

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 September 2007 (07.09.2007)

PCT

(10) International Publication Number
WO 2007/100908 A2

(51) International Patent Classification: **Not classified**

[US/US]; 146 Laidley Street, San Francisco, California 94131 (US).

(21) International Application Number:

PCT/US2007/005386

(74) Agents: **FANG, Carol, A.** et al.; TOWNSEND and TOWNSEND and CREW LLP, Two Embarcadero Center, Eighth Floor, San Francisco, California 94111-3834 (US).

(22) International Filing Date:

28 February 2007 (28.02.2007)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/778,026	28 February 2006 (28.02.2006)	US
60/801,645	19 May 2006 (19.05.2006)	US
60/802,992	22 May 2006 (22.05.2006)	US
60/821,492	4 August 2006 (04.08.2006)	US
60/846,658	22 September 2006 (22.09.2006)	US
60/848,195	28 September 2006 (28.09.2006)	US

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(71) Applicant (*for all designated States except US*): **WEST COAST BIOLOGICALS** [US/US]; 600 Townsend Street, Suite 120E, San Francisco, California 94103 (US).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

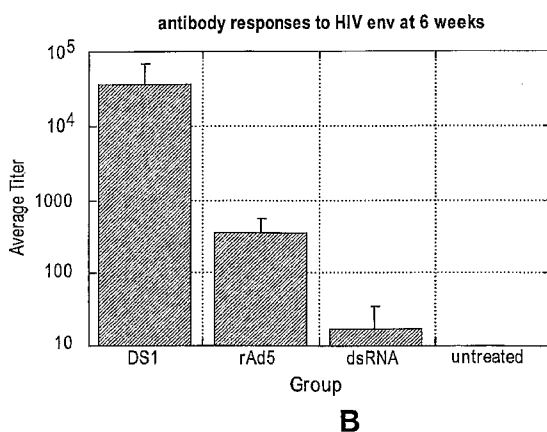
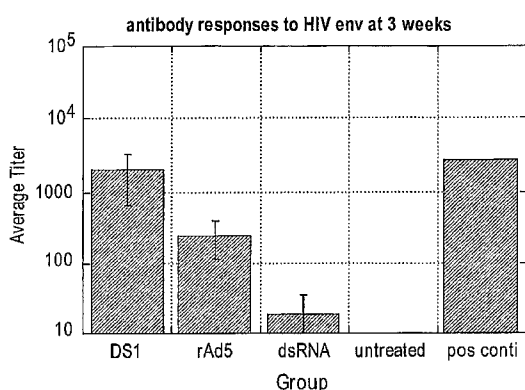
(72) Inventor; and

(75) Inventor/Applicant (*for US only*): **TUCKER, Sean, N.**

[Continued on next page]

(54) Title: CHIMERIC ADENOVIRAL VECTORS

(57) Abstract: The present invention provides chimeric adenoviral vectors and methods for using the vectors to elicit an immune response to an antigen of interest.



WO 2007/100908 A2



FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished
upon receipt of that report*

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

Chimeric Adenoviral Vectors

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] The present application claims priority to U.S. Provisional Patent Application No. 60/778026, filed February 28, 2006, U.S. Provisional Patent Application No. 60/801,645, filed May 19, 2006, U.S. Provisional Patent Application No. 60/802,992, filed May 22, 2006, U.S. Provisional Patent Application No. 60/821,492, filed August 4, 2006, U.S. Provisional Patent Application No. 60/846,658, filed September 22, 2006, and U.S. Provisional Patent Application No. 60/848,195, filed September 28, 2006), the disclosures of which are hereby incorporated by reference in their entirety for all purposes.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] Not applicable.

BACKGROUND OF THE INVENTION

[0003] Vaccines are an important means for preventing and/or treating a number of diseases and disorders (*e.g.*, viral infection, bacterial infection, and cancer). Nucleic acid-based vaccines have several advantages over protein or attenuated-live vaccines. Introduction of a nucleic acid that expresses an antigen into a target cell allows for rapid development of vaccine that generates an immune response against an antigen of interest. For protein vaccines, an effective and efficient method of protein purification needs to be developed each time a new vaccine is created. For live vaccines, a method of attenuation needs to be identified that doesn't completely stop the growth of the pathogen, yet proven to be completely safe in humans. Development of protein purification and attenuation methodologies are extremely time-consuming processes. In contrast, most nucleic acid-based vaccines can be manufactured very quickly using the same manufacturing techniques each time with just a quick change in the nucleic acid encoding the antigen of interest. Replication incompetent adenovirus is one nucleic acid-based vaccine system which is rapidly, predictably, and inexpensively made at high titer [Polo, J.M. and Dubensky, T.W., Jr., *Drug Discov Today*, 7(13), 719-727 (2002)]. However, the efficiency of the antigen-specific response following administration of adenoviral vectors known in the art is low. Thus, there

is a need in the art for new adenoviral vectors that can be used to efficiently elicit an immune response against an antigen of interest. The present invention satisfies these and other needs.

SUMMARY OF THE INVENTION

[0004] The present invention provides chimeric adenoviral vectors comprising nucleic acids encoding a heterologous polypeptide and methods for eliciting an immune response against the heterologous polypeptide.

[0005] One embodiment of the invention provides chimeric adenoviral expression vectors comprising an expression cassette comprising: (a) first promoter operably linked to a nucleic acid encoding a toll-like receptor (TLR)-3 agonist; and (b) a second promoter operably linked to a nucleic acid encoding a heterologous polypeptide. In some embodiments, the TLR-3 agonist is dsRNA. In some embodiments, the nucleic acid encoding the TLR agonist comprises a sequence selected from SEQ ID NOS: 3, 7, 8, 9, 10, 11, and 12. In some embodiments, the heterologous polypeptide is selected from an HIV envelope polypeptide (e.g., gp41, gp120 or gp160) and influenza HA polypeptide. In some embodiments, the first and second promoters are the same. In some embodiments, the first and second embodiments are different. In some embodiments, the promoters are selected from the beta actin promoter and the CMV promoter. The invention also provides immunogenic compositions comprising the expression vector.

[0006] A further embodiment of the invention provides methods of eliciting an immune response against the heterologous polypeptide by administering an immunogenically effective amount of the compositions to a mammalian subject (e.g., a rodent such as a mouse, a rat, or a guinea pig or a primate such as a chimpanzee, a rhesus macaque, or a human). In some embodiments, the vector is administered via any non-parenteral route (e.g., orally, intranasally, or mucosally). In some embodiments, the heterologous polypeptide is expressed in a cell selected from a dendritic cell, a microfold cell, and an intestinal epithelial cell.

[0007] A further embodiment of the invention provides immunogenic compositions comprising: (a) a chimeric adenoviral expression vector comprising a promoter operably linked to a nucleic acid encoding a heterologous polypeptide; and (b) a TLR-3 agonist (e.g., a dsRNA). In some embodiments, the TLR-3 agonist is encoded by a nucleic acid. The invention also provides methods of eliciting an immune response by administering the compositions to a mammalian subject (e.g., a rodent such as a mouse, a rat, or a guinea pig or

a primate such as a chimpanzee, a rhesus macaque, or a human) via any non-parenteral route (e.g., oral, intranasal, or mucosal).

[0008] Another embodiment of the invention provides an isolated nucleic acid comprising the sequence set forth in SEQ ID NOS:1, 2, 6, 7, 13, 14, 15, 16, or 17.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Figure 1 illustrates data demonstrating that a chimeric adenoviral vector of the invention (i.e., DS1) in combination with a TLR-3 agonist is more effective than a standard adenoviral vector (i.e., rAd5) at inducing an antigen specific immune response following oral vector delivery. Figure 1A illustrates data depicting the antibody titer to HIV envelope protein (i.e., gp120) at 3 weeks following oral delivery of the adenoviral vectors. Figure 1B illustrates data depicting the antibody titer to HIV envelope protein (i.e., gp120) at 6 weeks following oral delivery of the adenoviral vectors.

[0010] Figure 2 illustrates data demonstrating that a chimeric adenoviral vector of the invention (i.e., DS1b or DS1c) in combination with a TLR-3 agonist is more effective at inducing an antigen specific immune response than a standard adenoviral vector (i.e., rAd5). Figure 2A illustrates data depicting the anti-GFP IgG titer at 3 weeks following oral administration of the vectors. Figure 2B illustrates data depicting the CD8⁺ T cell response to GFP at 10 weeks following administration of the vector at 0, 4, and 8 weeks. Figure 2C illustrates data depicting the anti-HA antibody titer at 3 weeks following oral administration of the vectors.

[0011] Figure 3 illustrates data demonstrating that the chimeric adenoviral vectors of the invention are superior for eliciting immune responses when administered non-parenterally. Figure 3A illustrates data depicting the anti-gp120 antibody titer 3 weeks following intramuscular administration of DS1. Figure 3B illustrates data depicting the anti-HA antibody titer three weeks following intranasal administration of DS1c.

[0012] Figure 4 illustrates data demonstrating that the expressed TLR-3 ligand agonists can induce activation of antigen presenting cells. Figure 4A illustrates data depicting dendritic cell activation by the expressed dsRNA TLR-3 agonist luc1. Figure 4B illustrates data depicting dendritic cell activation by the expressed dsRNA TLR-3 agonists luc1 and m1.

[0013] Figure 5 is a graphic illustration of the chimeric adenoviral vectors of the invention, i.e., chimeric adenoviral vectors comprising nucleic acids encoding expressed ds RNA TLR-3 agonists.

[0014] Figure 6 illustrates data demonstrating that the chimeric adenoviral vectors of the invention are effective at inducing an antigen-specific immune response following oral delivery. Figure 6 illustrates data depicting anti-gp120 antibody titer 3 weeks following oral administration of a chimeric adenoviral comprising a nucleic acid sequence encoding the dsRNA TLR-3 agonist luc1.

[0015] Figure 7 illustrates data demonstrating that TLR-7/8 agonists have poor effectiveness in inducing an antigen-specific immune response.

[0016] Figure 8 illustrates data demonstrating that chimeric adenoviral vectors of the invention are effective at inducing an antigen-specific immune response following oral delivery. Figure 8A illustrates data depicting the anti-HA antibody titer 4 weeks following oral administration of a chimeric adenoviral comprising a nucleic acid sequence encoding the dsRNA TLR-3 agonist luc1. Figure 8B illustrates data depicting the anti-HA antibody titer 4 weeks or 7 weeks following administration of a chimeric adenoviral comprising a nucleic acid sequence encoding the dsRNA TLR-3 agonist luc1. Figure 8C illustrates data depicting the anti-HA antibody titer 3 weeks following oral or intranasal administration of a chimeric adenoviral comprising a nucleic acid sequence encoding the dsRNA TLR-3 agonist luc1.

BRIEF DESCRIPTION OF THE SEQUENCES

[0017] SEQ ID NO:1 sets forth the nucleotide sequence for the chimeric adenoviral vector DS1.

[0018] SEQ ID NO:2 sets forth the nucleotide sequence for the chimeric adenoviral vector DS2.

[0019] SEQ ID NO:3 sets forth a nucleotide sequence encoding a TLR-3 agonist.

[0020] SEQ ID NO:4 sets forth a nucleotide sequence encoding a TLR-3 agonist.

[0021] SEQ ID NO:5 sets forth a nucleotide sequence encoding a TLR-3 agonist.

[0022] SEQ ID NO:6 sets forth a nucleotide sequence for a chimeric adenoviral vector comprising a nucleic acid encoding influenza HA and a nucleic acid encoding a TLR-3 agonist (luc), wherein the influenza HA and the TLR-3 agonist are in the same orientation.

[0023] SEQ ID NO: 7 sets forth a nucleotide sequence for a chimeric adenoviral vector comprising a nucleic acid encoding influenza HA and a nucleic acid encoding a TLR-3 agonist (luc), wherein the influenza HA and the TLR-3 agonist are in the opposite orientation.

[0024] SEQ ID: 8 sets forth a nucleotide sequence encoding a short hairpin RNA TLR-3 agonist. Complementary portions of the sequence are shown in capital letters and the linker sequence is shown in lower case letters.

[0025] SEQ ID NO: 9 sets forth a nucleotide sequence encoding a short hairpin RNA TLR-3 agonist (g1). Complementary portions of the sequence are shown in capital letters and the linker sequence is shown in lower case letters.

[0026] SEQ ID NO: 10 sets forth a nucleotide sequence encoding a short hairpin RNA TLR-3 agonist (luc). Complementary portions of the sequence are shown in capital letters and the linker sequence is shown in lower case letters.

[0027] SEQ ID NO: 11 sets forth a nucleotide sequence encoding a short hairpin RNA TLR-3 agonist (m1). Complementary portions of the sequence are shown in capital letters and the linker sequence is shown in lower case letters.

[0028] SEQ ID NO: 12 sets forth a nucleotide sequence encoding a short hairpin RNA TLR-3 agonist. Complementary portions of the sequence are shown in capital letters and the linker sequence is shown in lower case letters.

[0029] SEQ ID NO: 13 sets forth the nucleotide sequence for the chimeric adenoviral vector DS1c. The sequence comprises a nucleotide encoding HA(PR8/34).

[0030] SEQ ID NO: 14 sets forth the nucleotide sequence for the chimeric adenoviral vector DS2beta-luc. The vector comprises a sequence encoding the TLR-3 agonist luc under the control of the beta actin promoter. The vector also comprises open cloning sites for insertion of nucleic acid sequence(s) encoding an antigen of interest.

[0031] SEQ ID NO: 15 sets forth the nucleotide sequence for the chimeric adenoviral vector DS2C-luc. The vector comprises a sequence encoding the TLR-3 agonist luc under the control of the CMV promoter. The vector also comprises open cloning sites for insertion of nucleic acid sequence(s) encoding an antigen of interest.

[0032] SEQ ID NO: 16 sets forth the nucleotide sequence for the pshuttle vector comprising a nucleic acid sequence encoding the TLR-3 agonist luc under the control of the CMV promoter and a nucleic acid sequence encoding HA (avian flu) under the control of a separate CMV promoter

[0033] SEQ ID NO: 17 sets forth the nucleotide sequence for the chimeric adenoviral vector ND1.1 214. The nucleic acid encoding the heterologous antigen is in bold text and is flanked by a Cla I recognition site on the 5' end and a Not I recognition site on the 3' end. The nucleic acid sequence encoding the TLR-3 agonists is in italic, with the linker sequence in bold.

DETAILED DESCRIPTION OF THE INVENTION

I. Introduction

[0034] The present invention provides novel chimeric adenoviral vectors that can be administered non-parenterally to elicit an immune response against an antigen of interest. The chimeric adenoviral vectors of the invention comprise a nucleic acid encoding a heterologous polypeptide and a nucleic acid encoding a TLR-3 agonist. The chimeric adenoviral vectors elicit strong and effective immune responses specific for the heterologous polypeptide, particularly when administered via a non-parenteral route (e.g., orally, intranasally, or mucosally).

[0035] The invention is based on the suprising discovery that administration of dsRNA TLR-3 agonists are effective adjuvants when administered in conjunction with viral vectors. In fact, the use of dsRNA as an adjuvant for viral vectors would be counterintuitive considering that the major proposed utility of the dsRNA mimetic poly I:C was as an antiviral agent [Nemes, *et al.*, *Proc Soc Exp Biol Med.* (1969) 132:776; Schafer, *et al.*, *Nature.* (1970) 226:449; Fenje, *et al.*, *Nature* (1970) 226:171.].

II. Definitions

[0036] The term “chimeric” or “recombinant” as used herein with reference, *e.g.*, to a nucleic acid, protein, or vector, indicates that the nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein. Thus, for example, chimeric and recombinant vectors include nucleic acid sequences that are not found within the native (non-chimeric or non-recombinant) form of the vector. A chimeric adenoviral expression vector refers to an adenoviral expression vector comprising a nucleic acid sequence encoding a heterologous polypeptide.

[0037] An “expression vector” is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

[0038] The terms “promoter” and “expression control sequence” are used herein to refer to an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of

transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. Promoters include constitutive and inducible promoters. A “constitutive” promoter is a promoter that is active under most environmental and developmental conditions. An “inducible” promoter is a promoter that is active under environmental or developmental regulation. The term “operably linked” refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

[0039] The terms “TLR agonist” or “Toll-like receptor agonist” as used herein refers to a compound that binds and stimulates a Toll-like receptor including, e.g., TLR-2, TLR-3, TLR-6, TLR-7, or TLR-8. TLR agonists are reviewed in MacKichan, *LAVI Report*. 9:1-5 (2005) and Abreu *et al.*, *J Immunol*, 174(8), 4453-4460 (2005). Agonists induce signal transduction following binding to their receptor.

[0040] The terms “TLR-3 agonist” or “Toll-like receptor 3 agonist” as used herein refers to a compound that binds and stimulates the TLR-3. TLR-3 agonists have been identified including double-stranded RNA, virally derived dsRNA, several chemically synthesized analogs to double-stranded RNA including polyinosine-polycytidylic acid (poly I:C) - polyadenylic-polyuridylic acid (poly A:U) and poly I:poly C, and antibodies (or cross-linking of antibodies) to TLR-3 that lead to IFN-beta production [Matsumoto, M, *et al*, *Biochem Biophys Res Commun* 24:1364 (2002), de Bouteiller, *et al*, *J Biol Chem* 18:38133-45 (2005)]. TLR-3 agonists also include expressed dsRNA (e.g., dsRNA encoded by a nucleic acid comprising a sequence set forth in SEQ ID NOS: 3, 7, 8, 9, 10, 11, or 12).

[0041] The terms “TLR-7/8 agonist” or “Toll-like receptor 7/8 agonist” as used herein refers to a compound that binds and stimulates either the TLR-7 or TLR-8 receptors; these receptors recognize several of same ligands. Several TLR-7/8 agonists have been identified such as viral single-stranded RNA, imiquimod, loxoribine, polyuridylic acid, or resiquimod.

[0042] The term “heterologous” when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises

two or more subsequences that are not found in the same relationship to each other in nature (*e.g.*, a fusion protein).

[0043] The terms “nucleic acid” and “polynucleotide” are used interchangeably herein to refer to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

[0044] Unless otherwise indicated, a particular nucleic acid sequence also encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, Nucleic Acid Res. 19:5081 (1991); Ohtsuka *et al.*, J. Biol. Chem. 260:2605-2608 (1985); Rossolini *et al.*, Mol. Cell. Probes 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

[0045] Antigen- refers to a protein or part of a polypeptide chain that can be recognized by T cell receptors and/or antibodies. Typically, antigens are derived from bacterial, viral, or fungal proteins.

[0046] An “immunogenically effective dose or amount” of the of the compositions of the present invention is an amount that elicits or modulates an immune response specific for the heterologous polypeptide. Immune responses include humoral immune responses and cell-mediated immune responses. An immunogenic composition can be used therapeutically or prophylactically to treat or prevent disease at any stage.

[0047] “Humoral immune responses” are mediated by cell free components of the blood, *i.e.*, plasma or serum; transfer of the serum or plasma from one individual to another transfers immunity.

[0048] “Cell mediated immune responses” are mediated by antigen specific lymphocytes; transfer of the antigen specific lymphocytes from one individual to another transfers immunity.

[0049] A “therapeutic dose” or “therapeutically effective amount” or “effective amount” of a chimeric adenoviral vector or a composition comprising a chimeric adenoviral vector is an amount of the vector or composition comprising the vector which prevents, alleviates, abates, or reduces the severity of symptoms of diseases and disorders associated with the source of the heterologous polypeptide (*e.g.*, a virus, bacteria, a parasite, or a cancer).

[0050] Antibody- refers to a polypeptide encoded by an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[0051] T cells- refer to a particular class of lymphocytes that express a specific receptor (T cell receptor) encoded by a family of genes. The recognized T cell receptor genes include alpha, beta, delta, and gamma loci, and the T cell receptors typically (but not universally) recognize a combination of MHC plus a short peptide.

[0052] Adaptive immune response- refers to T cell and/or antibody recognition of antigen.

[0053] Antigen presenting cells (APCs)- as used herein refers to cells that are able to present immunogenic peptides or fragments thereof to T cells to activate or enhance an immune response. APCs include dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects per se and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may be isolated from any of a variety of biological fluids and organs including bone marrow, peripheral blood, tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells. APCs typically utilize a receptor from the major histocompatibility (MHC) locus to present short polypeptides to T cells.

[0054] Adjuvant- is a non-specific immune response enhancer. Suitable adjuvants include, for example, cholera toxin, monophosphoryl lipid A (MPL), Freund’s Complete Adjuvant, Freund’s Incomplete Adjuvant, Quil A, and Al(OH). Adjuvants can also be those substances that cause APC activation and enhanced presentation of T cells through secondary signaling molecules like Toll-like receptors. Examples of Toll-like receptors include the receptors that recognize double-stranded RNA, bacterial flagella, LPS, CpG DNA, and bacterial lipopeptide (Reviewed recently in [Abreu *et al.*, *J Immunol*, 174(8), 4453-4460 (2005)]).

[0055] The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0056] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0057] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0058] “Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG,

which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

[0059] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[0060] The following eight groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
 - 2) Aspartic acid (D), Glutamic acid (E);
 - 3) Asparagine (N), Glutamine (Q);
 - 4) Arginine (R), Lysine (K);
 - 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
 - 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
 - 7) Serine (S), Threonine (T); and
 - 8) Cysteine (C), Methionine (M)
- (see, e.g., Creighton, *Proteins* (1984)).

[0061] The phrase “selectively (or specifically) hybridizes to” refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

[0062] The phrase “stringent hybridization conditions” refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, “Overview of principles of hybridization and the strategy of nucleic acid assays” (1993). Generally,

stringent conditions are selected to be about 5-10°C lower than the thermal melting point T_m for the specific sequence at a defined ionic strength Ph . The T_m is the temperature (under defined ionic strength, Ph , and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at Ph 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, optionally 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

[0063] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary “moderately stringent hybridization conditions” include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

[0064] “Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[0065] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kDa) and one “heavy” chain (about 50-70 kDa). The N-terminus of each

chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

[0066] The phrase “specifically (or selectively) binds” to an antibody or “specifically (or selectively) immunoreactive with,” when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to fusion proteins can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with fusion protein and not with individual components of the fusion proteins. This selection may be achieved by subtracting out antibodies that cross-react with the individual antigens. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988), for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity). Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

[0067] Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes an individual polypeptide or dsRNA or a portion thereof) or may comprise a variant of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded polypeptide is not diminished, relative to a polypeptide comprising native antigens. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the TLR-3 agonist activity of the encoded dsRNA is not diminished, relative to a dsRNA that does not contain the substitutions, additions, deletions and/or insertions. Variants preferably exhibit at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, to a polynucleotide sequence that encodes a native polypeptide or a portion thereof or to a polynucleotide sequence that encodes a dsRNA with TLR-3 agonist activity..

[0068] The terms “identical” or percent “identity,” in the context of two or more nucleic acids (*e.g.*, a dsRNA that is a TLR-3 agonist) or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more identity over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be “substantially identical.” This definition also refers to the complement of a test sequence. Optionally, the identity exists over a region that is at least about 10 to about 100, about 20 to about 75, about 30 to about 50 amino acids or nucleotides in length.

[0069] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[0070] A “comparison window”, as used herein, includes reference to a segment of any one of the number of contiguous positions from about 10 to about 500, about 25 to about 200, 50 to about 150, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat’l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

[0071] One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the

clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, *e.g.*, version 7.0 (Devereaux *et al.*, *Nuc. Acids Res.* 12:387-395 (1984).

[0072] Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1997) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-

scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

[0073] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.,* Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

III. Compositions of the Present Invention

[0074] The invention provides compositions comprising chimerical adenoviral vectors. In some embodiments, the chimeric adenoviral vectors of the invention comprise a first promoter operably linked to a nucleic acid encoding a heterologous polypeptide and a second promoter operably linked to a nucleic acid encoding a TLR3 agonist. The first and second promoters may be the same or different. In some embodiments, the first and second promoters are independently selected from: the beta actin promoter and the CMV promoter.

[0075] In some embodiments, the chimeric adenoviral vector comprises the adenoviral genome (minus the E1 and E3 genes) and a nucleic acid encoding a gene that activates IRF-3 and other signaling molecules downstream of TLR-3. The chimeric vector can be administered to a cell that expresses Ad's E1 gene such that recombinant adenovirus (rAd) is produced by the cell. This rAd can be harvested and is capable of a single round of infection that will deliver the transgenic composition to another cell within a mammal in order to elicit immune responses to the heterologous polypeptide.

A. Suitable Adenoviral Vectors

[0076] In some embodiments, the adenoviral vector is adenovirus 5, including, for example, Ad5 with deletions of the E1/E3 regions and Ad5 with a deletion of the E4 region. Other suitable adenoviral vectors include strains 2, orally tested strains 4 and 7, enteric adenoviruses 40 and 41, and other strains (*e.g.* Ad34) that are sufficient for delivering an antigen and eliciting an adaptive immune response to the transgene antigen [Lubeck *et al.*, *Proc Natl Acad Sci U S A*, 86(17), 6763-6767 (1989); Shen *et al.*, *J Virol*, 75(9), 4297-4307 (2001); Bailey *et al.*, *Virology*, 202(2), 695-706 (1994)]. In some embodiments, the adenoviral vector is a live, replication incompetent adenoviral vector (such as E1 and E3 deleted rAd5), live and attenuated adenoviral vector (such as the E1B55K deletion viruses), or a live adenoviral vector with wild-type replication.

[0077] The transcriptional and translational control sequences in expression vectors to be used in transforming vertebrate cells *in vivo* may be provided by viral sources. For example, commonly used promoters and enhancers are derived, *e.g.*, from beta actin, adenovirus, simian virus (SV40), and human cytomegalovirus (CMV). For example, vectors allowing expression of proteins under the direction of the CMV promoter, SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, transducer promoter, or other promoters shown effective for expression in mammalian cells are suitable. Further viral genomic promoter, control and/or signal sequences may be used, provided such control sequences are compatible with the host cell chosen.

B. Heterologous Polypeptides

[0078] Nucleic acids encoding suitable heterologous polypeptides may be derived from antigens, such as, for example, viral antigens, bacterial antigens, cancer antigens, fungal antigens, or parasite antigens.

[0079] Viral antigens may be derived from, for example, human immunodeficiency virus (*e.g.*, gag (p55 and p160), pol, env (gp120 and gp41) as set forth in Shiver *et al.* *Nature* 415(6869):331 (2002); the HIV genomic sequences set forth in Genbank Accession Nos. EF363127; EF363126; EF363125; EF363124; EF363123; EF363122; EF192592; and EF192591; the HIV gag sequences set forth in Genbank Accession Nos. EF396891; EF396890; EF396889; EF396888; EF396887; EF396886; EF396885; EF396884; EF396883; EF396882; EF396881; EF396880; EF396879; EF396878; EF396877; EF396876; EF39687;

EF396874; EF396873; and EF396872; the HIV pol sequences set forth in Genbank Accession Nos. EF396810; EF396809; EF396808; EF396807; EF396806; EF396805; EF396804; EF396803; EF396802; EF396801; EF396800; EF396799; EF396798; EF396797; EF396796; EF396795; EF396794; EF396793; EF396792; and EF396791; and the HIV env sequences set forth in Genbank Accession Nos. 9: EF367234; EF367233; EF367232; EF367231; EF367230; EF367229; EF367228; EF367227; EF367226; EF367225; EF367224; and EF367223, human papilloma virus (*e.g.*, capsid protein L1 as described in, *e.g.*, Donnelly *et al.* J Infect Dis. 173:314 (1996) and the sequences set forth in Genbank Accession Nos. EF362755; EF362754; NC_001694; NC_001693; NC_001691; NC_001690; NC_005134; NC_001458; NC_001457; NC_001354; NC_001352; NC_001526; and X94164), Epstein Barr virus, herpes simplex virus, human herpes virus, rhinoviruses, cocksackieviruses, enteroviruses, hepatitis A, B, C, and E (*e.g.*, hepatitis B surface antigen as described in *e.g.*, Lubeck *et al.*, PNAS USA 86:6763 (1989) and the sequences set forth in GenBank Accession Nos. AB236481; AB236471; AB206501; AB206489; AB206487; AB221788; AB221777; AB221773; AR933671; AR933670; AB236514; AB236513; AB236512; AB236511; AB236510; AB236509; AB236508; AB236507); hepatitis C NS5 (*see, e.g.*, Genbank Accession Nos. X59609; DQ911563; S71627; S70787; S70786; S70341; S62220; S70790; S70789; S70788; and AB204642)), mumps virus, rubella virus, measles virus, poliovirus, smallpox virus, rabies virus, and Variella-zoster virus.

Influenza antigens include, *e.g.*, hemagglutinin (HA), matrix protein 1 (M1), and nucleoprotein (NP) (*see, e.g.*, Donnelly, *et al.*, Vaccine 15:865 (1997) and the influenza HA sequences set forth in Genbank Accession Nos. AB294219; AB294217; AB294215; AB294213; EF102944; EF102943; EF102942; EF102941; EF102940; EF102939; EF102938; EF102937; EF102936; EF102935; EF102934; EF102933; DQ643982; DQ464354; CY019432; CY019424; CY019416; CY019408; CY019400; CY019392; CY019384; CY019376; CY019368; CY019360; CY019352; EF124794; EF110519; EF110518; EF165066; EF165065; EF165064; and EF165063; the influenza M1 sequences set forth in Genbank Accession Nos. AB292791; CY019980; CY019972; CY019964; CY019956; CY019948; CY019940; CY019628; CY019652; CY019644; CY019932; CY019924; CY019916; CY019908; CY019900; CY019892; CY019884; CY019876; CY019868; CY019860; and the influenza NP sequences set forth in Genbank Accession Nos. AB292790; CY019461; CY019974; CY019966; CY019958; CY019950; CY019942; CY019630; CY019654; CY019646; CY019934; CY019926; CY019918; CY019910; CY019902; CY019894; CY019886; CY019878; CY019870; and CY019862.

[0080] Suitable viral antigens also include, e.g., viral nonstructural proteins. The term "Viral nonstructural protein" as used herein refers to proteins encoded by viral nucleic acid that do not encode for structural polypeptides, such as those that make capsid or the protein surrounding a virus. Non-structural proteins include those proteins that promote viral nucleic acid replication and viral gene expression such as, for example, Nonstructural proteins 1, 2, 3, and 4 (NS1, NS2, NS3, and NS4, respectively) from Venezuelan Equine encephalitis (VEE), EEE, or Semliki Forest virus [Dubensky *et al.*, *J Virol*, 70(1), 508-519 (1996); Petrakova *et al* *J Virol* 2005 79(12): 7597-608; U.S. Patent Nos. 5,185,440; 5,739,026; 6,566,093; and 5,814,482. Several representative examples of suitable alphaviruses include Aura (ATCC VR-368), Bebaru virus (ATCC VR-600, ATCC VR-1240), Cabassou (Genbank Accession Nos. AF398387, ATCC VR-922), Chikungunya virus (ATCC VR-64, ATCC VR-1241), Eastern equine encephalomyelitis virus (Genbank Accession Nos. AY705241, AY705240, ATCC VR-65, ATCC VR-1242), Fort Morgan (ATCC VR-924), Getah virus (ATCC VR-369, ATCC VR-1243), Kyzylagach (ATCC VR-927), Mayaro (ATCC VR-66), Mayaro virus (ATCC VR-1277), Middleburg (ATCC VR-370), Mucambo virus (ATCC VR-580, ATCC VR-1244), Ndumu (ATCC VR-371), Pixuna virus (ATCC VR-372, ATCC VR-1245), Ross River virus (ATCC VR-373, ATCC VR-1246), Semliki Forest (Genbank Accession Nos. AJ251359, ATCC VR-67, ATCC VR-1247), Sindbis virus (Genbank Accession Nos. J02363, ATCC VR-68, ATCC VR-1248), Tonate (ATCC VR-925), Trinita (ATCC VR-469), Una (ATCC VR-374), Venezuelan equine encephalomyelitis (ATCC VR-69), Venezuelan equine encephalomyelitis virus (Genbank Accession Nos. AY986475, AY973944, NC 001449, ATCC VR-923, ATCC VR-1250 ATCC VR-1249, ATCC VR-532), Western equine encephalomyelitis (ATCC VR-70, ATCC VR-1251, ATCC VR-622, ATCC VR-1252), Whataroa (ATCC VR-926), and Y-62-33 (ATCC VR-375).

[0081] Bacterial antigens may be derived from, for example, *Staphylococcus aureus*, *Staphylococcus epidermis*, *Helicobacter pylori*, *Streptococcus bovis*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Listeria monocytogenes*, *Mycobacterium tuberculosis*, *Mycobacterium leprae*, *Corynebacterium diphtheriae*, *Borrelia burgdorferi*, *Bacillus anthracis*, *Bacillus cereus*, *Clostridium botulinum*, *Clostridium difficile*, *Salmonella typhi*, *Vibrio cholerae*, *Haemophilus influenzae*, *Bordetella pertussis*, *Yersinia pestis*, *Neisseria gonorrhoeae*, *Treponema pallidum*, *Mycoplasma sp.*, *Neisseria ransducer s*, *Legionella pneumophila*, *Rickettsia typhi*, *Chlamydia trachomatis*, and *Shigella dysenteriae*, *Vibrio cholera*(e.g., Cholera toxin subunit B as set forth in Genbank Accession Nos. U25679;

A09803; EF158842; X76391; AF390572; cholera toxin-coregulated pilus (TCP) as described in Wu *et al.*, *Infection and Immunity* Vol. 69(12):7695 (2001) and as set forth in Genbank Accession Nos. NC_002505 and AE004169); *Helicobacter pylori* (VacA as set forth in Genbank Accession Nos. AY848858; AF042737; AF042736; AF042735; AF042734; NC_000921; CagA as set forth in Genbank Accession Nos. AF043490; AF043489; AF043488; AF043487; NAP as set forth in Genbank Accession Nos. AF284121; AF284120; AF284119; AF284118; AF284117; AF284116; AB045143; AB045142; AF227081; AF227080; AF227079; AF227078; AF227077; AF227076; AF227075; AF227074; Hsp or catalase as set forth in Genbank Accession No. NC_000921; urease as set forth in Genbank Accession Nos. AM417610; AM417609; AM417608; AM417607; AM417606; AM417605; AM417604; AM417603; AM417602; AM417601; and AM417600; *E. coli* antigens as set forth in Genbank Accession Nos. NC_000913; U00096; NC_002655; BA000007; AE014075; including *E. coli* fimbrial antigens as set forth in Genbank Accession Nos. AB214865; AB214864; AB214863; AB214862; *E. coli* heat-labile enterotoxin as set forth in Genbank Accession Nos. X83966; V00275; X83966; J01646; V00275; M35581; M17873; M17874; K01995; M61015; M17894; M17101; K00433.

[0082] Parasite antigens may be derived from, for example, *Giardia lamblia*, *Leishmania sp.*, *Trypanosoma sp.*, *Trichomonas sp.*, *Plasmodium sp.* (e.g., *P. falciparum* surface protein antigens such as pfs25 sequences as set forth in Genbank Accession Nos. XM_001347551; X07802; AF193769; AF179423; AF154117; and AF030628, pfs28 sequences as set forth in Genbank Accession No. L25843, pfs45 sequences as set forth in Genbank Accession Nos. EF158081; EF158079; EF158078; EF158076; EF158075; and EF158085, pfs84, pfs 48/45 sequences as set forth in Genbank Accession Nos. AF356146; AF356145; AF356144; AF356143; AF356142; AF356141; AF356140; AF356139; AF356138; AF356137; AF356136; AF356135; AF356134; AF356133; AF356132; AF356131; AF356130; AF356129; AF356128; AF356127, pfs 230 sequences as set forth in Genbank Accession Nos. NC_000910; XM_001349564; AE001393; L22219; L08135; and AF269242, *P. vivax* antigens such as Pvs25 sequences as set forth in Genbank Accession Nos. DQ641509; DQ641508; DQ641507; AY639972; AY639971; AY639970; AY639969; AY639968; AY639967; AY639966; and AY639965; and Pvs28 sequences as set forth in Genbank Accession Nos. AB033364; AB033363; AB033362; AB033361; AB033360; AB033359; AB033358; AB033357; AB033356; B033355; AB033354; AB033353; AB033352; AB033351; AB033350; AB033349; AB033348; AB033347; AB033346; and AB033345), *Schistosoma sp.*, *Mycobacterium tuberculosis* (e.g., Ag85 sequences as set forth in Genbank

Accession Nos. AX253506; AX253504; AX253502; and AX211309; MPT64, ESAT-6, CFP10, R8307, MTB-32 MTB-39, CSP, LSA-1, LSA-3, EXP1, SSP-2, SALSA, STARP, GLURP, MSP-1, MSP-2, MSP-3, MSP-4, MSP-5, MSP-8, MSP-9, AMA-1, Type 1 integral membrane protein, RESA, EBA-175, and DBA sequences as set forth in Genbank Accession Nos. BX842572; BX842573; BX842574; BX842575; BX842576; BX842577; BX842578; BX842579; BX842580; BX842581; BX842582; BX842583; BX842584 and NC_000962, HSP65 sequences as set forth in Genbank Accession Nos. AY299175; AY299174; AY299144; AF547886; and AF547885).

[0083] Cancer antigens include, for example, antigens expressed, for example, in colon cancer, stomach cancer, pancreatic cancer, lung cancer, ovarian cancer, prostate cancer, breast cancer, skin cancer (*e.g.*, melanoma), leukemia, lymphoma, or myeloma, exemplary cancer antigens include, for example, HPV L1, HPV L2, HPV E1, HPV E2, placental alkaline phosphatase, AFP, BRCA1, Her2/neu, CA 15-3, CA 19-9, CA-125, CEA, Hcg, urokinase-type plasminogen activator (Upa), plasminogen activator inhibitor.

[0084] Fungal antigens may be derived from, for example, *Tinea pedis*, *Tinea corporis*, *Tinea cruris*, *Tinea unguium*, *Cladosporium carionii*, *Coccidioides immitis*, *Candida sp.*, *Aspergillus fumigatus*, and *Pneumocystis carinii*.

[0085] The nucleic acids encoding immunogenic polypeptides, are typically produced by recombinant DNA methods (see, *e.g.*, Ausubel, *et al.* ed. (2001) *Current Protocols in Molecular Biology*). For example, the DNA sequences encoding the immunogenic polypeptide can be assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, or amplified from cDNA using appropriate primers to provide a synthetic gene which is capable of being inserted in a recombinant expression vector (*i.e.*, a plasmid vector or a viral vector) and expressed in a recombinant transcriptional unit. Once the nucleic acid encoding an immunogenic polypeptide is produced, it may be inserted into a recombinant expression vector that is suitable for *in vivo* or *ex-vivo* expression.

[0086] Recombinant expression vectors contain a DNA sequence encoding an immunogenic polypeptide operably linked to suitable transcriptional or translational regulatory elements derived from mammalian or viral genes. Such regulatory elements include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. An origin of replication and a selectable marker to facilitate recognition of transformants may additionally be incorporated. The genes utilized in the recombinant expression vectors may be divided between more than one virus

such that the gene products are on two different vectors, and the vectors are used for co-transduction to provide all the gene products in trans. There may be reasons to divide up the gene products such as size limitations for insertions, or toxicity of the combined gene products to the virus produce cell-lines.

C. TLR Agonists

[0087] According to the methods of the invention, TLR agonists are used to enhance the immune response to the heterologous polypeptide. In some embodiments, TLR-3 agonists are used. In other embodiments, TLR 7/8 agonists are used. The TLR agonists described herein can be delivered simultaneously with the expression vector encoding an antigen of interest or delivered separately (*i.e.*, temporally or spatially) from the expression vector encoding an antigen of interest. For example, the expression vector may be administered via a non-parenteral route (*e.g.*, orally, intranasally, or mucosally), while the TLR-agonist is delivered by a parenteral route (*e.g.*, intramuscularly, intraperitoneally, or subcutaneously).

1. TLR-3 Agonists

[0088] In a preferred embodiment of the present invention, a TLR-3 agonist is used to stimulate immune recognition of an antigen of interest. TLR-3 agonists include, for example, short hairpin RNA, virally derived RNA, short segments of RNA that can form double-strands or short hairpin RNA, and short interfering RNA (siRNA). In one embodiment of the invention, the TLR-3 agonist is virally derived dsRNA, such as for example, a dsRNA derived from a Sindbis virus or dsRNA viral intermediates [Alexopoulou *et al*, *Nature* 413:732-8 (2001)]. In some embodiments, the TLR-3 agonists is a short hairpin RNA. Short hairpin RNA sequences typically comprise two complementary sequences joined by a linker sequence. The particular linker sequence is not a critical aspect of the invention. Any appropriate linker sequence can be used so long as it does not interfere with the binding of the two complementary sequences to form a dsRNA.

[0089] In some embodiments, the short hairpin RNA comprises a sequence set forth in SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11, or 12, a sequence with substantial identity to a sequence set forth in SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11, or 12, or a variant of a sequence set forth in SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11, or 12. In certain embodiments, dsRNA that is a TLR-3 agonist does not encode a particular polypeptide, but produces a pro-inflammatory cytokine (*e.g.* IL-6, IL-8, TNF-alpha, IFN-alpha, IFN-beta) when contacted with a responder cell (*e.g.*, a dendritic cell, a peripheral blood mononuclear cell, or a macrophage) *in vitro* or *in-vivo*. In

some cases, the nucleic acid encoding the TLR-3 agonist (*e.g.*, an expressed dsRNA) and the chimeric adenoviral vector comprising a nucleic acid encoding a heterologous antigen are administered in the same formulation. In other cases the nucleic acid encoding the TLR-3 agonist and the chimeric adenoviral vector comprising a nucleic acid encoding a heterologous polypeptide are administered in different formulations. When the nucleic acid encoding the TLR-3 agonist and the adenoviral vector comprising a nucleic acid encoding a heterologous antigen are administered in different formulations, their administration may be simultaneous or sequential. For example, the nucleic acid encoding the TLR-3 agonist may be administered first, followed by the chimeric adenoviral vector (*e.g.*, 1, 2, 4, 8, 12, 16, 20, or 24 hours, 2, 4, 6, 8, or 10 days later). Alternatively, the adenoviral vector may be administered first, followed by the nucleic acid encoding the TLR-3 agonist (*e.g.*, 1, 2, 4, 8, 12, 16, 20, or 24 hours, 2, 4, 6, 8, or 10 days later). In some embodiment, the nucleic acid encoding the TLR-3 agonist and the nucleic acid encoding the heterologous antigen are under the control of the same promoter. In other embodiments, the nucleic acid encoding the TLR-3 agonist and the nucleic acid encoding the heterologous antigen are under the control of different promoters.

[0090] Several chemically synthesized analogs to double-stranded RNA are commercially available. These include polyinosine-polycytidylic acid (poly I:C), polyadenylic:polyuridylic acid (poly A:U), and poly I:poly C. Antibodies (or cross-linking of antibodies) to TLR-3 can also lead to IFN- β or pro-inflammatory cytokine production [Matsumoto *et al*, *Biochem. Biophys. Res. Commun.* 24:1364 (2002), de Bouteiller *et al*, *J Biol. Chem.* 18:38133-45 (2005)]. Commercially available siRNA segments of any sequence can also be obtained through sources such as Invitrogen.

2. TLR7/8 Agonists

[0091] In some embodiments, the TLR agonists are TLR7/8 agonists. TLR7/8 ligands are typically single-stranded, virally derived RNA. Because the receptors are expressed in intracellular compartments such as the endosome, not all short segments of RNA will trigger the TLR7/8 signaling cascade because they need to reach the correct compartment. Some ligands that have been shown to trigger this through exogenous addition are polyuridylic acid, resiquimod, and imiquimod [Westwood, *et al*, *Vaccine* 24:1736-1745(2006)].

IV. Pharmaceutical Compositions

[0092] Pharmaceutical compositions comprising the vectors described herein may also contain other compounds, which may be biologically active or inactive. Polypeptides may, but need not, be conjugated to other macromolecules as described, for example, in U.S. Pat. Nos. 4,372,945 and 4,474,757. Pharmaceutical compositions may generally be used for prophylactic and therapeutic purposes. Pharmaceutical compositions may be composed of methods to protect against stomach degradation such that the administered chimeric adenoviral vector may reach the desired locations. For the oral environment, several of these are available including the Eudragit and the TimeClock release systems as well as other methods specifically designed for adenovirus [Lubeck *et al.*, *Proc Natl Acad Sci U S A*, 86(17), 6763-6767 (1989); Chourasia and Jain, *J Pharm Pharm Sci*, 6(1), 33-66 (2003)]. There are also several methods already described for microencapsulation of DNA and drugs for oral delivery (*see, e.g.*, U.S. Patent Publication No. 2004043952). In some embodiments, the Eudragit system will be used to deliver the chimeric adenoviral vector to the lower small intestine. However, delivery to other locations of the small intestine should also work.

[0093] As noted above, the chimeric adenoviral vectors on the invention may be delivered using any delivery systems known to those of ordinary skill in the art. Numerous gene delivery techniques are well known in the art, such as those described by Rolland (1998) *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, and references cited therein.

[0094] It will be apparent that an immunogenic composition may contain pharmaceutically acceptable salts of the polynucleotides encoding the heterologous polypeptides (*e.g.*, immunogenic polypeptides). Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts). Some particular examples of salts include phosphate buffered saline and saline for injection.

[0095] Any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention. Suitable carriers include, for example, water, saline, alcohol, a fat, a wax, a buffer, a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, or biodegradable microspheres (*e.g.*, polylactate polyglycolate). Suitable biodegradable microspheres are disclosed, for example, in U.S. Pat. Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883. The immunogenic polypeptide and/or carrier

virus may be encapsulated within the biodegradable microsphere or associated with the surface of the microsphere.

[0096] Such compositions may also comprise buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

[0097] In some embodiments of the present invention, the compositions further comprise an adjuvant. Suitable adjuvants include, for example, the lipids and non-lipid compounds, cholera toxin (CT), CT subunit B, CT derivative CTK63, *E. coli* heat labile enterotoxin (LT), LT derivative LTK63, $\text{Al}(\text{OH})_3$, and polyionic organic acids as described in *e.g.*, WO 04/020592, Anderson and Crowle, *Infect. Immun.* 31(1):413-418 (1981), Roterman *et al.*, *J. Physiol. Pharmacol.*, 44(3):213-32 (1993), Arora and Crowle, *J. Reticuloendothel.* 24(3):271-86 (1978), and Crowle and May, *Infect. Immun.* 38(3):932-7 (1982)). Suitable polyionic organic acids include for example, 6,6'-[3,3'-dimethyl[1,1'-biphenyl]-4,4'-diyl]bis(azo)bis[4-amino-5-hydroxy-1,3-naphthalene-disulfonic acid] (Evans Blue) and 3,3'-[1,1'-biphenyl]-4,4'-diylbis(azo)bis[4-amino-1-naphthalenesulfonic acid] (Congo Red). It will be appreciated by those of skill in the art that the polyionic organic acids may be used for any genetic vaccination method in conjunction with any type of administration.

[0098] Other suitable adjuvants include topical immunomodulators such as, members of the imidazoquinoline family such as, for example, imiquimod and resiquimod (see, *e.g.*, Hengge *et al.*, *Lancet Infect. Dis.* 1(3):189-98 (2001)). Expressed TLR-3 agonists (*e.g.*, dsRNA) and TLR-7 agonists (*e.g.*, ssRNA) could also be used with the invention

[0099] Additional suitable adjuvants are commercially available as, for example, additional alum-based adjuvants (*e.g.*, Alhydrogel, Rehydralgel, aluminum phosphate, Algammulin); oil based adjuvants (Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, Mich.), Specol, RIBI, TiterMax, Montanide ISA50 or Seppic MONTANIDE ISA 720); nonionic block copolymer-based adjuvants, cytokines (*e.g.*, GM-CSF or Flt3-ligand); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, N.J.); AS-2 (SmithKline Beecham, Philadelphia, Pa.); salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes;

biodegradable microspheres; monophosphoryl lipid A and Quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, are also suitable adjuvants. Hemocyanins (*e.g.*, keyhole limpet hemocyanin) and hemoerythrins may also be used in the invention. Polysaccharide adjuvants such as, for example, chitin, chitosan, and deacetylated chitin are also suitable as adjuvants. Other suitable adjuvants include muramyl dipeptide (MDP, N acetylmuramyl L alanyl D isoglutamine) bacterial peptidoglycans and their derivatives (*e.g.*, threonyl-MDP, and MTPPE). BCG and BCG cell wall skeleton (CWS) may also be used as adjuvants in the invention, with or without trehalose dimycolate. Trehalose dimycolate may be used itself (see, *e.g.*, U.S. Pat. No. 4,579,945). Detoxified endotoxins are also useful as adjuvants alone or in combination with other adjuvants (see, *e.g.*, U.S. Pat. Nos. 4,866,034; 4,435,386; 4,505,899; 4,436,727; 4,436,728; 4,505,900; and 4,520,019. The saponins QS21, QS17, QS7 are also useful as adjuvants (see, *e.g.*, U.S. Pat. No. 5,057,540; EP 0362 279; WO 96/33739; and WO 96/11711). Other suitable adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, Calif., United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2, SBAS-4 or SBAS-6 or variants thereof, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, Mont.), and RC-529 (Corixa, Hamilton, Mont.).

[0100] Superantigens are also contemplated for use as adjuvants in the present invention. Superantigens include Staphylococcus exoproteins, such as the α , β , γ and Δ enterotoxins from *S. aureus* and *S. epidermidis*, and the α , β , γ and Δ E. coli exotoxins. Common Staphylococcus enterotoxins are known as staphylococcal enterotoxin A (SEA) and staphylococcal enterotoxin B (SEB), with enterotoxins through E (SEE) being described (Rott *et al.*, 1992). *Streptococcus pyogenes* B (SEB), *Clostridium perfringens* enterotoxin (Bowness *et al.*, 1992), cytoplasmic membrane-associated protein (CAP) from *S. pyogenes* (Sato *et al.*, 1994) and toxic shock syndrome toxin 1 (TSST 1) from *S. aureus* (Schwab *et al.*, 1993) are further useful superantigens.

[0101] Within the pharmaceutical compositions provided herein, the adjuvant composition can be designed to induce, *e.g.*, an immune response predominantly of the Th1 or Th2 type. High levels of Th1-type cytokines (*e.g.*, IFN-gamma, TNF-alpha, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following oral delivery of a composition comprising an immunogenic polypeptide as provided herein, an immune response that includes Th1- and Th2-type responses will typically be elicited.

[0102] The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule or sponge that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (see, *e.g.*, Coombes *et al.* (1996) *Vaccine* 14:1429-1438). Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

[0103] Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), as well as polyacrylate, latex, starch, cellulose and dextran. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound (see, *e.g.*, WO 94/20078; WO 94/23701; and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

[0104] The pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

V. Therapeutic Uses of the Invention

[0105] One aspect of the present invention involves using the immunogenic compositions described herein to elicit an antigen specific immune response from a subject or patient with a disease such as, for example, a viral infection, bacterial infection, a parasitic infection, a fungal infection, or cancer. As used herein, a "subject" or a "patient" refers to any warm-blooded animal, such as, for example, a rodent, a feline, a canine, or a primate, preferably a human. The immunogenic compositions may be used to treat at any stage of the disease, *i.e.*, at the pre-cancer, cancer, or metastatic stages, or to prevent disease. For example, the compositions described herein may be used to treat a viral disease such as HIV or hepatitis or for prevention or treatment of cancer. Within such methods, pharmaceutical compositions

are typically administered to a patient. The patient may or may not be afflicted with the disease or disorder (e.g., a viral infection, a bacterial infection, or cancer). Accordingly, the above pharmaceutical compositions may be used to prevent the development of a disease or disorder (e.g., a viral infection, a bacterial infection, or cancer) or to treat a patient afflicted with the disease or disorder (e.g., a viral infection, a bacterial infection, or cancer). The disease or disorder may be diagnosed using criteria generally accepted in the art. For example, viral infection may be diagnosed by the measurement of viral titer in a sample from the patient, bacterial infection may be diagnosed by detecting the bacteria in a sample from the patient, and cancer may be diagnosed by detecting the presence of a malignant tumor. Pharmaceutical compositions may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs.

[0106] Immunotherapy is typically active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against, e.g., tumors or bacterially or virally infected cells, with the administration of immune response-modifying agents (compositions comprising nucleic acids encoding immunogenic polypeptides as provided herein).

[0107] Frequency of administration of the prophylactic or therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. Often between 1 and 10 doses may be administered over a 52 week period. Typically 3 doses are administered, at intervals of 1 month, more typically, 2-3 doses are administered every 2-3 months. It is possible that the intervals will be more like once a year for certain therapies. Booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients and particular diseases and disorders. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting, e.g., an anti-tumor, an anti-viral, or an antibacterial, immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic T cells capable of killing, e.g., the patient's tumor cells, the patient's virally infected cells, or the patient's bacterially infected cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (e.g., more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. Typically, the amount of the viral titers will be between 1.0×10^4 pfu/animal and 1.0×10^{15}

pfu/animal. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.01 ml to about 10 ml, more typically from about 0.025 to about 7.5 ml, most typically from about 0.05 to about 5 ml. Those of skill in the art will appreciate that the dose size may be adjusted based on the particular patient or the particular disease or disorder being treated. For oral administration, the chimeric adenoviral vector can conveniently be formulated in a pill.

[0108] In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays described above, which may be performed using samples obtained from a patient before and after treatment.

[0109] For example, detection of immunocomplexes formed between immunogenic polypeptides and antibodies in body fluid which are specific for immunogenic polypeptides may be used to monitor the effectiveness of therapy, which involves a particular immunogenic polypeptide, for a disease or disorder in which the immunogenic polypeptide is associated. Samples of body fluid taken from an individual prior to and subsequent to initiation of therapy may be analyzed for the immunocomplexes by the methodologies described above. Briefly, the number of immunocomplexes detected in both samples are compared. A substantial change in the number of immunocomplexes in the second sample (post-therapy initiation) relative to the first sample (pre-therapy) reflects successful therapy.

A. Administration of the Compositions of the Present Invention

[0110] According to the methods of the present invention, a composition comprising the chimeric adenoviral vector is administered by any non-parenteral route (*e.g.*, orally, intranasally, or mucosally via, for example, the vagina, lungs, salivary glands, nasal cavities, small intestine, colon, rectum, tonsils, or Peyer's patches). The composition may be administered alone or with an adjuvant as described above. In some embodiments, the adjuvants are encoded by a nucleic acid sequence (*e.g.*, a nucleic acid encoding IL-2, GM-CSF, IL-12, or bacterial flagellin). In some embodiments of the present invention, the adjuvant is administered at the same time as the composition. In other embodiments of the present invention, the adjuvant is administered after the composition, *e.g.*, 6, 12, 18, 24, 36, 48, 60, or 72 hours after administration of the composition.

B. Detection of an Immune Response to Antigens of Interest

[0111] An immune response to the heterologous polypeptide can be detected using any means known in the art including, for example detecting specific activation of CD4⁺ or CD8⁺ T cells or by detecting the presence of antibodies that specifically bind to the polypeptide.

[0112] Specific activation of CD4⁺ or CD8⁺ T cells associated with a mucosal, humoral, or cell-mediated immune response may be detected in a variety of ways. Methods for detecting specific T cell activation include, but are not limited to, detecting the proliferation of T cells, the production of cytokines (*e.g.*, lymphokines), or the generation of cytolytic activity (*i.e.*, generation of cytotoxic T cells specific for the immunogenic polypeptide). For CD4⁺ T cells, a preferred method for detecting specific T cell activation is the detection of the proliferation of T cells. For CD8⁺ T cells, a preferred method for detecting specific T cell activation is the detection of the generation of cytolytic activity using ⁵¹Cr release assays (see, *e.g.*, Brossart and Bevan, *Blood* 90(4): 1594-1599 (1997) and Lenz *et al.*, *J. Exp. Med.* 192(8):1135-1142 (2000)).

[0113] Detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring the rate of DNA synthesis. T cells which have been stimulated to proliferate exhibit an increased rate of DNA synthesis. A typical way to measure the rate of DNA synthesis is, for example, by pulse-labeling cultures of T cells with tritiated thymidine, a nucleoside precursor which is incorporated into newly synthesized DNA. The amount of tritiated thymidine incorporated can be determined using a liquid scintillation spectrophotometer. Other ways to detect T cell proliferation include measuring increases in interleukin-2 (IL-2) production, Ca²⁺ flux, or dye uptake, such as 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium. Alternatively, synthesis of lymphokines (*e.g.*, interferon-gamma) can be measured or the relative number of T cells that can respond to the immunogenic polypeptide may be quantified.

[0114] Antibody immune responses (aka Humoral immune responses or B cell responses), including mucosal antibody responses can be detected using immunoassays known in the art [Tucker *et al.*, *Mol Therapy*, 8, 392-399 (2003); Tucker *et al.*, *Vaccine*, 22, 2500-2504 (2004)]. Suitable immunoassays include the double monoclonal antibody sandwich immunoassay technique of David *et al.* (U.S. Pat. No. 4,376,110); monoclonal-polyclonal antibody sandwich assays (Wide *et al.*, in Kirkham and Hunter, eds., *Radioimmunoassay Methods*, E. and S. Livingstone, Edinburgh (1970)); the "western blot" method of Gordon *et al.* (U.S. Pat. No. 4,452,901); immunoprecipitation of labeled ligand (Brown *et al.* (1980) *J.*

Biol. Chem. 255:4980-4983); enzyme-linked immunosorbent assays (ELISA) as described, for example, by Raines *et al.* (1982) *J. Biol. Chem.* 257:5154-5160; immunocytochemical techniques, including the use of fluorochromes (Brooks *et al.* (1980) *Clin. Exp. Immunol.* 39:477); and neutralization of activity (Bowen-Pope *et al.* (1984) *Proc. Natl. Acad. Sci. USA* 81:2396-2400). In addition to the immunoassays described above, a number of other immunoassays are available, including those described in U.S. Pat. Nos. 3,817,827; 3,850,752; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; and 4,098,876.

EXAMPLES

[0115] The following examples are intended to illustrate, but not to limit the present invention.

Example 1: Construction of a Chimeric Adenoviral Vector (DSI)

[0116] To demonstrate that TLR-3 agonists can improve adaptive immune responses to expressed antigens of interest, several different chimeric adenoviral vectors were constructed that comprise nucleic acid sequences encoding several different antigens of interest. In this example, the nucleic acid encoding gp120 (from the NIH AIDS Reagent and Reference Reagent Program) was placed under control of a CMV promoter with a small intron just upstream of the start codon in the shuttle vector (pShuttle, Qbiogene). A poly A tail from bGH was placed downstream of the nucleic acid encoding gp120. The vector sequence is set forth in SEQ ID NO: 1. Homologous recombination with the vector pAd (Qbiogene) was performed to generate a vector capable of producing recombinant Ad (E1/E3 deleted) that contained the nucleic acid encoding gp120. DSI was generated by transfecting the new pAd-CMV-gp120 expression construct into 293 cells. Titers were measured by standard methods.

Example 2: DSI (vector plus TLR-3 agonist) is superior to standard rAd5 for inducing an antigen specific immune response

[0117] To determine whether the addition of TLR-3 agonist could improve adaptive immune responses, 10×10^7 PFU of either rAd-CMV-gp120 plus 5 ug/ml poly I:C (DS1) or rAd-CMV-gp120 alone (rAd5) were administered to animals by oral gavage on weeks 0 and 3. Both vectors express HIV gp120 under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to gp120 were measured in the plasma 3 and 6 weeks after the initial administration by anti-gp120 IgG ELISA as described in Tucker, *et al.*, *Mol Ther* 8:392 (2004)). As shown in Figure 1, DS1 performed significantly better than rAd5 in eliciting antibody responses to the protein gp120 both at 3 and 6 weeks post

initial oral administration. In particular, the average antibody titer to gp120 was 100 fold better with the DS1 group than with the rAd5 group at week 6. It also appears that the DS1 group was boosted by readministration at week 4 in that the average titer increased greater than 20 fold between weeks 3 and 6 whereas the rAd5 group showed only a slight increase in mean antibody titer. The results demonstrate that the addition of a TLR-3 agonist can greatly improve Ad5 mediated antibody responses to antigens of interest following oral administration of a chimeric adenoviral vector comprising a nucleic acid encoding the antigen of interest. As a positive control for the assay, sera from an animal injected subcutaneously with gp120 plus Complete Freund's Adjuvant was also measured in the anti-gp120 ELISA at week 3. Untreated animals and animals administered the dsRNA analog alone (dsRNA) served as negative and background controls respectively for the ELISA. Each group contained 6 animals.

Example 3: Construction of a Second Chimeric Adenoviral Vector (DS1b) and a Third Chimeric Adenoviral Vector (DS1c)

[0118] A nucleic acid encoding green fluorescent protein (GFP) was inserted into pShuttle-CMV (Qbiogene) using standard restriction enzyme digests. The plasmid pShuttleCMV-GFP was combined by homologous recombination with the vector pAd (Qbiogene) as described before in order to generate a vector capable of producing recombinant Ad (E1/E3 deleted) comprising a nucleic acid sequence encoding GFP. A nucleic acid encoding hemagglutinin (HA) from influenza A/PR/8/34 was cloned and placed in the pShuttle-CMV vector (Qbiogene) (SEQ ID NO: 13). The plasmid pShuttleCMV-HA (PR/8) was combined by homologous recombination with the vector pAd (Qbiogene) as described before in order to generate a vector capable of producing recombinant Ad (E1/E3 deleted) comprising a nucleic acid sequence encoding HA. Recombinant Ad was generated by transfecting the new pAd-CMV-GFP and pAd-CMV-HA expression construct into 293 cells. Titers were measured by standard methods.

Example 4: DS1b (Ad-CMV-GFP plus TLR-3 agonist) and DS1c (Ad-CMV-HA plus TLR-3 agonist) is superior to standard rAd5 for inducing an antigen specific immune response.

[0119] 1.0×10^7 PFU of either Ad-CMV-GFP plus 5 ug/ml poly I:C (DS1b) or Ad-CMV-GFP (rAd5) were administered to animals by oral gavage on week 0. Both viruses express the GFP under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to GFP were measured in the plasma 3 weeks after the initial virus

administration by anti-GFP IgG ELISA. As shown in in Figure 2, the DS1b group performed significantly better than rAd5 in eliciting antibody responses to the protein GFP at 3 weeks post initial oral administration.

[0120] The CD8⁺T cell responses to GFP were measured by tetramer staining of splenocytes. Animals were vaccinated on weeks 0, 4, 8 and spleens were harvested on week 10. The splenocytes were stained with CD8-FITC and the tetramer which recognizes the immunodominant epitope to GFP in Balb/c mice. Results show that oral administration of the DS1b vector was statistically better than rAd alone in inducing tetramer positive CD8 cells (Figure 2b).

[0121] 1.0×10^7 PFU of either Ad-CMV-HA plus 5 ug/ml poly I:C (DS1c) or Ad-CMV-HA (rAd5) were administered to animals by oral gavage on week 0. Both viruses express HA under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to HA were measured in the plasma 3 weeks after the initial virus administration by anti-PR8/34 IgG ELISA. The procedure for measuring antibody responses is similar to that described before with the exception that the ELISA plates were coated with 5 ug/ml of whole A/PR8/34 lysate (Advanced Biotechnology Incorporated, Gaithersburg, MD). As shown in Figure 2C, the DS1c group performed significantly better (approximate 100 fold better) than rAd5 in eliciting antibody responses to influenza at 3 weeks post initial oral administration. The results of these studies also demonstrate that the approach of using TLR-3 agonist along with a chimeric recombinant adnoviral vector can be generally applied to mulitple different heterologous antigens, with a 100 fold improvement in antibody titer.

Example 5 Non-Parenteral Routes of Delivery Are Superior to Parenteral Routes

[0122] Intramuscular delivery was tested by directly injecting 1.0×10^7 pfu of pAd-CMV-gp120 (DS10 +/- poly I:C at 5 ug/ml into the quadriceps of animals. Plasma serum IgG titers to GFP were measured as described before. Each group contained 6 animals. As shown in Figure 3A, significant antibody titers to gp120 were observed at 3 weeks post administration in the group with TLR-3 agonist (i.m. rAd+PI). (Figure 3a).

[0123] Intranasal administration was tested by administering 20 ul of 1.1×10^6 pfu of DS1c +/- 5 ug/ml of poly I:C into the nasal cavity of mice. The mice were lightly anesthetized with isoflurane before administering the virus formulated in sterile saline. The results show that the rAd-CMV-HA plus poly I:C (DS1c) had slightly higher antibody titers compared to animals given the standard rAd-CMV-HA. Results are plotted as individual animals for the

DS1c (N=6) and the rAd (N=5) groups. Untreated animals (N=4) are used for negative controls.

Example 6. Construction of an expressed TLR3 agonist.

[0124] A short 45 bp segment of DNA was synthesized by ordering of DNA oligos that when annealed together formed a 45 bp segment designed to make a hairpin of double-stranded RNA (GAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTC) (SEQ ID NO:10). This short segment (called luc1) was cloned into the plasmid pSK- containing the human beta actin promoter and a BGH poly A tail. This plasmid is called pSk-luc1.

Example 7. The pSK-luc 1 functions in dendritic cell cultures like poly I:C, the effects of poly I:C and rAd are additive.

[0125] To determine whether the expressed TLR-3 agonist of Example 6 above could function as an inducer of pro-inflammatory cytokines and dendritic cell maturation like the TLR-3 ligand poly I:C, an expressed dsRNA TLR-3 agonist was tested in dendritic cell cultures. Bone marrow from the femurs of Balb/c mice were cultured with flt-3 ligand (200 ng/ml), 5% serum, in DMEM media in order to make primary dendritic cell cultures. Five days after primary bone marrow cultures were set-up, 293 cells were transfected with either pSk-luc1, pSK-beta2 (a long segment of beta galactosidase that forms a 200 bp hairpin), or pcDNA3 (empty expression vector). On day 6, the transfected cells were treated by UV irradiation (20 seconds at 40 kJ/cm²) to cause apoptosis and these cells were given to the dendritic cells. Either poly I:C (1 ug/ml), rAd (1 pfu/cell), rAd + poly I:C, pSK-luc1 transfected cells, pSK-beta2 transfected cells, or pcDNA3 transfected cells were given to the dendritic cells and cultured overnight. As shown in Figure 4A, pSK-luc1 transfected cells can significantly improve dendritic cell activation as measured by the mouse IL-6 ELISA. The results of this experiment also show that the combination of rAd plus TLR3 ligand (poly I:C) together can greatly improve dendritic cell activity.

[0126] Additional ligands were also tested. The TLR-3 agonist set forth in SEQ ID NO: 11 (m1) also forms a dsRNA hairpin of approximately the same size as luc1. These were made by overlapping oligonucleotides and annealing them together before cloning into the pSK-vector under control of the human beta actin promoter. The vectors were transfected into 293 cells and given to primary dendritic cells as described before. As shown in Figure 4B, these additional ligands can activate dendritic cells similar to that of the ligand luc1 (Figure 4B).

Example 8. Construction of a Fourth Chimeric Adenoviral Vector (DS2) and rapid cloning vectors (DS2beta-luc and DS2C-luc).

[0127] A nucleic acid encoding gp120 (from the NIH AIDS Research and Reference Reagent Program)) was placed under control of a CMV promoter with a small intron just upstream of the start codon in the shuttle vector (pShuttleCMV, Qbiogene). A poly A tail from bGH was placed downstream of the nucleic acid encoding gp120. The dsRNA TLR-3 agonist luc1 under the control of the human beta actin promoter and poly A (described in example 5 above) was inserted into the gp120 pShuttle vector such that both the nucleic acid encoding gp120 and the nucleic acid encoding TLR-3 agonist were contained in a single vector under the control of two separate promoters. The orientation of the expression of the nucleic acid encoding the antigen of interest and the expression of the TLR-3 agonist is illustrated in Figure 5.

[0128] Two generic shuttle vectors called DS2beta-luc (SEQ ID NO: 14) and DS2C-luc (SEQ ID NO: 15) were also constructed such that a nucleic acid encoding any antigen of interest could be inserted under the CMV promoter and either the human beta actin promoter or the CMV promoter is used to drive expression of a dsRNA TLR-3 agonist. In particular, the vector DS2C-luc has a unique Kpn 1 site that a nucleic acid encoding an antigen of interest can easily be cloned into. The purpose of these vectors is to make subsequent vector construction much easier because a nucleic acid encoding any antigen of interest could be inserted into the cloning site to rapidly manufacture a vector capable of eliciting antibody and T cells responses against the antigen of interest. Homologous recombination of DS2 with the vector pAd (Qbiogene) was performed as before in order to generate a vector capable of producing recombinant Ad (E1/E3 deleted) that contained a nucleic acid encoding GFP and a nucleic acid encoding the dsRNA LR-3 agonist luc1. Recombinant Ad was generated by transfecting the new pAd-betaactin-luc1-CMV-gp120 expression construct into 293 cells. Titers were measured by standard methods.

Example 9: Induction of an antigen specific immune response following oral delivery of DS2.

[0129] 1.0×10^7 PFU of either pAd-CMV-gp120 plus the TLR-3 agonist luc1 (DS2) or pAd-CMV-gp120 (rAd5) were administered to animals by oral gavage on week. Both viruses express the gp120 under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to gp120 were measured in the plasma 3 weeks after virus administration by anti-gp120 IgG ELISA. The ELISA protocol has been described before (Tucker, et al, Mol Therapy 8:392 (2004)). . Results demonstrate that DS2 can induce

approximately a 2 log improvement in antibody titer to gp120, the heterologous antigen used in the experiment. The DS2 vector comprises a nucleic acid sequence encoding expressing gp120 and a nucleic acid sequence expressing a dsRNA TLR-3 agonist. As a positive control for the assay, sera from two animals injected subcutaneously with 10 micrograms gp120 protein plus Complete Freund's Adjuvant was also measured in the anti-gp120 ELISA. Untreated animals served as negative controls for the ELISA. Each group contained 6 animals. The results are illustrated in Figure 6.

Example 10: Induction of an antigen specific immune response following oral delivery of DS3.

[0130] 1.0×10^7 PFU of either pAd-CMV-influenza HA (from A/PR/8/34) plus the TLR7/8 ligand polyuridylic acid (DS3) or pAd-CMV-HA (rAd5) were administered to animals by oral gavage on week 0. Both viruses express influenza HA under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to HA were measured in the plasma 3 weeks after virus administration by anti-influenza HA IgG ELISA. Each group contained 6 animals. The results are illustrated in Figure 7.

Example 11: Construction of a fifth, sixth, and seventh chimeric adenoviral vector (DS2b, DS2b-for, and ND1.1 214).

[0131] The gene influenza HA (A/Indo/5/2005) was synthesized by CelTek (Nashville, TN) and placed into the vector pShuttleCMV (Qbiogene) which has a CMV promoter with a small intron just upstream of the start codon in the shuttle vector. The luc1 DNA with human beta actin promoter and poly A (described in example 5) were placed into the vector downstream of the antigen, in the orientation shown in figure 5 for DS2b. The sequence of luc1 is

(GAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTC) (SEQ ID NO:10) and the completed pShuttle vector is set forth in SEQ ID NO: 6. An alternative orientation of luc1 with promoter in a shuttle vector is described as SEQ ID NO: 7 and is designated DS2b-for. We have also constructed another pShuttle vector (called DS2bC-HA) (SEQ ID NO: 16) that comprises two separate CMV promoters driving expression of the TLR-3 agonist luc1 and influenza HA described above. Homologous recombination with the vector pAd (Qbiogene) was performed as before in order to generate vectors capable of producing recombinant Ad (E1/E3 deleted) that contained the nucleic acid encoding HA and the TLR-3 agonist luc1 under separate promoters. Recombinant Ad was generated by

transfecting the new pAd- constructs into 293 cells. Titers were measured by standard methods. The completed pAd vector containing DS2C-luc was named ND1.1 214 and deposited in the ATCC patent depository on Feb 22, 2007 (Manassus, VA). The nucleic acid sequence of this chimeric adenoviral vector is set forth in in SEQ ID NO: 17. The nucleic acid encoding the heterologous antigen is in bold text and is flanked by a Cla I recognition site on the 5' end and a Not 1 recognition site on the 3' end. The nucleic acid sequence encoding the TLR-3 agonists is in italic, with the linker sequence in bold. A nucleic acid sequence encoding any antigen of interest and a nucleic acid sequence encoding any suitable expressed TLR-3 agonist can be inserted into the chimeric adenoviral vector.

Example 12: Induction of an antigen specific immune response following oral delivery of DS2b.

[0132] 1.0×10^7 PFU of either pAd-CMV-HA plus the TLR-3 agonists luc1 in the reverse orientation (DS2b) or forward orientation (DS2b-for), or pAd-CMV-HA (rAd5) were administered to animals by oral gavage on week 0. These viruses express the antigen influenza HA under control of the CMV promoter and use recombinant E1/E3 deleted adenovirus type 5. Antibody titers to HA were measured in the plasma 3 weeks after virus administration by anti-HA IgG ELISA. Results demonstrate that the DS2b vector elicits an antibody responses to the protein HA greater than the standard rAd vector (rAd5). The DS2b vector contains rAd5 expressing HA as well as expresses a toll-like receptor 3 (TLR3) agonist, a hairpin of double-stranded RNA, demonstrating that the use of the encoded dsRNA ligand can improve adaptive immune responses to antigens of interest. As shown in Figure 8A and Figure 6, expressed dsRNA can improve adaptive immune responses to multiple different heterologous antigens. Untreated animals served as negative control for the ELISA. Each group contained 6 animals.

[0133] Vectors in the opposite orientation (DS2for) were examined for antibody responses following either oral or intramuscular administration of 1.0×10^7 pfu virus per animal at 0 and 5 weeks. Antibody responses to HA were measured at 4 and 7 weeks post initial administration. As shown in Figure 8B, the opposite orientation vector can also induce substantial antibody responses to heterologous antigens. The DS1b and DS1bfor vectors induced similar responses to HA at the 4 week time point. Significantly, the effect of boosting of the antibody response was demonstrated with the DS1bfor vector and showed that multiple doses could be used to increase antibody responses to the heterologous antigen.

[0134] Another example of potential of the chimeric adenoviral vector approach was demonstrated as well. The vector ND1.1 214 was given to animals by oral (1.0×10^7 pfu) or intranasal administration (3×10^6 pfu) and the antibody responses to the heterologous antigen were measured at week 3. As shown in Figure 8C, substantial antibody responses to HA were measured following oral administration, well beyond the typical values from a single oral administration of rAd vector.

[0135] All publications, patent publications, patents, and Genbank Accession Nos. applications cited in this specification are herein incorporated by reference in their entirety for all purposes as if each individual publication, patent publication, or patent were specifically and individually indicated to be incorporated by reference.

INFORMAL SEQUENCE LISTING

SEQ ID NO:1 (DS1)

taacatcatc aataatatac cttatTTTTgg attgaagcca atatgataat
 gaggggggtg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
 caggaagtga caattttcgc gcggttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtga
 tctgaataattttgtgttactcatagcgcgtaatactggtaccgcggccgcctcgagtctag
 agatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAAACGCCAGGGTTTCCC
 AGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTG
 GGTACTGGCCAcagagcttggccattgcatacgttgtatccatatcataatatgtacattt
 atattggctcatgtccaacattaccgccatgttgacattgattattgactagttattaatag
 taatcaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttac
 ggtaaatggccgcctggctgaccgcccacgacccccgccattgacgtcaataatgacgt
 atgttcccataagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacgg
 taaactgcccacttggcagtacatcaagtgtatcatatgccagtagccccctattgacgt
 caatgacggtaaattggccgcctggcattatgccagtagacacattatgggactttcccta
 cttggcagtacatctacgtattagtcacgtcattaccatgggtgatgcggttttggcagtag
 atcaatgggctggtatagcgggttgactcacggggatttccaagtctccacccattgacgt
 caatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactccg
 cccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatataagcagagctcgt
 ttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacctccatagaagaca
 ccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgaggccctgggcagg
 ttggtatcaagggttacaagacaggtttaaggagaccaatagaaactgggcatgtggagacag
 agaagactcttgggtttctgataggcactgactctctctgcctattggtctattttccacc
 cttaggctgctgggtctgagcctagGAGATCTCTCGAGGTCGACGGTATAGCTTCTAGAGATC
 CCTCGACCTCGAGATCCATTGTGCTCTAAAGGAGATACCCGGCCAGACACCCTCACCTGCGG
 TGCCAGCTGCCAGGCTGAGGCAAGAGAAGGCCAGAAACCATGCCATGGGGTCTCTGCAA
 CCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTGCCTTCCGTGCTAGCTGTGGAGAAGCT
 GTGGGTGACTGTATACTATGGGGTGCCTGTGTGGAAGGAGGCCACCACCACCCTGTTCTGTG
 CCTCTGATGCCAAGGCCTATGACACTGAGGTCCACAATGTCTGGGCCACCCATGCCTGTGTG
 CCCACTGACCCCAACCCTCAGGAGGTGGTGTGCTGGAGAATGTGACTGAGCACTTCAACATGTG
 GAAGAACAACATGGTGGAGCAGATGCAGGAGGACATCATCAGCCTGTGGGACCAGAGCCTGA
 AGCCCTGTGTGAAGCTGACCCCCCTGTGTGTGACCCTGAACTGCAAGGATGTGAATGCCACC
 AACACCACCAATGACTCTGAGGGCACTATGGAGAGGGGTGAGATCAAGAACTGCAGCTTCAA
 CATCACCACCAGCATCAGGGATGAGGTGCAGAAGGAGTATGCCCTGTTCTACAAGCTGGATG
 TGGTGCCCATTGACAACAACAACACCAGCTACAGGCTGATCAGCTGTGACACCTCTGTGATC
 ACCCAGGCCTGCCCCAAGATCAGCTTTGAGCCCATCCCCATCCACTACTGTGCCCTGCTGG
 CTTTGCCATCCTGAAGTGCAATGACAAGACCTTCAATGGCAAAGGCCCTTGCAAGAATGTGA
 GCACTGTGCAGTGCACCTCATGGCATCAGGCCTGTGGTGTGACACCCAGCTGCTGCTGAATGGC
 AGCCTGGCTGAGGAGGAGGTGGTGTGATCAGGTCTGACAACCTTACCAACAATGCCAAGACCAT
 CATTGTGCAGC
 TGAAGGAGTCTGTGGAGATCAACTGCACCAGGCCCAACAACAACACCAGGAAGAGCATTTCAC
 ATTGGCCCTGGCAGGGCCTTCTACACCACTGGGGAGATCATTGGGGACATCAGGCAGG
 CCCACTGCAACATCAGCAGGGCCAAGTGGAATGACACCCTGAAGCAGATTGTGATCAAGCTG
 AGGGAGCAGTTTGAGAACAAAGACCATTGTGTTCAATCACAGCTCTGGTGGTGTGATCCTGAGAT
 TGTGATGCACAGCTTCAACTGTGGTGGTGTGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCA
 ACAGCACCTGGAACAACAACACTGAGGGCAGCAACAACACTGAGGGCAACACCATCACCCCTG

CCTTGCAGGATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTATGCTCC
 TCCCATCAGGGGCCAGATCAGGTGCAGCAGCAACATCACTGGCCTGCTGCTGACCAGGGATG
 GTGGCATCAATGAGAATGGCACTGAGATTTTCAGGCCTGGTGGTGGGGACATGAGGGGACAAC
 TGGAGGTCTGAGCTGTACAAGTACAAGGTGGTGAAGATTGAGCCCCCTTGGTGTGGCTCCCAC
 CAAGGCTAAGCGCAGGGTGGTGCAGAGGGAGAAGcgcgctgtgggctgaggatccCGAGGGT
 GAGTGCTCCTGCCTGGACGCATCCCGGCTATGCAGCCCCAGTCCAGGGCAGCAAGGCAGGCC
 CCGTCTGCCTCTTACCCGGAGCCTCTGCCCCGCCCACTCATGCTCAGGGAGAGGGTCTTCT
 GGCTTTTTTCCCAGGCTCTGGGCAGGCACAGGCTAGGTGCCCTAACCCAGGGCCTGCACACA
 AAGGGGCAGGTGCTGGGCTCAGACCTGCCAAGAGCCATATCCGGGAGGACCCTGCCCTGAC
 CTAAGCCCCACCCCAAAGGCCAACTCTCCACTCCCTCAGCTCGGACACCTTCTCTCTCCCA
 GATTCCAGTAACTCCCAATCTTCTCTCTGCAGAGCCCAAATCTTGTGACAAAACCTCACACAT
 GCCCACCGTGCCAGGTAAGCCAGCCCAGGCCTCGCCCTCCAGCTCAAGGCGGGACAGGTGC
 CCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACACGTCCACCTCCATCT
 CTTCTCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAAACCCAAG
 GACACCCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACGA
 AGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAA
 AGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACCGTCTGCAC
 CAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCC
 CATCGAGAAAACCATCTCCAAAGCCAAAGGTGGGACCCGTGGGGTGCGAGGGGCCACATGGAC
 AGAGGCCGGCTCGGCCACCCCTCTGCCCTGAGAGTGACCGCTGTACCAACCTCTGTCTACA
 GGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAA
 CCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCCGTGGAGTGGG
 AGAGCAATGGGCAGCCGGAGAACAATAACAAGACCACGCCTCCCGTGCTGGACTCCGACGGC
 TCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTT
 CTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGT
 CTCGGGTAAATGAGTGCGACGGCCGAGGTAAAGCCAGCCCAGGCCTCGCCCTCCAGCTCAA
 GCGGGACAGGTGCCCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACAC
 GTCCACCTCCATCTCTTCTCAGGTCTGCCCGGGTGGCATCCCTGTGACCCCTCCCAGTGC
 CTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCTAATAAAATTAAGT
 TGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATG
 GAGCAAGGGGCCCAAGTTAACTTGTATTATGCAGCTTATAATGGTTACAAATAAAGCAATAG
 CATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC
 TCATCAATGTATCTTATCATGTCTGgatctggg cgtgggttaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctgtt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac

541 gcgcagtccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggtcg

601 ccccgctctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgcccgtt

661 ggagactgca gcctccgccc cgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac

721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga

781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt

841 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc

901 caatgcgggt taaaacataa ataaaaaacc agactctgtt
 tggatttggga tcaagcaagt

961 gtcttgctgt ctttatattag gggttttgcg cgcgcggtag
 gcccgggacc agcgggtctcg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcgccct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgtcat aggccatttt taciaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gcccaggggc gtagttaccc
 tcacagattt gcattttcca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca gggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggg cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttccgcggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctggtgctga agcgtgcgcg gtcttcgccc tgcgcgctcg
 ccaggtagca ttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcggtg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgcgcacg aggggcagtg cagacttttg agggcgtaga
 gcttggggcg gagaaatacc

2461 gattccgggg agtaggcac cgccgcgcag gccccgcaga
 cggctctcgca ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accagggttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttaa aacgaattca
 atagcttggt gcatgggagg
 2701 cgatataaaa tgcaagggtgc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaaccatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggagggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aaggggccaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca ccagaaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatacgtc acttccggtt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctagctc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc
 3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgtcttccg cttcctcgct
 3781 cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaag
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg

3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaaccc ccggttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggga agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggat gtaggcgggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgt cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtgggtc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcgggtt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc

5461 tgatgcgcgc gtgttccggc tgtcagcgca ggggcgccc
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcgtt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgcccagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgccc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatgggtg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcacgcgc ttctatcgcc ttcttgacga gttcttctga
 attttgttaa aatttttgtt
 6181 aaatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgaa
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
 taaagctcta aatcggaacc
 6421 ctaaagggag ccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID:2 (Vector Sequence DS2)

taacatcatc aataatatac cttatttttg attgaagcca atatgataat
 gagggggtg agtttgtgac gtggcgccgg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
 caggaagtga caattttcgc gcggttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtga tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTG**GCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGC

CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCAcagagccttgcccattgcatatcggttgatccatatcata
 atatgtacattttatattgggtcatgtccaacattaccgccatggttgacattgattattgact
 agttattaatagtaataattacgggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggcccgctggtgaccgcccacgacccccgcccattgacgt
 caataatgacgtatggtcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaaactgccacttggcagtacatcaagtgtatcatatgccaagtacgcc
 ccctattgacgtcaatgacggtaaatggcccgctggcattatgccagtacatgacctat
 gggacttttcctacttggcagtacatcta**cgtaattagtcacogctattaccat**ggtgatgcgg
 ttttggcagtacatcaatgggcgtggatagcggtttgactcacggggattttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
 gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtgggagacagagaagactccttggttttctgataggcactgactctctctgcctattggtc
 tattttcccacccttaggctgctggtctgag**cctagGAGATCTCTCGAGGTGACGGTATCG**
ATGGGTACAGCTTCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTCTAAAGGAGATACCC
 GGCCAGACACCCTCACCTGCGGTGCCAGCTGCCAGGCTGAGGCAAGAGAAGGCCAGAAAC
 CATGCCCATGGGGTCTCTGCAACCGCTGGCCACCTTGTACCTGCTGGGGATGCTGGTTCGCTT
 CCGTGCTAGCTGTGGAGAAGCTGTGGGTGACTGTATACTATGGGGTGCCTGTGTGGAAGGAG
 GCCACCACCACCCTGTTCTGTGCCTCTGATGCCAAGGCCTATGACACTGAGGTCCACAATGT
 CTGGGCCACCCATGCCTGTGTGCCACTGACCCCAACCCTCAGGAGGTGGTGTGGAAGATG
 TGAAGGAGTCTGTGGAGATCAACTGCACCAGGCCCAACAACAACACCAGGAAGAGCATTCAC
 AGCCTGTGGGACCAGAGCCTGAAGCCCTGTGTGAAGCTGACCCCCCTGTGTGTGACCCTGAA
 CTGCAAGGATGTGAATGCCACCAACACCACCAATGACTCTGAGGGCACTATGGAGAGGGGTG
 AGATCAAGAACTGCAGCTTCAACATCACACCAGCATCAGGGATGAGGTGCAGAAGGAGTAT
 GCCCTGTTCTACAAGCTGGATGTGGTGCCATTGACAACAACAACACCAGCTACAGGCTGAT
 CAGCTGTGACACCTCTGTGATCACCCAGGCCTGCCCCAAGATCAGCTTTGAGCCCATCCCCA
 TCCACTACTGTGCCCTGTGCTGGCTTTGCCATCCTGAAGTGCAATGACAAGACCTTCAATGGC
 AAAGGCCCTTGCAAGAATGTGAGCACTGTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT
 CACCCAGCTGCTGCTGAATGGCAGCCTGGCTGAGGAGGAGGTGGTGTATCAGGTCTGACAAC
 TCACCAACAATGCCAAGACCATCATTGTGCAGC
 TGAAGGAGTCTGTGGAGATCAACTGCACCAGGCCCAACAACAACACCAGGAAGAGCATTCAC
 ATTGGCCCTGGCAGGGCCTTCTACACCCTGGGGAGATCATTGGGGACATCAGGCAGG
 CCCACTGCAACATCAGCAGGGCCCAAGTGAATGACACCCTGAAGCAGATTGTGATCAAGCTG
 AGGGAGCAGTTTGAGAACAAGACCATTGTGTTCAATCACAGCTCTGGTGGTGATCCTGAGAT
 TGTGATGCACAGCTTCAACTGTGGTGGTGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCA
 ACAGCACCTGGAACAACAACACTGAGGGCAGCAACAACACTGAGGGCAACACCATCACCTG
 CCTTGAGGATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTATGCTCC
 TCCCATCAGGGGCCAGATCAGGTGCAGCAGCAACATCACTGGCCTGCTGCTGACCAGGGATG
 GTGGCATCAATGAGAATGGCACTGAGATTTTCAGGCCTGGTGGTGGGGACATGAGGGACAAC
 TGGAGGTCTGAGCTGTACAAGTACAAGGTGGTGAAGATTGAGCCCCCTGGTGTGGCTCCAC
 CAAGGCTAAGCGCAGGGTGGTGCAGAGGGAGAAGcgcgctgtgggctgaggatccCGAGGGT
 GAGTGCTCCTGCCTGGACGCATCCCGGCTATGCAGCCCCAGTCCAGGGCAGCAAGGCAGGCC
 CCGTCTGCCTCTTCAACCGGAGCCTCTGCCCCGCCCACTCATGCTCAGGGAGAGGGTCTTCT
 GGCTTTTTTCCCAGGCTCTGGGCAGGCACAGGCTAGGTGCCCTAACCAGGCCCTGCACACA
 AAGGGGCAGGTGCTGGGCTCAGACCTGCCAAGAGCCATATCCGGGAGGACCCTGCCCTGAC
 CTAAGCCCACCCAAAGGCCAACTCTCCACTCCCTCAGCTCGGACACCTTCTCTCCTCCCA
 GATTCCAGTAACTCCCAATCTTCTCTCTGCAGAGCCCAAATCTTGTGACAAAACCTCACACAT
 GCCCACCCTGCCAGGTAAGCCAGCCAGGCCTCGCCCTCCAGCTCAAGGCGGGACAGGTGC

CCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACACGTCCACCTCCATCT
 CTTCTCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAACCCAAG
 GACACCCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACGA
 AGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAA
 AGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTACGCGTCTCACCCTCCTGCAC
 CAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCC
 CATCGAGAAAACCATCTCCAAAGCCAAAGGTGGGACCCGTGGGGTGGCAGGGGCCACATGGAC
 AGAGGCCGGCTCGGCCCCACCCTCTGCCCTGAGAGTGACCGCTGTACCAACCTCTGTCCTACA
 GGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAA
 CCAGGTGAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG
 AGAGCAATGGGCAGCCGGAGAACAATAACAAGACCACGCCTCCCGTGGTGGACTCCGACGGC
 TCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTT
 CTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGT
 CTCCGGGTAAATGAGTGCGACGGCCGCAGGTAAGCCAGCCCAGGCCTCGCCCTCCAGCTCAA
 GGCGGGACAGGTGCCCTAGAGTAGCCTGCATCCAGGGACAGGCCCCAGCCGGGTGCTGACAC
 GTCCACCTCCATCTCTTCTCAGGTCTGCCCGGGTGGCATCCCTGTGACCCCTCCCAGTGC
 CTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCAACAGCCTTGTCTTAATAAAATTAAGT
 TGCATCATTTTGTCTGACTAGGTGTCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATG
 GAGCAAGGGGCCCAAGTTAACTTGTATTATGCAGCTTATAATGGTTACAAATAAAGCAATAG
 CATCACAATTTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC
 TCATCAATGTATCTTATCATGTCTGgatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctgtt
 ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt gtgagctcat

CGGCGGCCGCCCTATTCTATAGTGTACCTAAATGCTA**GAGCTC**GCTGATCAGCCTCGACTG
 TGCCTT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTTCTTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCCTATGGCTTCTGAGGCGGAAAGAACC**CA**ACCACCGCGG
 TGGCGGCCGCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATATTTTATTTCT
 AGAGAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCTGAGGAATT
 CGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGGGTTTTATAG
 GGCGCCGCCGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACGCTGATTGGC
 CCCGCGCCGCTCACTACCGGCTTCGCCGCACAGTGCAGCATTTTTTTTACCCCTCTCCCT
 CCTTTTGCAGAAAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGAGGAGTTTGT
 GCGTTGGCCGCTTGGGGTGGTGGCGTGCAGATATCTAAGGGCGAATTCGATATCAagct
 agct**Tgtcgact**cg aagatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctgtt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac

541 gcgcatgccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggctg

601 ccccggtcctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgccggt
 661 ggagactgca gcctccgccg ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggggaac ttaatgtcgt
 841 ttctcagcag ctggtggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctggt
 tggattttgga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcgggtctcg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccggt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgct atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg ggcgggcgct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgatcat aggccatttt tacaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gcccaggggc gtagttaccc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggcccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggt cacctgctct
 acggcatctc gatccagcat

2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggatcatgtc tttccacggg cgcaggggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggtcctg
 2281 ctgggtgctga agcgtgcgcg gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttaa aacgaattca
 atagcttggt gcattggcgc
 2701 cgatataaaa tgcaagggtg tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actgggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcagggttgat tcatcggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca ccagaaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccggtt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg

3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc
 3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgct
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caacccggtg agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gcccagtgcc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattggtgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacggtgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc

5101 ccg gat gaat gtc agt act ggg ctat ctg gaca aggg aa
 aac gca agc g caa agag aaaa
 5161 gcagg tagct tgc agt gggc ttac atggcg atag cttag ac
 tgg gcg gttt tat ggac agc
 5221 aag cga accg gaatt gccag ctgg ggcg cc ctct ggta ag
 gtt ggga agc cct gcaa agt
 5281 aaact ggatg gcttt tctgc cgcca aggat ctgat ggcgc
 aggg gatcaa gct ctgat ca
 5341 agag acagga tgagg atcgt ttc gcat gat tgaaca agat
 ggatt gcacg caggt tctcc
 5401 ggcc gcttg gtgg agaggc tatt cggcta tgact gggca
 caac agacaa tcgg ctgctc
 5461 tgat gccgc gtgt tccgc tgc agcgca ggg gcgccc g
 gtt ctttttg tca agaccga
 5521 cctgt ccgt gccct gaatg aact gcaaga cgagg cagcg
 cgg ctatcgt ggct ggccac
 5581 gacgg gcgt cctt gcgc ag ctgt gctcga cgtt gtcact
 gaag cgggaa ggg actggct
 5641 gctatt gggc gaagt gccgg ggc aggatct cctgt catct
 caccttgctc ctgcc agaa
 5701 agt atccatc atggct gatg caat gcggcg gctgc atacg
 ctt gatccgg ctacct gcc
 5761 attc gaccac caagc gaaac atc gcatcga gcg agcacgt
 actc ggatgg aagcc ggtct
 5821 tgc gatcag gatgat ctgg acga agagca tcagg ggcgc
 gcg ccagccg aact gttcgc
 5881 cagg ctcaag gcg agcatgc ccg acggcg ggatct cgtc
 gtg acccatg gcgat gcctg
 5941 cttgcc gaat atcat ggtgg aaaat ggccg ctttt ctgga
 ttc atcgact gtgg ccggct
 6001 ggggtgtggc gacc gctatc aggac atagc gttgg ctacc
 cgt gatattg ctga agagct
 6061 tggcggcgaa tgggct gacc gcttcctcgt gctttacgg
 atc gccgctc ccg attcga
 6121 gcg catcgcc ttct atcgcc ttcttgacga gttcttctga
 atttt gttaa aattttt gtt
 6181 aaatc agctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatag accgc gatagg ggtg agtggtgttc cagtttggaa
 caag agtcca ctattaaaga
 6301 acgtgg actc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggc cgagg tgccg
 taaagctcta aatcggaacc
 6421 ctaaagg gag ccccgattt agagcttgac ggggaaagcc
 ggcaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID No:3

GAAACGATATGGGCTGAATACTTAAGTATTAGCCCATATCGTTTC.

Seq id no:4

CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCGCCCTTAGATATC
GTCGACGCCAGCACCCCAAGGCGGCCAACGCCAAACTCTCCCTCCTCCTCTTCTCAATC
TCGCTCTCGCTCTTTTTTTTTTTCGCAAAGGAGGGGAGAGGGGGTAAAAAATGCTGCACT
GTGCGGCGAAGCCGGTGAGTGAGCGGCGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTG
CCTTTTATGGCTCGAGCGGCCGCGGCGGCCCTATAAAACCCAGCGGCGCGACGCGCCACC
ACCGCCGAGACATCGATGATATCTA

AAGGGCGAATTCCTGCAGCCCGGGGATCCACTAGTCTAGATGCATGCTCGAGCGGCCGCCA
GTGTGATGGATATCTGCAGAATTCGCCCTTCAGCTGCGGATCCATTCGCCATTCAGGCTGCG
CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAA
ACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCT
CGAGGTGACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCACTAGTT
TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGGCGGCCG
CCACCGCGGTGGAGCT

Atcgaattcaagcttgtcgactc gaagatc

CT AG ACTAGTGGAT CCCCCGGGCTGCAGGAATT

CGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGGGTTTTATAG
GGCGCCGCGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACGCTGATTGGC
CCCGCGCCGCTCACTACCGGCTTCGCCGCACAGTGACGATTTTTTTTACCCCTCTCCCCT
CCTTTTGCAGAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGAGAGTTTTG
GCGTTGGCCGCTTGGGGTGCTGGGCGTCGACGATATCTAAGGGCGAATT C
GATATCAAGC TTATCGATAC CGTCGACCTC GAGGGGGGC CCG

SEQ NO:5

CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC
GAATTCGCCCTTAGATATCGTCGACGCCAGCACCCCAAGGCGGCCA
ACGCCAAACTCTCCCTCCTCCTCTTCTCAATCTCGCTCTCGCTCTTTTTTTTTTCGCAA
AAGGAGGGGAGAGGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCG
CGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGGCGG
CGCCCTATAAAACCCAGCGGCGCGACGCGCCACCACCGCCGAGACATCGATGATATCTA
AAGGGCGAATTCCTGCAGCCCGGGGGATCCACTAGTCTAGATGCATGCTCGAGCGGCCGCCA
GTGTGATGGATATCTGCAGAATTCGCCCTT
CAGCTGCGGATCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGC
GATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACG
TTGTAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAAT
TGGGTACCGGGCCCCCCTCGAGGTGACGGTATCGATAAGCTTGATATC
GAATTCCTGCAGCCCGGGGGATCCACTAGTT
TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGGCGGCCG
CCACCGCGGTGGAGCTA

Tcgaattcaagcttgtcgactc gaagatcgata

caggaagtgaacaattttcgcgcggttttaggcggatgttg

tagtaaatttgggcgtaaccgagtaagatttggccattttcgcgggaaaactgaataaga

ggaagtgaaatctgaataatTTTTgtgttactcatagcgcgtaataactggtac
 CGGGCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATC
 GAATTCGCCCTTAGATATCGTCGACGCCCAGCACCCCAAGGCGGCCA
 ACGCCAAACTCTCCCTCCTCCTTCTCAATCTCGCTCTCGCTCTTTTTTTTTTCGCAA
 AAGGAGGGGAGAGGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCG
 CGGGGCCAATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGGCGG
 CGCCCTATAAAACCCAGCGGCGCGACGCGCCACCACCGCCGAGACATCGATGATATCTA
 AAGGGCGAATTCCTGCAGCCCCGGGGGATCCACTAGTCTAGAACTAGTGGATCCCCCGGGCTG
 CAGGAATTC
 GATATCAAGCTTATCGATAACCGTCGACCTCGAGGGGGGGCCCGGTACCCA
 ATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAA
 CGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTAATCGCCTTGCAGC
 ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC
 GCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGATCCGCAGCTG
 AAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCA
 TCTAGAAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGGCGGCCG
 CCACCGCGGTGGAGCTA

SEQ ID:6

taacatcatc aataatatac cttatTTTTgg attgaagcca atatgataat
 gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
 caggaagtga caatTTTcgc gcggtTTtag gcggtgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taataact**GCTAGAgatCT**GGCGAAAGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC
 CAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCAcagagccttggcccattgcatacgttgtatccatatcata
 atatgtacatttatatttggtcatgtccaacattaccgccatgttgacattgattattgact
 agttattaatagtaatcaattacgggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggcccgctggctgacggcccaacgacccccgcccattgacgt
 caataatgacgtatgttcccataagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaaactgccacttggcagtagcatcaagtgtatcatatgccaagtacgcc
 ccctattgacgtcaatgacggtaaatggcccgctggcattatgccagtagcatgaccttat
 gggactttcctacttggcagtagcatctac**cgtagtagtcatcgtagtagtattaccat**gggtgatgcgg
 ttttggcagtagcatcaatgggcgtggatagcggtttgactcacggggattttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgccgtggagacgcatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtggtgag
 gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtgggagacagagaagactcttgggtttctgataggcactgactctctctgcctattgggtc
 tattttcccacccttaggctgctggtctgag**cctagGAGATCTCTCGAGGTCGACGGTAT**cg
 atgcca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCT

51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
 101 AGCAGGTCGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
 151 GACATCTTGG AAAAGACCCA CAACGGCAAA CTTTGCAGACC TGGATGGAGT
 201 GAAGCCCCTG ATCCTCCGGG ACTGTTCACT CGCTGGTTGG CTGCTCGGGA
 251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
 301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA

351 TTACGAGGAA CTGAAACACC TGTTGTCTAG AATTAACCAC TTTGAAAAGA
 401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
 451 GTTAGCTCAG CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
 501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAAGTCTT
 551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTTGTGGGG TATCCATCAC
 601 CCGAACGACG CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA
 651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
 701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
 751 ACAATCTTAA AACCCAACGA TGCATAAAT TTCGAGTCCA ATGGCAATTT
 801 CATCGCCCTT GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA
 851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA
 901 CCAATGGGAG CCATTAATAG CTCAATGCCC TTCCATAATA TTCATCCATT
 951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCTCG
 1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
 1051 GGCCTGTTTG GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
 1101 GGTTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
 1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
 1201 AAGGTTAATA GTATTATAGA TAAATGAAC ACGCAATTTC AGGCCGTTGG
 1251 GAGGGAGTTT AACAATCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 1301 TGGAAGACGG CTTCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 1351 CTCATGGAGA ACGAGAGGAC CCTGGATTTC CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
 1451 GGAATGGATG CTTTCGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
 1501 AGTATACGCA ACGGGACCTA CAATTACCTT CAGTATAGCG AAGAGGCTCG
 1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
 1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
 1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG GAAGCCTTCA
 1701 GTGCCGTATT TGTATATGA c
GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAAC**CAA**

 CCACCGCGGTGGCGGCCGCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATAT
 TTTATTTCTAGAGAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCCT
 GCAGGAATTTCGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGG
 GTTTTATAGGGCGCCGCCGCGGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACG
 CTGATTGGCCCCGCGCCGCTCACTCACCGGCTTCGCCGCACAGTGCAGCATTTTTTTTACCCC
 CTCTCCCCTCCTTTTGCAGAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGG
 AGAGTTTTGGCGTTGGCCGCCTTGGGGTGCTGGGCGTCGACGATATCTAAGGGCGAATTCGA
 TATCA

AGCTAGCT **Tgtcgactcg** aagatctggg cgtggttaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctgtt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac
 541 gcgcatgccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggctg
 601 ccccgctcctg cccgcaaact ctactacett gacctacgag
 accgtgtctg gaacgcccgtt
 661 ggagactgca gcctccgccc ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt
 841 ttctcagcag ctggtggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctgtt
 tggatttga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccgt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggt ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tctaactacgt catagtgtg
 1561 ttccaggatg agatcgatc atggccatttt tacaagcgc
 gggcggaggg tgccagactg
 1621 cggtataatg gttccatccg gccagggggc gtagttacce
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcagttt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgcccagc
 gatagcagtt cttgcaagga

1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggc cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttgcggggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacggggcca gggtcacgtc tttccacggg cgcaggggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggtcctg
 2281 ctggtgctga agcgtgcgcg gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgcgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accagggttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg.
 2641 tccccgtata cagacttgag agggagttta aacgaattca
 atagcttggt gcatgggcgg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac cttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcagggtgat tcatcggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgtt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaacattt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 aaaaaaaca cccagaaaac cgcacgcgaa cctacgcca
 gaaacgaaag caaaaaac

3481 cacaacttcc tcaaatoctc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tcccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg
 3601 tccccacgcc ccgcgccacg tcacaaactc cccccctca
 ttatcatatt ggcttcaatc
 3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgtc
 3781 cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt tttccccctg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcgggt gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggat gtaggcgggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gcccagtgcc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtggctc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag

4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
gcagccatga gattatcaaa
5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
ccgcagaaac ggtgctgacc
5101 ccgatgaat gtcagctact gggctatctg gacaagggaa
aacgcaagcg caaagagaaa
5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
tgggcggttt tatggacagc
5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
gttgggaagc cctgcaaagt
5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
aggggatcaa gctctgatca
5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
ggattgcacg caggttctcc
5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
caacagacaa tcggctgctc
5461 tgatgccgcc gtgttccggc tgtcagcgca ggggcgccc
gttctttttg tcaagaccga
5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
cggctatcgt ggctggccac
5581 gacgggctt ccttgccgag ctgtgctcga cgttgtcact
gaagcgggaa gggactggct
5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
caccttgctc ctgccagaaa
5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
cttgatccgg ctacctgcc
5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
actcggatgg aagccggtct
5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
gcgccagccg aactgttcgc
5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
gtgacccatg gcgatgcctg
5941 cttgccgaat atcatgggtg aaaatggccg cttttctgga
ttcatcgact gtggccggct
6001 ggggtgtggcg gacgcctatc aggacatagc gttggctacc
cgtgatattg ctgaagagct
6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
atcgccgctc ccgattcgca
6121 gcgcacgcgc ttctatcgcc ttcttgacga gttcttctga
attttgttaa aatttttgtt
6181 aaatcagctc attttttaac caataggccg aaatcggcaa
catcccttat aaatcaaaag
6241 aatagaccgc gatagggttg agtggtgttc cagtttgga
caagagtcca ctattaaaga
6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
gggcgatggc ccactacgtg
6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
taaagctcta aatcggaacc
6421 ctaaaggag ccccgattt agagcttgac ggggaaagcc
ggcgaacgtg gcgagaaagg

6481 aaggggaagaa agcgaaagga gcgggcgcta gggcgctggc
aagtgtagcg gtcacgctgc
6541 gcgtaaccac cacacccgcg cgcttaatgc gccgctacag
ggcgcgctcca ttcgccattc
6601 aggatcgaat taattcttaa t

SEQ ID:7

taacatcatc aataatatac cttatthttgg attgaagcca atatgataat
gagggggttg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
caggaagtga caatthttgc gcggtthttag gcggatgttg tagtaaattt
ggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
ggaagtgaat tctgaataat tttgtgttac tcatagcgcg
taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatatcata
atatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattattgact
agttattaatagtaaatcaattacgggggtcattagttcatagcccatatatggagttccgcgt
tacataacttacggtaaatggcccgcttggctgaccgccaacgacccccgcccattgacgt
caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
gagtatttacggtaaatgcccacttggcagtagcatcaagtgtatcatatgccaagtacgcc
ccctattgacgtcaatgacggtaaatggcccgcttggcattatgccagtagcatgaccttat
gggactttccctacttggcagtagcatctaccgtaattagtcacgcctattaccat**gggtgatgcgg**
ttttggcagtagcatcaatgggcgtggatagcggtttgactcacggggatttccaagtctcca
ccccattgacgtcaatgggagttttgttttggcaccaaaatcaacgggactttccaaaatgtc
gtaacaactccgccccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatata
agcagagctcgthtttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacct
ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtgggtgag
gccctgggcaggttgggtatcaaggttacaagacaggtttaaggagaccaatagaaaactgggc****
atgtgggagacagagaagactccttgggtttctgataggcactgactctctctgcctattgggtc
tattttcccacccttaggctgctggtctgagcctagGAGATCTCTCGAGGTGACGGTAT**cg**
atgccca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCT

51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
101 AGCAGGTCGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
151 GACATCTTGG AAAAGACCCA CAACGGCAAA CTTTGCGACC TGGATGGAGT
201 GAAGCCCCTG ATCCTCCGGG ACTGTTCACT CGCTGGTTGG CTGCTCGGGA
251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA
351 TTACGAGGAA CTGAAACACC TGTGTCTAG AATTAACCAC TTTGAAAAGA
401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
451 GTTAGCTCAG CGTGTCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT
551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTGTGGGG TATCCATCAC
601 CCGAACGACG CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA
651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
751 ACAATCTTAA AACCCAACGA TGCGATAAAT TTCGAGTCCA ATGGCAATTT
801 CATCGCCCT GAATACGCCT ATAAGATCGT GAAAAGGGG GACTCTGCAA
851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA

901 CCAATGGGAG CCATTAATAG CTCAATGCCC TTCCATAATA TTCATCCATT
 951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCTCG
 1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
 1051 GGCCTGTTTG GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
 1101 GGTTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
 1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
 1201 AAGGTTAATA GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTTGG
 1251 GAGGGAGTTT AACAATCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 1301 TGGGAAGACGG CTTCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 1351 CTCATGGAGA ACGAGAGGAC CCTGGATTTC CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
 1451 GGAATGGATG CTTTCGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
 1501 AGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG AAGAGGCTCG
 1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
 1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
 1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG GAAGCCTTCA
 1701 GTGCCGTATT TGTATATGAg c
GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAAC**CAA**

TGATATC**GAATT**CGCCCTTAGATATCGTCGAC**GGCCAGCACCCCAAG**GCGGCCAACGCCAAA
 ACTCTCCCTCCTCCTCTTCTCAATCTCGCTCTCGCTCTTTTTTTTTTCGCAAAGGAGGG
 GAGAGGGGGTAAAAAATGCTGCACTGTGCGGCGAAGCCGGTGAGTGAGCGGCGCGGGG**CCA**
ATCAGCGTGCGCCGTTCCGAAAGTTGCCTTTTATGGCTCGAGCGGCCGCGCGGCCCTAT
AAAACCCAGCGGCGCGACGCGCCACCACCGCCGAGACATCGATGATATCTAAAGGGCGAATT
 CCTGCAGgaaacgatatgggctgaatacggatccgtattcagcccatatcgtttc**TCTAGAA**
 ATAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTG**GCGGCCGCCACCGC**
 GGTGG

agctTgtcgactcg aagatctggg cgtgggtaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
tgtatctggt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
gtgagctcat atttgacaac

541 gcgcatgccc ccatgggccc ggggtgcgtca gaatgtgatg
ggctccagca ttgatggctg

601 ccccgctctg cccgcaaact ctactacctt gacctacgag
accgtgtctg gaacgccgtt

661 ggagactgca gcctccgccc cgcttcagc cgctgcagcc
accgcccgcg ggattgtgac

721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
tcccgttcat ccgcccgcga

781 tgacaagttg acggctcttt tggcacaatt ggattctttg
acccgggaac ttaatgtcgt

841 ttctcagcag ctgttgatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctgtt
 tggatttgga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcgggtctcg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccggt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtgggt ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtgggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgct atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggccacg ggcggcgccct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgatcat aggccatttt taaaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gcccaggggc gtagttaccc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcgggt cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acgggtgaag ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctgggtgctga agcgtgcgcg gtcttcgccc tgcgcgtcgg
 ccaggtagca ttgaccatg

2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc ctggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcgca ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttta aacgaattca
 atagcttggt gcatgggcgg
 2701 cgatataaaa tgcaagggtgc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcagggtgat tcatcgggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aaggggccaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgcca
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaactcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc
 3661 caaataaagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttcgg ctctctcgtc
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc

3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgcttttctca
 4141 tagctcacgc tgtaggatc tcagttcggg tagggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcgggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gcccagtgcc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtgggtcc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca

5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcggt ccttgccgag ctgtgctcga cgttgtcact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgcggg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggctc
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacctatg gcgatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcgggcga tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttggtta aatttttggt
 6181 aatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgga
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggc cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggagg ccccgattt agagcttgac ggggaaagcc
 ggcgaaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 8 GATGGTGCTTCAAGCTAGTACttaaGTACTAGCTTGAAGCACCATC

SEQ ID: 9 GATGGTGCTTCAAGCTAGTACggatccGTACTAGCTTGAAGCACCATC

SEQ ID: 10 GAAACGATATGGGCTGAATACggatccGTATTCAGCCCATATCGTTTC

SEQ ID: 11 CCTAATAATTAT CAAAATGTggatccACATTTTGATAATTATTAGG

SEQ ID: 12 CCTAATAATTAT CAAAATGTaattACATTTTGATAATTATTAGG

SEQ ID: 13 (DS1c)

1 taacatcatc aataatatac cttatTTTTgg attgaagcca atatgataat
gaggggggtgg

61 agttttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
acgtagtagt gtggcggaag

121 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg
atgtggcaaa agtgacgttt

181 ttggtgtgcg ccggtgtaca caggaagtga caattttcgc
gcggttttag gcggatgttg

241 tagtaaattt gggcgtaacc gagtaagatt tggccatttt
cgcgggaaaa ctgaataaga

301 ggaagtgaat tctgaataat tttgtgttac tcatagcgcg
taatactgta atagtaatca

361 attacggggc cattagttca tagcccatat atggagttcc
gcgttacata acttacggta

421 aatggcccg cgtggtgacc gccaacgac ccccgcccat
tgacgtcaat aatgacgtat

481 gttcccatag taacgccaat agggactttc cattgacgtc
aatgggtgga gtatttacgg

541 taaactgcc acttggcagt acatcaagtg tatcatatgc
caagtacgcc ccctattgac

601 gtcaatgacg gtaaattggc cgctggcat tatgcccagt
acatgacctt atgggacttt

661 cctacttggc agtacatcta **cgtattagtc atcgctatta**
ccatggtgat gcggttttgg

721 cagtacatca atgggcgtgg atagcggttt gactcacggg
gatttccaag tctccacccc

781 attgacgtca atgggagttt gttttggcac caaatcaac
gggactttcc aaaatgtcgt

841 aacaactccg cccattgac gcaaatgggc ggtaggcgtg
tacggtggga ggtctatata

901 agcagagctg gtttagtgaa ccgtcagatc cgctagagat
ctGGTACCGAGCTCGGATCCGCCACCATGGAGGCAAACCTACTGGTCCTGTTATGTGCACTT
GCAGCTGCAGATGCAGACACAATATGTATAGGCTACCATGCGAACAATTCAACCGACACTGG
TGACACAGTA**ctcgaga**aga atgtgacagt **gacacactot gttaacctgc**

181 tcgaagacag ccacaacgga aaactatgta gattaaaagg
aatagcccca ctacaattgg

241 ggaaatgtaa catcgccgga tggctcttgg gaaaccaga
atgcgaccca ctgcttccag

301 tgagatcatg gtcctacatt gtagaaacac caaactctga
gaatggaata tgttatccag

361 gagatttcat cgactatgag gagctgaggg agcaattgag
ctcagtgtca tcattcgaaa

421 gattcgaaat atttccaaa gaaagctcat ggcccaacca
caacacaacc aaaggagtaa

481 cggcagcatg ctcccatgcg gggaaaagca gtttttacag
aaatttgcta tggctgacgg

541 agaaggaggg ctcataccca aagctgaaaa attcttatgt
 gaacaagaaa gggaaagaag
 601 tccttggtact gtgggggtatt catcacccgt ctaacagtaa
 ggatcaacag aatatctatc
 661 agaatgaaaa tgcttatgtc tctgtagtga cttcaaatta
 taacaggaga tttaccccg
 721 aatagcaga aagacccaaa gtaagagatc aagctgggag
 gatgaactat tactggacct
 781 tgctaaaacc cggagacaca ataatatattg aggcaaattg
 aaatctaata gcaccaaggt
 841 atgctttcgc actgagtaga ggctttgggt ccggcatcat
 cacctcaaac gcatcaatgc
 901 atgagtgtaa cacgaagtgt caaacacccc tgggagctat
 aaacagcagt ctccctttcc
 961 agaataata cccagtcaca ataggagagt gcccaaaata
 cgtcaggagt gccaaattga.
 1021 ggatgggttac aggactaagg aacattccgt ccattcaatc
 cagaggtcta tttggagcca
 1081 ttgccgggtt tattgaagg ggatggactg gaatgataga
 tggatgggtac gggtatcatc
 1141 atcagaatga acagggatca ggctatgcag cggatcaaaa
 aagcacacaa aatgccatta
 1201 acgggattac aaacaagggtg aactctgtta tcgagaaaat
 gaacattcaa ttcacagctg
 1261 tgggtaaaga attcaacaaa ttagaaaaaa ggatggaaaa
 tttaaataaa aaagttgatg
 1321 atggatttct ggacatttgg acatataatg cagaattggt
 agttctactg gaaaatgaaa
 1381 ggactctgga tttccatgac tcaaattgtga agaattctgta
 tgagaaagta aaaagccaat
 1441 taaagaataa tgccaaagaa atcggaaatg gatgttttga
 gttctaccac aagtgtgaca
 1501 atgaatgcat ggaaagtgtg agaaatggga cttatgatta
 tcccaaatat tcagaagagt
 1561 caaagttgaa cagggaaaag gtagatggag tgaaattgga
 atcaatgggg atctatcaga
 1621 ttctggcgat ctactcaact gtcgccagtt cactgggtgct
 tttggtctcc ctgggggcaa
 1681 tcagtttctg gatgtgttct aatggatctt tgcagtgag
 aatatgcac tgagattaga
 1741 atttcagaga tatgaggaaa aacacccttg tttctact
 CCCAAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAA
 CGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCAC
 CCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCTCTTGCGGGATGG
GCGGCCGCtcgagcct aagcttctag ataagatatc cgatccaccg gatctagata
 actgatcata
 1021 atcagccata ccacatttgt agaggtttta cttgctttta
 aaaacctccc acacctccc
 1081 ctgaacctga aacataaaat gaatgcaatt gttgtttgta
 acttgtttat tgcagcttat

1141 aatggttaca aataaagcaa tagcatcaca aatttcacaa
 ataaagcatt tttttcactg
 1201 cattctagtt gtggtttgtc caaactcatc aatgtatctt
 aacgcggatc tgggcgtggt
 1261 taagggtggg aaagaatata taagggtggg gtcttatgta
 gttttgtatc tgttttgcag
 1321 cagccgccgc cgccatgagc accaactcgt ttgatggaag
 cattgtgagc tcatatttga
 1381 caacgcgcat gcccccatgg gccgggggtgc gtcagaatgt
 gatgggctcc agcattgatg
 1441 gtgcgcccggt cctgccccga aactctacta ccttgaccta
 cgagaccgtg tctggaacgc
 1501 cgttggagac tgcagcctcc gccgccgctt cagccgctgc
 agccaccgcc cgcgggattg
 1561 tgactgactt tgctttcctg agcccgcttg caagcagtgc
 agcttcccgt tcatccgccc
 1621 gcgatgacaa gttgacggct cttttggcac aattggattc
 tttgaccggg gaacttaatg
 1681 tcgtttctca gcagctggtg gatctgcgcc agcaggtttc
 tgccctgaag gcttctctcc
 1741 ctcccaatgc ggtttaaaac ataaataaaa aaccagactc
 tgtttgatt tggatcaagc
 1801 aagtgtcttg ctgtctttat ttaggggttt tgcgcgcgcg
 gtagggccgg gaccagcggg
 1861 ctcggtcgtt gagggtcctg tgtatttttt ccaggacgtg
 gtaaagggtga ctctggatgt
 1921 tcagatacat gggcataagc ccgtctctgg ggtggaggta
 gcaccactgc agagcttcat
 1981 gctgcggggg ggtgtttag atgatccagt cgtagcagga
 gcgctgggcg tgggtgcctaa
 2041 aaatgtcttt cagtagcaag ctgattgcca ggggcaggcc
 cttggtgtaa gtgtttacaa
 2101 agcgggtaag ctgggatggg tgcatacgtg gggatatgag
 atgcatcttg gactgtattt
 2161 ttaggttggc tatgttccca gccatatccc tccggggatt
 catgttgtgc agaaccacca
 2221 gcacagtgtg tccggtgcac ttgggaaatt tgtcatgtag
 cttagaagga aatgcgtgga
 2281 agaacttgga gacgcccttg tgacctcaa gattttccat
 gcattcgtcc ataattgatg
 2341 caatgggccc acgggcggcg gcctgggcga agatatttct
 gggatcacta acgtcatagt
 2401 tgtgttccag gatgagatcg tcataggcca tttttacaaa
 gcgcgggccc agggtgccag
 2461 actgcggtat aatggttcca tccggcccag gggcgtagtt
 accctcacag atttgcatth
 2521 cccacgcttt gagttcagat ggggggatca tgtctacctg
 cggggcgatg aagaaaacgg
 2581 tttccggggg aggggagatc agctgggaag aaagcagggt
 cctgagcagc tgcgacttac

2641 cgcagccggt gggcccgtaa atcacaccta ttaccgggtg
 caactggttag ttaagagagc
 2701 tgcagctgcc gtcattccctg agcagggggg ccacttcggt
 aagcatgtcc ctgactcgca
 2761 tgttttccct gaccaaattcc gccagaaggc gctcgccgcc
 cagcgatagc agttcttgca
 2821 aggaagcaaa gtttttcaac ggtttgagac cgtccgcccgt
 aggcattgctt ttgagcggtt
 2881 gaccaagcag ttccaggcgg tcccacagct cggtcacctg
 ctctacggca tctcgatcca
 2941 gcatactctc tcgtttcgcg gggtggggcg gctttcgcgt
 tacggcagta gtcgggtgctc
 3001 gtccagacgg gccagggtca tgtctttcca cgggcgcagg
 gtccctcgtca gcgtagctctg
 3061 ggtcacgggtg aaggggtgctg ctccgggctg cgcgctggcc
 aggggtgcgt tgaggctggt
 3121 cctgctggtg ctgaagcgct gccggtcttc gccctgcgcg
 tcggccagggt agcatttgac
 3181 catggtgtca tagtccagcc cctccgcggc gtggcccttg
 gcgcgcagct tgcccttgga
 3241 ggaggcgccg cacgaggggc agtgcagact tttgaggggc
 tagagcttgg gcgcgagaaa
 3301 taccgattcc ggggagtagg catccgcgcc gcaggccccg
 cagacggtct cgcattccac
 3361 gagccagggtg agctctggcc gttcggggtc aaaaaccagg
 ttcccccat gctttttgat
 3421 gcgtttctta cctctggttt ccatgagccg gtgtccacgc
 tcggtgacga aaaggctgct
 3481 cgtgtccccg tatacagact tgagagggag tttaaagaa
 ttcaatagct tggtgcatgg
 3541 gcggcgatat aaaatgcaag gtgctgctca aaaaatcagg
 caaagcctcg cgcaaaaaag
 3601 aaagcacatc gtagtcatgc tcatgcagat aaaggcagg
 aagctccgga accaccacag
 3661 aaaaagacac cttttttctc tcaaacatgt ctgcggggtt
 ctgcataaac acaaaataaa
 3721 ataacaaaaa aacattttaa cattagaagc ctgtcttaca
 acaggaaaaa caacccttat
 3781 aagcataaga cggactacgg ccatgccggc gtgaccgtaa
 aaaaactggt caccgtgatt
 3841 aaaaagcacc accgacagct cctcgggtcat gtccggagtc
 ataattgtaag actcggtaaa
 3901 cacatcagggt tgattcatcg gtcagtgcta aaaagcgacc
 gaaatagccc gggggaatac
 3961 ataccgcag gcgtagagac aacattacag ccccatagg
 aggtataaca aaattaatag
 4021 gagagaaaaa cacataaaca cctgaaaaac cctcctgcct
 aggcataata gcaccctccc
 4081 gctccagaac aacatacagc gcttcacagc ggcagcctaa
 cagtcagcct taccagtaaa

4141 aaagaaaacc tattaataaaa acaccactcg acacggcacc
 agctcaatca gtcacagtgt
 4201 aaaaaagggc caagtgcaga gcgagtatat ataggactaa
 aaaatgacgt aacgggttaa
 4261 gtccacaaaa aacacccaga aaaccgcacg cgaacctacg
 cccagaaacg aaagccaaaa
 4321 aaccacaaac ttctcaaat cgtcacttcc gttttccac
 gttacgtaac ttccatttt
 4381 aagaaaacta caattcccaa cacatacaag ttactccgcc
 ctaaaaccta cgtcacccgc
 4441 cccgttccca cgccccgcgc cacgtcaca actccacccc
 ctcattatca tattggcttc
 4501 aatccaaaat aaggtatatt attgatgatg ttaattaaca
 tgcattgatc catatgcggt
 4561 gtgaaatacc gcacagatgc gtaaggagaa aataccgcat
 caggcgctct tccgcttcct
 4621 cgctcactga ctgcgtgcgc tcggtcgttc ggctgcggcg
 agcggtatca gctcactcaa
 4681 aggcggtaat acggttatcc acagaatcag gggataacgc
 aggaaagaac atgtgagcaa
 4741 aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt
 gctggcggtt ttccataggc
 4801 tccgcccccc tgacgagcat cacaaaaatc gacgctcaag
 tcagaggtgg cgaaaccgca
 4861 caggactata aagataccag gcgtttcccc ctggaagctc
 cctcgtgcgc tctcctgttc
 4921 cgaccctgcc gcttaccgga tacctgtccg cttttctccc
 ttccggaagc gtggcgcttt
 4981 ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt
 cgttcgtcc aagctgggct
 5041 gtgtgcacga accccccggt cagcccgacc gctgcgcctt
 atccggtaac tatcgtcttg
 5101 agtccaacc ggtaagacac gacttatcgc cactggcagc
 agccactggt aacaggatta
 5161 gcagagcgag gtatgtaggc ggtgctacag agttcttgaa
 gtgggtgcct aactacggct
 5221 aactagaag gacagtattt ggtatctgcg ctctgctgaa
 gccagttacc ttccgaaaaa
 5281 gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg
 tagcgggtgt tttttgttt
 5341 gcaagcagca gattacgcgc agaaaaaaag gatctcaaga
 agatcctttg atcttttcta
 5401 cggggtctga cgctcagtgg aacgaaaact cacgttaagg
 gattttggtc atgagattat
 5461 caaaaaggat cttcacctag atccttttaa attaaaaatg
 aagttttaaa tcaatctaaa
 5521 gtatatatga gtaaaacttg tctgacagtt accaatgctt
 aatcagtgag gcacctatct
 5581 cagcgatctg tctatttcgt tcatccatag ttgcctgact
 ccccgctcgt tagataacta

5641 cgatacggga gggcttacca tctggcccca gtgctgcaat
 gataccgcga gaccacgct
 5701 caccggctcc agatttatca gcaataaacc agccagccgg
 aagggccgag cgcagaagtg
 5761 gtcttgcaac tttatccgcc tccatccagt ctattaattg
 ttgcogggaa gctagagtaa
 5821 gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat
 tgctgcagcc atgagattat
 5881 caaaaaggat cttcacctag atccttttca cgtagaaagc
 cagtcgcag aaacggtgct
 5941 gaccccgat gaatgtcagc tactgggcta tctggacaag
 ggaaaacgca agcgcaaaga
 6001 gaaagcaggt agcttgcagt gggcttacat ggcgatagct
 agactgggag gttttatgga
 6061 cagcaagcga accggaattg ccagctgggg cgccctctgg
 taagggtggg aagccctgca
 6121 aagtaaaactg gatggctttc tcgccgcca ggatctgatg
 gcgcagggga tcaagctctg
 6181 atcaagagac aggatgagga tcgtttcgca tgattgaaca
 agatggattg cacgcagggt
 6241 ctccggccgc ttgggtggag aggctattcg gctatgactg
 ggcacaacag acaatcggct
 6301 gctctgatgc cgccgtgttc cggctgtcag cgcaggggag
 ccggttctt tttgtcaaga
 6361 ccgacctgtc cgggtgccctg aatgaactgc aagacgaggc
 agcgcggtta tcgtggctgg
 6421 ccacgacggg cgcttccttgc gcagctgtgc tcgacgttgt
 cactgaagcg ggaagggact
 6481 ggctgctatt gggcgaagtg ccggggcagg atctcctgtc
 atctcacctt gctcctgcc
 6541 agaaagtatc catcatggct gatgcaatgc ggcggctgca
 tacgcttgat ccggctacct
 6601 gccattcga ccaccaagcg aaacatcgca tcgagcgagc
 acgtactcgg atggaagccg
 6661 gtcttgctga tcaggatgat ctggacgaag agcatcaggg
 gctcgcgcca gccgaactgt
 6721 tcgccaggct caaggcgagc atgcccagcg gcgaggatct
 cgctgtgacc catggcgatg
 6781 cctgcttgcc gaatatcatg gtggaaaatg gccgcttttc
 tggattcatc gactgtggcc
 6841 ggctgggtgt ggcggaccgc tatcaggaca tagcgttggc
 taccgctgat attgctgaag
 6901 agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta
 cggatcgcg gctcccgaatt
 6961 cgcagcgcac cgccttctat cgccttcttg acgagttctt
 ctgaattttg ttaaaatttt
 7021 tgttaaatca gctcattttt taaccaatag gccgaaatcg
 gcaacatccc ttataaatca
 7081 aaagaataga ccgcgatagg gttgagtgtt gttccagttt
 ggaacaagag tccactatta

7141 aagaacgtgg actccaacgt caaagggcga aaaaccgtct
 atcagggcga tggcccacta
 7201 cgtgaacat caccacaatc aagttttttg cggtcgaggt
 gccgtaaagc tctaaatcgg
 7261 aaccctaaag ggagcccccg atttagagct tgacggggaa
 agccggcgaa cgtggcgaga
 7321 aaggaagga agaaagcgaa aggagcgggc gctagggcgc
 tggcaagtgt agcggtcacg
 7381 ctgcgcgtaa ccaccacacc cgcgcgctta atgcgccgct
 acagggcgcg tccattcgcc
 7441 attcaggatc gaattaattc ttaat

SEQ ID:14 DS2beta-luc

taacatcatc aataatatac cttatttttg attgaagcca atatgataat
 gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgcg ccggtgtaca
 caggaagtga caattttcgc gcggttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatatcata
 atatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattattgact
 agttattaatagtaatcaattacggggtcattagttcatagcccatatatggaggtccgcgt
 tacataacttacggtaaatggccgcctggctgaccgccaacgacccccgcccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtattttacggtaaaactgccacttggcagtacatcaagtgtatcatatgccaaagtacgcc
 ccctattgacgtcaatgacggtaaatggccgcctggcattatgccagttacatgacctat
 gggacttttctacttggcagttacatctac**gtattagtcacogctattaccat**gggtgatgcgg
 ttttggcagttacatcaatgggcgtggatagcggtttgactcacggggattttccaagtctcca
 cccattgacgtcaatgggagttttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttggtggtgag
 gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtggagacagagaagactcttgggtttctgataggcactgactctctctgcctattgggtc
 tattttcccacccttaggtgctggtctgag**cctagGAGATCTCTCGAGGTCGACGGTATCG**
ATGGGTACCGGCGGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGC
CTCGACTGTGCCTT
 CTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCCAC
 TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCTATGGCTTCTGAGGCGGAAAGAACC**CA**ACCACCGCGG
 TGGCGGCCGCCACACAAAAAACACACAGATGTAATGAAAATAAAGATATTTTATTTCT
 AGAGAAACGATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCCTGCAGGAATT
 CGCCCTTTAGATATCATCGATGTCTCGGCGGTGGTGGCGCGTCGCGCCGCTGGGTTTTATAG

GGCGCCGCGCGCGCCGCTCGAGCCATAAAAGGCAACTTTCGGAACGGCGCACGCTGATTGGC
 CCCGCGCCGCTCACTCACC GGCTTCGCCGCACAGTGCAGCATTTTTTTTACCCCTCTCCCT
 CCTTTTGC GAAAAAAGAGCGAGAGCGAGATTGAGGAAGAGGAGGAGGGAGAGTTTGG
 GCGTTGGCCGCTTGGGGTGCTGGGCGTCGACGATATCTAAGGGCGAATTCGATATCAgct
 agct**Tgtcgact**cg aagatctggg cgtgggttaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctgtt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac

541 gcgcatgccc ccatgggccg ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggctg

601 ccccgctctg ccgcgcaact ctactacctt gacctacgag
 accgtgtctg gaacgccgtt

661 ggagactgca gcctccgccg ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac

721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga

781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggaac ttaatgtcgt

841 ttctcagcag ctggtggatc tgccgagca ggtttctgcc
 ctgaaggctt cctcccctcc

901 caatgcggtt taaaacataa ataaaaaacc agactctgtt
 tggatttgga tcaagcaagt

961 gtcttgctgt ctttatcttag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg

1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag

1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg

1141 cgggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggctggtt gcctaaaaat

1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg

1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttgact gtatcttttag

1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac

1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa

1441 cttggagacg cccttgtagc ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat

1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagtgtg

1561 ttccaggatg agatcgatcat aggccatttt tacaaagcgc
 gggcggaggg tgccagactg

1621 cggtataatg gttccatccg gcccgagggc gtagttaccc
 tcacagattt gcatttccca

1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc

1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca

1801 gccggtgggc ccgtaaataca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca gggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacggtt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcggtccc acagctcggc cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttgcgcgggtt ggggcggcctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggtcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctggtgctga agcgcgtgcc gtcttcgccc tgcgcgtcgg
 ccaggtagca ttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcggtg cccttgggcg
 gcagcttgcc cttggaggag
 2401 gcgcgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcg cagaaatacc
 2461 gattccgggg agtaggcac cgcgcgcag gccccgcaga
 cggctctgca ttccacgagc
 2521 caggtgagct ctggccgctc ggggtcaaaa accaggtttc
 ccccatgctt ttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttggt gcatgggcg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaaagcctgt cttacaacag
 gaaaaacaac ccttataagc
 2941 ataagacgga ctacggccat gccggcggtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcgggtc gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggagggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgtc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag

3301 aaaacctatt aaaaaaacac cactogacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgcca
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatecgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tcccaacaca tacaagttac tccgcctaa
 aacctacgtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccocctca
 ttatcatatt ggcttcaatc
 3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgtc
 3781 cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctcgg
 3961 ccccccgtac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgcgcgtt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcgggt gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgagggtat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgtggttagc
 ggtgggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat

4801 acgggagggc ttaccatctg gcccagtg gcgaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtgggtcc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cggaagcta gagtaagtag
 4981 ttgccaggtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccgatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcg
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttccatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttcgggc tgtcagcgca ggggcgccc
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cgctatcgt ggctggccac
 5581 gacgggcgtt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcgatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatgggtg aaaatggccg cttttctgga
 ttcacgact gtggccggct
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcataccc ttctatcgcc ttcttgacga gttcttctga
 attttggttaa aatttttgtt
 6181 aaatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgga
 caagagtcca ctattaaaga

6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgaggtgccg
 taaagctcta aatcggaacc
 6421 cttaaaggag ccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 15 taacatcatc aataatatac cttatthttgg attgaagcca
 atatgataat gagggggtgg agtttgtgac gtggcgcggg gcgtgggaac
 ggggcgggtg acgtagtagt gtggcggaag tgtgatgttg caagtgtggc
 ggaacacatg taagcgacgg atgtggcaaa agtgacgttt ttgggtgtgcg
 ccggtgtaca caggaagtga caatthttcgc gcggtthttag gcggatgttg
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
 ctgaataaga ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCAcagagcttggcccattgcatacgttgtatccatatcata
 atatgtacattttatatattggctcatgtccaacattaccgccatgttgacattgattattgact
 agttattaatagtaatacaattacggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggccgcctggctgaccgcccacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtattttacggtaaaactgccacttggcagtagcatcaagtgtatcatatgccaagtacgcc
 ccctattgacgtcaatgacggtaaatggccgcctggcattatgccagtagcatgaccttat
 gggactttcctacttggcagtagcatc**cgatttagtcacgcctattaccatgggtgatgcgg**
 ttttggcagtagcatcaatgggcgtggatagcggtttgactcacggggattttccaagtctcca
 ccccatgacgtcaatgggagttttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggtaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtggtag
 gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaact**gggc**
atgtggagacagagaagactcttgggtttctgataggcactgactctctctgcctattggctc
 tattttcccacccttaggctgctggctctgag**cctagGAGATCTCTCGAGGTCGACGGTATCG**
ATGGGTACCGCGGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGC
CTCGACTGTGCCTT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAAC**CA**

AAGCTTAGGCTCGAGCGGCC
 GCCACACAAAAACCAACACACAGATGTAATGAAAATAAAGATATTTTATTTCTAGAGAAAC
 GATATGGGCTGAATACGGATCCGTATTCAGCCCATATCGTTTCCCAGATCTCTAGCGGATCT
 GACGGTTCATAAACCAGCTCTGCT

TATATAGACCTCCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTT
 ACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAAT
 GGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTG
 CCAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGG
 AAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGAC
 GTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTA
 CCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAAC
 ATACGTCATTATTGACGTCAATGGGCGGGGGTTCGTTGGGCGGTCAGCCAGGCGGGGCCATT
 TACCGTAAGTTATGTAACGCGGAACTCCATATATGGGCTATGAACTAATGACCCCGTAAT
 TGATTACTATTACAGTATTACGCGCTATGAGTAACACAAAATTATTCAGATTTCACTTCC
 TCTTATTCAGTTTTTCCCGCGAAAATGGCCAAATCTTACTCGGTTACGCCCAAATTTACTA
 CAACATCCGCCTAAAACCGCGCGAAAATTGTCACTTCCTGTGTACACCGGCGCACACCAA
 AAACGTCACTTTTGCCACATCCGTCGTTACATGTGTTCCGCCCACTTGCAACATCACA
 CTTCCGCCCACTACTACGTCACCCGCCCCGTTCACGCCCCGCGCCACGTCACAACT
 CCACCCCCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTGATGATGTTAAG
 CTT

gtcgactcg aagatctggg cgtggttaag

421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctgtt ttgcagcagc

481 cgccgcccgc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac

541 ggcgatgccc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggtcg

601 ccccgctcctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgcccgtt

661 ggagactgca gcctccgccc ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac

721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga

781 tgacaagttg acggctcttt tggcacaatt ggattccttg
 acccggaac ttaatgtcgt

841 ttctcagcag ctgttggatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc

901 caatgcggtt taaaacataa ataaaaaacc agactctgtt
 tggatttgga tcaagcaagt

961 gtcttgctgt ctttatattag gggttttgcg cgcgcggtag
 gcccgggacc agcggctctg

1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag

1081 atacatgggc ataagcccgt ctctggggtg gaggtagcac
 cactgcagag cttcatgctg

1141 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat

1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg

1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag

1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac

1381 agtgtatccg gtgcacttgg gaaatttgtc atgtagctta
 gaaggaaatg cgtggaagaa

1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagttgtg
 1561 ttccaggatg agatcgatcat aggccatttt taciaagcgc
 gggcggaggg tgccagactg
 1621 cgggtataatg gttccatccg gcccaggggc gtagttaccc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttaccgca
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca ggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgctc gccgccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcggc cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc
 2161 agacgggcca gggatcatgtc tttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctcctg
 2281 ctggtgctga agcgtgccc gtcttcgccc tgcgcgtcgg
 ccaggtagca tttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctgca ttccacgagc
 2521 cagggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt tttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagtta aacgaattca
 atagcttggt gcatgggcgg
 2701 cgatataaaa tgcaagggtc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaag
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt tttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc

2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcaggttgat tcatcggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggaggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aagggccaaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgcccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaategctc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tcccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg
 3601 tttccacgcc ccgcgccacg tcacaaactc caccocctca
 ttatcatatt ggcttcaatc
 3661 caaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgtc
 3781 cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt tttcccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgcgcgtt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcgggt gtaggtcgtt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caaccggta agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt

4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggttttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gccccagtgc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtgggtcc
 4921 tgcaacttta tccgcctcca tccagtctat taattggtgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg caaagagaaa
 5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttcgcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg
 gttcttttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggcgtt ctttgccgag ctgtgctcga cgttgtcact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatoga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg

5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcgact gtggccgggt
 6001 ggggtgtggcg gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcacgcgc ttctatcgcc ttcttgacga gttcttctga
 attttggttaa aatttttgtt
 6181 aaatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtggtgttc cagtttgga
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggag ccccgattt agagcttgac ggggaaagcc
 ggcaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag
 ggcgctcca ttcgccattc
 6601 aggatcgaat taattcttaa t

SEQ ID: 16 shuttle Vector HA, CMV-luc

taacatcatc aataatatac cttattttgg attgaagcca atatgataat
 gaggggggtg agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
 acgtagtagt gtggcggaag tgtgatgttg caagtgtggc ggaacacatg
 taagcgacgg atgtggcaaa agtgacgttt ttggtgtgag ccggtgtaca
 caggaagtga caattttcgc gcggttttag gcggatgttg tagtaaattt
 gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa ctgaataaga
 ggaagtgaat tctgaataat tttgtgttac tcatagcgcg
 taatact**GCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGC**
CAGGGTTTTCCCAGTCACGACGTTGTAAACGACGGCCAGTGAATTGTAATACGACTCACTA
TAGGGCGAATTGGGTACTGGCCACagagcttggcccattgcatacgttgtatccatatcata
atatgtacattttatattggctcatgtccaacattaccgccatgttgacattgattattgact
agttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttcgcgct
tacataacttacggtaaatggccgcctggctgaccgcccacgacccccgcccattgacgt
caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
gagtatttacggtaaaactgccacttggcagtacatcaagtgtatcatatgccaagtacgcc
ccctattgacgtcaatgacggtaaatggccgcctggcattatgccagttacatgaccttat
gggactttcctacttggcagttacatctacgtattagtcacgtattaccatgggtgatgcgg
ttttggcagttacatcaatgggcgtggatagcggtttgactcacggggattttccaaagtctcca
ccccattgacgtcaatgggagttttgtttggcaccâaaatcaacgggactttccaaaatgtc
gtaacaactccgccccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatata
agcagagctcgttttagtgaaccgtcagatcgcttgagacgcatccacgctgttttgacct
ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtgggtgag
gcctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaactgggc
atgtggagacagagaagactcttgggtttctgataggcactgactctctctgcctattgggtc

tatccccccacccttaggctgctggtctgagcctagGAGATCTCTCGAGGTCGACGGTATcg
atgccca ccATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCT

51 AGTGAAGAGC GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAG
101 AGCAGGTCGA TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA
151 GACATCTTGG AAAAGACCCA CAACGGCAA CTTTGCAGACC TGGATGGAGT
201 GAAGCCCCTG ATCCTCCGGG ACTGTTTCACT CGCTGGTTGG CTGCTCGGGA
251 ACCCTATGTG TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG
301 GAGAAGGCTA ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA
351 TTACGAGGAA CTGAAACACC TGTGTCTAG AATTAACCAC TTTGAAAAGA
401 TACAGATTAT ACCCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC
451 GTTAGCTCAG CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT
501 GGTCTGGTTA ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT
551 ATAACAATAC CAATCAGGAG GACCTGCTCG TGTGTGGGG TATCCATCAC
601 CCGAACGACG CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA
651 CATCAGTATT GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG
701 CAACACGGAG CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG
751 ACAATCTTAA AACCCAACGA TGCATAAAT TTCGAGTCCA ATGGCAATTT
801 CATCGCCCCT GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA
851 TTATGAAGTC CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA
901 CCAATGGGAG CCATTAATAG CTCAATGCC TTCCATAATA TTCATCCATT
951 GACCATTGGG GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCTCTG
1001 CAACCGGTCT AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT
1051 GGCCTGTTTG GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT
1101 GGTCGATGGT TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT
1151 ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
1201 AAGGTTAATA GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTTGG
1251 GAGGGAGTTT AACAACTCTG AACGCCGGAT CGAAAATCTG AATAAGAAAA
1301 TGGAAGACGG CTTCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
1351 CTCATGGAGA ACGAGAGGAC CCTGGATTTC CACGATAGCA ACGTGAAGAA
1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGG
1451 GGAATGGATG CTTGAGTTT TACCACAAAT GTGACAATGA GTGCATGGAA
1501 AGTATACGCA ACGGGACCTA CAATTACCT CAGTATAGCG AAGAGGCTCG
1551 GCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACAT
1601 ATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
1651 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTCCAATG GAAGCCTTCA
1701 GTGCCGTATT TGTATATGAg c

GGCCGCCCTATTCTATAGTGTCACCTAAATGCTA**GAGCTC**GCTGATCAGCCTCGACTGTGCC
TT

CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCAC

TGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
TGGGGGGT

GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCT

CTATGGCTTCTGAGGCGGAAAGAAC**CAAAGCT**taacatcatc aataatatac
cttatttttgg attgaagcca atatgataat gaggggggtgg

61 agtttgtgac gtggcgcggg gcgtgggaac ggggcgggtg
acgtagtagt gtggcggaag

121 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg
atgtggcaaa agtgacgttt

181 ttggtgtgcg cgggtgtaca caggaagtga caattttcgc
 gcggttttag gcgcatgttg
 241 tagtaaattt gggcgtaacc gagtaagatt tggccatttt
 cgcgggaaaa ctgaataaga
 301 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg
 taatactgta atagtaatca
 361 attacgggggt cattagttca tagcccatat atggagttcc
 gcgttacata acttacggta
 421 aatggcccg cgtggtgacc gcccaacgac ccccgcccat
 tgacgtcaat aatgacgtat
 481 gttcccatag taacgccaat agggactttc cattgacgtc
 aatgggtgga gtattttacgg
 541 taaactgcc acttggcagt acatcaagtg tatcatatgc
 caagtacgcc cctattgac
 601 gtcaatgacg gtaaattggcc cgcctggcat tatgcccagt
 acatgacctt atgggacttt
 661 cctacttggc agtacatcta cgtattagtc atcgctatta
 ccatggtgat gcggttttgg
 721 cagtacatca atgggcgtgg atagcggttt gactcacggg
 gatttccaag tctccacccc
 781 attgacgtca atgggagttt gttttggcac caaatcaac
 gggactttcc aaaatgtcgt
 841 aacaactccg cccattgac gcaaatgggc ggtaggcgtg
 tacggtggga ggtctatata
 901 agcagagctg gtttagtgaa ccgtcagatc
 cgctag**AGATCTGG**gaaacgatatgggctgaatac**ggatccc**gtattcagcccatatcgtttc
T CT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT GGC
GG CCGCtgcagcct AAGCTTctag ataagatatc cgatccaccg gatctagata
 actgatcata
 1021 atcagccata ccacatttgt agaggtttta ctigtcttta
 aaaacctccc acacctcccc
 1081 ctgaacctga aacataaaat gaatgcaatt gttgttggtta
 acttgtttat tgcagcttat
 1141 aatgggttaca aataaagcaa tagcatcaca aatttcacaa
 ataaagcatt tttttcactg
 1201 cattctagtt gtggtttgtc caaactcatc aatgtatctt
 aacgcggatc tgggcgtggg
 1261 taagggtggg aaagaatata taagggtggg gtcttatgta
 gtttt**gtatc tgttttgcag**
 1321 **cagcgcgcgc** cgccatgagc accaactcgt ttgatggaag cattgtg
AGCTTgtcgactcg aagatctggg cgtgggttaag
 421 ggtgggaaag aatatataag gtgggggtct tatgtagttt
 tgtatctggt ttgcagcagc
 481 cgccgccgcc atgagcacca actcgtttga tggaagcatt
 gtgagctcat atttgacaac
 541 gcgcatgcc ccatgggccc ggggtgcgtca gaatgtgatg
 ggctccagca ttgatggtcg
 601 ccccgctctg cccgcaaact ctactacctt gacctacgag
 accgtgtctg gaacgccgtt

661 ggagactgca gcctccgccc ccgcttcagc cgctgcagcc
 accgcccgcg ggattgtgac
 721 tgactttgct ttcctgagcc cgcttgcaag cagtgcagct
 tcccgttcat ccgcccgcga
 781 tgacaagttg acggctcttt tggcacaatt ggattctttg
 acccggggaac ttaatgtcgt
 841 ttctcagcag ctggttgatc tgcgccagca ggtttctgcc
 ctgaaggctt cctcccctcc
 901 caatgcggtt taaaacataa ataaaaaacc agactctggt
 tggattttgga tcaagcaagt
 961 gtcttgctgt ctttatttag gggttttgcg cgcgcggtag
 gcccgggacc agcgggtctcg
 1021 gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa
 aggtgactct ggatgttcag
 1081 atacatgggc ataagcccggt ctctgggggtg gaggtagcac
 cactgcagag cttcatgctg
 1141 cgggggtggtg ttgtagatga tccagtcgta gcaggagcgc
 tgggcgtggt gcctaaaaat
 1201 gtctttcagt agcaagctga ttgccagggg caggcccttg
 gtgtaagtgt ttacaaagcg
 1261 gttaagctgg gatgggtgca tacgtgggga tatgagatgc
 atcttggact gtatttttag
 1321 gttggctatg ttcccagcca tatccctccg gggattcatg
 ttgtgcagaa ccaccagcac
 1381 agtgtatccg gtgcacttgg gaaatttgct atgtagctta
 gaaggaaatg cgtggaagaa
 1441 cttggagacg cccttgtgac ctccaagatt ttccatgcat
 tcgtccataa tgatggcaat
 1501 gggcccacgg gcggcggcct gggcgaagat atttctggga
 tcactaacgt catagtgtg
 1561 ttccaggatg agatcgatcat aggccatttt taaaaagcgc
 gggcggagggtg tgccagactg
 1621 cggtataatg gttccatccg gccagggggc gtagttaccc
 tcacagattt gcatttccca
 1681 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg
 gcgatgaaga aaacggtttc
 1741 cggggtaggg gagatcagct gggaagaaag caggttcctg
 agcagctgcg acttacgcga
 1801 gccggtgggc ccgtaaatca cacctattac cgggtgcaac
 tggtagttaa gagagctgca
 1861 gctgccgtca tccctgagca gggggggccac ttcgttaagc
 atgtccctga ctgcgatgtt
 1921 ttccctgacc aaatccgcca gaaggcgtc gccgcccagc
 gatagcagtt cttgcaagga
 1981 agcaaagttt ttcaacgggt tgagaccgtc cgccgtaggc
 atgcttttga gcgtttgacc
 2041 aagcagttcc aggcgggtccc acagctcgggt cacctgctct
 acggcatctc gatccagcat
 2101 atctcctcgt ttcgcgggtt ggggcggctt tcgctgtacg
 gcagtagtcg gtgctcgtcc

2161 agacggggcca gggatcatgtc ttccacggg cgcagggtcc
 tcgtcagcgt agtctgggtc
 2221 acggtgaagg ggtgcgctcc gggctgcgcg ctggccaggg
 tgcgcttgag gctggctctg
 2281 ctgggtgctga agcgtgcgcg gtcttcgccc tgcgcgtcgg
 ccaggtagca ttgaccatg
 2341 gtgtcatagt ccagcccctc cgcggcgtgg cccttggcgc
 gcagcttgcc cttggaggag
 2401 gcgccgcacg aggggcagtg cagacttttg agggcgtaga
 gcttgggcgc gagaaatacc
 2461 gattccgggg agtaggcac cgcgccgcag gccccgcaga
 cggctctcga ttccacgagc
 2521 caggtgagct ctggccgttc ggggtcaaaa accaggtttc
 ccccatgctt ttgatgcgt
 2581 ttcttacctc tggtttccat gagccggtgt ccacgctcgg
 tgacgaaaag gctgtccgtg
 2641 tccccgtata cagacttgag agggagttaa aacgaattca
 atagcttggt gcattggcgcg
 2701 cgatataaaa tgcaagggtgc tgctcaaaaa atcaggcaaa
 gcctcgcgca aaaaagaaa
 2761 cacatcgtag tcatgctcat gcagataaag gcaggtaagc
 tccggaacca ccacagaaaa
 2821 agacaccatt ttctctcaa acatgtctgc gggtttctgc
 ataaacacaa aataaaataa
 2881 caaaaaaaca tttaaacatt agaagcctgt cttacaacag
 gaaaaacaac ccttataagc
 2941 ataagacgga ctacggccat gccggcgtga ccgtaaaaaa
 actggtcacc gtgattaaaa
 3001 agcaccaccg acagctcctc ggtcatgtcc ggagtcataa
 tgtaagactc ggtaaacaca
 3061 tcagggttgat tcatcggtca gtgctaaaaa gcgaccgaaa
 tagcccgggg gaatacatac
 3121 ccgcaggcgt agagacaaca ttacagcccc cataggagggt
 ataacaaaat taataggaga
 3181 gaaaaacaca taaacacctg aaaaaccctc ctgcctaggc
 aaaatagcac cctcccgctc
 3241 cagaacaaca tacagcgctt cacagcggca gcctaacagt
 cagccttacc agtaaaaaag
 3301 aaaacctatt aaaaaaacac cactcgacac ggcaccagct
 caatcagtca cagtgtaaaa
 3361 aaggggccaag tgcagagcga gtatatatag gactaaaaaa
 tgacgtaacg gttaaagtcc
 3421 acaaaaaaca cccagaaaac cgcacgcgaa cctacgccc
 gaaacgaaag ccaaaaaacc
 3481 cacaacttcc tcaaatcgtc acttccgttt tcccacgtta
 cgtaacttcc cattttaaga
 3541 aaactacaat tccaacaca tacaagttac tccgccctaa
 aacctacgtc acccgccccg
 3601 ttcccacgcc ccgcgccacg tcacaaactc caccacctca
 ttatcatatt ggcttcaatc

3661 caaaataagg tatattattg atgatgttaa ttaacatgca
 tggatccata tgcggtgtga
 3721 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg
 cgctcttccg cttcctcgct
 3781 cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg
 gtatcagctc actcaaaggc
 3841 ggtaatacgg ttatccacag aatcagggga taacgcagga
 aagaacatgt gagcaaaagg
 3901 ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg
 gcgtttttcc ataggctccg
 3961 cccccctgac gagcatcaca aaaatcgacg ctcaagtcag
 aggtggcgaa acccgacagg
 4021 actataaaga taccaggcgt ttccccctgg aagctccctc
 gtgcgctctc ctgttccgac
 4081 cctgccgctt accggatacc tgtccgcctt tctcccttcg
 ggaagcgtgg cgctttctca
 4141 tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt
 cgctccaagc tgggctgtgt
 4201 gcacgaacc cccgttcagc ccgaccgctg cgccttatcc
 ggtaactatc gtcttgagtc
 4261 caacccggtg agacacgact tatcgccact ggcagcagcc
 actggtaaca ggattagcag
 4321 agcgaggtat gtaggcgggtg ctacagagtt cttgaagtgg
 tggcctaact acggctacac
 4381 tagaaggaca gtatttggtg tctgcgctct gctgaagcca
 gttaccttcg gaaaaagagt
 4441 tggtagctct tgatccggca aacaaaccac cgctggtagc
 ggtgggtttt ttgtttgcaa
 4501 gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat
 cctttgatct tttctacggg
 4561 gtctgacgct cagtggaaacg aaaactcacg ttaagggatt
 ttggtcatga gattatcaaa
 4621 aaggatcttc acctagatcc ttttaaatta aaaatgaagt
 tttaaatcaa tctaaagtat
 4681 atatgagtaa acttgggtctg acagttacca atgcttaatc
 agtgaggcac ctatctcagc
 4741 gatctgtcta tttcgttcat ccatagttgc ctgactcccc
 gtcgtgtaga taactacgat
 4801 acgggagggc ttaccatctg gcccagtgcc tgcaatgata
 ccgcgagacc cacgctcacc
 4861 ggctccagat ttatcagcaa taaaccagcc agccggaagg
 gccgagcgca gaagtgggtcc
 4921 tgcaacttta tccgcctcca tccagtctat taattgttgc
 cgggaagcta gagtaagtag
 4981 ttcgccagtt aatagtttgc gcaacgttgt tgccattgct
 gcagccatga gattatcaaa
 5041 aaggatcttc acctagatcc ttttcacgta gaaagccagt
 ccgcagaaac ggtgctgacc
 5101 ccggatgaat gtcagctact gggctatctg gacaagggaa
 aacgcaagcg .caaagagaaa

5161 gcaggtagct tgcagtgggc ttacatggcg atagctagac
 tgggcggttt tatggacagc
 5221 aagcgaaccg gaattgccag ctggggcgcc ctctggtaag
 gttgggaagc cctgcaaagt
 5281 aaactggatg gctttctcgc cgccaaggat ctgatggcgc
 aggggatcaa gctctgatca
 5341 agagacagga tgaggatcgt ttogcatgat tgaacaagat
 ggattgcacg caggttctcc
 5401 ggccgcttgg gtggagaggc tattcggcta tgactgggca
 caacagacaa tcggctgctc
 5461 tgatgccgcc gtgttcgggc tgtcagcgca ggggcgccc
 gttctttttg tcaagaccga
 5521 cctgtccggt gccctgaatg aactgcaaga cgaggcagcg
 cggctatcgt ggctggccac
 5581 gacgggctt ccttgccgag ctgtgctcga cgttgctact
 gaagcgggaa gggactggct
 5641 gctattgggc gaagtgccgg ggcaggatct cctgtcatct
 caccttgctc ctgccgagaa
 5701 agtatccatc atggctgatg caatgcggcg gctgcatacg
 cttgatccgg ctacctgcc
 5761 attcgaccac caagcgaaac atcgcatcga gcgagcacgt
 actcggatgg aagccggtct
 5821 tgtcgatcag gatgatctgg acgaagagca tcaggggctc
 gcgccagccg aactgttcgc
 5881 caggctcaag gcgagcatgc ccgacggcga ggatctcgtc
 gtgacccatg gcgatgcctg
 5941 cttgccgaat atcatggtgg aaaatggccg cttttctgga
 ttcacgcact gtggccggct
 6001 ggggtgtggc gaccgctatc aggacatagc gttggctacc
 cgtgatattg ctgaagagct
 6061 tggcggcgaa tgggctgacc gcttcctcgt gctttacggt
 atcgccgctc ccgattcgca
 6121 gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga
 attttggttaa aatttttgtt
 6181 aatcagctc attttttaac caataggccg aaatcggcaa
 catcccttat aaatcaaaag
 6241 aatagaccgc gatagggttg agtgttgttc cagtttgga
 caagagtcca ctattaaaga
 6301 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
 gggcgatggc ccactacgtg
 6361 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg
 taaagctcta aatcggaacc
 6421 ctaaaggag ccccgattt agagcttgac ggggaaagcc
 ggcgaacgtg gcgagaaagg
 6481 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc
 aagtgtagcg gtcacgctgc
 6541 gcgtaaccac cacacccgcg cgcttaatgc gccgctacag
 ggcgcgtcca ttcgccattc
 6601 aggatcgaat taattcttaa t

1 taaggatccn nncctgtect cgaccgatgc ccttgagagc cttcaaccca
 gtcagctcct
 61 tccggtgggc gcggggcatg actatcgctg ccgcacttat
 gactgtcttc tttatcatgc
 121 aactcgtagg acaggtgccg gcagcgctct gggtcatttt
 cggcgaggac cgctttcgct
 181 ggagcgcgac gatgatcggc ctgtcgcttg cggatttcgg
 aatcttgcac gccctcgctc
 241 aagccttcgt cactggtccc gccaccaaac gtttcggcga
 gaagcaggcc attatcgccg
 301 gcatggcggc cgacgcgctg ggctacgtct tgctggcggt
 cgcgacgcga ggctggatgg
 361 ccttccccat tatgattctt ctcgcttccg gcggcatcgg
 gatgcccgcg ttgcaggcca
 421 tgctgtccag gcaggtagat gacgaccatc agggacagct
 tcaaggatcg ctcgcggtc
 481 ttaccagcct aacttcgatc actggaccgc tgatcgtcac
 ggcgatttat gccgcctcgg
 541 cgagcacatg gaacggggtg gcatggattg taggcgccgc
 cctatacctt gtctgcctcc
 601 ccgcgttgcg tcgcggtgca tggagccggg ccacctcgac
 ctgaatggaa gccggcggca
 661 cctcgctaac ggattcacca ctccaagaat tggagccaat
 caattcttgc ggagaactgt
 721 gaatgcgcaa accaaccctt ggcagaacat atccatcgcg
 tccgccatct ccagcagccg
 781 cacgcggcgc atctcgggca gcgttgggtc ctggccacgg
 gtgcgcata tctgtctcct
 841 gtcgttgagg acccggttag gctggcgggg ttgccttact
 ggtagcaga atgaatcacc
 901 gatacgcgag cgaacgtgaa gcgactgctg ctgcaaaacg
 tctgcgacct gagcaacaac
 961 atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa
 acgcggaagt cagcgcctg
 1021 caccattatg ttccggatct gcatcgcagg atgctgctgg
 ctaccctgtg gaacacctac
 1081 atctgtatta acgaagcgct ggcattgacc ctgagtgatt
 tttctctggg cccgccgcat
 1141 ccataccgcc agttgtttac cctcacaacg ttccagtaac
 cgggcatgtt catcatcagt
 1201 aacccgatc gtgagcatcc tctctcgttt catcggtatc
 attaccccca tgaacagaaa
 1261 ttccccctta cacggaggca tcaagtgacc aaacaggaaa
 aaaccgccct taacatggcc
 1321 cgctttatca gaagccagac attaacgctt ctggagaaac
 tcaacgagct ggacgcggat
 1381 gaacaggcag acatctgtga atcgcttcac gaccacgctg
 atgagcttta ccgcagctgc
 1441 ctgcgcgctt tcggatgatga cggtgaaaac ctctgacaca
 tgcagctccc ggagacggtc

1501 acagcttgtc tgtaagcgga tgccgggagc agacaagccc
 gtcaggggcgc gtcagcgggt
 1561 gttggcgggt gtcggggcgc agccatgacc cagtcacgta
 gcgatagcgg agtgtatact
 1621 ggcttaacta tgccggcatca gagcagattg tactgagagt
 gcaccatatg cgggtgtgaaa
 1681 taccgcacag atgcgtaagg agaaaatacc gcatcaggcg
 ctcttccgct tcctcgctca
 1741 ctgactcgct gcgctcggtc gttcggctgc ggcgagcgggt
 atcagctcac tcaaaggcgg
 1801 taatacgggt atccacagaa tcaggggata acgcaggaaa
 gaacatgtga gcaaaaggcc
 1861 agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc
 gtttttccat aggctccgcc
 1921 cccctgacga gcatcacaaa aatcgacgct caagtcagag
 gtggcgaaac ccgacaggac
 1981 tataaagata ccaggcggtt cccctggaa gctccctcgt
 gcgctctcct gttccgaccc
 2041 tgccgcttac cggatacctg tccgccttcc tcccttcggg
 aagcgtggcg ctttctcaat
 2101 gctcacgctg taggtatctc agttcgggtg aggtcgttcg
 ctccaagctg ggctgtgtgc
 2161 acgaaccccc cgttcagccc gaccgctgcg ccttatccgg
 taactatcgt cttgagtcca
 2221 acccggttaag acacgactta tcgccactgg cagcagccac
 tggtaacagg attagcagag
 2281 cgaggatatgt aggcgggtgct acagagttct tgaagtgggtg
 gcctaactac ggctacacta
 2341 gaaggacagt atttgggtatc tgcgctctgc tgaagccagt
 taccttcgga aaaagagttg
 2401 gtagctcttg atccggcaaa caaaccaccg ctggtagcgg
 tgggtttttt gtttgcaagc
 2461 agcagattac gcgcagaaaa aaaggatctc aagaagatcc
 tttgatcttt tctacgggggt
 2521 ctgacgctca gtggaacgaa aactcacgtt aagggatttt
 ggtcatgaga ttatcaaaaa
 2581 ggatcttcac ctagatcctt ttaaattaaa aatgaagttt
 taaatcaatc taaagtatat
 2641 atgagtaaac ttggtctgac agttaccaat gcttaatcag
 tgaggcacct atctcagcga
 2701 tctgtctatt tcgttcaccc atagttgcct gactccccgt
 cgtgtagata actacgatac
 2761 gggagggtt accatctggc cccagtgcgt caatgatacc
 gcgagaccca cgctcacccg
 2821 ctccagattt atcagcaata aaccagccag ccggaagggc
 cgagcgcaga agtgggtcctg
 2881 caactttatc cgcctccatc cagtctatta attggtgccg
 ggaagctaga gtaagtagtt
 2941 cgccagttaa tagtttgccg aacgttggtg ccattgctgc agccatga
 gattatcaaa

aaggatcttc acctagatcc ttttcacgta gaaagccagt ccgcagaaac
ggtgctgacc
ccggatgaat gtcagctact gggctatctg gacaagggaa aacgcaagcg
caaagagaaa
gcaggtagct tgcagtgggc ttacatggcg atagctagac tgggcggttt
tatggacagc
aagcgaaccg gaattgccag ctggggcgcc ctctggtaag gttgggaagc
cctgcaaagt
aaactggatg gctttctcgc cgccaaggat ctgatggcgc aggggatcaa
gctctgatca
agagacagga tgaggatcgt ttcgcatgat tgaacaagat ggattgcacg
caggttctcc
ggcgccttgg gtggagaggc tattcggcta tgactgggca caacagacaa
tcggctgctc
tgatgccgcc gtgttcgggc tgtcagcgca ggggcgcccg gttctttttg
tcaagaccga
cctgtccggt gccctgaatg aactgcaaga cgaggcagcg cggctatcgt
ggctggccac
gacgggcgtt ccttgccgag ctgtgctcga cgttgctact gaagcgggaa
gggactggct
gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttgctc
ctgccgagaa
agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg
ctacctgccc
attcgaccac caagcgaaac atcgcatcga gcgagcacgt actcggatgg
aagccggtct
tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg
aactgttcgc
caggctcaag gcgagcatgc ccgacggcga ggatctcgtc gtgacccatg
gcgatgcctg
cttgccgaat atcatggtgg aaaatggccg cttttctgga ttcacgact
gtggccggct
gggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg
ctgaagagct
tggcggcgaa tgggctgacc gcttcctcgt gctttacggt atcgccgctc
ccgattcgca
gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga attttgtaa
aatttttgtt
aatcagctc attttttaac caataggccg aaatcggcaa catcccttat
aaatcaaaag
aatagaccgc gataggggtt agtggtgttc cagtttgga caagagtcca
ctattaaaga
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc
ccactacgtg
aaccatcacc caaatcaagt tttttgcggt cgagggtgccg taaagctcta
aatcgggaacc
ctaaagggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg
gcgagaaagg
aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg
gtcacgctgc

90

ATGCAGCAGA TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC
 AAGGTTAATA GTATTATAGA TAAAATGAAC ACGCAATTCG AGGCCGTTGG
 GAGGGAGTTT AACAACTCTGG AACGCCGGAT CGAAAATCTG AATAAGAAAA
 TGGAAGACGG CTTCTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA
 CTCATGGAGA ACGAGAGGAC CCTGGATTTC CACGATAGCA ACGTGAAGAA
 1401 CCTTTACGAC AAGGTGAGAC TTCAGCTCCG AGACAACGCC
 AAGGAGCTGGGGAATGGATG CTTCGAGTTT TACCACAAAT GTGACAATGA
 GTGCATGGAAAGTATACGCA ACGGGACCTA CAATTACCCT CAGTATAGCG
 AAGAGGCTCGGCTCAAACGC GAAGAGATAA GCGGGGTGAA ATTGGAATCA
 ATCGGAACATATCAAATCCT GTCCATCTAT TCCACCGTCG CCTCTTCGCT
 GGCCTCGCT
 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCGAATG
 GAAGCCTTCAGTGCCGTATT TGTATATGA_g c
GGCCGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
 TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCAC
 TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAAC**CAAGCT**taacatcatc aataatatac
 cttattttgg attgaagcca atatgataat gagggggtgg
 agtttgtgac gtggcgcggg gcgtgggaac gggcggggtg acgtagtagt
 gtggcggaag
 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg atgtggcaaa
 agtgacgttt
 ttggtgtgcg ccggtgtaca caggaagtga caattttcgc gcggttttag
 gcggatgttg
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
 ctgaataaga
 ggaagtgaat tctgaataat tttgtgttac tcatagcgcg taatactgta
 atagtaatca
 attacggggt cattagttca tagcccatat atggagttcc gcgttacata
 acttacggta
 aatggcccg cgtggtgacc gcccaacgac ccccgcccat tgacgtcaat
 aatgacgtat
 gttcccatag taacgccaat agggaacttt cattgacgtc aatgggtgga
 gtattttacgg
 taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc
 ccctattgac
 gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt
 atgggacttt
 cctacttggc agtacatcta cgtattagtc atcgtatta ccatggtgat
 gcggttttgg
 cagtacatca atgggctgg atagcggttt gactcacggg gatttccaag
 tctccacccc
 attgacgtca atgggagttt gttttggcac caaatcaac gggactttcc
 aaaatgtcgt
 aacaactccg cccattgac gcaaatgggc ggtaggcgtg tacggtggga
 ggtctatata

agcagagctg gtttagtgaa ccgtcagatc
 cgctag**AGATCTGG**gaaacgatatgggctgaatac**ggatcc**gtattcagcccatatcgtttc
TCT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT **GGC**
GG CCGCtcgagcct **AAGCTT**ctag ataagatatc cgatccaccg gatctagata
 actgatcata
 atcagccata ccacatttgt agagggttta cttgctttaa aaaacctccc
 acacctcccc
 ctgaacctga aacataaaat gaatgcaatt gttgttgta acttgtttat
 tgcagcttat
 aatggttaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt
 tttttcactg
 cattctagtt gtggtttgc caaactcatc aatgtatctt aacgcggatc
 tgggcgtggt
 taagggtggg aaagaatata taagggtggg gtcttatgta gtttt**gtatc**
tgttttgcag
cagccgccgc cgccatgagc accaactcgt ttgatggaag cattgtg
AGCTTgtcgactcg aagatctggg cgtgggttaag
 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctggt
 ttgcagcagc
 cgccgcccgc atgagcacca actcgtttga tggaagcatt gtgagctcat
 atttgacaac
 gcgcatgcc ccattgggccc ggggtgcgtca gaatgtgatg ggctccagca
 ttgatggctg
 ccccgctctg ccgcgaaact ctactacctt gacctacgag accgtgtctg
 gaacgcccgtt
 ggagactgca gcctccgccc cgcttcagc cgctgcagcc accgcccgcg
 ggattgtgac
 tgactttgct ttctgagcc cgcttgcaag cagtgcagct tcccgttcat
 ccgcccgcga
 tgacaagttg acggctcttt tggcacaatt ggattctttg acccggggaa
 ttaatgtcgt
 ttctcagcag ctggttgatc tgcgccagca ggtttctgcc ctgaaggctt
 cctccccctc
 caatgcggtt taaaacataa ataaaaaacc agactctggt tggatttgga
 tcaagcaagt
 gtcttgctgt ctttatcttag gggttttgcg cgcgcggtag gcccgggacc
 agcgggtctcg
 gtcgttgagg gtccgtgtga ttttttccag gacgtggtaa aggtgactct
 ggatgttcag
 atacatgggc ataagcccgt ctctggggtg gaggtagcac cactgcagag
 cttcatgctg
 cggggtggtg ttgtagatga tccagtcgta gcaggagcgc tgggcgtggt
 gcctaaaaat
 gtctttcagt agcaagctga ttgccagggg caggcccttg gtgtaagtgt
 ttacaaagcg
 gttaagctgg gatgggtgca tacgtgggga tatgagatgc atcttggaact
 gtatttttag
 gttggctatg ttcccagcca tatccctccg gggattcatg ttgtgcagaa
 ccaccagcac

agtgtatccg gtgcacttgg gaaatttgtc atgtagctta gaaggaaatg
 cgtggaagaa
 cttggagacg cccttgtgac ctccaagatt ttccatgcat tcgtccataa
 tgatggcaat
 gggccacagg gcggcgccct gggcgaagat atttctggga tctaactg
 catagtgtg
 ttccaggatg agatcgatc aggccatttt tacaagcgc gggcggaggg
 tgccagactg
 cggataatg gttccatccg gccaggggc gtagttacc tcacagattt
 gcatttccca
 cgctttgagt tcagatgggg ggatcatgtc tacctgcggg gcgatgaaga
 aaacggtttc
 cgggtaggg gagatcagct gggaagaaag caggttcctg agcagctgcg
 acttaccgca
 gccggtgggc ccgtaaata cacctattac cgggtgcaac tggtagttaa
 gagagct
g cagctgccg catccctgag caggggggccc
 acttcgttaa gcatgtccct gactcgcag ttttccctga ccaaatccgc
 cagaaggcgc
 tcgcccga gcatagcag ttcttgcaag gaagcaaagt ttttcaacgg
 tttgagaccg
 tccgccgtag gcatgctttt gagcgtttga ccaagcagtt ccaggcggtc
 ccacagctcg
 gtcacctgct ctacggcatc tcgatccagc atatctctc gtttcgcggg
 ttggggcggc
 tttcgctgta cggcagtagt cggtgctcgt ccagacgggc cagggtcag
 tctttccacg
 ggcgcagggt cctcgtcagc gtagtctggg tcacggtgaa ggggtgcgct
 ccgggctgcg
 cgctggccag ggtgcgcttg aggctgggtcc tgctgggtgct gaagcgctgc
 cgggtcttcgc
 cctgcgcgct gccaggtag catttgacca tggtgtcata gtccagcccc
 tccgcggcgt
 ggccttggc gcgcagcttg cccttgagg aggcgcgcga cgaggggcag
 tgcagacttt
 tgagggcgta gagcttgggc gcgagaaata ccgattccgg ggagtaggca
 tccgcgcgcg
 agccccgca gacgggtctcg cattccacga gccaggtgag ctctggccgt
 tcgggggtcaa
 aaaccaggtt tcccccatgc tttttgatgc gtttcttacc tctggtttcc
 atgagccggt
 gtccacgctc ggtgacgaaa aggctgtccg tgtccccgta tacagacttg
 agaggcctgt
 cctcgagcgg tgttccgcgg tcctcctcgt atagaaactc ggaccactct
 gagacaaagg
 ctgcgtcca gccagcacg aaggaggcta agtgggaggg gtagcggctc
 ttgtccacta
 ggggggtccac tcgctccagg gtgtgaagac acatgtcgcc ctcttcggca
 tcaaggaagg
 tgattggttt gtaggtgtag gccacgtgac cgggtgttcc tgaagggggg
 ctataaaagg

gggtgggggc gcgttcgtcc tcaactctctt ccgcatcgct gtctgcgagg
gccagctggt

ggggtgagta ctccctctga aaagcgggca tgacttctgc gctaagattg
tcagtttcca

6301 aaaacgagga ggatttgata ttcacctggc ccgcggtgat
gcctttgagg gtggccgcat

6361 ccatctggtc agaaaagaca atctttttgt tgtcaagctt
ggtggcaaac gacccgtaga

6421 gggcgttgga cagcaacttg gcgatggagc gcagggtttg
gtttttgtcg cgatcggcgc

6481 gctccttggc cgcgatgttt agctgcacgt attcgcgcg
aacgcacgc cattcgggaa

6541 agacggtggt gcgctcgctc ggcaccagggt gcacgcgcca
accgcggttg tgcagggtga

6601 caaggtcaac gctggtggct acctctccgc gtaggcgctc
gttggtccag cagaggcggc

6661 cgcccttgcc cgagcagaat ggcggtagggt ggtctagctg
cgtctcgctc ggggggtctg

6721 cgtccacggt aaagaccccg ggcagcaggc gcgcgtcgaa
gtagtctatc ttgcatcctt

6781 gcaagtctag cgctgctgc catgcgcggt cggaagcgc
gcgctcgat gggttgagt

6841 ggggacccca tggcatgggt tgggtgagcg cggaggcgta
catgccgcaa atgtcgtaaa

6901 cgtagagggg ctctctgagt attccaagat atgtagggta
gcatcttcca ccgcggatgc

6961 tggcgcgcac gtaatcgat agttcgtgcg agggagcgag
gaggtcggga ccgaggttgc

7021 tacgggaggg ctgctctgct cggaagacta tctgcctgaa
gatggcatgt gagttggatg

7081 atatggttgg acgctggaag acgttgaagc tggcgtctgt
gagacctacc gcgtcacgca

7141 cgaaggaggc gtaggagtcg cgcagcttgt tgaccagctc
ggcggtgacc tgcacgtcta

7201 gggcgagta gtccagggtt tccttgatga tgtcatactt
atcctgtccc ttttttttcc

7261 acagctcgcg gttgaggaca aactcttcgc ggtctttcca
gtactcttgg atcggaacc

7321 cgtcggcctc cgaacggtaa gagcctagca tgtagaactg
gttgacggcc tggtaggcgc

7381 agcatccctt ttctacgggt agcgcgtatg cctgcgcggc
cttccggagc gaggtgtggg

7441 tgagcgcaaa ggtgtccctg accatgactt tgaggtactg
gtatttgaag tcagtgtcgt

7501 cgcattccgc ctgctcccag agcaaaaagt ccgtgcgctt
tttggaaacgc ggatttgga

7561 gggcgaagggt gacatcggtg aagagtatct ttcccgcgcg
aggcataaag ttgcgtgtga

7621 tgcggaagggt tcccggcacc tcggaacgggt tgtaattac
ctgggogggc agcacgatct

7681 cgtcaaagcc gttgatgttg tggcccacaa tgtaaagttc
 caagaagcgc gggatgccct
 7741 tgatggaagg caatttttta agttcctcgt aggtgagctc
 ttcaggggag ctgagcccgt
 7801 gctctgaaag ggcccagtc t gcaagatgag ggttgaagc
 gacgaatgag ctccacaggt
 7861 cacggggccat tagcatttgc aggtggtcgc gaaaggctct
 aaactggcga cctatggcca
 7921 ttttttctgg ggtgatgcag tagaaggtaa gcgggtcttg
 ttcccagcgg tcccatccaa
 7981 ggttcgcggc taggtctcgc gcggcagtc ctagaggctc
 atctccgcgc aacttcatga
 8041 ccagcatgaa gggcacgagc tgcttccaa agggcccat
 ccaagtatag gtctctacat
 8101 cgtaggtgac aaagagacgc tcggtgcgag gatgcgagcc
 gatcggaag aactggatct
 8161 cccgccacca attggaggag tggctattga tgtggtgaaa
 gtagaagtcc ctgcgacggg
 8221 ccgaacactc gtgctggctt ttgtaaaaac gtgcgcagta
 ctggcagcgg tgcacgggct
 8281 gtacatcctg cacgaggttg acctgacgac cgcgcacaa
 gaagcagagt gggaatttga
 8341 gccctcgc tggcgggtt ggctgggtgt cttctacttc
 ggctgcttgt ccttgaccgt
 8401 ctggctgctc gaggggagtt acggtggatc ggaccaccac
 gccgcgcgag ccaaagtcc
 8461 agatgtccgc gcgcggcgg t cggagcttga tgacaacatc
 gcgcagatgg gagctgtcca
 8521 tggcttgag ctcccgcggc gtcaggtcag gcgggagctc
 ctgcaggttt acctcgcata
 8581 gacgggtcag ggcgcgggct agatccaggt gataccta
 ttccaggggc tggttggtg
 8641 cggcgtcgat ggcttgcaag aggccgcac cccgcggcgc
 gactacggta ccgcgcggc
 8701 ggcggtggc gcggggggtg tccttgatg atgcatctaa
 aagcggtgac gcgggcgagc
 8761 ccccgaggt agggggggct ccggaccgc cgggagaggg
 ggcaggggca cgtcggcgc
 8821 gcgcgcggc aggagctgg tctgcgcgc taggttgctg
 gcgaacgcga cgacgcggc
 8881 gttgatctcc tgaatctggc gcctctgcgt gaagacgacg
 ggcccgtga gcttgagcct
 8941 gaaagagagt tcgacagaat caatttcggt gtcgttgacg
 gcggcctggc gcaaaatctc
 9001 ctgcacgtct cctgagttgt cttgataggc gatctcggcc
 atgaactgct cgatctcttc
 9061 ctcttgaga tctccgcgtc cggctcgtc cacggtggcg
 gcgaggtcgt tggaatgcg
 9121 ggccatgagc tgcgagaagg cgttgaggcc tccctcgctc
 cagacgcggc ttagaccac

9181 gcccccttcg gcatcgcggg cgcgcatgac cacctgcgcg
 agattgagct ccacgtgccg
 9241 ggcgaagacg gcgtagtttc gcaggcgctg aaagaggtag
 ttgagggtgg tggcgggtgtg
 9301 ttctgccacg aagaagtaca taaccacgcg tcgcaacgtg
 gattcgttga tatcccccaa
 9361 ggctcaagg cgctccatgg cctcgtagaa gtccacggcg
 aagttgaaaa actgggagtt
 9421 gcgcgccgac acggttaact cctcctccag aagacggatg
 agctcggcga cagtgtcgcg
 9481 cacctcgcg ctaaaggcta caggggcctc ttcttcttct
 tcaatctcct ctcccataag
 9541 ggctcccct tcttcttctt ctggcgggcg tgggggaggg
 gggacacggc ggcgacgacg
 9601 gcgcaccggg aggcggtcga caaagcgctc gatcatctcc
 ccgcggcgac ggcgcatggt
 9661 ctcggtgacg gcgcggccgt tctcgcgggg gcgcagttgg
 aagacgccgc ccgtcatgtc
 9721 ccggttatgg gttggcgggg ggctgccatg cggcagggat
 acggcgctaa cgatgcatct
 9781 caacaattgt tgtgtaggta ctccgccgcc gagggacctg
 agcgagtccg catcgaccgg
 9841 atcggaaaac ctctcgagaa aggcgtctaa ccagtcacag
 tcgcaaggta ggctgagcac
 9901 cgtggcgggc ggcagcgggc ggcggtcggg gttgtttctg
 gcggaggtgc tgctgatgat
 9961 gtaattaaag taggcggtct tgagacggcg gatggtcgac
 agaagcacca tgccttggg
 10021 tccggcctgc tgaatgcgca ggcggtcggc catgccccag
 gcttcgtttt gacatcggcg
 10081 caggtctttg tagtagtctt gcatgagcct ttctaccggc
 acttcttctt ctcttctc
 10141 ttgtcctgca tctcttgcac ctatcgctgc ggcggcgggc
 gagtttgcc gtaggtggcg
 10201 ccctcttct cccatgcgtg tgaccccgaa gccctcctc
 ggctgaagca ggcctaggtc
 10261 ggcgacaacg cgctcggcta atatggcctg ctgcacctgc
 gtgagggtag actggaagtc
 10321 atccatgtcc acaaagcggg ggtatgcgcc cgtgttgatg
 gtgtaagtgc agttggccat
 10381 aacggaccag ttaacggtct ggtgaccgg ctgcgagagc
 tcggtgtacc tgagacgcga
 10441 gtaagccctc gagtcaaata cgtagtcgtt gcaagtccgc
 accaggtact ggtatccac
 10501 caaaaagtgc ggcggcggt ggcggtagag gggccagcgt
 aggggtggccg gggctccggg
 10561 ggcgagatct tccaacataa ggcgatgata tccgtagatg
 tacctggaca tccaggtgat
 10621 gccggcgggc gtggtggagg cgcgcgaaa gtcgcgggacg
 cggttccaga tgttgcgcag

10681 cggcaaaaag tgctccatgg tcgggacgct ctggccgggtc
 aggcgcgcgc aatcgttgac
 10741 gctctaccgt gcaaaaggag agcctgtaag cgggcactct
 tccgtggtct ggtggataaa
 10801 ttcgcaaggg tatcatggcg gacgaccggg gttcgagccc
 cgtatccggc cgtccgccgt
 10861 gatccatgcg gttaccgccc gcggtgcgaa cccaggtgtg
 cgacgtcaga caacggggga
 10921 gtgctccttt tggcttcctt ccaggcgcgg cggctgctgc
 gctagctttt ttggccactg
 10981 gccgcgcgca gcgtaagcgg ttaggctgga aagcgaaagc
 attaagtggc tcgctccctg
 11041 tagccggagg gttattttcc aagggttgag tcgcgggacc
 cccggttcga gtctcggacc
 11101 ggccggactg cggcgaaagg gggtttgctt ccccgctcatg
 caagaccccc cttgcaaatt
 11161 cctccggaaa cagggacgag cccctttttt gcttttccca
 gatgcatccg gtgctgcggc
 11221 agatgcgccc ccctcctcag cagcggcaag agcaagagca
 gcggcagaca tgcagggcac
 11281 cctccccctc tcctaccgcg tcaggagggg cgacatccgc
 ggttgacgcg gcagcagatg
 11341 gtgattacga acccccgcgg cgccggggccc ggcactacct
 ggacttgag gagggcgagg
 11401 gcctggcgcg gctaggagcg ccctctcctg agcggtagcc
 aagggtgcag ctgaagcgtg
 11461 atacgcgtga ggcgtacgtg ccgcggcaga acctgtttcg
 cgaccgcgag ggagaggagc
 11521 ccgaggagat gcgggatcga aagtccacg cagggcgcga
 gctgcggcat ggcctgaatc
 11581 gcgagcgggt gctgcgcgag gaggactttg agcccgacgc
 gcgaaccggg attagtcccc
 11641 cgcgcgca cgtggcgggc gccgacctgg taaccgcata
 cgagcagacg gtgaaccagg
 11701 agattaactt tcaaaaaagc ttttaacaacc acgtgcgtac
 gcttgtggcg cgcgaggagg
 11761 tggctatagg actgatgcat ctgtgggact ttgtaagcgc
 gctggagcaa aaccctaaata
 11821 gcaagccgct catggcgag ctgttcctta tagtgcagca
 cagcaggac aacgaggcat
 11881 tcagggatgc gctgctaaac atagtagagc ccgagggccg
 ctggctgctc gatttgataa
 11941 acatcctgca gagcatagtg gtgcaggagc gcagcttgag
 cctggctgac aagggtggcg
 12001 ccatcaacta ttccatgctt agcctgggca agttttacgc
 ccgcaagata taccataccc
 12061 cttacgttcc catagacaag gaggtaaaga tcgaggggtt
 ctacatgcgc atggcgctga
 12121 aggtgcttac cttgagcgac gacctgggog tttatcgcaa
 cgagcgcac cacaaggccg

12181 tgagcgtgag cgggcggcgc gagctcagcg accgcgagct
 gatgcacagc ctgcaaaggg
 12241 ccctggctgg cacgggcagc ggcgatagag aggccgagtc
 ctactttgac gcggggcgctg
 12301 acctgcgctg ggccccaagc cgacgcgccc tggaggcagc
 tggggccgga cctgggctgg
 12361 cgggtggcacc cgcgcgcgct ggcaacgtcg gcggcgctgga
 ggaatatgac gaggacgatg
 12421 agtacgagcc agaggacggc gagtactaag cggtgatggt
 tctgatcaga tgatgcaaga
 12481 cgcaacggac ccggcggtgc gggcggcgct gcagagccag
 ccgtccggcc ttaactccac
 12541 ggacgactgg cgccaggtca tggaccgcat catgtcgctg
 actgcgcgca atcctgacgc
 12601 gttccggcag cagccgcagg ccaaccggct ctccgcaatt
 ctggaagcgg tgggtccggc
 12661 gcgcgcaaac ccacgcacg agaaggtgct ggcgatcgta
 aacgcgctgg ccgaaaacag
 12721 ggccatccgg ccgcagcagg ccggcctggt ctacgacgcg
 ctgcttcagc gcgtggctcg
 12781 ttacaacagc ggcaacgtgc agaccaacct ggaccggctg
 gtgggggatg tgcgcgaggc
 12841 cgtggcgagc cgtgagcgcg cgcagcagca gggcaacctg
 ggctccatgg ttgcactaaa
 12901 cgccttcctg agtacacagc ccgccaacgt gccgcgggga
 caggaggact acaccaactt
 12961 tgtgagcgca ctgcggctaa tggtgactga gacaccgcaa
 agtgagggtg accagtctgg
 13021 gccagactat tttttccaga ccagtagaca aggcctgcag
 accgtaaacc tgagccaggc
 13081 tttcaaaaac ttgcaggggc tgtgggggggt gcgggctccc
 acaggcgacc gcgcgaccgt
 13141 gtctagcttg ctgacgcca actcgcgcct gttgctgctg
 ctaatagcgc ccttcacgga
 13201 cagtggcagc gtgtcccggg acacatacct aggtcacttg
 ctgacactgt accgcgaggc
 13261 cataggtcag gcgcatgtgg acgagcatac tttccaggag
 attacaagtg tcagccgcgc
 13321 gctggggcag gaggacacgg gcagcctgga ggcaacccta
 aactacctgc tgaccaaccg
 13381 gcggcagaag atcccctcgt tgcacagttt aaacagcgag
 gaggagcgca ttttgcgcta
 13441 cgtgcagcag agcgtgagcc ttaacctgat gcgcgacggg
 gtaacgcca gcgtggcgct
 13501 ggacatgacc gcgcgcaaca tggaaccggg catgtatgcc
 tcaaaccggc cgtttatcaa
 13561 ccgcctaattg gactacttgc atcgcgcggc cgccgtgaac
 cccgagtatt tcaccaatgc
 13621 catcttgaac ccgcactggc taccgcccc tggtttctac
 accgggggat tcgaggtgcc

13681 cgagggtaac gatggattcc tctgggacga catagacgac
 agcgtgtttt ccccgcaacc
 13741 gcagaccctg ctagagttgc aacagcgcgga gcaggcagag
 gcggcgctgc gaaaggaaag
 13801 cttccgcagg ccaagcagct tgtccgatct aggcgctgcg
 gccccgcggt cagatgctag
 13861 tagcccattht ccaagcttga taggggtctct taccagcact
 cgcaccaccc gcccgcgctt
 13921 gctgggagag gaggagtacc taaacaactc gctgctgcag
 ccgcagcgcg aaaaaaacct
 13981 gcctccggca tttcccaaca acgggataga gagcctagt
 gacaagatga gtagatggaa
 14041 gacgtacgcg caggagcaca gggacgtgcc aggcccgcg
 ccgcccaccc gtcgtcaaag
 14101 gcacgaccgt cagcggggtc tgggtgtggga ggacgatgac
 tcggcagacg acagcagcgt
 14161 cctggatttg ggaggagtg gcaaccggtt tgcgcacctt
 cgccccaggc tggggagaat
 14221 gttttaaaaa aaaaaaagca tgatgcaaaa taaaaaactc
 accaaggcca tggcaccgag
 14281 cggttggtttt cttgtattcc ccttagtatg cggcgcgcg
 cgatgtatga ggaaggtcct
 14341 cctccctcct acgagagtgt ggtgagcgcg gcgccagtgg
 cggcgcgcgct gggttctccc
 14401 ttcgatgctc ccctggaccc gccgtttgtg cctccgcggt
 acctgcggcc taccgggggg
 14461 agaaacagca tccgttactc tgagttggca cccctattcg
 acaccacccg tgtgtacctg
 14521 gtggacaaca agtcaacgga tgtggcatcc ctgaactacc
 agaacgacca cagcaacttt
 14581 ctgaccacgg tcattcaaaa caatgactac agccccgggg
 aggcaagcac acagaccatc
 14641 aatcttgacg accggtcgca ctggggcggc gacctgaaaa
 ccattctgca taccaacatg
 14701 ccaaagtga acgagttcat gtttaccat aagtttaagg
 cgcggtgat ggtgtcgcgc
 14761 ttgcctacta aggacaatca ggtggagctg aaatacaggt
 ggggtggagtt cagctgccc
 14821 gagggcaact actccgagac catgaccata gaccttatga
 acaacgcgat cgtggagcac
 14881 tacttgaaag tgggcagaca gaacgggggt ctggaaagcg
 acatcggggt aaagtttgac
 14941 acccgcaact tcagactggg gtttgacccc gtcactggtc
 ttgtcatgcc tggggatatat
 15001 acaaacgaag ccttccatcc agacatcatt ttgctgccag
 gatgcgggggt ggacttcacc
 15061 cacagccgcc tgagcaactt gttgggcatc cgcaagcggc
 aacccttcca ggagggttt
 15121 aggatcacct acgatgatct ggagggtggg aacattccc
 cactgttgga tgtggacgcc

15181 taccaggcga gcttgaaaga tgacaccgaa cagggcgggg
 gtggcgagcagg cggcagcaac
 15241 agcagtggca gcggcgcgga agagaactcc aacgcggcag
 ccgcggcaat gcagccggtg
 15301 gaggacatga acgatcatgc cattcgcggc gacacctttg
 ccacacgggc tgaggagaag
 15361 cgcgctgagg ccgaagcagc ggccgaagct gccgccccg
 ctgcgcaacc cgaggctcag
 15421 aagcctcaga agaaaccggt gatcaaacc ctgacagagg
 acagcaagaa acgcagttac
 15481 aacctaataa gcaatgacag caccttcacc cagtaccgca
 gctggtacct tgcatacaac
 15541 tacggcgacc ctgagaccgg aatccgctca tggaccctgc
 tttgcactcc tgacgtaacc
 15601 tgcggtcgg agcaggtcta ctggctcgtt ccagacatga
 tgcaagaccc cgtgacctc
 15661 cgctccacgc gccagatcag caactttccg gtgggtgggcg
 ccgagctgtt gcccgcgac
 15721 tccaagagct tctacaacga ccaggccgct tactcccaac
 tcatccgcca gtttacctc
 15781 ctgaccacg tgttcaatcg ctttcccgag aaccagattt
 tggcgcgccc gccagcccc
 15841 accatcacca ccgtcagtga aaacgttcc gctctcacag
 atcacgggac gctaccgctg
 15901 cgcaacagca tcggaggagt ccagcgagt accattactg
 acgccagacg ccgcacctgc
 15961 ccctacgttt acaaggccct gggcatagtc tcgccgcgcg
 tcctatcgag ccgcactttt
 16021 tgagcaagca tgtccatcct tatatcgccc agcaataaca
 caggctgggg cctgcgctc
 16081 ccaagcaaga tgtttggcgg ggccaagaag cgctccgacc
 aacaccagtg gcgcgtgcgc
 16141 gggcactacc gcgcgccctg gggcgcgcac aaacgcggcc
 gcactgggcg caccaccgctc
 16201 gatgaccca tcgacgcggt ggtggaggag gcgcgcaact
 acacgcccac gccgccacca
 16261 gtgtccacag tggacgcggc cattcagacc gtggtgcgcg
 gagcccggcg ctatgctaaa
 16321 atgaagagac ggcgaggcg cgtagcacgt cgccaccgcc
 gccgaccg cactgccgc
 16381 caacgcgcg gcgcggccct gcttaaccgc gcacgtcgca
 ccggccgacg ggcggccatg
 16441 cgggcccgtc gaaggctggc cgcggtatt gtcactgtgc
 ccccaggctc caggcgacga
 16501 gcggccgccc cagcagccgc ggccattagt gctatgactc
 agggctcag gggcaacgtg
 16561 tattgggtgc gcgactcggg tagcgccctg cgcggtgccg
 tgcgcacccg cccccgcgc
 16621 aactagattg caagaaaaaa ctacttagac tcgtactgtt
 gtatgtatcc agcggcggcg

16681 gcgcgcaacg aagctatgtc caagcgcaaa atcaaagaag
 agatgctcca ggtcatcgcg
 16741 ccggagatct atggccccc gaagaaggaa gagcaggatt
 acaagccccg aaagctaaag
 16801 cgggtcaaaa agaaaaagaa agatgatgat gatgaacttg
 acgacgaggt ggaactgctg
 16861 cacgctaccg cgcccaggcg acgggtacag tggaaaggtc
 gacgcgtaaa acgtgttttg
 16921 cgaccgggca ccaccgtagt ctttacgccc ggtgagcgct
 ccaccgcac ctacaagcgc
 16981 gtgtatgatg aggtgtacgg cgacgaggac ctgcttgagc
 aggccaacga gcgcctcggg
 17041 gagtttgcct acggaaagcg gcataaggac atgctggcgt
 tgccgctgga cgagggcaac
 17101 ccaacaccta gcctaaagcc cgtaacactg cagcagggtgc
 tgcccgcgct tgcaccgtcc
 17161 gaagaaaagc gcggcctaaa gcgcgagtct ggtgacttgg
 caccaccgt gcagctgatg
 17221 gtaccaagc gccagcgact ggaagatgtc ttggaaaaaa
 tgaccgtgga acctgggctg
 17281 gagcccgagg tccgcgtgcg gccaatcaag caggtggcgc
 cgggactggg cgtgcagacc
 17341 gtggacgttc agataccac taccagtagc accagtattg
 ccaccgccac agagggcatg
 17401 gagacacaaa cgtccccggt tgcctcagcg gtggcggatg
 ccgcggtgca ggcggtcgct
 17461 gcggccgcgt ccaagacctc tacggagggtg caaacggacc
 cgtggatggt tcgcgtttca
 17521 gccccccggc gcccgcgcgg ttcgaggaag tacggcgccg
 ccagcgcgct actgcccga
 17581 tatgccctac atccttccat tgccgctacc cccggctatc
 gtggctacac ctaccgcccc
 17641 agaagacgag caactaccg acgccgaacc accactggaa
 cccgcgccc cgtcgcgct
 17701 gccagcccc tgetggcccc gatttccgtg cgcagggtgg
 ctgcggaagg aggcaggacc
 17761 ctggtgctgc caacagcgcg ctaccacccc agcatcgttt
 aaaagccggt ctttgtggtt
 17821 cttgcagata tggccctcac ctgcgcctc cgtttcccgg
 tgccgggatt ccgaggaaga
 17881 atgcaccgta ggaggggcat ggccggccac ggcctgacgg
 gcggcatgcg tcgtgcgcac
 17941 caccggcggc ggcgcgcgtc gcaccgtcgc atgcgcggcg
 gtatcctgcc cctccttatt
 18001 ccaactgatcg ccgcggcgat tggcgccgtg cccggaattg
 catccgtggc cttgcaggcg
 18061 cagagacact gattaanaac aagttgcatg tggaaaaatc
 aaaataaaaa gtctggactc
 18121 tcacgctcgc ttggtcctgt aactattttg tagaatggaa
 gacatcaact ttgcgtctct

18181 ggccccgcga cacggctcgc gcccgttcat gggaaactgg
 caagatatcg gcaccagcaa
 18241 tatgagcggg ggcgccttca gctggggctc gctgtggagc
 ggcattaaaa atttcggttc
 18301 caccgttaag aactatggca gcaaggcctg gaacagcagc
 acaggccaga tgctgaggga
 18361 taagttgaaa gagcaaaatt tccaacaaaa ggtggtagat
 ggcctggcct ctggcattag
 18421 cgggggtggtg gacctggcca accaggcagt gcaaaataag
 attaacagta agcttgatcc
 18481 ccgccctccc gtagaggagc ctccaccggc cgtggagaca
 gtgtctccag aggggcgtgg
 18541 cgaaaagcgt ccgcgccccg acagggaaga aactctggtg
 acgcaaatag acgagcctcc
 18601 ctcgtagcag gaggcactaa agcaaggcct gccaccacc
 cgtcccatcg cgcccatggc
 18661 taccggagtg ctggggccagc acacacccgt aacgctggac
 ctgcctcccc ccgcccagac
 18721 ccagcagaaa cctgtgctgc caggcccagc cgccgttggt
 gtaaccgctc ctagccgcgc
 18781 gtccctgcgc cgcgcgcgca gcggtccgcg atcggtgcgg
 ccgtagcca gtggcaactg
 18841 gcaaagcaca ctgaacagca tcgtgggtct gggggtgcaa
 tccctgaagc gccgacgatg
 18901 cttctgaata gctaacgtgt cgtatgtgtg tcatgtatgc
 gtccatgtcg ccgccagagg
 18961 agctgctgag ccgcgcgcgc cccgctttcc aagatggcta
 ccccttcgat gatgccgcag
 19021 tggctttaca tgcacatctc gggccaggac gcctcggagt
 acctgagccc cgggctgggt
 19081 cagttttgcc gcgccaccga gacgtacttc agcctgaata
 acaagtttag aaaccccacg
 19141 gtggcgccta cgcacgacgt gaccacagac cgggtcccagc
 gtttgacgct gcggttcac
 19201 cctgtggacc gtgaggatac tgcgtactcg tacaaggcgc
 gggttcaccct agctgtgggt
 19261 gataaccgtg tgctggacat ggcttccacg tactttgaca
 tccgcggcgt gctggacagg
 19321 ggccctactt ttaagcccta ctctggcact gcctacaacg
 ccctggctcc caagggtgcc
 19381 ccaaatacct gcgaatggga tgaagctgct actgctcttg
 aaataaacct agaagaagag
 19441 gacgatgaca acgaagacga agtagacgag caagctgagc
 agcaaaaaac tcacgtatct
 19501 gggcaggcgc cttattcttg tataaatatt acaaaggagg
 gtattcaaat aggtgtcgaa
 19561 ggtcaaacac ctaaataatgc cgataaaaca tttcaacctg
 aacctcaaat aggagaatct
 19621 cagtggtagc aaactgaaat taatcatgca gctgggagag
 tccttaaaaa gactacccca

19681 atgaaacccat gttacggttc atatgcaaaa cccacaaatg
 aaaatggagg gcaaggcatt
 19741 cttgttaaagc aacaaaatgg aaagctagaa agtcaagtgg
 aaatgcaatt tttctcaact
 19801 actgaggcga cgcaggcaa tgggtgataac ttgactccta
 aagtgggtatt gtacagtga
 19861 gatgtagata tagaaacccc agacactcat atttcttaca
 tgcccactat taaggaagg
 19921 aactcacgag aactaatggg ccaacaatct atgcccaaca
 ggcctaatta cattgctttt
 19981 agggacaatt ttattgggtct aatgtattac aacagcacgg
 gtaatatggg tgttctggcg
 20041 ggccaagcat cgcagttgaa tgctgttgta gatttgcaag
 acagaaacac agagctttca
 20101 taccagcttt tgcttgatgc cattgggtgat agaaccagg
 acttttctat gtggaatcag
 20161 gctgttgaca gctatgatcc agatgttaga attattgaaa
 atcatggaac tgaagatgaa
 20221 cttccaaatt actgctttcc actgggagggt gtgattaata
 cagagactct taccaaggta
 20281 aaacctaata caggtcagga aaatggatgg gaaaaagatg
 ctacagaatt ttcagataaa
 20341 aatgaaataa gagttggaaa taattttgcc atggaaatca
 atctaaatgc caacctgtgg
 20401 agaaatttcc tgtactccaa catagcgctg tatttgcccg
 acaagctaaa gtacagtcct
 20461 tccaacgtaa aaattttctga taaccctaac acctacgact
 acatgaacaa gcgagtggtg
 20521 gctcccggt tagtggactg ctacattaac cttggagcac
 gctgggtccct tgactatatg
 20581 gacaacgtca acccatTTAA ccaccaccgc aatgctggcc
 tgcgctaccg ctcaatgttg
 20641 ctgggcaatg gtcgctatgt gcccttcac atccagggtgc
 ctcaagaatt ctttgccatt
 20701 aaaaacctcc ttctcctgcc gggctcatac acctacgagt
 ggaacttcag gaaggatgtt
 20761 aacatggttc tgcagagctc cctaggaaat gacctaaagg
 ttgacggagc cagcattaag
 20821 tttgatagca ttgacctta cgccacctc tccccatgg
 cccacaacac cgctccacg
 20881 cttgaggcca tgcttagaaa cgacaccaac gaccagtcct
 ttaacgacta tctctccgccc
 20941 gccacatgc tctaccctat acccgccaac gctaccaacg
 tgcccatatc catccctcc
 21001 cgcaactggg cggttttccg cggtggggcc ttcacgcgccc
 ttaagactaa ggaaacccca
 21061 tcaactgggt cggtctacga cccttattac acctactctg
 gctctatacc ctacctagat
 21121 ggaacctttt acctcaacca cacctttaag aagggtggcca
 ttacctttga ctcttctgtc

21181 agctggcctg gcaatgaccg cctgcttacc cccaacgagt
 ttgaaattaa ggcgtcagtt
 21241 gacggggagg gttacaacgt tgcccagtggt aacatgacca
 aagactgggtt cctgggtacaa
 21301 atgctagcta actacaacat tggctaccag ggcttctata
 tcccagagag ctacaaggac
 21361 cgcattgtact ccttcttttag aaacttccag cccatgagcc
 gtcagggtggt ggatgatact
 21421 aaatacaagg actaccaaca ggtgggcatc ctacaccaac
 acaacaactc tggatttggt
 21481 ggctaccttg cccccaccat ggcggaagga caggcctacc
 ctgctaactt cccctatccg
 21541 cttataggca agaccgcagt tgacagcatt acccagaaaa
 agtttctttg cgatcgcacc
 21601 ctttggcgca tcccattctc cagtaacttt atgtccatgg
 gcgcactcac agacctgggc
 21661 caaaaccttc tctacgcaa ctccgcccac gcgctagaca
 tgacttttga ggtggatccc
 21721 atggacgagc ccacccttct ttatgttttg tttgaagtct
 ttgacgtggt ccgtgtgcac
 21781 cggccgcacc gcggcgatcat cgaaaccgtg tacctgcgca
 cgcccttctc ggccggcaac
 21841 gccacaacat aaagaagcaa gcaacatcaa caacagctgc
 cgccatgggc tccagtgcac
 21901 aggaactgaa agccattgtc aaagatcttg gttgtggggc
 atattttttg ggcacctatg
 21961 acaagcgctt tccaggcttt gtttctccac acaagctcgc
 ctgcgccata gtcaatacgg
 22021 ccggtcgcga gactgggggc gtacactgga tggcctttgc
 ctggaacccg cactcaaaaa
 22081 catgctacct ctttgagccc tttggctttt ctgaccagcg
 actcaagcag gtttaccagt
 22141 ttgagtacga gtcactcctg cgccgtagcg ccattgcttc
 ttcccccgac cgctgtataa
 22201 cgctggaaaa gtccacccaa agcgtacagg ggcccaactc
 ggccgcctgt ggactattct
 22261 gctgcatgtt tctccacgcc tttgccaact ggccccaac
 tcccatggat cacaacccca
 22321 ccatgaacct tattaccggg gtacccaact ccatgctcaa
 cagtccccag gtacagccca
 22381 ccctgcgtcg caaccaggaa cagctctaca gcttcctgga
 gcgccactcg ccctacttcc
 22441 gcagccacag tgccgcagatt aggagcgcca cttctttttg
 tcacttgaaa aacatgtaaa
 22501 aataatgtac tagagacact ttcaataaag gcaaagtgtt
 ttatttgtac actctcgggt
 22561 gattatttac cccaccctt gccgtctgcg ccgtttaaaa
 atcaaagggg ttctgcgcg
 22621 catcgctatg cgccactggc agggacacgt tgcgatactg
 gtgttttagtg ctccacttaa

22681 actcaggcac aaccatccgc ggcagctcgg tgaagttttc
 actccacagg ctgcgcacca
 22741 tcaccaacgc gtttagcagg tcgggcgccg atatcttgaa
 gtcgcagttg gggcctccgc
 22801 cctgcgcgcg cgagttgcga tacacagggg tgcagcactg
 gaacactatc agcgccgggt
 22861 ggtgcacgct ggccagcacg ctcttgctcg agatcagatc
 cgcgtccagg tcctccgcgt
 22921 tgctcagggc gaacggagtc aactttggta gctgccttcc
 caaaaagggc gcgtgccag
 22981 gctttgagtt gcactcgcac cgtagtgga tcaaaagggtg
 accgtgcccg gtctgggcgt
 23041 taggatacag cgcctgcata aaagccttga tctgcttaaa
 agccacctga gcctttgcgc
 23101 cttcagagaa gaacatgccg caagacttgc cggaaaactg
 attggccgga caggccgcgt
 23161 cgtgcacgca gcaccttgcg tcgggtgttg agatctgcac
 cacatttcgg cccaccgggt
 23221 tcttcacgat cttggccttg ctagactgct ccttcagcgc
 gcgtgcccg ttttcgctcg
 23281 tcacatccat ttcaatcacg tgctccttat ttatcataat
 gcttccgtgt agacacttaa
 23341 gctgccttc gatctcagcg cagcggtgca gccacaacgc
 gcagcccgtg ggctcgtgat
 23401 gcttgtaggt cacctctgca aacgactgca ggtacgcctg
 caggaatcgc cccatcatcg
 23461 tcacaaagggt cttgttgctg gtgaagggtca gctgcaaccc
 gcggtgctcc tcgttcagcc
 23521 aggtcttgca tacggccgcc agagcttcca cttggtcagg
 cagtagtttg aagttcgcct
 23581 ttagatcgtt atccacgtgg tacttggtcca tcagcgcgcg
 cgcagcctcc atgcccttct
 23641 cccacgcaga cacgatcggc aactcagcg ggttcatcac
 cgtaatttca ctttccgctt
 23701 cgtgaggctc ttcctcttcc tcttgctcc gcataaccag
 cgccactggg tcgtcttcat
 23761 tcagccgcgc cactgtgcgc ttacctcctt tgccatgctt
 gattagcacc ggtgggttgc
 23821 tgaaaccac catttgtagc gccacatctt ctctttcttc
 ctgcgtgtcc acgattacct
 23881 ctggtgatgg cgggcgctcg ggcttgggag aagggcgctt
 ctttttcttc ttgggcgcaa
 23941 tggccaaatc cgccgccgag gtcgatggcc gcgggctggg
 tgtgcgcggc accagcgcgt
 24001 cttgtgatga gtcttcctcg tcctcggact cgatacgccg
 cctcatccgc ttttttgggg
 24061 gcgcccgggg aggcggcggc gacggggacg gggacgacac
 gtctccatg gttgggggac
 24121 gtcgcgccgc accgcgtccg cgctcggggg tggtttcgcg
 ctgctcctct tcccagactgg

24181 ccattttcctt ctcttatagg cagaaaaaga tcatggagtc
 agtcgagaag aaggacagcc
 24241 taaccgcccc ctctgagttc gccaccaccg cctccaccga
 tgccgccaac gcgcctacca
 24301 ccttccccgt cgaggcacc cgccttgagg aggaggaagt
 gattatcgag caggacccag
 24361 gttttgtaag cgaagacgac gaggaccgct cagtaccaac
 agaggataaa aagcaagacc
 24421 aggacaacgc agaggcaaac gaggaacaag tcggggcgggg
 ggacgaaagg catggcgact
 24481 acctagatgt gggagacgac gtgctgttga agcatctgca
 gcgccagtgc gccattatct
 24541 gcgacgcgtt gcaagagcgc agcgatgtgc ccctcgccat
 agcggatgtc agccttgcc
 24601 acgaacgcca cctattctca ccgcgcgtac ccccaaacg
 ccaagaaaac ggcacatgcg
 24661 agcccaaccc gcgcctcaac ttctaccccg tatttgccgt
 gccagaggtg cttgccacct
 24721 atcacatctt ttccaaaac tgcaagatac ccctatcctg
 ccgtgccaac cgcagccgag
 24781 cggacaagca gctggccttg cggcagggcg ctgtcatacc
 tgatatcgcc tcgctcaacg
 24841 aagtgccaaa aatctttgag ggtcttgagc gcgacgagaa
 gcgcgcggca aacgctctgc
 24901 aacaggaaaa cagcgaaaat gaaagtcact ctggagtgtt
 ggtggaactc gagggtgaca
 24961 acgcgcgcct agccgtacta aaacgcagca tcgagggtcac
 ccactttgcc tacccggcac
 25021 ttaacctacc cccaaggctc atgagcacag tcatgagtga
 gctgatcgtg cgcctgctgc
 25081 agcccctgga gagggatgca aatttgcaag aacaaacaga
 ggagggccta cccgcagttg
 25141 gcgacgagca gctagcgcgc tggcttcaaa cgcgcgagcc
 tgccgacttg gaggagcgac
 25201 gcaaaactaat gatggccgca gtgctcgcta ccgtggagct
 tgagtgcattg cagcggttct
 25261 ttgctgacct ggagatgcag cgcaagctag aggaaacatt
 gcactacacc ttctgacagg
 25321 gctacgtacg ccaggcctgc aagatctcca acgtggagct
 ctgcaacctg gtctcctacc
 25381 ttggaatttt gcacgaaaac cgccttgggc aaaacgtgct
 tcattccacg ctcaaggcg
 25441 aggcgcgcgc cgactacgtc cgcgactgcg tttacttatt
 tctatgctac acctggcaga
 25501 cggccatggg cgtttggcag cagtgccttg aggagtgcaa
 cctcaaggag ctgcagaaac
 25561 tgctaaagca aaacttgaag gacctatgga cggccttcaa
 cgagcgctcc gtggccgcgc
 25621 acctggcgga catcattttc cccgaacgcc tgcttaaaac
 cctgcaacag ggtctgcag

25681 acttcaccag tcaaagcatg ttgcagaact ttaggaactt
 tatecctagag cgctcaggaa
 25741 tcttgcccgc cacctgctgt gcacttccta gcgactttgt
 gccattaag taccgcgaat
 25801 gccctccgcc gctttggggc cactgctacc ttctgcagct
 agccaactac cttgcctacc
 25861 actctgacat aatggaagac gtgagcgggtg acggtctact
 ggagtgtcac tgcgctgca
 25921 acctatgcac cccgcaccgc tccctgggtt gcaattcgca
 gctgcttaac gaaagtcaaa
 25981 ttatcggtac ctttgagctg cagggtcctt cgcctgacga
 aaagtccgcg gctccggggg
 26041 tgaaactcac tccgggggctg tggacgtcgg cttaccttcg
 caaatattgta cctgaggact
 26101 accacgcca cgagattagg ttctacgaag accaatcccc
 cccgccaat gcggagctta
 26161 ccgcctgcgt cattaccag ggccacattc ttggccaatt
 gcaagccatc aacaaagccc
 26221 gccaaagatt tctgctacga aaggagcggg gggtttactt
 ggacccccag tccggcgagg
 26281 agctcaacc aatccccccg ccgcccgcag cctatcagca
 gcagccgcgg gcccttgctt
 26341 cccaggatgg cacccaaaaa gaagctgcag ctgccgccgc
 caccacgga cgaggaggaa
 26401 tactgggaca gtcaggcaga ggaggttttg gacgaggagg
 aggaggacat gatggaagac
 26461 tgggagagcc tagacgagga agcttccgag gtcgaagagg
 tgtcagacga aacaccgtca
 26521 ccctcggtcg cattcccctc gccggcgccc cagaaatcgg
 caaccggttc cagcatggct
 26581 acaacctccg ctctcaggc gccgccggca ctgcccgttc
 gccgacccaa cgtagatgg
 26641 gacaccactg gaaccagggc cggtaagtcc aagcagccgc
 cgccgttagc ccaagagcaa
 26701 caacagcgcc aaggctaccg ctcatggcgc gggcacaaga
 acgcatagt tgcttgcttg
 26761 caagactgtg ggggcaacat ctcttccgcc cgccgctttc
 ttctctacca tcacggcgtg
 26821 gccttcccc gtaacatcct gcattactac cgtcatctct
 acagcccata ctgcaccggc
 26881 ggcagcggca gcggcagcaa cagcagcggc cacacagaag
 caaaggcgac cggatagcaa
 26941 gactctgaca aagcccaaga aatccacagc ggcggcagca
 gcaggaggag gagcgtgcg
 27001 tctggcgccc aacgaacccg tatcgacccg cgagcttaga
 aacaggattt ttcccactct
 27061 gtatgctata tttcaacaga gcaggggcca agaacaagag
 ctgaaaataa aaaacaggtc
 27121 tctgcgatcc ctaccggca gctgctgta tcacaaaagc
 gaagatcagc ttcggcgcac

27181 gctggaagac gcgagggtc tcttcagtaa atactgcgcg
 ctgactctta aggactagtt
 27241 tcgcgccctt tctcaaattt aagcgcgaaa actacgtcat
 ctccagcggc cacacccggc
 27301 gccagcacct gtcgtcagcg ccattatgag caaggaaatt
 cccacgccct acatgtggag
 27361 ttaccagcca caaatgggac ttgcggttg agctgcccaa
 gactactcaa cccgaataaa
 27421 ctacatgagc gcgggacccc acatgatatc ccgggtcaac
 ggaatccgcg cccaccgaaa
 27481 ccgaattctc ttggaacagg cggctattac caccacacct
 cgtaataacc ttaatccccg
 27541 tagttggccc gctgccctgg tgtaccagga aagtcccgt
 cccaccactg tgggtacttc
 27601 cagagacgcc caggccgaag ttcagatgac taactcaggg
 gcgcagcttg cgggcggctt
 27661 tcgtcacagg gtgcggtcgc ccgggcaggg tataactcac
 ctgacaatca gagggcgagg
 27721 tattcagctc aacgacgagt cgggtgagctc ctcgcttggt
 ctccgtccgg acgggacatt
 27781 tcagatcggc ggcgccggcc gtccttcatt cacgcctcgt
 caggcaatcc taactctgca
 27841 gacctcgtcc tctgagccgc gctctggagg cattggaact
 ctgcaattta ttgaggagtt
 27901 tgtgccatcg gtctacttta accccttctc gggacctccc
 ggccactatc cggatcaatt
 27961 tattcctaac ttgacgcgg taaaggactc ggcggacggc
 tacgactgaa tggttaagtgg
 28021 agaggcagag caactgcgcc tgaaacacct ggtccactgt
 cgccgccaca agtgctttgc
 28081 ccgcgactcc ggtgagtttt gctactttga attgcccag
 gatcatatcg agggcccggc
 28141 gcacggcgtc cggcttaccg cccagggaga gcttgcccgt
 agcctgattc gggagtttac
 28201 ccagcgcccc ctgctagttg agcgggacag gggaccctgt
 gttctcactg tgatttgcaa
 28261 ctgtcctaac cttggattac atcaagatcc tctagttata
 actagagtac ccggggatct
 28321 tattcccttt aactaataaa aaaaaataat aaagcatcac
 ttacttaaaa tcagtttagca
 28381 aatttctgtc cagtttatcc agcagcacct ccttgcctc
 ctcccagctc tgggtattgca
 28441 gcttcctcct ggctgcaaac tttctccaca atctaaatgg
 aatgtcagtt tcctcctggt
 28501 cctgtccatc cgcacccact atcttcatgt tggtgcagat
 gaagcgcgca agaccgtctg
 28561 aagatacctt caaccccggtg tatccatatg acacggaaac
 cggtcctcca actgtgcctt
 28621 ttcttactcc tccctttgta tcccccaatg gggttcaaga
 gagtccccct ggggtactct

28681 ctttgcgctc atccgaacct ctagttacct ccaatggcat
 gcttgcgctc aaaatgggca
 28741 acggcctctc tctggacgag gccggcaacc ttacctccca
 aaatgtaacc actgtgagcc
 28801 cacctctcaa aaaaaccaag tcaaacataa acctggaaat
 atctgcaccc ctcacagtta
 28861 cctcagaagc cctaactgtg gctgccgccg cacctctaata
 ggtcgcgggc aacacactca
 28921 ccatgcaatc acaggccccg ctaaccgtgc acgactccaa
 acttagcatt gccacccaag
 28981 gaccctctac agtgtcagaa ggaaagctag ccctgcaaac
 atcaggcccc ctcaccacca
 29041 cccatagcag tacccttact atcactgcct caccctctct
 aactactgcc actggtagct
 29101 tgggcattga cttgaaagag cccattttata cacaaaatgg
 aaaactagga ctaaagtacg
 29161 gggctccttt gcatgtaaca gacgacctaa acactttgac
 cgtagcaact ggtccagggtg
 29221 tgactattaa taatacttcc ttgcaaacta aagttactgg
 agccttgggt tttgattcac
 29281 aaggcaatat gcaacttaat gtagcaggag gactaaggat
 tgattctcaa aacagacgcc
 29341 ttatacttga tgtagttat ccgtttgatg ctcaaaacca
 actaaatcta agactaggac
 29401 agggccctct ttttataaac tcagcccaca acttggatat
 taactacaac aaaggccttt
 29461 acttgtttac agcttcaaac aattccaaaa agcttgagggt
 taacctaagc actgccaagg
 29521 gggtgatgtt tgacgctaca gccatagcca ttaatgcagg
 agatgggctt gaatttggtt
 29581 cacctaattgc accaaacaca aatcccctca aaacaaaaat
 tggccatggc ctagaatttg
 29641 attcaaacaa ggctatgggt cctaaactag gaactggcct
 tagttttgac agcacagggtg
 29701 ccattacagt aggaacaaaa aataatgata agctaacttt
 gtggaccaca ccagctccat
 29761 ctctaactg tagactaaat gcagagaaag atgctaaact
 cactttggtc ttaacaaaat
 29821 gtggcagtca aatacttgct acagtttcag ttttggctgt
 taaaggcagt ttggctccaa
 29881 tatctggaac agttcaaagt gctcatctta ttataagatt
 tgacgaaaat ggagtgtac
 29941 taaacaattc ctctctggac ccagaatatt ggaactttag
 aatggagat cttactgaag
 30001 gcacagccta tacaacgct gttggattta tgcctaacct
 atcagcttat ccaaaatctc
 30061 acggtaaaac tgccaaaagt aacattgtca gtcaagttta
 cttaaacgga gacaaaacta
 30121 aacctgtaac actaaccatt acactaaacg gtacacagga
 aacaggagac acaactccaa

30181 gtgcatactc tatgtcattt tcatgggact ggtctggcca
 caactacatt aatgaaatat
 30241 ttgccacatc ctcttacact ttttcataca ttgccaaga
 ataaagaatc gtttgtgtta
 30301 tgtttcaacg tgtttatttt tcaattgcag aaaatttcaa
 gtcatttttc attcagtagt
 30361 atagcccccac caccacatag cttatacaga tcaccgtacc
 ttaatcaaac tcacagaacc
 30421 ctagtattca acctgccacc tccctcccaa cacacagagt
 acacagtcct ttctccccgg
 30481 ctggccttaa aaagcatcat atcatgggta acagacatat
 tcttaggtgt tatattccac
 30541 acggtttcct gtcgagccaa acgctcatca gtgatattaa
 taaactcccc gggcagctca
 30601 cttaagttca tgctgctgtc cagctgctga gccacaggct
 gctgtccaac ttgcgggtgc
 30661 ttaacgggcg gcgaaggaga agtccacgcc tacatggggg
 tagagtcata atcgtgcatc
 30721 aggatagggc ggtgggtgctg cagcagcgcg cgaataaact
 gctgccgccc ccgctccgtc
 30781 ctgcaggaat acaacatggc agtgggtctcc tcagcgatga
 ttgcgaccgc ccgcagcata
 30841 aggcgccttg tcctccgggc acagcagcgcc accctgatct
 cacttaaatac agcacagtaa
 30901 ctgcagcaca gcaccacaat attgttcaaa atcccacagt
 gcaaggcgct gtatccaaag
 30961 ctcatggcgg ggaccacaga acccagctgg ccatcatacc
 acaagcgcag gtagattaag
 31021 tggcgacccc tcataaacac gctggacata aacattacct
 cttttggcat gttgtatattc
 31081 accacctccc ggtaccatat aaacctctga ttaaaccatgg
 cgccatccac caccatccta
 31141 aaccagctgg ccaaaacctg cccgccggct atacactgca
 gggaaccggg actggaacaa
 31201 tgacagtgga gagcccagga ctcgtaacca tggatcatca
 tgctcgatcat gatatcaatg
 31261 ttggcacaac acaggcacac gtgcatacac ttcttcagga
 ttacaagctc ctcccgcggtt
 31321 agaaccatat cccagggaac aacctattcc tgaatcagcg
 taaatcccac actgcaggga
 31381 agacctcgca cgtaactcac gttgtgcatt gtcaaagtgt
 tacattcggg cagcagcgga
 31441 tgatcctcca gtatggtagc gcgggtttct gtctcaaaag
 gaggtagacg atccctactg
 31501 tacggagtgc gccgagacaa ccgagatcgt gttggtcgta
 gtgtcatgcc aaatggaacg
 31561 ccggacgtag tcatatttcc tgaagcaaaa ccagggtgcgg
 gcgtgacaaa cagatctgcg
 31621 tctccgggtct cgccgcttag atcgctctgt gtagtagttg
 tagtatatcc actctctcaa

31681 agcatccagg cgccccctgg cttcggggtc tatgtaaact
 ccttcattgcg ccgctgcccct
 31741 gataacatcc accaccgcag aataagccac acccagccaa
 cctacacatt cgttctgcga
 31801 gtcacacacg ggaggagcgg gaagagctgg aagaaccatg
 tttttttttt tattccaaaa
 31861 gattatccaa aacctcaaaa tgaagatcta ttaagtgaac
 gcgctcccct ccggtggcgt
 31921 ggtcaaactc tacagccaaa gaacagataa tggcatttgt
 aagatgttgc acaatggctt
 31981 ccaaaaggca aacggccctc acgtccaagt ggacgtaaag
 gctaaaccct tcagggtgaa
 32041 tctcctctat aaacattcca gcaccttcaa ccatgcccaa
 ataattctca tctcgccacc
 32101 ttctcaatat atctctaagc aaatcccga tattaagtcc
 ggccattgta aaaatctgct
 32161 ccagagcgcc ctccaccttc agcctcaagc agcgaatcat
 gattgcaaaa attcagggtc
 32221 ctacagacc tgtataagat tcaaaagcgg aacattaaca
 aaaataccgc gatcccgtag
 32281 gtcccttcgc agggccagct gaacataatc gtgcagggtc
 gcacggacca gcgcggccac
 32341 ttccccgcca ggaaccttga caaaagaacc cacactgatt
 atgacacgca tactcggagc
 32401 tatgctaacc agcgtagccc cgatgtaagc tttgttgcat
 gggcgcgcat ataaaatgca
 32461 aggtgctgct caaaaaatca ggcaaagcct cgcgcaaaaa
 agaaagcaca tcgtagtcat
 32521 gctcatgcag ataaaggcag gtaagctccg gaaccaccac
 agaaaaagac accatthttc
 32581 tctcaaacat gtctgcgggt ttctgcataa acacaaaata
 aaataacaaa aaaacattta
 32641 aacattagaa gcctgtctta caacaggaaa aacaaccctt
 ataagcataa gacggactac
 32701 ggccatgccg gcgtgaccgt aaaaaaactg gtcaccgtga
 ttaaaaagca ccaccgacag
 32761 ctctcgggtc atgtccggag tcataatgta agactcggta
 aacacatcag gttgattcat
 32821 cggtcagtgc taaaaagcga ccgaaatagc ccgggggaat
 acatacccgcc aggcgtagag
 32881 acaacattac agccccata ggaggtataa caaaattaat
 aggagagaaa aacacataaa
 32941 cacctgaaaa accctcctgc ctaggcaaaa tagcaccctc
 ccgctccaga acaacataca
 33001 gcgcttcaca gcggcagcct aacagtcagc cttaccagta
 aaaaagaaaa cctattaaaa
 33061 aaacaccact cgacacggca ccagctcaat cagtcacagt
 gtaaaaaagg gccaaagtga
 33121 gagcgagtat atataggact aaaaaatgac gtaacggtta
 aagtccacaa aaaacaccca

33181 gaaaaccgca cgcgaaaccta cgcccagaaa cgaaagccaa
 aaaaccacaca acttcctcaa
 33241 atcgctcactt ccgtttttccc acgttacgta acttcccatt
 ttaagaaaac tacaattccc
 33301 aacacataca agttactccg ccctaaaacc tacgtcaccc
 gcccgcgttcc cagcggccgc
 33361 gccacgtcac aaactccacc ccctcattat catattgggt
 tcaatccaaa ataaggtata
 33421 ttattgatga tnnnnnttaa t

SEQ ID NO:17 (ND1.1 214)

1 taaggatccn nncctgtcct cgaccgatgc ccttgagagc cttcaaccca
 gtcagctcct
 61 tccgggtgggc gcggggcatg actatcgctg ccgcacttat
 gactgtcttc tttatcatgc
 121 aactcgtagg acaggtgccg gcagcgctct gggtcatttt
 cggcgaggac cgctttcgct
 181 ggagcgcgac gatgatcggc ctgtcgcttg cggtatccgg
 aatcttgcac gccctcgctc
 241 aagccttcgt cactgggtccc gccaccaaac gtttcggcga
 gaagcaggcc attatcgccc
 301 gcatggcggc cgacgcgctg ggctacgtct tgctggcggt
 cgcgacgcga ggctggatgg
 361 ccttccccat tatgattctt ctgccttccg gcggcatcgg
 gatgccgcg ttgcaggcca
 421 tgctgtccag gcaggtagat gaagaccatc agggacagct
 tcaaggatcg ctgcggctc
 481 ttaccagcct aacttcgata actggaccgc tgatcgtcac
 ggcgatttat gccgcctcgg
 541 cgagcacatg gaacgggttg gcatggattg taggcgcgcg
 cctatacctt gtctgcctcc
 601 ccgcgttgcg tcgcggtgca tggagccggg ccacctcgac
 ctgaatggaa gccggcggca
 661 cctcgctaac ggattcacca ctccaagaat tggagccaat
 caattcttgc ggagaactgt
 721 gaatgcgcaa accaaccctt ggcagaacat atccatcgcg
 tccgccatct ccagcagccc
 781 cacgcggcgc atctcgggca gcgttgggtc ctggccacgg
 gtgcgcatga tcgtgctcct
 841 gtcgttgagg acccggtag gctggcgggg ttgccttact
 ggtagcaga atgaatcacc
 901 gatacgcgag cgaacgtgaa gcgactgctg ctgcaaaacg
 tctgcgacct gagcaacaac
 961 atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa
 acgcggaagt cagcgcctg
 1021 caccattatg ttccggatct gcatcgagc atgctgctgg
 ctaccctgtg gaacacctac
 1081 atctgtatta acgaagcgct ggcattgacc ctgagtgatt
 tttctctggt cccgccgcac

1141 ccataccgcc agttgtttac cctcacaacg ttccagtaac
 cgggcatggt catcatcagt
 1201 aaccggtatc gtgagcatcc tctctcgttt catcggtatc
 attaccccca tgaacagaaa
 1261 ttccccctta cacggaggca tcaagtgacc aaacaggaaa
 aaaccgccct taacatggcc
 1321 cgctttatca gaagccagac attaacgctt ctggagaaac
 tcaacgagct ggacgcggat
 1381 gaacaggcag acatctgtga atcgcttcac gaccacgctg
 atgagcttta ccgagctgc
 1441 ctgcgcgctt tcggtgatga cggtgaaaac ctctgacaca
 tgcagctccc ggagacggtc
 1501 acagcttgct tgtaagcggg tgccgggagc agacaagccc
 gtcaggggcg gtcagcgggt
 1561 gttggcgggt gtcggggcg cagcatgacc cagtcacgta
 gcgatagcgg agtgataact
 1621 ggcttaacta tgcggcatca gagcagattg tactgagagt
 gcaccatatg cgggtgtgaaa
 1681 taccgcacag atgcgtaagg agaaaatacc gcatcaggcg
 ctcttcgct tctcgcctca
 1741 ctgactcgtt ggcctcggtc gttcggctgc ggcgagcggg
 atcagctcac tcaaaggcgg
 1801 taatacgggt atccacagaa tcaggggata acgcaggaaa
 gaacatgtga gcaaaaggcc
 1861 agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc
 gtttttccat aggtccgccc
 1921 cccctgacga gcatcacaaa aatcgacgct caagtcagag
 gtggcgaaac ccgacaggac
 1981 tataaagata ccaggcggtt cccctggaa gctccctcgt
 gcgctctcct gttccgaccc
 2041 tgccgcttac cggatacctg tccgccttcc tcccttcggg
 aagcgtggcg ctttctcaat
 2101 gtcacgctg taggtatctc agttcgggtg aggtcgttcg
 ctccaagctg ggctgtgtgc
 2161 acgaaccccc cgttcagccc gaccgctgcg ccttatccgg
 taactatcgt cttgagtcca
 2221 acccggtgag acacgactta tcgccactgg cagcagccac
 tggtaacagg attagcagag
 2281 cgaggatgt aggcgggtgct acagagttct tgaagtgggtg
 gcctaactac ggctacacta
 2341 gaaggacagt atttgggtatc tgcgctctgc tgaagccagt
 taccttcgga aaaagagttg
 2401 gtagctcttg atccggcaaa caaaccaccg ctggtagcgg
 tgggtttttt gtttgcaagc
 2461 agcagattac gcgcagaaaa aaaggatctc aagaagatcc
 tttgatcttt tctacggggt
 2521 ctgacgctca gtggaacgaa aactcacggt aagggatttt
 ggtcatgaga ttatcaaaaa
 2581 ggatcttcac ctagatcctt ttaaattaaa aatgaagttt
 taaatcaatc taaagtatat

2641 atgagtaaac ttgggtctgac agttaccaat gcttaatcag
tgaggcacct atctcagcga
2701 tctgtctatt tcgttcatcc atagttgcct gactccccgt
cgtgtagata actacgatac
2761 gggagggctt accatctggc cccagtgtg caatgatacc
gcgagaccca cgctcaccgg
2821 ctccagatct atcagcaata aaccagccag ccggaagggc
cgagcgcaga agtggctctg
2881 caactttatc cgcctccatc cagtctatta attgttgccg
ggaagctaga gtaagtagtt
2941 cgccagttaa tagtttgccg aacgttggtg ccattgctgc agccatga
gattatcaaa
aaggatcttc acctagatcc ttttcacgta gaaagccagt ccgcagaaac
ggtgctgacc
ccggatgaat gtcagctact gggctatctg gacaagggaa aacgcaagcg
caaagagaaa
gcaggtagct tgcagtgggc ttacatggcg atagctagac tgggcggttt
tatggacagc
aagcgaaccg gaattgccag ctggggcgcc ctctggtaag gttgggaagc
cctgcaaagt
aaactggatg gctttctcgc cgccaaggat ctgatggcgc aggggatcaa
gctctgatca
agagacagga tgaggatcgt ttgcgatgat tgaacaagat ggattgcacg
caggttctcc
ggcgcgttgg gtggagaggc tattcggcta tgactgggca caacagacaa
tcggctgctc
tgatgcgcgc gtgttcggc tgtcagcgca ggggcgccc gttctttttg
tcaagaccga
cctgtccggt gccctgaatg aactgcaaga cgaggcagcg cggctatcgt
ggctggccac
gacgggcgtt ccttgccgag ctgtgctcga cgttgctact gaagcgggaa
gggactggct
gctattgggc gaagtgccgg ggcaggatct cctgtcatct caccttgctc
ctgccgagaa
agtatccatc atggctgatg caatgcggcg gctgcatacg cttgatccgg
ctacctgcc
attcgaccac caagcgaaac atcgcatcga gcgagcacgt actcggatgg
aagccggtct
tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg
aactgttcgc
caggetcaag gcgagcatgc ccgacggcga ggatctcgtc gtgacccatg
gcgatgcctg
cttgccgaat atcatggtgg aaaatggccg cttttctgga ttcatcgact
gtggccggct
gggtgtggcg gaccgctatc aggacatagc gttggctacc cgtgatattg
ctgaagagct
tggcggcgaa tgggctgacc gcttctcgt gctttacggg atcgccgctc
ccgattcgca
gcgcategcc ttctatcgcc ttcttgacga gttcttctga attttggttaa
aatttttgtt

aaatcagctc attttttaac caataggccg aaatcggcaa catcccttat
 aaatcaaaag
 aatagaccgc gatagggttg agtggtgttc cagtttggaa caagagtcca
 ctattaaaga
 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc
 ccactacgtg
 aaccatcacc caaatcaagt tttttgcggt cgagggtgccg taaagctcta
 aatcgggaacc
 ctaaaggagag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg
 gcgagaaagg
 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg
 gtcacgctgc
 gcgtaaccac cacaccgcg cgcttaatgc gccgctacag ggcgcgtcca
 ttcgccattc
 aggatcgaat taattcttaa ttaacatcatc aataatatac cttattttgg
 attgaagcca atatgataat gagggggttg agtttgtgac gtggcgcggg
 gcgtgggaac ggggcgggtg acgtagtagt gtggcggaag tgtgatgttg
 caagtgtggc ggaacacatg taagcgacgg atgtggcaaa agtgacgttt
 ttggtgtgcg ccggtgtaca caggaagtga caattttcgc gcggttttag
 gcggtatgttg tagtaaattt gggcgtaacc gagtaagatt tggccatttt
 cgcgggaaaa ctgaataaga ggaagtgaat tctgaataat tttgtgttac
 tcatagcgcg
 taatactGCTAGAgatCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGC
 CAGGGTTTTCCCAGTCACGACGTTGTAAACGACGGCCAGTGAATTGTAATACGACTCACTA
 TAGGGCGAATTGGGTACTGGCCACagagccttgcccattgcatacgttgtatccatatcata
 atatgtacatttatattggctcatgtccaacattaccgccatggttgacattgattattgact
 agttattaatagtaatacaattacggggtcattagttcatagcccatatatggagttccgcgt
 tacataacttacggtaaatggcccgcttggtgaccgcccacgacccccgccattgacgt
 caataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtg
 gagtatttacggtaaactgccacttggcagtacatcaagtgtatcatatgccaaagtacgcc
 ccctattgacgtcaatgacggtaaatggcccgcttggtcattatgccagtacatgacctat
 gggactttcctacttggcagtacatctacgtattagtcacgtattaccatgggtgatgcgg
 ttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctcca
 cccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtc
 gtaacaactccgccccattgacgcaaatgggcggttaggcgtgtacgggtgggaggtctatata
 agcagagctcgttttagtgaaccgtcagatcgccctggagacgccatccacgctgttttgacct
 ccatagaagacaccgggaccgatccagcctgactctagCctAGCTCtgaagttgggtggtgag
 gccctgggcaggttggtatcaaggttacaagacaggtttaaggagaccaatagaaactgggc
 atgtggagacagagaagactccttgggtttctgataggcactgactctctctgcctattgggc
 tattttcccacccttaggctgctggtctgagcctagGAGATCTCTCGAGGTCGACGGTATcg
 atgccca cc**ATGGAGAA AATCGTCCTG TTGCTCGCTA TTGTGTCTCTAGTGAAGAGC**
GATCAAATTT GTATCGGCTA CCATGCCAAT AACTCAACAGAGCAGGTCTGA
TACTATCATG GAGAAAAACG TAACAGTTAC TCATGCCCAA GACATCTTGG
AAAAGACCCA CAACGGCAAA CTTTGCGACC TGGATGGAGT GAAGCCCCTG
ATCCTCCGGG ACTGTTCACT CGCTGGTTGG CTGCTCGGGA ACCCTATGTG
TGATGAGTTT ATCAACGTGC CTGAATGGTC TTACATTGTG GAGAAGGCTA
ACCCTACCAA TGACCTCTGC TATCCTGGGT CATTTAACGA TTACGAGGAA
CTGAAACACC TGTGTCTAG AATTAACCAC TTTGAAAAGA TACAGATTAT
ACCAAGTCT AGTTGGAGTG ATCACGAAGC CTCCTCAGGC GTTAGCTCAG
CGTGTCCCTA TCTGGGCTCT CCATCCTTCT TTAGAAATGT GGTCTGGTTA
ATCAAAAAGA ACAGTACCTA CCCAACCATC AAAAAGTCTT ATAACAATAC

CAATCAGGAG GACCTGCTCG TGTGTGTTGGG TATCCATCAC CCGAACGACG
 CCGCTGAACA GACTAGGCTG TATCAGAACC CCACTACATA CATCAGTATT
 GGCACGAGTA CTCTGAACCA GCGATTAGTG CCAAAGATTG CAACACGGAG
 CAAAGTAAAT GGGCAATCTG GCAGGATGGA GTTTTTCTGG ACAATCTTAA
 AACCCAACGA TGCATAAAT TTCGAGTCCA ATGGCAATTT CATCGCCCCT
 GAATACGCCT ATAAGATCGT GAAAAAGGGG GACTCTGCAA TTATGAAGTC
 CGAATTAGAG TATGGCAATT GCAACACGAA GTGCCAGACA CCAATGGGAG
 CCATTAATAG CTCAATGCCC TTCCATAATA TTCATCCATT GACCATTGGG
 GAGTGCCCAA AGTACGTGAA GTCCAACCGC CTGGTCCTCG CAACCGGTCT
 AAGAAATAGC CCGCAGAGAG AATCGCGGAG GAAGAAACGT GGCCTGTTTG
 GCGCGATTGC CGGATTCATC GAGGGAGGCT GGCAGGGTAT GGTGATGGT
 TGGTACGGAT ACCACCATAG CAACGAACAG GGGTCCGGCT ATGCAGCAGA
 TAAGGAGAGC ACTCAGAAAG CTATTGACGG AGTTACAAAC AAGGTTAATA
 GTATTATAGA TAAATGAAC ACGCAATTCTG AGGCCGTTGG GAGGGAGTTT
 AACAACTCTG AACGCCGGAT CGAAAATCTG AATAAGAAAA TGAAGACGG
 CTTCTTTGAC GTGTGGACTT ATAATGCAGA GCTGCTTGTA CTCATGGAGA
 ACGAGAGGAC CCTGGATTTT CACGATAGCA ACGTGAAGAA CCTTTACGAC
 AAGGTGAGAC TTCAGCTCCG AGACAACGCC AAGGAGCTGGGGAATGGATG
 CTTGAGTGTG TACCACAAAT GTGACAATGA GTGCATGGAAAGTATACGCA
 ACGGGACCTA CAATTACCTT CAGTATAGCG AAGAGGCTCGGCTCAAACGC
 GAAGAGATAA GCGGGGTGAA ATTGGAATCA ATCGGAACATATCAAATCCT
 GTCCATCTAT TCCACCGTCG CCTCTTCGCT GGCCCTCGCT
 ATCATGATGG CTGGTCTGTC CCTATGGATG TGTTCCAATG
 GAAGCCTTCAGTGCCGTATT TGTATATGA_{g c}
 GGCCGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCC
 TT
 CTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
 ACTCCCAC
 TGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGT
 GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
 GGTGGGCT
 CTATGGCTTCTGAGGCGGAAAGAACCAAAGCTtaacatcatc aataatatac
 cttattttgg attgaagcca atatgataat gagggggtgg.
 agtttgtgac gtggcgcggg gcgtgggaac ggggcggggtg acgtagtagt
 gtggcggaag
 tgtgatgttg caagtgtggc ggaacacatg taagcgacgg atgtggcaaa
 agtgacgttt
 ttggtgtgcg ccggtgtaca caggaagtga caattttcgc gcggttttag
 gcggatgttg
 tagtaaattt gggcgtaacc gagtaagatt tggccatttt cgcgggaaaa
 ctgaataaga
 ggaagtgaaa tctgaataat tttgtgttac tcatagcgcg taatactgta
 atagtaatca
 attacgggggt cattagttca tagcccatat atggagttcc gcgttacata
 acttacggta
 aatggcccg cctggctgacc gcccaacgac ccccgcccat tgacgtcaat
 aatgacgtat
 gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga
 gtattttacgg

taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc
 ccctattgac
 gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt
 atgggacttt
 cctacttggc agtacatcta cgtattagtc atcgctatta ccatgggtgat
 gcggttttgg
 cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag
 tctccacccc
 attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc
 aaaatgtcgt
 aacaactccg cccattgac gcaaattggc ggtaggcgtg tacgggtggga
 ggtctatata
 agcagagctg gtttagtgaa ccgtcagatc
 cgctagAGATCTGGgaaacgatatgggctgaatac**cgatccgtattcagcccatatcgtttc**
 TCT AGA AAT AAA ATA TCT TTA TTT TCA TTA CAT CTG TGT GTT GGT
 TTT TTG TGT GGC
 GG CCGCtcgagcct AAGCTTctag ataagatatc cgatccaccg gatctagata
 actgatcata
 atcagccata ccacatttgt agaggtttta cttgctttaa aaaacctccc
 acacctccc
 ctgaacctga aacataaaat gaatgcaatt gttgttgta acttgtttat
 tgcagcttat
 aatgggtaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt
 tttttcactg
 cattctagtt gtggtttgtc caaactcacc aatgtatctt aacgcggatc
 tgggcgtggt
 taagggtggg aaagaatata taagggtggg gtcttatgta gttttgtatc
 tgttttgcag
 cagccgcgcg cgccatgagc accaactcgt ttgatggaag cattgtg
 AGCTTgtcgactcg aagatctggg cgtggttaag
 ggtgggaaag aatatataag gtgggggtct tatgtagttt tgtatctggt
 ttgcagcagc
 cgccgcccgc atgagcacca actcgtttga tggaagcatt gtgagctcat
 atttgacaac
 gcgcatgccc ccatgggccg ggggtgcgtca gaatgtgatg ggctccagca
 ttgatggtcg
 cccgctcctg cccgcaaact ctactacctt gacctacgag accgtgtctg
 gaacgccgtt
 ggagactgca gcctccgccg ccgcttcagc cgctgcagcc accgcccgcg
 ggattgtgac
 tgactttgct ttctgagcc cgcttgcaag cagtgcagct tcccgttcat
 ccgcccgcga
 tgacaagttg acggctcttt tggcacaatt ggattctttg acccggaac
 ttaatgtcgt
 ttctcagcag ctggttgatc tgcgccagca ggtttctgcc ctgaaggctt
 cctccccctc
 caatgcgggt taaaacataa ataaaaaacc agactctggt tggatttgga
 tcaagcaagt
 gtcttgctgt ctttatcttag gggttttgcg cgccggttag gcccgggacc
 agcggctcgt

gtcgttgagg gtcctgtgta ttttttccag gacgtggtaa aggtgactct
ggatgttcag
atacatgggc ataagcccggt ctctgggggtg gaggtagcac cactgcagag
cttcatgctg
cgggggtgggtg ttgtagatga tccagtcgta gcaggagcgc tgggcgtgggt
gcctaataaat
gtctttcagt agcaagctga ttgccagggg caggcccttg gtgtaagtgt
ttacaaagcg
gttaagctgg gatgggtgca tacgtgggga tatgagatgc atcttggact
gtatttttag
gttggctatg ttcccagcca tatccctccg gggattcatg ttgtgcagaa
ccaccagcac
agtgtatccg gtgcacttgg gaaatttgtc atgtagctta gaaggaaatg
cgtggaagaa
cttgagagcg ccttgtgac ctccaagatt ttccatgcat tcgtccataa
tgatggcaat
gggcccacgg gcggcggcct gggcgaagat atttctggga tcactaacgt
catagtgtg
ttccaggatg agatcgatcat aggccatttt taaaaagcgc gggcggaggg
tgccagactg
cggataaatg gttccatccg gccaggggc gtagttaccc tcacagattt
gcatttccca
cgctttgagt tcagatgggg ggatcatgtc tacctgcggg gcgatgaaga
aaacggtttc
cggggtaggg gagatcagct ggaagaaaag caggttcctg agcagctgcg
acttacgca
gccggtgggc ccgtaaatca cacctattac cgggtgcaac tggtagttaa
gagagct
g cagctgccgt catccctgag cagggggggc
acttcgttaa gcatgtccct gactcgcagt ttttccctga ccaaataccg
cagaaggcgc
tcgcgcacca gcgatagcag ttcttgcaag gaagcaaagt ttttcaacgg
tttgagaccg
tccgccgtag gcatgctttt gagcgtttga ccaagcagtt ccaggcggtc
ccacagctcg
gtcacctgct ctacggcatc tcgatccagc atatctctc gtttcgcggg
ttggggcggc
tttcgctgta cggcagtagt cgggtgctcg ccagacgggc cagggtcagt
tctttccacg
ggcgcagggg cctcgtcagc gtagtctggg tcacggtgaa ggggtgcgct
ccgggctgcg
cgctggccag ggtgcgcttg aggctggccc tgctgggtgct gaagcgctgc
cggctcttcg
cctgcgcgtc ggccaggtag catttgacca tggtgtcata gtccagcccc
tccgcggcgt
ggcccttggc gcgcagcttg cccttggagg aggcgccgca cgaggggag
tgcagacttt
tgagggcgta gagcttgggc gcgagaaata ccgattccgg ggagtaggca
tccgcgcgc
aggccccgca gacgggtctcg cattccacga gccaggtgag ctctggccgt
tcgggggtcaa

aaaccagggtt tcccccatgc tttttgatgc gtttcttacc tctggtttcc
 atgagccggt
 gtccacgctc ggtgacgaaa aggctgtccg tgtccccgta tacagacttg
 agaggcctgt
 cctcgagcgg tggtccgagg tctcctcgt atagaaactc ggaccactct
 gagacaaagg
 ctgcggtcca ggccagcacg aaggaggcta agtgggaggg gtagcggtcg
 ttgtccacta
 ggggggtccac tcgctccagg gtgtgaagac acatgtcgcc ctcttcggca
 tcaaggaagg
 tgattggttt gtaggtgtag gccacgtgac cgggtgttcc tgaagggggg
 ctataaaagg
 ggggtgggggc gcgttcgtcc tcaactctctt ccgcacgtct gtctgcgagg
 gccagctgtt
 ggggtgagta ctccctctga aaagcgggca tgacttctgc gctaagattg
 tcagtttcca
 6301 aaaacgagga ggatttgata ttcacctggc ccgcgggtgat
 gcctttgagg gtggccgcat
 6361 ccatctggtc agaaaagaca atctttttgt tgtcaagctt
 ggtggcaaac gaccgtaga
 6421 gggcggttga cagcaacttg gcgatggagc gcagggtttg
 gtttttgtcg cgatcggcgc
 6481 gtccttggc cgcatgttt agctgcacgt attcgcgcgc
 aacgcaccgc cattcgggaa
 6541 agacgggtgt gcgctcgctc ggcaccagggt gcacgcgcca
 accgcggttg tgcagggtga
 6601 caaggtcaac gctggtggct acctctccgc gtaggcgctc
 gttggtccag cagaggcggc
 6661 cgcccttgcg cgagcagaat ggcggtaggg ggtctagctg
 cgtctcgctc ggggggtctg
 6721 cgtccacggt aaagaccccg ggcagcaggc gcgcgtcgaa
 gtagtctatc ttgcatcctt
 6781 gcaagtctag cgctgctgc catgcgcggg cggcaagcgc
 gcgctcgat gggttgagt
 6841 ggggaccca tggcatgggg tgggtgagcg cggaggcgta
 catgccgcaa atgtcgtaaa
 6901 cgtagagggg ctctctgagt attccaagat atgtagggta
 gcatcttcca ccgcggatgc
 6961 tggcgcgcac gtaatcgat agttcgtgcg agggagcgag
 gaggtcggga ccgaggttgc
 7021 tacgggcggg ctgctctgct cggaagacta tctgcctgaa
 gatggcatgt gagttggatg
 7081 atatggttg acgctggaag acgttgaagc tggcgtctgt
 gagacctacc gcgtcacgca
 7141 cgaaggaggc gtaggagtcg cgcagcttgt tgaccagctc
 ggcggtgacc tgcacgtcta
 7201 gggcgcagta gtccagggtt tcttgatga tgtcatactt
 atcctgtccc ttttttttcc
 7261 acagctcgcg gttgaggaca aactcttcgc ggtctttcca
 gtactcttgg atcggaacc

7321 cgtcggcctc cgaacggtaa gagcctagca tgtagaactg
 gttgacggcc tggtaggcgc
 7381 agcatccctt ttctacgggt agcgcgtatg cctgcgcggc
 ctcccgagc gaggtgtggg
 7441 tgagcgcaaa ggtgtccctg accatgactt tgaggtactg
 gtatttgaag tcagtgtcgt
 7501 cgcacccgcc ctgctcccag agcaaaaagt ccgtgcgctt
 tttggaacgc ggatttggca
 7561 gggcgaagggt gacatcggtg aagagtatct ttcccgcgcg
 aggcataaag ttgcgtgtga
 7621 tgcggaagggt tcccggcacc tcggaacgggt tgttaattac
 ctgggcggcg agcacgatct
 7681 cgtcaaagcc gttgatggtg tggcccacaa tgtaaagtcc
 caagaagcgc gggatgccct
 7741 tgatggaagg caatttttta agttcctcgt aggtgagctc
 ttcaggggag ctgagcccgt
 7801 gctctgaaag ggcccagtct gcaagatgag ggttggaagc
 gacgaatgag ctccacaggt
 7861 cacggggccat tagcatttgc aggtggtcgc gaaaggctct
 aaactggcga cctatggcca
 7921 ttttttctgg ggtgatgcag tagaaggtaa gcgggtcttg
 ttcccgaggc tcccatccaa
 7981 ggctcgcgcc taggtctcgc gcggcagtca ctagaggctc
 atctccgccc aacttcatga
 8041 ccagcatgaa gggcacgagc tgcttcccaa aggcccccat
 ccaagtatag gtctctacat
 8101 cgtagggtgac aaagagacgc tcggtgcgag gatgcgagcc
 gatcggaag aactggatct
 8161 cccgccacca attggaggag tggctattga tgtggtgaaa
 gtagaagtcc ctgcgacggg
 8221 ccgaacactc gtgctgggctt ttgtaaaaac gtgcgcagta
 ctggcagcgg tgcacgggct
 8281 gtacatcctg cacgagggtg acctgacgac cgcgcacaa
 gaagcagagt gggaatttga
 8341 gccctcgcgc tggcggggtt ggctgggtgt cttctacttc
 ggctgcttgt ccttgaccgt
 8401 ctggctgctc gaggggaggt acggtggatc ggaccaccac
 gccgcgcgag ccaaagtcc
 8461 agatgtccgc gcgcggcgggt cggagcttga tgacaacatc
 gcgcagatgg gagctgtcca
 8521 tggctctggag ctcccgcggc gtcaggctcag gcgggagctc
 ctgcagggtt acctcgcata
 8581 gacgggtcag ggcgcgggct agatccaggt gataccta
 ttccaggggc tggttggtg
 8641 cggcgtcgat ggcttgcaag aggcgcgcat cccgcggcgc
 gactacggta ccgcgcggcg
 8701 ggcggtgggc cgcgggggtg tccttggtat atgcatctaa
 aagcggtgac gcgggcgagc
 8761 ccccgagggt agggggggct ccggaccgcg cgggagaggg
 ggcaggggca cgtcggcgcc

8821 gcgcgcgggc aggagctggt gctgcgcgcg taggttgctg
 gcgaacgcga cgacgcggcg
 8881 gttgatctcc tgaatctggc gcctctgcgt gaagacgacg
 ggcccgggtga gcttgagcct
 8941 gaaagagagt tcgacagaat caatttcggt gtcgttgacg
 gcggcctggc gcaaaatctc
 9001 ctgcacgtct cctgagttgt cttgataggc gatctcggcc
 atgaactgct cgatctcttc
 9061 ctccctggaga tctccgcgtc cggctcgcctc cacgggtggcg
 gcgaggtcgt tggaaatgcg
 9121 ggccatgagc tgcgagaagg cgttgaggcc tccctcgttc
 cagacgcggc tgtagaccac
 9181 gccccttcg gcacgcggg gcgcacatgac cacctgcgcg
 agattgagct ccacgtgccg
 9241 ggccaagacg gcgtagtctt gcaggcgctg aaagaggtag
 ttgaggggtg tggcgggtgtg
 9301 ttctgccacg aagaagtaca taaccacgctc tcgcaacgtg
 gattcgttga tatcccccaa
 9361 ggctcaagg cgctccatgg cctcgtagaa gtccacggcg
 aagttgaaaa actgggagtt
 9421 gcgcgcgcgac acgggttaact cctcctccag aagacggatg
 agctcggcga cagtgtcgcg
 9481 cacctcgcgc tcaaaggcta caggggcctc ttcttcttct
 tcaatctcct cttccataag
 9541 ggctcccct tcttcttctt ctggcggcg tgggggaggg
 gggacacggc ggcgacgacg
 9601 gcgcaccggg aggcggctga caaagcgctc gatcatctcc
 ccgcggcgac ggcgcatggt
 9661 ctcggtgacg gcgcggccgt tctcgcgggg gcgcagttgg
 aagacgccgc ccgtcatgtc
 9721 ccggttatgg gttggcgggg ggctgccatg cggcagggat
 acggcgctaa cgatgcacat
 9781 caacaattgt tgtgtaggta ctccgccgcc gagggacctg
 agcgagtccg catcgaccgg
 9841 atcggaaaac ctctcgagaa aggcgtctaa ccagtcacag
 tcgcaaggta ggctgagcac
 9901 cgtggcgggc ggcagcgggc ggcggtcggg gttgtttctg
 gcggaggtgc tgctgatgat
 9961 gtaattaaag taggcggtct tgagacggcg gatggtcgac
 agaagcacca tgtccttggg
 10021 tccggcctgc tgaatgcgca ggcggtcggc catgccccag
 gcttcgtttt gacatcggcg
 10081 caggtctttg tagtagtctt gcatgagcct ttctaccggc
 acttcttctt ctcttcctc
 10141 ttgtcctgca tctcttgcac ctatcgctgc ggcgggcgcg
 gagtttggcc gtaggtggcg
 10201 ccctcttcct cccatgcgtg tgaccccgaa gccctcatc
 ggctgaagca gggctaggtc
 10261 ggcgacaacg cgctcggcta atatggcctg ctgcacctgc
 gtgagggtag actggaagtc

10321 atccatgtcc acaaagcggg ggtatgcgcc cgtgttgatg
 gtgtaagtgc agttggccat
 10381 aacggaccag ttaacgggtct ggtgacccgg ctgcgagagc
 tcggtgtacc tgagacgcga
 10441 gtaagccctc gagtcaaata cgtagtcgtt gcaagtcgcg
 accaggtact ggtatccac
 10501 caaaaagtgc ggcggcggct ggcggtagag gggccagcgt
 aggggtggccg gggctccggg
 10561 ggcgagatct tccaacataa ggcgatgata tccgtagatg
 tacctggaca tccaggtgat
 10621 gccggcggcg gtggtggagg cgcgcggaaa gtcgcggacg
 cggttccaga tgttgccgag
 10681 cggcaaaaag tgctccatgg tcgggacgct ctggccggtc
 aggcgcgcgc aatcgttgac
 10741 gctctaccgt gcaaaaggag agcctgtaag cgggcactct
 tccgtggtct ggtggataaa
 10801 ttcgcaaggg tatcatggcg gacgaccggg gttcgagccc
 cgtatccggc cgtccgccgt
 10861 gatccatgcg gttaccgccc gcgtgtcgaa cccaggtgtg
 cgacgtcaga caacggggga
 10921 gtgctccttt tggcttcctt ccaggcgcgg cggctgctgc
 gctagctttt ttggccactg
 10981 gccgcgcgca gcgtaagcgg ttaggctgga aagcgaaagc
 attaagtggc tcgctccctg
 11041 tagccggagg gttattttcc aagggttgag tcgcgggacc
 cccggttcga gtctcggacc
 11101 ggccggactg cggcgaacgg gggtttgct ccccgatcag
 caagaccccg cttgcaaatt
 11161 cctccgaaa cagggacgag cccctttttt gcttttccca
 gatgcatccg gtgctgcggc
 11221 agatgcgccc ccctcctcag cagcggcaag agcaagagca
 gcggcagaca tgcagggcac
 11281 cctcccctcc tcctaccgcg tcaggagggg cgacatccgc
 ggttgacgcg gcagcagatg
 11341 gtgattacga accccgcgg cgcggggccc ggcactacct
 ggacttggag gagggcgagg
 11401 gcctggcgcg gctaggagcg ccctctcctg agcggtaccc
 aagggtgcag ctgaagcgtg
 11461 atacgcgtga ggcgtacgtg ccgcggcaga acctgtttcg
 cgaccgcgag ggagaggagc
 11521 ccgaggagat gcgggatcga aagttccacg cagggcgcga
 gctgcggcat ggcctgaatc
 11581 gcgagcgggt gctgcgcgag gaggactttg agcccgacgc
 gcgaaccggg attagtcccc
 11641 cgcgcgcaca cgtggcggcc gccgacctgg taaccgcata
 cgagcagacg gtgaaccagg
 11701 agattaactt tcaaaaaagc tttaacaacc acgtgcgtac
 gcttggtggcg cgcgaggagg
 11761 tggctatagg actgatgcat ctgtgggact ttgtaagcgc
 gctggagcaa aacccaaata

11821 gcaagccgct catggcgag ctgttcctta tagtgcagca
 cagcagggac aacgaggcat
 11881 tcagggatgc gctgctaaac atagtagagc ccgagggccg
 ctggctgctc gatttgataa
 11941 acatcctgca gagcatagtgt gtgcaggagc gcagcttgag
 cctggctgac aaggtggccg
 12001 ccatcaacta ttccatgctt agcctgggca agttttacgc
 ccgcaagata taccataccc
 12061 cttacgttcc catagacaag gaggtaaaga tcgagggggt
 ctacatgcgc atggcgctga
 12121 aggtgcttac cttgagcgac gacctgggcg tttatcgcaa
 cgagcgcatc cacaaggccg
 12181 tgagcgtgag ccggcgggcg gagctcagcg accgcgagct
 gatgcacagc ctgcaaaggg
 12241 ccctggctgg cacgggcagc ggcgatagag aggccgagtc
 ctactttgac gcgggcgctg
 12301 acctgcgctg ggccccaagc cgacgcgccc tggaggcagc
 tggggccgga cctgggctgg
 12361 cggtggcacc cgcgcgcgct ggcaacgctg gcggcggtga
 ggaatatgac gaggacgatg
 12421 agtacgagcc agaggacggc gagtactaag cggtgatggt
 tctgatcaga tgatgcaaga
 12481 cgcaacggac ccggcggtgc gggcgggcgt gcagagccag
 ccgtccggcc ttaactccac
 12541 ggacgactgg cgccaggtea tggaccgcat catgtcgctg
 actgcgcgca atcctgacgc
 12601 gttccggcag cagccgcagg ccaaccggct ctccgcaatt
 ctggaagcgg tgggtccggc
 12661 gcgcgcaaac ccacgcacg agaaggtgct ggcgatcgta
 aacgcgctgg ccgaaaacag
 12721 ggccatccgg cccgacgagg ccggcctggt ctacgacgcg
 ctgcttcagc gcgtggctcg
 12781 ttacaacagc ggcaacgtgc agaccaacct ggaccggctg
 gtgggggatg tgcgcgaggc
 12841 cgtggcgag cgtgagcgcg cgcagcagca gggcaacctg
 ggctccatgg ttgcactaaa
 12901 cgcttcctg agtacacagc ccgccaacgt gccgcgggga
 caggaggact acaccaactt
 12961 tgtgagcgca ctgcggctaa tggtgactga gacaccgcaa
 agtgaggtgt accagtctgg
 13021 gccagactat tttttccaga ccagtagaca aggctgcag
 accgtaaacc tgagccaggc
 13081 tttcaaaaac ttgcaggggc tgtgggggggt gcgggctccc
 acaggcgacc gcgcgaccgt
 13141 gtctagcttg ctgacgcca actcgcgctt gttgctgctg
 ctaatagcgc ccttcacgga
 13201 cagtggcagc gtgtcccggg acacatacct aggtcacttg
 ctgacactgt accgcgaggc
 13261 cataggtcag gcgcagtgtg acgagcatac tttccaggag
 attacaagtg tcagccgcgc

13321 gctggggcag gaggacacgg gcagcctgga ggcaacccta
 aactacctgc tgaccaaccg
 13381 gcggcagaag atcccctcgt tgcacagttt aaacagcgag
 gaggagcgca ttttgcgcta
 13441 cgtgcagcag agcgtgagcc ttaacctgat gcgcgacggg
 gtaacgcccc gcgtggcgct
 13501 ggacatgacc gcgcgcaaca tggaaccggg catgtatgcc
 tcaaaccggc cgtttatcaa
 13561 ccgcctaattg gactacttgc atcgcgcggc cgccgtgaac
 cccgagtatt tcaccaatgc
 13621 catcttgaac ccgcaactggc taccgcccc tggtttctac
 accgggggat tcgaggtgcc
 13681 cgagggtaac gatggattcc tctgggacga catagacgac
 agcgtgtttt ccccgcaacc
 13741 gcagaccctg ctagagttgc aacagcgcgga gcaggcagag
 gcggcgctgc gaaaggaaag
 13801 cttccgcagg ccaagcagct tgtccgatct aggcgctgcg
 gccccgcggt cagatgctag
 13861 tagcccattht ccaagcttga taggggtctct taccagcact
 cgcaccaccc gccgcgcct
 13921 gctgggagcag gaggagtacc taaacaactc gctgctgcag
 ccgcagcgcg aaaaaaacct
 13981 gcctccggca tttcccaaca acgggataga gagcctagtg
 gacaagatga gtagatggaa
 14041 gacgtacgcg caggagcaca gggacgtgcc agggccgcgc
 ccgcccaccc gtcgtcaaag
 14101 gcacgaccgt cagcgggggtc tgggtgtggga ggacgatgac
 tcggcagacg acagcagcgt
 14161 cctggattttg ggagggagtg gcaaccctgt tgcgcacctt
 cgccccaggc tggggagaat
 14221 gtttttaaaaa aaaaaaagca tgatgcaaaa taaaaaactc
 accaaggcca tggcaccgag
 14281 cgttgggtttt cttgtattcc ccttagtatg cggcgcgcg
 cgatgtatga ggaaggtcct
 14341 cctccctcct acgagagtgt ggtgagcgcg gcgccagtgg
 cggcggcgct gggttctccc
 14401 ttcgatgctc ccctggaccc gccgtttgtg cctccgcggt
 acctgcggcc taccgggggg
 14461 agaaacagca tccgttactc tgagttggca cccctattcg
 acaccacccg tgtgtacctg
 14521 gtggacaaca agtcaacgga tgtggcatcc ctgaactacc
 agaacgacca cagcaacttt
 14581 ctgaccacgg tcattcaaaa caatgactac agcccggggg
 aggcaagcac acagaccatc
 14641 aatcttgacg accggtcgca ctggggcggc gacctgaaaa
 ccattctgca taccaacatg
 14701 ccaaattgtga acgagttcat gtttaccat aagtttaagg
 cgcgggtgat ggtgtcgcgc
 14761 ttgcctacta aggacaatca ggtggagctg aaatacgagt
 ggggtggagtt cacgctgccc

14821 gagggcaact actccgagac catgaccata gaccttatga
 acaacgcgat cgtggagcac
 14881 tacttgaaag tgggcagaca gaacgggggtt ctggaaagcg
 acatcgggggt aaagtttgac
 14941 acccgcaact tcagactggg gtttgacccc gtcactggtc
 ttgtcatgcc tggggatatat
 15001 acaaacgaag cttccatcc agacatcatt ttgctgccag
 gatgcgggggt ggacttcacc
 15061 cacagccgcc tgagcaactt gttgggcatc cgcaagcggc
 aacccttcca ggagggtttt
 15121 aggatcacct acgatgatct ggagggtgggt aacattccccg
 cactgttgga tgtggacgcc
 15181 taccaggcga gcttgaaaga tgacaccgaa cagggcgggg
 gtggcgcagg cggcagcaac
 15241 agcagtggca gcggcgcgga agagaactcc aacgcggcag
 ccgcggcaat gcagccgggtg
 15301 gaggacatga acgatcatgc cattcgcggc gacaccttg
 ccacacgggc tgaggagaag
 15361 cgcgctgagg ccgaagcagc ggccgaagct gccgcccccg
 ctgcgcaacc cgaggtcgag
 15421 aagcctcaga agaaaccggt gatcaaacc ctgacagagg
 acagcaagaa acgcagttac
 15481 aacctaataa gcaatgacag caccttcacc cagtaccgca
 gctggtacct tgcatacaac
 15541 tacggcgacc ctgagaccgg aatccgctca tggaccctgc
 tttgcactcc tgacgtaacc
 15601 tgcggtcgg agcaggctta ctggtcggtg ccagacatga
 tgcaagaccc cgtgacctc
 15661 cgctccacgc gccagatcag caactttccg gtggtgggcg
 ccgagctgtt gcccggtgac
 15721 tccaagagct tctacaacga ccaggccgtc tactcccaac
 tcatccgcca gtttacctct
 15781 ctgaccacg tgttcaatcg ctttcccgag aaccagattt
 tggcgcgccc gccagcccc
 15841 accatcacca ccgtcagtga aaacgttcct gctctcacag
 atcacgggac gctaccgctg
 15901 cgcaacagca tcggaggagt ccagcgagtg accattactg
 acgccagacg ccgcacctgc
 15961 ccctacgttt acaaggccct gggcatagtc tcgccgcgcg
 tcctatcgag ccgcactttt
 16021 tgagcaagca tgtccatcct tatatcgccc agcaataaca
 caggctgggg cctgcgcttc
 16081 ccaagcaaga tgtttggcgg ggccaagaag cgctccgacc
 aacacccagt gcgcgtgcgc
 16141 gggcactacc gcgcgccctg gggcgcgcac aaacgcggcc
 gactgggcg caccaccgtc
 16201 gatgacgcca tcgacgcggt ggtggaggag gcgcgcaact
 acacgcccac gccgccacca
 16261 gtgtccacag tggacgcggc cattcagacc gtggtgcgcg
 gagcccgcg ctatgctaaa

16321 atgaagagac ggcgaggcg cgtagcacgt cgccaccgcc
 gccgaccg cactgcgcc
 16381 caacgcgcg cgcgccct gcttaaccgc gcacgtcgca
 ccggccgacg ggcgccatg
 16441 cgggcccgtc gaaggctggc cgcggtatt gtcactgtgc
 cccccaggtc caggcgacga
 16501 gcggccgcg cagcagccgc ggccattagt gctatgactc
 agggtcgcag gggcaacgtg
 16561 tattgggtgc gcgactcggg tagcgccctg cgcggtgccg
 tgcgcacccg ccccccgcgc
 16621 aactagattg caagaaaaaa ctacttagac tcgtactggt
 gtatgtatcc agcggcgcg
 16681 gcgcgcaacg aagctatgtc caagcgcaaa atcaaagaag
 agatgctcca ggtcatcgcg
 16741 ccggagatct atggccccc gaagaaggaa gagcaggatt
 acaagccccg aaagctaaag
 16801 cgggtcaaaa agaaaaagaa agatgatgat gatgaacttg
 acgacgagggt ggaactgctg
 16861 cacgctaccg cgcccaggcg acgggtacag tggaaaggtc
 gacgcgtaaa acgtgttttg
 16921 cgaccgcgca ccaccgtagt ctttacgccc ggtgagcgct
 ccaccgcac ctacaagcg
 16981 gtgtatgatg aggtgtacgg cgacgaggac ctgcttgagc
 aggccaacga gcgcctcggg
 17041 gagtttgct acggaaagcg gcataaggac atgctggcgt
 tgccgctgga cgagggcaac
 17101 ccaacaccta gcctaaagcc cgtaacactg cagcagggtgc
 tgcccgcgct tgcaccgtcc
 17161 gaagaaaagc gcggcctaaa gcgcgagtct ggtgacttgg
 caccaccgt gcagctgatg
 17221 gtaccaaacg gccagcgact ggaagatgtc ttggaaaaaa
 tgaccgtgga acctgggctg
 17281 gagcccgagg tccgcgtgcg gccaatcaag cagggtggcg
 cgggactggg cgtgcagacc
 17341 gtggacgttc agataccac taccagtagc accagtattg
 ccaccgccac agagggcatg
 17401 gagacacaaa cgtccccggt tgcctcagcg gtggcggatg
 ccgcggtgca ggcggtcgt
 17461 gcggccgcgt ccaagacctc tacggagggtg caaacggacc
 cgtggatgtt tcgcgtttca
 17521 gccccccggc gcccgcgcg ttcgaggaag tacggcgccg
 ccagcgcgct actgcccga
 17581 tatgccctac atccttccat tgcgcctacc cccggctatc
 gtggctacac ctaccgcccc
 17641 agaagacgag caactaccg acgccgaacc accactggaa
 ccgcccgcg ccgtcgccgt
 17701 cgccagcccg tgetggcccc gatttcggtg cgcagggtgg
 ctgcgaagg aggcaggacc
 17761 ctggtgctgc caacagcgcg ctaccacccc agcatcgttt
 aaaagccggt ctttgtggtt

17821 cttgcagata tggccctcac ctgccgctc cgtttcccgg
 tgccgggatt ccgaggaaga
 17881 atgcaccgta ggaggggcat ggccggccac ggcctgacgg
 gcggcatgcg tcgtgcgcac
 17941 caccggcggc ggcgcgcgtc gcaccgtcgc atgcgcggcg
 gtatcctgcc cctccttatt
 18001 ccactgatcg ccgcggcgat tggcgccgtg cccggaattg
 catccgtggc cttgcaggcg
 18061 cagagacact gattaaaaac aagttgcatg tggaaaaatc
 aaaataaaaa gtctggactc
 18121 tcacgctcgc ttggtcctgt aactattttg tagaatggaa
 gacatcaact ttgcgtctct
 18181 ggccccgcga cacggctcgc gcccgttcat gggaaactgg
 caagatatcg gcaccagcaa
 18241 tatgagcggg ggcgccctca gctggggctc gctgtggagc
 ggcattaaaa atttcgggtc
 18301 caccgttaag aactatggca gcaaggcctg gaacagcagc
 acaggccaga tgcctgaggga
 18361 taagttgaaa gagcaaaatt tccaacaaaa ggtggtagat
 ggcctggcct ctggcattag
 18421 cgggggtggg gacctggcca accaggcagt gcaaaataag
 attaacagta agcttgatcc
 18481 ccgccctccc gtagaggagc ctccaccggc cgtggagaca
 gtgtctccag aggggcgtgg
 18541 cgaaaagcgt ccgcgccccg acagggaaga aactctgggtg
 acgcaaatag acgagcctcc
 18601 ctcgtaacgag gaggcactaa agcaaggcct gccaccacc
 cgtecccatcg cgcccatggc
 18661 taccggagtg ctgggcccagc acacaccgt aacgctggac
 ctgcctcccc ccgccgacac
 18721 ccagcagaaa cctgtgctgc caggcccgac cgccgttggt
 gtaaccgctc ctagccgcgc
 18781 gtccctgcgc cgcgccgcca gcggtccgcg atcgttgcgg
 cccgtagcca gtggcaactg
 18841 gcaaagcaca ctgaacagca tcgtgggtct gggggtgcaa
 tccctgaagc gccgacgatg
 18901 cttctgaata gctaacgtgt cgtatgtgtg tcatgtatgc
 gtccatgtcg ccgccagagg
 18961 agctgctgag ccgccgcgcg cccgctttcc aagatggcta
 ccccttcgat gatgcgcgag
 19021 tgggtcttaca tgcacatctc gggccaggac gcctcggagt
 acctgagccc cgggctgggtg
 19081 cagtttgccc gcgccaccga gacgtacttc agcctgaata
 acaagtttag aaaccccacg
 19141 gtggcgccta cgcacgacgt gaccacagac cggteccagc
 gtttgacgct gcggttcac
 19201 cctgtggacc gtgaggatac tgcgtactcg tacaaggcgc
 ggttcaccct agctgtgggt
 19261 gataaccgtg tgcctggacat ggcttccacg tactttgaca
 tccgcggcgt gctggacagg

19321 ggccctactt ttaagcccta ctctggcact gcctacaacg
 ccctggctcc caaggggtgcc
 19381 ccaaatcctt gcgaatggga tgaagctgct actgctcttg
 aaataaacct agaagaagag
 19441 gacgatgaca acgaagacga agtagacgag caagctgagc
 agcaaaaaac tcacgtatatt
 19501 gggcaggcgc cttattcttg tataaatatt acaaaggagg
 gtattcaaatt aggtgtcgaa
 19561 ggtcaaacac cttaaataatgc cgataaaaca tttcaacctg
 aacctcaaatt aggagaatct
 19621 cagtgggtacg aaactgaaat taatcatgca gctgggagag
 tccttaaaaa gactacccca
 19681 atgaaacctat gttacgggtc atatgcaaaa cccacaaatg
 aaaatggagg gcaaggcatt
 19741 cttgtaaagc aacaaaatgg aaagctagaa agtcaagtgg
 aaatgcaatt tttctcaact
 19801 actgaggcga ccgcaggcaa tgggtgataac ttgactccta
 aagtgggtatt gtacagtga
 19861 gatgtagata tagaaacccc agacactcat atttcttaca
 tgcccactat taaggaaggt
 19921 aactcacgag aactaatggg ccaacaatct atgcccacaa
 ggccctaatta cattgctttt
 19981 agggacaatt ttattgggtct aatgtattac aacagcacgg
 gtaatatggg tgttctggcg
 20041 ggccaagcat cgcagttgaa tgctgttgta gatttgcaag
 acagaaacac agagctttca
 20101 taccagcttt tgcttgattc cattgggtgat agaaccaggt
 acttttctat gtggaatcag
 20161 gctgttgaca gctatgatcc agatgttaga attattgaaa
 atcatggaac tgaagatgaa
 20221 cttccaaatt actgctttcc actgggaggt gtgattaata
 cagagactct taccaaggta
 20281 aaacctaata caggtcagga aaatggatgg gaaaaagatg
 ctacagaatt ttcagataaa
 20341 aatgaaataa gagttggaaa taattttgcc atggaaatca
 atctaaatgc caacctgtgg
 20401 agaaatttcc tgtactccaa catagcgctg tatttgccccg
 acaagctaaa gtacagtcct
 20461 tccaacgtaa aaatttctga taaccacaaac acctacgact
 acatgaacaa gcgagtgggtg
 20521 gctcccgggt tagtggactg ctacattaac cttggagcac
 gctggtcctt tgactatatg
 20581 gacaacgtca acccatttaa ccaccaccgc aatgctggcc
 tgcgctaccg ctcaatgttg
 20641 ctgggcaatg gtcgctatgt gcccttcac atccaggtgc
 ctcagaagtt ctttgccatt
 20701 aaaaacctcc ttctcctgcc gggctcatac acctacgagt
 ggaacttcag gaaggatgtt
 20761 aacatgggtc tgcagagctc cctaggaaat gaçctaaggg
 ttgacggagc cagcattaag

20821 tttgatagca tttgccttta cgccaccttc ttccccatgg
 cccacaacac cgcctccacg
 20881 cttgaggcca tgcttagaaa cgacaccaac gaccagtcct
 ttaacgacta tctctccgcc
 20941 gccaacatgc tctaccctat acccgccaac gctaccaacg
 tgcccatatc catcccctcc
 21001 cgcaactggg cggcttttcg cggctgggccc ttcacgcgcc
 ttaagactaa ggaaacccca
 21061 tcaactgggt cgggctacga cccttattac acctactctg
 gctctatacc ctacctagat
 21121 ggaacctttt acctcaacca cacctttaag aaggtggcca
 ttacctttga ctcttctgtc
 21181 agctggcctg gcaatgaccg cctgcttacc cccaacgagt
 ttgaaattaa gcgctcagtt
 21241 gacggggagg gttacaacgt tgcccagtggt aacatgacca
 aagactgggt cctggtacaa
 21301 atgctagcta actacaacat tggctaccag ggcttctata
 tcccagagag ctacaaggac
 21361 cgcagtact ccttcttttag aaacttccag cccatgagcc
 gtcaggtgggt ggatgatact
 21421 aaatacaagg actaccaaca ggtgggcatc ctacaccaac
 acaacaactc tggatttggt
 21481 ggctaccttg cccccaccat gcgcgaagga caggcctacc
 ctgctaactt cccctatccg
 21541 cttataggca agaccgcagt tgacagcatt acccagaaaa
 agtttctttg cgatcgcacc
 21601 ctttggcgca tcccattctc cagtaacttt atgtccatgg
 gcgcactcac agacctgggc
 21661 caaaaccttc tctacgcaa ctccgcccac gcgctagaca
 tgacttttga ggtggatccc
 21721 atggacgagc ccacccttct ttatgttttg tttgaagtct
 ttgacgtgggt ccgtgtgcac
 21781 cggccgcacc gcggcgctcat cgaaaccgtg tacctgcgca
 cgcccttctc ggccgggcaac
 21841 gccacaacat aaagaagcaa gcaacatcaa caacagctgc
 cgccatgggc tccagtgagc
 21901 aggaactgaa agccattgtc aaagatcttg gttgtgggccc
 atattttttg ggcacctatg
 21961 acaagcgctt tccaggcttt gtttctccac acaagctcgc
 ctgcgccata gtcaatacgg
 22021 ccggtcgcga gactgggggc gtacactgga tggcctttgc
 ctggaacccg cactcaaaaa
 22081 catgctacct ctttgagccc tttggctttt ctgaccagcg
 actcaagcag gtttaccagt
 22141 ttgagtacga gtcactcctg cgccgtagcg ccattgcttc
 ttcccccgac cgctgtataa
 22201 cgctggaaaa gtccacccaa agcgtacagg ggcccaactc
 ggccgcctgt ggactattct
 22261 gctgcatggt tctccacgcc tttgccaact ggccccaaac
 tcccattgat cacaacccca

22321 ccatgaacct tattaccggg gtacccaact ccatgctcaa
 cagtcgccag gtacagccca
 22381 ccctgcgtcg caaccaggaa cagctctaca gcttcctgga
 gcgccactcg ccctacttcc
 22441 gcagccacag tgcgcagatt aggagcgcca cttctttttg
 tcacttgaaa aacatgtaaa
 22501 aataatgtac tagagacact ttcaataaag gcaaagtctt
 ttatttgtac actctcgggt
 22561 gattatttac cccaccctt gccgtctgcg ccgtttaaaa
 atcaaagggg ttctgccgcg
 22621 catcgctatg cgccactggc agggacacgt tgcgatactg
 gtgttttagtg ctccacttaa
 22681 actcaggcac aaccatccgc ggcagctcgg tgaagttttc
 actccacagg ctgcgcacca
 22741 tcaccaacgc gtttagcagg tcgggcgccg atatcttgaa
 gtcgcagttg gggcctccgc
 22801 cctgcgcgcg cgagttgcga tacacagggg tgcagcactg
 gaacactatc agcgccgggt
 22861 ggtgcacgt ggccagcacg ctcttgctcg agatcagatc
 cgcgctccagg tcctccgcgt
 22921 tgctcagggc gaacggagtc aactttggta gctgccttcc
 caaaaagggc gcgtgccag
 22981 gctttgagtt gcactcgcac cgtagtgga tcaaaagggt
 accgtgccg gtctgggcgt
 23041 taggatacag cgccctgcata aaagccttga tctgcttaaa
 agccacctga gcctttgcg
 23101 cttcagagaa gaacatgccg caagacttgc cggaaaactg
 attggccgga caggccgcgt
 23161 cgtgcacgca gcaccttgcg tcgggtgttg agatctgcac
 cacatttcgg cccaccgggt
 23221 tcttcacgat cttggccttg ctagactgct ccttcagcgc
 gcgctgccg ttttcgctcg
 23281 tcacatccat ttcaatcacg tgctccttat ttatcataat
 gcttccgtgt agacacttaa
 23341 gctgccttc gatctcagcg cagcgggtgca gccacaacgc
 gcagcccgtg ggctcgtgat
 23401 gcttgtaggt cacctctgca aacgactgca ggtacgcctg
 caggaatcgc cccatcatcg
 23461 tcacaaagggt cttgttgctg gtgaagggtca gctgcaaccc
 gcggtgctcc tcgttcagcc
 23521 aggtcttgca tacggccgcc agagcttcca cttggtcagg
 cagtagtttg aagttcgct
 23581 ttagatcgtt atccacgtgg tacttggtcca tcagcgcgcg
 cgcagcctcc atgcccttct
 23641 cccacgcaga cagatcggc aactcagcg ggttcatcac
 cgtaatttca ctttccgctt
 23701 cgctgggctc ttctcttcc tcttgctcc gcataaccag
 cgccactggg tcgtcttcat
 23761 tcagccgcg cactgtgcg ttacctcctt tgccatgctt
 gattagcacc ggtgggttgc

23821 tgaaacccac catttgtagc gccacatctt ctctttcttc
 ctcgctgtcc acgattacct
 23881 ctggtgatgg cgggcgctcg ggcttgggag aagggcgctt
 ctttttcttc ttgggcgcaa
 23941 tggccaaatc cgccgccgag gtcgatggcc gcgggctggg
 tgtgcgcggc accagcgcgt
 24001 cttgtgatga gtcttctctg tcctcggact cgatacgccg
 cctcatccgc ttttttgggg
 24061 gcgcccgggg aggcggcggc gacggggacg gggacgacac
 gtcttccatg gttgggggac
 24121 gtcgcgccgc accgcgtccg cgctcggggg tggtttcgcg
 ctgctcctct tcccgaactg
 24181 ccattttcctt ctcttatagg cagaaaaaga tcatggagtc
 agtcgagaag aaggacagcc
 24241 taaccgcccc ctctgagttc gccaccaccg cctccaccga
 tgccgccaac gcgcctacca
 24301 ccttccccgt cgaggcacc cgccttgagg aggaggaagt
 gattatcgag caggacccag
 24361 gttttgtaag cgaagacgac gaggaccgct cagtaccaac
 agaggataaa aagcaagacc
 24421 aggacaacgc agaggcaaac gaggaacaag tcgggcgggg
 ggacgaaagg catggcgact
 24481 acctagatgt gggagacgac gtgctgttga agcatctgca
 gcgccagtgc gccattatct
 24541 gcgacgcgtt gcaagagcgc agcgatgtgc ccctcgccat
 agcggatgtc agccttgctt
 24601 acgaacgcca cctattctca ccgcgcgtac cccccaacg
 ccaagaaaac ggcacatgcg
 24661 agcccaaccc gcgcctcaac ttctaccccg tatttgccgt
 gccagaggtg cttgccacct
 24721 atcacatctt ttccaaaac tgcaagatac ccctatcctg
 ccgtgccaac cgcagccgag
 24781 cggacaagca gctggccttg cggcaggggc ctgtcatacc
 tgatatcgcc tcgctcaacg
 24841 aagtgccaaa aatctttgag ggtcttgac gcgacgagaa
 gcgcgcggca aacgctctgc
 24901 aacaggaaaa cagcgaaaat gaaagtcact ctggagtgtt
 ggtggaactc gagggtgaca
 24961 acgcgcgcct agccgtacta aaacgcagca tcgaggtcac
 ccactttgcc taccggcac
 25021 ttaacctacc cccaaggctc atgagcacag tcatgagtga
 gctgatcgtg cgccgtgcgc
 25081 agcccctgga gagggatgca aatttgcaag aacaaacaga
 ggagggccta cccgcagttg
 25141 gcgacgagca gctagcgcgc tggcttcaaa cgcgcgagcc
 tgccgacttg gaggagcgac
 25201 gcaaactaat gatggccgca gtgctcgta ccgtggagct
 tgagtgcatt cagcggttct
 25261 ttgctgacct ggagatgcag' cgcaagctag aggaaacatt
 gcactacacc ttctgacagg

25321 gctacgtacg ccaggcctgc aagatctcca acgtggagct
ctgcaacctg gtctcctacc
25381 ttggaatttt gcacgaaaac cgccttgggc aaaacgtgct
tcattccacg ctcaagggcg
25441 aggcgcgccg cgactacgtc cgcgactgcg tttacttatt
tctatgctac acctggcaga
25501 cggccatggg cgtttggcag cagtgccttg aggagtgcaa
cctcaaggag ctgcagaaac
25561 tgctaaagca aaacttgaag gacctatgga cggccttcaa
cgagcgctcc gtggccgcg
25621 acctggcgga catcattttc cccgaacgcc tgcttaaaac
cctgcaacag ggtctgccag
25681 acttcaccag tcaaagcatg ttgcagaact ttaggaactt
tattcctagag cgctcaggaa
25741 tcttgcccg cacttgctgt gcacttccta gcgactttgt
gccattaag taccgcgaat
25801 gccctccgcc gctttggggc cactgctacc ttctgcagct
agccaactac cttgcctacc
25861 actctgacat aatggaagac gtgagcgggtg acggtctact
ggagtgtcac tgtcgtgca
25921 acctatgcac ccgcaccgc tccctggttt gcaattcgca
gctgcttaac gaaagtcaaa
25981 ttatcggtac ctttgagctg cagggtccct cgcctgacga
aaagtccgcg gctccgggggt
26041 tgaaactcac tccggggctg tggacgtcgg cttaccttgc
caaatttgta cctgaggact
26101 accacgcca cgagattagg ttctacgaag accaatcccg
cccgccaaat gcggagctta
26161 ccgcctgcgt cattaccag ggccacattc ttggccaatt
gcaagccatc acaaaagccc
26221 gccaagagtt tctgctacga aaggacggg gggtttactt
ggacccccag tccggcgagg
26281 agctcaacc aatcccccg ccgcccgcgc cctatcagca
gcagccgcgg gcccttgctt
26341 ccaggatgg cacccaaaaa gaagctgcag ctgccgcgcg
cacccacgga cgaggaggaa
26401 tactgggaca gtcaggcaga ggaggttttg gacgaggagg
aggaggacat gatggaagac
26461 tgggagagcc tagacgagga agcttccgag gtcgaagagg
tgtcagacga aacaccgtca
26521 ccctcggtcg cattcccctc gccggcgccc cagaaatcgg
caaccggttc cagcatggct
26581 acaacctcg ctctcaggc gccggcggca ctgcccgttc
gccgaccxaa ccgtagatgg
26641 gacaccactg gaaccagggc cggtaagtcc aagcagccgc
cgccgttagc ccaagagcaa
26701 caacagcgcc aaggctaccg ctcatggcg gggcacaaga
acgccatagt tgcttgcttg
26761 caagactgtg ggggcaacat ctcttcgcc cgccgctttc
ttctctacca tcacggcgtg

26821 gccttcccc gtaacatcct gcattactac cgtcatctct
 acagcccata ctgcacggc
 26881 ggcagcggca gcggcagcaa cagcagcggc cacacagaag
 caaaggcgac cggatagcaa
 26941 gactctgaca aagcccaaga aatccacagc ggcggcagca
 gcaggaggag gagcgctgcg
 27001 tctggcgccc aacgaacccg tatcgacccg cgagcttaga
 aacaggattt ttcccactct
 27061 gtatgctata tttcaacaga gcaggggcca agaacaagag
 ctgaaaataa aaaacaggtc
 27121 tctgcatcc ctcacccgca gctgcctgta tcacaaaagc
 gaagatcagc ttcggcgcac
 27181 gctggaagac gcggaggctc tcttcagtaa atactgcgag
 ctgactctta aggactagtt
 27241 tcgcgccctt tctcaaattt aagcgcgaaa actacgtcat
 ctccagcggc cacacccggc
 27301 gccagcacct gtcgtcagcg ccattatgag caaggaaatt
 cccacgccct acatgtggag
 27361 ttaccagcca caaatgggac ttgctggctgg agctgcccac
 gactactcaa cccgaataaa
 27421 ctacatgagc gcgggacccc acatgatata ccgggtcaac
 ggaatccgag cccaccgaaa
 27481 ccgaattctc ttggaacagg cggctattac caccacacct
 cgtaataacc ttaatccccg
 27541 tagttggccc gctgccctgg tgtaccagga aagtcccgt
 cccaccactg tggacttcc
 27601 cagagacgcc caggccgaag ttcagatgac taactcaggg
 gcgcagcttg cgggcggctt
 27661 tcgtcacagg gtgcggctgc ccgggcaggg tataactcac
 ctgacaatca gagggcgagg
 27721 tattcagctc aacgacgagt cggtagctc ctgccttgg
 ctccgtccg acgggacatt
 27781 tcagatcggc ggcgccggcc gtccttcatt cagcctcgt
 caggcaatcc taactctgca
 27841 gacctcgtcc tctgagccgc gctctggagg cattggaact
 ctgcaattta ttgaggagtt
 27901 tgtgccatcg gtctacttta accccttctc gggacctccc
 ggccactata cggatcaatt
 27961 tattcctaac tttgacgcgg taaaggactc ggcggacggc
 tacgactgaa tgttaagtgg
 28021 agaggcagag caactgcgcc tgaaacacct ggtccactgt
 cgccgccaca agtgctttgc
 28081 ccgcgactcc ggtgagtttt gctactttga attgcccag
 gatcatatcg agggcccggc
 28141 gcacggcgctc cggttaccg ccagggaga gcttgcccgt
 agcctgattc gggagtttac
 28201 ccagcgcccc ctgctagttg agcgggacag gggacctgt
 gttctcactg tgatttgcaa
 28261 ctgtcctaac cttggattac atcaagatcc tctagttata
 actagagtac ccggggatct

28321 tattcccttt aactaataaa aaaaaataat aaagcatcac
 ttacttaaaa tcagtttagca
 28381 aattttctgtc cagttttattc agcagcacct ccttgccctc
 ctcccagctc tgggtattgca
 28441 gcttcctcct ggctgcaaac tttctccaca atctaaatgg
 aatgtcagtt tcctcctggt
 28501 cctgtccatc cgcacccact atcttcatgt tgttgcagat
 gaagcgcgca agaccgtctg
 28561 aagatacctt caaccccggtg tatccatatg acacggaaac
 cggtcctcca actgtgcctt
 28621 ttcttactcc tccctttgta tcccccaatg gggttcaaga
 gagtccccct ggggtactct
 28681 ctttgcgctt atccgaacct ctagttacct ccaatggcat
 gcttgcgctc aaaatgggca
 28741 acggcctctc tctggacgag gccggcaacc ttacctcca
 aaatgtaacc actgtgagcc
 28801 cacctctcaa aaaaaccaag tcaaacataa acctggaaat
 atctgcaccc ctacagtta
 28861 cctcagaagc cctaactgtg gctgccgccg cacctcta
 ggtcgcgggc aacacactca
 28921 ccatgcaatc acaggccccg ctaaccgtgc acgactccaa
 acttagcatt gccacccaag
 28981 gacccctcac agtgtcagaa ggaaagctag ccctgcaaac
 atcaggcccc ctaccacca
 29041 cogatagcag tacccttact atcactgcct cccccctct
 aactactgcc actggttagct
 29101 tgggcattga cttgaaagag cccatttata cacaaaatgg
 aaaactagga ctaaagtacg
 29161 gggctccttt gcatgtaaca gacgacctaa acactttgac
 cgtagcaact ggtccagggtg
 29221 tgactattaa taatacttcc ttgcaaacta aagttactgg
 agccttgggt tttgattcac
 29281 aaggcaatat gcaacttaat gtagcaggag gactaaggat
 tgattctcaa aacagacgcc
 29341 ttatacttga tgtttagttat ccgtttgatg ctcaaaacca
 actaaatcta agactaggac
 29401 agggccctct ttttataaac tcagcccaca acttggatat
 taactacaac aaaggccttt
 29461 acttgtttac agcttcaaac aattccaaaa agcttgagggt
 taacctaagc actgccaagg
 29521 ggttgatggt tgacgctaca gccatagcca ttaatgcagg
 agatgggctt gaatttggtt
 29581 cacctaatgc accaaacaca aatcccctca aaacaaaaat
 tggccatggc ctagaatttg
 29641 attcaaaca ggctatgggt cctaaactag gaactggcct
 tagttttgac agcacagggt
 29701 ccattacagt aggaacaaa aataatgata agctaacttt
 gtggaccaca ccagctccat
 29761 ctcttaactg tagactaaat gcagagaaag atgctaaact
 cactttggct ttaacaaaat

29821 gtggcagtc aatacttgct acagtttcag ttttggctgt
 taaaggcagt ttggctccaa
 29881 tatctggaac agttcaaagt gctcatctta ttataagatt
 tgacgaaaat ggagtgtac
 29941 taaacaattc ctctctggac ccagaatatt ggaactttag
 aaatggagat cttactgaag
 30001 gcacagccta tacaaacgct gttggattta tgcctaacct
 atcagcttat ccaaaatctc
 30061 acggtaaaac tgccaaaagt aacattgtca gtcaagttta
 cttaaaccga gacaaaacta
 30121 aacctgtaac actaaccatt aactaaacg gtacacagga
 aacaggagac acaactccaa
 30181 gtgcatactc tatgtcattt tcatgggact ggtctggcca
 caactacatt aatgaaatat
 30241 ttgccacatc ctcttacact ttttcataca ttgcccaaga
 ataaagaatc gtttgtgtta
 30301 tgtttcaacg tgtttatfff tcaattgcag aaaatttcaa
 gtcattttttc attcagtagt
 30361 atagcccccac caccacatag cttatacaga tcaccgtacc
 ttaatcaaac tcacagaacc
 30421 ctagtattca acctgccacc tccctcccaa cacacagagt
 acacagtcct ttctccccg
 30481 ctggccttaa aaagcatcat atcatgggta acagacatat
 tcttaggtgt tatattccac
 30541 acggtttcct gtcgagccaa acgctcatca gtgatattaa
 taaactcccc gggcagctca
 30601 cttaagttca tgtcgctgtc cagctgctga gccacaggct
 gctgtccaac ttgcggttgc
 30661 ttaacgggcg gcgaaggaga agtccacgcc tacatggggg
 tagagtcata atcgtgcac
 30721 aggatagggc ggtggtgctg cagcagcgcg cgaataaact
 gctgccgccc ccgctccgctc
 30781 ctgcaggaat acaacatggc agtggctctc tcagcgatga
 ttgcgaccgc ccgcagcata
 30841 aggcgccttg tcctccgggc acagcagcgc accctgatct
 cacttaaate agcacagtaa
 30901 ctgcagcaca gcaccacaat attggttcaa atcccacagt
 gcaaggcgct gtatccaaag
 30961 ctcatggcgg ggaccacaga acccagctgg ccatcatacc
 acaagcgcag gtagattaag
 31021 tggcgacccc tcataaacac gctggacata aacattacct
 cttttggcat gttgtaattc
 31081 accacctccc ggtaccatat aaacctctga ttaaaccatgg
 cgccatccac caccatccta
 31141 aaccagctgg ccaaaacctg cccgcccggct atacactgca
 gggaaccggg actggaacaa
 31201 tgacagtgga gagcccagga ctcgtaacca tggatcatca
 tgctcgtcat gatatcaatg
 31261 ttggcacaac acaggcacac gtgcatacac ttcctcagga
 ttacaagctc ctcccgcgtt

31321 agaaccatat cccagggaac aacccattcc tgaatcagcg
 taaatcccac actgcaggga
 31381 agacctcgca cgtaactcac gttgtgcatt gtcaaagtgt
 tacattcggg cagcagcgga
 31441 tgatcctcca gtatggtagc gcgggtttct gtctcaaaag
 gaggtagacg atccctactg
 31501 tacggagtgc gccgagacaa ccgagatcgt gttggtcgta
 gtgtcatgcc aaatggaacg
 31561 ccggacgtag tcatatttcc tgaagcaaaa ccagggtgcgg
 gcgtgacaaa cagatctgcg
 31621 tctccggtct cgccgcttag atcgctctgt gtagtagttg
 tagtatatcc actctctcaa
 31681 agcatccagg cgccccctgg cttcggggttc tatgtaaact
 ccttcatgcg ccgctgccct
 31741 gataacatcc accaccgcag aataagccac acccagccaa
 cctacacatt cgttctgcga
 31801 gtcacacacg ggaggagcgg gaagagctgg aagaaccatg
 tttttttttt tattccaaaa
 31861 gattatccaa aacctcaaaa tgaagatcta ttaagtgaac
 gcgctccccct ccggtggcgt
 31921 ggtcaaactc tacagccaaa gaacagataa tggcatttgt
 aagatgttgc acaatggcct
 31981 ccaaaaggca aacggccctc acgtccaagt ggacgtaaag
 gctaaaccct tcagggtgaa
 32041 tctcctctat aaacattcca gcaccttcaa ccatgcccaa
 ataattctca tctcgccacc
 32101 ttctcaatat atctctaagc aaatcccga tattaagtcc
 ggccattgta aaaatctgct
 32161 ccagagcgcc ctccaccttc agcctcaagc agcgaatcat
 gattgcaaaa attcagggtc
 32221 ctacacagacc tgtataagat tcaaaagcgg aacattaaca
 aaaataccgc gatcccgtag
 32281 gtcccttcgc agggccagct gaacataatc gtgcaggctc
 gcacggacca gcgcggccac
 32341 ttccccgcca ggaaccttga caaaagaacc cacactgatt
 atgacacgca tactcggagc
 32401 tatgctaacc agcgtagccc cgatgtaagc tttgttgc
 gggcggcgat ataaaatgca
 32461 aggtgctgct caaaaaatca ggcaaagcct cgcgcaaaaa
 agaaagcaca tcgtagtcat
 32521 gtcattgcag ataaaggcag gtaagctccg gaaccaccac
 agaaaaagac accatttttc
 32581 tctcaaacat gtctgcgggt ttctgcataa acacaaaata
 aaataacaaa aaaacattta
 32641 aacattagaa gcctgtctta caacaggaaa aacaaccctt
 ataagcataa gacggactac
 32701 ggccatgccg gcgtgaccgt aaaaaaactg gtcaccgtga
 ttaaaaagca ccaccgacag
 32761 ctctcgggtc atgtccggag tcataatgta agactcggta
 aacacatcag gttgattcat

32821 cggtcagtgc taaaaagcga ccgaaatagc ccgggggaat
acatacccg c aggcgtagag
32881 acaacattac agccccata ggaggtataa caaaattaat
aggagagaaa aacacataaa
32941 cacctgaaaa accctcctgc ctaggcaaaa tagcaccctc
ccgctccaga acaacataca
33001 gcgcttcaca gcggcagcct aacagtcagc cttaccagta
aaaaagaaaa cctattaataa
33061 aaacaccact cgacacggca ccagctcaat cagtcacagt
gtaaaaaagg gccaaagtgc
33121 gagcgagtat atataggact aaaaaatgac gtaacgggta
aagtccacaa aaaacaccca
33181 gaaaaccgca cgcgaaaccta cgcccagaaa cgaaagccaa
aaaaccacaa acttcctcaa
33241 atcgtcactt ccgttttccc acgttacgta acttcccatt
ttaagaaaaac tacaattccc
33301 aacacataca agttactcgc ccctaaaacc tacgtcaccc
gccccgttcc cagccccgc
33361 gccacgtcac aaactccacc ccctcattat catattggct
tcaatccaaa ataaggtata
33421 ttattgatga tnnnnnttaa t

WHAT IS CLAIMED IS:

1. A chimeric adenoviral expression vector, said vector comprising an expression cassette comprising the following elements:
 - (a) a first promoter operably linked to a nucleic acid encoding a toll-like receptor-3 (TLR-3) agonist; and
 - (b) a second promoter operably linked to a nucleic acid encoding a heterologous polypeptide.
2. The chimeric adenoviral expression vector of claim 1, wherein the nucleic acid encoding the TLR-3 agonist comprises a nucleic acid encoding a dsRNA.
3. The chimeric adenoviral expression vector of claim 1, wherein the nucleic acid encoding the TLR-3 agonist comprises a sequence selected from the group consisting of: SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11, and, 12.
4. The chimeric adenoviral expression vector of claim 1, wherein the heterologous polypeptide is an HIV env polypeptide.
5. The chimeric adenoviral expression vector of claim 4, wherein the HIV env polypeptide is selected from the group consisting of: gp41, gp120, and gp160.
6. The chimeric adenoviral expression vector of claim 1, wherein the heterologous polypeptide is an influenza HA polypeptide.
7. The chimeric adenoviral expression vector of claim 1, wherein the first promoter and the second promoter are the same.
8. The chimeric adenoviral expression vector of claim 7, wherein the first promoter and the second promoter are each a CMV promoter.
9. An immunogenic composition comprising the expression vector of claim 1 and a pharmaceutically acceptable carrier.
10. A method for eliciting an immune response, said method comprising administering an immunogenically effective amount of the vector of claim 1 to a mammalian subject, wherein the immune response is directed against the heterologous polypeptide, and

wherein the route of administration is selected from the group consisting of: oral, intranasal, and mucosal.

11. The method of claim 10, wherein the heterologous polypeptide is expressed in a cell selected from the group consisting of: a dendritic cell, a microfold cell, and an intestinal epithelial cell.

12. The method of claim 10, wherein said mammal is a human.

13. An immunogenic composition, said composition comprising:

(a) a chimeric adenoviral expression vector comprising a promoter operably linked to a nucleic acid encoding a heterologous polypeptide;

(b) a TLR-3 agonist; and

(c) a pharmaceutically acceptable carrier.

14. The composition of claim 13, wherein the promoter is a CMV promoter.

15. The composition of claim 13, wherein the TLR-3 agonist is selected from the group consisting of: dsRNA and poly I:C.

16. The composition of claim 13, wherein the heterologous polypeptide is an HIV env polypeptide.

17. The composition of claim 16, wherein the HIV env polypeptide is selected from the group consisting of: gp41, gp120, and gp160.

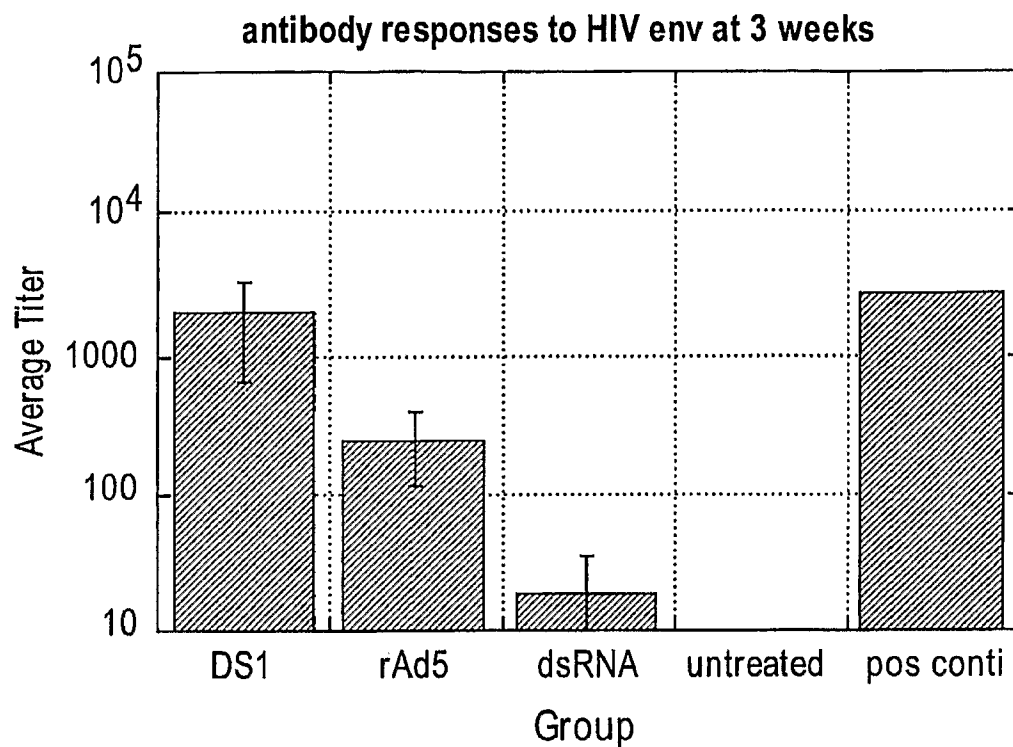
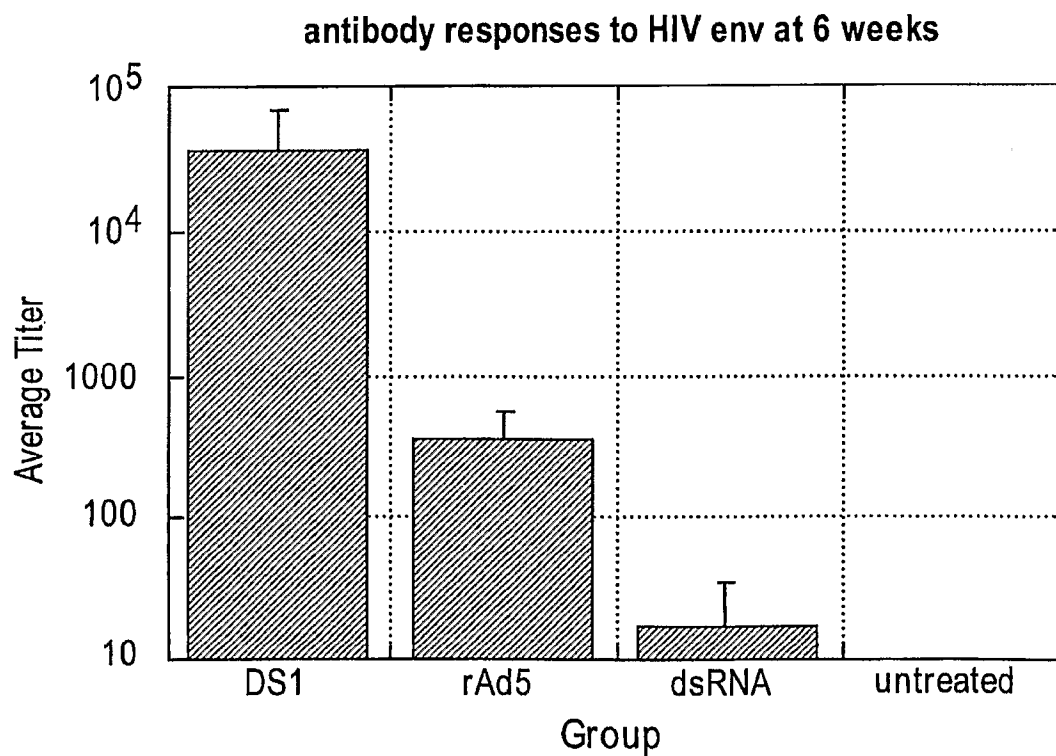
18. The composition of claim 13, wherein the heterologous polypeptide is an influenza HA polypeptide.

19. The composition of claim 13, wherein the TLR-3 agonist is administered within 48 hours of the administration of the heterologous polypeptide.

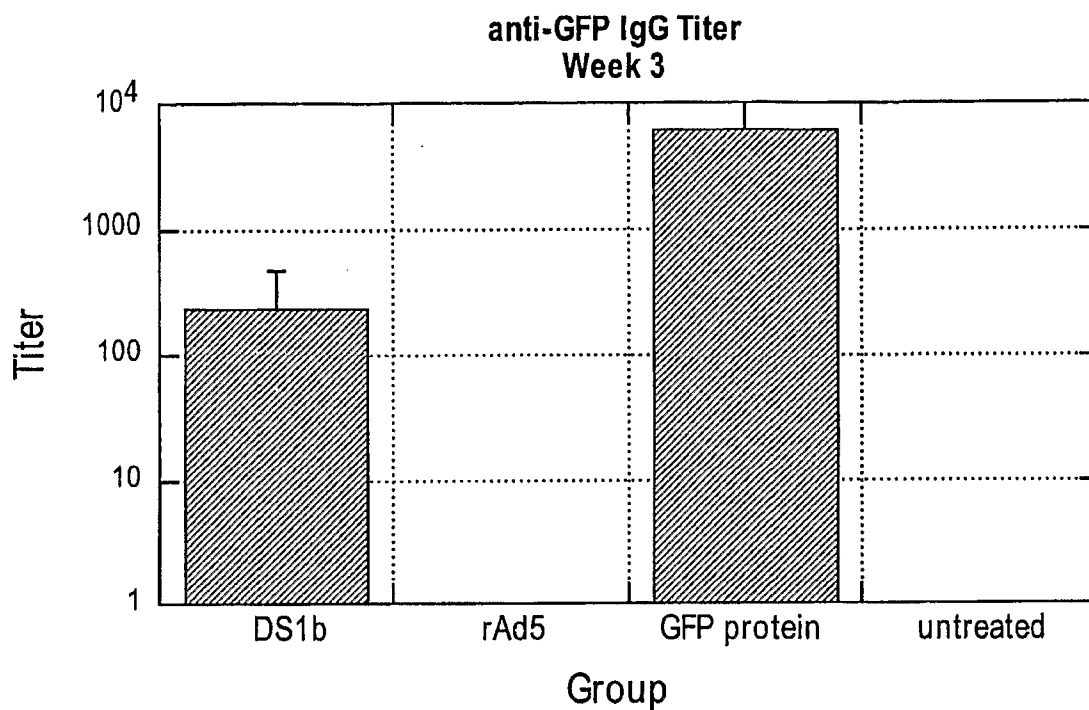
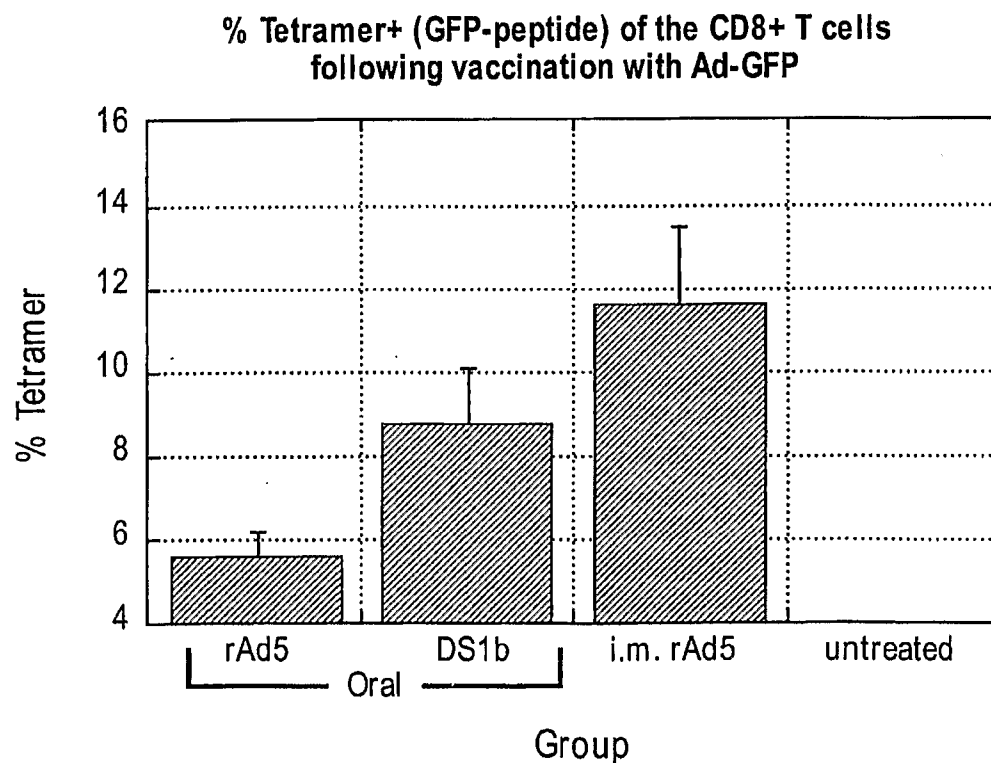
20. The composition of claim 13, wherein the TLR-3 agonist is a nucleic acid comprising a sequence selected from the group consisting of: SEQ ID NOS: 3, 4, 5, 8, 9, 10, 11, and, 12.

21. A method for eliciting an immune response, the method comprising administering to a mammalian subject
- (a) a chimeric adenoviral expression vector comprising a promoter operably linked to a nucleic acid encoding a heterologous polypeptide;
 - (b) a TLR-3 agonist,
- wherein the immune response is directed against the heterologous polypeptide.
22. The method of claim 21, wherein the promoter is selected from the group consisting of the: CMV promoter and the human beta actin promoter.
23. The method of claim 21, wherein the heterologous polypeptide is an HIV env polypeptide.
24. The method of claim 23, wherein the HIV env polypeptide is selected from the group consisting of: gp41, gp120 and gp160.
25. The method of claim 21, wherein the heterologous polypeptide is an influenza HA polypeptide.
26. The method of claim 21, wherein the TLR-3 agonist is a dsRNA.
27. The method of claim 21, wherein the TLR-3 agonist is administered within 48 hours of the administration of the chimeric adenoviral expression vector.
28. An isolated nucleic acid comprising the sequence set forth in SEQ ID NO: 1, 2, 3, 7, 14, 15, 16, or 17.

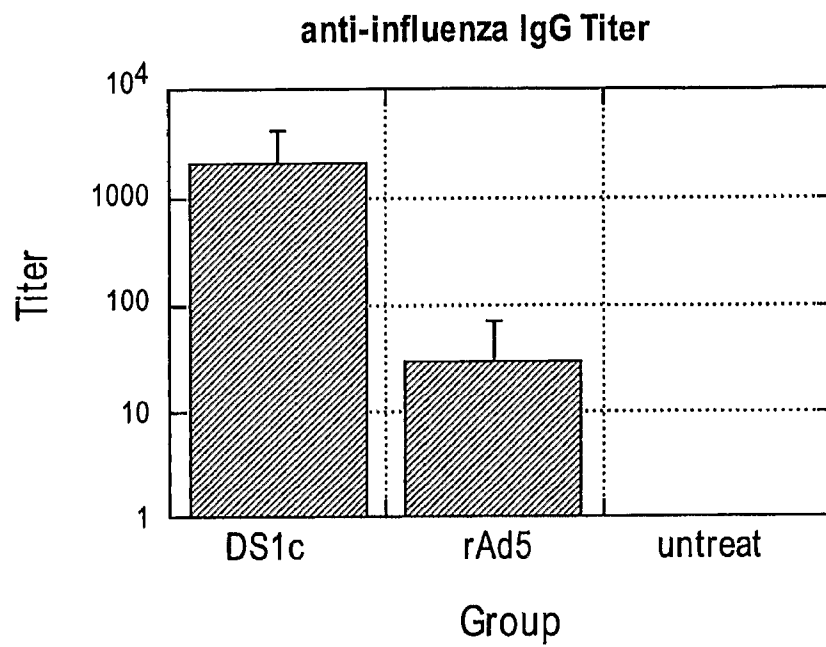
1/10

**FIG. 1A****FIG. 1B**

2/10

**FIG. 2A****FIG. 2B**

3/10

**FIG. 2C**

4/10

anti-gp120 IgG Titer

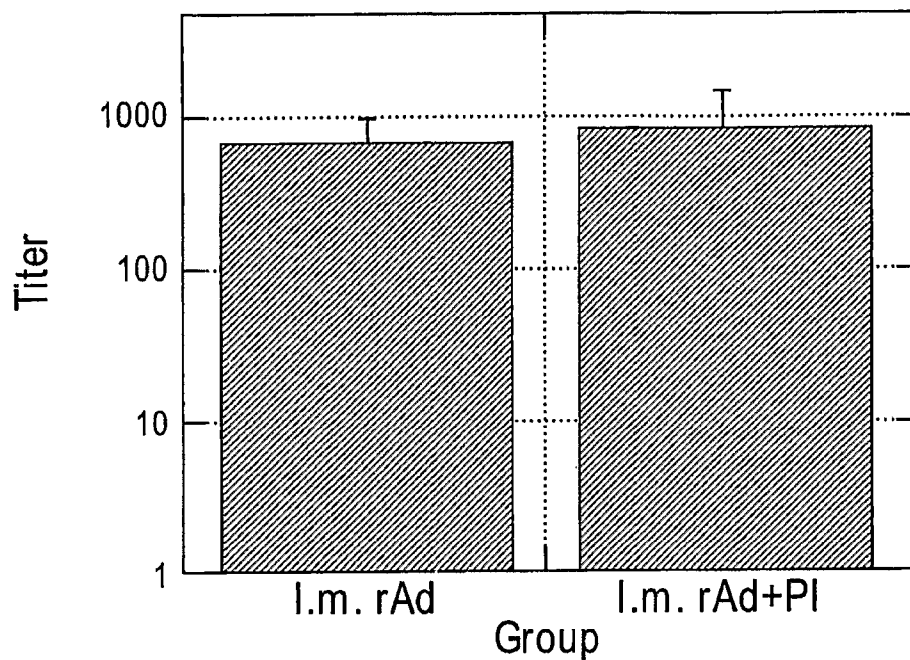


FIG. 3A

Anti-PR8/34 IgG Titer

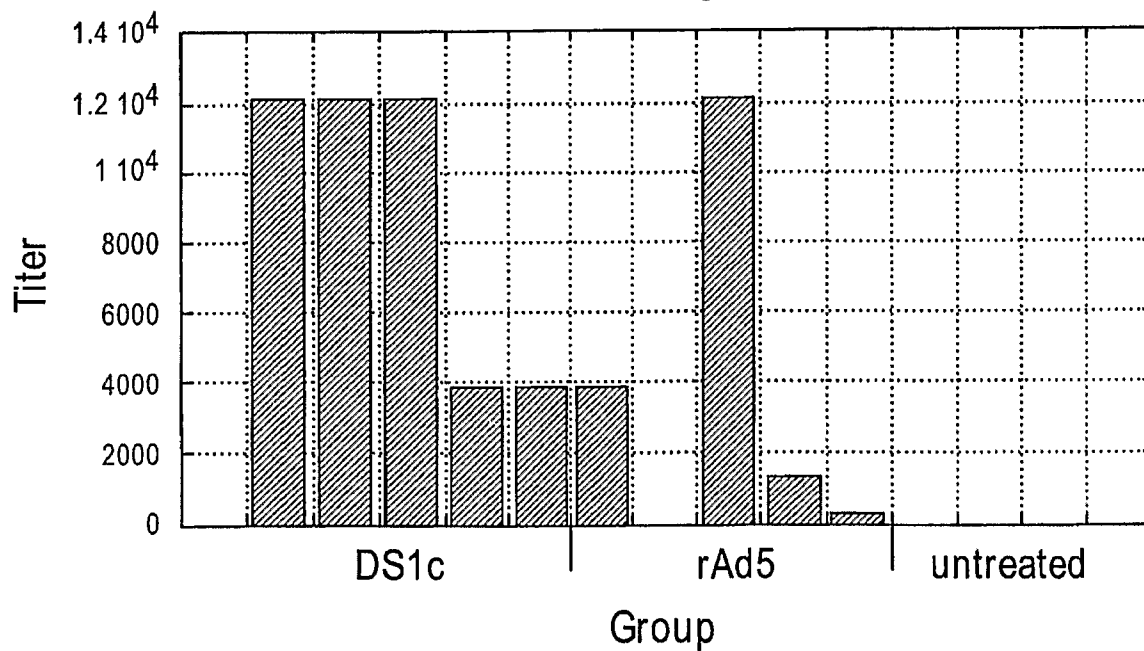


FIG. 3B

5/10

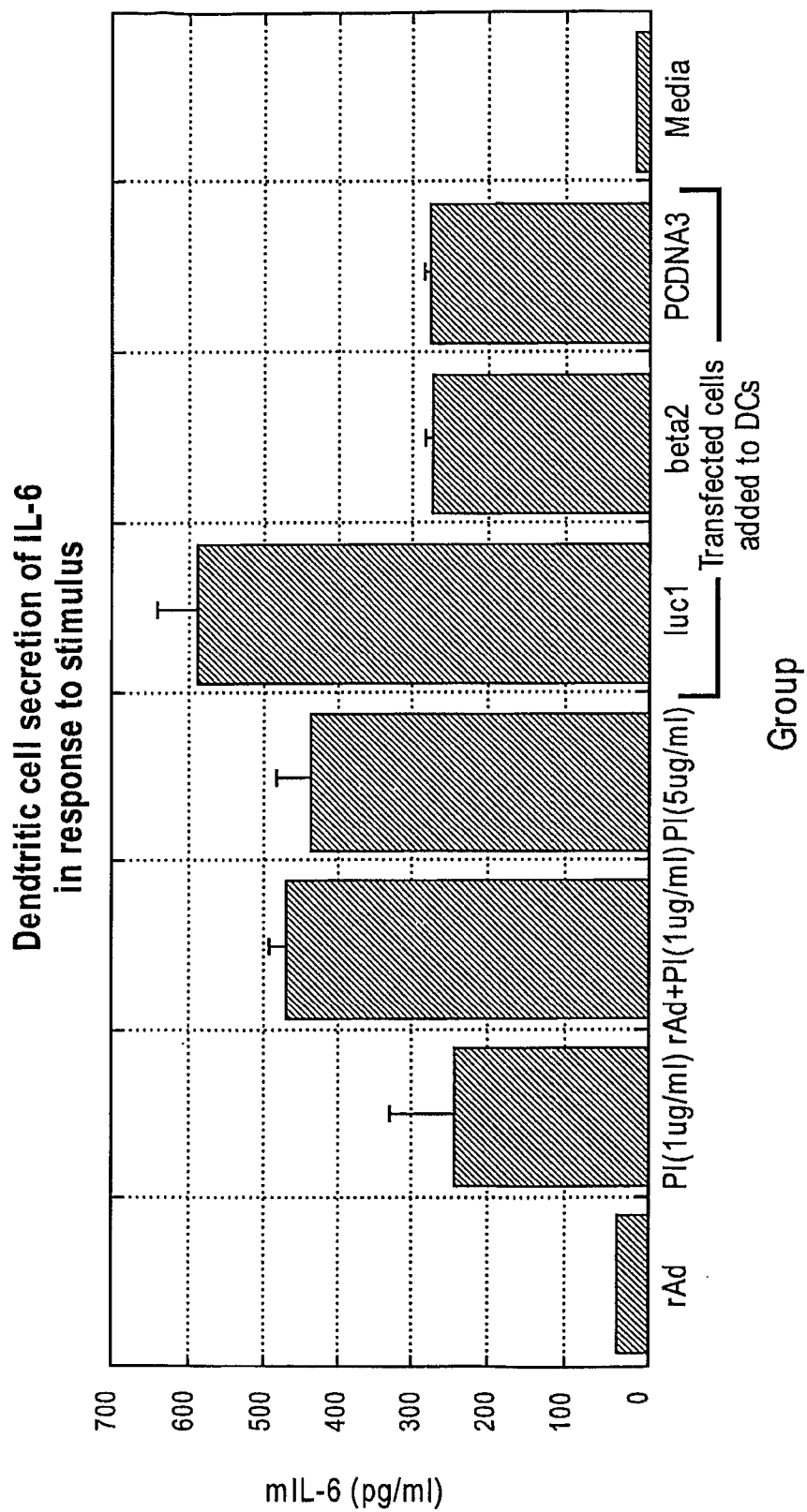
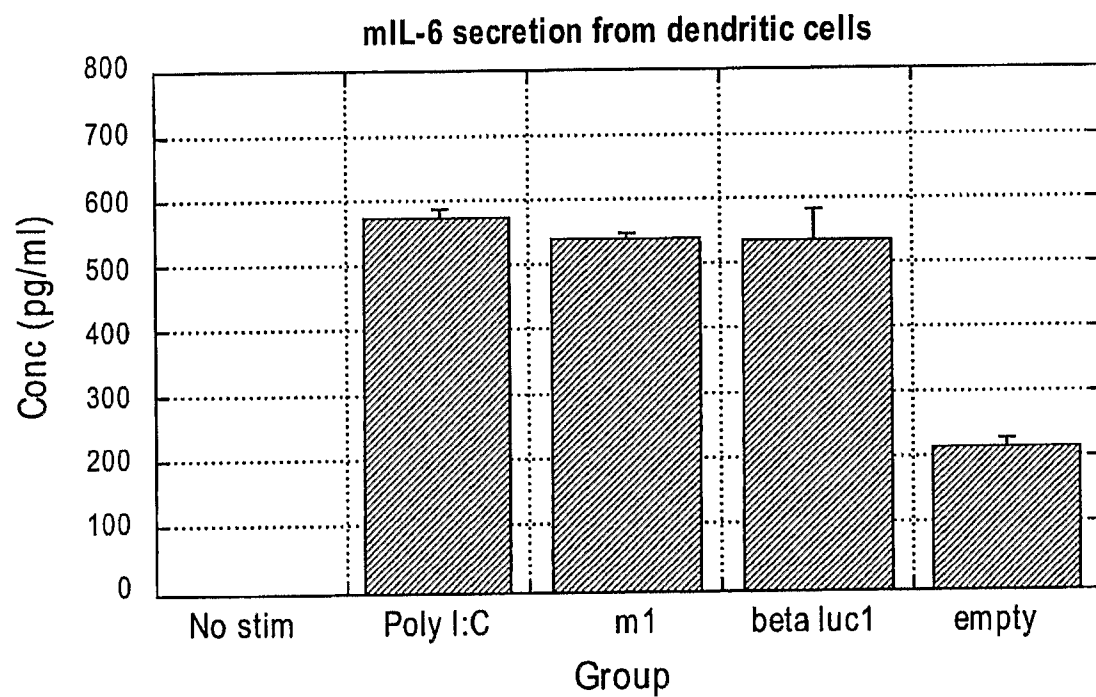


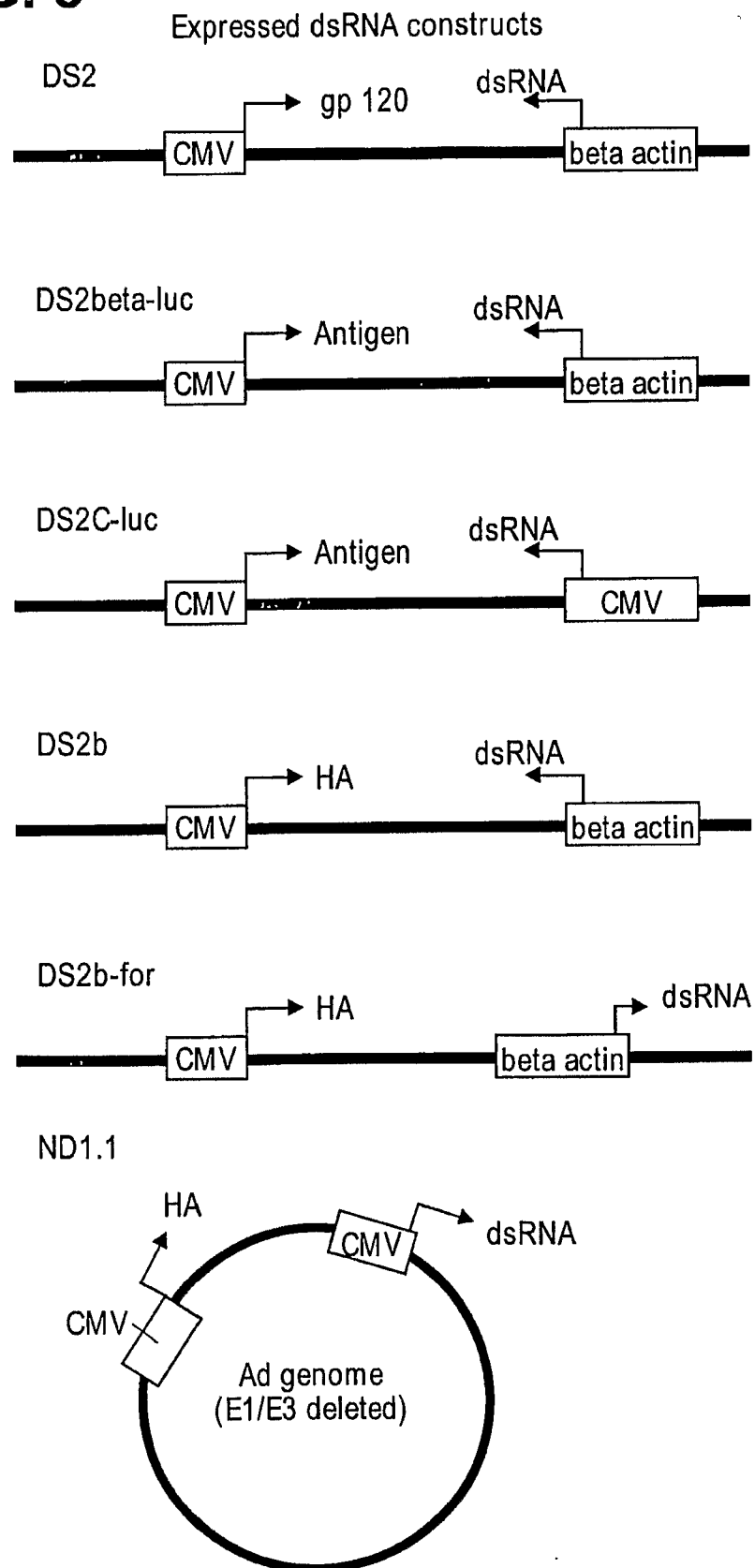
FIG. 4A

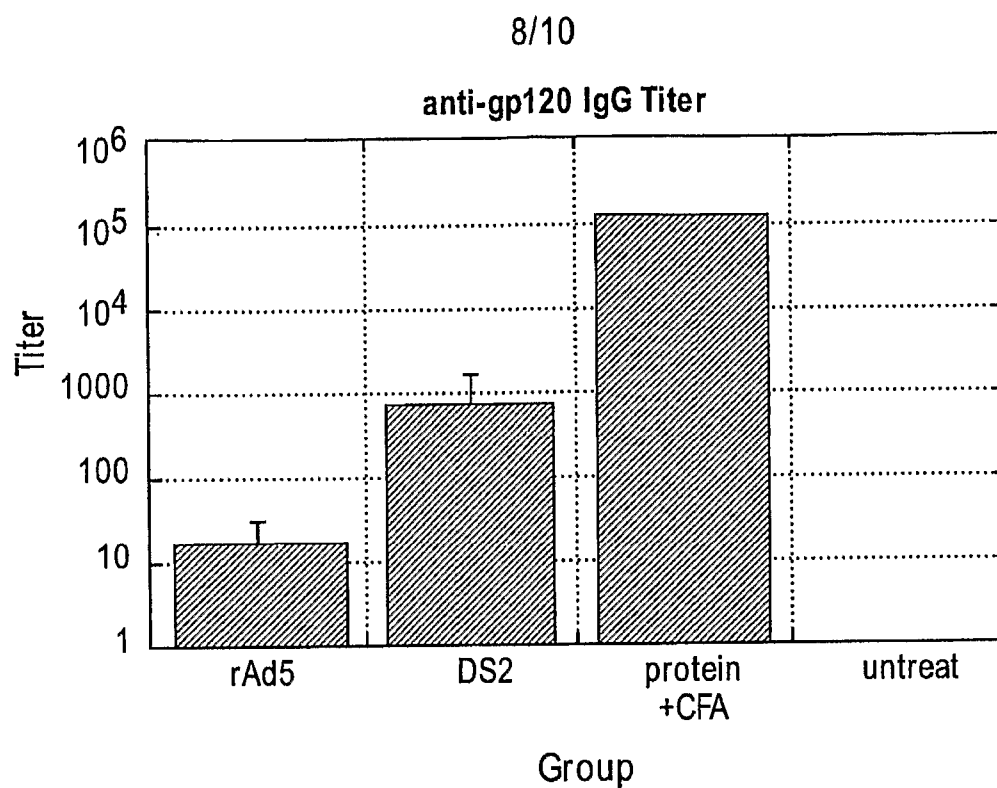
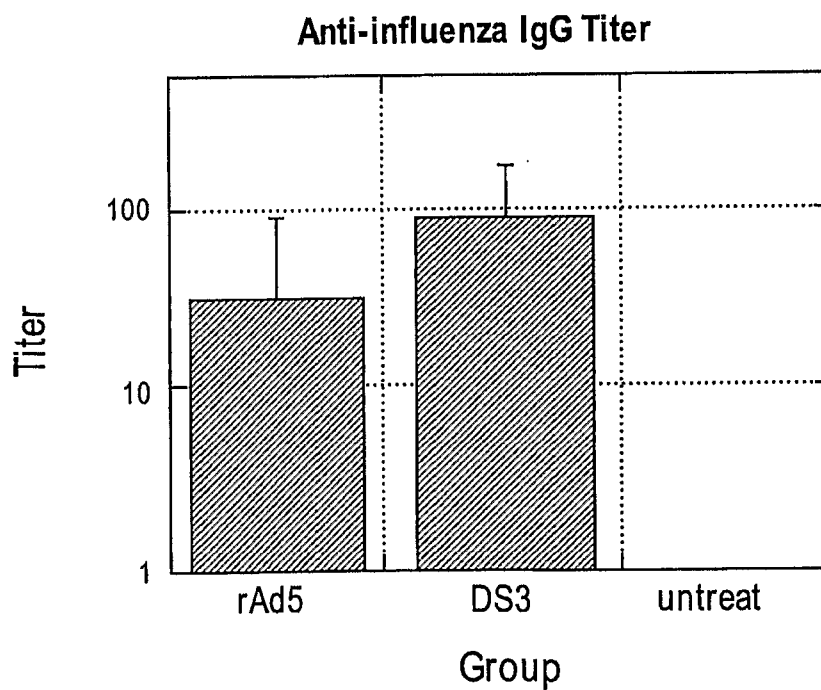
6/10

**FIG. 4B**

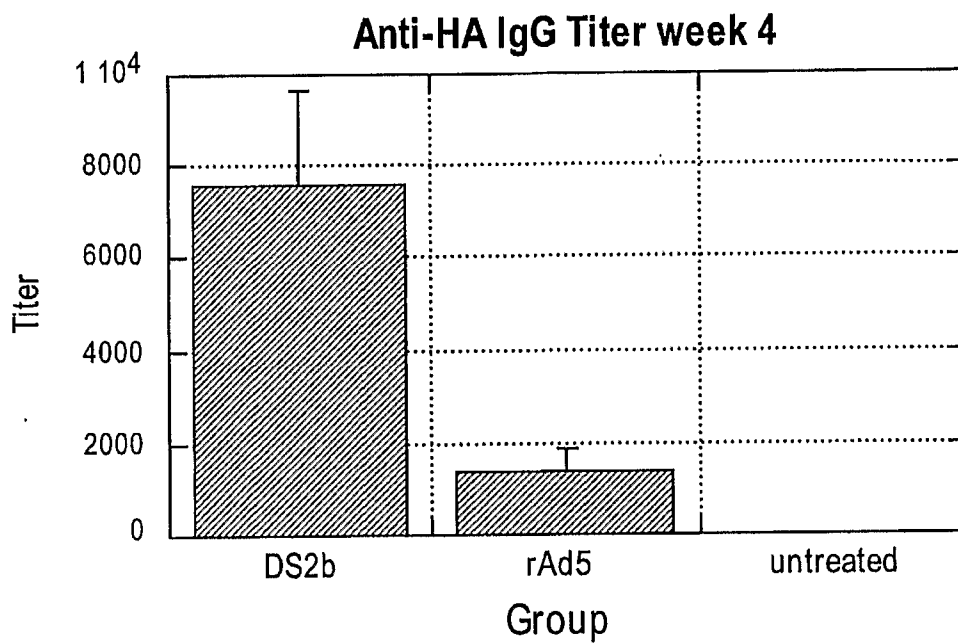
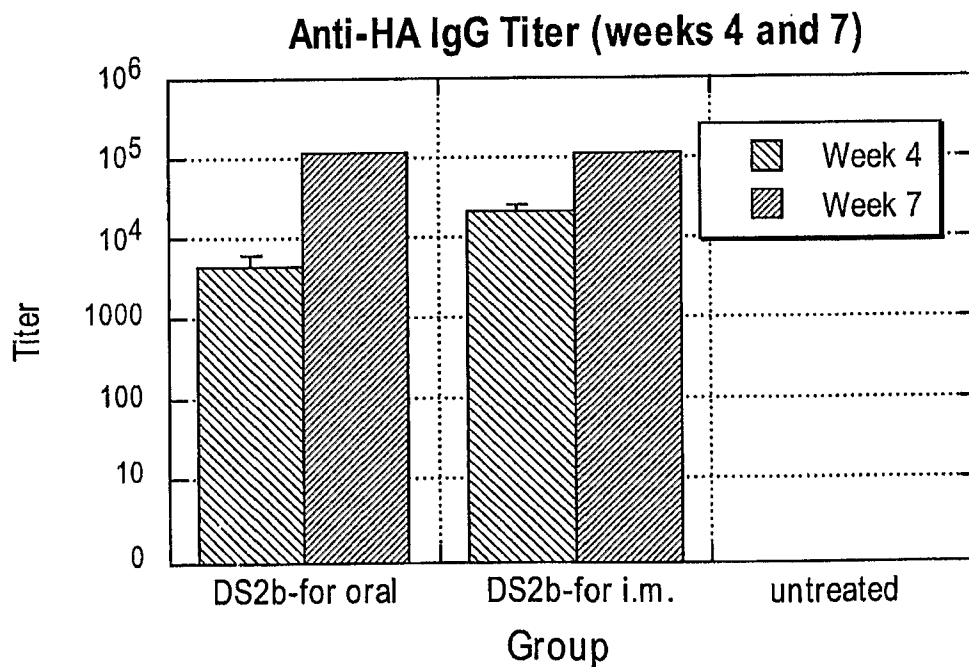
7/10

FIG. 5

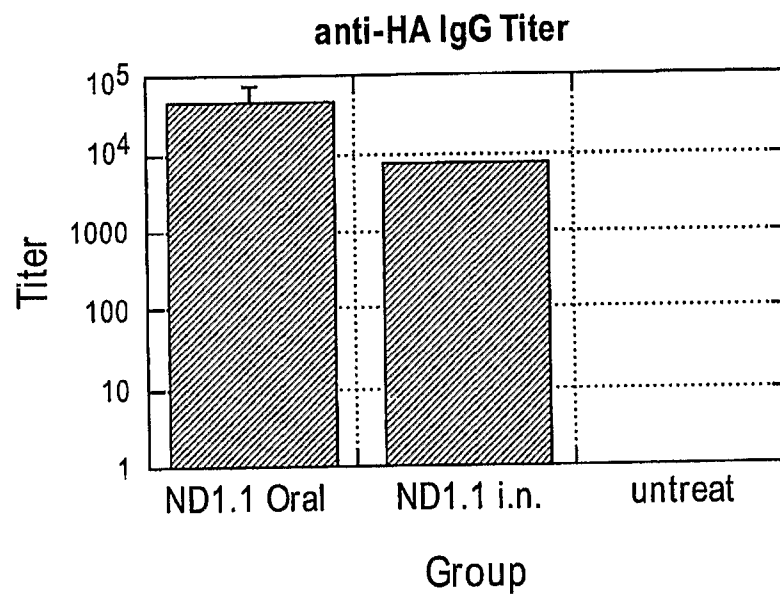


**FIG. 6****FIG. 7**

9/10

**FIG. 8A****FIG. 8B**

10/10

**FIG. 8C**