagatgtg	tcccaggcag	4500
ggaccaagac	aggtgaacca	tgttgttaca	ctctatttgt	aacaagggga	aagagagtgg	4560
acgcgcagac	cagcggactc	cactggttgt	ctctaacacc	cccgaataat	aaacggggct	4620
ccacgccaat	ggggcccata	aacaaagaca	agtgggccact	cttttttttg	aaattgtgga	4680
gtggggggc	gcgtcagccc	ccacacgcgc	ccctgcgggt	ttggactgta	aaataagggt	4740
gtaataaact	ggctgattgt	aaccccgcta	accactgcgg	tcaaacact	tgcccacaaa	4800
accactaatg	gcaccccg	gaatacctgc	ataagtaggt	gggcgggcca	agataggggc	4860
gcgattgctg	cgatctggag	gacaaattac	acacacttgc	gcctgagcgc	caagcacagg	4920
gttgttggtc	ctcatattca	cgaggtcgct	gagagcacgg	tgggctaatt	tgccatggg	4980
tagcatatac	tacccaaata	tctggatagc	atatgctatc	ctaactctata	tctgggtagc	5040
ataggctatc	ctaactctata	tctgggtagc	atatgctatc	ctaactctata	tctgggtagt	5100
atatgctatc	ctaatttata	tctgggtagc	ataggctatc	ctaactctata	tctgggtagc	5160
atatgctatc	ctaactctata	tctgggtagt	atatgctatc	ctaactctata	tccgggtagc	5220
atatgctatc	ctaataagaga	ttagggtagt	atatgctatc	ctaatttata	tctgggtagc	5280
atatactacc	caaatactctg	gatagcatat	gctatcctaa	tctatatctg	ggtagcatat	5340
gctatcctaa	tctatatctg	ggtagcatag	gctatcctaa	tctatatctg	ggtagcatat	5400
gctatcctaa	tctatatctg	ggtagtatat	gctatcctaa	tttatatctg	ggtagcatag	5460
gctatcctaa	tctatatctg	ggtagcatat	gctatcctaa	tctatatctg	ggtagtatat	5520
gctatcctaa	tctgtatccg	ggtagcatat	gctatcctca	tgataagctg	tcaaacatga	5580
gaattttctt	gaagacgaaa	gggcctcggtg	atacgcttat	ttttataggt	taatgtcatg	5640
ataataatgg	tttcttagac	gtcaggtggc	acttttcggg	gaaatgtgcg	cggaaaccct	5700

-continued

atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga	5760
taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt cctgtgctgcc	5820
cttattccct tttttcgccg attttgcctt cctgtttttg ctcaccacaga aacgctgggtg	5880
aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc	5940
aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact	6000
tttaaagttc tgctatgtgg cgcggtatta tcccgtgttg acgccgggca agagcaactc	6060
ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag	6120
catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat	6180
aacactgcgg ccaacttact tctgacaaag atcggaggac cgaaggagct aaccgctttt	6240
ttgcacaaca tgggggcatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa	6300
gccataccaa acgacgagcg tgacaccacg atgcctgcag caatggcaac aacgttgccg	6360
aaactattaa ctggcgaaact acttactcta gcttcccgcc aacaattaat agactggatg	6420
gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt	6480
gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc actggggcca	6540
gatggtaagc cctcccgat cgtagttatc tacacgacgg ggagtcaggc aactatggat	6600
gaacgaata gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca	6660
gaccaagttt actcatatat acttttagatt gatttaaaac ttcatTTTTA atttaaaag	6720
atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg	6780
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt	6840
ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg	6900
ccgatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	6960
caaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa cctgttagca	7020
ccgcctacat acctcgctct gctaatectg ttaccagtgg ctgctgccag tggcgataag	7080
tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	7140
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	7200
tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaaa ggcggacagg	7260
tatccggtaa gcggcagggg cggaacagga gagcgacga gggagcttcc agggggaaac	7320
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg	7380
tgatgctcgt cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg	7440
ttcctggcct tttgctggcc ttttgcctac atgttctttc ctgcgttate cctgattct	7500
gtggataaac gtattaccgc ctttgagtga gctgataccg ctcgccgag ccgaacgacc	7560
gagcgcagcg agtcagttag cgaggagcg gaagagcgcc caatacgcaa accgcctctc	7620
cccgcgctt ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg	7680
ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta	7740
cactttatgc ttccggctcg tatgttgtg ggaattgtga gcgataaca atttcacaca	7800
ggaaacagct atgaccatga ttacgccaaag ctctagctag aggtcgagtc cctccccagc	7860
aggcagaagt atgcaaaagca tgcattctca ttagtcagca accatagtcc cgccctaac	7920
tccgcccac cgcgccctaa ctccgccag ttccgcccat tctccgcccc atggctgact	7980

-continued

aatttttttt atttatgcag aggccgaggc cgctcggcc tctgagctat tccagaagta	8040
gtgaggaggc ttttttgag gcctaggctt ttgcaaaaag ctttgcaaag atggataaag	8100
ttttaaacag agaggaatct ttgcagctaa tggaccttct aggtcttgaa agg	8153

<210> SEQ ID NO 27

<211> LENGTH: 7244

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pHybC-mBR3-mCg2a

<400> SEQUENCE: 27

ttgacattga ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag	60
cccatatatg gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc	120
caacgacccc cgccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg	180
gactttccat tgacgtcaat ggggtggagta tttacggtaa actgccact tggcagtaca	240
tcaagtgtat catatgccaa gtccgcccc tattgacgtc aatgacggta aatggccgcg	300
ctggcattat gccagtaga tgacctacg ggactttcct acttggcagt acatctacgt	360
attagtcacg gctattacca tgggtgatgcg gttttggcag tacaccaatg ggcgtggata	420
gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt	480
ttggcaccaa aatcaacggg actttccaaa atgtcgtaat aaccccgccc cgttgacgca	540
aatgggcggt aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg	600
tcagatcctc actctcttcc gcatcgctgt ctgcgagggc cagctgttgg gctcgcggtt	660
gaggacaaac tcttcgcggt ctttccagta ctcttgatc ggaaacccgt cggcctccga	720
acggtactcc gccaccgagg gacctgagcg agtccgcac gaccggatcg gaaaacctct	780
cgagaaaggc gtctaaccag tcacagtcgc aaggtaggct gagcacgtg gcgggaggca	840
gcgggtggcg gtcgggggtg tttctggcgg aggtgctgct gatgatgtaa ttaaagtagg	900
cggtcttgag acggcgagtg gtcgaggtga ggtgtggcag gcttgagatc cagctgttgg	960
ggtagtgact ccctctcaaa agcgggcatt actctgcgc taagattgtc agtttccaaa	1020
aacgaggagg atttgatatt cacctggccc gatctggcca tacacttgag tgacaatgac	1080
atccactttg cctttctctc cacaggtgtc cactcccagg tccaagtttg ggcgccacca	1140
tggagtttgg gctgagctgg ctttttcttg tcgcgatttt aaaaggtgtc cagtgcggcg	1200
ccaggagact cgggttccga agccagagga gccgggacag ctcggtgccc acccagtgca	1260
atcagaccga gtgcttcgac cctctggtga gaaactgcgt gtctgtgag ctcttccaca	1320
cgccggacac tggacataca agcagcctgg agcctgggac agctctgcag cctcaggagg	1380
gctccgcgct gagaccgcag gtggcggagc ccagagggcc cacaatcaag cctgtcctc	1440
catgcaaagc cccagcacct aacctcttgg gtggaccatc cgtcttcac ttcctccaa	1500
agatcaagga tgtaactcat atctccctga gcccatagt cacatgtgtg gtgggtggatg	1560
tgagcgagga tgaccagat gtccagatca gctggtttgt gaacaacgtg gaagtacaca	1620
cagctcagac acaaacccat agagaggatt acaacagtac tctccgggtg gtcagtgcc	1680
tccccatcca gcaccaggac tggatgagtg gcaaggagtt caaatgcaag gtcaacaaca	1740
aagacctccc agcgcctac gagagaacca tctcaaaacc caaaggttca gtaagagctc	1800
cacaggtata tgtcttgctt ccaccagaag aagagatgac taagaaacag gtcactctga	1860

-continued

cctgcatggt cacagacttc atgcctgaag acatttacgt ggagtggacc aacaacggga	1920
aaacagagct aaactacaag aacactgaac cagtcctgga ctctgatggt tcttacttca	1980
tgtacagcaa gctgagagt gaaaagaaga actgggtgga aagaaatagc tactcctgtt	2040
cagtgggtcca cgagggtctg cacaatcacc acacgactaa gagcttctcc cggactccgg	2100
gtaaataagc ggccgctcga ggccggcaag gccggatccc ccgacctcga cctctggcta	2160
ataaaggaaa ttatttttca ttgcaatagt gtgttggaat tttttgtgtc tctcactcgg	2220
aaggacatat gggagggcaa atcatttggt cgagatccct cggagatctc tagctagagg	2280
atcgatcccc gccccggacg aactaaacct gactacgaca tctctgcccc ttcttcgagg	2340
ggcagtgcac gtaatccctt cagttggttg gtacaacttg ccaactgggc cctgttcac	2400
atgtgacacg gggggggacc aaacacaaa gggttctctg actgtagtgt acatccttat	2460
aaatggatgt gcacatttgc caacactgag tggctttcat cctggagcag actttgcagt	2520
ctgtggactg caacacaaca ttgcctttat gtgtaactct tggtggaagc tcttacacca	2580
atgtggggg acatgtacct ccagggggc caggaagact acgggaggct acaccaacgt	2640
caatcagagg ggctgtgta gctaccgata agcggacct caagagggca ttagcaatag	2700
tgtttataag gcccccttgt taaccctaaa cgggttagcat atgcttcccg ggtagtagta	2760
tatactatcc agactaaccc taattcaata gcatatgtta cccaacggga agcatatgct	2820
atcgaattag ggtagtaaa agggctcctaa ggaacagcga tatctccac ccatgagct	2880
gtcacggttt tatttacatg gggtcaggat tccacgagg tagtgaacca ttttagtcac	2940
aagggcagt gctgaagatc aaggagcggg cagtgaactc tctgaaatct tcgctgctt	3000
cttcattctc ctctgtttag ctaatagaat aactgctgag ttgtgaacag taagggtgat	3060
gtgaggtgct cgaacaacag gtttcagggt acgccccag aataaaattt ggacgggggg	3120
ttcagtgggt gcattgtgct atgacaccaa tataaccctc acaaacctt tgggaataa	3180
atactagtgt aggaatgaaa cattctgaat atctttaaca atagaaatcc atggggtggg	3240
gacaagccgt aaagactgga tgtccatctc acacgaattt atggctatgg gcaacacata	3300
atcctagtgc aatatgatac tggggttatt aagatgtgtc ccaggcaggg accaagacag	3360
gtgaaccatg ttgttacact ctatttgtaa caaggggaaa gagagtggac gccgacagca	3420
gcggactcca ctggttgtct ctaacacccc cgaaaattaa acggggctcc acgccaatgg	3480
ggcccataaa caaagacaag tggccactct tttttttgaa attgtggagt gggggcacgc	3540
gtcagcccc acacgccgcc ctgcggtttt ggactgtaaa ataagggtgt aataacttgg	3600
ctgattgtaa ccccgctaac cactgcggtc aaaccacttg cccacaaaac cactaatggc	3660
acccggggga atacctgcat aagtaggtgg gcgggccaag ataggggcgc gattgtgcg	3720
atctggagga caaattacac acacttgcgc ctgagcgcca agcacagggt tgttggtcct	3780
catattcacg aggtcgctga gagcacgggt ggctaattgt gccatgggta gcatatacta	3840
cccaaatatc tggatagcat atgctatcct aatctatatac tgggtagcat aggtatcct	3900
aatctatatac tgggtagcat atgctatcct aatctatatac tgggtagcat atgctatcct	3960
aatttatatac tgggtagcat aggtatcct aatctatatac tgggtagcat atgctatcct	4020
aatctatatac tgggtagcat atgctatcct aatctatatac cgggtagcat atgctatcct	4080
aatagagatt agggtagtat atgctatcct aatttatatac tgggtagcat atactacca	4140

-continued

aatatctgga tagcatatgc tatcctaato tatatctggg tagcatatgc tatcctaato	4200
tatatctggg tagcataggc tatcctaato tatatctggg tagcatatgc tatcctaato	4260
tatatctggg tagtatatgc tatcctaatt tatatctggg tagcataggc tatcctaato	4320
tatatctggg tagcatatgc tatcctaato tatatctggg tagtatatgc tatcctaato	4380
tgtatccggg tagcatatgc tatcctcatg ataagctgtc aaacatgaga attttcttga	4440
agacgaaagg gcctcgtgat acgcctatct ttataggtta atgtcatgat aataatggtt	4500
tcttagacgt cagggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttatct	4560
ttctaaatac attcaaatat gtatccgctc atgagacaat aacctgata aatgcttcaa	4620
taatatgaa aaaggaagag tatgagtatt caacatttcc gtgtcgcctt tattcccttt	4680
tttgcgcat tttgccttcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat	4740
gctgaagatc agttgggtgc acgagtgggt tacatcgaa tggtatctaa cagcggtaag	4800
atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagtcttg	4860
ctatgtggcg cgggtattatc ccgtgttgac gccgggcaag agcaactcgg tcgccgcata	4920
cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat	4980
ggcatgacag taagagaatt atgcagtgtc gccataacca tgagtataa cactgcggcc	5040
aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg	5100
ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac	5160
gacgagcgtg acaccacgat gcctgcagca atggcaacaa cgttgcgcaa actattaact	5220
ggcgaaactac ttactctagc tttccggcaa caattaatag actggatgga ggccgataaa	5280
gttgacggac cacttctcgc ctccggcctt ccggctggct ggtttattgc tgataaatct	5340
ggagccgggt agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaaagcc	5400
ttccgtatcg tagttatcta caccgacggg agtcaggcaa ctatggatga acgaaataga	5460
cagatcgtg agataggtgc ctactgatt aagcattggg aactgtcaga ccaagtttac	5520
tcatatatac tttagattga tttaaaactt catttttaac ttaaaaggat ctagggtgaag	5580
atcctttttg ataactctat gacaaaaac ccttaacgtg agttttcgtt ccactgagcg	5640
tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaac	5700
tgctgcttgc aaacaaaaa accacgcgta ccagcgggtg tttgtttgcc ggatcaagag	5760
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt	5820
cttctagtgt agccgtagt aggccaccac ttcaagaact ctgtagcacc gcctacatac	5880
ctcgtctcgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc	5940
gggttggaact caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt	6000
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt	6060
gagctatgag aaagcgccac gcttcccgaa gggagaaaag cggacaggta tccggtaagc	6120
ggcaggggtc gaacaggaga gcgcacgag gagcttcag ggggaaacgc ctgggtatctt	6180
tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca	6240
ggggggcgga gcctatggaa aaacgccagc aacgcggcct tttacgggtt cctggccttt	6300
tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaacgt	6360
attaccgcct ttgagtgagc tgataccgct cgcgcagacc gaacgaccga gcgcagcgag	6420
tcagtgagcg aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg	6480

-continued

```

ccgattcatt aatgcagctg gcacgacagg ttccccgact ggaaagcggg cagtgcgcgc 6540
aacgcaatta atgtgagtta gctcactcat taggcacccc aggcctttaca ctttatgctt 6600
ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat 6660
gaccatgatt acgccaagct ctagctagag gtgcagtcct tccccagcag gcagaagtat 6720
gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 6780
gcccctaact ccgcccagtt ccgcccattc tccgcccctt ggctgactaa ttttttttat 6840
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 6900
ttttggaggc ctaggctttt gcaaaaagct ttgcaaagat ggataaagtt ttaaacagag 6960
aggaatcttt gcagctaatt gaccttctag gtcttgaaag gagctcgacc aattctcatg 7020
tttgacagct tatcatcgca gatccgggca acgttggtgc cattgctgca ggcgcagAAC 7080
tggtaggtat ggaagatcta tacattgaat caatattggc aattagccat attagtcatt 7140
ggttatatag cataaatcaa tattggctat tggccattgc atacgttgta tctatatcat 7200
aatatgtaca tttatattgg ctcatgtcca atatgaccgc catg 7244

```

<210> SEQ ID NO 28

<211> LENGTH: 7104

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pHybE-mBR3-mCg2a

<400> SEQUENCE: 28

```

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga 60
gaagttgggg ggagggtgct gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa 120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gaggggtggg gagaaccgta 180
tataagtga gtagtcgccc tgaacgttct ttttcgcaac gggtttgccg ccagaacaca 240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctt 300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg 360
aagtgggtgg gagagttcga ggccttgccc ttaaggagcc ccttcgcctc gtgcttgagt 420
tgaggccttg cctgggcgct ggggcgcgag cgtgcgaatc tggtggaacc ttcgcgcctg 480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct 540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt 600
ttttggggcc gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg 660
gggcctgcga gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc 720
tctgggtgct ggctcgcgc gcgcgtgtat cgcgccgccc tgggcggcaa ggctggcccc 780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc 840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag tcaccacac aaaggaaaag 900
ggcctttccg tcctcagccg tcgcttcatg tgactccacg gactaccggg cgcgctccag 960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt 1020
ttatgcgatg gagtttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca 1080
cttgatgtaa ttctccttgg aatttgccct ttttgagttt ggatcttggg tcattctcaa 1140
gcctcagaca gtgggtcaaa gtttttttct tccatttcag gtgtcgtgag gaattctcta 1200

```

-continued

gagatccctc gacctcgaga tccattgtgc ccgggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaaggt gtccagtgcg gcgccaggag actccgggtc	1320
cgaagccaga ggagccggga cagctcgggtg cccacccagt gcaatcagac cgagtgcctc	1380
gaccctctgg tgagaaactg cgtgtcctgt gagctcttcc acacgccgga cactggacat	1440
acaagcagcc tggagcctgg gacagctctg cagcctcagg agggctccgc gctgagaccc	1500
gacgtggcgg agcccagagg gccacaatc aagcctgtc ctccatgcaa atgccagca	1560
cctaacctct tgggtggacc atccgtcttc atcttccctc caaagatcaa ggatgtactc	1620
atgatctccc tgagcccat agtcacatgt gtggtggtgg atgtgagcga ggatgaccca	1680
gatgtccaga tcagctgggt tgtgaacaac gtggaagtac acacagctca gacacaaacc	1740
catagagagg attacaacag tactctccgg gtggtcagtg cctccccc atccagcaccag	1800
gactggatga gtggcaagga gttcaaatgc aaggtcaaca acaagacct cccagcgccc	1860
atcgagagaa ccatctcaaa acccaaaggg tcagtaagag ctccacaggt atatgtcttg	1920
cctccaccag aagaagagat gactaagaaa caggtcactc tgacctgcat ggtcacagac	1980
ttcatgacct aagacattta cgtggagtgg accaacaacg ggaaaacaga gctaaactac	2040
aagaacactg aaccagtcct ggactctgat gggtcttact tcatgtacag caagctgaga	2100
gtggaaaaga agaactgggt ggaaagaaat agctactcct gttcagtggt ccacgagggg	2160
ctgcacaatc accacacgac taagagcttc tcccggactc cgggtaata agcggccgct	2220
cgaggccggc aaggccggat ccccccagct cgacctctgg ctaataaagg aaatttattt	2280
tcattgcaat agtgtgttg aattttttgt gtctctcact cggaaggaca tatgggaggg	2340
caaatcattt ggtcgagatc cctcggagat ctctagctag aggatcgatc cccgccccgg	2400
acgaactaaa cctgactacg acatctctgc ccttctctcg cggggcagtg catgtaatcc	2460
cttcagttgg ttggtacaac ttgccaaactg ggccctgttc cacatgtgac acgggggggg	2520
acaaaacaca aaggggttct ctgactgtag ttgacatcct tataaatgga tgtgcacatt	2580
tgccaacact gagtggcttt catcctggag cagactttgc agtctgtgga ctgcaacaca	2640
acattgcctt tatgtgtaac tcttggtga agctcttaca ccaatgctgg gggacatgta	2700
cctcccaggg gcccggaag actacgggag gctacaccaa cgtcaatcag aggggcctgt	2760
gtagctaccg ataagcggac cctcaagagg gcattagcaa tagtgtttat aaggccccct	2820
tgtaaacctt aaacgggtag catatgcttc ccgggtagta gtatatacta tccagactaa	2880
ccctaattca atagcatatg ttacccaacg ggaagcatat gctatcgaat tagggttagt	2940
aaaagggctc taaggaacag cgatatctcc caccocatga gctgtcacgg ttttatttac	3000
atggggctcag gattccacga gggtagtgaa ccattttagt cacaagggca gtggctgaag	3060
atcaaggagc gggcagtgaa ctctcctgaa tcttcgctg cttcttcatt ctcttcggt	3120
tagctaatag aataactgct gagttgtgaa cagtaagggt tatgtgaggt gctcgaaaac	3180
aagggttcag gtgaccccc cagaataaaa tttggacggg gggttcagtg gtggcattgt	3240
gctatgacac caatataacc ctacaaaacc ccttgggcaa taaatactag ttaggaatg	3300
aaacattctg aatatcttta acaatagaaa tccatggggg ggggacaagc cgtaagact	3360
ggatgtccat ctacacgaa tttatggcta tgggcaacac ataactctag tgcaatatga	3420
tactggggtt attaatgtgt gtcccaggca gggaccaaga caggtgaacc atgttggtac	3480
actctatttg taacaagggg aaagagagtg gacgccgaca gcagcggact cactgggtg	3540

-continued

tctctaacac ccccgaaaat taaacggggc tccacgcaa tggggcccat aaacaaagac	3600
aagtggccac tctttttttt gaaattgtgg agtgggggca cgcgtcagcc cccacacgcc	3660
gccttgcggt tttggactgt aaaataaggg tgtaataact tggtgattg taaccccgct	3720
aaccactgcg gtcaaacac ttgccacaa aaccactaat ggcaccccg ggaatacctg	3780
cataagtagg tgggcggggc aagatagggg cgcgattgct gcgatctgga ggacaaatta	3840
cacacacttg cgcttgagcg ccaagcacag ggttgttggt cctcatattc acgaggtcgc	3900
tgagagcagc gtgggctaata gttgccatgg gtagcatata ctacccaaat atctggatag	3960
catatgctat cctaacttat atctgggtag cataggctat cctaacttat atctgggtag	4020
catatgctat cctaacttat atctgggtag tatatgctat cctaatttat atctgggtag	4080
cataggctat cctaacttat atctgggtag catatgctat cctaacttat atctgggtag	4140
tatatgctat cctaacttgt atccgggtag catatgctat cctaataagag attagggtag	4200
tatatgctat cctaatttat atctgggtag catatactac ccaaatatct ggatagcata	4260
tgctatccta atctatatct gggtagcata tgctatccta atctatatct gggtagcata	4320
ggctatccta atctatatct gggtagcata tgctatccta atctatatct gggtagtata	4380
tgctatccta atttatatct gggtagcata ggctatccta atctatatct gggtagcata	4440
tgctatccta atctatatct gggtagtata tgctatccta atctgtatcc gggtagcata	4500
tgctatccta atgataagct gtcaaacatg agaattttct tgaagacgaa agggcctcgt	4560
gatacgcta tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcagggtg	4620
cacttttcgg ggaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa	4680
tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa	4740
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttcggg cattttgcct	4800
tcctgttttt gctcaccag aaacgtggg gaaagtaaaa gatgctgaag atcagttggg	4860
tgacagagtg gggtacatcg aactggatct caacagcggg aagatccttg agagttttcg	4920
ccccgaagaa cgttttccaa tgatgagcac ttttaagtt ctgctatgtg gcgcgggtatt	4980
atcccggtgt gacgcggggc aagagcaact cggtcgcgcg atacactatt ctacagaatga	5040
cttggttgag tactcaccag tcacagaaaa gcacattacg gatggcatga cagtaagaga	5100
attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac	5160
gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg	5220
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac	5280
gatgcctgca gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct	5340
agcttcccg caacaattaa tagactggat ggaggcgat aaagtgcag gaccacttct	5400
gcgctcggcc ctcccgctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg	5460
gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagtatt	5520
ctacacgagc gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg	5580
tgctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat	5640
tgatttaaaa cttcatTTTT aatttaaaa gatctagggtg aagatccttt ttgataatct	5700
catgacaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa	5760
gatcaaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaaaaa	5820

-continued

aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc	5880
gaaggtaaact ggcttcagca gagcgcagat accaaatact gttcttctag tgtagccgta	5940
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgtc tgtaatcct	6000
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg	6060
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag	6120
cttgagcgca acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc	6180
cacgcttccc gaaggagaaa aggcggacag gtatccggtg agcggcaggg tcggaacagg	6240
agagcgacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt	6300
tcgcacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg	6360
gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca	6420
catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg	6480
agctgatacc gctcgcgcga gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc	6540
ggaagagcgc ccaatacgca aaccgcctct cccgcgcgtg tgcccgattc attaatgcag	6600
ctggcacgac aggtttcccg actggaaaagc gggcagtgag cgcaacgcaa ttaatgtgag	6660
ttagctcact cattaggcac ccaggtttt aacttttatg cttccggtc gtatgttggtg	6720
tggaattgtg agcggataac aatttcacac aggaacacgc tatgacctg attacgcaa	6780
gctctagcta gaggtcgagt cctccccag caggcagaag tatgcaaagc atgcatctca	6840
attagtcagc aaccatagtc ccgccctaa ctccgccat cccgcccta actccgccca	6900
gttcgcccc ttctccgcc catggctgac taattttttt tatttatgca gaggcgagg	6960
ccgctcggc ctctgagcta ttccagaagt agtgaggagg cttttttgga ggcttaggt	7020
tttgcaaaaa gctttgcaaa gatggataaa gttttaaaca gagaggaatc tttgcagcta	7080
atggaccttc taggtcttga aagg	7104

<210> SEQ ID NO 29

<211> LENGTH: 6985

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pHybC-E7-hCk

<400> SEQUENCE: 29

ttgacattga ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag	60
cccatatatg gagttccggt ttacataact tacggtaaat ggcccgcctg gctgaccgcc	120
caacgacccc cgccattgca cgtcaataat gacgtatgtt cccatagtaa cgccaatagg	180
gactttccat tgacgtcaat ggggtggagta tttacggtaa actgcccact tggcagtaca	240
tcaagtgtat catatgcccc gtccgcccc tattgacgtc aatgacggta aatggccccg	300
ctggcattat gccagctaca tgaccttacg ggactttcct acttggcagt acatctacgt	360
attagtcacg gctattacca tgggtgatgcg gttttggcag tacaccaatg ggcggtggata	420
gcgggttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttggt	480
ttggcaccac aatcaacggg actttccaaa atgtcgtaat aaccccgccc cgttgacgca	540
aatggggcgt aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg	600
tcagatcctc actctcttcc gcacgcgtgt ctgcgagggc cagctgttgg gctcgcgggt	660
gaggacaaac tcttcgcggt ctttccagta ctcttgatc ggaaccgcgt cgccctccga	720

-continued

acggtactcc gccaccgagg gacctgagcg agtccgcacg gaccggatcg gaaaacctct	780
cgagaaaggc gtctaaccag tcacagtcgc aaggtaggct gagcaccgtg gcgggaggca	840
gcgggtggcg gtcgggggtg tttctggcgg aggtgctgct gatgatgtaa ttaaagtagg	900
cggctcttag acggcggatg gtcgaggtag ggtgtggcag gcttgagatc cagctgttgg	960
ggtgagtact ccctctcaaa agcgggcatt acttctgcgc taagattgtc agtttccaaa	1020
aacgaggagg atttgatatt cacctggccc gatctggcca tacacttgag tgacaatgac	1080
atccactttg cctttctctc cacagggtgc cactcccagg tccaagtttg ggcgcaccat	1140
ggacatgcgc gtgcccgcgc agctgctggg cctgctgctg ctgtggttcc ccggctcgcg	1200
atcgacatc cagatgaccc agtctccatc ctccctgtct gcacttatag gggacagagt	1260
caccatcact tgtcgggcaa gtcagggcat cagaaattac ttagcctggt atcagcaaaa	1320
accagggaaa gcccctaagc tcctgatcta tgcctgatcc actttgcaat caggggtccc	1380
atctcggttc agtggcagtg gatctgggac agatttcact ctcaccatca gcagcctaca	1440
gcctgaagat gttgcaactt attactgtca aaggtataac cgtgccccgt acacttttgg	1500
ccaggggacc aaggtggaaa tcaaacgtac ggtggctgca ccactgtctc tcatcttccc	1560
gccatctgat gagcagttga aatctggaac tgctctgtt gtgtgcctgc tgaataactt	1620
ctatcccaga gaggccaaag tacagtggaa ggtggataac gccctccaat cgggtaactc	1680
ccaggagagt gtcacagagc aggacagcaa ggacagcacc tacagcctca gcagcacct	1740
gacgctgagc aaagcagact acgagaaaac caaagtctac gcctgcgaag tcaacctca	1800
gggcctgagc tcgccctgca caaagagctt caacagggga gagtgttgag cggccgctcg	1860
aggccggcaa ggccggatcc cccgacctcg acctctggct aataaaggaa atttattttc	1920
attgcaatag tgtgttgga ttttttgtgt ctctcactcg gaaggacata tgggagggca	1980
aatcatttgg tcgagatccc tcggagatct ctacttagag gatcgatccc cggcccgagc	2040
gaactaaacc tgactacgac atctctgccc cttcttcgcg gggcagtgca tgtaatccct	2100
tcagttgggt ggtacaactt gccaaactggg ccctgttcca catgtgacac ggggggggac	2160
caaacacaaa ggggttctct gactgtagtt gacatcctta taaatggatg tgcacatttg	2220
ccaacactga gtggctttca tcctggagca gactttgcag tctgtggact gcaacacaac	2280
attgccttta tgtgtaactc ttggctgaag ctcttacacc aatgctgggg gacatgtacc	2340
tcccaggggc ccaggaagac tacgggaggc tacaccaacg tcaatcagag gggcctgtgt	2400
agctaccgat aagcggagcc tcaagagggc attagcaata gtgtttataa ggcccccttg	2460
ttaaccctaa acgggtagca tatgttcccc gggtagtagt atatactatc cagactaacc	2520
ctaattcaat agcatatgtt acccaacggg aagcatatgc tatcgaatta gggtagtaa	2580
aagggtccta aggaacagcg atatctccca cccatgagc tgtcacgggt ttatttacat	2640
ggggtcagga ttccacgagg gtagtgaacc attttagtca caagggcagt ggctgaagat	2700
caaggagcgg gcagtgaact ctccatgaac ttgcctgctc tcttcattct ccttcgttta	2760
gctaatagaa taactgctga gttgtgaaca gtaagggtga tgtgaggtgc tcgaaaacaa	2820
ggtttcaggt gacgccccca gaataaaatt tggacggggg gttcagtggt ggcattgtgc	2880
tatgacacca atataaccct cacaaccccc ttgggcaata aatactagtg taggaatgaa	2940
acattctgaa tatctttaac aatagaaatc catgggggtg ggacaagccg taaagactgg	3000

-continued

atgtccatct	cacacgaatt	tatggctatg	ggcaacacat	aatcctagtg	caatatgata	3060
ctggggttat	taagatgtgt	cccaggcagg	gaccaagaca	ggtgaaccat	gttggtacac	3120
tctatttgta	acaaggggaa	agagagtggg	cgccgacagc	agcggactcc	actggtgtgc	3180
tctaacaccc	cgaaaaatta	aacggggctc	cacgccaatg	gggcccataa	acaagacaa	3240
gtggccactc	ttttttttga	aattgtggag	tgggggcacg	cgtcagcccc	cacacgccgc	3300
cctcggtttt	tggactgtaa	aataaggggtg	taataacttg	gctgattgta	accccgctaa	3360
ccactgcggt	caaaccactt	gccacaaaa	ccactaatgg	cacccggggg	aatacctgca	3420
taagtaggtg	ggcgggccaa	gataggggag	cgattgctgc	gatctggagg	acaaattaca	3480
cacacttgcc	cctgagcgcc	aagcacaggg	ttgttggtcc	tcattatcac	gaggtcgctg	3540
agagcacggt	gggctaagt	tgccatgggt	agcatatact	acccaaatat	ctggatagca	3600
tatgctatcc	taatctatat	ctgggtagca	taggctatcc	taatctatat	ctgggtagca	3660
tatgctatcc	taatctatat	ctgggtagta	tatgctatcc	taatttatat	ctgggtagca	3720
taggctatcc	taatctatat	ctgggtagca	tatgctatcc	taatctatat	ctgggtagta	3780
tatgctatcc	taatctgtat	cgggtagca	tatgctatcc	taatagagat	tagggtagta	3840
tatgctatcc	taatttatat	ctgggtagca	tatactaccc	aatatctgg	atagcatatg	3900
ctatccta	ctatatctgg	gtagcatatg	ctatccta	ctatatctgg	gtagcatagg	3960
ctatccta	ctatatctgg	gtagcatatg	ctatccta	ctatatctgg	gtagtatatg	4020
ctatccta	ttatatctgg	gtagcatagg	ctatccta	ctatatctgg	gtagcatatg	4080
ctatccta	ctatatctgg	gtagtatatg	ctatccta	ctgtatccgg	gtagcatatg	4140
ctatcctcat	gataagctgt	caaacatgag	aattttcttg	aagacgaaag	ggcctcgtga	4200
tacgcctatt	tttatagggt	aatgtcatga	taataatggg	ttcttagacg	tcagggtggca	4260
cttttcgggg	aaatgtgcgc	ggaaccccta	ttgttttatt	tttctaaata	cattcaaata	4320
tgtatccgct	catgagacaa	taaccctgat	aatgcttca	ataatattga	aaaaggaaga	4380
gtatgagtat	tcaacatttc	cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	4440
ctgtttttgc	tcaccagaa	acgctgggtg	aagtaaaaga	tgctgaagat	cagttgggtg	4500
cacgagtggt	ttacatcgaa	ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	4560
ccgaagaacg	ttttccaatg	atgagcactt	ttaaagttct	gctatgtggc	gcggtattat	4620
cccggtgtga	cgccgggcaa	gagcaactcg	gtcgccgcat	acactattct	cagaatgact	4680
tggttgagta	ctcaccagtc	acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	4740
tatgcagtgc	tgccataacc	atgagtgtga	acactgcggc	caacttactt	ctgacaacga	4800
tcggaggacc	gaaggagcta	accgcttttt	tgcaacaacat	gggggatcat	gtaactcgcc	4860
ttgatcggtg	ggaaccggag	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	4920
tgctgcagc	aatggcaaca	acgttgcgca	aactattaac	tggcgaacta	cttactctag	4980
cttcccgcca	acaattaata	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	5040
gctcgccctt	tccggtgggc	tggtttattg	ctgataaatc	tggagccggt	gagcgtgggt	5100
ctcgcggtat	cattgcagca	ctggggccag	atggtaagcc	ctcccgatc	gtagttatct	5160
acacgacggg	gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagatagggt	5220
cctcactgat	taagcattgg	taactgtcag	accaagttta	ctcatatata	ctttagattg	5280
atttaaaact	tcatttttaa	tttaaaagga	tctaggtgaa	gacctttttt	gataatctca	5340

-continued

```

tgacaaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga 5400
tcaaaggatc ttcttgagat ccttttttct tgcgcgtaat ctgctgcttg caaacaaaaa 5460
aaccacccgt acccagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga 5520
aggtaactgg cttcagcaga gcgcagatag caaatactgt tcttctagt tagccgtagt 5580
taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt 5640
taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat 5700
agttaccgga taaggcgag cggtcgggt gaacgggggg ttcgtgcaca cagcccagct 5760
tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca 5820
cgcttcccga agggagaaa ggcgacaggt atccggtaag cggcagggtc ggaacaggag 5880
agcgcagcag ggagcttcca ggggaaacg cctggatatc ttatagtcct gtcgggtttc 5940
gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcg agcctatgga 6000
aaaacgccag caacgcggcc tttttacggt tccctggcct ttgctggcct tttgctcaca 6060
tgttcttttc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag 6120
ctgataccgc tcgccgcagc cgaacgacgg agcgcagcga gtcagttagc gaggaagcgg 6180
aagagcgcgc aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatgcagct 6240
ggcacgacag gtttcccagc tggaaagcgg gcagttagcg caacgcaatt aatgtgagtt 6300
agctcactca ttaggcaccc caggctttac actttatgct tccggctcgt atgttggtgtg 6360
gaattgtgag cggataacaa tttcacacag gaaacagcta tgaccatgat tacgccaagc 6420
tctagctaga ggtcgagtcc ctcccagca ggcagaagta tgcaaagcat gcatctcaat 6480
tagtcagcaa ccatagtcct gcccttaact ccgcccatcc cggccctaac tccgccagct 6540
tccgccatt ctccgcccc tggtgacta atttttttta tttatgcaga ggccgaggcc 6600
gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt 6660
tgcaaaaagc tttgcaaga tggataaagt tttaaacaga gaggaatctt tgcagctaatt 6720
ggaccttcta ggtcttgaaa ggagctcgac caattctcat gtttgacagc ttatcatcgc 6780
agatccgggc aacgttgttg ccattgctgc aggcgcagaa ctggtaggta tggaagatct 6840
atacattgaa tcaatattgg caattagcca tattagtcac tggttatata gcataaatca 6900
atattggcta ttggccattg catacgttgt atctatatca taatatgtac atttatattg 6960
gctcatgtcc aatatgaccg catg 6985

```

<210> SEQ ID NO 30

<211> LENGTH: 7688

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pHybC-D2-hCg1,z,a

<400> SEQUENCE: 30

```

ttgacattga ttattgacta gttattaata gtaatcaatt acgggggtcat tagttcatag 60
cccatatatg gagttccgcg ttacataact tacggtaaat gggccgcctg gctgaccgcc 120
caacgacccc cggccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg 180
gactttccat tgacgtcaat ggggtggagta tttacggtaa actgccactc tggcagtaca 240
tcaagtgtat catatgccaa gtccgcccc tattgacgtc aatgacggta aatggccgc 300

```

-continued

ctggcattat gccagttaca tgaccttacg ggacttttcc acttggcagt acatctacgt	360
attagtcacg gctattacca tgggtgatgc gttttggcag tacaccaatg ggcgtggata	420
gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt	480
ttggcaccaa aatcaacggg actttccaaa atgtcgtaat aaccccgccc cgttgacgca	540
aatggggcgt aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg	600
tcagatcctc actctcttcc gcacgcgtgt ctgcgagggc cagctgttgg gctcgcggtt	660
gaggacaaac tcttcgcggt ctttccagta ctcttgatc ggaaaccggt cggcctccga	720
acggtactcc gccaccgagg gacctgagcg agtcgcgacg gaccggatcg gaaaacctct	780
cgagaaaggc gtctaaccag tcacagtcgc aaggtaggtc gagcacgtg gcgggcggca	840
gcgggtggcg gtcgggggtt tttctggcgg aggtgctgct gatgatgtaa ttaaagtagg	900
cggctcttgc acggcggtg gtcgaggtga ggtgtggcag gcttgagatc cagctgttgg	960
ggtgagtact ccctctcaaa agcgggcatt actcttcgcg taagattgtc agtttccaaa	1020
aacgaggagg atttgatatt cacctggccc gatctggcca tacacttgag tgacaatgac	1080
atccactttg cctttctctc cacaggtgtc cactcccagg tccaagtttg ggcgccacca	1140
tggagtttgg gctgagctgg ctttttcttg tcgcgatttt aaaaggtgtc cagtgtgagg	1200
tgacagctgt ggagctcggg ggaggcttgg tacagcccgg caggctccctg agactctcct	1260
gtcggcctc tggattcacc tttgatgatt atgccatgca ctgggtccgg caagctccag	1320
ggaaggccct ggaatgggtc tcagctatca cttggaatag tggtcacata gactatgcgg	1380
actctgtgga ggcccgatcc accatctcca gagacaacgc caagaactcc ctgtatctgc	1440
aaatgaacag tctgagagct gaggatacgg ccgtatatta ctgtgcgaaa gtctcgtacc	1500
ttagcacccg gtctctccct gactatttgg gccaaagtac cctggtcacc gtctcgagt	1560
cgtcgaccaa ggcccacg gtctctcccc tggcacccct cccaagagc acctctggg	1620
gcacagcggc cctgggctgc ctggccaagg actacttccc cgaaccgggt acgggtgtcgt	1680
ggaactcagg cgcctcgacc agcggcgtgc acaccttccc ggctgtccta cagtcctcag	1740
gactctactc cctcagcagc gtggtgacgg tgccctccag cagcttgggc acccagacct	1800
acatctgcaa cgtgaatcac aagcccagca acaccaagggt ggacaagaaa gttgagccca	1860
aatcttctga caaaactcac acatgccac cgtgccacgc acctgaactc ctggggggac	1920
cgtcagtcct cctcttcccc ccaaaaccca aggacacct catgatctcc cggacccctg	1980
aggtcacatg cgtgggtgtg gacgtgagcc acgaagaccc tgaggtcaag ttcaactggt	2040
acgtggacgg cgtggagggt cataatgcc agacaaagcc gcgggaggag cagtacaaca	2100
gcacgtaccg tgtggtcagc gtctctcacc tcctgcacca ggactggctg aatggcaagg	2160
agtacaagtg caaggctctc aacaaagccc tcccagcccc catcgagaaa accatctcca	2220
aagccaaagg gcagccccga gaaccacagg tgtacaccct gccccatcc cgggatgagc	2280
tgaccaagaa ccaggtcagc ctgacctgcc tggcacaagg cttctatccc agcgacatcg	2340
ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccacg cctcccgtgc	2400
tggactccga cggctccttc ttcctctaca gcaagctcac cgtggacaag agcaggtggc	2460
agcaggggaa cgtcttctca tgctccgtga tgcagagggc tctgcacaac cactacacgc	2520
agaagagcct ctccctgtct ccgggtaaat gagcggccgc tcgaggccgg caaggccgga	2580
tcccccgacc tcgacctctg gctaataaag gaaatttatt ttcattgcaa tagtgtgttg	2640

-continued

gaatTTTTtg	tgtctctcac	tcggaaggac	atatgggagg	gcaaatcatt	tggtcgagat	2700
ccctcggaga	tctctagcta	gaggatcgat	ccccgccccg	gacgaactaa	acctgactac	2760
gacatctctg	ccccctcttc	gcggggcagt	gcatgtaatc	ccttcagttg	gttggtacaa	2820
cttgccaact	gggccttgtt	ccacatgtga	cacggggggg	gaccaaacac	aaaggggttc	2880
tctgactgta	gttgacatcc	ttataaatgg	atgtgcacat	tgccaacac	tgagtggcct	2940
tcacctcgga	gcagactttg	cagtctgtgg	actgcaacac	aacattgcct	ttatgtgtaa	3000
ctcttggtg	aagctcttac	accaatgctg	ggggacatgt	acctcccagg	ggcccaggaa	3060
gactacggga	ggctacacca	acgtcaatca	gaggggcctg	tgtagctacc	gataagcgga	3120
ccctcaagag	ggcattagca	atagtgttta	taaggccccc	ttgttaaccc	taaacgggta	3180
gcatatgctt	cccggttagt	agtatatact	atccagacta	accctaattc	aatagcatat	3240
gttacccaac	gggaagcata	tgctatcgaa	ttagggttag	taaaagggtc	ctaaggaaaca	3300
gcgatatctc	ccaccccatg	agctgtcacg	gttttattta	catggggtea	ggattccacg	3360
agggtagtga	accatttttag	tcacaagggc	agtggctgaa	gatcaaggag	cgggcagtga	3420
actctctga	atcttcgct	gcttcttcat	tctccttcgt	ttagctaata	gaataactgc	3480
tgagttgtga	acagtaaggt	gtatgtgagg	tgctcgaaaa	caaggtttca	ggtgacgccc	3540
ccagaataaa	atttgagcgg	ggggttcagt	ggtggcattg	tgctatgaca	ccaatataac	3600
cctcacaaac	cccttgggca	ataaatacta	gtgtaggaat	gaaacattct	gaatatcttt	3660
aacaatagaa	atccatgggg	tggggacaag	ccgtaaagac	tggtatgtcca	tctcacacga	3720
atttatggct	atgggcaaca	cataatccta	gtgcaatatg	atactggggg	tattaagatg	3780
tgtoccaggc	agggaccaag	acaggtgaac	catgttgta	cactctatct	gtaacaaggg	3840
gaaagagagt	ggacgcccac	agcagcggac	tccactggtt	gtctctaaca	ccccgaaaa	3900
ttaaacgggg	ctccacgcca	atggggccca	taaacaaaga	caagtggcca	ctcttttttt	3960
tgaaattgtg	gagtgggggc	acgcgtcagc	ccccacacgc	cgccctgcgg	ttttggactg	4020
taaaataagg	gtgtaataac	ttggctgatt	gtaaccccg	taaccactgc	ggtcaaacca	4080
cttgcccaca	aaaccactaa	tggcaccccc	gggaatacct	gcataagtag	gtgggcgggc	4140
caagataggg	gcgcgattgc	tgcgatctgg	aggacaaatt	acacacactt	gcgcctgagc	4200
gccaaagcaca	gggttggttg	tcctcatatt	cacgaggtcg	ctgagagcac	ggtgggctaa	4260
tggtgccatg	ggtagcatat	actacccaaa	tatctggata	gcatatgcta	tcctaactcta	4320
tatctgggta	gcataggcta	tcctaactcta	tatctgggta	gcatatgcta	tcctaactcta	4380
tatctgggta	gtatatgcta	tcctaattta	tatctgggta	gcataggcta	tcctaactcta	4440
tatctgggta	gcatatgcta	tcctaactcta	tatctgggta	gtatatgcta	tcctaactctg	4500
tatccgggta	gcatatgcta	tcctaataga	gattagggta	gtatatgcta	tcctaattta	4560
tatctgggta	gcataacta	cccaaatatc	tggatagcat	atgctatcct	aatctatatc	4620
tgggtagcat	atgctatcct	aatctatatc	tgggtagcat	aggctatcct	aatctatatc	4680
tgggtagcat	atgctatcct	aatctatatc	tgggtagtat	atgctatcct	aatctatatc	4740
tgggtagcat	aggctatcct	aatctatatc	tgggtagcat	atgctatcct	aatctatatc	4800
tgggtagtat	atgctatcct	aatctgtatc	cgggtagcat	atgctatcct	catgataagc	4860
tgtaaaacat	gagaattttc	ttgaagacga	aagggcctcg	tgatacgctt	atttttatag	4920

-continued

gttaatgtca tgataataat ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg	4980
cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga	5040
caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat	5100
ttccgtgtcg cccttattcc cttttttgcg gcattttgccc ttccgtttt tgctcaccca	5160
gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc	5220
gaactggatc tcaacagcgg taagatcctt gagagttttc gccccgaaga acgtttttcca	5280
atgatgagca cttttaaaagt tctgctatgt ggcgcggtat tatcccggtg tgacgcgggg	5340
caagagcaac tcggtcgcgc catacactat tctcagaatg acttggttga gtactcacca	5400
gtcacagaaa agcatccttac ggatggcatg acagtaagag aattatgcag tgctgccata	5460
accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag	5520
ctaaccgctt ttttgacaaa catgggggat catgtaactc gccttgatcg ttgggaaccg	5580
gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgc agcaatggca	5640
acaacgttgc gcaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta	5700
atagactgga tggagcgga taaagttgca ggaccacttc tgcgctcggc ccttcgggt	5760
ggctggttta ttgctgataa atctggagcc ggtgagcgtg ggtctcgcg tatcattgca	5820
gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag	5880
gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat	5940
tggttaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt	6000
taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa	6060
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga	6120
gatccttttt ttctgcgcgt aatctgctgc ttgcaacaa aaaaaccacc gctaccagcg	6180
gtggtttgtt tgccggatca agagctacca actcttttcc cgaaggtaac tggcttcagc	6240
agagcgcaga taccaaaatac tgttcttcta gtgtagccgt agttaggcca ccacttcaag	6300
aactctgtag caaccgctac atacctcgt ctgctaatac tgttaccagt ggctgctgcc	6360
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg	6420
cagcggtcgg gctgaacggg ggggttcgtc acacagccca gcttgagcg aacgacctac	6480
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga	6540
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt	6600
ccagggggaa acgcttggt tctttatagt cctgtcgggt ttccccacct ctgacttgag	6660
cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg	6720
gcctttttac ggttctcggc cttttgctgg ccttttgctc acatgttctt tctgctgta	6780
tccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc	6840
agccgaacga ccgagcgag cgagtcagt agcgaggaag cggaagagcg cccaatacgc	6900
aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga cagggttccc	6960
gactggaaag cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattaggca	7020
ccccaggctt tacactttat gcttcgggt cgtatgttgt gtggaattgt gagcgataa	7080
caatttcaca caggaaacag ctatgacct gattacgcca agctctagct agaggtcgag	7140
tccctcccca gcaggcagaa gtatgcaaa catgcatctc aattagtcag caaccatagt	7200
cccgcccta actccgcca tcccgccct aactccgcc agttccgcc attctccgcc	7260

-continued

ccatggctga ctaatttttt ttatttatgc agaggccgag gccgcctcgg cctctgagct	7320
attccagaag tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa agctttgcaa	7380
agatggataa agtttttaac agagaggaat ctttgcagct aatggacctt ctaggtcttg	7440
aaaggagctc gaccaattct catgtttgac agcttatcat cgcagatccg ggcaacgttg	7500
ttgccattgc tgcaggcgca gaactggtag gtatggaaga tctatacatt gaatcaatat	7560
tggcaattag ccatattagt cattgggtat atagcataaa tcaatattgg ctattggcca	7620
ttgcatacgt tgtatctata tcataatatg tacatttata ttggctcatg tccaatatga	7680
ccgccatg	7688

<210> SEQ ID NO 31
 <211> LENGTH: 7548
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic construct: pHybE-D2-hCg1,z,a
 <400> SEQUENCE: 31

agtgggaatt ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtcgccga	60
gaagttgggg ggagggtcg gcaattgaac cgggtgcctag agaaggtggc gcggggtaaa	120
ctgggaaagt gatgtcgtgt actggctccg cctttttccc gagggtgagg gagaaccgta	180
tataagtga gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca	240
ggtaagtgcc gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgctt	300
gccttgaatt acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg	360
aagtgggtgg gagagtcca ggccttgccg ttaaggagcc ccttcgcctc gtgcttgagt	420
tgaggcctgg cctgggcgct ggggcccgcg cgtgcgaatc tggtaggcacc ttcgcgcctg	480
tctcgtgctt ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct	540
ttttttctgg caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt	600
ttttggggcc gcggggcgcg acggggcccg tgcgtcccag cgcacatggt cggcgaggcg	660
gggcctgcga gcgcggccac cgagaatcgg acgggggttag tctcaagctg gccggcctgc	720
tctgggtgct ggcctcgccg cgcctgtgat cgcgccgccc tgggcggcaa ggcgtgcccg	780
gtcggcacca gttgcgtgag cggaaagatg gccgcttccc ggccctgctg caggagagctc	840
aaaatggagg acgcggcgct cgggagagcg ggcgggtgag taccacacac aaaggaaaag	900
ggcctttccg tcctcagccg tcgcttcctg tgactccacg gactaccggg cgccttcag	960
gcacctcgat tagttctcga gcttttgag tacgtcgtct ttaggttggg gggaggggtt	1020
ttatgcgatg gagtttcccc aactgagtg ggtggagact gaagttaggc cagcttgcca	1080
cttgatgtaa ttctccttgg aatttgcct ttttgagttt ggatcttggg tcattctcaa	1140
gcctcagaca gtggttcaaa gttttttct tccatttcag gtgtcgtgag gaattctcta	1200
gagatccctc gaacctcaga tccattgtgc cggggcgcca ccatggagtt tgggctgagc	1260
tggctttttc ttgtcgcgat tttaaaagg gtccagtgtg aggtgcagct ggtggagtct	1320
gggggaggct tggtagaccc cggcagggtc ctgagactct cctgtgcggc ctctggattc	1380
acctttgatg attatgcat gcaactgggc cggcaagctc cagggaaggg cctggaatgg	1440
gtctcagcta tcacttgaa tagtggtcac atagactatg cggactctgt ggagggccga	1500

-continued

ttcaccatct	ccagagacaa	cgccaagaac	tccctgtatc	tgcaaatgaa	cagtctgaga	1560
gctgaggata	cgcccgata	ttactgtgcg	aaagtctcgt	accttagcac	cgcgtcctcc	1620
cttgactatt	ggggccaagg	taccctggtc	accgtctcga	gtgcgtcgac	caagggccca	1680
tcggctcttc	ccctggcacc	ctcctccaag	agcacctctg	ggggcacagc	ggccctgggc	1740
tgctgggtca	aggactactt	ccccgaaccg	gtgacggtgt	cgtggaactc	aggcgccctg	1800
accagcggcg	tgacacacct	cccggtgtgc	ctacagtcct	caggactcta	ctccctcagc	1860
agcgtgggtga	ccgtgccctc	cagcagcttg	ggcaccacga	cctacatctg	caacgtgaat	1920
cacaagccca	gcaacaccaa	ggtggacaag	aaagttgagc	ccaaatcttg	tgacaaaact	1980
cacacatgcc	caccgtgccc	agcacctgaa	ctcctggggg	gaccgtcagt	cttcctcttc	2040
cccccaaac	ccaagagacac	cctcatgata	tcccggaacc	ctgaggtcac	atgcgtgggtg	2100
gtggacgtga	gccacgaaga	ccctgaggtc	aagttcaact	ggtacgtgga	cggcgtggag	2160
gtgcataatg	ccaagacaaa	gccgcgggag	gagcagtaca	acagcacgta	ccgtgtggtc	2220
agcgtcctca	ccgtcctgca	ccaggactgg	ctgaatggca	aggagtacaa	gtgcaaggtc	2280
tccaacaaag	ccctcccagc	ccccatcgag	aaaaccatct	ccaaagccaa	agggcagccc	2340
cgagaaccac	aggtgtacac	cctgccccca	tcccgggatg	agctgaccaa	gaaccaggtc	2400
agcctgacct	gcctgggtcaa	aggcttctat	cccagcgaca	tcgccgtgga	gtgggagagc	2460
aatgggcagc	cggagaacaa	ctacaagacc	acgcctcccg	tgctggactc	cgcggctccc	2520
ttcttctct	acagcaagct	caccgtggac	aagagcaggt	ggcagcaggg	gaacgtcttc	2580
tcattgctcg	tgatgcatga	ggctctgcac	aaccactaca	cgcagaagag	cctctccctg	2640
tctccgggta	aataagcggc	cgctcgaggg	cggcaaggcc	ggatcccccg	acctcgacct	2700
ctggctaata	aaggaaattt	attttcattg	caatagtgtg	ttggaatttt	ttgtgtctct	2760
cactcggaag	gacatatggg	agggcaaatc	atttggtcga	gatccctcgg	agatctctag	2820
ctagaggatc	gatccccgcc	ccggaacgaac	taaacctgac	tacgacatct	ctgcccttcc	2880
ttcgcggggc	agtgcatgta	atcccttcag	ttggttggtg	caacttgcca	actgggccct	2940
gttcacatg	tgacacgggg	ggggacccaa	cacaaagggg	ttctctgact	gtagttgaca	3000
tccttataaa	tgatgttgca	catttgccaa	cactgagtg	ctttcatcct	ggagcagact	3060
ttgcagtctg	tgagatgcaa	cacaacattg	cctttatgtg	taactcttgg	ctgaagctct	3120
tacaccaatg	ctggggggaca	tgtacctccc	agggggccag	gaagactacg	ggaggctaca	3180
ccaacgtcaa	tcagaggggc	ctgtgtagct	accgataagc	ggaccctcaa	gagggcatta	3240
gcaatagtgt	ttataaggcc	cccttggtta	ccctaaacgg	gtagcatatg	cttcccggtg	3300
agtagtatat	actatccaga	ctaaccctaa	ttcaatagca	tatgttacct	aacgggaagc	3360
atatgctatc	gaattagggg	tagtaaaagg	gtcctaagga	acagcgatat	ctcccacccc	3420
atgagctgtc	acgggtttat	ttacatgggg	tcaggattcc	acgagggtag	tgaaccattt	3480
tagtcacaag	ggcagtggtc	gaagatcaag	gagcgggcag	tgaactctcc	tgaatcttcg	3540
cctgcttctt	cattctcctt	cgttttagcta	atagaataac	tgctgagttg	tgaacagtaa	3600
ggtgtatgtg	aggtgtctga	aaacaagggt	tcaggtgacg	cccccagaat	aaaatttgga	3660
cgggggggtc	agtggtgcca	ttgtgctatg	acaccaatat	aacctcaca	aacctcttgg	3720
gcaataaata	ctagtgtagg	aatgaaacat	tctgaatatc	tttaacaata	gaaatccatg	3780
gggtggggac	aagccgtaaa	gactggatgt	ccatctcaca	cgaatttatg	gctatgggca	3840

-continued

acacataatc	ctagtgc	aatatgata	ctgggtattaag	atgtgtccca	ggcagggacc	3900
aagacaggtg	aaccatgttg	ttacactcta	tttgaacaa	ggggaaagag	agtggacgcc	3960
gacagcagcg	gactccactg	gttgtctcta	acacccccga	aaattaaacg	gggctccacg	4020
ccaatggggc	ccataaacia	agacaagtgg	ccactctttt	ttttgaaatt	gtggagtggg	4080
ggcacgcgtc	agccccaca	cgccgccctg	cggttttgga	ctgtaaaata	aggggtgta	4140
aacttggtcg	attgtaaccc	cgctaaccac	tgcggtcaaa	ccacttgccc	acaaaaccac	4200
taatggcacc	ccggggaata	cctgcataag	taggtgggag	ggccaagata	ggggcgcgat	4260
tgtgtcgatc	tggaggacaa	attacacaca	cttgccgctg	agcgccaagc	acagggttgt	4320
tggctcctcat	attcacgagg	tgcgtgagag	cacggtgggc	taatgttgcc	atgggtagca	4380
tatactaccc	aaatatctgg	atagcatatg	ctatccta	ctatatctgg	gtagcatagg	4440
ctatccta	ctatatctgg	gtagcatatg	ctatccta	ctatatctgg	gtagtatatg	4500
ctatccta	ttatatctgg	gtagcatatg	ctatccta	ctatatctgg	gtagcatatg	4560
ctatccta	ctatatctgg	gtagtatatg	ctatccta	ctgtatccg	gtagcatatg	4620
ctatccta	agagattagg	gtagtatatg	ctatccta	ttatatctgg	gtagcatata	4680
ctacccaaat	atctggatag	catatgctat	cctaacttat	atctgggtag	catatgctat	4740
cctaacttat	atctgggtag	cataggttat	cctaacttat	atctgggtag	catatgctat	4800
cctaacttat	atctgggtag	tatatgctat	cctaatttat	atctgggtag	cataggttat	4860
cctaacttat	atctgggtag	catatgctat	cctaacttat	atctgggtag	tatatgctat	4920
cctaacttgt	atccgggtag	catatgctat	cctcatgata	agctgtcaaa	catgagaatt	4980
ttcttgaa	cgaaagggcc	tctgtatagc	cctattttta	taggttaatg	tcatgataat	5040
aatggtttct	tagacgtcag	gtggcacttt	tccgggaaat	gtgcgcggaa	ccctattttg	5100
tttttttttc	taataacatt	caaataatgta	tccgctcatg	agacaataac	cctgataaat	5160
gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	tgcgcccttat	5220
tccccctttt	gcggcatttt	gccttccctg	ttttgctcac	ccagaaacgc	tggtgaaagt	5280
aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	atctcaacag	5340
cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttaa	5400
agttctgcta	tgtggcgctg	tattatcccg	tgttgacgcc	gggcaagagc	aactcggtcg	5460
ccgcatacac	tattctcaga	atgacttggt	tgagtactca	ccagtcacag	aaaagcatct	5520
tacggatggc	atgacagtaa	gagaattatg	cagtgcctgc	ataaccatga	gtgataaac	5580
tgcggccaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	cttttttgca	5640
caacatgggg	gatcatgtaa	ctcgccttga	tcggtgggaa	ccggagctga	atgaagccat	5700
accaaaccgac	gagcgtgaca	ccacgatgcc	tgcagcaatg	gcaacaacgt	tgcgcaaact	5760
attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	ggatggaggc	5820
ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttccg	gctggctggt	ttattgctga	5880
taaatctgga	gccgggtgagc	gtgggtctcg	cgggtatcatt	gcagcactgg	ggccagatgg	5940
taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	caggcaacta	tggatgaacg	6000
aaatagacag	atcgctgaga	taggtgcctc	actgattaag	cattggtaac	tgtcagacca	6060
agtttactca	tatatacttt	agattgattt	aaaacttcat	ttttaattta	aaaggatcta	6120

-continued

ggatgaagatc	ctttttgata	atctcatgac	caaaatccct	taacgtgagt	tttcgttcca	6180
ctgagcgta	gaccccgtag	aaaagatcaa	aggatcttct	tgagatcctt	ttttctgcg	6240
cgtaatctgc	tgcttgcaaa	caaaaaaacc	accgctacca	gcggtggttt	gtttgcggga	6300
tcaagagcta	ccaactcttt	ttccgaaggt	aactggcttc	agcagagcgc	agataccaaa	6360
tactgttctt	ctagtgtagc	cgtagttagg	ccaccacttc	aagaactctg	tagcacccgc	6420
tacatacttc	gctctgctaa	tcctgttacc	agtggctgct	gccagtggcg	ataagtcgtg	6480
tcttacccgg	ttggactcaa	gacgatagtt	accggataag	gcgcagcggg	cgggctgaac	6540
ggggggttcg	tgacacacgc	ccagcttgga	gcgaacgacc	tacaccgaac	tgagatacct	6600
acagcgtgag	ctatgagaaa	gcgccacgct	tcccgaaggg	agaaaggcgg	acaggtatcc	6660
ggtaagcggc	agggtcggaa	caggagagcg	cacgaggag	cttccagggg	gaaacgcctg	6720
gtatctttat	agtctgtcgc	ggtttcgcc	cctctgactt	gagcgtcgat	ttttgtgatg	6780
ctcgtcaggg	gggcggagcc	tatggaaaaa	cgccagcaac	gcggcctttt	tacggttcct	6840
ggccttttgc	tgcccttttg	ctcacatgtt	cttctctgcg	ttatccctcg	attctgtgga	6900
taaccgtatt	accgcctttg	agtgaactga	taccgctcgc	cgacgccgaa	cgaccgagcg	6960
cagcgagtca	gtgagcgagg	aagcggaaga	gcgcccaata	cgcaaaccgc	ctctcccgcg	7020
gcgttgcccg	attcattaat	gcagctggca	cgacaggttt	cccgaactga	aagcgggcag	7080
tgagcgcaac	gcaattaatg	tgagttagct	caactcattg	gcaccccgag	ctttacactt	7140
tatgcttccg	gctcgtatgt	tgtgtggaat	tgtgagcgga	taacaatttc	acacaggaaa	7200
cagctatgac	catgattacg	ccaagctcta	gctagaggtc	gagtcctctc	ccagcaggca	7260
gaagtatgca	aagcatgcat	ctcaattagt	cagcaacccat	agtcgccgcc	ctaactccgc	7320
ccatcccgcc	cctaactccg	cccagttccg	cccattcttc	gccccatggc	tgactaattt	7380
tttttattta	tcagagggcc	gaggccgcct	cgccctctga	gctattccag	aagtagtgag	7440
gaggcttttt	tggaggccca	ggcttttgca	aaaagctttg	caaagatgga	taaagtttta	7500
aacagagagg	aatctttgca	gctaattggac	cttctaggtc	ttgaaagg		7548

<210> SEQ ID NO 32

<211> LENGTH: 6845

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic construct: pHybE-E7-hCk

<400> SEQUENCE: 32

agctttgcaa	agatggataa	agttttaaac	agagaggaat	ctttgcagct	aatggacctt	60
ctaggtcttg	aaaggagtgg	gaattggctc	cggtgcccg	cagtgggcag	agcgcacatc	120
gcccacagtc	cccagagaagt	tggggggagg	ggtcggcaat	tgaaccggtg	cctagagaag	180
gtggcgccgg	gtaaactggg	aaagtgatgt	cgtgtactgg	ctccgccttt	ttcccagggg	240
tgggggagaa	ccgtatataa	gtgcagtagt	cgccgtgaac	gttctttttc	gcaacggggt	300
tgccgccaga	acacaggtaa	gtgccgtgtg	tggttccgcg	gggcctggcc	tctttacggg	360
ttatggccct	tgctgtcctt	gaattacttc	cacctggctg	cagtacgtga	ttcttgatcc	420
cgagcttcgg	gttggaagtg	ggtgggagag	ttcgaggcct	tgcgcttaag	gagccccttc	480
gcctcgtgct	tgagttgagg	cctggcctgg	gcgctggggc	cgccgcgtgc	gaatctgggtg	540
gcaccttcgc	gcctgtctcg	ctgctttcga	taagtctcta	gccatttaaa	atttttgatg	600

-continued

acctgctgcg acgctttttt tctggcaaga tagtcttgta aatgcggggc aagatctgca	660
cactggtatt tcggtttttg gggccgcggg cggcgacggg gcccggtcgt cccagcgcac	720
atgttcggcg aggcggggcc tgcgagcgcg gccaccgaga atcggacggg ggtagtctca	780
agctggccgg cctgctctgg tgccctggcct cgcgcgcgcg tgtatcgccc cgccctgggc	840
ggcaaggctg gcccggtcgg caccagttgc gtgagcggaa agatggccgc tccccggccc	900
tgctgcaggg agctcaaaat ggaggacgcg gcgctcggga gacggggcgg gtgagtcacc	960
cacacaaagg aaaagggcct ttcctgcctc agccgtcgtc tcatgtgact ccacggagta	1020
ccgggcgcgc tcaggcacc tcgattagtt ctcgagcttt tggagtacgt cgtctttagg	1080
ttggggggag gggttttatg cgatggagtt tccccacact gagtgggtgg agactgaagt	1140
taggccagct tggcacttga tgtaattctc cttggaattt gccctttttg agtttggatc	1200
ttggttcatt ctcaagcctc agacagtggg tcaaagtttt tttcttccat ttcagggtgc	1260
gtgaggaatt ctctagagat cccctgcacct cgagatccat tgtgccccgg cgcaccatgg	1320
acatgcgcgt gcccgcaccg ctgctgggccc tgctgctgct gtggttcccc ggctcgcgat	1380
gcgacatcca gatgaccag tctccatcct ccctgtctgc atctgtaggg gacagagtca	1440
ccatcacttg tcgggcaagt cagggcatca gaaattactt agcctggtat cagcaaaaac	1500
cagggaagc ccctaagctc ctgatctatg ctgcatccac tttgcaatca ggggtcccat	1560
ctcggttcag tggcagtggg tctgggacag atttactctc caccatcagc agcctacagc	1620
ctgaagatgt tgcaacttat tactgtcaaa ggtataaccg tgcaccgtat acttttggcc	1680
aggggaccaa ggtggaaatc aaacgtacgg tggctgcacc atctgtcttc atcttcccgc	1740
catctgatga gcagtgtaaa tctggaactg cctctgttgt gtgcctgctg aataacttct	1800
atcccagaga ggccaaaagta cagtgggaagg tggataacgc cctccaatcg ggtaactccc	1860
aggagagtgt cacagagcag gacagcaagg acagcaccta cagcctcagc agcaccctga	1920
cgctgagcaa agcagactac gagaaacaca aagtctacgc ctgcgaagtc acccatcagg	1980
gcctgagctc gcccgtcaca aagagcttca acaggggaga gtgttgagcg gccgctcgag	2040
gccggcaagg ccggatcccc cgacctcgac ctctggctaa taaaggaaat ttattttcat	2100
tgcaatagtg tgttggaatt ttttgtgtct ctactcggg aggcacatag ggagggcaaa	2160
tcatttggtc gagatccctc ggagatctct agctagaggg tcgatccccg ccccggaaga	2220
actaaacctg actacgacat ctctgcccct tcttcgcggg gcagtgcagt taatcccttc	2280
agttggttgg tacaacttgc caactgggcc ctgttccaca tgtgacacgg ggggggacca	2340
aacacaaagg ggttctctga ctgtagtga catccttata aatggatgtg cacatttgcc	2400
aacactgagt ggctttcctc ctggagcaga ctttgagtc tgtggactgc aacacaacat	2460
tgcccttatg tgtaactctt ggctgaagct cttacaccaa tgctggggga catgtacctc	2520
ccagggggcc aggaagacta cgggaggcta caccaacgtc aatcagaggg gcctgtgtag	2580
ctaccgataa gcggaccctc aagagggcat tagcaatagt gtttataagg ccccttgtt	2640
aaccctaaac gggtagcata tgcttcccgg gtagtagtat atactatcca gactaacctc	2700
aattcaatag catatgttac ccaacgggaa gcatatgcta tcgaattagg gttagtaaaa	2760
gggtcctaag gaacagcgat atctcccacc coatgagctg tcacgggtttt atttacatgg	2820
ggtcaggatt ccacgagggt agtgaacctt tttagtcaca agggcagtgg ctgaagatca	2880

-continued

aggagcgggc	agtgaactct	cctgaatctt	cgctgcttc	ttcattctcc	ttcgtttagc	2940
taatagaata	actgctgagt	tgtgaacagt	aaggtgtatg	tgagggtctc	gaaaacaagg	3000
tttcagggtga	cgccccaga	ataaaatttg	gacggggggg	tcagtgggtg	cattgtgcta	3060
tgacaccaat	ataaccctca	caaaccctt	gggcaataaa	tactagtgtg	ggaatgaaac	3120
attctgaata	tctttaacaa	tagaaatcca	tgggggtggg	acaagccgta	aagactggat	3180
gtccatctca	cacgaattta	tggctatggg	caacacataa	tcctagtgtg	atatgatact	3240
gggggttatta	agatgtgtcc	caggcaggga	ccaagacagg	tgaaccatgt	tgttacactc	3300
tatttgtaac	aaggggaaa	agagtggacg	cgcacagcag	cggactccac	tggttgtctc	3360
taacaccccc	gaaaattaaa	cggggtctca	cgccaatggg	gcccataaac	aaagacaagt	3420
ggccactctt	tttttga	ttgtggagt	ggggcacg	tcagccccc	cacgcgccc	3480
tgcggtttt	gactgtaaaa	taagggtgtg	ataacttggc	tgattgtaac	cccgttaacc	3540
actgcggtca	aaccacttgc	ccacaaaacc	actaatggca	ccccggggaa	tacctgcata	3600
agtaggtggg	cgggccaaga	taggggcg	attgctgcga	tctggaggac	aaattacaca	3660
cacttgccc	tgagcgcaa	gcacagggtt	gttggtctc	atattcacga	ggcgtgtgag	3720
agcaggtgg	gctaattgtg	ccatgggtag	catatactac	ccaaatatct	ggatagcata	3780
tgctatccta	atctatatct	gggtagcata	ggctatccta	atctatatct	gggtagcata	3840
tgctatccta	atctatatct	gggtagtata	tgctatccta	atttatatct	gggtagcata	3900
ggctatccta	atctatatct	gggtagcata	tgctatccta	atctatatct	gggtagtata	3960
tgctatccta	atctgtatcc	gggtagcata	tgctatccta	atagagatta	gggtagtata	4020
tgctatccta	atttatatct	gggtagcata	tactacccaa	atatctggat	agcatatgct	4080
atcctaactct	atatctgggt	agcatatgct	atcctaactct	atatctgggt	agcataggct	4140
atcctaactct	atatctgggt	agcatatgct	atcctaactct	atatctgggt	agtatatgct	4200
atcctaattt	atatctgggt	agcataggct	atcctaactct	atatctgggt	agcatatgct	4260
atcctaactct	atatctgggt	agtatatgct	atcctaactct	gtatccgggt	agcatatgct	4320
atcctcatga	taagctgtca	aacatgagaa	tttcttgaa	gacgaaagg	cctcgtgata	4380
cgctatttt	tataggttaa	tgtcatgata	ataatggtt	cttagacgtc	aggtggcact	4440
tttcggggaa	atgtgcgcg	aaccctatt	tgtttat	tctaaataca	ttcaaata	4500
tatccgctca	tgagacaata	accctgataa	atgcttcaat	aatattgaaa	aaggaagagt	4560
atgagtattc	aacatttccg	tgtcgccctt	attccctttt	ttgcggcatt	ttgccttcct	4620
gtttttgtctc	accagaaa	gctggtgaaa	gtaaaagatg	ctgaagatca	gttgggtgca	4680
cgagtgggtt	acatcgaa	ggatctcaac	agcggtaaga	tccttgagag	tttcgcccc	4740
gaagaacgtt	ttccaatgat	gagcactttt	aaagttctgc	tatgtggcgc	ggtattatcc	4800
cgtgttgacg	cggggcaaga	gcaactcgg	cgcgcatac	actattctca	gaatgacttg	4860
gttgagtact	caccagtcac	agaaaagcat	cttacggatg	gcattgacagt	aagagaatta	4920
tgcatgtctg	ccataaccat	gagtgataac	actgcggcca	acttacttct	gacaacgatc	4980
ggaggaccga	aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	aactcgccct	5040
gatcgttggg	aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	5100
cctgcagcaa	tggcaacaac	gttgcgcaaa	ctattaactg	gcgaactact	tactctagct	5160
tcccggcaac	aattaataga	ctggatggag	gcggataaag	ttgcaggacc	acttctgcgc	5220

-continued

tcggcccttc cggtggctg gtttattgct gataaatctg gagccggtga gctgggtct	5280
cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac	5340
acgacgggga gtcagcaac tatggatgaa cgaaatagac agatcgctga gatagggtgcc	5400
tcaactgatta agcattggta actgtcagac caagtttact catatatact ttagattgat	5460
ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga taatctcatg	5520
acaaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccg agaaaagatc	5580
aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa	5640
ccaccgctac cagcgggtgtt ttgtttgccg gatcaagagc taccaactct tttccgaag	5700
gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtga gccgtagtta	5760
ggccaccact tcaagaactc tgtagaccg cctacatacc tcgctctgct aatcctgtta	5820
ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc aagacgatag	5880
ttaccggata aggcgcagcg gtcgggtga acggggggtt cgtgcacaca gccagcttg	5940
gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg	6000
cttcccgaag ggagaaagc ggacaggtat ccggaagcg gcagggtcgg aacaggagag	6060
cgacgaggg agcttccagg gggaaacgcc tggatatctt atagtcctgt cgggtttcgc	6120
cacctctgac ttgagcgtc atttttgtga tgctcgtcag gggggcggag cctatggaaa	6180
aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg	6240
ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgaact	6300
gataccgctc gccgcagccg aacgaccgag cgcagcaggt cagtgcagca ggaagcggaa	6360
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg	6420
cacgacaggt ttcccagctg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag	6480
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga	6540
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagctc	6600
tagctagagg tcgagtcctt ccccagcagg cagaagtatg caaagcatgc atctcaatta	6660
gtcagcaacc atagtcctgc ccctaactcc gccatcccg ccctaactc cgcccagttc	6720
cgccattct ccgccccatg gctgactaat tttttttatt tatgcagagg ccgaggccgc	6780
ctcgccctct gagctattcc agaagtagtg aggaggcttt tttggaggcc taggcttttg	6840
caaaa	6845

1. An expression vector comprising:

- (a) an OriP origin of replication derived from Epstein-Barr virus (EBV);
- (b) an SV40 origin of replication;
- (c) an insertion site for inserting a gene of interest;
- (d) a promoter operably linked to the insertion site, wherein the promoter is an EF-1 α promoter or a CMV promoter; and, optionally,
- (e) a nucleic acid sequence encoding an antibody heavy or light chain constant region that is operably linked to the insertion site.

2. The expression vector of claim **1**, wherein the gene of interest is an antibody heavy or light chain variable region.

3. The expression vector of claim **2**, wherein the antibody heavy or light chain variable region is selected from the group consisting of: murine, humanized, chimeric, and human.

4. The expression vector of claim **2**, wherein the antibody heavy chain variable region is the heavy chain variable region of an antibody selected from the group consisting of: an anti-TNF α antibody, an anti-IL-18 antibody, and an anti-IL-12 antibody.

5. The expression vector of claim **2**, wherein the antibody light chain variable region is the light chain variable region of an antibody selected from the group consisting of: an anti-TNF α antibody, an anti-IL-18 antibody, and an anti-IL-12 antibody.

6. The expression vector of claim **1**, wherein the antibody heavy chain constant region is murine or human.

7. The expression vector of claim 1, wherein the antibody heavy chain constant region is selected from the group consisting of: gamma 1, z, a; gamma 1, z, non-a; gamma 2, n+; gamma 2, n-; and gamma 4.

8. The expression vector of claim 7, wherein the gamma 1, z, non-a antibody heavy chain constant region further comprises an alanine mutation at position 234 of the heavy chain constant region.

9. The expression vector of claim 8, wherein the expression vector further comprises an alanine mutation at either position 235 or 237 of the heavy chain constant region.

10. The expression vector of claim 1, wherein the antibody light chain constant region is either a human kappa isotype or a human lambda isotype.

11. The expression vector of claim 1, wherein the antibody heavy chain constant region is either a murine gamma 1 isotype or a murine gamma 2a isotype.

12. The expression vector of claim 1, wherein the antibody light chain constant region is a murine kappa isotype.

13. The expression vector of claim 1, wherein the antibody heavy chain constant region is an Fc domain.

14. The expression vector of claim 2, wherein the heavy or light chain variable region is 5' to the insertion site.

15. (canceled)

16. The expression vector of claim 1, wherein the expression vector further comprises a selectable marker.

17. The expression vector of claim 16, wherein the selectable marker is an ampicillin resistance gene.

18. (canceled)

19. (canceled)

20. (canceled)

21. The expression vector of claim 1, wherein the CMV promoter comprises nucleotides 1 to 608 of SEQ ID NO: 1.

22. The expression vector of claim 1, wherein the EF-1 α promoter is human.

23. (canceled)

24. (canceled)

25. (canceled)

26. The expression vector of claim 1, wherein the EF-1 α promoter comprises nucleotides 76 to 1267 of SEQ ID NO: 2.

27. (canceled)

28. (canceled)

29. The expression vector of claim 1, wherein the OriP origin of replication comprises nucleotides 1795 to 3545 of SEQ ID NO: 1.

30. (canceled)

31. (canceled)

32. (canceled)

33. The expression vector of claim 1, wherein the SV40 origin of replication comprises nucleotides 5834 to 6140 of SEQ ID NO: 1.

34. (canceled)

35. (canceled)

36. (canceled)

37. An expression vector comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, and 32.

38. (canceled)

39. The expression vector of claim 1, wherein the expression vector further comprises a nucleic acid sequence encoding a signal peptide.

40. A kit comprising the expression vector of claim 1.

41. A host cell comprising the expression vector of claim 1.

42. (canceled)

43. (canceled)

44. (canceled)

45. A method for producing a recombinant protein comprising introducing the expression vector of claim 1 into a host cell, culturing the host cell under suitable conditions so as to express the protein, and recovering the protein.

46. A kit comprising the expression vector of claim 37.

47. A host cell comprising the expression vector of claim 37.

48. A method for producing a recombinant protein comprising introducing the expression vector of claim 37 into a host cell, culturing the host cell under suitable conditions so as to express the protein, and recovering the protein.

* * * * *