HU000033576T2

(19) **HU**

(11) Lajstromszám: **E 033 576**

(13) **T2**

MAGYARORSZÁG
Szellemi Tulajdon Nemzeti Hivatala

EURÓPAI SZABADALOM
SZÖVEGÉNEK FORDÍTÁSA

(21) Magyar ügyszám: **E 09 176105**
(22) A bejelentés napja: **2005. 01. 18.**

(96) Az európai bejelentés bejelentési száma:
EP 20050176105
(97) Az európai bejelentés közzétételi adatai:
EP 2163260 A2 **2010. 03. 17.**
(97) Az európai szabadalom megadásának meghirdetési adatai:
EP 2163260 B1 **2017. 03. 15.**

(51) Int. Cl.:
A61K 38/00 *(2006.01)*
A61K 45/06 *(2006.01)*
C12N 15/86 *(2006.01)*
C07K 14/705 *(2006.01)*
C12N 7/00 *(2006.01)*
A61K 39/00 *(2006.01)*
C07K 14/005 *(2006.01)*

(30) Elsőbbségi adatok:
538799 P **2004. 01. 23.** **US**

(72) Feltaláló(k):
Cirillo, Agostino, 00040 Rome (IT)
Colloca, Stefano, 00040 Rome (IT)
Ercole, Bruno Bruni, 00040 Rome (IT)
Meola, Annalisa, 00040 Rome (IT)
Nicosia, Alfreda, 00040 Rome (IT)
Sporeno, Elisabetta, 00040 Rome (IT)

(73) Jogosult(ak):
MSD Italia S.r.l., 00189 Rome (IT)

(74) Képviselő:
Danubia Szabadalmi és Jogi Iroda Kft., Budapest

(54) **Csimpánz adenovírus vakcina hordozók**

Az európai szabadalom ellen, megadásának az Európai Szabadalmi Közlönyben való meghirdetésétől számított kilenc hónapon belül, felszólalást lehet benyújtani az Európai Szabadalmi Hivatalnál. (Európai Szabadalmi Egyezmény 99. cikk(1))

A fordítást a szabadalmas az 1995. évi XXXIII. törvény 84/H. §-a szerint nyújtotta be. A fordítás tartalmi helyességét a Szellemi Tulajdon Nemzeti Hivatala nem vizsgálta.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 163 260 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
15.03.2017 Bulletin 2017/11

(21) Application number: **09176105.6**

(22) Date of filing: **18.01.2005**

(51) Int Cl.:
A61K 38/00 *(2006.01)* ***C07K 14/705*** *(2006.01)*
C07K 14/005 *(2006.01)* ***C12N 15/86*** *(2006.01)*
A61K 39/00 *(2006.01)* ***C12N 7/00*** *(2006.01)*
A61K 45/06 *(2006.01)*

(54) **Chimpanzee adenovirus vaccine carriers**

Impfstoffträger für Schimpansen-Adenovirus

Supports de vaccin adénovirus de chimpanzé

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR
Designated Extension States:
LV

(30) Priority: **23.01.2004 US 538799 P**

(43) Date of publication of application:
17.03.2010 Bulletin 2010/11

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
05701091.0 / 1 711 518

(73) Proprietor: **MSD Italia S.r.l.**
00189 Rome (IT)

(72) Inventors:
- **Cirillo, Agostino**
  00040 Rome (IT)
- **Colloca, Stefano**
  00040 Rome (IT)
- **Ercole, Bruno Bruni**
  00040 Rome (IT)
- **Meola, Annalisa**
  00040 Rome (IT)
- **Nicosia, Alfreda**
  00040 Rome (IT)
- **Sporeno, Elisabetta**
  00040 Rome (IT)

(74) Representative: **Buchan, Gavin MacNicol**
Merck & Co., Inc.
European Patent Department
Hertford Road
Hoddesdon, Hertfordshire EN11 9BU (GB)

(56) References cited:
WO-A2-03/000851

- **ARIGA TOSHIHIDE ET AL: "New genome type of adenovirus serotype 4 caused nosocomial infections associated with epidemic conjunctivitis in Japan." JOURNAL OF CLINICAL MICROBIOLOGY AUG 2004, vol. 42, no. 8, August 2004 (2004-08), pages 3644-3648, XP002565439 ISSN: 0095-1137**
- **FARINA S F ET AL: "Replication-defective vector based on a Chimpanzee adenovirus" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 75, no. 23, 1 December 2001 (2001-12-01), pages 11603-11613, XP002957497 ISSN: 0022-538X**
- **CRAWFORD-MIKSZA L ET AL: "ANALYSIS OF 15 ADENOVIRUS HEXON PROTEINS REVEALS THE LOCATION AND STRUCTURE OF SEVEN HYPERVARIABLE REGIONS CONTAINING SEROTYPE-SPECIFIC RESIDUES" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 70, no. 3, 1 March 1996 (1996-03-01), pages 1836-1844, XP002071016 ISSN: 0022-538X**
- **GALL J G D ET AL: "Construction and characterization of hexon-chimeric adenoviruses: Specification of adenovirus serotype" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 72, no. 12, 1 December 1998 (1998-12-01), pages 10260-10264, XP002327347 ISSN: 0022-538X**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 2 163 260 B1

Description

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 60/538,799, filed January 23, 2004.

FIELD OF THE INVENTION

[0002] The present invention relates to the field of recombinant vectors and more specifically to the production and use of recombinant replication-defective chimpanzee adenoviral vectors to elicit immune responses in mammalian hosts.

BACKGROUND OF THE INVENTION

[0003] The adenoviruses (Ads) comprise a large family of double-stranded DNA viruses found in amphibians, avians, and mammals which have a nonenveloped icosahedral capsid structure (Straus, Adenovirus infections in humans. In The Adenoviruses. 451-498, 1984; Hierholzer et al., J. Infect. Dis., 158: 804-813, 1988; Schnurr and Dondero, Intervirology., 36: 79-83, 1993; Jong et al., J Clin Microbiol., 37:3940-3945:1999). In contrast to retroviruses, adenoviruses can transduce numerous cell types of several mammalian species, including both dividing and nondividing cells, without integrating into the genome of the host cell.

[0004] Generally speaking, adenoviral DNA is typically very stable and remains episomal (e.g., extrachromosomal), unless transformation or tumorigenesis has occurred. In addition, adenoviral vectors can be propagated to high yields in well-defined production systems which are readily amenable to pharmaceutical scale production of clinical grade compositions. These characteristics and their well-characterized molecular genetics make recombinant adenoviral vectors good candidates for use as vaccine carriers. Typically, the production of recombinant adenoviral vectors relies on the use of a packaging cell line which is capable of complementing the functions of adenoviral gene products that have been either deleted or engineered to be nonfunctional.

[0005] Presently, two well-characterized human subgroup C adenovirus serotypes (i.e., hAd2 and hAd5) are widely used as the sources of the viral backbone for most of the adenoviral vectors that are used for gene therapy. Replication-defective human adenoviral vectors have also been tested as vaccine carriers for the delivery of a variety of immunogens derived from a variety of infectious agents (e.g., viruses, parasites, or bacterial pathogens) and tumor cells, including tumor-associated antigens (TAAs). Studies conducted in experimental animals (e.g., rodents, canines and nonhuman primates) indicate that recombinant replication-defective human adenoviral vectors carrying transgenes encoding immunogens derived from the E6 and E7 oncoproteins of human papillomavirus (HPV-16) (He, Z et al., (2001) Virology, 270:3583-3590, the rabies virus glycoprotein (Xiang, Z. et al (1996) Virolgy, 219:220-227), the circumsporozoite protein of *Plasmodium falciparum* Rodriguez, E. et al. (1997) J. Immunol. 158:1268-1274) as well as other heterologous antigens elicit both humoral and cell-mediated immune responses against the transgene product. Generally speaking, investigators have reported success using human adenoviral vectors as vaccine carriers in nonhuman experimental systems by either using an immunization protocols that utilizes high doses of recombinant adenoviral vectors that are predicted to elicit immune responses; or by using immunization protocols which employ the sequential administration of adenoviral vectors that are derived from different serotypes but which carry the same transgene product as boosting immunizations (Mastrangeli, et al., Human Gene Therapy, 7: 79-87 (1996).

[0006] However, it is predicted that vaccine carriers derived from ubiquitous human serotypes, such as types 2 and 5, will encounter preexisting humoral and cellular immunity in the human population. Thus, although replication-defective recombinant human adenoviruses have been successfully employed as vaccine carriers in experimental systems employing rodent, canine, and nonhuman primate hosts; human innate and adaptive immunity is expected to significantly limit the utility of these serotypes as vaccine carriers. This expectation is based on the fact that subgroup C, which includes type 2 and type 5, adenoviral infection is endemic in the human population. As a consequence, the majority of humans seroconvert within the first five years of life as the result of a natural infection. Thus, vectors derived from viruses that naturally infect and replicate in humans may not be optimal candidates for use as vaccine carriers.

[0007] Another problem associated with the use of human adenoviral-derived vectors is the risk that the production method used to propagate the recombinant viruses will give rise to vector stocks that are contaminated with replication competent adenovirus (RCA). This is caused by homologous recombination between overlapping sequences from the recombinant vector and the adenoviral genes that are present in the E1-complementing helper cell lines such as human 293 (Graham, F.L. et al, (1977) J. Gen. Virol. 36:59-72.) cells. The presence of RCA in vector stocks prepared for use in clinical trials constitutes a safety risk because it can promote the mobilization and spread of the replication defective virus. Spread of the defective virus can aggravate the host immune response and cause other adverse immunopathological consequences (Fallux, F. J., et al. Human Gene Therapy 9: 1909-1917 (1998). Accordingly, the Food and Drug Administration (FDA) and other regulatory bodies have promulgated guidelines which establish limits on the levels of

RCA that can be present in vector preparations intended for clinical use. The intent of imposing RCA limits is to ensure limited exposure of patients to replicating adenovirus in compositions that are used in clinical trials.

[0008] Thus, there continues to be a need for the development of adenoviral vaccine carriers that are suitable for use in mammalian hosts which are: easy to manipulate, amenable to pharmaceutical scale production and long term storage, capable of high-level replication in human complementation cell lines, highly immunogenic, devoid of neutralizing B cell epitopes that cross-react with the common serotypes of human adenoviruses, comply with the safety RCA standards promulgated by regulatory agencies, and which are amenable for use in prime/boost protocols that are suitable for use in humans.

SUMMARY OF THE INVENTION

[0009] The scope of the present invention relates to recombinant replication-defective adenovirus vectors derived from chimpanzee according to claim 1, comprising SEQ ID No. 116 and a transgene and methods for generating chimpanzee adenoviral vectors in human E1-expressing cell lines. The application discloses methods for generating clinical grade vector stocks suitable for use in humans and means for using the disclosed vectors as vaccine carriers to elicit protective and/or therapeutic immune responses. The invention further provides methods for using the recombinant chimpanzee adenoviruses of the invention to prepare vaccine compositions designed to delivery, and direct the expression of, transgenes encoding immunogens. In one embodiment, the invention contemplates the use of the disclosed vectors as vaccine carriers for the administration of vaccines comprising transgenes encoding immunogens derived from an infectious agent. The invention discloses the use of the disclosed vectors to prepare and administer cancer vaccines, in particular the preparation and administration of a cancer vaccine comprising a transgene encoding a TAA.

[0010] The invention discloses the complete genomic sequence of five chimpanzee adenoviruses (ChAds), referred to herein as ChAd3 (SEQ ID NO: 1) (Figures 5A-5K), ChAd6 (SEQ ID NO: 2) (Figures 6A-6K, CV32 (SEQ ID NO: 3) (Figures 7A-7K), CV33 (SEQ ID NO: 4) (Figures 8A-8K), and CV23 (SEQ ID NO: 5) (Figures 9A-9J).

[0011] ChAd3 and ChAd6 represent novel adenoviruses isolated according to the methods disclosed herein. The genomes of the ChAd3 and ChAd6 are 37741 and 36648 base pairs in length, respectively. The ChAd3 hexon gene (SEQ ID NO: 41) comprises nucleotides (nt) 19086-21965 of SEQ ID NO: 1 (exclusive of stop codon) and the ChAd3 fiber gene (SEQ ID NO: 42) comprises nt 32805-34487 of SEQ ID NO: 1 (exclusive of stop codon). The ChAd6 hexon gene comprises nt 18266-21124 (SEQ ID NO: 43) of SEQ ID NO: 2 (exclusive of stop codon) and its fiber gene (SEQ ID NO: 44) comprises nt 32218-33552of SEQ ID NO:2 (exclusive of stop codon). Based on sequence homology deduced from a multiple sequence alignment of full-length hexon peptides, ChAd3 has been classified into human subgroup C and ChAd6 has been classified into human subgroup E.

[0012] The genomes of the CV32, CV33 and CV23 adenoviruses are 36,606, 36,535, and 32,020 base pairs in length, respectively. CV32 (Pan 6) (ATCC N. VR-592), CV33 (Pan 7) (ATCC N. VR-593) and CV23 (Pan 5) (Esoterix Inc.,) have all been determined to be related to human Ad4 (hAd4) (subgroup E) (Wigand, R et al. Intervirology 1989, 30:1-9). However, based on hexon sequence alignment CV32
has subsequently characterized as being more closely analogous to human subgroup D members than to hAd4.

[0013] The application discloses nucleotide sequences for the fiber and hexon genes of 21 additional chimpanzee adenoviruses (ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82) isolated according to the methods disclosed herein.

[0014] The fiber gene nucleotide sequences for ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, are set forth in Figures 10-19, respectively, and are referred to herein as SEQ ID NOS: 6- 15: (SEQ ID NO: 6, ChAd20); (SEQ ID NO: 7, ChAd4); (SEQ ID NO: 8, ChAd5); (SEQ ID NO: 9, ChAd7); (SEQ ID NO: 10, ChAd9); (SEQ ID NO: 11, ChAd10); (SEQ ID NO: 12, ChAd11); (SEQ ID NO: 13, ChAd16) (SEQ ID NO: 14, ChAd17) and (SEQ ID NO: 15, ChAd19).

[0015] The fiber gene nucleotide sequences for ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 referred to herein as: (SEQ ID NO: 58, ChAd8), (SEQ ID NO: 60, ChAd22), (SEQ ID NO: 62, ChAd24), (SEQ ID NO: 64, ChAd26), (SEQ ID NO: 66, ChAd30), (SEQ ID NO: 68, ChAd31), (SEQ ID NO: 70, ChAd37), (SEQ ID NO: 72, ChAd38), (SEQ ID NO: 74, ChAd44), (SEQ ID NO: 76, ChAd63) and (SEQ ID NO: 78, ChAd82) and are set forth in the sequence listing.

[0016] The hexon gene nucleotide sequences for ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, are set forth in Figures 21-30, respectively, and are referred to herein as SEQ ID NOS: 16-25: (SEQ ID NO: 16, ChAd20); (SEQ ID NO: 17, ChAd4); (SEQ ID NO: 18, ChAd5); (SEQ ID NO: 19, ChAd7); (SEQ ID NO: 20, ChAd9); (SEQ ID NO: 21, ChAd10); (SEQ ID NO: 22, ChAd11); (SEQ ID NO: 23, ChAd16); (SEQ ID NO: 24, ChAd17) and (SEQ ID NO: 25, ChAd19).

[0017] The hexon gene nucleotide sequences for ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 referred to herein as: (SEQ ID NO: 97, ChAd8), (SEQ ID NO: 99, ChAd22),

(SEQ ID NO: 101, ChAd24), (SEQ ID NO: 103, ChAd26), (SEQ ID NO: 105, ChAd30), (SEQ ID NO: 107, ChAd31), (SEQ ID NO: 109, ChAd37), (SEQ ID NO: 111, ChAd38), (SEQ ID NO: 113, ChAd44), (SEQ ID NO: 115, ChAd63) and (SEQ ID NO: 117, ChAd82) and are set forth in the sequence listing.

[0018] Disclosed are amino acid sequences for the fiber and hexon proteins of 21 additional chimpanzee adenoviruses (ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82) isolated according to the methods disclosed herein.

[0019] The fiber proteins which are disclosed as are referred to as: (SEQ ID NO: 83, ChAd3), (SEQ ID NO: 84, ChAd6), (SEQ ID NO: 48, ChAd20), (SEQ ID NO: 49, ChAd4), (SEQ ID NO: 50, ChAd5), (SEQ ID NO: 51, ChAd7), (SEQ ID NO: 52, ChAd9), (SEQ ID NO: 53, ChAd10), (SEQ ID NO: 54, ChAd11), (SEQ ID NO: 55, ChAd16), (SEQ ID NO: 56, ChAd17), (SEQ ID NO: 57, ChAd19), (SEQ ID NO: 59, ChAd8), (SEQ ID NO: 61, ChAd22), (SEQ ID NO: 63, ChAd24), (SEQ ID NO: 65, ChAd26), (SEQ ID NO: 67, ChAd30), (SEQ ID NO: 69, ChAd31), (SEQ ID NO: 71, ChAd37), (SEQ ID NO: 73, ChAd38), (SEQ ID NO: 75, ChAd44), (SEQ ID NO: 77, ChAd63) and (SEQ ID NO: 79, ChAd82). Figures 20A-20G provides an alignment comparing the amino acid sequences of the fiber proteins disclosed with the amino acid sequences of the fiber proteins of: C1 (SEQ ID NO: 85), CV68 (SEQ ID NO: 86), Pan5 (alternatively referred to as CV23) (SEQ ID NO: 80), Pan6 (alternatively referred to as CV32) (SEQ ID NO: 81), and Pan7 (alternatively referred to as CV33) (SEQ ID NO: 82).

[0020] The hexon proteins which are disclosed as are referred to as: (SEQ ID NO: 122, ChAd3), (SEQ ID NO: 123, ChAd6), (SEQ ID NO: 87, ChAd20), (SEQ ID NO: 88, ChAd4), (SEQ ID NO: 89, ChAd5), (SEQ ID NO: 90, ChAd7), (SEQ ID NO: 91, ChAd9), (SEQ ID NO: 92, ChAd10), (SEQ ID NO: 93, ChAd11), (SEQ ID NO: 94, ChAd16), (SEQ ID NO: 95, ChAd17), (SEQ ID NO: 96, ChAd19), (SEQ ID NO: 98, ChAd8), (SEQ ID NO: 100, ChAd22), (SEQ ID NO: 102, ChAd24), (SEQ ID NO: 104, ChAd26), (SEQ ID NO: 106, ChAd30), (SEQ ID NO: 108, ChAd31), (SEQ ID NO: 110, ChAd37), (SEQ ID NO: 112, ChAd38), (SEQ ID NO: 114, ChAd44), (SEQ ID NO: 116, ChAd63) and (SEQ ID NO: 118, ChAd82). Figures 31A-31J provide a comparison of the amino acid sequences of the hexon proteins disclosed with the amino acid sequences of the hexon proteins of: C1 (SEQ ID NO: 124), CV68 (SEQ ID NO: 125), Pan5 (alternatively referred to as CV23) (SEQ ID NO: 119), Pan6 (alternatively referred to as CV32) (SEQ ID NO: 120), and Pan7 (alternatively referred to as CV33) (SEQ ID NO: 121). A multiple sequence alignment of hexon proteins allows an artisan to perform a phylogenetic analysis of that is consistent with the proposed classification of human adenoviral serotypes (Rux, J.J., et al (2003) J. Virol. 77:9553-9566).

[0021] Disclosed are 21 additional chimpanzee adenovirus isolates. Samples comprising ChAd20, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17 and ChAd19 were deposited on December 12, 2003 with the European Collection of Cell Cultures (ECACC, Porton Down, Salisbury, Wiltshire, SP4 0JG, United Kingdom) as an original deposit under the Budapest Treaty. The deposits were assigned accession numbers: 03121201 (ChAd4), 03121202 (ChAd5), 03121203 (ChAd7), 03121204 (ChAd9), 03121205 (ChAd10), 03121206 (ChAd11), 03121207 (ChAd16), 03121208 (ChAd17), 03121209 (ChAd19) and 03121210 (ChAd20).

[0022] Samples comprising ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 were deposited with the ECACC (Porton Down, Salisbury, Wiltshire, SP4 0JG, United Kingdom) as an original deposit under the Budapest Treaty on January 12, 2005. These deposits were assigned accession numbers: 05011201 (ChAd8), 05011202 (ChAd22), 05011203 (ChAd24), 05011204 (ChAd26), 05011205 (ChAd30), 05011206 (ChAd31), 05011207 (ChAd37), 05011208 (ChAd38), 05011209 (ChAd44), 05011210 (ChAd63) and 05011211 (ChAd82).

[0023] These deposits will be maintained under the terms of the *Budapest Treaty* on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. These deposits were made merely as a convenience for those of skill in the art and are not an admission that a deposit is required under 35 U. S.C. § 112. All restrictions on the availability to the public of the deposited material will be irrevocably removed, except for the requirements specified in 37 C.F.R. § 1.808(b), upon the granting of a patent.

[0024] In an additional aspect, the invention also provides replication-defective recombinant chimpanzee adenoviral vectors which are capable of infecting mammalian cells, preferably human cells, and directing expression of encoded transgene product(s). As demonstrated herein, the disclosed vectors are suitable for use as vaccine carriers for the delivery of transgenes comprising immunogens against which an immune response is desired. In particular embodiments, the invention provides recombinant replication-defective chimpanzee adenoviral vectors that are capable of high-level replication in human E1-expressing (i.e., packaging) cell lines. In one embodiment, the invention provides recombinant chimpanzee adenoviruses that are capable of replicating in PER.C6™ cells.

[0025] Generally speaking, the recombinant vectors encompassed by the invention provide vaccine carriers that will evade pre-existing immunity to the adenovirus serotypes that are typically encountered in the human population. More specifically, the recombinant vectors of the invention comprise vector backbone sequences which are shown herein to be devoid of neutralizing B epitopes that cross-react with the common serotypes of human adenoviral derived vectors.

[0026] Disclosed aregroup-specific shuttle vectors that include an adenoviral portion and a plasmid portion, wherein

said adenoviral portion generally comprises: a) viral left end (ITR and packaging signal), part of the pIX gene and viral genome right end; and b) a gene expression cassette. The group-specific shuttle vectors are designed to exploit the nucleotide sequence homology which is observed between adenoviruses that are assigned to the same serotype subgroup (i.e., subgroups A, B, C, D or E), and can be used to manipulate the nucleotide sequences disclosed herein and/or to clone other chimpanzee adenoviruses belonging to the same subgroup generating an adenovirus pre-plasmid containing a chimp adenoviral genome deleted of E1 region.

[0027] Other aspects of this invention include host cells comprising the chimpanzee adenoviral vaccine vectors, methods of producing the vectors comprising introducing the chimpanzee adenoviral vaccine vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant adenoviral vaccine vectors. Further disclosed is a method of producing a replication-defective chimpanzee adenoviral vector comprising introducing one of the disclosed adenoviral vectors into an adenoviral E-1 expressing human cell, and harvesting the resulting recombinant adenoviruses.

[0028] Disclosed are vaccine compositions which comprise an adenoviral vector of the invention. Compositions comprising recombinant chimpanzee adenoviral vectors may be administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen. These compositions can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. As shown herein, administration of the disclosed vaccine compositions, either alone or in a combined modality, such as a prime boost regimen or multiple injections of serologically distinct Ad vectors results in the induction of an immune response in a mammal that is capable of specifically recognizing the immunogen encoded by the transgene.

[0029] One of the methods disclosed and claimed herein, comprises administering to a mammal (that is either naïve or primed to be immunoreactive to a target antigen), a sufficient amount of a recombinant chimpanzee adenoviral vector, containing at least a functional deletion of its wild-type E1 gene, carrying a sequence comprising a promoter capable of directing expression of a nucleotide sequence encoding the least one target antigen, wherein administration of the recombinant vector elicits (or primes) an antigen-specific immune response.

[0030] In one embodiment, the invention provides a method designed to induce an immune response (prophylactic or therapeutic) against an infectious agent (e.g., a viral or bacterial pathogen or a mammalian parasite). Further disclosed is a method designed to induce an immune response in a mammal that will break tolerance to a self-antigen, such as a TAA. This aspect of the application contemplates the use of the disclosed vectors as a vaccine carrier for the preparation and administration of cancer vaccines.

[0031] Yet other embodiments and advantages of the present invention will be readily apparent from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0032]

Figure 1 is a schematic drawing which summarizes the cloning strategy used to construct a ChAd6 shuttle vector (pARS ChAd6-3).

Figure 2 is a schematic drawing which illustrates the cloning strategy used to clone the ChAd6 viral genome by homologous recombination in *E.coli* strain BJ5183.

Figure 3 is a schematic drawing illustrating the elements of various ChAd6 shuttle plasmids including: pARS ChAd6-3 GAG; pARS ChAd6-3 SEAP; pARS ChAd6-3 EGFP; and pARS ChAd6-3 NS MUT.

Figure 4 is a schematic drawing which illustrates the homologous recombination scheme utilized to clone the ChAd6 ΔE1 expression vectors.

Figures 5A-5K provides the genomic nucleotide sequence of ChAd3 (SEQ ID NO: 1).

Figures 6A-6K provides the genomic nucleotide sequence of ChAd6 (SEQ ID NO: 2).

Figures 7A-7K provides the genomic nucleotides sequence of CV32 (SEQ ID NO: 3).

Figures 8A-8K provides the genomic nucleotide sequence of CV33 (SEQ ID NO: 4).

Figures 9A-9J provides the genomic nucleotide sequence of CV23 (SEQ ID NO: 5).

Figure 10 provides the nucleotide sequence of the fiber gene of ChAd20 (SEQ ID NO: 6).

Figure 11 provides the nucleotide sequence of the fiber gene of ChAd4 (SEQ ID NO: 7).

Figure 12 provides the nucleotide sequence of the fiber gene of ChAd5 (SEQ ID NO: 8).

Figure 13 provides the nucleotide sequence of the fiber gene of ChAd7 (SEQ ID NO: 9).

Figure 14 provides the nucleotide sequence of the fiber gene of ChAd9 (SEQ ID NO: 10).

Figure 15 provides the nucleotide sequence of the fiber gene of ChAd10 (SEQ ID NO: 11).

Figure 16 provides the nucleotide sequence of the fiber gene of ChAd11 (SEQ ID NO: 12).

Figure 17 provides the nucleotide sequence of the fiber gene of ChAd16 (SEQ ID NO: 13).

Figure 18 provides the nucleotide sequence of the fiber gene of ChAd17 (SEQ ID NO: 14).

Figure 19 provides the nucleotide sequence of the fiber gene of ChAd19 (SEQ ID NO: 15).

Figures 20A-20G provides a comparison of the amino acid sequences of the fiber proteins of: ChAd3, ChAd4, ChAd5, ChAd6, ChAd7, ChAd8,ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, ChAd20, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 with the reference fiber protein sequences from C1 (SEQ ID NO: 85), CV68 (SEQ ID NO: 86), PAN5 (also referred to as CV23) (SEQ ID NO: 80), PAN6 (also referred to as CV32) (SEQ ID NO: 81) and Pan7 (also referred to as CV33) (SEQ ID NO: 82).

Figure 21 provides the nucleotide sequence of the hexon gene of ChAd20 (SEQ ID NO: 16).

Figure 22 provides the nucleotide sequence of the hexon gene of ChAd4 (SEQ ID NO: 17).

Figure 23 provides the nucleotide sequence of the hexon gene of ChAd5 (SEQ ID NO: 18).

Figure 24 provides the nucleotide sequence of the hexon gene of ChAd7 (SEQ ID NO: 19).

Figure 25 provides the nucleotide sequence of the hexon gene of ChAd9 (SEQ ID NO: 20).

Figure 26 provides the nucleotide sequence of the hexon gene of ChAd10 (SEQ ID NO: 21).

Figure 27 provides the nucleotide sequence of the hexon gene of ChAd11 (SEQ ID NO: 22).

Figure 28 provides the nucleotide sequence of the hexon gene of ChAd16 (SEQ ID NO: 23).

Figure 29 provides the nucleotide sequence of the hexon gene of ChAd17 (SEQ ID NO: 24).

Figure 30 provides the nucleotide sequence of the hexon gene of ChAd19 (SEQ ID NO: 25).

Figures 31A-31J provides a comparison of the amino acid sequences of the hexon proteins of ChAd3, ChAd4, ChAd5, ChAd6, ChAd7, ChAd8,ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, ChAd20, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 with the reference fiber protein sequences from C1 (SEQ ID NO: 124), CV68 (SEQ ID NO: 125), PAN5 (also referred to as CV23) (SEQ ID NO: 119), PAN6 (also referred to as CV32) (SEQ ID NO: 120) and Pan7 (also referred to as CV33) (SEQ ID NO: 121).

Figure 32 provides a listing of the artificial sequences SEQ ID NOS: 26-40 and SEQ ID NOS: 45 and 46, including oligomers and primers, disclosed herein.

Figure 33 is a graphic representation of the immunization break-point of ChAd vectors belonging to different serotype subgroups (i.e., subgroups C, E and D). The lowest dose eliciting a measurable immune response was determined by performing titration experiments in mice immunized with gag-expressing ChAd3, ChAd11, ChAd20, CV33, CV68, ChAd6, ChAd9, ChAd10, CV32, ChAd4, ChAd7 and ChAd16 vectors.

Figure 34 provides a graphic representation of a CEA-specific T cell response elicited in rhesus macaques immunized sequentially with a human adenoviral vector (MRKAd5 RhCEA) followed by a chimpanzee adenoviral vector (CV33 RhCEA) after 12 week interval. The immune responses were evaluated by IFN-γ ELISPOT assay, and the data illustrate the number of spot-forming cells (SFC) per million peripheral blood mononuclear cells (PBMC) following incubation in the absence (DMSO) and presence of rhesus CEA C and D peptide pools.

Figure 35 provides a phylogenetic tree of human and chimpanzee adenoviruses of deduced from a multiple sequence alignment of full-length hexon peptide sequences using PAUPSEARCH (Wisconsin Package Version 10.3, Accelrys Inc.) and visualized and manipulated with TREEVIEW.

Figure 36 is a graphic representation of immunization results obtained in response to the administration of ChAd3 and hAd5 gag vectors to mice which were pre-exposed to hAd5. Cell-mediated immunity was evaluated 3 weeks post-immunization by IFN-γ ELISPOT using purified splenocytes.

Figure 37 is a graphic representation of kinetics of anti-CEA CMI elicited in human CEA transgenic mice immunized with ChAd3hCEA and Ad5hCEA. CMI was evaluated by ICS of PBMC stimulated with CEA peptide pool. The results are expressed as % of IFNγ^+ CD8$^+$/total PBMC.

Figures 38 A-D is a graphic representation of the efficiency of infection of different human primary cells exposed to moi 50, 250 and 1250 of different ChAd vectors expressing EGFP and belonging to different subgroups (B, C, D, E). The results are expressed as % of fluorescent cells /on total cells.

DETAILED DESCRIPTION OF THE INVENTION

[0033] As used throughout the specification and appended claims, the following definitions and abbreviations apply:

The term "cassette" refers to a nucleic acid molecule which comprises at least one nucleic acid sequence that is to be expressed, along with its transcription and translational control sequences. Changing the cassette, will cause the vector into which is incorporated to direct the expression of different sequence or combination of sequences. In the context of the present invention, the nucleic acid sequences present in the cassette will usually encode an immunogen. Because of the restriction sites engineered to be present at the 5' and 3' ends, the cassette can be easily inserted, removed or replaced with another cassette.

[0034] The term "*cis*-acting element" refers to nucleotide sequences which regulate genes to which they are attached. *Cis*-acting elements present in DNA regulate transcription, and those transcribed into mRNA can regulate RNA processing, turnover and protein synthesis.

[0035] The term "vector" refers to some means by which DNA fragments can be introduced into a host organism or host tissue. There are various types of vectors including plasmid, virus (including adenovirus), bacteriophages and cosmids.

[0036] The term "promoter" refers to a recognition site on a DNA strand to which an RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to initiate and drive transcriptional activity. The complex can be modified by activating sequences such as enhancers, or inhibiting sequences such as silencers.

[0037] The term "pharmaceutically effective amount" refers to an amount of recombinant adenovirus that is effective in a particular route of administration to transduce host cells and provide sufficient levels of transgene expression to elicit an immune response.

[0038] The term "replication-competent" recombinant adenovirus (AdV) refers to an adenovirus with intact or functional essential early genes (i.e., E1A, E1B, E2A, E2B and E4). Wild type adenoviruses are replication competent.

[0039] The term "replication-defective" recombinant AdV refers to an adenovirus that has been rendered to be incapable of replication because it has been engineered to have at least a functional deletion, or a complete removal of, a gene product that is essential for viral replication. The recombinant chimpanzee adenoviral vectors of the invention are replication-defective.

[0040] The term "mammalian" refers to any mammal, including a human being.

[0041] The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome (e.g., about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, e.g., the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g. of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

[0042] Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

[0043] In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example, "pChAd3 ΔE1gag" refers to a plasmid construct which comprises a ChAd3 chimpanzee adenoviral genome deleted of the E1 region. In this plasmid, the E1 region is replaced by an immunogen expression cassette comprising an HIV gag gene under the control of a human CMV promoter followed by a bovine growth hormone polyadenylation signal. Similarly, pCV33DE1-E3 NSmut, refers to a second plasmid construct disclosed herein which comprises a CV33 chimpanzee adenoviral genome, deleted of the E1 and E3 regions, which is replaced by an immunogen expression cassette comprising HCV non-structural genes under the control a human CMV promoter followed by a bovine growth hormone polyadenylation signal.

[0044] The abbreviation "Ag" refers to an antigen.

[0045] As used throughout the specification and in the appended claims, the singular forms "a," "an," and "the" include the plural reference unless the context clearly dictates otherwise.

[0046] Adenoviruses (Ads) are noneveloped, icosahedral viruses that have been identified in several avian and mammalian hosts. Human Ads (hAd) belong to the Mastadenovirus genus which includes all known human and many Ads of animal (e.g., bovine, porcine, canine, murine, equine, simian and ovine) origin. Human adenoviruses are divided into six subgroups (A-F) based on a number of biological, chemical, immunological and structural criteria which include hemagglutination properties of rat and rhesus monkey erythrocytes, DNA homology, restriction enzyme cleavage patterns, percentage G+C content and oncogenicity (Straus, 1984, In *The Adenoviruses,* ed. H. Ginsberg, pps. 451-498, New York: Plenus Press, and Horwitz, 1990 In *Virology,* eds. B.N. Fields and D.M. Knipe, pps. 1679-1721). To date, 51 distinct serotypes have been recognized and grouped into subgroups on the basis of their hemagglutination properties and biophysical and biochemical criteria.

[0047] The adenoviral virion has an icosahedral symmetry and, depending on the serotype, a diameter of 60-90 nm. The icosahedral capsid consists three major proteins, hexon (II), penton base (III) and a knobbed fiber (IV) as well as a number of minor proteins (i.e., VI, VIII, IX, IIIa and IVa2) (W.C. Russel, J. Gen. Virol., 81: 2573-2604 (2000). One aspect of the preexisting immunity that is observed in humans is humoral immunity, which can result in the production and persistence of antibodies that are specific for viral proteins. The humoral response elicited by adenovirus is mainly directed against the major structural proteins: hexon, penton and fiber.

[0048] Published reports have established that titers comprising antibodies against multiple serotypes are common (Dambrosio, E. (1982) J. Hyg. (London) 89: 209-219) and that a substantial portion of the preexisting titers have neutralizing activity. Neutralizing immunity to adenovirus is type specific, and infection with a particular serotype of adenovirus

confers immunity only to that serotype. Several reports have suggested that antibodies directed towards the hexon are the strongest and the most neutralizing (Toogood, C.I.A., Crompton, J. and Hay R.T. (1992) J. Gen. Virol. 73, 1429-1435). Therefore, it is reasonable to assume that the epitopes responsible for type-specific neutralization are located within seven hypervariable regions identified by alignment of the hexon sequences deriving from different serotypes. (Crawford-Miksza, L and D.P.Schnurr. (1996) J. Virol.70:1836-1844).

[0049] A direct correlation between the presence of type-specific neutralizing antibodies and the inability to elicit an immune response with a vector based on the same serotype has been established by different methods including the passive transfer of immune sera from treated to naive animals. Generally speaking, preexisting humoral immunity for a specific viral serotype reduces the therapeutic efficacy of the vector administration. Moreover, the administration of a vector based on a specific viral serotype elicits an immune-response against the vector that prevents the re-administration of the same serotype.

[0050] The application discloses a method of circumventing the adverse effects associated with the consequences of preexisting immunity to common serotypes of hAds. More specifically, the invention contemplates the use of chimpanzee adenoviral vectors characterized by a serotype that does not circulate in humans. Accordingly, the invention discloses adenoviral (Chad) vectors which lack neutralizing B-cell epitopes that cross react with those of common human serotypes as a vaccine carrier.

[0051] Although it has been reported that adenoviral-specific cell mediated immunity (CMI) can be cross-reactive, vaccination studies based on repeated injections of multiple serotypes demonstrated a higher efficiency than immunization schedules based on a single vector. These experiments further demonstrate that the main limitation of a vector administration for vaccine purposes is the humoral pre-existing immunity against the vector. Potential solutions to the problems associated with the use of a human adenovirus as a vaccine carrier include the administration of a higher dose of an adenovirus (e.g., a subgroup C serotype) that is predicted to encounter a preexisting immune response, and the use of vectors based on rare human serotypes. However, the use of higher doses of vaccine increases the cost of the vaccine and risk of undesirable side effects and the results of preclinical testing suggest that human alternate serotypes are less immunogenic than hAd5 and hAd6.

[0052] In an attempt to avoid the problems of host humoral and cellular immune responses against the adenoviral backbone elements of the vector, and to minimize the risk of using human adenovirus-derived vector stocks that may be contaminated with replication-competent adenoviruses (RCA), several nonhuman adenoviruses have been characterized and developed as vaccine carriers (Soudois, C. et al (2000) J. Virology, 74:10639-10649; Farina, S.F. et al (2001) J. Virology, 75:11603-11613; Cohen, C.J. et al (2002) J. Gen. Virology, 83:151-155.) The premise underlying the use of nonhuman adenoviral sequences to circumvent the problems associated with preexisting immunity is based on the observation that neutralizing antibodies to common human adeonvirus serotypes are unlikely to cross-neutralize non-human viruses. However, the incompatibility of viral and cellular factors imposes a practical limitation on the vast majority of alternative vector systems (bovine, ovine, canine) which are characterized by the disadvantage of having to be propagated in non-human cell lines.

[0053] Wilson *et al.* have published a report describing the characterization of a replication-defective vector based on chimpanzee adenovirus type 68 (CV68) C68, which was originally isolated from a mesenteric lymph node of a chimpanzee (Basnight, M., et. al. (1971) Am. J. Epidemiol. 94:166-171.), CV68 was fully sequenced and found to be similar in overall structure to human adenoviruses (Farina, S. F. et al., J. Virol. 75(23): 11603-11613 (2001). The genome of the virus is 36,521 base pairs in length and has been described as being most similar to subgroup E of human adenoviruses, with 90% identity to most human Ad4 open reading frames that have been sequenced. The CV68 ITRs are 130 base pairs in length, and all of the major adenoviral early and late genes are present. CV68 is characterized by a serotype that does not circulate in humans and which lacks neutralizing B cell epitopes that cross-react with those of common human serotypes. Although Chimpanzee adenonviruses are similar to human adenoviruses cross-reactive neutralizing immunity against chimpanzee serotypes has not been documented in humans (Farina, S. F. et al. J. Virol. (2001) 75(23):11603-13).

[0054] The recombinant vectors derived from CV68 are described as being sufficiently similar to human serotypes to support transduction of cells expressing the coxsackievirus and adenovirus receptor (Cohen, C. et al., J. Gen. Virol. 83: 151-155 (2002). Significantly, CV68 is characterized by a sufficient level of similarity to human adenoviruses to support its replication 293 cells which harbor E1 from human adenovirus type 5 (Farina, S. F. et al., J. Virol. 75(23): 11603-11613 (2001). Furthermore, based on the observation that the flanking sequences of the human serotype 5 E1 are nonhomologous with those of the CV68-derived vector sequences, it is predicted that homologous recombination will not occur. Thus, it has been predicted that there is a low likelihood that CV68-derived vaccine stocks will be contaminated with RCA.

[0055] The same group of investigators subsequently reported the use of CV68-derived adenoviral sequences as a vaccine carrier for induction of antibodies to the rabies virus glycoprotein in mice. A replication-defective version of CV68 was created by replacing the E1A and E1B genes with a minigene cassette. Mice immunized with an E1-deletion-containing adenoviral recombinant (AdC68rab.gp) comprising a transgene product encoding the rabies virus glycoprotein developed protective immunity to rabies virus and remained resistant to challenge with an otherwise lethal dose of rabies virus (Xiang, Z et al., J. Virol. 76(5): 2667-2675 (2002). A second CV68 construct expressing a codon-optimized, truncated

form of gag of HIV-1 was recently reported to induce a vigorous gag-specific $CD8^+$ T cell response in mice. The vaccine-induced response was shown to provide protection to challenge with a vaccinia gag recombinant virus (Fitzgerald, J. C. et al., J. Immunol. 170: 1416-1422 (2003). Experimental vaccination of mice preimmunized to human adenovirus serotype 5 with CV68gag or Ad5gag vectors demonstrated a more pronounced reduction of gag-specific T cells and protection against viral challenge elicited by Ad5 than by CV68 vaccine. The reduction in efficacy of C68gag vaccine was attributed to a cross-reactivity of Ad5-specific CD8+ T cells (*Id.*).

[0056] Considered together this data suggests that simian-derived replication-defective adenoviral vectors may be more suitable for use as human vaccine carriers than vectors based on common human serotypes. As shown herein, the results of experiments in which mice that were strongly immunized against human Ad5 (Figure 36) can be immunized with ChAd3-gag adenoviral vectors indicate the preexisiting anti-human Ad5 immunity did not reduce the gag-specific CMI response elicited by the ChAd vectors. These results are consistent with the conclusion that human Ad5 cross-reactive B and T-cell epitopes are not present in ChAd3- or ChAd6 vectors.

[0057] Generally speaking, the adenoviral genome is very well characterized and despite the existence of several distinct serotypes, there is some general conservation in the overall organization of the adenoviral genome with specific functions being similarly positioned. The nucleotide sequences of the chimpanzee adenoviruses C1 and CV68 disclosed by Wilson *et al.,* and the location of the E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of each virus are provided in U.S. Patent No. 6,083,716 (Chimpanzee Adenovirus Vectors), and PCT published application WO 03/000851 (Methods for Rapid Screening of Bacterial Transformants and Novel Simion Adenoviral Proteins).

[0058] Each extremity of the adenoviral genome comprises a sequence known as an inverted terminal repeat (ITRs), which is necessary for viral replication. The virus also comprises a virus-encoded protease, which is necessary for processing some of the structural proteins required to produce infectious virions. The structure of the adenoviral genome is described on the basis of the order in which the viral genes are expressed following host cell transduction. More specifically, the viral genes are referred to as early (E) or late (L) genes according to whether transcription occurs prior to or after onset of DNA replication. In the early phase of transduction, the E1, E2, E3 and E4 genes of adenovirus are expressed to prepare the host cell for viral replication. The virus can be rendered replication defective by deletion of the essential early-region 1(E1) of the viral genome. Brody et al, 1994 Ann N YAcad Sci., 716:90-101. During the late phase, expression of the late genes L1-L5, which encode the structural components of the virus particles is switched on. All of the late genes are under the control of a single promoter and encode proteins including the penton (L2), the hexon (L3), the 100 kDa scaffolding protein (L4), and the fiber protein (L5), which form the new virus particle into which the adenoviral DNA becomes encapsidated. Depending on the serotype of the virus, 10,000-100,000 progeny adenovirus particles can be generated in a single host cell. Ultimately, the adenoviral replication process causes lysis of the cells.

[0059] The replication-defective adenoviral vectors disclosed herein were constructed by deletion of specific nucleotide sequences from the disclosed chimpanzee nucleic acid sequences and insertion of sequences derived other DNA sequences that are useful for transgene insertion, expression or other genetic manipulations. Accordingly, the recombinant chimpanzee adenoviruses described herein may contain adenoviral sequences derived from one or more chimpanzee adenoviruses, or sequences from a chimpanzee adenovirus and from a human adenovirus. Suitable polynucleotide sequences can be produced recombinantly, synthetically or isolated from natural sources. Adenoviral sequences suitable for use in particular aspects of the invention include sequences which lack neutralizing B-cell epitopes that are cross-reactive with common human serotypes.

[0060] At a minimum, the recombinant chimpanzee adenovirus (e.g., vector) of the invention contain the chimpanzee adenovirus *cis*-acting elements necessary for replication and virion encapsidation, in combination with at least one immunogen expression cassette. Typically, the *cis*-acting elements flank the expression cassette which comprises a transgene that encodes at least one antigen. More specifically, the vectors of the invention contain the requisite *cis*-acting 5' inverted terminal repeat (ITR) sequences of the adenoviruses (which function as origins of replication), 3' ITR sequences, packaging/enhancer domains, and a nucleotide sequence encoding a heterologous molecule. Regardless of whether the recombinant vector comprises only the minimal adenoviral sequences or an entire adenoviral genome with only functional deletions in particular genes (e.g., the E1 and/or E3 or E4 regions), the vectors of the invention comprise a chimpanzee adenovirus capsid.

[0061] Generally, speaking the adenoviral vectors disclosed herein comprise a replication-defective adenoviral genome, wherein the adenoviral genome does not have a functional E1 gene, and an immunogen expression cassette which comprises: a) a nucleic acid encoding at least one immunogen against which an immune response is desired; and b) a heterologous (i.e., with respect to the adenoviral sequence) promoter operatively linked to the nucleic acid sequence encoding the immunogen(s); and a transcription terminator.

[0062] More specifically, the invention provides replication-defective chimpanzee adenoviral vectors that consist of a recombinant adenoviral genome that is devoid of at least one early gene selected from the group consisting of E1, E2, E3, and E4. In one embodiment, a replication-defective vector is prepared by replacing, or disrupting, the E1 gene of one of the adenoviral isolates disclosed herein (e.g., ChAd3, ChAd6, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19,ChAd20, ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44,

ChAd63 or ChAd82) with an immunogen expression cassette. For example, a vector can be prepared by deleting/disrupting the E1 gene of ChAd 3 (SEQ ID NO:1) or ChA6 (SEQ ID NOS: 2). Alternatively, a replication-defective vector can be prepared from any one of the other adenovirus isolates disclosed herein, including ChAd3, ChAd6, ChAd4, ChAd5, ChAd7, ChAd9, ChAd10, ChAd11, Chad16, Chad17, ChAd19, ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 or ChAd20. Replication-defective vectors comprising an adenoviral genome derived from one of the ChAds disclosed herein has been optionally engineered to lack a functional E3 gene. It is to be understood that the chimpanzee adenoviral sequences disclosed herein can be rendered replication-defective by either completely removing an early gene or by rendering the gene inoperative or nonfunctional.

[0063] It is to be understood that the invention encompasses vectors that are characterized as having modifications, such as a "functional deletion" which destroys the ability of the adenovirus to express one or more selected gene products. The phrase "functional deletion" as used herein broadly encompasses modifications that have the effect of rendering a particular gene product nonfunctional. Generally speaking, functional deletions take the form of a partial or total deletion of an adenoviral gene. However, one of skill in the art will readily acknowledge that other manipulations, including but not limited to making a modification which introduces a frame shift mutation, will also achieve a functional deletion. For example, the recombinant chimpanzee adenoviral vectors of the invention can be rendered replication-defective by introducing a modification that is designed to interfere with, or to functionally delete, the ability of the virus to express adenoviral E1A and/or E1B.

[0064] It is well-known that replication-defective adenoviral vectors can be obtained by introducing a modification that is designed to interfere with, or to functionally delete the expression of one o more genes from the group of E2 genes. More in detail, a replication-defective vector can be constructed by inactivating the polymerase gene, or the pre-terminal protein gene or the DNA binding protein gene. Moreover deletion or inactivation of genes expressed by E4 region is an alternative strategy to construct replication-defective chimp Ad vectors. Early gene deletion or inactivation can be combined in order to produce more attenuated vectors. Alternatively, replication-defective ChAd vectors can also comprise additional modifications in other viral genes, such as the late genes L1 through L5. In addition, novel adenoviral vaccine carriers can be generated by combining hexon and fiber genes obtained from different serotypes. The utilization of a hexon and fiber gene shuffling strategy will also allow an investigator to change the biological properties of a ChAd and facilitate the production of vectors with a different tropism or with new serological characteristics.

[0065] It is to be understood that the present invention encompasses recombinant chimpanzee adenoviral vectors comprising deletions of entire genes or portions thereof which effectively destroy the biological activity of the modified gene either alone or in any combination. Alternatively, the function of the adenoviral delayed early E3 gene may be eliminated; however because the function of E3 is not necessary for the production of a recombinant adenoviral particle it is not necessary to replace this gene product in order to produce a recombinant that is capable of packaging a virus useful in the invention.

[0066] In one embodiment of this invention, the replication- defective adenoviral vector used is a chimpanzee subgroup C adenovirus containing deletions in E1 and optionally in E3. For example, for ChAd3, a suitable E1 deletion/disruption can be introduced in the region from bp 460 to bp 3542 (with reference to SEQ ID NO: 1). For ChAd6, a suitable E1 deletion/disruption can be introduced in the region from bp 457 to bp 3425 (with reference to SEQ ID NO: 2). For CV32, the E1 deletion is preferably from bp 456 to bp 3416 (with reference to SEQ ID NO: 3); for CV33, the E1 deletion is preferably from bp 456 to bp 3425 (with reference to SEQ ID NO: 4) and for CV23, the E1 deletion is preferably from bp 456 to bp 3415 (with reference to SEQ ID NO: 5). E3 deletions for CV32 and CV33 are preferably from bp 27446 to bp 31911 (with reference to SEQ ID NO: 3); from bp 27146 to bp 31609 (with reference to SEQ ID NO: 4) respectively. Those of skill in the art can easily determine the equivalent sequences for other chimpanzee isolates based on sequence homologies and multiple sequence alignments.

[0067] One of skill in the art will readily acknowledge that in order to construct an E1-deleted adenoviral vector a number of decisions must be made regarding the structure of the vector backbone and the composition of the nucleic acid sequence comprising the transgene. For example, an investigator must determine if the size of the E1 deletion will accommodate the size of the transgene. If not, then additional deletions will have to be introduced into the backbone of the vector.

[0068] The nucleic acid sequence embodying the transgene can be a gene, or a functional part of a gene and will typically exist in the form of an expression cassette. Typically a gene expression cassette includes: (a) nucleic acid encoding a protein or antigen of interest; (b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and (c) a transcription termination signal. The nucleic acid can be DNA and/or RNA, can be double or single stranded. The nucleic acid can be codon-optimized for expression in the desired host (*e.g.,* a mammalian host).

[0069] Decisions must also be made regarding the site within the backbone where the transgene will be introduced and the orientation of the transgene. More specifically, the transgene can be inserted in an E1 parallel (transcribed 5' to 3') or anti-parallel (transcribed in a 3' to 5' direction relative to the vector backbone) orientation. In addition, appropriate transcriptional regulatory elements that are capable of directing expression of the transgene in the mammalian host cells that the vector is being prepared for use as a vaccine carrier in need to be identified and operatively linked to the

transgene. "Operatively linked" sequences include both expression control sequences that are contiguous with the nucleic acid sequences that they regulate and regulatory sequences that act in *trans,* or at a distance to control the regulated nucleic acid sequence.

[0070] Regulatory sequences include: appropriate expression control sequences, such as transcription initiation, termination, enhancer and promoter sequences; efficient RNA processing signals, such as splicing and polyadenylation signals; sequences that enhance translation efficiency (e.g., Kozak consensus sequences); sequences that enhance protein stability, and optionally sequences that promote protein secretion. Selection of these and other common vector elements are conventional and many suitable sequences are well known to those of skill in the art (see, e.g., Sambrook *et al,* and references cited therein at, for example, pages 3.18-3.26 and 16.17-16.27 and Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989).

[0071] In specific embodiments, the promoter is a heterologous promoter (i.e., with respect to the adenovirus sequences) which is recognized by an eukaryotic RNA polymerase. In a preferred embodiment, the promoter is a "strong" or "efficient" promoter. An example of a strong promoter is the immediate early human cytomegalovirus promoter (Chapman et al, 1991 Nucl. Acids Res 19:3979-3986, which is incorporated by reference). The human CMV promoter can be used without (CMV) or with the intron A sequence (CMV-intA), although those skilled in the art will recognize that any of a number of other known promoters, such as the strong immunoglobulin, or other eukaryotic gene promoters may be used, including the EF1 alpha promoter, the murine CMV promoter, Rous sarcoma virus (RSV) promoter, SV40 early/late promoters and the beta-actin promoter.

[0072] Further examples of promoters that can be used in the present invention are the strong immunoglobulin promoter, the EF1 alpha promoter, the murine CMV promoter, the Rous Sarcoma Virus promoter, the SV40 early/late promoters and the beta actin promoter, albeit those of skill in the art can appreciate that any promoter capable of effecting expression in the intended host can be used in accordance with the methods of the present invention. The promoter may comprise a regulatable sequence such as the Tet operator sequence. Sequences such as these that offer the potential for regulation of transcription and expression are useful in instances where repression of gene transcription is sought.

[0073] Suitable gene expression cassettes will also comprise a transcription termination sequence. A preferred transcriptional terminator is the bovine growth hormone terminator. The promoter/transcription termination combination of CMVintA-BGH terminator is particularly preferred although other promoter/terminator combinations may also be used. As shown herein, the bovine growth hormone termination/polyadenylation signal (bGHpA) or short synthetic polyA signal (SPA) of 50 nucleotides in length defined as follows:

AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTG (SEQ ID NO: 26).

Generally speaking, exemplify suitable termination sequences. The polyA signal is inserted following the nucleic acid sequence which comprises the transgene and before the 3' adenovirus ITR sequence.

[0074] The recombinant chimpanzee adenoviral vectors described herein may contain adenoviral sequences derived from one or more strain of adenovirus. Suitable sequences may be obtained from natural sources, produced recombinantly, synthetically, or by other genetic engineering or chemical methods. In a particular embodiment, the recombinant chimpanzee adenovirus is a chimeric recombinant comprising non-chimpanzee adenoviral polynucleotide sequences. Suitable non-chimpanzee adenoviral sequences can be obtained from human adenoviral strains. For example, the native E4 region can be replaced by hAd5 E4 (Ad5 nt 32816 to nt 35619) or by Ad5E4orf6 (Ad5 nt 33193 to nt 34077) (Ad5 GenBank Accession No: M73260).

[0075] Generally speaking, the immunogen (antigenic molecule) delivered by the recombinant adenoviral vector of the invention comprises a polypeptide, protein, or enzyme product which is encoded by a transgene in combination with a nucleotide sequence which provides the necessary regulatory sequences to direct transcription and/or translation of the encoded product in a host cell. The composition of the transgene depends upon the intended use of the vector. For example, if the immunogenic composition is being designed to elicit an antibody response or a cell-mediated immune response in a mammalian host which is specific for an infectious agent, then it is appropriate to utilize a nucleic acid sequence encoding at least one immunogenic product that is predicted to confer pathogen-specific immunity to the recipient. Alternatively, if the composition is being prepared for use as a cancer vaccine, a suitable transgene may comprise an immunogenic portion of a self-antigen, such as a TAA, which has been selected with the goal of eliciting a protective immune response of sufficient potency to both break host tolerance to a particular TAA and to elicit a long-lived (e.g., memory) response that will be sufficient to prevent the initiation of cancer or to prevent tumor progression. Accordingly, suitable immunogenic gene products may be obtained from a wide variety of pathogenic agents (such as, but not limited to viruses, parasites, bacteria and fungi) that infect mammalian hosts, or from a cancer or tumor cell. Although, the invention is illustrated herein with a particular set of test immunogens it is to be understood that the invention is not limited to the use of the antigens exemplified herein. More specifically, the invention contemplates the use of both heterologous and self-antigens as immunogens, including but not limited to TAAs.

[0076] In one embodiment, the invention provides an immunogenic composition (e.g., a vaccine) for inducing an

immune response against antigens (i.e., immunogens) expressed by an infectious agent. For example, it is desirable to elicit an immune response against a virus infecting humans and/or non-human animal species. Examples of virus families against which a prophylactic and/or therapeutic immune response would be desirable include the *Picornaviridae* family which includes six different genera such as Aphtovirus, Cardiovirus, Enterovirus, Hepatovirus, Parechovirus, Rhinovirus. Examples of Picornavirus against which an immune response would be desirable are: Foot-and-mouth disease viruses, Encephalomyocarditis viruses, Polioviruses, Coxackieviruses, Human hepatitis A virus, Human parechoviruses, Rhinoviruses. *Caliciviridae* family includes different genera associated with epidemic gastroenteritis in humans caused by the Norwalk group of viruses and other syndromes in animals like the hemorrhagic disease in rabbits associated with rabbit hemorrhagic disease virus or respiratory disease in cats caused by feline calicivirus.

[0077] Another family of viruses, against which it may be desirable to elicit an immune response is the *Astroviridae* which comprises viruses isolated from humans as well as many different animal species. Human astroviruses are associated with gastroenteritis and young children diarrhea. Alternatively, it may be desirable to confer mammalian hosts with immunity to members of the *Togaviridae* family of viruses which comprises two genera: alphavirus and rubivirus. Alphaviruses are associated with human and veterinary diseases such as arthritis (i.e. Chikungunya virus, Sindbis virus) or encephalitis (i.e. Eastern Equine Encephalitis Virus, Western Equine Encephalitis Virus).

[0078] Rubella virus provides an alternative viral target against which is the only member of the Rubivirus genus is responsible for outbreaks of a mild exanthematic disease associated with fever and lymphoadenopathy. Rubella virus infection is also associated with fetus abnormalities when acquired by mother during in early pregnancy. *Flaviviridae* is another virus family consisting of three genera: the flaviviruses, the pestiviruses and the hepaciviruses that includes important human as well as animal pathogens. Many of the flavivirus genus members are arthropod-borne human pathogens causing a variety of diseases including fever, encephalitis and hemorrhagic fevers. Dengue Fever Viruses, Yellow Fever Virus, Japanese Encephalitis Virus, West Nile Fever Virus, Tick-borne Encephalitis Virus are pathogens of major global concern or of regional (endemic) concern. Pestivirus genus includes animal pathogens of major economic importance such as Bovine Viral Diarrhea Virus, Classical Swine Fever Virus, Border Disease Virus. Hepatitis C Virus is the only member of the Hepacivirus genus responsible for acute and chronic hepatitis. HCV proteins expressed by a recombinant adenovirus can elicit a protective as well as therapeutic immune response limiting the consequences of a viral infection affecting 170 million people worldwide.

[0079] Alternatively, antigens derived from members of the *Coronaviridae* family can be expressed by recombinant adenovirus vectors in order to obtain protection against infection. Protection against the severe acute respiratory syndrome coronavirus (SARS-Co Virus) can be obtained by immunizing with one or more chimpanzee adenovirus chosen from the group including ChAd3, 4, 5,6, 7,9,10,11,16,17,19, 20, ChAd8, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82 expressing one or more SARS-CoV protein including without limitations nucleocapsid (N) protein, polymerase (P) protein, membrane (M) glycoprotein, spike (S) glycoprotein, small envelope (E) protein or any other polypeptide expressed by the virus. *Rhabdoviridae* family members including rabies virus can be target of recombinant vaccine expressing viral proteins.

[0080] Other possible targets include the *Filoviridae* family comprising Ebola-like viruses and Marburg-like viruses genera, that is responsible of outbreaks of severe hemorrhagic fever; the *Paramyxoviridae* family comprising some of the most prevalent virus known in humans like measles, respiratory syncytial, parainfluenza viruses and viruses of veterinary interest like Newcastle disease and rinderpest viruses; the *Orthomyxoviridae* family including Influenza A,B,C viruses; *Bunyaviridae* family mainly transmitted by arthropod to vertebrate hosts comprising important human pathogens like Rift valley fever, Sin Nombre, Hantaan, Puumala viruses; *Arenaviridae* family comprising Lymphocytic choriomeningitis, Lassa fever, Argentine Hemorragic fever, Bolivian Hemorragic fever viruses; *Bornaviridae* family comprising viruses causing central nervous system diseases mainly in horses and sheep; *Reoviridae* family including rotaviruses, the most important cause of severe diarrheal illness in infants and young children worldwide, orbiviruses that can affect both humans and other mammals (bluetongue, epizootic hemorrhagic disease viruses); Retroviridae family, a large group of viruses comprising important human pathogens like human immunodeficiency virus 1 and 2 (HIV-1 and HIV-2) and human t-cell leukemia virus type 1 and 2 (HTLV 1 and 2) as well as non-human lentivirus such as Maedi/Visna viruses affecting sheep and goats, Equine infectious anemia virus affecting horses, bovine immunodeficiency virus affecting cattle, feline immunodeficiency virus affecting cats; Polyomaviridae family groups small DNA oncogenic viruses, prototype viruses are polyoma and SV40 infecting mouse and rhesus monkey respectively, (BK and JC viruses closely related to SV40 were isolated from human patients); *Papillomaviridae* family consists of a group of DNA viruses infecting higher vertebrates including humans generating warts and condylomas. Papilloma viral infection is associated with the development of cancer in both humans and animals. Human papilloma viruses are associated with cervical cancer, vaginal cancer and skin cancer. The herpesviridae families includes subfamilies in which are classified a number of important pathogens for humans and other mammals. Suitable sources of antigens can be but are not limited to herpes simplex viruses 1 and 2, varicella-zoster virus, Epstein-Barr virus, Cytomegalovirus, human herpesviruses 6A,6B and 7, Kaposi's sarcoma-associated herpesvirus. Further suitable source of antigens are members of the Poxviridae family like Monkeypox virus, Molluscum contagiusum virus, smallpox virus; Hepatitis B virus, the prototype member of the

hepadnaviridae family as well as other virus causing acute and/or chronic hepatitis like hepatitis delta virus, hepatitis E virus.

[0081] The adenoviral vectors of the present invention are also suitable for the preparation of immunogenic compositions designed to stimulate an immune response in humans or animals against protein expressed by non-viral pathogens including bacteria, fungi, parasites pathogens. For example, the vectors disclosed herein can be used to prepare vaccines against, but not limited to: *Staphylococcus aureus, Streptococcus pyogenes, Streptococcus pneumoniae, Vibrio cholerae, Clostridium tetani, Neisseria meningitis, Corynebacterium diphteriae, Mycobacteria tuberculosis and leprae, Listeria* it *monocytogenes,* and *Legionella pneumofila.* Examples of fungi and mammalian parasites for which it may be desirable to prepare prophylactic or therapeutic vaccines include: *Candida albicans, Aspergillus fumigatus, Histoplasma capsulatum, Plasmodium malariae, Leishmania major, Trypanosome cruzi and brucei, Schistosoma haematobium, mansoni and japonicum; Entamoeba histolytica,* and numerous species of Filaria known to be responsible for human filariasis.

[0082] Cancer typically involves the deregulation of genes that encode polypeptides which contribute to maintaining cell cycle or controlling cell proliferation (e.g., growth factors, oncogenes, receptors and tumor suppressors). The products of many of the genes implicated in cancer are expressed on the surface of a wide variety of tumor cells. A variety of tumor antigens that may be recognized by T and B lymphocytes have been identified in human and animal cancer. The vast majority of human tumor-associated antigens (TAAs) that are suitable for use in an anticancer vaccine trial are described as "self-antigens" due to the fact that in addition to being expressed on tumor cells they also are expressed on normal tissue and/or during fetal development. Immunotolerance of the target population to TAAs may explain why many cancer vaccines have proven to be ineffective.

[0083] Tumor antigens can be produced by oncogenic mutants of normal cellular genes altered proto-oncogenes or tumor suppressor genes such as Ras, p53 or Bcr-Abl protein are examples of altered cellular proteins that can stimulate T/B cell response. Tumor antigens can be normal cellular proteins that are overexpressed in tumor cells (tyrosinase, GP100, MART are normally expressed at low levels in melanocytes and overexpressed in melanoma) or aberrantly expressed in tumor cells (MAGE, BAGE, GAGE expressed in melanomas and many carcinomas but normally expressed in the testis and placenta). Tumor antigens can be products of oncogenic viruses: papillomavirus E6 and E7 proteins expressed by cervical carcinomas; EBV EBNA-1 protein produced by EBV+ lymphomas and nasopharyngeal carcinomas; SV40 T antigen in SV40 induced experimental tumors. Oncofetal antigens are expressed to high levels on cancer cells and in normal developing (fetal) tissues but not in adult tissues. Carcinoembryonic antigen (CEA) and alpha-fetoprotein (AFP) are examples of well characterized oncofetal antigens.

[0084] Recent evidence supports the existence of TAAs that are capable of eliciting an immune response, thus making this class of antigens suitable immunogens for vaccine therapy. However, as a class of antigens TAAs are notoriously poor immunogens and T cells that are highly specific for TAAs are either deleted or anergized during T-cell development. Accordingly, there is an expectation that the immune response of a tumor-bearing host to a particular TAA will be extremely weak. Because of the inherent need to break host tolerance to a target TAA experimental clinical vaccine studies are particularly focused on developing immunization strategies that will enhance TAA-specific T-cell responses. Generally, speaking an effective cancer vaccine must both overcome immunotolerance and enhance host's immune response to a level that is preventative and/or protective. Anti-tumor effects in many experimental vaccine studies have been correlated with T-cell responses to TAAs.

[0085] Disclosed within the application is an immunogenic composition (e.g., a cancer vaccine) which can be used to induce an immune response against tumor antigens. A suitable composition would contain a recombinant chimpanzee adenovirus comprising nucleic acid sequence encoding a tumor antigen and a physiologically acceptable carrier. The coding sequence element of the cassette may encode a single immunogen, such as an immunogenic peptide sequence derived from a self-antigen, such as a tumor-associated antigen. The nucleic acid sequence encoding the immunogen (i.e., the transgene) may be codon optimized for expression in a particular mammalian species. The coding sequence may encode more than one immunogen, such as one or more codon optimized tumor antigens. For example, a cancer vaccine utilizing the disclosed adenoviral vectors may encode a combination of self-antigens such as: HER2/neu, CEA, Hepcam, PSA, PSMA, Telomerase, gp100, Melan-A/MART-1, Muc-1, NY-ESO-1, Survivin, Stromelysin 3, Tyrosinase, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, hKLP2, VEGF.

[0086] Development of an effective cancer vaccine requires the identification of a strategy that will elicit antigen-specific immunity in vaccinated patients and the generation of an immune response that will persist after active immunization has ended. The success of the strategy will depend on whether a measurable immune response directed against a target antigen will correlate with protection against cancer occurrence or relapse. The effector mechanisms of both cell-mediated immunity and humoral immunity have been show to kill tumor cells. However, data from experimental systems suggest that antigen-specific T cells represent the most powerful immunologic mechanism for the elimination of tumor cells. Recognition of tumor-specific antigens (e.g., TAAs) by effector T-cells is predicted to allow the TAA to function as a tumor-rejection antigen. Published studies suggest that stimulation of $CD8^+$ and $CD4^+$ helper T-cell responses are important for achieving optimal tumor clearance ((Greenberg, P. D. (1991) Adv. Immunol. 49: 281-355; Pardoll, D. M. et al. (1998) Curr. Opin. Immunol. 10: 588-94). Clinical response (i.e., efficacy) has been associated with increases in

interferon γ-secreting cytotoxic T cells. The advent of assays, such as the ELISPOT assay used herein, to demonstrate the efficacy of the instant vaccine carriers, allows investigators to measure T-cell responses to vaccination regimens and thereby facilitates the development of cancer vaccines.

[0087] Cancer vaccines can be either prophylactic or therapeutic. The general assumption underlying the prophylactic use of cancer vaccines is that TAAs are extremely weak immunogens or functionally nonimmunogenic in tumor-bearing subjects. More specifically, in the field of cancer immunology, vaccines can be used as immunotherapy in patients afflicted with cancer. Accordingly, cancer vaccines can be designed to elicit an immune response that is that is directed against a TAA that is expressed by a pre-existing tumor or malignancy. Thus, therapeutic cancer vaccines are intended for use in tumor-bearing patients who have developed resistance to conventional regimens of treatment or who have a high probability of developing a recurrence following conventional treatment.

[0088] The high immunogenicity of adenoviruses, make adenoviral vectors particularly good candidates for use in the context of a vaccine carrier designed to break host tolerance to a self-antigen. The phenomenon of epitope or determinant spreading, which was first described in autoimmune diseases, has been associated with both MHC class I- and MHC class II-restricted responses and correlated to the development of HER-2/neu protein-specific T-cell immunity. Epitope spreading represents the generation of an immune response to a particular portion of an immunogenic protein followed by the natural spread of immunity to other antigenic determinants present on the same protein. For example, Disis *et al.* observed epitope spreading in 84% of patients afflicted with HER-2/neu overexpressing malignancies who were administered vaccines comprising peptides derived from potential T-helper epitopes of the HER-2 protein mixed with granulocyte-macrophage colony stimulating factor (J. Clin. Oncol. (2002) 20(11): 2624-2632). Importantly, epitope spreading was correlated with the generation of a HER-2/neu protein domain response and suggests that immunization effectively circumvented immunologic tolerance.

[0089] TAAs that are suitable for use in the disclosed adenoviral vectors and methods as a target for a cancer vaccine should possess a number of characteristics. For example, a target TAA must have a favorable expression profile, meaning that it should be preferentially expressed or overexpressed in the tumor or malignant tissue as compared with normal tissue. In addition, because TAAs that play a role in tumorigenesis are more likely to be retained during the different stages of cancer progression, a suitable target TAA should also preserved throughout tumor progression and metastases. Suitable target TAAs should also be expressed homogenously within the tumor. Third, suitable target TAAs must not be subject to absolute immunologic tolerance. More specifically, there should be some evidence that T cells which can both recognize and respond to the TAA of interest have not been entirely deleted from the host's T-cell repertoire (Berinstein, N. L., J. Clin. Oncol. 29(8): 2197 (2002).

[0090] Carcinoembryonic antigen (CEA) has many characteristics which make it an attractive TAA for use as a target antigen for an anticancer vaccine. It is a member of the Ig superfamily which is characterized by a favorable expression pattern. It is expressed in more than 50% of all human cancers and has been implicated in the tumorigenesis process, which suggests that its expression may be selected and conserved throughout cancer progression. In addition, it has been established that immunologic tolerance to CEA is not absolute. Published studies establish that human T cells can recognize, become activated to, and lyse cancer cells that express CEA (Berinstein, N. L., J. Clin. Oncol. 29(8): 2197 (2002). For example, the immunization of patients with recombinant vaccinia virus expressing CEA, combined with subsequent peptide-based in vitro stimulations, generated CD8+ MHC-restricted CTLs capable of lysing autologous tumors (Tsang, K. Y. et al. J. Natl. Cancer Inst., (1995) 87:982-990). Alternatively, immunization of colorectal carcinoma patients after surgery with recombinant CEA was reported to induce weak antibody and cellular responses to recombinant CEA (Samanci, A., et al. (1998) Cancer Immunol. Immunother. 47: 131-142.) Further, the administration of anti-CEA anti-idiotypic antibody to patients diagnosed with colorectal cancer generated anti-CEA antibodies and idiotype-specific T-cell proliferation (Foon, L, A. et al. (1995) J. Clin. Invest.., 96: 334-342). The literature also indicates that tolerance to CEA in cancer patients can be overcome with several different vaccination approaches (i.e., vaccination with recombinant CEA or recombinant orthopox or avipox-CEA viruses, administration of anti-idiotype antibodies, pulsing dendritic cells with CEA agonist epitopes).

[0091] CEA is an oncofetal glycoprotein that is expressed in normal fetal colon and to a much lesser extent in normal colonic mucosa. It is also overexpressed in the vast majority of adenocarcinomas, particularly those of the colon, pancreas, breast, lung, rectum and stomach. Many colorectal cancers and some carcinomas produce high levels of CEA that are measurable in sera, which makes it one of the most widely used serological markers of malignancy, especially in patients with colorectal cancer.

[0092] A second TAA which provides a suitable immunogen for use in the compositions and methods of the invention is product of the HER2/erb-2 (also called neu) proto-oncogene. Like, CEA, HER2/neu has a favorable expression pattern and is not subject to absolute tolerance. More specifically, low levels of expression of the HER2/neu transcript, and the 185 kD polypeptide product, are detected in normal adult epithelial cells of various tissues, including the skin and breast, and tissues of the gastrointestinal, reproductive, and urinary tracts; higher levels of expression are detected in the corresponding fetal tissues during embryonic development (Press et al., Oncogene 5: 953-962 (1990). Several lines of evidence suggest a link between the amplification of HER-2 and neoplastic transformation in human breast, lung, prostate,

ovarian, endometrial and colorectal tumors (Disis and Cheever, Adv. Cancer Research 71: 343-371(1997). Generally speaking, overexpression of HER2/neu correlates with a poor prognosis and a higher relapse rate for cancer patients (Slamon et al., Science 244: 707-712 (1989). Thus, a vaccine specific for the HER-2/neu protein could have wide application and utility in the prevention of disease recurrence in many different human malignancies.

[0093] HER2/neu encodes a transmembrane glycoprotein possessing intrinsic tyrosine kinase activity and displaying extensive homology to the epidermal growth factor (EGF) receptor (Akiyama, T et al., (1986) Science 232: 1644-1646). One of the first clinical studies which utilized HER2 as target for cancer immunotherapy employed the HER-2-specific monoclonal antibody Herceptin for the treatment of breast cancer (Goldenberg MM (1999) Clin. Ther. 21: 309-318). This led to subsequent efforts which focused on the use of HER-2 as a target for the T-cell arm of the immune system to elicit effective antitumor responses, including the use of recombinant fusion proteins comprising HER-2 domains to activate autologous antigen presenting cells. Published reports establish that numerous cancer patients afflicted with neu-expressing mammary and ovarian cancers mount immune responses (e.g., produce antigen-specific antibody and T-cells) against the protein product of the HER2/neu oncogene.

[0094] Assembly of the recombinant adenoviral sequences, transgene and other vector elements into various intermediate plasmids and shuttle vectors, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques as described in standard textbooks that are well known to those of skill in the art (Sambrook et al, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989). Such techniques include, but are not limited to conventional cDNA cloning techniques, use of overlapping oligonucleotde sequences derived from the adenoviral genome, homologous recombination, polymerase chain reaction, standard transfection techniques, plaquing of viruses in agar overlay and other related methodologies.

[0095] To assist in preparation of polynucleotides in prokaryotic cells, a plasmid version of the adenovirus vector is often prepared (adenovirus pre-plasmid). The adenovirus pre-plasmid contains an adenoviral portion and a plasmid portion. The adenoviral portion is essentially the same as the adenoviral portion contained in the adenoviral vectors of the invention (containing adenoviral sequences with non-functional or deleted E1 and optionally E3 regions) and an immunogen expression cassette, flanked by convenient restriction sites.

[0096] The plasmid portion of the adenovirus pre-plasmid often contains an antibiotic resistance marker under transcriptional control of a prokaryotic promoter so that expression of the antibiotic does not occur in eukaryotic cells. Ampicillin resistance genes, neomycin resistance genes and other pharmaceutically acceptable antibiotic resistance markers may be used. To aid in the high level production of the polynucleotide by fermentation in prokaryotic organisms, it is advantageous for the adenovirus pre-plasmid to contain a prokaryotic origin of replication and be of high copy number. A number of commercially available prokaryotic cloning vectors provide these benefits. It is desirable to remove non-essential DNA sequences. It is also desirable that the vectors not be able to replicate in eukaryotic cells. This minimizes the risk of integration of polynucleotide vaccine sequences into the recipients' genome. Tissue-specific promoters or enhancers may be used whenever it is desirable to limit expression of the polynucleotide to a particular tissue type.

[0097] Adenovirus pre-plasmids (plasmids comprising the genome of the replication-defective adenovirus with desired deletions and insertions) can be generated by homologous recombination using adenovirus backbones DNA and an appropriate shuttle vector (designed to target-in specific deletions and incorporate desired restriction sites into the resultant plasmid). Shuttle vectors of use in this process can be generated using general methods widely understood and appreciated in the art, *e.g.,* PCR of the adenoviral terminal ends taking into account the desired deletions, and the sequential cloning of the respective segments into an appropriate cloning plasmid. The adenoviral pre-plasmid can then be digested and transfected into the complementing cell line via calcium phosphate co-precipitation or other suitable means. Virus replication and amplification then occurs, a phenomenon made evident by notable cytopathic effect. Infected cells and media are then harvested after viral replication is complete (generally, 7-10 days post-transfection).

[0098] Generally speaking, following the construction and assembly of the desired adenovirus pre-plasmids, adenovirus pre-plasmids are rescued into virus by transfecting an adenoviral E1-expressing human cell line. Complementation between the packaging cell line and the viral genes of the vector permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the production of recombinant adenoviruses. The resulting viruses may be isolated and purified by any of a variety of methods known to those of skill in the art for use in the methods of the invention.

[0099] It will be readily apparent to those of skill in the art that when one or more selected deletions of chimpanzee adenoviral genes are introduced into a viral vector, the function of the deleted gene product can be supplied during the production process by sequences present in the production cell line. Thus, the function of the manipulated genes can be provided by a permanently transformed cell line that is characterized by some or all of the adenoviral functions which are required for packaging but which are not functional in the vector (e.g., any of E1A, E1B, E2A, E2B E4). Alternatively, the requisite adenoviral functions can be provided to a suitable packaging cell line by infecting or transiently transfecting a suitable cell with a construct comprising the requisite gene to provide the function.

[0100] Accordingly, disclosed is a method of producing chimpanzee adenoviral vectors in E1-expressing human cell

lines. More specifically, the disclosed vectors can be propagated in an E1 complementing cell lines, including the known cell lines 293 and PER.C6™. Both these cell lines express the adenoviral E1 gene product. PER.C6™ is described in WO 97/00326, published January 3, 1997, which is hereby incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient first generation adenoviruses, but is designed to prevent generation of replication competent adenovirus by homologous recombination. 293 cells are described in Graham et al (1977) J. Gen. Virol 36:59-72, which is also hereby incorporated by reference. One of skill in the art will recognize the term "first generation adenovirus" refers to a replication deficient adenovirus which has either a non-functional or deleted E1 region, and optionally a non-functional or deleted E3 region.

[0101] Batches of replication-defective adenoviral vectors that are intended for use as a vaccine composition in a clinical trial should be proven to be free of RCA (Fallaux, F.J. et al (1998) Human Gene Therapy, 9:1909-1917). In practice, this is a labor intensive process which requires establishing and utilizing an expensive screening program. One of skill in the art will acknowledge that a high frequency of RCA generation not only results in a high failure rate for the batches produced, but also severely limits scale-up efforts. Elimination of sequence homology between the nucleotide sequence of the vector and the adenoviral sequences present in the genome of the helper production/packaging cell line should eliminate the possibility of producing batches of vector that are contaminated with RCAs produced by homologous recombination.

[0102] Typically, recombinant replication-defective adenoviral vectors are propagated in cell lines that provide E1 gene products *in trans.* Supplementation of the essential E1 gene products *in* trans is very effective when the vectors are from the same or a very similar serotype. For example, it is well-known that E1-deleted (i.e. ΔE1) group C serotype (Ad2 and Ad5) vectors, can be propagated in 293 or PER.C6 cells which contain and express the Ad5 E1 region. However, it has been observed that Ad5 E1 sequences present in the 293 and PER.C6 production cells may not always fully complement the replication of non-group C serotypes. Accordingly, E1-deleted serotypes outside of subgroup C, for example those from subgroups A, B, D, E, and F may replicate with a lower efficiency respect to the corresponding wt virus or may not replicate at all in 293 or PER.C6 cells. This may be due to the inability of the Ad5 (group C) E1B 55K gene product to establish a functional interaction with the E4 orf6 gene product of the non-group C serotypes.

[0103] The decrease in replication efficiency in cells expressing Ad5 E1 is variable considering vectors of different subgroups. While ΔE1 vectors deriving from subgroup D and E adenovirus can be rescued and propagated in 293 and Per.C6™ cells with variable efficiency, the propagation ΔE1 vectors of subgroup B is completely impaired (Vogels R, et. al. (2003) Aug. Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. J Virol. 77 (15):8263-71).

[0104] Although the interaction between Ad5 E1b 55k and vector-expressing E4 orf6 protein is conserved within members of the same subgroup, it may be not sufficiently stable when E4 orf6 protein of a non-C serotype is expressed. This inefficient or unstable formation of E1B-55K/E4-orf6 complex lead to an absent of reduced propagation of the ΔE1 vector. Accordingly, it has been empirically determined that in order to successfully and efficiently rescue recombinant adenovirus of groupB serotypes, a cell line expressing the E1 region of the serotype of interest may need to be generated. In cells expressing Ad5E1 like 293 or Per.C6™, the expression can be limited to E1b 55K protein. Alternatively, a suitable Ad5E1-expressing cell lines could be modified to express the entire Ad5 E4 region (or E4 orf6 only) in addition to Ad5E1. The generation of cell lines expressing both Ad5 E1 and orf6 are useful in complementing alternative adenovirus serotypes; *see, e.g.,* Abrahamsen et al., 1997 J. Virol. 8946-8951. The incorporation of E4 (orf6) into Ad5 complementing cell lines, is known, as is the generation of serotype-specific cell lines providing a serotype-specific E1 gene product(s) *in trans.* Alternatively, the efficiency of non-group C vector propagation may be improved by modification of the viral backbone by substituting the native E4 region with Ad5 orf6. Similar results can be achieved by substituting the only the native orf6 with orf6 deriving from Ad5 or other subgroup C viruses (Ad1, Ad2, Ad6). U.S. Patent No. 5,849,561 discloses complementation of an E1-deleted non-group C adenovirus vector in an Ad5-E1 complementing cell line which also expresses portions of the Ad5-E4 gene.

[0105] U.S. Patent No. 6,127,175, issued to Vigne, et al., discloses a stably transfected mammalian cell line which expresses a portion of the E4 region of adenovirus, preferably orf6/orf6/7. Such a cell line is useful for complementation of recombinant Ad genomes deficient in the E4 region.

[0106] Compositions, including vaccine compositions, comprising the disclosed adenoviral vectors are disclosed. These compositions can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. Potential hosts/vaccinees include but are not limited to primates and especially humans and non-human primates, and include any non-human mammal of commercial or domestic veterinary importance. Compositions comprising recombinant chimpanzee adenoviral vectors may be administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen.

[0107] The disclosed vectors may be used in an immunization protocol designed to break host tolerance to a self-antigen or a tumor-associated antigen. The identification of a number of TAA has enabled the development of active vaccination approaches for the therapy of cancer. Both cell surface antigens and intracellular antigens that are processed and presented provide useful targets. Generally speaking, the disclosed method of breaking host tolerance to a self-

antigen comprises: (a) stimulating an antigen-specific response to a self-antigen by administering a first vaccine composition comprising a first ChAd vector or a plasmid vector carrying a nucleotide sequence encoding the self-antigen against which an antigen-specific immune response is desired, and (b) sustaining and expanding the immune response of (a) by administering a second vaccine composition comprising a recombinant ChAd vector of a different serotype containing at least a functional deletion of its genomic E1 gene, and in the site of the E1 gene, a sequence comprising a promoter capable of directing the expression of DNA encoding the same self-antigen delivered in the priming step, whereby the host mounts an immune response which has the effect of breaking tolerance to the self-antigen.

[0108] Accordingly, a skilled artisan can utilize this disclosure to design several different immunization protocols that may be suitable for use to break host tolerance. For example, it may be possible to utilize a protocol in which the first, or priming immunization comprises plasmid DNA which encodes a particular self-antigen, such as a TAA, and any subsequent immunizations comprise a ChAd vector. Plasmid DNA sequences comprising nucleotide sequences that encode self-antigens, may be delivered intramuscularly, with or without electrostimulation, in one or more injections. For example, an immunization protocol based on multiple (e.g., 3 or 4 or 5) intramuscular injections of plasmid DNA encoding a TAA via electroporation followed by one or more intramuscular injections of a ChAd vector comprising a transgene encoding the same TAA is encompassed by the general method disclosed and claimed herein.

[0109] Alternatively, a suitable protocol to break tolerance could involve one or more priming immunizations with a ChAd or hAd vector comprising a transgene encoding a self antigen, followed by one or more boosting immunizations with either the same, or a different ChAd vector that is know to be non cross-reactive with the vector used for the priming immunization(s). For example, an immunization protocol using ChAd3 for priming and ChAd6 for boosting, or ChAd3 for priming followed by ChAd6 and ChAd9 for boosting could be used to break host tolerance. In particular embodiments, the invention contemplates the use of self-antigens comprising at least one tumor associated antigen selected from the group consisting of: HER2/neu, CEA, EpCAM, PSA, PSMA, Telomerase, gp100, Melan-A/MART-1, Muc-1, NY-ESO-1, Survivin, Stromelysin 3, Tyrosinase, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, hKLP2, VEGF. In a particular embodiment, the invention provides a method for inducing an immune response (e.g., humoral or cell-mediated) to a tumor-associated antigen which is specific for a selected malignancy by delivering a recombinant chimpanzee adenovirus encoding the TAA to a mammal afflicted with cancer. In a preferred embodiment of this aspect of the invention the elicited immune response constitutes an immune response characterized by the production of antigen-specific CD4+ and CD8+ T cells.

[0110] The immunogenic compositions can be administered to mammalian hosts, preferably human hosts, in either a prophylactic or therapeutic setting. Potential hosts/vaccinees include but are not limited to primates and especially humans and non-human primates, and include any non-human mammal of commercial or domestic veterinary importance. Compositions comprising recombinant chimpanzee adenoviral vectors may be administered alone or in combination with other viral- or non-viral-based DNA/protein vaccines. They also may be administered as part of a broader treatment regimen. Suitable compositions, for use in the methods of the invention may comprise the recombinant viral vectors of the invention in combination with physiologically acceptable components, such as buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use. Optionally, a vaccine composition of the invention may be formulated to contain other components, such as but not limited to, an adjuvant, a stabilizer, a pH adjusting agent, or a preservative. Such components are well known to those of skill in the art.

[0111] It is envisioned that the recombinant chimpanzee adenoviruses of the invention will be administered to human or veterinary hosts in an "effective amount," that is an amount of recombinant virus which is effective in a chosen route of administration to transduce host cells and provide sufficient levels of expression of the transgene to invoke an immune response which confers a therapeutic benefit or protective immunity to the recipient/vaccine.

[0112] The amount of viral particles in the vaccine composition to be introduced into a vaccine recipient will depend on the strength of the transcriptional and translational promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of $1x10^7$ to $1x10^{12}$ particles (i.e., $1x10^7$, $2x10^7$, $3x10^7$, $5x10^7$, $1x10^8$, $2x10^8$, $3x10^8$, $5x10^8$ or $1x10^9$, $2x10^9$, $3x10^9$, $5x10^9$) and preferably about $1x10^{10}$ to $1x10^{11}$ particles is administered directly into muscle tissue. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, or inhalation delivery are also contemplated.

[0113] The recombinant chimpanzee adenoviral vectors of the present invention may be administered alone, as part of a mixed modality prime/boost vaccination regimen or in a vaccination regimen based on combination of multiple injections of different vector serotypes. Typically, a priming dose(s) comprising at least one immunogen is administered to a mammalian host in need of an effective immune response to a particular pathogen or self-antigen. This dose effectively primes the immune response so that, upon subsequent identification of the antigen(s), the host is capable of immediately mounting an enhanced or boosted immune response to the immunogen. A mixed modality vaccination scheme which utilized alternative formulations for the priming and boosting can result in an enhanced immune response. Prime-boost administrations typically involve priming the subject (by viral vector, plasmid, protein, *etc.*) at least one time, allowing a predetermined length of time to pass, and then boosting (by viral vector, plasmid, protein, *etc.*). Multiple

immunizations, typically 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, albeit other time frames may be used as one of ordinary skill in the art will appreciate. Multiple injection of each vector can be administered within approximately a 2 weeks time frame, before neutralizing immunity becomes evident.

[0114] In some embodiments of this invention, a vaccine is given more than one administration of chimpanzee adenovirus vaccine vector, and it may be given in a regiment accompanied by the administration of a plasmid vaccine. Suitable plasmid vaccines for use in combination with the vectors disclosed herein comprise a plasmid encoding at least one immunogen against which a primed or boosted immune response is desired, in combination with a heterologous promoter, which is capable of directing expression of the nucleic acid sequences encoding the immunogen(s), operably linked to the immunogen coding sequence, and a transcription terminator sequence.

[0115] For example, a dosing regimen which utilizes multiple injection of different serotypes of recombinant replication-defective chimpanzee adenoviral vectors can be used. Alternatively, an individual may be given a first dose (i.e., a priming dose) of a plasmid vaccine, and a second dose (i.e., a boosting dose) which comprises a replication-defective recombinant chimpanzee adenoviral vector which comprises a coding sequence for the same immunogen that was delivered in the plasmid vaccine. Alternatively, the individual may be given a first dose of a human adenovirus vaccine vector encoding at least one immunogen, followed by a second dose comprising a replication-defective recombinant chimpanzee adenoviral vector disclosed herein, which comprises a coding sequence for the same immunogen that was delivered in the priming dose. In a second alternative embodiment a vaccine composition comprising a vector of the invention may be administered first, followed by the administration of a plasmid vaccine. In any of these embodiments, an individual may be given multiple doses of the same immunogen in either viral vector or plasmid form. There may be a predetermined minimum amount of time separating the administrations.

[0116] In addition to a single protein or antigen of interest being delivered by the recombinant, replication-defective chimpanzee adenovirus vectors of the present invention, two or more proteins or antigens can be delivered either via separate vehicles or delivered *via* the same vehicle. Multiple genes/functional equivalents may be ligated into a proper shuttle plasmid for generation of a adenovirus pre-plasmid comprising multiple open reading frames. Open reading frames for the multiple genes/functional equivalents can be operatively linked to distinct promoters and transcription termination sequences.

[0117] As shown herein, suitable immunization regimens can employ different adenoviral serotypes. One example of such a protocol would be a priming dose(s) comprising a recombinant adenoviral vector of a first serotype, for example a ChAd3 or ChAd6 followed by a boosting dose comprising a recombinant chimpanzee adenoviral vector of a second serotype. In an alternative embodiment, the priming dose can comprise a mixture of separate chimpanzee adenoviral vehicles each comprising a gene encoding for a different protein/antigen. In such a case, the boosting dose would also comprise a mixture of vectors each comprising a gene encoding a separate protein/antigen, provided that the boosting dose(s) administers recombinant viral vectors comprising genetic material encoding for the same or similar set of antigens that were delivered in the priming dose(s). These multiple gene/vector administration modalities can further be combined. It is further within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen.

[0118] Use of recombinant vectors derived from chimpanzee adenoviruses that are not neutralized by preexisting immunity directed against the viral elements of human vector offers an alternative to the use of human Ad vectors as vaccine carriers. Because adenoviruses are highly immunogenicity, adenoviral vectors are particularly good candidates for use in the context of a vaccine carrier designed to break host tolerance to a self-antigen. Furthermore, the ability to propagate the chimp viruses in human cells, particularly in the Per.C6™ cell line, with an efficiency comparable to human viruses, offers considerable advantages both from a regulatory point of view and for the large scale production of therapeutics or vaccines. Accordingly, the instant invention provides a collection of chimpanzee adenoviral sequences, vectors and plasmids that allow the preparation of recombinant virus which may be used, alone or in combination, as a vaccine carrier for genetic vaccination.

[0119] The following examples illustrate, but do not limit the invention.

Example 1: Isolation, Cloning, Sequencing And Characterization Of ChAds Chimpanzee Adenovirus Isolation

[0120] Stool specimens were collected in viral transport medium (VTM; Microtest M4-R Multi-Microbe Transport Medium, Remel Inc.) then frozen or frozen directly at -70°C at NIRC (New Iberia Research Center 4401 W. Admiral Doyle Drive New Iberia, LA 70560). The specimens were kept frozen at < -70°C until they were processed for inoculation into cell cultures. At that time, the specimens were thawed and then vortexed in excess of chilled viral transport medium. After the specimens had dissociated into suspensions, they were centrifuged for 10 min at 1500-1800 rpm. The supernatants were filtered through 0.8 and 0.2 μm syringe filters in series and then the filtered material was inoculated into cell cultures (200-250 μL into shell vials and 250-300 μL into tube cultures). Each processed specimen was inoculated into tube cultures and shell vial cultures seeded with 293 cells or A549 cells.

[0121] Control (positive and negative) cultures were prepared each time a set of samples was inoculated. Once all of the shell vials in a set-up had been inoculated, they were centrifuged at room temperature for 60 ± 10 min at 2000 rpm (900 x g). The vials were removed from the centrifuge immediately after the rotor stopped spinning to prevent heat damage in the cultures. After centrifugation, the inocula were aspirated from the shell vials, using a fresh sterile Pasteur pipette in each vial to prevent cross-contamination. The cultures were washed three times using 1.0-mL fresh culture medium for each wash. Fresh medium (1.0 mL) was pipetted into each vial after the third wash and the shell vials were placed in an incubator at 35-37°C for three to four days (approx. 96 hr).

[0122] At the end of the culture period, the supernatants were aspirated from the cultures and the cell layer in each vial was washed twice with Immunofluorescence Assay (IFA) Buffer using approximately 1.0 mL buffer with each wash. The cells were fixed by adding 1.0 mL refrigerated acetone to each vial (10 min at 2-8°C. Acetone-cleaned slides were labeled with the specimen identification number(s) associated with the shell vial coverslips. The shell vial coverslips were processed for fluorescence labeling of Adenovirus-infected cells using a primary mouse anti-adenovirus antibody [MAB8052, Chemicon]. The slides are evaluated with the aid of a fluorescence microscope. Each preparation was scanned using the 10X objective noting the extent of immunofluorescence coverage across the well (1+ to 4+). The presence or absence of specific immunofluorescence was confirmed using the 40X objective. Tube cultures were inoculated in the same sequence as described for the shell vials (e.g., negative control first, followed by clinical specimens and positive controls). The inocula were allowed to adsorb for 60-120 min at 36-38°C. After the adsorption period, the specimens/controls were aspirated from the tubes and replaced by fresh culture medium.

[0123] Three to four days post-inoculation, and once a week thereafter, the media was aspirated from the culture tubes and replaced with 1.5 mL fresh media. Culture tubes were visually monitored for CPE at least every other day for at least 21 days after inoculation. Cultures inoculated with chimp specimens were compared against the controls and rated by observing the CPE extent. Cultures showing no CPE were passed to fresh tube cultures after 14 days; culture tubes that were negative for CPE after 21 days were considered negative. Culture tubes with 3-4+ CPE were vortexed for 10 seconds. The cells were scraped from the wall of the tube using a sterile 1.0 mL serological pipette and suspended in the culture supernatant. After labeling a 5 mL snap cap tube with the specimen identification number and date and stored at -70°C. 500 μL of the cell suspension was transferred from the culture tube into the snap cap tube and stored for up to one day at 2-8°C until it was processed using an indirect immunofluorescent antibody technique to detect adenovirus (equivalent to procedure for staining shell vials).

Chimpanzee Adenovirus Amplification

[0124] Wild type chimp adenoviruses CV32, CV33, CV23 and CV68 purchased from the ATCC (ATCC Accession Numbers: CV32, VR-592; CV-33, VR-593;) or from Esoterix Inc. Austin, Texas and original isolates were propagated as follows by using the human E1-expressing cell line PER.C6™ or 293. Briefly, cells were cultivated in Dulbecco's Modified Eagles Medium (DMEM; GibcoBRL, Life Technologies) supplemented with 10% Fetal Bovine Serum (FBS GibcoBRL, Life Technologies), 1% Penicillin-Streptomycin, 2mM Glutamine and 10mM $MgCl_2$ (Per.C6™). Adenovirus infection was carried out in DMEM supplemented with 5% Horse Serum (GibcoBRL, Life Technologies). Infected cells and medium were collected when 100% of the cells exhibited virus-induced cytopathic effect (CPE) and lysed by three cycles of freezing and thawing.

[0125] All wild type chimp adenoviral (CV) stocks were cloned by infecting 293 cells seeded in 96-well plates, after the first passage of amplification. The virus cloning was performed by limiting dilution of the cell lysate obtained at the first passage of the virus amplification. Five isolated clones were picked up and serially propagated. After 3-4 serial passages of amplification, a large-scale preparation of adenovirus was performed on cells planted on 5 two-layer cell-factories (NUNC) (200 millions of cells/cell factory). Purified viral particles were obtained from cell lysate by two ultra-centrifugation steps on cesium chloride density gradients.

Sequencing of Viral Genomic DNA

[0126] Genomic DNA was isolated from 3 X 10^{12} pp of purified virus preparation by digestion with Proteinase K (0.5 mg/ml) in 1% SDS-TEN (2 hrs at 55°C). After a Phenol-Chloroform extraction and Ethanol precipitation, the genomic DNA was resuspended in water and submitted for genomic sequencing.

[0127] For full length Ad genome sequencing, the purified viral DNA was nebulized to produce randomly sheared fragments. The DNA fragments were blunt-ended with the klenow fragment of E.coli DNA polymerase and polynucleotide kinase. The blunt end fragment were run on a low melting point agarose gel to purify the fragments in the size range of 1-3 kb and cloned into the SmaI site of pUC19 vector to create a shotgun library. The ligations were used to transform competent XL1-Blue MRF'. Positive colonies were identified by white/blue screening on LB agar containing X-gal and IPTG. Three to four 96-well block of plasmid DNA were isolated from the library and sequenced with pUC forward and reverse primers. All sequencing reads were screened for quality and vector sequence using the Phred-Phrap software

package. The reads that passed the screening were assembled into contigs. Primers were designed to directly sequence the adenoviral DNA for closing the gaps and determine the DNA sequence of both ends.

[0128] Complete viral genome sequencing was obtained for selected viruses including ChAd3 (SEQ ID NO: 1), ChAd6 (SEQ ID NO: 2), CV32 (SEQ ID NO: 3), CV33 (SEQ ID NO: 4), and CV23 (SEQ ID NO: 5). Table 1 provides data summarizing the percentage of identity between the nucleotide sequences of ChAd3, ChAd6, Pan5 (CV23), Pan6 (CV32), Pan7 (CV33), C1 and C68 adenoviral genomes. Alignments were calculated using the ALIGN program as part of the FASTA package version 2 (William R. Penson, University of Virginia; Myers & Miller, CABIOS 1989, 4:11-17).

Table 1. Percentage of Nucleotide Sequence Identity Between Chimpanzee Adenovirus Genomes

	ChAd3	ChAd6	Pan5	Pan6	Pan7	C1	C68
ChAd3	100	68.1	68.5	68.2	68.3	64.2	68.0
ChAd6		100	95.5	94.5	95.5	73.6	91.4
Pan5			100	94.9	96.7	73.9	92.7
Pan6				100	95.1	73.6	91.3
Pan7					100	73.8	93.0
C1						100	74.3
C68							100

[0129] To characterize the new adenoviral isolates (e.g., ChAd20, ChAd4, ChAd5, ChAd7, ChAd8, ChAd9, ChAd10, ChAd11, ChAd16, ChAd17, ChAd19, ChAd22, ChAd24, ChAd26, ChAd30, ChAd31, ChAd37, ChAd38, ChAd44, ChAd63 and ChAd82) the nucleotide sequence of the hexon and fiber genes were also determined by primer walking. Fiber gene: SEQ ID NOS: 6- 15: (SEQ ID NO: 6, ChAd20); (SEQ ID NO: 7, ChAd4); SEQ ID NO: 8, ChAd5); SEQ ID NO: 9, ChAd7); SEQ ID NO: 10, ChAd9); SEQ ID NO: 11, ChAd10); SEQ ID NO: 12, ChAd11); SEQ ID NO: 13, ChAd16) SEQ ID NO: 14, ChAd17), SEQ ID NO: 15, ChAd19), and (SEQ ID NO: 58, ChAd8), (SEQ ID NO: 60, ChAd22), (SEQ ID NO: 62, ChAd24), (SEQ ID NO: 64, ChAd26), (SEQ ID NO: 66, ChAd30), (SEQ ID NO: 68, ChAd31), (SEQ ID NO: 70, ChAd37), (SEQ ID NO: 72, ChAd38), (SEQ ID NO: 74, ChAd44), (SEQ ID NO: 76, ChAd63) and (SEQ ID NO: 78, ChAd82). Figures 20A-20G provide a comparison of the amino acid sequences of the fiber proteins of the ChAd isolates disclosed and claimed herein.

[0130] The hexon gene sequences are set forth in SEQ ID NOS: 16-25: (SEQ ID NO:16, ChAd20); SEQ ID NO: 17, ChAd4); SEQ ID NO: 18, ChAd5); SEQ ID NO: 19, ChAd7); SEQ ID NO: 20, ChAd9); SEQ ID NO: 21, ChAd10); SEQ ID NO: 22, ChAd11); SEQ ID NO: 23, ChAd16); SEQ ID NO: 24, ChAd17) SEQ ID NO: 25, ChAd19), (SEQ ID NO: 97, ChAd8), (SEQ ID NO: 99, ChAd22), (SEQ ID NO:101, ChAd24), (SEQ ID NO: 103, ChAd26), (SEQ ID NO: 105, ChAd30), (SEQ ID NO: 107, ChAd31), (SEQ ID NO: 109, ChAd37), (SEQ ID NO: 111, ChAd38), (SEQ ID NO: 113, ChAd44), (SEQ ID NO: 115, ChAd63) and (SEQ ID NO: 117, ChAd82). Figures 31A-31J provide a comparison of the amino acid sequences of the hexon proteins of the ChAd isolates disclosed and claimed herein.

Chimpanzee Adenovirus Classification

[0131] Classification of the different chimp adenoviral strains follows the already proposed classification of human adenovirus serotypes into 6 subgroups (Horowitz, MS (1990) Adenoviridae and their replication. In Virology B.N. Fields and D.M. Knipe, eds (Raven Press, New York) pp.1679-1740) and it was obtained by amino acid and nucleotide sequence alignment by using Align X program (Informax, Inc).

[0132] An initial classification of the new isolates was obtained by looking at the restriction pattern of the viral genome with different restriction endonucleases and by sequence analysis of the hypervariable region 7 (HVR7) of the hexon gene. To this end two primers were designed on the highly conserved regions flanking HVR7: TGTCCTACCARCTCTT-GCTTGA (SEQ ID NO. 45) and GTGGAARGGCACGTAGCG (SEQ ID NO. 46). The HVR7 was amplified by PCR using purified viral DNA or crude 293 lysate as template and then sequenced. Based on HVR7 sequence analysis we classified the new isolated viruses into the subgroups (A-F) proposed for human Ad viruses (Horowitz, MS (1990) Adenoviridae and their replication. In Virology B.N. Fields and D.M. Knipe, eds (raven Press, New York) pp.1679-1740).

[0133] The phylogenetic tree presented in Figure 35 was obtained by alignment of human and chimp adenovirus hexon amino acid sequences. The results are consistent with the initial classification based on nucleotide sequence alignment limited to hexon HVR7 by using Align X program (Informax, Inc). The tree was deduced from a multiple sequence alignment of full-length hexon peptide sequences using a PAUPSEARCH (Wisconsin Package Version 10.3, Accelrys Inc.) and visualized and manipulated with TREEVIEW. Bootstrap confidence analysis was performed using the PAUP-SEARCH program as implemented in the Wisconsin Package. For each of the alignments the program was run on 1000 replicates using "Heuristic Search" as search criterion and Maximum Parsimony as the optimality criterion and confidence

values reported were taken from a 50% majority-rule consensus.

Example 2: ChAd Shuttle Vector and Expression Vector Construction and Rescue Vector Construction and Rescue

[0134] Genomic viral DNA was cloned into a standard plasmid vector by homologous recombination with an appropriate shuttle vector containing viral DNA sequences derived from both left and right end of viral genome (Figure 2). As described more fully below, the sequence homology observed between viruses classified in the same serotype subgroup was exploited to develop group-specific shuttle vectors. Genomic viral DNA of Chimp adenovirus classified into subgroup D and E resulted to be sufficiently homologous to allow the construction of a common shuttle vector in order to clone viruses belonging to both subgroups.

Construction of a Subgroup D/E Shuttle Vector

[0135] The ChAd6 viral genome was fully sequenced (SEQ ID NO: 2) and the information obtained was used to construct a shuttle vector to facilitate cloning by homologous recombination of subgroup D and E chimpanzee adenovirus.

[0136] Construction of the ChAd6 shuttle vector, referred to herein as pARS ChAd6-3 is described in Figure 1. Figure 32 provides a list of the oligonucleotide sequences (SEQ ID NOS: 26-40 and SEQ ID NOS: 45-46) used in the cloning experiments described herein. Briefly, 457 bp deriving from the left end of ChAd6 DNA were amplified by PCR with the oligonucleotides 5'-ATGGAATTCGTTTAAACCATCATCAATAATATACCTC-3 (SEQ ID NO: 27) and 5'-CGCTGGCACT-CAAGAGTGGCCTC-3' (SEQ ID NO: 28) digested with EcoRI and SnaBI and cloned into pNEBAd35-2 cut EcoRI-SnaBI, generating pNEBChAd6-LI. The right ChAd6 ITR (bp 36222 to bp 36648) was amplified by PCR using the oligonucleotides: 5'- ATGAAGCTTGTTTAAACCCAT CATCAATAATATACCT-3 '(SEQ ID NO: 29) and 5'- ATCTAGACAGCGTCCAT-AGCTTACCG-3'(SEQ ID NO: 30) digested with restriction enzymes HindIII and XbaI and cloned into pNEBChAd6-LI HindIII-XbaI digested thus generating pNEBChAd6-RLI. Finally, the DNA fragment corresponding to nucleotides 3426-3813 of the ChAd6 genomic DNA sequence was amplified with the oligonucleotides: 5' ATGCTACGTAGCGATCGCGTGAGTAGTGTTTGGGGGTGGGTGGG-3' (SEQ ID NO: 31) and 5'-TAG-GCGCGCCGCTTCTCCTCGTTCAGGCTGGCG-3' (SEQ ID NO: 32), digested with SnaBI and AscI then ligated with SnaBI-AscI digested pNEBChAd6-RLI thus generating pNEBChAd6-RLIdE1.

[0137] To improve the efficiency of recombination and plasmid propagation in DH5a *E. coli* strain, the 1306 bp fragment containing both left and right ITRs of ChAd6 as well as pIX gene fragment was excised by PmeI digestion from pNEBChAd6-RLIdE1 and transferred to a different plasmid vector obtained by PCR amplification with the olinucleotides 5'-GATCTAGTTAGTTTAAACGAATTCGGATCTGCGACGCG-3' (SEQ ID NO: 33) and 5' TTCGATCATGTTTAAAC-GAAATTAAGAATTCGGATCC-3' (SEQ ID NO: 34) from pMRKAd5SEAP. This final ligation step generated the ChAd6 shuttle vector pARSChAd6-3.

Construction of a Subgroup C Shuttle Vector

[0138] The ChAd3 viral genome was fully sequenced (SEQ ID NO: 1) and the information obtained was used to construct a shuttle vector to facilitate cloning by homologous recombination of subgroup C chimpanzee adenovirus.

[0139] Briefly, the shuttle vector used to clone subgroup C chimp adenovirus, referred to herein as pChAd3EGFP was constructed as follows: a ChAd3 DNA fragment (nt 3542-4105) containing pIX coding region was amplified by PCR with the oligonucleotides 5'-TATTCTGCGATCGCTGAGGTGGGTGAGTGGGCG-3' (SEQ ID NO: 35) and 5'-TAG-GCGCGCCCTTAAACGGCATTTGTGGGAG-3' (SEQ ID NO: 36) digested with SgfI- AscI then cloned into pARSCV32-3 digested with SgfI- AscI, generating pARS-ChAd3D. ChAd3 right end (nt 37320-37441) was amplified by PCR with oligonucleotides 5'-CGTCTAGAAGACCCGAGTCTTACCAGT-3' (SEQ ID NO: 37) and 5'-CGGGATCCGTTTAAACCAT-CATCAATAATATACCTTATT-3' (SEQ ID NO: 38) digested with XbaI and BamHI then ligated to pARS-ChAd3D restricted with XbaI and BamHI, generating pARS-ChAd3RD. ChAd3 viral DNA left end (nt 1-460) was amplified by PCR with oligonucleotides 5'-ATGGAATTCGTTTAAACCATCATCAATAATATACCTT-3' (SEQ ID NO: 39) and 5'-ATGACGCGATCGCTGATATCCTATAATAATAAAACGCAGACTTTG-3', (SEQ ID NO: 40) digested with EcoRI and SgfI then cloned pARS-ChAd3RD digested with EcoRI and SgfI, thus generating pARS-ChAd3RLD. The viral DNA cassette was also designed to contain restriction enzyme sites (PmeI) located at the end of both ITR's so that digestion will release viral DNA from plasmid DNA.

Construction of a Subgroup B Shuttle Vector

[0140] The construction of subgroup B shuttle followed the already described strategy for subgroup C and D/E shuttle constructions. In brief, pARS-ChAd3RLD was modified by substituting the left end, the pIX region, the right end with the

corresponding fragments of ChAd30. In addition the E4 region of ChAd30 was substituted with Ad5 E4orf6 that was cloned under the ChAd30 E4 promoter control. The shuttle plasmid was denominated pChAd30 EGFP shuttle vector.

Construction of ΔE1 Chimp Adenoviral Vectors

[0141] Subgroup B: Subgroup B chimp adenovirus vectors were constructed by homologous recombination in *E. coli* strain BJ5183. BJ5183 cells were co-transformed with pChAd30EGFP shuttle vector digested with BstEII and Bst1107I and ChAd8 and ChAd30, purified viral DNA. Homologous recombination between pIX genes, right ITR DNA sequences present at the ends of linearized pChAd30EGFP shuttle and viral genomic DNA allowed its insertion in the plasmid vector, deleting at the same time the E1 region that was substituted by EGFP expression cassette. Expression cassettes based on human cytomegalovirus (HCMV) promoter and bovine growth hormone polyadenylation signal (Bgh polyA) were constructed to express secreted alkaline phosphatase (SEAP), EGFP, HIV gag, HCV NS region (as described in fig.3 for ChAd6 shuttle vectors) as well as tumor-associated antigens like CEA and HER2/neu from human and Rhesus monkey origin. All expression cassette were inserted into ChAd30 vectors by homologous recombination.

[0142] Subgroup C: Subgroup C chimp adenovirus vectors were constructed by homologous recombination in *E. coli* strain BJ5183. BJ5183 cells were co-transformed with pChAd3EGFP shuttle vector digested with BstEII and Bst1107I and ChAd3, ChAd11, ChAd19 and ChAd20 purified viral DNA. Homologous recombination between pIX genes, right ITR DNA sequences present at the ends of linearized pChAd3EGFP and viral genomic DNA allowed its insertion in the plasmid vector, deleting at the same time the E1 region that was substituted by EGFP expression cassette. Expression cassettes based on human cytomegalovirus (HCMV) promoter and bovine growth hormone polyadenylation signal (Bgh polyA) were constructed to express secreted alkaline phosphatase (SEAP), EGFP, HIV gag, HCV NS region (as described in fig.3 for ChAd6 shuttle vectors) as well as tumor-associated antigens like CEA and HER2/neu from human and Rhesus monkey origin.

[0143] Subgroups D and E: In order to construct ΔE1 vectors based on subgroup D and E chimp adenovirus, the shuttle vector pARS ChAd6-3 was digested with AscI and co-transformed into *E. coli* strain BJ5183 with CV32, CV33, CV68, ChAd4, ChAd5, ChAd6, ChAd7, ChAd9, ChAd10 and ChAd16 purified viral DNA. Homologous recombination between DNA sequences from pIX genes and right ITR present at the ends of linearized pARS ChAd6-3 and viral genomic DNA allowed its insertion in the plasmid vector, deleting at the same time the E1 region (Figures 2 and 4). Expression cassettes based on human cytomegalovirus (HCMV) promoter and bovine growth hormone poly-adenylation signal (Bgh polyA) were constructed to express secreted alkaline phosphatase (SEAP), EGFP, HIV gag, HCV NS genes (Figure 3) as well as tumor-associated antigens like CEA and HER2/neu of human and Rhesus monkey origin. All the expression cassette were inserted into the single SnaBI site of pARS ChAd6-3 vector to be transferred by homologous recombination into the ΔE1 adenovirus pre-plasmids as described in Figure 4.

Rescue and Amplification of ΔE1 Vectors

[0144] $5X10^6$ PER.C6™ cells planted on 6cm cell culture dishes were transfected with 10 micrograms of cloned viral vector released from plasmid sequences by endonuclease digestion. DNA transfection was performed using Lipofectamine (Invitrogen). Transfected cells and culture medium were collected 5-10 days post-transfection and lysed by freeze-thaw. Rescued vectors were then amplified by serial passaging on 293 or PER.C6™ cells. A large-scale amplification was performed by infecting cells planted on 5-10 cell-factories (NUNC, Inc.) on a total of $1\text{-}2x10^9$ cells. A purified vector preparation was obtained on cesium chloride gradient by two ultra-centrifuge runs, dialyzed against PBS containing 10% glycerol and stored at -70°C in aliquots.

Example 3: Neutralization Studies

[0145] Neutralization assays were carried out in order to evaluate the prevalence in human sera of neutralizing antibodies against the chimpanzee adenoviruses disclosed herein. The assay evaluated the effects of serum preincubation on the ability of chimp adenoviruses carrying the gene for secreted alkaline phosphatase (SEAP) to transduce human 293 cells. The neutralization titer is defined as the dilution of serum giving a 50% reduction of the SEAP activity observed in the positive control with the virus alone.

[0146] From $2x10^6$ to $1.5x10^7$ physical particles of CV33-SEAP, CV32-SEAP and ChAd3-SEAP vector were diluted in 100 μl of complete medium and added to an equal volume of human or chimp serum diluted in complete medium. Each serum samples was tested at various dilutions (five 4-fold increments starting from 1/18 dilution through 1:4608). Samples were pre-incubated for one hour at 37°C and then added to 293 cells seeded into 96-well plates ($3x10^4$ cells/well). The inoculum was removed after one hour of incubation, the cells were re-fed with fresh medium and, 24 hours later, 50 μl of medium was removed and the SEAP activity was measured by a chemiluminescent assay. The neutralization titer is defined as the dilution of serum giving a 50% reduction of the SEAP activity observed in the positive

control with the virus alone. A panel of 100 human sera was tested for ChAd neutralization activity. In parallel the same panel was tested on Ad5 SEAP vector.

Table 2. Prevalence of Neutralizing Antibodies Against Chimpanzee Adenovirus

	Virus						
titer	hAd5	CV32	CV33	ChAd3	ChAd30	ChAd9	ChAd10
< 200	77%	96%	100%	92%	100%	92%	100%
> 200	33%	4%	0%	8%	0%	6%	0%

[0147] The result provided in Table 2 indicates that a very low prevalence in human sera of neutralizing antibodies directed against vector derived from chimpanzee adenoviruses. Only four sera showed a titer over the threshold of 200 on CV32 vector while 8 showed a titer over 200 on ChAd3 SEAP vector. On the contrary, the panel of chimp sera examined showed a very high prevalence of anti-Chimp Ad immunity. These findings confirm that as expected, vectors based on chimp Ads have a very little chance to be neutralized in humans. Therefore they represents an ideal solution to the problem of the pre-existing anti-human Ad immunity that limits the administration of viral vectors based on common human Ad serotypes such as Ad5.

Example 4: ChAd Vector Tropism

[0148] Gene transfer efficacy mediated by Ad5 and ChAd vectors was assessed by EGFP expression on a panel of human primary cells of different histological origin. Human chondrocytes, osteoblasts, keratinocytes, melanocytes, skeletal muscle cells and melanocytes were cultivated according to manufacturer indication. Human monocytes, immature and mature dendritic cells (DC) were obtained as described (Romani, N. et al. 1996, J. Immunol. Methods, 196,137.). Transduced, fluorescent cells were detected by FACS analysis. The panel of human primary cells tested includes cells that are important target cells for different therapeutic strategies based on *in vivo* as well as *ex vivo* gene transfer in the field of cardiovascular disease, rheumatoid arthritis, tissue engineering (bone, skin, and cartilage), and vaccination. The results presented in Figure 38A-D suggests that different chimp adenoviruses can recognize receptors alternative to CAR as demonstrated by the differential efficiency of infection of the different cell types.

MURINE IMMUNIZATION STUDIES

METHODS AND MATERIALS

Immunization Protocols and Splenocyte /PBMC Preparation

[0149] Immunizations: Mice were immunized with the selected adenoviruses diluted in 0.1 ml of buffer. Each vector dose was divided in two aliquot of 50 μl and injected in both quadriceps of mice.

[0150] Splenocyte Preparation: Mice were sacrificed 3 weeks post-injection and their spleens excised and transferred in 10 ml of R10 (10% FCS, 55mM 2-mercaptoethanol, 1M HEPES buffer, 2mM L-glutamine, 1X penicillin-streptomicine solution in RPMI medium 1640). Spleens were minced through a steel screen and, after the screen was washed with 2 ml of R10, splenocytes were transferred in a 50 ml Falcon tube and centrifuged at 1200 rpm, 10 min, room temperature (rt). Supernatant was removed and 3 ml of ACK lysis buffer (Gibco BRL Formulation#79-0422DG) were added. Cells were incubated 5 min, rt. 45 ml of 1X PBS were added and tubes were centrifuged as above. After washing with 30 ml of R10, cells were resuspended in 5 ml of R10, filtered through a 70 m Nylon cell strainer (Falcon 2350). 10μl of cells were diluted with 990 μl Turk's solution (Merck 040417345) and counted. Cells were finally diluted to 10^7 cells /ml in R^{10}.

[0151] Peripheral blood mononuclear cell (PBMC) preparation: Mice blood samples (150 ul) were transferred to 2ml eppendorf tubes with 50ul PBS/2% EDTA. 1 ml ACK buffer was added to each tube. Gently mixed and incubated at RT for 5 min. Samples were centrifuged at 1500rpm in microcentrifuge for 5 min. Supernatant was discharged white cell pellets deriving from the same immunized cohorts were combined. ACK buffer incubation was repeated then pellets of PBMC were resuspended in 1 ml of R10 medium.

IFN-γ ELISPOT Assay

[0152] Millipore MAIP 45 plates were coated with 100 μl/well of purified rat anti-mouse IFN-y monoclonal antibody (Pharmingen, cat. 551216) diluted at 2.5 μg/ml in PBS and incubated over-night (o/n) at 4°C. Plates were washed 2X

with sterile PBS and un-specific binding sites were blocked by incubation for 2hrs in the CO_2 incubator with 200 μl/well of R10. In the immunization experiments with Ad vectors expressing HIV gag, a 9-mer peptide (AMQMLKETI, a CD8 HIV gag epitope mapped in Balb/C mice) (SEQ ID NO: 47) was diluted to 2 μg/ml in R10 and added to the wells in the amount of 50 μl/well. In immunization experiments conducted with HCV-NS expressing vectors, a pool of peptides covering NS3 helicase domain as well a 9-mer peptide representing a mapped CD8 epitope comprised in helicase domain were used. Immunization experiments with ChAds expressing human CEA antigen were evaluated by pools of overlapping 15-mer peptides covering the entire amino acid sequence. As controls DMSO and Concanavalin A were used. Cells were added to each well at the amount of $5X10^5$ and $2.5X10^5$. After an o/n incubation in the CO_2 incubator, plates were washed with 0.05% Tween 20/PBS and 50 μl/ well of biotinylated rat anti-mouse IFN-γ monoclonal antibody (PharMingen cat. 554410) diluted 1/250 in assay buffer (5%FBS, 0.005% Tween20, PBS) were added. Plates were incubated o/n at 4°C and washed as above. Streptavidin-alkaline phosphatase conjugate (BD554065) was diluted 1 /2500 in assay buffer and added in the amount of 50 μl/well for 2 hrs rt. After washing, plates were developed adding 50 μl/well of BCIP/NBT1-step solution (Pierce 34042). Reaction was stopped by washing wells with deionized water. Spots were automatically counted by an ELISPOT reader.

Murine IFN-γ Intracellular Staining (ICS)

[0153] Splenocytes were diluted at $2X10^6$ cells in 1 ml of R10 and stimulated with the same antigens described above at the concentration of 2 μg/ml. As controls, DMSO and Staphylococcal Enterotoxin B (SEB) were used. After an overnight incubation in the CO_2 incubator, cells were washed with FACS buffer (1% FCS, 0.01% NaN3, PBS) and purified anti-mouse CD16/CD32 Fc block (clone 2.4G2, Pharmingen cat. 553142) was diluted 1/25, added in the amount of 100 μl/sample and incubated for 15min at 4°C. Cells were washed in FACS buffer and APC conjugated anti-mouse CD3e (clone 145-2C11, Pharmingen #553066), PE conjugated anti-mouse CD4 (clone L3T4, BD Pharmingen cat. 553142) and PerCP conjugated anti-mouse CD8a (clone 53-6.7, Pharmingen cat. 553036) diluted 1:50 in FACS buffer were added in the amount of 100 μl/sample. Cells were incubated 30 min rt, washed, fixed and permeabilized (Becton Dickinson, FACS Perm 2) and incubated with FITC conjugated anti-mouse IFN- y □Pharmingen cat.554411) diluted 1:50 in PermWash (100 ul/sample) for 30 min at RT. After washing cells were resuspended in 500 ul 1% formaldehyde/PBS and intracellular cytokine staining (ICS) analyzed on a FACS-Calibur flow cytometer, using CellQuest software (Becton Dickinson).

Example 5: ChAd Vectors Elicit Strong CMI Responses in Mice

[0154] The ability of the ChAd vectors disclosed herein to elicit a cell-mediated immune response (CMI) was evaluated in mice using vectors expressing an HIV gag transgene. Briefly, groups of 5 Balb/C mice were injected with ten-fold increasing doses of the different vectors starting from 10^5 up to 10^{10} vp/mouse.

[0155] The strength of the immune response was determined three weeks after the injection by quantifying gag-specific CD8+ T cells in the splenocytes. The number of IFN-γ secreting CD8+ T cells was determined by ELISPOT assay or by IFN-γ intracellular staining and FACS analysis after stimulation *in vitro* with a peptide reproducing a gag CD8+ T cell epitope mapped in Balb/C mice.

[0156] The results obtained from the 5 immunized animals, reported in Table 3, are expressed as spot forming cells per 10^6 splenocytes. Shown are the number of spot forming cells per million splenocytes following incubation with 9-mer CD8+ gag epitope or with gag peptide pool. The gag peptide pool consisted of 20-aa peptide overlapping by 10aa encompassing the entire gag sequence. Positive values are reported in bold.

[0157] The data provided in Table 3 indicate that the administration of the ChAd vectors disclosed and claimed herein elicits a strong cell mediated immune response which is comparable to the response elicited by hAd5. By looking at the lowest vector dose resulting in a positive immunization result (immunization breakpoint), we ranked the potency of the different vectors being subgroup C ChAd3gag the most potent with a breakpoint at 10^6 pp vector dose. Ranking by immunization break-points is shown in Figure 33.

Table 3. Gag-Specific T Cell Response in Balbc Mice Immunized with Chimpanzee Ad Vectors

Vaccination	10^5 vp		10^6 vp		10^7 vp		10^8 vp		10^9 vp		10^10 vp	
	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag
ChAd3DE1gag	1	neg	1	944	1	**1298**	1	**1258**	NT	NT	NT	NT
	3	neg	1	1039	1	**1958**	1	**1962**	NT	NT	NT	NT
	1	neg	1	859	1	**1923**	1	**1931**	NT	NT	NT	NT
	1	neg	1	1620	1	**1386**	1	**1369**	NT	NT	NT	NT
	1	neg	1	1529	5	**1442**	4	**1436**	NT	NT	NT	NT
CV33DE1gag	NT	NT	1	neg	2	**475**	1	**2910**	NT	NT	NT	NT
	NT	NT	1	neg	1	**433**	1	**401**	NT	NT	NT	NT
	NT	NT	1	neg	1	**243**	1	**634**	NT	NT	NT	NT
	NT	NT	1	neg	1	**505**	2	**3457**	NT	NT	NT	NT
	NT	NT	1	neg	1	**683**	2	**1684**	NT	NT	NT	NT
CV68DE1gag	NT	NT	3	neg	1	**340**	2	**332**	0	**406**	2	**635**
	NT	NT	1	neg	1	**512**	0	**536**	1	**256**	3	**1172**
	NT	NT	0	neg	2	**458**	3	**944**	2	**462**	2	**505**
	NT	NT	7	neg	0	**148**	1	**519**	0	**488**	2	**1184**
	NT	NT	0	neg	2	**1418**	1	**243**	0	**240**	1	**789**
ChAd9DE1gag	NT	NT	1	neg	7	**369**	1	**609**	NT	NT	NT	NT
	NT	NT	1	neg	1	**508**	1	**739**	NT	NT	NT	NT
	NT	NT	1	neg	1	**299**	16	**291**	NT	NT	NT	NT
	NT	NT	1	neg	2	**507**	8	**926**	NT	NT	NT	NT
	NT	NT	0.5	neg	1	**36**	40	**1034**	NT	NT	NT	NT
ChAd10DE1gag	NT	NT	1	neg	1	**83**	1	**822.5**	NT	NT	NT	NT
	NT	NT	1	neg	1	**42.5**	1	**1033**	NT	NT	NT	NT
	NT	NT	1	neg	1	**48**	1	**1339.5**	NT	NT	NT	NT
	NT	NT	1	neg	4	**51**	1	**1132**	NT	NT	NT	NT
	NT	NT	1	neg	1	**466.5**	1	**521.5**	NT	NT	NT	NT
ChAd6DE1gag	NT	NT	1	neg	1	**34**	1	**721**	NT	NT	NT	NT
	NT	NT	1	neg	10	**4**	1	**560**	NT	NT	NT	NT
	NT	NT	1	neg	1	**24**	1	**624**	NT	NT	NT	NT
	NT	NT	1	neg	1	**225**	3	**3002**	NT	NT	NT	NT
	NT	NT	1	neg	1	**276**	4	**1738**	NT	NT	NT	NT

(continued)

Vaccination	10^5 vp		10^6 vp		10^7 vp		10^8 vp		10^9 vp		10^10 vp	
	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag	mock	Gag
ChAd11DE1gag	1	neg	1	neg	0	**573**	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	0	**919**	NT	NT	NT	NT	NT	NT
	0	neg	1	neg	1	**1438**	NT	NT	NT	NT	NT	NT
	2	neg	0	neg	0	**0**	NT	NT	NT	NT	NT	NT
	1	neg	1	neg	0	**456**	NT	NT	NT	NT	NT	NT
ChAd20DE1gag	0	neg	0	neg	0	**1**	NT	NT	NT	NT	NT	NT
	2	neg	0	neg	0	**408**	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	0	**414**	NT	NT	NT	NT	NT	NT
	1	neg	0	neg	0	**2**	NT	NT	NT	NT	NT	NT
	0	neg	0	neg	1	**311**	NT	NT	NT	NT	NT	NT
ChAd7DE1gag	NT	NT	1	neg	1	neg	1	**1044**	NT	NT	NT	NT
	NT	NT	3	neg	1	neg	1	**606**	NT	NT	NT	NT
	NT	NT	1	neg	8	neg	1	**407**	NT	NT	NT	NT
	NT	NT	1	neg	1	neg	2	**567**	NT	NT	NT	NT
	NT	NT	1	neg	3	neg	1	**1677**	NT	NT	NT	NT
CV32DE1gag	NT	NT	NT	NT	1	neg	0	**83**	0	**291**	0	**194**
	NT	NT	NT	NT	3	neg	0	**382**	0	**805**	2	**380**
	NT	NT	NT	NT	0	neg	1	**97**	0	**136**	1	**501**
	NT	NT	NT	NT	1	neg	5	**96**	4	**1162**	0	**1115**
	NT	NT	NT	NT	2	neg	1	**328**	NT	**NT**	0	**596**
ChAd4DE1gag	NT	NT	1	neg	0	neg	0	**0**	NT	NT	NT	NT
	NT	NT	0	neg	0	neg		**159**	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	0	**1**	NT	NT	NT	NT
	NT	NT	1	neg	0	neg	0	**234**	NT	NT	NT	NT
	NT	NT	1	neg	0		1	**0**	NT	NT	NT	NT
ChAd16DE1gag	NT	NT	0	neg	0	neg	0	**243**	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	1	**296**	NT	NT	NT	NT
	NT	NT	0	neg	2	neg	1	**68**	NT	NT	NT	NT
	NT	NT	0	neg	0	neg	1	**433**	NT	NT	NT	NT
	NT	NT	1	neg	0	neg	1	**28**	NT	NT	NT	NT

Example 6: ChAd3 and CV33 GAG Vectors Elicit a CMI Response Characterized by GAG-Specific CD8+ T Cells

[0158] In order to characterize the CMI response elicited in response to the ChAd vectors comprising HIV gag transgene, splenocytes pooled from cohorts of five mice immunized with different doses of vector were analyzed by intracellular IFN-γ staining. The data shown in table 3 and table 4 were collected in separate experiments.

[0159] Splenocytes were diluted at $2X10^6$ cells in 1 ml of R10 and stimulated with the same antigens described above at the concentration of 2 μg/ml. As controls, DMSO and SEB (Staphylococcal Enterotoxin B) were used. After an o/n incubation in the CO_2 incubator, cells were washed with FACS buffer (1% FCS, 0.01% NaN3, PBS) and purified anti-mouse CD16/CD32 Fc block (clone 2.4G2, Pharmingen cat. 553142) was diluted 1/25, added in the amount of 100 μl/sample and incubated for 15min at 4°C. Cells were washed in FACS buffer and APC conjugated anti-mouse CD3e (clone 145-2C11, Pharmingen #553066), PE conjugated anti-mouse CD4 (clone L3T4, BD Pharmingen cat. 553142) and PerCP conjugated anti-mouse CD8a (clone 53-6.7, Pharmingen cat. 553036) diluted 1:50 in FACS buffer were added in the amount of 100 μl/sample. Cells were incubated 30 min rt, washed, fixed and permeabilized (Becton Dickinson, FACS Perm 2) and incubated with FITC conjugated anti-mouse IFN- y □ □Pharmingen cat.554411) diluted 1:50 in PermWash (100 ul/sample) for 30 min at RT. After washing cells were resuspended in 500 ul 1% formaldehyde/PBS and analyzed on a FACS-Calibur flow cytometer, using CellQuest software (Becton Dickinson).

[0160] Table 4 provides data summarizing the percentage of gag-specific CD3+T cells that were either gag-specific CD8+ or CD4+ T cells. Positive results are reported in bold. The data provided herein indicate that the cellular profile of the immune response elicited by ChAd vectors derived from viruses classified into different serotype subgroups (i.e., subgroups C, D and E) are similar and all of the gag-specific responses characterized predominantly by CD8+ T cells. In addition, it is noted that at high vector doses a gag-specific CD4+ response becomes evident in all immunization experiments. The ICS assay confirmed that ChAd3 vector can stimulate anti-gag CD8+ response at 10^6 vector dose.

Table 4. Characterization of Gag-Specific T Cells in Mice Immunized with Chimp Adenovirus Vectors of Different Subgroups

vaccine		10^5		10^6		10^7		10^8		10^9	
		DMSO	gag	DMSO	gag	DMSO	gag	DMSO	gag	DMSO	gag
ChAd3DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.01%	**4.65%**	0.01%	**17.15%**	0.04%	**24.71%**	NT	NT
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.07%	0.03%	0.08%	0.04%	**0.28%**	NT	NT
CV33DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.02%	0.01%	0.01%	**0.83%**	0.03%	**8.69%**	NT	NT
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.00%	0.00%	0.04%	0.01%	**0.10%**	NT	NT
ChAd9DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.02%	0.01%	0.01%	**0.68%**	NT	NT	0.04%	**4.73%**
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.00%	0.00%	0.00%	NT	NT	0.00%	0.01%
ChAd10DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.02%	0.01%	0.01%	**0.57%**	NT	NT	0.02%	**5.04%**
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.00%	0.00%	0.00%	NT	NT	0.00%	0.01%
ChAd6DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.00%	0.01%	0.00%	**0.59%**	0.01%	**14.28%**	NT	NT
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.00%	0.00%	0.05%	0.01%	**0.12%**	NT	NT
ChAd7DE1gag	%CD8$^+$CD3$^+$	NT	NT	0.01%	0.02%	0.01%	0.00%	0.02%	**5.00%**	NT	NT
	%CD4$^+$CD3$^+$	NT	NT	0.00%	0.01%	0.00%	0.00%	0.01%	**0.21%**	NT	NT

Example 7: ChAd Vectors Elicit HCV NS-Specific T Cell Response

[0161] The potency of CV32-NSmut and CV33-NSmut vectors was evaluated in C57/Black6 mice relative to the potency of MRKAd6NSmut. The animals were injected with 10-fold increasing doses of vector starting from 10^7 up to 10^9 vp/mouse. CMI was analyzed 3 weeks after a single injection by IFN-γ ELISPOT and IFN-γ intracellular staining by stimulating T cells with a 9-mer peptide reproducing a CD8+ T cell epitope mapped in the helicase domain of NS3 protein. The data provided in Table 5 summarize the number of spot-forming cells per million splenocytes following incubation in absence (mock) or in presence of NS3 9-mer peptide.

[0162] The data indicate that both CV32 and CV33 vectors expressing HCV-NS stimulate strong T cell responses. Based on the observation that the first positive result for the CV32 vector was obtained by injecting 10^9 vp/dose, the immunization potency of CV32DE1E3 NSmut vector appears to be approximately 100-fold lower than human subgroup C Ad6DE1E3 NSmut vector. The parallel experiment with MRKAd6NSmut indicated that a dose of 10^7 vp/animal was sufficient to stimulate cell mediated immunity. Therefore, these results confirm the lower immunization potency of CV32-derived vectors relative to human subgroup C vectors (such as hAd5 and hAd6) that was also observed in the experiment with gag expressing vectors (see Table 3).

Table 5. HCV NS-Specific T Cell Response in Mice Immunized with Mrkad6 Nsmut, CV32NSmut or CV33NSmut

Vaccination	10^7 vp		10^8 vp		10^9 vp		10^10 vp	
	Mock	NS3	mock	NS3	mock	NS3	mock	NS3
MRKAd6NSmut	1	**345**	1	**449**	NT	NT	NT	NT
	1	**248**	1	**1590**	NT	NT	NT	NT
	1	**1**	1	**549**	NT	NT	NT	NT
	1	**262**			NT	NT	NT	NT
					NT	NT	NT	NT
CV33NSmut	1	1	1	**195**	2	**338**	NT	NT
	1	2	1	**409**	1	**1136**	NT	NT
	1	1	1	**396**	1	**497**	NT	NT
	1	2	2	**172**	1	**344**	NT	NT
	1	**237**			1	**163**	NT	NT
CV32NSmut	neg	neg	1	**181**	1	**118**	1	**176**
	neg	neg	1	**71**	1	**239**	1	**238**
	neg	neg	1	**56**	1	**862**	1	**555**
	neg	neg	1	**459**	1	**219**	1	**545**
	neg	neq	1	**195**	1	**123**	1	**578**

Example 8: Anti -Ad5 Pre-Existing Immunity Does Not Abrogate Anti-GAG CMI Elicited by ChAd3gag

[0163] To evaluate the impact on ChAd3 immunization of the pre-existing immunity against the high seroprevalent Ad5, 4 cohorts of 5 BalbC mice were pre-immunized with two injection of 10^{10} vp of Ad5 wt in the quadriceps at week 0 and 2. As control, 2 cohorts of 5 mice were injected at the same time points with buffer only. Cohorts of Ad5 pre-immunized mice were then immunized with 10^6 and 10^7 vp/mouse of either Ad5gag or ChAd3gag vectors. Cohorts of control (naïve) mice were immunized with 106 vp/mouse of Ad5gag or ChAd3gag vectors.

[0164] Anti-Ad5 and ChAd3 neutralizing immunity was evaluated at week 4 by the neutralization assay described above using Ad5 and ChAd3 SEAP vectors. Anti-gag immunity was evaluated by ELISPOT analysis on purified splenocytes stimulated with gag 9-mer peptide containing a gag epitope mapped in BalbC mice. The results reported in figure 36 demonstrated that Anti-Ad5 immunity does not abrogate anti-gag CMI elicited by ChAd3gag while, as expected, anti-Ad5 immunity completely block Ad5gag immunization.

Example 9: ChAd3hCEA Immunization Elicits a Strong CEA-Specific Immune Response in Transgenic Mice Expressing Human CEA

[0165] The ability of the ChAd vectors disclosed and claimed herein to elicit an immune response against a self-antigen therefore breaking the tolerance was also evaluated in transgenic mice expressing human CEA (Clarke, P et al. Cancer Res. (1998) 58(7):1469-77.)

[0166] Cohorts of 8 mice were injected in the quadriceps with 10^10 vp of ChAd3hCEA or Ad5hCEA as already described. The immune response against CEA was followed weekly up to day 75 on PBMC stimulated with a pool of 15-mer peptides encompassing human CEA amino acid sequence from aa 497 to the end (aa 703). Anti-CEA immunity was evaluated by ICS determining CD4-CD8+ T cells secreting interferon-γ in response to CEA peptide pool incubation.

[0167] The results reported in figure 37 demonstrate that ChAd3hCEA vector immunization stimulate a more sustained CD8+ T cell response against human CEA than Ad5 expressing the same transgene.

PRIMATE IMMUNIZATION STUDIES

METHODS AND MATERIALS

Immunization Protocol

[0168] The ability of the ChAd vectors disclosed and claimed herein to elicit CMI in Rhesus macaques (referred to herein as monkeys) was also evaluated. The macaques were anesthetized (ketamine/xylazine) and the vaccines were delivered i.m. in 0.5-mL aliquots into both deltoid muscles using tuberculin syringes (Becton-Dickinson). In all cases the macaques were between 3-10 kg in weight, and the total dose of each vaccine was administered in 1 mL of buffer.

[0169] Sera and peripheral blood mononuclear cells (PBMC) were prepared from blood samples collected at several time points during the immunization regimen. All animal care and treatment were in accordance with standards approved by the Institutional Animal Care and Use Committee according to the principles set forth in the Guide for Care and Use of Laboratory Animals, Institute of Laboratory Animal Resources, National Research Council.

ELISPOT Assay

[0170] The IFN-γ ELISPOT assays for rhesus macaques were conducted following a previously described protocol (Allen et al., 2001 J. Virol. 75(2):738-749), with some modifications. For gag-specific stimulation, a peptide pool was prepared from 20-aa peptides that encompass the entire HIV-1 gag sequence with 10-aa overlaps (Synpep Corp., Dublin, CA). For HCV NS-specific stimulation 6 peptide pools were prepared from 15-aa peptides that encompass the entire HCV-NS sequence from NS3 to NS5b with 10-aa overlaps.
HER2/neu and CEA-specific stimulations were performed with 15-aa peptides that encompass the entire protein sequence with 10-aa overlaps.

[0171] To each well, 50 μL of 2-4 x 10^5 peripheral blood mononuclear cells (PBMCs) were added; the cells were counted using Beckman Coulter Z2 particle analyzer with a lower size cut-off set at 80 fL. Either 50 μL of media or the gag peptide pool at 8 μg/mL concentration per peptide was added to the PBMC. The samples were incubated at 37°C, 5% CO_2 for 20-24 hrs. Spots were developed accordingly and the plates were processed using custom-built imager and automatic counting subroutine based on the ImagePro platform (Silver Spring, MD); the counts were normalized to 10^6 cell input.

Intracellular Cytokine Staining (ICS)

[0172] To 1 ml of 2 x 10^6 PBMC/mL in complete RPMI media (in 17x100mm round bottom polypropylene tubes (Sarstedt, Newton, NC)), anti-hCD28 (clone L293, Becton-Dickinson) and anti-hCD49d (clone L25, Becton-Dickinson) monoclonal antibodies were added to a final concentration of 1 μg/mL. For gag-specific stimulation, 10 μL of the peptide pool (at 0.4 mg/mL per peptide) were added. Similar conditions were used for HCV NS-specific stimulation. The tubes were incubated at 37 °C for 1 hr., after which 20 μL of 5 mg/mL of brefeldin A (Sigma) were added. The cells were incubated for 16 hr at 37 °C, 5% CO_2, 90% humidity. 4 mL cold PBS/2%FBS were added to each tube and the cells were pelleted for 10 min at 1200 rpm. The cells were re-suspended in PBS/2%FBS and stained (30 min, 4 °C) for surface markers using several fluorescent-tagged mAbs: 20 μL per tube anti-hCD3-APC, clone FN-18 (Biosource); 20 μL anti-hCD8-PerCP, clone SK1 (Becton Dickinson, Franklin Lakes, NJ); and 20 μL anti-hCD4-PE, clone SK3 (Becton Dickinson). Sample handling from this stage was conducted in the dark. The cells were washed and incubated in 750 μL 1xFACS Perm buffer (Becton Dickinson) for 10 min at room temperature. The cells were pelleted and re-suspended in PBS/2%FBS and 0.1 μg of FITC-anti-hIFN-γ, clone MD-1 (Biosource) was added. After 30 min incubation, the cells were washed and

resuspended in PBS. Samples were analyzed using all four color channels of the Becton Dickinson FACSCalibur instrument. To analyze the data, the low side- and forward-scatter lymphocyte population was initially gated; a common fluorescence cut-off for cytokine-positive events was used for both $CD4^+$ and $CD8^+$ populations, and for both mock and gag-peptide reaction tubes of a sample.

Example 10: A Homologous Prime-Boost Regimen Using ChAd ΔE1-gag Vectors Elicits Gag-Specific T Cells in Monkeys

[0173] Cohorts of 3 animals were given intramuscular injection at week 0 and week 4 of either of the following constructs: 10^10 vp of CV-32ΔE1-gag; or 10^10 vp CV33ΔE1-gag; or 10^10 vp and 10^8 vp MRKAd5ΔE1gag. PBMCs collected at regular 4-wks intervals were analyzed in an ELISPOT assay. The results provided in Table 6, which indicate the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of Gag peptide pool establish that both CV32ΔE1-gag and CV-33ΔE1gag are able to induce significant levels of gag-specific T cells in non-human primates. It is interesting to note that after a single dose (wk 4), the CV32ΔE1-gag responses were comparable to MRKAd5 ΔE1-gag 10^8 vp dose and lower than that of MRKAd5-gag 10^10 vp/dose. CV33ΔE1-gag 10^10 vp/dose induces a response comparable to that of MRKAd5-gag 10^10 vp/dose. This result was confirmed at week 8 after the second dose.

Table 6. Gag-Specific T Cell Response in Monkeys Immunized with Mrkad5 ΔE1-Gag, CV32ΔE1-Gag, CV33ΔE1-Gag

Vaccination T=0	vector dose	Monk #	Pre-bleed		T=4		T=8	
			Mock	Gag	Mock	Gag	Mock	Gag
CV32ΔE1gag	10^10 vp	01C023	1	0	14	353	3	278
		01C029	1	3	13	605	3	419
		01C032	1	0	5	274	1	179
CV33ΔE1gag	10^10 vp	01C033	0	0	9	1545	1	659
		01C036	4	5	4	1540	13	881
		01D303	0	3	19	949	10	628
MRKAd5gag	10^8 vp	01D267	0	0	4	473	0	341
		01D279	1	4	44	831	6	336
		01D284	4	5	4	264	5	129
MRKAd5gag	10^10 vp	99C218	0	3	5	2500	0	1580
		99C227	6	1	4	529	5	365
		99D185	ND	ND	0	425	0	310

Example 11: ChAd Vectors Elicit a HCV NS-Specific T-Cell Response in a Heterologous Prime-Boost Regimen

[0174] In a separate experiment, groups of two and three monkeys were given immunization at week 0, 4 of MRK Ad6NSoptmut vector at 10^8 or 10^10 vp per animal. The animals were boosted with the same virus at the same dose at week 24 and then boosted again at week 104 with CV33-NSmut at 10^10 vp per animal. The results are presented in Tables 7 and 8 which summarize the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of HCV NS peptide pool.

[0175] T cell immunity, as assessed by IFN-γ ELISPOT, showed a peak response at week 4 after the first dose in the animals injected with 10^10 vp (Table 8) and at week 8 (post-dose 2) in the animals injected at 10^8 (Table 7). The response was not boosted by the injection at week 24 ("homologous boost"), while a strong boost effect was observed after the injection with CV33-NSmut ("heterologous boost").

Table 7. HCV NS-Specific T Cell Response in Monkeys Immunized with MRK Ad6NSoptmut At 10^8 Vp/Animal and Boosted with CV33-Nsmut

Vaccine	MRKAd6NSoptmut 10^8 vp						CV33-NSmut10^10 vp			
time point	post-priming I dose T=4		post-priming II dose T=8		pre-homologous boost T=24		post-homologous boost T=28		pre-heterologous boost T=104		post-heterologous boost T=108	
monkey	*95116*	*138T*	*95116*	*138T*	*95116*	*138T*	*95116*	*138T*	*95116*	*138T*	*95116*	*138T*
poolF	44	112	77	**124**	115	176	105	55	120	150	**188**	**2228**
pool G	20	**2110**	86	**1975**	201	**1105**	94	**884**	120	**192**	96	**4590**
poolH	12	18	54	22	169	221	28	9	81	33	**447**	**543**
poolI	14	53	62	47	163	189	96	18	80	67	71	**515**
poolL	33	**86**	58	44	353	**608**	**235**	33	110	131	**224**	**308**
poolM	**184**	75	**168**	**138**	204	**336**	67	44	55	46	**2028**	**1570**
DMSO	14	3	44	7	104	79	33	6	57	40	33	65

Table 8. HCV NS-Specific T Cell Response In Monkeys Immunized MRK Ad6NSoptmut At And 10^10 vp/Animal And Boosted With CV33-Nsmut

Vaccine	**MRKAd6NSoptmut 10^10 vp**												**CV33-NSmut 10^10**					
time point	**post-priming dose** T=4			**post-priming II dose** T=8			**pre-homologous boost** T=24			**post-homologous boost** T=28			**pre-heterologous boost** T=104			**post-heterologous boost** T=108		
monkey	***98D209***	***106Q***	***113Q***	***98D209***	***106Q***	***113Q***	***98D209***	***106Q***	***113Q***	***98D209***	***106Q***	***113Q***	***98D209***	***106Q***	***113Q***	***98D209***	***106Q***	***113Q***
poolF	**3110**	**263**	**404**	**1340**	**300**	**723**	**678**	**61**	**583**	**321**	**123**	**1438**	**204**	192	**326**	**1581**	**1525**	**1714**
poolG	**2115**	**642**	**1008**	**1070**	**316**	**2205**	**685**	**71**	**701**	**251**	**178**	**1758**	**166**	106	**625**	**1118**	**524**	**4238**
poolH	**373**	**72**	19	**358**	**43**	43	424	24	42	51	23	18	**92**	45	55	**413**	58	**211**
poolI	**103**	37	**347**	**80**	36	**531**	237	39	**169**	12	35	**485**	66	79	**376**	**459**	**85**	**2738**
poolL	**149**	22	10	**93**	36	29	279	46	48	11	49	51	**89**	109	**73**	**199**	**76**	**431**
poolM	**314**	**428**	19	**153**	**243**	20	333	**81**	38	38	**134**	11	41	81	9	**228**	**1440**	**227**
DMSO	0	1	3	16	16	5	128	8	9	8	10	16	20	51	12	18	13	5

[0176] The efficiency of heterologous boost with chimp Ad vectors was evaluated in a second experiment. Cohorts of three monkeys were immunized at week 0 and week 4 with MRKAd5gag (10^10 vp/animal), MRKAd6NSmut (10^10 vp/animal) or with the combination of both vectors (10^10 vp/animal each vector) then boosted with the same immunogen at week 24 (homologous boost). Homologous boost was performed with the same immunogens; heterologous boost was performed with CV33gag, CV32 NSmut or with the two vectors in combination. The results provided in Table 9 summarize the number of spot-forming cells per million PBMC following incubation in absence (mock) or presence of HCV NS peptide pool.

The same cohorts were boosted again at week 51 with CV33gag (10^10 vp/animal), CV32NSmut (10^10 vp/animal) and with the combination of the two vectors (10^10 vp/animal each vector). The results provided in Table 9 further indicate that the homologous boost was not efficient since the responses are below the peak observed at week 4 after the injection of the first dose of vaccine. A strong boosting effect was measured by IFN-γ ELISPOT at week 54 after immunization with heterologous chimp vectors.

Table 9. Immunization with Chimp Ad vectors efficiently boost Gag and HCV NS-specific T cell response in monkeys immunized with MRK Ad5gag or MRK Ad6NSoptmut at 10^10 vp/animal

Vaccine	MRKAd5gag												CV33gag					
time point	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299	00D105	00D076	00D299
poolF	18	35	60	16	29	14	37	76	40	37	8	14	37	27	44	43	44	70
poolG	16	23	49	4	28	31	54	95	106	81	2	46	36	27	37	84	**108**	109
poolH	45	51	57	18	31	42	55	88	55	47	11	32	69	36	60	85	58	120
poolI	21	21	48	4	26	11	19	54	26	38	6	6	22	11	32	33	26	24
poolL	15	21	58	9	31	20	71	183	128	**106**	6	27	61	21	65	28	45	44
poolM	39	24	49	26	14	49	38	93	39	59	6	19	62	23	38	27	19	14
Gag	**1764**	**2208**	**2762**	**574**	**1906**	**1959**	**391**	**935**	**702**	**2123**	**336**	**736**	**485**	**833**	**1384**	**4003**	**4333**	**3863**
DMSO	9	13	37	7	14	13	16	76	33	26	3	11	28	19	39	23	16	53

Vaccine	MRK Ad5gag + MRKAd6NSmut												CV33gag + CV32NSmut					
time point	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240	00D088	00D099	00D240
poolF	**438**	**118**	**105**	**720**	**116**	**154**	**206**	**108**	**242**	**408**	**99**	**219**	**778**	**135**	**56**	**1701**	**1121**	**424**
poolG	21	**784**	**1483**	44	**362**	**940**	19	**234**	**548**	47	**781**	**844**	**78**	**363**	**265**	**228**	**3180**	**2770**
poolH	24	53	8	46	27	19	13	66	93	49	41	**87**	**115**	50	28	**97**	**291**	**104**
poolI	**83**	28	9	**90**	24	8	**16**	40	68	33	16	42	**56**	19	8	**165**	**145**	22
poolL	13	14	13	16	17	9	28	101	140	39	27	78	**59**	28	15	**137**	**815**	**463**
poolM	39	31	6	**101**	27	16	**21**	73	107	44	26	78	**114**	28	10	**219**	**109**	21
Gag	**2138**	**1044**	**1063**	**2260**	**505**	**819**	**454**	**241**	**456**	**1100**	**368**	**716**	**1542**	**237**	**161**	**4460**	**2908**	**1764**
DMSO	5	6	3	8	5	1	10	18	43	9	13	28	14	18	12	9	21	6

Vaccine	MRKAd6 NSmut												CV32NSmut					
time point	post-dose 1 (T=4)			post-dose 2 (T=8)			pre-homol. boost (T=24)			post-homol. boost (T=28			pre-heterol. boost (T=51)			post-heterol. boost (T=54)		
animal ID	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159	00D065	00D116	00D159
poolF	**139**	44	**82**	**92**	**121**	**63**	**62**	**116**	**54**	44	42	23	57	**85**	53	**313**	**385**	**261**
poolG	**154**	**253**	**119**	**77**	**156**	**108**	**93**	**165**	**126**	**104**	**59**	39	44	**198**	48	**196**	**764**	**559**
poolH	**1284**	41	**211**	**768**	35	**124**	**394**	84	**77**	24	**817**	48	**624**	31	**116**	**3758**	**90**	**925**
poolI	**302**	22	**1174**	**221**	16	**1069**	**134**	31	**561**	18	**133**	**478**	**84**	16	**362**	**485**	51	**2951**
poolL	28	16	48	35	32	21	141	113	78	19	48	17	46	33	46	**379**	**339**	**541**
poolM	**1329**	**1007**	36	**579**	**392**	30	**314**	**293**	43	**558**	**398**	22	**159**	**369**	33	**1278**	**1750**	16
Gag	15	9	7	13	5	2	36	33	36	9	23	14	16	8	10	37	9	26
DMSO	16	4	5	9	6	4	23	17	8	1	9	3	23	8	6	26	9	10

Example 12: Vaccination with a ChAd Vector Comprising a TAA Breaks Tolerance and Elicits a TAA-Specific T Cell Response in Monkeys

[0177] Experiments designed to determine whether chimpanzee adenoviral vectors are sufficiently immunogenic to break the tolerance to a self-antigen and to document the utility of chimpanzee vectors for boosting an immune response primed with a human adenoviral vector were performed in cohorts of four monkeys. Animals were immunized with three injection at week 0, 2 and 4 of Ad5DE1 RhCEA (10^11 vp), comprising the tumor associated antigen CEA, followed by vaccination at week 16, 18 and 20 with CV33DE1 RhCEA (10^11 vp). T cell response was measured by IFNγ ELISPOT with rhesus CEA peptides.

[0178] The results reported in figure 34, which provide the number of spot-forming cells per million PBMC following

incubation in absence (DMSO) or in presence of rhesus CEA C and D peptides pools, establish that an immunization protocol based on vaccination with two different Ad serotypes leads to a sustained T cell response against CEA in non-human primates.

SEQUENCE LISTING

[0179]

<110> Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A

<120> CHIMPANZEE ADENOVIRUS VACCINE CARRIERS

<130> ITR0048Y PCT

<150> 60/538,799
<151> 2004-01-23

<160> 125

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 37741
<212> DNA
<213> Chimpanzee Adenovirus- ChAd3 Genomic

<400> 1

```
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg 60
cgaggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcggtgtg 120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag 180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttcccgc ggtttttacc 240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact 300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta 360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat 420
ttccgcgttc cgggtcaaag tctccgtttt attattatag tcagctgacg cggagtgtat 480
ttataccctc tgatctcgtc aagaggccac tcttgagtgc cagcgagtag agttttctcc 540
tctgccgctc tccgctccgc tccgctcggc tctgacaccg gggaaaaaat gagacatttc 600
acctacgatg gcggtgtgct caccggccag ctggctgctg aggtcctgga caccctgatc 660
gaggaggtat tggccgataa ttatcctccc tcgactcctt ttgagccacc tacacttcac 720
gaactatacg atctggatgt ggtggggccc agcgatccga acgagcaggc ggtttccagt 780
ttttttccag agtccatgtt gttggccagc caggaggggg tcgaacttga gacccctcct 840
ccgatcgtgg attcccccga tccgccgcag ctgactaggc agcccgagcg ctgtgcggga 900
cctgagacta tgccccagct gctacctgag gtgatcgatc tcacctgtaa tgagtctggt 960
tttccaccca gcgaggatga ggacgaagag ggtgagcagt ttgtgttaga ttctgtggaa 1020
caacccgggc gaggatgcag gtcttgtcaa tatcaccgga aaaacacagg agactcccag 1080
attatgtgtt ctctgtgtta tatgaagatg acctgtatgt ttatttacag taagtttatc 1140
atcggtgggc aggtgggcta tagtgtgggt ggtggtcttt ggggggtttt ttaatatatg 1200
tcaggggtta tgctgaagac ttttttattg tgatttttaa aggtccagtg tctgagcccg 1260
agcaagaacc tgaaccggag cctgagcctt ctcgccccag gagaaagcct gtaatcttaa 1320
ctagacccag cgcaccggta gcgagaggcc tcagcagcgc ggagaccacc gactccggtg 1380
cttcctcatc acccccggag attcaccccc tggtgccccc atgtcccgtt aagcccgttg 1440
ccgtgagagt cagtgggcgg cggtctgctg tggagtgcat tgaggacttg ctttttgatt 1500
cacaggaacc tttggacttg agcttgaaac gccccaggca ttaaacctgg tcacctggac 1560
tgaatgagtt gacgcctatg tttgcttttg aatgacttaa tgtgtataga taataaagag 1620
tgagataatg ttttaattgc atggtgtgtt taacttgggc ggagtctgct gggtatataa 1680
gcttccctgg gctaaacttg gttacacttg acctcatgga ggcctgggag tgtttggaga 1740
actttgccgg agttcgtgcc ttgctggacg agagctctaa caatacctct tggtggtgga 1800
ggtatttgtg gggctctccc cagggcaagt tagtttgtag aatcaaggag gattacaagt 1860
gggaatttga agagcttttg aaatcctgtg gtgagctatt ggattctttg aatctaggcc 1920
```

accaggctct cttccaggag aaggtcatca ggactttgga tttttccaca ccggggcgca 1980
ttgcagccgc ggttgctttt ctagcttttt tgaaggatag atggagcgaa gagacccact 2040
tgagttcggg ctacgtcctg gattttctgg ccatgcaact gtggagagca tggatcagac 2100
acaagaacag gctgcaactg ttgtcttccg tccgcccgtt gctgattccg gcggaggagc 2160
aacaggccgg gtcagaggac cgggcccgtc gggatccgga ggagagggca ccgaggccgg 2220
gcgagaggag cgcgctgaac ctgggaaccg ggctgagcgg ccatccacat cgggagtgaa 2280
tgtcgggcag gtggtggatc tttttccaga actgcggcgg attttgacta ttagggagga 2340
tgggcaattt gttaagggtc ttaagaggga gaggggggct tctgagcata acgaggaggc 2400
cagtaattta gcttttagct tgatgaccag acaccgtcca gagtgcatca cttttcagca 2460
gattaaggac aattgtgcca atgagttgga tctgttgggt cagaagtata gcatagagca 2520
gctgaccact tactggctgc agccgggtga tgatctggag gaagctatta gggtgtatgc 2580
taaggtggcc ctgcggcccg attgcaagta caagctcaag gggctggtga atatcaggaa 2640
ttgttgctac atttctggca acggggcgga ggtggagata gagaccgaag acagggtggc 2700
tttcagatgc agcatgatga atatgtggcc gggggtgctg ggcatggacg gggtggtgat 2760
tatgaatgtg aggttcacgg ggcccaactt taacggcacg gtgtttttgg ggaacaccaa 2820
cctggtcctg cacggggtga gcttctatgg gtttaacaac acctgtgtgg aggcctggac 2880
cgatgtgaag gtccgcggtt gcgcctttta tggatgttgg aaggccatag tgagccgccc 2940
taagagcagg agttccatta agaaatgctt gtttgagagg tgcaccttgg ggatcctggc 3000
cgagggcaac tgcagggtgc gccacaatgt ggcctccgag tgcggttgct tcatgctagt 3060
caagagcgtg gcggtaatca agcataatat ggtgtgcggc aacagcgagg acaaggcctc 3120
acagatgctg acctgcacgg atggcaactg ccacttgctg aagaccatcc atgtaaccag 3180
ccacagccgg aaggcctggc ccgtgttcga gcacaacttg ctgacccgct gctccttgca 3240
tctgggcaac aggcgggggg tgttcctgcc ctatcaatgc aactttagtc acaccaagat 3300
cttgctagag cccgagagca tgtccaaggt gaacttgaac ggggtgtttg acatgaccat 3360
gaagatctgg aaggtgctga ggtacgacga gaccaggtcc cggtgcagac cctgcgagtg 3420
cgggggcaag catatgagga accagcccgt gatgctggat gtgaccgagg agctgaggac 3480
agaccacttg gttctggcct gcaccagggc cgagtttggt tctagcgatg aagacacaga 3540
ttgaggtggg tgagtgggcg tggcctgggg tggtcatgaa aatatataag ttgggggtct 3600
tagggtctct ttatttgtgt tgcagagacc gccggagcca tgagcgggag cagcagcagc 3660
agcagtagca gcagcgcctt ggatggcagc atcgtgagcc cttatttgac gacgcggatg 3720
ccccactggg ccggggtgcg tcagaatgtg atgggctcca gcatcgacgg ccgacccgtc 3780
ctgcccgcaa attccgccac gctgacctat gcgaccgtcg cggggacgcc gttggacgcc 3840
accgccgccg ccgccgccac cgcagccgcc tcggccgtgc gcagcctggc cacggacttt 3900
gcattcctgg gaccactggc gacaggggct acttctcggg ccgctgctgc cgccgttcgc 3960
gatgacaagc tgaccgccct gctggcgcag ttggatgcgc ttactcggga actgggtgac 4020
ctttctcagc aggtcatggc cctgcgccag caggtctcct ccctgcaagc tggcgggaat 4080
gcttctccca caaatgccgt ttaagataaa taaaaccaga ctctgtttgg attaaagaaa 4140
agtagcaagt gcattgctct ctttatttca taattttccg cgcgcgatag gccctagacc 4200
agcgttctcg gtcgttgagg gtgcggtgta tcttctccag gacgtggtag aggtggctct 4260
ggacgttgag atacatgggc atgagcccgt cccgggggtg gaggtagcac cactgcagag 4320
cttcatgctc cggggtggtg ttgtagatga tccagtcgta gcaggagcgc tgggcatggt 4380
gcctaaaaat gtccttcagc agcaggccga tggccagggg gaggcccttg gtgtaagtgt 4440
ttacaaaacg gttaagttgg gaagggtgca ttcggggaga gatgatgtgc atcttggact 4500
gtatttttag attggcgatg tttccgccca gatcccttct gggattcatg ttgtgcagga 4560
ccaccagtac agtgtatccg gtgcacttgg ggaatttgtc atgcagctta gagggaaaag 4620
cgtggaagaa cttggagacg cccttgtggc ctcccagatt ttccatgcat tcgtccatga 4680
tgatggcaat gggcccgcgg gaggcagctt gggcaaagat atttctgggg tcgctgacgt 4740
cgtagttgtg ttccagggtg aggtcgtcat aggccatttt tacaaagcgc gggcggaggg 4800
tgcccgactg ggggatgatg gtcccctctg gccctggggc gtagttgccc tcgcagatct 4860
gcatttccca ggccttaatc tcggaggggg gaatcatatc cacctgcggg gcgatgaaga 4920
aaacggtttc cggagccggg gagattaact gggatgagag caggtttcta agcagctgtg 4980
attttccaca accggtgggc ccataaataa cacctataac cggttgcagc tggtagttta 5040
gagagctgca gctgccgtcg tcccggagga ggggggccac ctcgttgagc atgtccctga 5100
cgcgcatgtt ctccccgacc agatccgcca gaaggcgctc gccgcccagg gacagcagct 5160
cttgcaagga agcaaagttt ttcagcggct tgaggccgtc cgccgtgggc atgtttttca 5220

gggtctggct cagcagctcc aggcggtccc agagctcggt gacgtgctct acggcatctc 5280
tatccagcat atctcctcgt ttcgcgggtt ggggcgactt tcgctgtagg gcaccaagcg 5340
gtggtcgtcc agcggggcca aagtcatgtc cttccatggg cgcagggtcc tcgtcagggt 5400
ggtctgggtc acggtgaagg ggtgcgctcc gggctgagcg cttgccaagg tgcgcttgag 5460
gctggttctg ctggtgctga agcgctgccg gtcttcgccc tgcgcgtcgg ccaggtagca 5520
tttgaccatg gtgtcatagt ccagcccctc cgcggcgtgt cccttggcgc gcagcttgcc 5580
cttggaggtg gcgccgcacg aggggcagag caggctcttg agcgcgtaga gcttgggggc 5640
gaggaagacc gattcggggg agtaggcgtc cgcgccgcag accccgcaca cggtctcgca 5700
ctccaccagc caggtgagct cggggcgcgc cgggtcaaaa accaggtttc ccccatgctt 5760
tttgatgcgt ttcttacctc gggtctccat gaggtggtgt ccccgctcgg tgacgaagag 5820
gctgtccgtg tctccgtaga ccgacttgag gggtcttttc tccaggggggg tccctcggtc 5880
ttcctcgtag aggaactcgg accactctga gacgaaggcc cgcgtccagg ccaggacgaa 5940
ggaggctatg tgggaggggt agcggtcgtt gtccactagg gggtccacct tctccaaggt 6000
gtgaagacac atgtcgcctt cctcggcgtc caggaaggtg attggcttgt aggtgtaggc 6060
cacgtgaccg ggggttcctg acgggggggt ataaaagggg gtgggggcgc gctcgtcgtc 6120
actctcttcc gcatcgctgt ctgcgagggc cagctgctgg ggtgagtatt ccctctcgaa 6180
ggcgggcatg acctccgcgc tgaggttgtc agtttccaaa aacgaggagg atttgatgtt 6240
cacctgtccc gaggtgatac ctttgagggt acccgcgtcc atctggtcag aaaacacgat 6300
ctttttattg tccagcttgg tggcgaacga cccgtagagg gcgttggaga gcagcttggc 6360
gatggagcgc agggtctggt tcttgtccct gtcggcgcgc tccttggccg cgatgttgag 6420
ctgcacgtac tcgcgcgcga cgcagcgcca ctcggggaag acggtggtgc gctcgtcggg 6480
caccaggcgc acgcgccagc cgcggttgtg cagggtgacc aggtccacgc tggtggcgac 6540
ctcgccgcgc aggcgctcgt tggtccagca gagacggccg cccttgcgcg agcagaaggg 6600
gggcaggggg tcgagctggg tctcgtccgg ggggtccgcg tccacggtga aaaccccggg 6660
gcgcaggcgc gcgtcgaagt agtctatctt gcaaccttgc atgtccagcg cctgctgcca 6720
gtcgcgggcg gcgagcgcgc gctcgtaggg gttgagcggc gggccccagg gcatggggtg 6780
ggtgagtgcg gaggcgtaca tgccgcagat gtcatagacg tagaggggct cccgcaggac 6840
cccgatgtag gtggggtagc agcggccgcc gcggatgctg gcgcgcacgt agtcatacag 6900
ctcgtgcgag ggggcgagga ggtcggggcc caggttggtg cgggcggggc gctccgcgcg 6960
gaagacgatc tgcctgaaga tggcatgcga gttggaagag atggtggggc gctggaagac 7020
gttgaagctg gcgtcctgca ggccgacggc gtcgcgcacg aaggaggcgt aggagtcgcg 7080
cagcttgtgt accagctcgg cggtgacctg cacgtcgagc gcgcagtagt cgagggtctc 7140
gcggatgatg tcatatttag cctgcccctt ctttttccac agctcgcggt tgaggacaaa 7200
ctcttcgcgg tctttccagt actcttggat cgggaaaccg tccggttccg aacggtaaga 7260
gcctagcatg tagaactggt tgacggcctg gtaggcgcag cagcccttct ccacggggag 7320
ggcgtaggcc tgcgcggcct tgcggagcga ggtgtgggtc agggcgaagg tgtccctgac 7380
catgactttg aggtactggt gcttgaagtc ggagtcgtcg cagccgcccc gctcccagag 7440
cgagaagtcg gtgcgcttct tggagcgggg gttgggcaga gcgaaggtga catcgttgaa 7500
gaggattttg cccgcgcggg gcatgaagtt gcgggtgatg cggaagggcc ccggcacttc 7560
agagcggttg ttgatgacct gggcggcgag cacgatctcg tcgaagccgt tgatgttgtg 7620
gcccacgatg tagagttcca ggaagcgggg ccggcccttt acggtgggca gcttctttag 7680
ctcttcgtag gtgagctcct cgggcgaggc gaggccgtgc tcggccaggg cccagtccgc 7740
gaggtgcggg ttgtctctga ggaaggactc ccagaggtcg cgggccagga gggtctgcag 7800
gcggtccctg aaggtcctga actggcggcc cacggccatt ttttcggggg tgatgcagta 7860
gaaggtgagg gggtcttgct gccagcggtc ccagtcgagc tgcagggcga ggtcgcgcgc 7920
ggcggtgacc aggcgctcgt cgcccccgaa tttcatgacc agcatgaagg gcacgagctg 7980
ctttccgaag gcccccatcc aagtgtaggt ctctacatcg taggtgacaa agaggcgctc 8040
cgtgcgagga tgcgagccga tcgggaagaa ctggatctcc cgccaccagt tggaggagtg 8100
gctgttgatg tggtggaagt agaagtcccg tcgccgggcc gaacactcgt gctggctttt 8160
gtaaaagcga gcgcagtact ggcagcgctg cacgggctgt acctcctgca cgagatgcac 8220
ctttcgcccg cgcacgagga agccgagggg aaatctgagc cccccgcctg gctcgcggca 8280
tggctggtgc tcttctactt tggatgcgtg tccgtctccg tctggctcct cgaggggtgt 8340
tacggtggag cggaccacca cgccgcgcga gccgcaggtc cagatatcgg cgcgcggcgg 8400
tcggagtttg atgacgacat cgcgcagctg ggagctgtcc atggtctgga gctcccgcgg 8460
cggcggcagg tcagccggga gttcttgcag gttcacctcg cagagtcggg ccagggcgcg 8520

gggcaggtct aggtggtacc tgatctctag gggcgtgttg gtggcggcgt cgatggcttg 8580
caggagcccg catccccggg gggcgacgac ggtgccccgc ggggtggtgg tggtggtggt 8640
ggtggtggtg gtggcggtgc agctcagaag cggtgccgcg ggcgggcccc cggaggtagg 8700
gggggctccg gtcccgccgg caggggcggc agcggcacgt cggcgtggag cgcgggcagg 8760
agttggtgct gtgcccggag gttgctggcg aaggcgacga cgcggcggtt gatctcctgg 8820
atctggcgcc tctgcgtgaa gacgacgggc ccggtgagct tgaacctgaa agagagttcg 8880
acagaatcaa tctcggtgtc attgaccgcg gcctggcgca ggatctcctg cacgtctccc 8940
gagttgtctt ggtaggcgat ctcggccatg aactgctcga tctcttcctc ctggaggtct 9000
ccgcgtccgg cgcgttccac ggtggccgcc aggtcgttgg agatgcgccc catgagctgc 9060
gagaaggcgt tgagtccgcc ctcgttccag actcggctgt agaccacgcc cccctggtca 9120
tcgcgggcgc gcatgaccac ctgcgcgagg ttgagctcca cgtgccgcgc gaagacggcg 9180
tagttgcgca gacgctggaa gaggtagttg agggtggtgg cggtgtgctc ggccacgaag 9240
aagttcatga cccagcggcg caacgtggat tcgttgatgt cccccaaggc ctccagccgt 9300
tccatggcct cgtagaagtc cacggcgaag ttgaaaaact gggagttgcg cgccgacacg 9360
gtcaactcct cctccagaag acggatgagc tcggcgacgg tgtcgcgcac ctcgcgctcg 9420
aaggctatgg ggatctcttc ctccgctagc atcaccacct cctcctcttc ctcctcttct 9480
ggcacttcca tgatggcttc ctcctcttcg gggggcggcg gcggcggcgg tgggggaggg 9540
ggcgctctgc gccggcggcg gcgcaccggg aggcggtcca cgaagcgcgc gatcatctcc 9600
ccgcggcggc ggcgcatggt ctcggtgacg gcgcggccgt tctcccgggg gcgcagttgg 9660
aagacgccgc cggacatctg gtgctggggc gggtggccgt gaggcagcga aacggcgctg 9720
acgatgcatc tcaacaattg ctgcgtaggt acgccgccga gggacctgag ggagtccata 9780
tccaccggat ccgaaaacct ttcgaggaag gcgtctaacc agtcgcagtc gcaaggtagg 9840
ctgagcaccg tggcgggcgg cgggggtgg ggggagtgtc tggcggaggt gctgctgatg 9900
atgtaattga agtaggcgga cttgacacgg cggatggtcg acaggagcac catgtccttg 9960
ggtccggcct gctggatgcg gaggcggtcg gctatgcccc aggcttcgtt ctggcatcgg 10020
cgcaggtcct tgtagtagtc ttgcatgagc ctttccaccg gcacctcttc tccttcctct 10080
tctgcttctt ccatgtctgc ttcggccctg gggcggcgcc gcgccccccc gccccccatg 10140
cgcgtgaccc cgaacccccc gagcggttgg agcagggcca ggtcggcgac gacgcgctcg 10200
gccaggatgg cctgctgcac ctgcgtgagg gtggtttgga agtcatccaa gtccacgaag 10260
cggtggtagg cgcccgtgtt gatggtgtag gtgcagttgg ccatgacgga ccagttgacg 10320
gtctggtggc ccggttgcga catctcggtg tacctgagtc gcgagtaggc gcgggagtcg 10380
aagacgtagt cgttgcaagt ccgcaccagg tactggtagc ccaccaggaa gtgcggcggc 10440
ggctggcggt agaggggcca gcgcagggtg gcgggggctc cgggggccag gtcttccagc 10500
atgaggcggt ggtaggcgta gatgtacctg gacatccagg tgatacccgc ggcggtggtg 10560
gaggcgcgcg ggaagtcgcg cacccggttc cagatgttgc gcaggggcag aaagtgctcc 10620
atggtaggcg tgctctgtcc agtcagacgc gcgcagtcgt tgatactcta gaccagggaa 10680
aacgaaagcc ggtcagcggg cactcttccg tggtctggtg aatagatcgc aagggtatca 10740
tggcggaggg cctcggttcg agccccgggt ccgggccgga cggtccgcca tgatccacgc 10800
ggttaccgcc cgcgtgtcga acccaggtgt gcgacgtcag acaacggtgg agtgttcctt 10860
ttggcgtttt tctggccggg cgccggcgtc gcgtaagaga ctaagccgcg aaagcgaaag 10920
cagtaagtgg ctcgctcccc gtagccggag ggatccttgc taagggttgc gttgcggcga 10980
accccggttc gaatcccgta ctcgggccgg ccggacccgc ggctaaggtg ttggattggc 11040
ctccccctcg tataaagacc ccgcttgcgg attgactccg gacacgggga cgagcccctt 11100
ttatttttgc tttccccaga tgcatccggt gctgcggcag atgcgccccc cgccccagca 11160
gcagcaacaa caccagcaag agcggcagca acagcagcgg gagtcatgca gggccccctc 11220
acccaccctc ggcgggccgg ccacctcggc gtccgcggcc gtgtctggcg cctgcggcgg 11280
cggcgggggg ccggctgacg accccgagga gcccccgcgg cgcagggcca gacactacct 11340
ggacctggag gagggcgagg gcctggcgcg gctgggggcg ccgtctcccg agcgccaccc 11400
gcgggtgcag ctgaagcgcg actcgcgcga ggcgtacgtg cctcggcaga acctgttcag 11460
ggaccgcgcg ggcgaggagc ccgaggagat gcgggacagg aggttcagcg cagggcggga 11520
gctgcggcag gggctgaacc gcgagcggct gctgcgcgag gaggactttg agcccgacgc 11580
gcggacgggg atcagccccg cgcgcgcgca cgtggcggcc gccgacctgg tgacggcgta 11640
cgagcagacg gtgaaccagg agatcaactt ccaaaagagt ttcaacaacc acgtgcgcac 11700
gctggtggcg cgcgaggagg tgaccatcgg gctgatgcac ctgtgggact ttgtaagcgc 11760
gctggtgcag aaccccaaca gcaagcctct gacggcgcag ctgttcctga tagtgcagca 11820

cagcagggac aacgaggcgt ttagggacgc gctgctgaac atcaccgagc ccgagggtcg 11880
gtggctgctg gacctgatta acatcctgca gagcatagtg gtgcaggagc gcagcctgag 11940
cctggccgac aaggtggcgg ccatcaacta ctcgatgctg agcctgggca agttttacgc 12000
gcgcaagatc taccagacgc cgtacgtgcc catagacaag gaggtgaaga tcgacggttt 12060
ttacatgcgc atggcgctga aggtgctcac cctgagcgac gacctgggcg tgtaccgcaa 12120
cgagcgcatc cacaaggccg tgagcgtgag ccggcggcgc gagctgagcg accgcgagct 12180
gatgcacagc ctgcagcggg cgctggcggg cgccggcagc ggcgacaggg aggcggagtc 12240
ctacttcgat gcgggggcgg acctgcgctg ggcgcccagc cggcgggccc tggaggccgc 12300
gggggtccgc gaggactatg acgaggacgg cgaggaggat gaggagtacg agctagagga 12360
gggcgagtac ctggactaaa ccgcgggtgg tgtttccggt agatgcaaga cccgaacgtg 12420
gtggacccgg cgctgcgggc ggctctgcag agccagccgt ccggccttaa ctcctcagac 12480
gactggcgac aggtcatgga ccgcatcatg tcgctgacgg cgcgtaaccc ggacgcgttc 12540
cggcagcagc cgcaggccaa caggctctcc gccatcctgg aggcggtggt gcctgcgcgc 12600
tcgaacccca cgcacgagaa ggtgctggcc atagtgaacg cgctggccga gaacagggcc 12660
atccgcccgg acgaggccgg gctggtgtac gacgcgctgc tgcagcgcgt ggcccgctac 12720
aacagcggca acgtgcagac caacctggac cggctggtgg gggacgtgcg cgaggcggtg 12780
gcgcagcgcg agcgcgcgga tcggcagggc aacctgggct ccatggtggc gctgaatgcc 12840
ttcctgagca cgcagccggc caacgtgccg cgggggcagg aagactacac caactttgtg 12900
agcgcgctgc ggctgatggt gaccgagacc ccccagagcg aggtgtacca gtcgggcccg 12960
gactacttct tccagaccag cagacagggc ctgcagacgg tgaacctgag ccaggctttc 13020
aagaacctgc gggggctgtg gggcgtgaag gcgcccaccg gcgaccgggc gacggtgtcc 13080
agcctgctga cgcccaactc gcgcctgctg ctgctgctga tcgcgccgtt cacggacagc 13140
ggcagcgtgt cccgggacac ctacctgggg cacctgctga ccctgtaccg cgaggccatc 13200
gggcaggcgc aggtggacga gcacaccttc caggagatca ccagcgtgag ccgcgcgctg 13260
gggcaggagg acacgagcag cctggaggcg actctgaact acctgctgac caaccggcgg 13320
cagaagattc cctcgctgca cagcctgacc tccgaggagg agcgcatctt gcgctacgtg 13380
cagcagagcg tgagcctgaa cctgatgcgc gacggggtga cgcccagcgt ggcgctggac 13440
atgaccgcgc gcaacatgga accgggcatg tacgccgcgc accggcctta catcaaccgc 13500
ctgatggact acctgcatcg cgcggcggcc gtgaaccccg agtactttac caacgccatc 13560
ctgaacccgc actggctccc gccgcccggg ttctacagcg ggggcttcga ggtcccggag 13620
gccaacgatg gcttcctgtg ggacgacatg gacgacagcg tgttctcccc gcggccgcag 13680
gcgctggcgg aagcgtccct gctgcgtccc aagaaggagg aggaggaggc gagtcgccgc 13740
cgcggcagca gcggcgtggc ttctctgtcc gagctggggg cggcagccgc cgcgcgcccc 13800
gggtccctgg gcggcagccc ctttccgagc ctggtggggt ctctgcacag cgagcgcacc 13860
acccgccctc ggctgctggg cgaggacgag tacctgaata actccctgct gcagccggtg 13920
cgggagaaaa acctgccccc cgccttcccc aacaacggga tagagagcct ggtggacaag 13980
atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcccgcgct ccggccgccc 14040
acgcggcgcc agcgccacga ccggcagcgg gggctggtgt gggatgacga ggactccgcg 14100
gacgatagca gcgtgctgga cctgggaggg agcggcaacc cgttcgcgca cctgcgcccc 14160
cgcctgggga ggatgtttta aaaaaaaaaa aagcaagaag catgatgcaa aattaaataa 14220
aactcaccaa ggccatggcg accgagcgtt ggtttcttgt gttcccttca gtatgcggcg 14280
cgcggcgatg taccaggagg gacctcctcc ctcttacgag agcgtggtgg gcgcggcggc 14340
ggcggcgccc tcttctccct ttgcgtcgca gctgctggag ccgccgtacg tgcctccgcg 14400
ctacctgcgg cctacggggg ggagaaacag catccgttac tcggagctgg cgcccctgtt 14460
cgacaccacc cgggtgtacc tggtggacaa caagtcggcg gacgtggcct ccctgaacta 14520
ccagaacgac cacagcaatt ttttgaccac ggtcatccag aacaatgact acagcccgag 14580
cgaggccagc acccagacca tcaatctgga tgaccggtcg cactggggcg gcgacctgaa 14640
aaccatcctg cacaccaaca tgcccaacgt gaacgagttc atgttcacca ataagttcaa 14700
ggcgcgggtg atggtgtcgc gctcgcacac caaggaagac cgggtggagc tgaagtacga 14760
gtgggtggag ttcgagctgc cagagggcaa ctactccgag accatgacca ttgacctgat 14820
gaacaacgcg atcgtggagc actatctgaa agtgggcagg caaaacgggg tcctggagag 14880
cgacatcggg gtcaagttcg acaccaggaa cttccgcctg gggctggacc ccgtgaccgg 14940
gctggttatg cccggggtgt acaccaacga ggccttccat cccgacatca tcctgctgcc 15000
cggctgcggg gtggacttca cttacagccg cctgagcaac ctcctgggca tccgcaagcg 15060
gcagcccttc caggagggct tcaggatcac ctacgaggac ctggaggggg gcaacatccc 15120

cgcgctcctc gatgtggagg cctaccagga tagcttgaag gaaaatgagg cgggacagga 15180
ggataccacc cccgccgcct ccgccgccgc cgagcagggc gaggatgctg ctgacaccgc 15240
ggccgcggac ggggcagagg ccgaccccgc tatggtggtg gaggctcccg agcaggagga 15300
ggatatgaat gacagtgcgg tgcgcggaga caccttcgtc acccgggggg aggaaaagca 15360
agcggaggcc gaggccgcgg ccgaggaaaa gcaactggcg gcagcagcgg cggcggcggc 15420
gttggccgcg gcggaggctg agtctgaggg gaccaagccc gccaaggagc ccgtgattaa 15480
gcccctgacc gaagatagca agaagcgcag ttacaacctg ctcaaggaca gcaccaacac 15540
cgcgtaccgc agctggtacc tggcctacaa ctacggcgac ccgtcgacgg gggtgcgctc 15600
ctggaccctg ctgtgcacgc cggacgtgac ctgcggctcg gagcaggtgt actggtcgct 15660
gcccgacatg atgcaagacc ccgtgacctt ccgctccacg cggcaggtca gcaacttccc 15720
ggtggtgggc gccgagctgc tgcccgtgca ctccaagagc ttctacaacg accaggccgt 15780
ctactcccag ctcatccgcc agttcacctc tctgacccac gtgttcaatc gctttcctga 15840
gaaccagatt ctggcgcgcc cgcccgcccc caccatcacc accgtcagtg aaaacgttcc 15900
tgctctcaca gatcacggga cgctaccgct gcgcaacagc atcggaggag tccagcgagt 15960
gaccgttact gacgccagac gccgcacctg cccctacgtt tacaaggcct tgggcatagt 16020
ctcgccgcgc gtcctttcca gccgcacttt ttgagcaaca ccaccatcat gtccatcctg 16080
atctcaccca gcaataactc cggctgggga ctgctgcgcg cgcccagcaa gatgttcgga 16140
ggggcgagga agcgttccga gcagcacccc gtgcgcgtgc gcgggcactt ccgcgccccc 16200
tggggagcgc acaaacgcgg ccgcgcgggg cgcaccaccg tggacgacgc catcgactcg 16260
gtggtggagc aggcgcgcaa ctacaggccc gcggtctcta ccgtggacgc ggccatccag 16320
accgtggtgc ggggcgcgcg gcggtacgcc aagctgaaga gccgccggaa gcgcgtggcc 16380
cgccgccacc gccgccgacc cggggccgcc gccaaacgcg ccgccgcggc cctgcttcgc 16440
cgggccaagc gcacgggccg ccgcgccgcc atgagggccg cgcgccgctt ggccgccggc 16500
atcaccgccg ccaccatggc cccccgtacc cgaagacgcg cggccgccgc cgccgccgcc 16560
gccatcagtg acatggccag caggcgccgg ggcaacgtgt actgggtgcg cgactcggtg 16620
accggcacgc gcgtgcccgt gcgcttccgc cccccgcgga cttgagatga tgtgaaaaaa 16680
caacactgag tctcctgctg ttgtgtgtat cccagcggcg gcggcgcgcg cagcgtcatg 16740
tccaagcgca aaatcaaaga agagatgctc caggtcgtcg cgccggagat ctatgggccc 16800
ccgaagaagg aagagcagga ttcgaagccc cgcaagataa agcgggtcaa aaagaaaaag 16860
aaagatgatg acgatgccga tggggaggtg gagttcctgc gcgccacggc gcccaggcgc 16920
ccggtgcagt ggaagggccg gcgcgtaaag cgcgtcctgc gccccggcac cgcggtggtc 16980
ttcacgcccg gcgagcgctc cacccggact ttcaagcgcg tctatgacga ggtgtacggc 17040
gacgaagacc tgctggagca ggccaacgag cgcttcggag agtttgctta cgggaagcgt 17100
cagcgggcgc tggggaagga ggacctgctg gcgctgccgc tggaccaggg caaccccacc 17160
cccagtctga agcccgtgac cctgcagcag gtgctgccga gcagcgcacc ctccgaggcg 17220
aagcggggtc tgaagcgcga gggcggcgac ctggcgccca ccgtgcagct catggtgccc 17280
aagcggcaga ggctggagga tgtgctggag aaaatgaaag tagaccccgg tctgcagccg 17340
gacatcaggg tccgtcccat caagcaggtg gcgccgggcc tcggcgtgca gaccgtggac 17400
gtggtcatcc ccaccggcaa ctcccccgcc gccaccacca ctaccgctgc ctccacggac 17460
atggagacac agaccgatcc cgccgcagcc gcagccgccg ccgcagccgc gacctcctcg 17520
gcggaggtgc agacggaccc ctggctgccg ccggcgatgt cagctccccg cgcgcgccgc 17580
ggacgcagaa agtacggcgc cgccaacgcg ctcctgcccg agtacgcctt gcatccttcc 17640
atcgcgccca cccccggcta ccgaggctat acctaccgcc cgcgaagagc caagggttcc 17700
acccgccgtc cccgccgacg cgccgccgcc accacccgcc gccgccgccg cagacgccag 17760
cccgcactgg ctccagtctc cgtgaggaga gtggcgcgcg acggacacac cctggtgctg 17820
cccagggcgc gctaccaccc cagcatcgtt taaaagcctg ttgtggttct tgcagatatg 17880
gccctcactt gccgcctccg tttcccggtg ccgggatacc gaggaggaag atcgcgccgc 17940
aggaggggtc tggccggccg cggcctgagc ggaggcagcc gccgcgcgca ccggcggcga 18000
cgcgccacca gccgacgcat gcgcggcggg gtgctgcccc tgttaatccc cctgatcgcc 18060
gcggcgatcg gcgccgtgcc cgggatcgcc tccgtggcct tgcaagcgtc ccagaggcat 18120
tgacagactt gcaaacttgc aaatatggaa aaaaaaaaaa aaccccaata aaaagtctag 18180
actctcacgc tcgcttggtc ctgtgactat tttgtagaat ggaagacatc aactttgcgt 18240
cgctggcccc gcgtcacggc tcgcgcccgt tcctgggaca ctggaacgat atcggcacca 18300
gcaacatgag cggtggcgcc ttcagttggg gctctctgtg gagcggcatt aaaagtatcg 18360
ggtctgccgt taaaaattac ggctcccggg cctggaacag cagcacgggc cagatgttga 18420

gagacaagtt gaaagagcag aacttccagc agaaggtggt ggagggcctg gcctccggca 18480
tcaacggggt ggtggacctg gccaaccagg ccgtgcagaa taaaatcaac agcagactgg 18540
acccccggcc gccggtggag gaggtgccgc cggcgctgga gacggtgtcc cccgatgggc 18600
gtggcgagaa gcgcccgcgg cccgataggg aagagaccac tctggtcacg cagaccgatg 18660
agccgccccc gtatgaggag gccctaaagc aaggtctgcc caccacgcgg cccatcgcgc 18720
ccatggccac cggggtggtg ggccgccaca cccccgccac gctggacttg cctccgcccg 18780
ccgatgtgcc gcagcagcag aaggcggcac agccgggccc gcccgcgacc gcctcccgtt 18840
cctccgccgg tcctctgcgc cgcgcggcca gcggcccccg cgggggggtc gcgaggcacg 18900
gcaactggca gagcacgctg aacagcatcg tgggtctggg ggtgcggtcc gtgaagcgcc 18960
gccgatgcta ctgaatagct tagctaacgt gttgtatgtg tgtatgcgcc ctatgtcgcc 19020
gccagaggag ctgctgagtc gccgccgttc gcgcgcccac caccaccgcc actccgcccc 19080
tcaagatggc gaccccatcg atgatgccgc agtggtcgta catgcacatc tcgggccagg 19140
acgcctcgga gtacctgagc cccgggctgg tgcagttcgc ccgcgccacc gagagctact 19200
tcagcctgag taacaagttt aggaacccca cggtggcgcc cacgcacgat gtgaccaccg 19260
accggtctca gcgcctgacg ctgcggttca ttcccgtgga ccgcgaggac accgcgtact 19320
cgtacaaggc gcggttcacc ctggccgtgg gcgacaaccg cgtgctggac atggcctcca 19380
cctactttga catccgcggg gtgctggacc ggggtcccac tttcaagccc tactctggca 19440
ccgcctacaa ctccctggcc cccaagggcg ctcccaactc ctgcgagtgg gagcaagagg 19500
aaactcaggc agttgaagaa gcagcagaag aggaagaaga agatgctgac ggtcaagctg 19560
aggaagagca agcagctacc aaaaagactc atgtatatgc tcaggctccc ctttctggcg 19620
aaaaaattag taaagatggt ctgcaaatag gaacggacgc tacagctaca gaacaaaaac 19680
ctatttatgc agaccctaca ttccagcccg aaccccaaat cggggagtcc cagtggaatg 19740
aggcagatgc tacagtcgcc ggcggtagag tgctaaagaa atctactccc atgaaaccat 19800
gctatggttc ctatgcaaga cccacaaatg ctaatggagg tcagggtgta ctaacggcaa 19860
atgcccaggg acagctagaa tctcaggttg aaatgcaatt cttttcaact tctgaaaacg 19920
cccgtaacga ggctaacaac attcagccca aattggtgct gtatagtgag gatgtgcaca 19980
tggagacccc ggatacgcac ctttcttaca agcccgcaaa aagcgatgac aattcaaaaa 20040
tcatgctggg tcagcagtcc atgcccaaca gacctaatta catcggcttc agagacaact 20100
ttatcggcct catgtattac aatagcactg gcaacatggg agtgcttgca ggtcaggcct 20160
ctcagttgaa tgcagtggtg gacttgcaag acagaaacac agaactgtcc taccagctct 20220
tgcttgattc catgggtgac agaaccagat acttttccat gtggaatcag gcagtggaca 20280
gttatgaccc agatgttaga attattgaaa atcatggaac tgaagacgag ctccccaact 20340
attgtttccc tctgggtggc ataggggtaa ctgacactta ccaggctgtt aaaaccaaca 20400
atggcaataa cgggggccag gtgacttgga caaaagatga aacttttgca gatcgcaatg 20460
aaatagggggt gggaaacaat ttcgctatgg agatcaacct cagtgccaac ctgtggagaa 20520
acttcctgta ctccaacgtg gcgctgtacc taccagacaa gcttaagtac aacccctcca 20580
atgtggacat ctctgacaac cccaacacct acgattacat gaacaagcga gtggtggccc 20640
cggggctggt ggactgctac atcaacctgg gcgcgcgctg gtcgctggac tacatggaca 20700
acgtcaaccc cttcaaccac caccgcaatg cgggcctgcg ctaccgctcc atgctcctgg 20760
gcaacgggcg ctacgtgccc ttccacatcc aggtgcccca gaagttcttt gccatcaaga 20820
acctcctcct cctgccgggc tcctacacct acgagtggaa cttcaggaag gatgtcaaca 20880
tggtcctcca gagctctctg ggtaacgatc tcagggtgga cggggccagc atcaagttcg 20940
agagcatctg cctctacgcc accttcttcc ccatggccca caacacggcc tccacgctcg 21000
aggccatgct caggaacgac accaacgacc agtccttcaa tgactacctt tccgccgcca 21060
acatgctcta ccccataccc gccaacgcca ccaacgtccc catctccatc ccctcgcgca 21120
actgggcggc cttccgcggc tgggccttca cccgcctcaa gaccaaggag acccccctccc 21180
tgggctcggg attcgacccc tactacacct actcgggctc tattccctac ctggacggca 21240
ccttctacct caaccacact ttcaagaagg tctcggtcac cttcgactcc tcggtcagct 21300
ggccgggcaa cgaccgtctg ctcaccccca acgagttcga gatcaagcgc tcggtcgacg 21360
gggaaggcta caacgtggcc cagtgcaaca tgaccaagga ctggttcctg gtccagatgc 21420
tggccaacta caacatcggc taccagggct tctacatccc agagagctac aaggacagga 21480
tgtactcctt cttcaggaac ttccagccca tgagccggca ggtggtggac cagaccaagt 21540
acaaggacta ccaggaggtg ggcatcatcc accagcacaa caactcgggc ttcgtgggct 21600
acctcgcccc caccatgcgc gagggacagg cctaccccgc caacttcccc tacccgctca 21660
taggcaagac cgcggtcgac agcatcaccc agaaaaagtt cctctgcgac cgcaccctct 21720

ggcgcatccc cttctccagc aacttcatgt ccatgggtgc gctctcggac ctgggccaga 21780
acttgctcta cgccaactcc gcccacgccc tcgacatgac cttcgaggtc gaccccatgg 21840
acgagcccac cctttctctat gttctgttcg aagtctttga cgtggtccgg gtccaccagc 21900
cgcaccgcgg cgtcatcgag accgtgtacc tgcgtacgcc cttctcggcc ggcaacgcca 21960
ccacctaaag aagcaagccg cagtcatcgc cgcctgcatg ccgtcgggtt ccaccgagca 22020
agagctcagg gccatcgtca gagacctggg atgcgggccc tatttttgg gcaccttcga 22080
caagcgcttc cctggctttg tctccccaca caagctggcc tgcgccatcg tcaacacggc 22140
cggccgcgag accgggggcg tgcactggct ggcctttgcc tggaacccgc gctccaaaac 22200
atgcttcctc tttgacccct tcggcttttc ggaccagcgg ctcaagcaaa tctacgagtt 22260
cgagtacgag ggcttgctgc gtcgcagcgc catcgcctcc tcgcccgacc gctgcgtcac 22320
cctcgaaaag tccacccaga ccgtgcaggg gcccgactcg gccgcctgcg gtctcttctg 22380
ctgcatgttt ctgcacgcct ttgtgcactg gcctcagagt cccatggacc gcaaccccac 22440
catgaacttg ctgacggggg tgcccaactc catgctccaa agcccccagg tcgagcccac 22500
cctgcgccgc aaccaggagc agctctacag cttcctggag cgccactcgc cctacttccg 22560
ccgccacagc gcacagatca ggagggccac ctccttctgc cacttgcaag agatgcaaga 22620
agggtaataa cgatgtacac actttttct caataaatgg catttttttt ttatttatac 22680
aagctctctg gggtattcat ttcccaccac caccacccgc cgttgtcgcc atctggctct 22740
atttagaaat cgaaagggtt ctgccgggag tcgccgtgcg ccacgggcag ggacacgttg 22800
cgatactggt agcgggtgcc ccacttgaac tcgggcacca ccaggcgagg cagctcgggg 22860
aagttttcgc tccacaggct gcgggtcagc accagcgcgt tcatcaggtc gggcgccgag 22920
atcttgaagt cgcagttggg gccgccgccc tgcgcgcgcg agttgcggta caccgggttg 22980
cagcactgga acaccaacag cgccgggtgc ttcacgctgg ccagcacgct gcggtcggag 23040
atcagctcgg cgtccaggtc ctccgcgttg ctcagcgcga acggggtcat cttgggcact 23100
tgccgcccca ggaagggcgc gtgccccggt ttcgagttgc agtcgcagcg cagcgggatc 23160
agcaggtgcc cgtgcccgga ctcggcgttg gggtacagcg cgcgcatgaa ggcctgcatc 23220
tggcggaagg ccatctgggc cttggcgccc tccgagaaga acatgccgca ggacttgccc 23280
gagaactggt ttgcggggca gctggcgtcg tgcaggcagc agcgcgcgtc ggtgttggcg 23340
atctgcacca cgttgcgccc ccaccggttc ttcacgatct tggccttgga cgattgctcc 23400
ttcagcgcgc gctgcccgtt ctcgctggtc acatccatct cgatcacatg ttccttgttc 23460
accatgctgc tgccgtgcag acacttcagc tcgccctccg tctcggtgca gcggtgctgc 23520
cacagcgcgc agcccgtggg ctcgaaagac ttgtaggtca cctccgcgaa ggactgcagg 23580
taccctgca aaaagcggcc catcatggtc acgaaggtct tgttgctgct gaaggtcagc 23640
tgcagcccgc ggtgctcctc gttcagccag gtcttgcaca cggccgccag cgcctccacc 23700
tggtcgggca gcatcttgaa gttcaccttc agctcattct ccacgtggta cttgtccatc 23760
agcgtgcgcg ccgcctccat gcccttctcc caggccgaca ccagcggcag gctcacgggg 23820
ttcttcacca tcaccgtggc cgccgcctcc gccgcgcttt cgctttccgc cccgctgttc 23880
tcttcctctt cctcctcttc ctcgccgccg cccactcgca gcccccgcac cacggggtcg 23940
tcttcctgca ggcgctgcac cttgcgcttg ccgttgcgcc cctgcttgat gcgcacgggc 24000
gggttgctga agcccaccat caccagcgcg gcctcttctt gctcgtcctc gctgtccaga 24060
atgacctccg gggaggggggg gttggtcatc ctcagtaccg aggcacgctt ctttttcttc 24120
ctgggggcgt tcgccagctc cgcggctgcg gccgctgccg aggtcgaagg ccgagggctg 24180
ggcgtgcgcg gcaccagcgc gtcttgcgag ccgtcctcgt cctcctcgga ctcgagacgg 24240
aggcgggccc gcttcttcgg gggcgcgcgg ggcggcggag gcggcggcgg cgacggagac 24300
ggggacgaga catcgtccag ggtgggtgga cggcgggccg cgccgcgtcc gcgctcgggg 24360
gtggtttcgc gctggtcctc ttcccgactg gccatctccc actgctcctt ctcctatagg 24420
cagaaagaga tcatggagtc tctcatgcga gtcgagaagg aggaggacag cctaaccgcc 24480
ccctctgagc cctccaccac cgccgccacc accgccaatg ccgccgcgga cgacgcgccc 24540
accgagacca ccgccagtac caccctcccc agcgacgcac ccccgctcga gaatgaagtg 24600
ctgatcgagc aggacccggg ttttgtgagc ggagaggagg atgaggtgga tgagaaggag 24660
aaggaggagg tcgccgcctc agtgccaaaa gaggataaaa agcaagacca ggacgacgca 24720
gataaggatg agacagcagt cgggcggggg aacggaagcc atgatgctga tgacggctac 24780
ctagacgtgg gagacgacgt gctgcttaag cacctgcacc gccagtgcgt catcgtctgc 24840
gacgcgctgc aggagcgctg cgaagtgccc ctggacgtgg cggaggtcag ccgcgcctac 24900
gagcggcacc tcttcgcgcc gcacgtgccc cccaagcgcc gggagaacgg cacctgcgag 24960
cccaacccgc gtctcaactt ctacccggtc ttcgcggtac ccgaggtgct ggccacctac 25020

cacatcttct tccaaaactg caagatcccc ctctcctgcc gcgctaaccg cacccgcgcc 25080
gacaaaaccc tgaccctgcg gcagggcgcc cacatacctg atattgcctc tctggaggaa 25140
gtgcccaaga tcttcgaggg tctcggtcgc gacgagaaac gggcggcgaa cgctctgcac 25200
ggagacagcg aaaacgagag tcactcgggg gtgctggtgg agctcgaggg cgacaacgcg 25260
cgcctggccg tactcaagcg cagcatagag gtcacccact ttgcctaccc ggcgctcaac 25320
ctgcccccca aggtcatgag tgtggtcatg ggcgagctca tcatgcgccg cgctcagccc 25380
ctggccgcgg atgcaaactt gcaagagtcc tccgaggaag gcctgcccgc ggtcagcgac 25440
gagcagctag cgcgctggct ggagacccgc gaccccgcgc agctggagga gcggcgcaag 25500
ctcatgatgg ccgcggtgct ggtcaccgtg gagctcgagt gtctgcagcg cttcttcgcg 25560
gaccccgaga tgcagcgcaa gctcgaggag accctgcact acaccttccg ccagggctac 25620
gtgcgccagg cctgcaagat ctccaacgtg gagctctgca acctggtctc ctacctgggc 25680
atcctgcacg agaaccgcct cgggcagaac gtcctgcact ccaccctcaa aggggaggcg 25740
cgccgcgact acatccgcga ctgcgcctac ctcttcctct gctacacctg gcagacggcc 25800
atgggggtct ggcagcagtg cctggaggag cgcaacctca aggagctgga aaagctactc 25860
aagcgcaccc tcagggacct ctggacgggc ttcaacgagc gctcggtggc cgccgcgctg 25920
gcggacatca tcttccccga gcgcctgctc aagaccctgc agcagggcct gcccgacttc 25980
accagccaga gcatgctgca gaactttagg actttcatcc tggagcgctc gggcatcctg 26040
cctgccactt gctgcgcgct gcccagcgac ttcgtgccca tcaagtacag ggagtgcccg 26100
ccgccgctct ggggccactg ctacctcttc cagctggcca actacctcgc ctaccactcg 26160
gacctcatgg aagacgtgag cggcgagggc ctgctcgagt gccactgccg ctgcaacctc 26220
tgcacgcccc accgctctct agtctgcaac ccgcagctgc tcagcgagag tcagattatc 26280
ggtaccttcg agctgcaggg tccctcgcct gacgagaagt ccgcggctcc ggggctgaaa 26340
ctcactccgg ggctgtggac ttccgcctac ctacgcaaat ttgtacctga ggactaccac 26400
gcccacgaga tcaggttcta cgaagaccaa tcccgcccgc ccaaggcgga gctcaccgcc 26460
tgcgtcatca cccaggggca catcctgggc caattgcaag ccatcaacaa agcccgccga 26520
gagttcttgc tgaaaaaggg tcgggggtg tacctggacc cccagtccgg cgaggagcta 26580
aacccgctac ccccgccgcc gccccagcag cgggaccttg cttcccagga tggcacccag 26640
aaagaagcag cagccgccgc cgccgcagcc atacatgctt ctggaggaag aggaggagga 26700
ctgggacagt caggcagagg aggtttcgga cgaggagcag gaggagatga tggaagactg 26760
ggaggaggac agcagcctag acgaggaagc ttcagaggcc gaagaggtgg cagacgcaac 26820
accatcaccc tcggtcgcag ccccctcgcc ggggcccctg aaatcctccg aacccagcac 26880
cagcgctata acctccgctc ctccggcgcc ggcgccaccc gcccgcagac ccaaccgtag 26940
atgggacacc acaggaaccg gggtcggtaa gtccaagtgc ccgccgccgc caccgcagca 27000
gcagcagcag cgccagggct accgctcgtg gcgcgggcac aagaacgcca tagtcgcctg 27060
cttgcaagac tgcgggggca acatctcttt cgcccggcgc ttcctgctat tccaccacgg 27120
ggtcgccttt ccccgcaatg tcctgcatta ctaccgtcat ctctacagcc cctactgcag 27180
cggcgaccca gaggcggcag cggcagccac agcggcgacc accacctagg aagatatcct 27240
ccgcgggcaa gacagcggca gcagcggcca ggagacccgc ggcagcagcg gcgggagcgg 27300
tgggcgcact gcgcctctcg cccaacgaac ccctctcgac ccgggagctc agacacagga 27360
tcttccccac tttgtatgcc atcttccaac agagcagagg ccaggagcag gagctgaaaa 27420
taaaaaacag atctctgcgc tccctcaccc gcagctgtct gtatcacaaa agcgaagatc 27480
agcttcggcg cacgctggag gacgcggagg cactcttcag caaatactgc gcgctcactc 27540
ttaaagacta gctccgcgcc cttctcgaat ttaggcggga gaaaactacg tcatcgccgg 27600
ccgccgccca gcccgcccag ccgagatgag caaagagatt cccacgccat acatgtggag 27660
ctaccagccg cagatgggac tcgcggcggg agcggcccag gactactcca cccgcatgaa 27720
ctacatgagc gcgggacccc acatgatctc acaggtcaac gggatccgcg cccagcgaaa 27780
ccaaatactg ctggaacagg cggccatcac cgccacgccc cgccataatc tcaacccccg 27840
aaattggccc gccgccctcg tgtaccagga aacccccctcc gccaccaccg tactacttcc 27900
gcgtgacgcc caggccgaag tccagatgac taactcaggg gcgcagctcg cgggcggctt 27960
tcgtcacggg gcgcggccgc tccgaccagg tataagacac ctgatgatca gaggccgagg 28020
tatccagctc aacgacgagt cggtgagctc ttcgctcggt ctccgtccgg acggaacttt 28080
ccagctcgcc ggatccggcc gctcttcgtt cacgccccgc caggcgtacc tgactctgca 28140
gacctcgtcc tcggagcccc gctccggagg catcggaacc ctccagttcg tggaggagtt 28200
cgtgccctcg gtctacttca acccttctc gggacctccc ggacgctacc ccgaccagtt 28260
cattccgaac tttgacgcgg tgaaggactc ggcggacggc tacgactgaa tgtcaggtgc 28320

cgaggcagag cagcttcgcc tgagacacct cgagcactgc cgccgccaca agtgcttcgc 28380
ccgcggttcc ggtgagttct gctactttca gctacccgag gagcataccg aggggccggc 28440
gcacggcgtc cgcctgacca cccagggcga ggttacctgt tccctcatcc gggagttcac 28500
cctccgtccc ctgctagtgg agcgggagcg gggtccctgt gtcctaacta tcgcctgcaa 28560
ctgccctaac cctggattac atcaagatct ttgctgtcat ctctgtgctg agtttaataa 28620
acgctgagat cagaatctac tggggctcct gtcgccatcc tgtgaacgcc accgtcttca 28680
cccaccccga ccaggcccag gcgaacctca cctgcggtct gcatcggagg gccaagaagt 28740
acctcacctg gtacttcaac ggcacccccc ttgtggttta caacagcttc gacggggacg 28800
gagtctccct gaaagaccag ctctccggtc tcagctactc catccacaag aacaccaccc 28860
tccaactctt ccctccctac ctgccgggaa cctacgagtg cgtcaccggc cgctgcaccc 28920
acctcacccg cctgatcgta aaccagagct ttccgggaac agataactcc ctcttcccca 28980
gaacaggagg tgagctcagg aaactccccg gggaccaggg cggagacgta ccttcgaccc 29040
ttgtggggtt aggatttttt attaccgggt tgctggctct tttaatcaaa gcttccttga 29100
gatttgttct ttccttctac gtgtatgaac acctcagcct ccaataactc tacccttttct 29160
tcggaatcag gtgacttctc tgaaatcggg cttggtgtgc tgcttactct gttgattttt 29220
ttccttatca tactcagcct tctgtgcctc aggctcgccg cctgctgcgc acacatctat 29280
atctactgct ggttgctcaa gtgcaggggt cgccacccaa gatgaacagg tacatggtcc 29340
tatcgatcct aggcctgctg gccctggcgg cctgcagcgc cgccaaaaaa gagattacct 29400
ttgaggagcc cgcttgcaat gtaactttca agcccgaggg tgaccaatgc accaccctcg 29460
tcaaatgcgt taccaatcat gagaggctgc gcatcgacta caaaaacaaa actggccagt 29520
ttgcggtcta tagtgtgttt acgcccggag acccctctaa ctactctgtc accgtcttcc 29580
agggcggaca gtctaagata ttcaattaca ctttcccttt ttatgagtta tgcgatgcgg 29640
tcatgtacat gtcaaaacag tacaacctgt ggcctccctc tccccaggcg tgtgtggaaa 29700
atactgggtc ttactgctgt atggctttgg caatcactac gctcgctcta atctgcacgg 29760
tgctatacat aaaattcagg cagaggcgaa tctttatcga tgaaaagaaa atgccttgat 29820
cgctaacacc ggctttctat ctgcagaatg aatgcaatca cctccctact aatcaccacc 29880
accctccttg cgattgccca tgggttgaca cgaatcgaag tgccagtggg gtccaatgtc 29940
accatggtgg gccccgccgg caattccacc ctcatgtggg aaaaatttgt ccgcaatcaa 30000
tgggttcatt tctgctctaa ccgaatcagt atcaagccca gagccatctg cgatgggcaa 30060
aatctaactc tgatcaatgt gcaaatgatg gatgctgggt actattacgg gcagcgggga 30120
gaaatcatta attactggcg accccacaag gactacatgc tgcatgtagt cgaggcactt 30180
cccactacca cccccactac cacctctccc accaccacta ccaccactac tactactact 30240
actaccacta ccgctgcccg ccatacccgc aaaagcacca tgattagcac aaagcccccc 30300
cgtgctcact cccacgccgg cgggcccatc ggtgcgacct cagaaaccac cgagctttgc 30360
ttctgccaat gcactaacgc cagcgctcat gaactgttcg acctggagaa tgaggatgcc 30420
cagcagagct ccgcttgcct gacccaggag gctgtggagc ccgttgccct gaagcagatc 30480
ggtgattcaa taattgactc ttcttctttt gccactcccg aataccctcc cgattctact 30540
ttccacatca cgggtaccaa agaccctaac ctctctttct acctgatgct gctgctctgt 30600
atctctgtgg tctcttccgc gctgatgtta ctggggatgt tctgctgcct gatctgccgc 30660
agaaagagaa aagctcgctc tcagggccaa ccactgatgc ccttccccta ccccccggat 30720
tttgcagata acaagatatg agctcgctgc tgacactaac cgctttacta gcctgcgctc 30780
taacccttgt cgcttgcgac tcgagattcc acaatgtcac agctgtggca ggagaaaatg 30840
ttactttcaa ctccacggcc gatacccagt ggtcgtggag tggctcaggt agctacttaa 30900
ctatctgcaa tagctccact tcccccagca tatccccaac caagtaccaa tgcaatgcca 30960
gcctgttcac cctcatcaac gcttccaccc tggacaatgg actctatgta ggctatgtac 31020
cctttggtgg gcaaggaaag acccacgctt acaacctgga agttcgccag cccagaacca 31080
ctacccaagc ttctcccacc accaccacca ccaccaccac caccatcacc agcagcagca 31140
gcagccacag cagcagcagc agattattga ctttggtttt ggccagctca tctgccgcta 31200
cccaggccat ctacagctct gtgcccgaaa ccactcagat ccaccgccca gaaacgacca 31260
ccgccaccac cctacacacc tccagcgatc agatgccgac caacatcacc cccttggctc 31320
ttcaaatggg acttacaagc cccactccaa aaccagtgga tgcggccgag gtctccgccc 31380
tcgtcaatga ctgggcgggg ctgggaatgt ggtggttcgc cataggcatg atggcgctct 31440
gcctgcttct gctctggctc atctgctgcc tccaccgcag gcgagccaga ccccccatct 31500
atagacccat cattgtcctg aaccccgata atgatgggat ccatagattg gatggcctga 31560
aaaacctact tttttctttt acagtatgat aaattgagac atgcctcgca ttttcttgta 31620

catgttcctt ctcccacctt ttctggggtg ttctacgctg gccgctgtgt ctcacctgga 31680
ggtagactgc ctctcaccct tcactgtcta cctgctttac ggattggtca ccctcactct 31740
catctgcagc ctaatcacag taatcatcgc cttcatccag tgcattgatt acatctgtgt 31800
gcgcctcgca tacttcagac accacccgca gtaccgagac aggaacattg cccaacttct 31860
aagactgctc taatcatgca taagactgtg atctgccttc tgatcctctg catcctgccc 31920
accctcacct cctgccagta caccacaaaa tctccgcgca aaagacatgc ctcctgccgc 31980
ttcacccaac tgtggaatat acccaaatgc tacaacgaaa agagcgagct ctccgaagct 32040
tggctgtatg gggtcatctg tgtcttagtt ttctgcagca ctgtctttgc cctcatgatc 32100
taccctact ttgatttggg atggaacgcg atcgatgcca tgaattaccc caccttccc 32160
gcacccgaga taattccact gcgacaagtt gtacccgttg tcgttaatca acgccccca 32220
tcccctacgc ccactgaaat cagctacttt aacctaacag gcggagatga ctgacgccct 32280
agatctagaa atggacggca tcagtaccga gcagcgtctc ctagagaggc gcaggcaggc 32340
ggctgagcaa gagcgcctca atcaggagct ccgagatctc gttaacctgc accagtgcaa 32400
aagaggcatc ttttgtctgg taaagcaggc caaagtcacc tacgagaaga ccggcaacag 32460
ccaccgcctc agttacaaat tgcccaccca gcgccagaag ctggtgctca tggtgggtga 32520
gaatcccatc accgtcaccc agcactcggt agagaccgag gggtgtctgc actctccctg 32580
tcggggtcca gaagacctct gcaccctggt aaagaccctg tgcggtctca gagatttagt 32640
ccccttaac taatcaaaca ctggaatcaa taaaaagaat cacttactta aaatcagaca 32700
gcaggtctct gtccagttta ttcagcagca cctccttccc ctcctcccaa ctctggtact 32760
ccaaacgcct tctggcggca aacttcctcc acaccctgaa gggaatgtca gattcttgct 32820
cctgtccctc cgcacccact atcttcatgt tgttgcagat gaagcgcacc aaaacgtctg 32880
acgagagctt caaccccgtg tacccctatg acacggaaag cggccctccc tccgtccctt 32940
tcctcacccc tcccttcgtg tctcccgatg gattccaaga aagccccccc ggggtcctgt 33000
ctctgaacct ggccgagccc ctggtcactt cccacggcat gctcgccctg aaaatgggaa 33060
gtggcctctc cctggacgac gctggcaacc tcacctctca agatatcacc accgctagcc 33120
ctcccctcaa aaaaaccaag accaacctca gcctagaaac ctcatccccc ctaactgtaa 33180
gcacctcagg cgccctcacc gtagcagccg ccgctcccct ggcagtggcc ggcacctccc 33240
tcaccatgca atcagaggcc cccctgacag tacaggatgc aaaactcacc ctggccacca 33300
aaggcccccc gaccgtgtct gaaggcaaac tggccttgca aacatcggcc ccgctgacgg 33360
ccgctgacag cagcaccctc accgttagcg ccacaccacc aattaatgta agcagtggaa 33420
gtttaggctt agacatggaa gaccctatgt atactcacga tggaaaactg ggaataagaa 33480
ttgggggtcc actaagagta gtagacagct tgcacacact cactgtagtt accggaaatg 33540
gactaactgt agataacaat gccctccaaa ctagagttac gggcgcccta ggttatgaca 33600
catcaggaaa tctacaattg agagctgcag gaggtatgcg aattgatgca aatggccaac 33660
ttatccttaa tgtggcatac ccatttgatg ctcagaacaa tctcagcctt agacttggtc 33720
agggacccct gtatataaac acagaccaca acctggattt gaattgcaac agaggtctaa 33780
ccacaactac caccaacaac acaaaaaaac ttgagactaa aattagctca ggcttagact 33840
atgacaccaa tggtgctgtc attattaaac ttggcactgg tctaagcttc gacaacacag 33900
gcgccctaac tgtgggaaac actggtgatg ataaactgac tctgtggacg accccagacc 33960
catctccaaa ttgcagaatt cactcagaca aagactgcaa gtttactcta gtcctaacta 34020
agtgtggaag ccaaatcctg gcctctgtcg ccgccctagc ggtatcagga aatctggctt 34080
cgataacagg caccgttgcc agcgttacca tctttctcag atttgatcag aatggagtgc 34140
ttatggaaaa ctcctcgcta gacaggcagt actggaactt cagaaatggc aactcaacta 34200
acgctgcccc ctacaccaat gcagttgggt tcatgccaaa cctcgcagca taccccaaaa 34260
cgcaaagcca gactgctaaa aacaacattg taagtcaggt ttacttgaat ggagacaaat 34320
ccaaacccat gacccttacc atcaccctca atggaactaa tgaatccagt gaaactagcc 34380
aggtgagtca ctactccatg tcatttacat gggcttggga aagtgggcaa tatgccactg 34440
aaacctttgc caccaactcc ttcacctttt cttacattgc tgaacaataa aaagcatgac 34500
actgatgttc atttctgatt cttattttat tattttcaaa cacaacaaaa tcattcaagt 34560
cattcttcca tcttagctta atagacacag tagcttaata gacccagtag tgcaaagccc 34620
cattctagct tatagatcag acagtgataa ttaaccacca ccaccaccat acctttgat 34680
tcaggaaatc atgatcatca caggatccta gtcttcaggc cgccccctcc ctcccaagac 34740
acagaataca cagtcctctc cccccgactg gctttaaata acaccatctg gttggtcaca 34800
gacatgttct taggggttat attccacacg gtctcctgcc gcgccaggcg ctcgtcggtg 34860
atgttgataa actctcccgg cagctcgctc aagttcacgt cgctgtccag cggctgaacc 34920

```
tccggctgac gcgataactg tgcgaccggc tgctggacaa acggaggccg cgcctacaag 34980
ggggtagagt cataatcctc ggtcaggata gggcggtgat gcagcagcag cgagcgaaac 35040
atctgctgcc gccgccgctc cgtccggcag gaaaacaaca agccggtggt ctcctccgcg 35100
ataatccgca ccgcccgcag catcagcttc ctcgttctcc gcgcgcagca cctcaccctg 35160
atctcgctca agtcggcgca gtaggtacag cacagcacca cgatgttatt catgatccca 35220
cagtgcaggg cgctgtatcc aaagctcatg ccgggaacca ccgcccccac gtggccatcg 35280
taccacaagc gcacgtaaat taagtgtcga cccctcatga acgtgctgga cacaaacatt 35340
acttccttgg gcatgttgta attcaccacc tcccggtacc agataaacct ctggttaaac 35400
agggcacctt ccaccaccat cctgaaccaa gaggccagaa cctgcccacc ggctatgcac 35460
tgcagggaac ccgggttgga acaatgacaa tgcagactcc aaggctcgta accgtggatc 35520
atccggctgc tgaaggcatc gatgttggca caacacagac acacgtgcat gcactttctc 35580
atgattagca gctcttccct cgtcaggatc atatcccaag gaataaccca ttcttgaatc 35640
aacgtaaaac ccacacagca gggaaggcct cgcacataac tcacgttgtg catggtcagc 35700
gtgttgcatt ctggaaacag cggatgatcc tccagtatcg aggcgcgggt ctccttctca 35760
cagggaggta aagggtccct gctgtacgga ctgcgccggg acgaccgaga tcgtgttgag 35820
cgtagtgtca tggaaaaggg aacgccggac gtggtcatac ttcttgaagc agaaccaggt 35880
tcgcgcgtgg caggcctcct tgcgtctgcg gtctcgccgt ctagctcgct ccgtgtgata 35940
gttgtagtac agccactccc gcagagcgtc gaggcgcacc ctggcttccg gatctatgta 36000
gactccgtct tgcaccgcgg ccctgataat atccaccacc gtagaataag caacacccag 36060
ccaagcaata cactcgctct gcgagcggca gacaggagga gcgggcagag atgggagaac 36120
catgataaaa aacttttttt aaagaatatt ttccaattct tcgaaagtaa gatctatcaa 36180
gtggcagcgc tcccctccac tggcgcggtc aaactctacg gccaaagcac agacaacggc 36240
atttctaaga tgttccttaa tggcgtccaa aagacacacc gctctcaagt tgcagtaaac 36300
tatgaatgaa aacccatccg gctgattttc caatatagac gcgccggcgg cgtccaccaa 36360
acccagataa ttttcttctc tccagcggtt tagaatctgt ctaagcaaat cccttatatc 36420
aagtccggcc atgccaaaaa tctgctcaag agcgccctcc accttcatga ccaagcagcg 36480
catcatgatt gcaaaaattc aggttcttca gagacctgta taagattcaa aatgggaaca 36540
ttaacaaaaa ttcctctgtc gcgcagatcc cttcgcaggg caagctgaac ataatcagac 36600
aggtctgaac ggaccagtga ggccaaatcc ccaccaggaa ccagatccag agaccctata 36660
ctgattatga cgcgcatact cggggctatg ctgaccagcg tagcgccgat gtaggcgtgc 36720
tgcatgggcg gcgagataaa atgcaaagtg ctggttaaaa aatcaggcaa agcctcgcgc 36780
aaaaaagcta acacatcata atcatgctca tgcaggtagt tgcaggtaag ctcaggaacc 36840
aaaacggaat aacacacgat tttcctctca aacatgactt cgcggatact gcgtaaaaca 36900
aaaattataa ataaaaaatt aattaactta aacattggaa gcctgtctca caacaggaaa 36960
aaccacttta atcaacataa gacgggccac gggcatgccg gcatagccgt aaaaaaattg 37020
gtccccgtga ttaacaagta ccacagacag ctccccggtc atgtcggggg tcatcatgtg 37080
agactctgta tacacgtctg gattgtgaac atcagacaaa caaagaaatc gagccacgta 37140
gcccggaggt ataatcaccc gcaggcggag gtacagcaaa acgaccccca taggaggaat 37200
cacaaaatta gtaggagaaa aaaatacata aacaccagaa aaaccctgtt gctgaggcaa 37260
aatagcgccc tcccgatcca aaacaacata aagcgcttcc acaggagcag ccataacaaa 37320
gacccgagtc ttaccagtaa aagaaaaaag atctctcaac gcagcaccag caccaacact 37380
tcgcagtgta aaaggccaag tgccgagaga gtatatatag gaataaaaag tgacgtaaac 37440
gggcaaagtc caaaaaacgc ccagaaaaac cgcacgcgaa cctacgcccc gaaacgaaag 37500
ccaaaaaaca ctagacactc ccttccggcg tcaacttccg ctttcccacg ctacgtcact 37560
tgccccagtc aaacaaacta catatcccga acttccaagt cgccacgccc aaaacaccgc 37620
ctacacctcc ccgcccgccg gcccgccccc aaacccgcct cccgccccgc gccccgcctc 37680
gcgccgccca tctcattatc atattggctt caatccaaaa taaggtatat tattgatgat 37740
g                                                                  37741
```

<210> 2
<211> 36648
<212> DNA
<213> Chimpanzee Adenvirus-ChAd6 Genomic

<400> 2

catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg agccgtttga 60
atttggggat gcggggcgct gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtgtt tgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg 660
gtgacgaccc tcctgagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgtatg 720
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggcc gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgca gcgagcgagg gagtgaaagc tgcgggcgag agctttagcc 1020
tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
acagtaagtg tgattaactt tagctgggaa ggcagagggt gactgggtgc tgactggttt 1200
atttatgtat atgtttttta tgtgtaggtc ccgtctctga cgtagatgag acccccactt 1260
cagagtgcat ttcatcaccc ccagaaattg gcgaggaacc gcccgaagat attattcata 1320
gaccagttgc agtgagagtc accgggcgga gagcagctgt ggagagtttg gatgacttgc 1380
tacagggtgg ggatgaacct ttggacttgt gtacccggaa acgccccagg cactaagtgc 1440
cacacatgtg tgtttactta aggtgatgtc agtatttata gggtgtggag tgcaataaaa 1500
tccgtgttga ctttaagtgt gtggtttatg actcaggggt ggggactgtg ggtatataag 1560
caggtgcaga cctgtgtggt cagttcagag caggactcat ggagatctgg acggtcttgg 1620
aagactttca ccagactaga cagctgctag agaactcatc ggaggaagtc tcttacctgt 1680
ggagattttg cttcggtggg gctctagcta agctagtcta tagggccaaa caggattata 1740
aggatcaatt tgaggatatt ttgagagagt gtcctagtat ttttgactct ctcaacttgg 1800
gccatcagtc tcactttaac cagagtattc tgagagccct tgacttttct actcctggca 1860
gaactaccgc cgcggtagcc ttttttgcct ttattcttga caaatggagt caagaaaccc 1920
atttcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt 1980
gccagcgcct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga 2040
tcctgagtct ccagtcaccc caggaacacc aacgccgcca gcagccgcag caggagcagc 2100
agcaagagga ggaggaggag gaggaccgag aagagaaccc gagagccggt ctggaccctc 2160
cggtggcgga ggaggaggag tagctgactt gtttcccgag ctgcgccggg tgctgactag 2220
gtcttccagt ggacgggaga gggggattaa gcgggagagg catgaggaga ctagtcacag 2280
aactgaactg actgtcagtc tgatgagccg caggcgccca gaatcggtgt ggtggcatga 2340
ggttcagtcg caggggatag atgaggtctc ggtaatgcat gagaaatatt ccctagaaca 2400
agtcaagact tgttggttgg agcccgagga tgattgggag gtagccatca ggaattatgc 2460
caagctggct ctgaggccag acaagaagta caagattacc aaactgatta atatcagaaa 2520
ttcctgctac atttcgggga atggggccga ggtggagatc agtacccagg agagggtggc 2580
cttcagatgt tgtatgatga atatgtaccc gggggtggtg ggcatggagg gagtcacctt 2640
tatgaacgcg aggtttaggg gtgatgggta taatggggtg gtctttatgg ccaacaccaa 2700
gctgacagtg cacggatgct ccttctttgg cttcaataac atgtgcatcg aggcctgggg 2760
cagtgtttca gtgaggggat gcagcttttc agccaactgg atgggggtcg tgggcagaac 2820
caagagcgtg gtgtcagtga agaaatgcct gttcgagagg tgccacctgg gggtgatgag 2880
cgagggcgaa gccaaagtca aacactgcgc ctctaccgag acgggctgct ttgtgatgat 2940
caagggcaat gccaaagtca agcataacat gatttgtggg gcctcggatg agcgcggcta 3000
ccagatgctg acctgtgccg gtgggaacag ccatatgctg gccaccgtgc atgtggcctc 3060
gcaccccgc aagacatggc ccgagttcga gcacaacgtc atgacccgct gcaatgtgca 3120
cctggggtcc cgccgaggca tgttcatgcc ctaccagtgc aacatgcaat ttgtgaaggt 3180
gctgctggag cccgatgcca tgtccagagt gagcctggtg ggggtgtttg acatgaatgt 3240
ggaggtgtgg aaaattctga gatatgatga atccaagacc aggtgccggg cctgcgaatg 3300

cggaggcaag cacgccaggc ttcagcccgt gtgtgtggag gtgacggagg acctgcgacc 3360
cgatcatttg gtgttgtcct gcaacgggac ggagttcggc tccagcgggg aagaatctga 3420
ctagagtgag tagtgtttgg gggtgggtgg gagtctgcat gatgggcaga atgactaaaa 3480
tctgtgtttt tctgcgcagc agcatgagcg gaagcgcctc ctttgaggga ggggtattca 3540
gcccttatct gacggggcgt ctcccctcct gggcgggagt gcgtcagaat gtgatgggat 3600
ccacggtgga cggccggccc gtgcagcccg cgaactcttc aaccctgacc tacgcgaccc 3660
tgagctcctc gtccgtggac gcagctgccg ccgcagctgc tgcttccgcc gccagcgccg 3720
tgcgcggaat ggccttgggc gccggctact acagctctct ggtggccaac tcgagttcca 3780
ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg 3840
ccctgaccca gcgcctgggc gagctgaccc agcaggtggc tcagctgcag gcggagacgc 3900
gggccgcggt tgccacggtg aaaaccaaat aaaaaatgaa tcaataaata aacggagacg 3960
gttgttgatt ttaacacaga gtcttgatct ttatttgatt tttcgcgcgc ggtaggccct 4020
ggaccaccgg tctcgatcat tgagcacccg gtggattttt tccaggaccc ggtagaggtg 4080
ggcttggatg ttgaggtaca tgggcatgag cccgtcccgg gggtggaggt agctccattg 4140
cagggcctcg tgctcggggg tggtgttgta aatcacccag tcatagcagg ggcgcagggc 4200
gtggtgctgc acgatgtcct tgaggaggag actgatggcc acgggcagcc ccttggtgta 4260
ggtgttgacg aacctgttga gctgggaggg atgcatgcgg ggggagatga gatgcatctt 4320
ggcctggatc ttgagattgg cgatgttccc gcccagatcc cgccgggggt tcatgttgtg 4380
caggaccacc agcacggtgt atccggtgca cttggggaat ttgtcatgca acttggaagg 4440
gaaggcgtga aagaatttgg agacgccctt gtgaccgccc aggttttcca tgcactcatc 4500
catgatgatg gcgatgggcc cgtgggcggc ggcctgggca aagacgtttc gggggtcgga 4560
cacatcgtag ttgtggtcct gggtgagctc gtcataggcc attttaatga atttggggcg 4620
gagagtgccc gactgggggа cgaaggtgcc ctcgatcccg ggggcgtagt tcccctcgca 4680
gatctgcatc tcccaggcct tgagctcgga ggggggggatc atgtccacct gcggggcgat 4740
gaaaaaaacg gtttccgggg cgggggagat gagctgggcc gaaagcaggt tccggagcag 4800
ctgggacttg ccgcagccgg tgggaccgta gatgaccccg atgaccggct gcaggtggta 4860
gttgagggag agacagctgc catcctcgcg gaggaggggg gccacctcgt tcatcatctc 4920
gcgcacatgc atgttctcgc gcacgagttc cgccaggagg cgctcgcccc ccagcgagag 4980
gagctcttgc agcgaggcga agtttttcag cggcttgagc ccgtcggcca tgggcatttt 5040
ggagagggtc tgttgcaaga gttccagacg gtcccagagc tcggtgatgt gctctagggc 5100
atctcgatcc agcagacctc ctcgtttcgc gggttggggc gactgcggga gtagggcacc 5160
aggcgatggg cgtccagcga ggccagggtc cggtccttcc agggtcgcag ggtccgcgtc 5220
agcgtggtct ccgtcacggt gaaggggtgc gcgccgggct gggcgcttgc gagggtgcgc 5280
ttcaggctca tccggctggt cgagaaccgc tcccggtcgg cgccctgcgc gtcggccagg 5340
tagcaattga gcatgagttc gtagttgagc gcctcggccg cgtggcccтt ggcgcggagc 5400
ttacctttgg aagtgtgtcc gcagacggga cagaggaggg acttgagggc gtagagcttg 5460
ggggcgagga agacggactc gggggcgtag gcgtccgcgc cgcagctggc gcagacggtc 5520
tcgcactcca cgagccaggt gaggtcgggg cggtcggggt caaaaacgag gtttcctccg 5580
tgctttttga tgcgtttctt acctctggtc tccatgagct cgtgtccccg ctgggtgaca 5640
aagaggctgt ccgtgtcccc gtagaccgac tttatgggcc ggtcctcgag cggggtgccg 5700
cggtcctcgt cgtagaggaa ccccgcccac tccgagacga aggcccgggt ccaggccagc 5760
acgaaggagg ccacgtggga ggggtagcgg tcgttgtcca ccagcgggtc caccttctcc 5820
agggtatgca agcacatgtc cccctcgtcc acatccagga aggtgattgg cttgtaagtg 5880
taggccacgt gaccgggggt cccggccggg ggggtataaa agggggcggg cccctgctcg 5940
tcctcactgt cttccggatc gctgtccagg agcgccagct gttggggtag gtattccctc 6000
tcgaaggcgg gcatgacctc ggcactcagg ttgtcagttt ctagaaacga ggaggatttg 6060
atattgacgg tgccgttgga gacgcctttc atgagcccct cgtccatctg gtcagaaaag 6120
acgatctttt tgttgtcgag cttggtggcg aaggagccgt agagggcgtt ggagaggagc 6180
ttggcgatgg agcgcatggt ctggttcttt tccttgtcgg cgcgctcctt ggcggcgatg 6240
ttgagctgca cgtactcgcg cgccacgcac ttccattcgg ggaagacggt ggtgagctcg 6300
tcgggcacga ttctgacccg ccagccgcgg ttgtgcaggg tgatgaggtc cacgctggtg 6360
gccacctcgc cgcgcagggg ctcgttggtc cagcagaggc gcccgccctt gcgcgagcag 6420
aaggggggca gcgggtccag catgagctcg tctggggggt cggcgtccac ggtgaagatg 6480
ccgggcagga gctcggggtc gaagtagctg atggaagtgg ccagatcgtc cagggaagct 6540
tgccagtcgc gcacggccag cgcgcgctcg taggggctga ggggcgtgcc ccagggcatg 6600

gggtgcgtga gcgcggaggc gtacatgccg cagatgtcgt agacgtagag gggctcctcg 6660
aggatgccga tgtaggtggg gtagcagcgc cccccgcgga tgctggcgcg cacgtagtcg 6720
tacagctcgt gcgagggcgc gaggagcccc gtgccgagat tggagcgctg cggcttttcg 6780
gcgcggtaga cgatctggcg gaagatggcg tgggagttgg aggagatggt gggcctctgg 6840
aagatgttga agtgggcgtg gggcaggccg accgagtccc tgatgaagtg ggcgtaggag 6900
tcctgcagct tggcgacgag ctcggcggtg acgaggacgt ccagggcgca gtagtcgagg 6960
gtctcttgga tgatgtcata cttgagctgg cccttctgct tccacagctc gcggttgaga 7020
aggaactctt cgcggtcctt ccagtactct tcgaggggga acccgtcctg atcggcacgg 7080
taagagccca ccatgtagaa ctggttgacg gccttgtagg cgcagcagcc cttctccacg 7140
gggagggcgt aagcttgcgc ggccttgcgc agggaggtgt gggtgagggc gaaggtgtcg 7200
cgcaccatga ccttgaggaa ctggtgcttg aagtcgaggt cgtcgcagcc gccctgctcc 7260
cagagttgga agtccgtgcg cttcttgtag gcggggttgg gcaaagcgaa agtaacatcg 7320
ttgaagagga tcttgcccgc gcggggcatg aagttgcgag tgatgcggaa aggctggggc 7380
acctcggccc ggttgttgat gacctgggcg gcgaggacga tctcgtcgaa gccgttgatg 7440
ttgtgcccga cgatgtagag ttccacgaat cgcgggcggc ccttgacgtg gggcagcttc 7500
ttgagctcgt cgtaggtgag ctcggcgggg tcgctgagcc cgtgctgttc gagggcccag 7560
tcggcgacgt gggggttggc gctgaggaag gaagtccaga gatccacggc cagggcggtc 7620
tgcaagcggt cccggtactg acggaactgc tggcccacgg ccatttttc gggggtgacg 7680
cagtagaagg tgcgggggtc gccgtgccag cggtcccact tgagctggag ggcgaggtcg 7740
tgggcgagct cgacgagcgg tgggtccccg gagagtttca tgaccagcat gaaggggacg 7800
agctgcttgc cgaaggaccc catccaggtg taggtttcca catcgtaggt gaggaagagc 7860
ctttcggtgc gaggatgcga gccgatgggg aagaactgga tctcctgcca ccagttggag 7920
gaatggctgt tgatgtgatg gaagtagaaa tgccgacggc gcgccgagca ctcgtgcttg 7980
tgtttataca agcgtccgca gtgctcgcaa cgctgcacgg gatgcacgtg ctgcacgagc 8040
tgtacctgag ttcctttgac gaggaatttc agtgggcagt ggagcgctgg cggctgcatc 8100
tggtgctgta ctacgtcctg gccatcggcg tggccatcgt ctgcctcgat ggtggtcatg 8160
ctgacgaggc cgcgcgggag gcaggtccag acctcggctc ggacgggtcg gagagcgagg 8220
acgagggcgc gcaggccgga gctgtccagg gtcctgagac gctgcggagt caggtcagtg 8280
ggcagcggcg gcgcgcggtt gacttgcagg agcttttcca gggcgcgcgg gaggtccaga 8340
tggtacttga tctccacggc gccgttggtg gcgacgtcca cggcttgcag ggtcccgtgc 8400
ccctggggcg ccaccaccgt gccccgtttc ttcttgggcg ctggttccat gccggtcaga 8460
agcggcggcg aggacgcgcg ccgggcggca ggggcggctc ggggcccgga ggcaggggcg 8520
gcaggggcac gtcggcgccg cgcgcgggca ggttctggta ctgcgcccgg agaagactgg 8580
cgtgagcgac gacgcgacgg ttgacgtcct ggatctgacg cctctgggtg aaggccacgg 8640
gacccgtgag tttgaacctg aaagagagtt cgacagaatc aatctcggta tcgttgacgg 8700
cggcctgccg caggatctct tgcacgtcgc ccgagttgtc ctggtaggcg atctcggtca 8760
tgaactgctc gatctcctcc tcctgaaggt ctccgcggcc ggcgcgctcg acggtggccg 8820
cgaggtcgtt ggagatgcgg cccatgagct gcgagaaggc gttcatgccg gcctcgttcc 8880
agacgcggct gtagaccacg gctccgtcgg ggtcgcgcgc gcgcatgacc acctgggcga 8940
ggttgagctc gacgtggcgc gtgaagaccg cgtagttgca gaggcgctgg tagaggtagt 9000
tgagcgtggt ggcgatgtgc tcggtgacga agaagtacat gatccagcgg cggagcggca 9060
tctcgctgac gtcgcccagg gcttccaagc gctccatggc ctcgtagaag tccacggcga 9120
agttgaaaaa ctgggagttg cgcgccgaga cggtcaactc ctcctccaga agacggatga 9180
gctcggcgat ggtggcgcgc acctcgcgct cgaaggcccc ggggggctcc tcttcttcca 9240
tctcctcctc ctcttcctcc tccactaaca tctcttctac ttcctcctca ggaggcggtg 9300
gcgggggagg ggccctgcgt cgccggcggc gcacgggcag acggtcgatg aagcgctcga 9360
tggtctcccc gcgccggcga cgcatggtct cggtgacggc gcgcccgtcc tcgcggggcc 9420
gcagcgtgaa gacgccgccg cgcatctcca ggtggccgcc gggggggtct ccgttgggca 9480
gggagagggc gctgacgatg catcttatca attggcccgt agggactccg cgcaaggacc 9540
tgagcgtctc gagatccacg ggatccgaaa accgctgaac gaaggcttcg agccagtcgc 9600
agtcgcaagg taggctgagc ccggtttctt cgggtatttg gtcgggaggc gggcgggcga 9660
tgctgctggt gatgaagttg aagtaggcgg tcctgagacg gcggatggtg gcgaggagca 9720
ccaggtcctt gggcccggct tgctggatgc gcagacggtc ggccatgccc caggcgtggt 9780
cctgacacct ggcgaggtcc ttgtagtagt cctgcatgag ccgctccacg ggcacctcct 9840
cctcgcccgc gcggccgtgc atgcgcgtga gcccgaaccc gcgctgcggc tggacgagcg 9900

ccaggtcggc gacgacgcgc tcggcgagga tggcctgctg gatctgggtg agggtggtct 9960
ggaagtcgtc gaagtcgacg aagcggtggt aggctccggt gttgatggtg tatgagcagt 10020
tggccatgac ggaccagttg acggtctggt ggccggggcg cacgagctcg tggtacttga 10080
ggcgcgagta ggcgcgcgtg tcgaagatgt agtcgttgca ggtgcgcacg aggtactggt 10140
atccgacgag gaagtgcggc ggcggctggc ggtagagcgg ccatcgctcg gtggcggggg 10200
cgccgggcgc gaggtcctcg agcatgaggc ggtggtagcc gtagatgtac ctggacatcc 10260
aggtgatgcc ggcggcggtg gtggaggcgc gcgggaactc gcggacgcgg ttccagatgt 10320
tgcgcagcgg caggaagtag ttcatggtgg ccgcggtctg gcccgtgagg cgcgcgcagt 10380
cgtggatgct ctatacgggc aaaaacgaaa gcggtcagcg gctcgactcc gtggcctgga 10440
ggctaagcga acgggttggg ctgcgcgtgt accccggttc gaatctcgaa tcaggctgga 10500
gccgcagcta acgtggtact ggcactcccg tctcgaccca agcctgcaca aaacctccag 10560
gatacggagg cgggtcgttt tgcaactttt tgaggccgga aatgaaacta gtaagcgcga 10620
aaagcggccg accgcgatgg ctcgctgccg tagtctggag aagaatcgcc agggttgcgt 10680
tgcggtgtgc cccggttcga ggccggccgg attccgcggc taacgagggc gtggctgccc 10740
cgtcgtttcc aagaccccta gccagccgac ttctccagtt acggagcgag cccctctttt 10800
gttttgtttg tttttgccag atgcatcccg tactgcggca gatgcgcccc caccaccctc 10860
caccgcaaca acagcccact ccacagccgg cgcttctgcc cccgccccag cagcagcaac 10920
ttccagccac gaccgccgcg gccgccgtga gcggggctgg acagacttct cagtatgacc 10980
acctggcctt ggaagagggc gaggggctgg cgcgcctggg ggcgtcgtcg ccggagcggc 11040
acccgcgcgt gcagatgaaa cgggacgctc gcgaggccta cgtgcccaag cagaacctgt 11100
tcagagacag gagcggcgag gagcccgagg agatgcgcgc ggcccggttc cacgcggggc 11160
gggagctgcg gcgcggcctg gaccgaaaga gggtgctgag ggacgaggat ttcgaggcgg 11220
acgagctgac ggggatcagc cccgcgcgcg cgcacgtggc cgcggccaac ctggtcacgg 11280
cgtacgagca gaccgtgaag gaggagagca acttccaaaa atccttcaac aaccacgtgc 11340
gcaccctgat cgcgcgcgag gaggtgaccc tgggcctgat gcacctgtgg gacctgctgg 11400
aggccatcgt gcagaacccc accagcaagc cgctgacggc gcagctgttc ctggtggtgc 11460
agcacagtcg ggacaacgag gcgttcaggg aggcgctgct gaatatcacc gagcccgagg 11520
gccgctggct cctggacctg gtgaacattc tgcagagcat cgtggtgcag gagcgcgggc 11580
tgccgctgtc cgagaagctg gcggccatca acttctcggt gctgagtctg ggcaagtact 11640
acgctaggaa gatctacaag accccgtacg tgcccataga caaggaggtg aagatcgacg 11700
ggttttacat gcgcatgacc ctgaaagtgc tgaccctgag cgacgatctg ggggtgtacc 11760
gcaacgacag gatgcaccgc gcggtgagcg ccagccgccg gcgcgagctg agcgaccagg 11820
agctgatgca cagcctgcag cgggccctga ccggggccgg gaccgagggg gagagctact 11880
ttgacatggg cgcggacctg cgctggcagc ccagccgccg ggccttggaa gctgccggcg 11940
gttcccccta cgtggaggag gtggacgatg aggaggagga gggcgagtac ctggaagact 12000
gatggcgcga ccgtattttt gctagatgca gcaacagcca ccgccgcctc ctgatcccgc 12060
gatgcgggcg gcgctgcaga gccagccgtc cggcattaac tcctcggacg attggaccca 12120
ggccatgcaa cgcatcatgg cgctgacgac ccgcaatccc gaagccttta gacagcagcc 12180
tcaggccaac cggctctcgg ccatcctgga ggccgtggtg ccctcgcgct cgaaccccac 12240
gcacgagaag gtgctggcca tcgtgaacgc gctggtggag aacaaggcca tccgcggcga 12300
cgaggccggg ctggtgtaca acgcgctgct ggagcgcgtg gcccgctaca acagcaccaa 12360
cgtgcagacg aacctggacc gcatggtgac cgacgtgcgc gaggcggtgt cgcagcgcga 12420
gcggttccac cgcgagtcga acctgggctc catggtggcg ctgaacgcct tcctgagcac 12480
gcagcccgcc aacgtgcccc ggggccagga ggactacacc aacttcatca gcgcgctgcg 12540
gctgatggtg gccgaggtgc cccagagcga ggtgtaccag tcggggccgg actacttctt 12600
ccagaccagt cgccagggct tgcagaccgt gaacctgagc caggctttca agaacttgca 12660
gggactgtgg ggcgtgcagg ccccggtcgg ggaccgcgcg acggtgtcga gcctgctgac 12720
gccgaactcg cgcctgctgc tgctgctggt ggcgcccttc acggacagcg gcagcgtgag 12780
ccgcgactcg tacctgggct acctgcttaa cctgtaccgc gaggccatcg ggcaggcgca 12840
cgtggacgag cagacctacc aggagatcac ccacgtgagc cgcgcgctgg gccaggagga 12900
cccgggcaac ctggaggcca ccctgaactt cctgctgacc aaccggtcgc agaagatccc 12960
gccccagtac gcgctgagca ccgaggagga gcgcatcctg cgctacgtgc agcagagcgt 13020
ggggctgttc ctgatgcagg agggggccac gcccagcgcc gcgctcgaca tgaccgcgcg 13080
caacatggag cccagcatgt acgcccgcaa ccgcccgttc atcaataagc tgatggacta 13140
cttgcatcgg gcggccgcca tgaactcgga ctactttacc aacgccatct tgaacccgca 13200

ctggctcccg ccgcccgggt tctacacggg cgagtatgac atgcccgacc ccaacgacgg 13260
gttcctgtgg gatgacgtgg acagcagcgt gttctcgccg cgccccgcca ccaccgtgtg 13320
gaagaaagag ggcggggacc ggcggccgtc ctcggcgctg tccggtcgcg cgggtgctgc 13380
cgcggcggtg cccgaggccg ccagcccctt cccgagcctg ccctttcgc tgaacagcgt 13440
gcgcagcagc gagctgggac ggctgacgcg gccgcgcctg ctgggcgagg aggagtacct 13500
gaacgactcc ttgttgaggc ccgagcgcga gaagaacttc cccaataacg ggatagagag 13560
cctggtggac aagatgagcc gctggaagac gtacgcgcac gagcacaggg acgagccgcg 13620
agctagcagc agcaccggcg cccgtagacg ccagcggcac gacaggcagc ggggactggt 13680
gtgggacgat gaggattccg ccgacgacag cagcgtgttg gacttgggtg ggagtggtgg 13740
tggtaacccg ttcgctcacc tgcgcccccg tatcgggcgc ctgatgtaag aatctgaaaa 13800
aataaaaaac ggtactcacc aaggccatgg cgaccagcgt gcgttcttct ctgttgtttg 13860
tagtagtatg atgaggcgcg tgtacccgga gggtcctcct ccctcgtacg agagcgtgat 13920
gcagcaggcg gtggcggcgg cgatgcagcc cccgctggag gcgccttacg tgcccccgcg 13980
gtacctggcg cctacggagg ggcggaacag cattcgttac tcggagctgg caccttgta 14040
cgataccacc cggttgtacc tggtggacaa caagtcggcg gacatcgcct cgctgaacta 14100
ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt tcacccccac 14160
ggaggccagc acccagacca tcaactttga cgagcgctcg cggtggggcg gccagctgaa 14220
aaccatcatg cacaccaaca tgcccaacgt gaacgagttc atgtacagca acaagttcaa 14280
ggcgcgggtc atggtctcgc gcaagacccc caacggggtc gcggtagggg atgattatga 14340
tggtagtcag gacgagctga cctacgagtg ggtggagttt gagctgcccg agggcaactt 14400
ctcggtgacc atgaccatcg atctgatgaa caacgccatc atcgacaatt acttggcggt 14460
gggacggcag aacggggtgc tggagagcga catcggcgtg aagttcgaca cgcgcaactt 14520
ccggctgggc tgggaccccg tgaccgagct ggtgatgccg ggcgtgtaca ccaacgaggc 14580
cttccacccc gacatcgtcc tgctgcccgg ctgcggcgtg gacttcaccg agagccgcct 14640
cagcaacctg ctgggcatcc gcaagcggca gcccttccag gagggcttcc agatcctgta 14700
cgaggacctg gaggggggca acatccccgc gctcttggat gtcgaagcct atgaagaaag 14760
taaggaaaaa gcagaggctg aggcaactgc agccgtggct actgccgctg tcaccgatgc 14820
agatgcagct actaccaggg gcgatacatt cgccactgtg gctgaagaag cagccgccgt 14880
agcggcgacc gatgatagtg aaagtaagat agtcatcaag ccggtggaga aggacagcaa 14940
gaacaggagc tacaacgttc tatcggatgg aaagaacacc gcctaccgca gctggtacct 15000
ggcctacaac tacggcgacc ccgagaaggg cgtgcgctcc tggacgctgc tcaccacctc 15060
ggacgtcacc tgcggcgtgg agcaagtcta ctggtcgctg cccgacatga tgcaagaccc 15120
ggtcaccttc cgctccacgc gtcaagttag caactacccg gtggtgggcg ccgagctcct 15180
gcccgtctac tccaagagct tcttcaacga gcaggccgtc tactcgcagc agctgcgcgc 15240
cttcacctcg ctcacgcacg tcttcaaccg cttccccgag aaccagatcc tcgtccgccc 15300
gcccgcgccc accattacca ccgtcagtga aaacgttcct gctctcacag atcacgggac 15360
cctgccgctg cgcagcagta tccggggagt ccagcgcgtg accgtcactg acgccagacg 15420
ccgcacctgc ccctacgtct acaaggccct gggcgtagtc gcgccgcgcg tcctctcgag 15480
ccgcaccttc taaaaaatgt ccattctcat ctcgcccagt aataacaccg gttggggcct 15540
gcgcgcgccc agcaagatgt acggaggcgc tcgccaacgc tccacgcaac accccgtgcg 15600
cgtgcgcggg cacttccgcg ctccctgggg cgccctcaag ggccgcgtgc gctcgcgcac 15660
caccgtcgac gacgtgatcg accaggtggt ggccgacgcg cgcaactaca cgcccgccgc 15720
cgcgcccgtc tccaccgtgg acgccgtcat cgacagcgtg gtggccgacg cgcgccggta 15780
cgcccgcgcc aagagccggc ggcggcgcat cgcccggcgg caccggagca cccccgccat 15840
gcgcgcggcg cgagccttgc tgcgcagggc caggcgcacg ggacgcaggg ccatgctcag 15900
ggcggccaga cgcgcggcct ctggcagcag cagcgccggc aggacccgca gacgcgcggc 15960
cacggcggcg gcggcggcca tcgccagcat gtcccgcccg cggcgcggca acgtgtactg 16020
ggtgcgcgac gccgccaccg gtgtgcgcgt gcccgtgcgc acccgccccc ctcgcacttg 16080
aagatgctga cttcgcgatg ttgatgtgtc ccagcggcga ggaggatgtc caagcgcaaa 16140
ttcaaggaag agatgctcca ggtcatcgcg cctgagatct acggccccgc ggcggcggtg 16200
aaggaggaaa gaaagccccg caaactgaag cgggtcaaaa aggacaaaaa ggaggaggaa 16260
gatgtggacg gactggtgga gtttgtgcgc gagttcgccc cccggcggcg cgtgcagtgg 16320
cgcgggcgga aagtgaaacc ggtgctgcga cccggcacca ccgtggtctt cacgcccggc 16380
gagcgttccg gctccgcctc caagcgctcc tacgacgagg tgtacgggga cgaggacatc 16440
ctcgagcagg cggccgagcg tctgggcgag tttgcttacg gcaagcgcag ccgccccgcg 16500

ccctttgaaag aggaggcggt gtccatcccg ctggaccacg gcaaccccac gccgagtctg 16560
aagccggtga ccctgcagca ggtgctgccg agcgcggcgc cgcgccgggg cttcaagcgc 16620
gagggcggcg aggatctgta cccgaccatg cagctgatgg tgcccaagcg ccagaagctg 16680
gaggacgtgc tggagcacat gaaggtggac cccgaggtgc agcccgaggt caaggtgcgg 16740
cccatcaagc aggtggcccc gggcctgggc gtgcagaccg tggacatcaa gatccccacg 16800
gagcccatgg aaacgcagac cgagcccgtg aagcccagca ccagcaccat ggaggtgcag 16860
acggatccct ggatgccggc gccggcttcc accaccacca ctcgccgaag acgcaagtac 16920
ggcgcggcca gcctgctgat gcccaactac gcgctgcatc cttccatcat ccccacgccg 16980
ggctaccgcg gcacgcgctt ctaccgcggc tacagcagcc gccgcaagac caccacccgc 17040
cgccgccgtc gtcgcacccg ccgcagcagc accgcgactt ccgccgcctt ggtgcggaga 17100
gtgtaccgca gcgggcgcga gcctctgacc ctgccgcgcg cgcgctacca cccgagcatc 17160
gccatttaac tctgccgtcg cctcctactt gcagatatgg ccctcacatg ccgcctccgc 17220
gtccccatta cgggctaccg aggaagaaag ccgcgccgta gaaggctgac ggggaacggg 17280
ctgcgtcgcc atcaccaccg gcggcggcgc gccatcagca agcggttggg gggaggcttc 17340
ctgcccgcgc tgatccccat catcgccgcg gcgatcgggg cgatccccgg catagcttcc 17400
gtggcggtgc aggcctctca gcgccactga gacacagctt ggaaaatttg taataaaaaa 17460
tggactgacg ctcctggtcc tgtgatgtgt gtttttagat ggaagacatc aatttttcgt 17520
ccctggcacc gcgacacggc acgcggccgt ttatgggcac ctggagcgac atcggcaaca 17580
gccaactgaa cgggggcgcc ttcaattgga gcagtctctg gagcgggctt aagaatttcg 17640
ggtccacgct caaaacctat ggcaacaagg cgtggaacag cagcacaggg caggcgctga 17700
gggaaaagct gaaagagcag aacttccagc agaaggtggt cgatggcctg gcctcgggca 17760
tcaacggggt ggtggacctg gccaaccagg ccgtgcagaa acagatcaac agccgcctgg 17820
acgcggtccc gcccgcgggg tccgtggaga tgccccaggt ggaggaggag ctgcctcccc 17880
tggacaagcg cggcgacaag cgaccgcgtc ccgatgcaga ggagacgctg ctgacgcaca 17940
cggacgagcc gcccccgtac gaggaggcgg tgaaactggg tctgcccacc acgcggcccg 18000
tggcgcctct ggccaccggg gtgctgaaac ccagcagcag cagccagccc gcgaccctgg 18060
acttgcctcc gcctgcttcc cgcccctcca cagtggctaa gcccctgccg ccggtggccg 18120
tcgcgtcgcg cgccccccga ggccgccccc aggcgaactg gcagagcact ctgaacagca 18180
tcgtgggtct gggagtgcag agtgtgaagc gccgccgctg ctattaaaag acactgtagc 18240
gcttaacttg cttgtctgtg tgtatatgta tgtccgccga ccagaaggaa gaggcgcgtc 18300
gccgagttgc aagatggcca ccccatcgat gctgccccag tgggcgtaca tgcacatcgc 18360
cggacaggac gcttcggagt acctgagtcc gggtctggtg cagttcgccc gcgccacaga 18420
cacctacttc agtctgggga acaagtttag gaaccccacg gtggcgccca cgcacgatgt 18480
gaccaccgac cgcagccagc ggctgacgct gcgcttcgtg cccgtggacc gcgaggacaa 18540
cacctactcg tacaaagtgc gctacacgct ggccgtgggc gacaaccgcg tgctggacat 18600
ggccagcacc tactttgaca tccgcggcgt gctggatcgg ggccccagct tcaaacccta 18660
ctccggcacc gcctacaaca gcctggctcc caagggagcg cccaacacct cacagtggat 18720
aaccaaagac aatggaactg ataagacata cagttttgga aatgctccag tcagaggatt 18780
ggacattaca gaagagggtc tccaaatagg acccgatgag tcaggggtg aaagcaagaa 18840
aatttttgca gacaaaacct atcagcctga acctcagctt ggagatgagg aatggcatga 18900
tactattgga gctgaagaca agtatggagg cagagcgctt aaacctgcca ccaacatgaa 18960
accctgctat gggtctttcg ccaagccaac taatgctaag ggaggtcagg ctaaaagcag 19020
aaccaaggac gatggcacta ctgagcctga tattgacatg gccttctttg acgatcgcag 19080
tcagcaagct agtttcagtc cagaacttgt tttgtatact gagaatgtcg atctggacac 19140
cccggatacc cacattattt acaaacctgg cactgatgaa acaagttctt ctttcaactt 19200
gggtcagcag tccatgccca acagacccaa ctacatcggc ttcagagaca actttatcgg 19260
tctcatgtac tacaacagta ctggcaatat gggtgtacta gctggacagg cctcccagct 19320
gaatgctgtg gtggacttgc aggacagaaa cactgaactg tcctaccagc tcttgcttga 19380
ctctctgggt gacagaacca ggtatttcag tatgtggaac caggcggtgg acagctacga 19440
ccccgatgtg cgcattattg aaaatcacgg tgtggaggat gaactaccca actattgctt 19500
ccctttgaat ggtgtgggct ttacagatac attccaggga attaaggtta aaactaccaa 19560
taacggaaca gcaaatgcta cagagtggga atctgatacc tctgtcaata atgctaatga 19620
gattgccaag ggcaatcctt tcgccatgga gatcaacatc caggccaacc tgtggcggaa 19680
cttcctctac gcgaacgtgg cgctgtacct gcccgactcc tacaagtaca cgccggccaa 19740
catcacgctg cccgccaaca ccaacaccta cgattacatg aacggccgcg tggtagcgcc 19800

ctcgctggtg gacgcctaca tcaacatcgg ggcgcgctgg tcgctggacc ccatggacaa 19860
cgtcaacccc ttcaaccacc accgcaacgc gggcctgcgc taccgctcca tgctcctggg 19920
caacgggcgc tacgtgccct tccacatcca ggtgccccaa aagtttttcg ccatcaagag 19980
cctcctgctc ctgcccgggt cctacaccta cgagtggaac ttccgcaagg acgtcaacat 20040
gatcctgcag agctccctcg gcaacgacct gcgcacggac ggggcctcca tcgccttcac 20100
cagcatcaac ctctacgcca ccttcttccc catggcgcac aacaccgcct ccacgctcga 20160
ggccatgctg cgcaacgaca ccaacgacca gtccttcaac gactacctct cggcggccaa 20220
catgctctac cccatcccgg ccaacgccac caacgtgccc atctccatcc cctcgcgcaa 20280
ctgggccgcc ttccgcggct ggtccttcac gcgcctcaag acccgcgaga cgccctcgct 20340
cggctccggg ttcgacccct acttcgtcta ctcgggctcc atcccctacc tcgacggcac 20400
cttctacctc aaccacacct tcaagaaggt ctccatcacc ttcgactcct ccgtcagctg 20460
gcccggcaac gaccgcctcc tgacgcccaa cgagttcgaa atcaagcgca ccgtcgacgg 20520
agaggggtac aacgtggccc agtgcaacat gaccaaggac tggttcctgg ttcagatgct 20580
ggcccactac aacatcggct accagggctt ctacgtgccc gagggctaca aggaccgcat 20640
gtactccttc ttccgcaact tccagcccat gagccgccag gtcgtggacg aggtcaacta 20700
caaggactac caggccgtca ccctggccta ccagcacaac aactcgggct tcgtcggcta 20760
cctcgcgccc accatgcgcc agggacagcc ctaccccgcc aactaccct acccgctcat 20820
cggcaagagc gccgtcgcca gcgtcaccca gaaaaagttc ctctgcgacc gggtcatgtg 20880
gcgcatcccc ttctccagca acttcatgtc catgggcgcg ctcaccgacc tcggccagaa 20940
catgctctac gccaactccg cccacgcgct agacatgaat ttcgaagtcg accccatgga 21000
tgagtccacc cttctctatg ttgtcttcga agtcttcgac gtcgtccgag tgcaccagcc 21060
ccaccgcggc gtcatcgagg ccgtctacct gcgcacgccc ttctcggccg gtaacgccac 21120
cacctaagcc ccgctcttgc ttcttgcaag atgacggcct gtgcgggctc cggcgagcag 21180
gagctcaggg ccatcctccg cgacctgggc tgcgggccct gcttcctggg caccttcgac 21240
aagcgcttcc cgggattcat ggccccgcac aagctggcct gcgccatcgt caacacggcc 21300
ggccgcgaga ccgggggcga gcactggctg gccttcgcct ggaacccgcg ctcccacacc 21360
tgctacctct tcgacccctt cgggttctcg aacgagcgcc tcaagcagat ctaccagttc 21420
gagtacgagg gcctgctgcg ccgcagcgcc ctggccaccg aggaccgctg cgtcaccctg 21480
gaaaagtcca cccagaccgt gcagggtccg cgctcggccg cctgcgggct cttctgctgc 21540
atgttcctgc acgccttcgt gcactggccc gaccgcccca tggacaagaa ccccaccatg 21600
aacttgctga cgggggtgcc caacggcatg ctccagtcgc cccaggtgga acccaccctg 21660
cgccgcaacc aggaagcgct ctaccgcttc ctcaacgccc actccgccta ctttcgctcc 21720
caccgcgcgc gcatcgagaa ggccaccgcc ttcgaccgca tgaatcaaga catgtaaacc 21780
gtgtgtgtat gtgaatgctt tattcataat aaacagcaca tgtttatgcc accttctctg 21840
aggctctgac tttatttaga aatcgaaggg gttctgccgg ctctcggcat gccccgcggg 21900
cagggatacg ttgcggaact ggtacttggg cagccacttg aactcgggga tcagcagctt 21960
gggcacgggg aggtcgggga acgagtcgct ccacagcttg cgcgtgagtt gcagggcgcc 22020
cagcaggtcg ggcgcggaga tcttgaaatc gcagttggga cccgcgttct gcgcgcgaga 22080
gttgcggtac acggggttgc agcactggaa caccatcagg gccgggtgct tcacgctcgc 22140
cagcaccgtc gcgtcggtga tgccctccac gtccagatcc tcggcgttgg ccatcccgaa 22200
gggggtcatc ttgcaggtct gccgccccat gctgggcacg cagccgggct tgtggttgca 22260
atcgcagtgc agggggatca gcatcatctg ggcctgctcg gagctcatgc ccgggtacat 22320
ggccttcatg aaagcctcca gctggcggaa ggcctgctgc gccttgccgc cctcggtgaa 22380
gaagaccccg caggacttgc tagagaactg gttggtggcg cagcccgcgt cgtgcacgca 22440
gcagcgcgcg tcgttgttgg ccagctgcac cacgctgcgc ccccagcggt tctgggtgat 22500
cttggcccgg tcggggttct ccttcagcgc gcgctgtccg ttctcgctcg ccacatccat 22560
ctcgatcgtg tgctccttct ggatcatcac ggtcccgtgc aggcaccgca gcttgctctc 22620
ggcctcggtg cacccgtgca gccacagcgc gcagccggtg ctctcccagt tcttgtgggc 22680
gatctgggag tgcgagtgca cgaagccctg caggaagcgg cccatcatcg cggtcagggt 22740
cttgttgctg gtgaaggtca gcgggatgcc gcggtgctcc tcgttcacat acaggtggca 22800
gatgcggcgg tacacctcgc cctgctcggg catcagctgg aaggcggact tcaggtcgct 22860
ctccacgcgg taccggtcca tcagcagcgt catgacttcc atgcccttct cccaggccga 22920
aacgatcggc aggctcaggg ggttcttcac cgttgtcatc ttagtcgccg ccgccgaggt 22980
caggggggtcg ttctcgtcca gggtctcaaa cactcgcttg ccgtccttct cgatgatgcg 23040
cacgggggggg aagctgaagc ccacggccgc cagctcctcc tcggcctgcc tttcgtcctc 23100

gctgtcctgg ctgatgtctt gcaaaggcac atgcttggtc ttgcggggtt tctttttggg 23160
cggcagaggc ggcggcggag acgtgctggg cgagcgcgag ttctcgctca ccacgactat 23220
ttcttcttct tggccgtcgt ccgagaccac gcggcggtag gcatgcctct tctggggcag 23280
aggcggaggc gacgggctct cgcggttcgg cgggcggctg gcagagcccc ttccgcgttc 23340
gggggtgcgc tcctggcggc gctgctctga ctgacttcct ccgcggccgg ccattgtgtt 23400
ctcctaggga gcaacaacaa gcatggagac tcagccatcg tcgccaacat cgccatctgc 23460
ccccgccgcc gacgagaacc agcagaatga aagcttaacc gccccgccgc ccagccccac 23520
ctccgacgcc gcggccccag acatgcaaga gatggaggaa tccatcgaga ttgacctggg 23580
ctacgtgacg cccgcggagc acgaggagga gctggcagcg cgcttttcag ccccggaaga 23640
gaaccaccaa gagcagccag agcaggaagc agagagcgag cagaaccagg ctgggctcga 23700
gcatggcgac tacctgagcg gggcagagga cgtgctcatc aagcatctga cccgccaatg 23760
catcatcgtc aaggacgcgc tgctcgaccg cgccgaggtg cccctcagcg tggcggagct 23820
cagccgcgcc tacgagcgca acctcttctc gccgcgcgtg ccccccaagc gccagcccaa 23880
cggcacctgc gagcccaacc cgcgcctcaa cttctacccg gtcttcgcgg tgcccgaggc 23940
cctggccacc taccacctct ttttcaagaa ccaaaggatc cccgtctcct gccgcgccaa 24000
ccgcacccgc gccgacgccc tgctcaacct gggcccccggc gcccgcctac ctgatatcac 24060
ctccttggaa gaggttccca agatcttcga gggtctgggc agcgacgaga ctcgggccgc 24120
gaacgctctg caaggaagcg gagaggaaca tgagcaccac agcgccctgg tggagttgga 24180
aggcgacaac gcgcgcctgg cggtgctcaa gcgcacggtc gagctgaccc acttcgccta 24240
cccggcgctc aacctgcccc ccaaggtcat gagcgccgtc atggaccagg tgctcatcaa 24300
gcgcgcctcg cccattgagg acatgcagga ccccgagagc tcggacgagg gcaagcccgt 24360
ggtcagcgac gagcagctgg cgcgctggct gggagcgagt agcacccccc agagcctgga 24420
agagcggcgc aagctcatga tggccgtggt cctggtgacc gtggagctgg agtgtctgcg 24480
ccgcttcttc gccgacgcag agaccctgcg caaggtcgag gagaacctgc actacctctt 24540
caggcacggg ttcgtgcgcc aggcctgcaa gatctccaac gtggagctga ccaacctggt 24600
ctcctacatg ggcatcctgc acgagaaccg cctggggcag aacgtgctgc acaccaccct 24660
gcgcggggag gcccgccgcg actacatccg cgactgcgtc tacctgtacc tctgccacac 24720
ctggcagacg ggcatgggcg tgtggcagca gtgcctggag gagcagaacc tgaaagagct 24780
ctgcaagctc ctgcagaaga acctcaaggc cctgtggacc gggttcgacg agcgcaccac 24840
cgcctcggac ctggccgacc tcatcttccc cgagcgcctg cggctgacgc tgcgcaacgg 24900
gctgcccgac tttatgagcc aaagcatgtt gcaaaacttt cgctctttca tcctcgaacg 24960
ctccgggatc ctgcccgcca cctgctccgc gctgccctcg gacttcgtgc cgctgacctt 25020
ccgcgagtgc cccccgccgc tctggagcca ctgctacttg ctgcgcctgg ccaactacct 25080
ggcctaccac tcggacgtga tcgaggacgt cagcggcgag ggtctgctgg agtgccactg 25140
ccgctgcaac ctctgcacgc cgcaccgctc cctggcctgc aacccccagc tgctgagcga 25200
gacccagatc atcggcacct tcgagttgca aggccccggc gaggagggca agggggtct 25260
gaaactcacc ccggggctgt ggacctcggc ctacttgcgc aagttcgtgc ccgaggacta 25320
ccatcccttc gagatcaggt tctacgagga ccaatcccag ccgcccaagg ccgagctgtc 25380
ggcctgcgtc atcacccagg gggccatcct ggcccaattg caagccatcc agaaatcccg 25440
ccaagaattt ctgctgaaaa agggccacgg ggtctacttg gacccccaga ccggagagga 25500
gctcaacccc agcttccccc aggatgcccc gaggaagcag caagaagctg aaagtggagc 25560
tgccgccgcc ggaggatttg gaggaagact gggagagcag tcaggcagag gaggagatgg 25620
aagactggga cagcactcag gcagaggagg acagcctgca agacagtctg gaggaggaag 25680
acgaggtgga ggaggaggag gcagaggaag aagcagccgc cgccagaccg tcgtcctcgg 25740
cggagaaagc aagcagcacg gataccatct ccgctccggg tcggggtcgc ggcggccggg 25800
cccacagtag gtgggacgag accgggcgct tcccgaaccc caccacccag accggtaaga 25860
aggagcggca gggatacaag tcctggcggg ggcacaaaaa cgccatcgtc tcctgcttgc 25920
aagcctgcgg gggcaacatc tccttcaccc ggcgctacct gctcttccac cgcggggtga 25980
acttcccccg caacatcttg cattactacc gtcacctcca cagcccctac tactgtttcc 26040
aagaagaggc agaaacccag cagcagcaga aaaccagcga cagcggcagc agctagaaaa 26100
tccacagcgg caggtggact gaggatcgcg gcgaacgagc cggcgcagac ccgggagctg 26160
aggaaccgga tctttcccac cctctatgcc atcttccagc agagtcgggg gcaggagcag 26220
gaactgaaag tcaagaaccg ttctctgcgc tcgctcaccc gcagttgtct gtatcacaag 26280
agcgaagacc aacttcagcg cactctcgag gacgccgagg ctctcttcaa caagtactgc 26340
gcgctcactc ttaaagagta gcccgcgccc gcccacacac ggaaaaaggc gggaattacg 26400

tcaccacctg cgcccttcgc ccgaccatca tcatgagcaa agagattccc acgccttaca 26460
tgtggagcta ccagccccag atgggtctgg ccgccggcgc cgcccaggac tactccaccc 26520
gcatgaactg gctcagtgcc gggcccgcga tgatctcacg ggtgaatgac atccgcgccc 26580
atcgaaacca gatactccta gaacagtcag cgatcaccgc cacgccccgc catcacctta 26640
atccgcgtaa ttggcccgcc gccctggtgt accaggaaat tccccagccc acgaccgtac 26700
tacttccgcg agacgcccag gccgaagtcc agctgactaa ctcaggtgtc cagctggccg 26760
gcggcgccgc cctgtgtcgt caccgccccg ctcagggtat aaagcggctg gtgatccgag 26820
gcagaggcac acagctcaac gacgaggtgg tgagctcttc gctgggtctg cgacctgacg 26880
gagtcttcca actcgccgga tcggggagat cttccttcac gcctcgtcag gccgtcctga 26940
ctttggagag ttcgtcctcg cagccccgct cgggcggcat cggcactctc cagttcgtgg 27000
aggagttcac tccctcggtc tacttcaacc ccttctccgg ctcccccggc cactacccgg 27060
acgagttcat cccgaacttc gacgccatca gcgagtcggt ggacggctac gattgaatgt 27120
cccatggtgg cgcagctgac ctagctcggc ttcgacacct ggaccactgc cgccgcttcc 27180
gctgcttcgc tcgggatctc gccgagtttg cctactttga gctgcccgag gagcaccctc 27240
agggcccggc ccacggagtg cggatcatca tcgaaggggg cctcgactcc cacctgcttc 27300
ggatcttcag ccagcgaccg atcctggtcg agcgcgagca aggacagacc cgtctgaccc 27360
tgtactgcat ctgcaaccac cccggcctgc atgaaagtct ttgttgtctg ctgtgtactg 27420
agtataataa aagctgagat cagcgactac tccggactcg attgtggtgt tcctgctatc 27480
aaccggtccc tgttcttcac cgggaacgag accgagctcc agctccagtg taagccccac 27540
aagaagtacc tcacctggct gttccagggc tctccgatcg ccgttgtcaa ccactgcgac 27600
aacgacggag tcctgctgag cggccctgcc aaccttactt tttccacccg cagaagcaag 27660
ctccagctct tccaaccctt cctccccggg acctatcagt gcgtctcggg accctgccat 27720
cacaccttcc acctgatccc gaataccaca gcgccgctcc ccgctactaa caaccaaact 27780
acccaccaac gccaccgtcg cgacctttcc tctgaatcta atactaccac ccacaccgga 27840
ggtgagctcc gaggtcgacc aacctctggg atttactacg gcccctggga ggtggtgggg 27900
ttaatagcgc taggcctagt tgtgggtggg cttttggctc tctgctacct atacctccct 27960
tgctgttcgt acttagtggt gctgtgttgc tggtttaaga aatggggaag atcaccctag 28020
tgagctgcgg tgtgctggtg gcggtgttgc tttcgattgt gggactgggc ggcgcggctg 28080
tagtgaagga ggagaaggcc gatccctgct tgcatttcaa tcccgacaaa tgccagctga 28140
gttttcagcc cgatggcaat cggtgcacgg tgctgatcaa gtgcggatgg gaatgtgaga 28200
acgtgagaat cgagtacaat aacaagactc ggaacaatac tctcgcgtcc gtgtggcaac 28260
ccggggaccc cgagtggtac accgtctctg tccccggtgc tgacggctcc ccgcgcaccg 28320
tgaataatac tttcattttt gcgcacatgt gcgacacggt catgtggatg agcaagcagt 28380
acgatatgtg gccccccacg aaggagaaca tcgtggtctt ctccatcgct tacagcgtgt 28440
gcacggcgct aatcaccgct atcgtgtgcc tgagcattca catgctcatc gctattcgcc 28500
ccagaaataa tgccgaaaaa gagaaacagc cataacacgt tttttcacac acctttttca 28560
gaccatggcc tctgttactg ccctaattat ttttttgggt ctcgtgggca ctagcagcac 28620
ttttcagcat ataaacaaaa ctgtttatgc tggttctaat tctgtattac ctgggcatca 28680
atcacaccag aaagtttcat ggtactggta tgataaaaat aacacgccag tcacactctg 28740
caagggtcat caaacaccca taaaccgtag tggaattttt tttaaatgta atcataataa 28800
tattacacta ctttcaatta caaagcacta ttctggtact tactatggaa ccaattttaa 28860
cataaaacag gacacttact atagtgtcac agtattggat ccaactactc ctagaacaac 28920
tacaaaaccc acaactacta agaggcacac taaacctaaa actaccaaga aaaccactgt 28980
caaaacaaca actaggacca ccacaactac agaggctacc accagcacaa cacttgctgc 29040
aactacacac acacacactg agctaacctt acagaccact aatgatttga tagccctgtt 29100
gcaaaagggg gataacagca ccacttccga tgaggaaata cccaaatcca tgattggcat 29160
tattgttgct gtagtggtgt gcatgttgat catcgccttg tgcatggtgt actatgcctt 29220
ctgctacaga aagcacagac tgaacgacaa gctggaacac ttactaagtg ttgaatttta 29280
atttttaga accatgaaga tcctaggcct tttagttttt tctatcatta cctctgctct 29340
ttgtgaatca gtgaataaag atgttactat taccactggt tctaattata cactgaaagg 29400
gccaccctca ggtatgcttt cgtggtattg ctattttgga actgacactg atcaaactga 29460
attatgcaat tttcaaaaag gcaaaacctc aaactctaaa atctctaatt atcaatgcaa 29520
tggcactgat ctgatactac tcaatgtcac gaaagcatat ggtggcagtt attcttgccc 29580
tggacaaaac actgaagaaa tgatttttta caaagtggaa gtggttgatc ccactactcc 29640
acccaccacc acaactactc acaccacaca cacagaacaa accacagcag aggaggcagc 29700

aaagttagcc ttgcaggtcc aagacagttc atttgttggc attaccccta cacctgatca 29760
gcggtgtccg gggctgctag tcagcggcat tgtcggtgtg ctttcgggat tagcagtcat 29820
aatcatctgc atgttcattt ttgcttgctg ctatagaagg ctttaccgac aaaaatcaga 29880
cccactgctg aacctctatg tttaattttt tccagagcca tgaaggcagt tagcactcta 29940
gtttttgtt ctttgattgg cattgttttt agtgctgggt ttttgaaaaa tcttaccatt 30000
tatgaaggtg agaatgccac tctagtgggc atcagtggtc aaaatgtcag ctggctaaaa 30060
taccatctag atgggtggaa agacatttgc gattggaatg tcactgtgta tacatgtaat 30120
ggagttaacc tcaccattac taatgccacc caagatcaga atggtaggtt taagggtcag 30180
agtttcacta gaaataatgg gtatgaatcc cataacatgt ttatctatga cgtcactgtc 30240
atcagaaatg agaccgccac caccacacag atgcccacta cacacagttc taccactact 30300
accaagcaaa ccacacagac aaccactttt tatacatcaa ctcagcatat gaccaccact 30360
acagcagcaa agccaagtag cgcagcgcct cagccacagg ctttggcttt gaaagctgca 30420
caacctagta caactactaa gaccaatgag cagactactg atttttgtc cactgtcgag 30480
agccacacca cagctacctc cagtgccttc tctagcaccg ccaatctctc ctcgctttcc 30540
tctacaccaa tcagtcccgc tactactcct agccccgctc ctcttcccac tcccctgaag 30600
caaacagacg gcggcatgca atggcagatc accctgctca ttgtgatcgg gttggtcatc 30660
ctggccgtgt tgctctacta catcttctgc cgccgcattc ccaacgcgca ccgcaagccg 30720
gtctacaagc ccatcgttgt cgggcagccg gagccgcttc aggtggaagg gggtctaagg 30780
aatcttctct tctcttttac agtatggtga ttgaactatg attcctagac aattcttgat 30840
cactattctt atctgcctcc tccaagtctg tgccaccctc gctctggtgg ccaacgccag 30900
tccagactgt attgggccct tcgcctccta cgtgctcttt gccttcatca cctgcatctg 30960
ctgttgtagc atagtctgcc tgcttatcac cttcttccag ttcattgact ggatctttgt 31020
gcgcatcgcc tacctgcgcc accaccccca gtaccgcgac cagcgagtgg cgcgactgct 31080
caggctcctc tgataagcat gcgggctctg ctacttctcg cgcttctgct gttagtgctc 31140
ccccgtcccg tcgacccccg gtcccccgag gaggtccgca aatgcaaatt ccaagaaccc 31200
tggaaattcc tcaaatgcta ccgccaaaaa tcagacatgc atcccagctg gatcatgatc 31260
attgggatcg tgaacattct ggcctgcacc ctcatctcct ttgtgattta cccctgcttt 31320
gactttggtt ggaactcgcc agaggcactc tatctcccgc ctgagcctga cacaccacca 31380
cagcagcaac ctcaggcaca cgcactacca ccaccacagc ctaggccaca atacatgccc 31440
atattagact atgaggccga gccacagcga cccatgctcc ccgctattag ttacttcaat 31500
ctaaccggcg gagatgactg acccactggc caacaacaac gtcaacgacc ttctcctgga 31560
catggacggc cgcgcctcgg agcagcgact cgcccaactc cgcatccgcc agcagcagga 31620
gagagccgtc aaggagctgc aggatgcggt ggccatccac cagtgcaaga aaggcatctt 31680
ctgcctggtg aagcaggcca agatctccta cgaggtcacc cagaccgacc atcgcctctc 31740
ctacgagctc ctgcagcagc gccagaagtt cacctgcctg gtcggagtca accccatcgt 31800
catcacccag cagtcgggcg ataccaaggg gtgcatccac tgctcctgcg actcccccga 31860
gtgcgttcac accatgatca agaccctctg cggcctccgc gacctcctcc ccatgaacta 31920
atcacccccc tatccagtga aataaagatc atattgatga tgatttaaat aaaaaaataa 31980
tcatttgatt tgaaataaag atacaatcat attgatgatt tgagtttaac aaaaataaag 32040
aatcacttac ttgaaatctg ataccaggtc tctgtccatg ttttctgcca acaccacctc 32100
actccctct tcccagctct ggtactgcag gccccggcgg gctgcaaact tcctccacac 32160
gctgaagggg atgtcaaatt cctcctgtcc ctcaatcttc attttctctt ctatcagatg 32220
tccaaaaagc gcgcgcgggt ggatgatgac ttcgaccccg tgtaccccta cgatgcagac 32280
aacgcaccga ctgtgccctt catcaaccct cccttcgtct cttcagatgg attccaagaa 32340
aagccctgg gggtgttgtc cctgcgactg gccgatcccg tcaccaccaa gaacggggct 32400
gtcaccctca agctgggggga gggggtggac ctcgacgact cgggaaaact catctccaaa 32460
aatgccacca aggccactgc ccctctcagt atttccaaca acaccatttc ccttaacatg 32520
gataccctc tttacaacaa caatggaaag ctaggtatga aggtaaccgc accattaaag 32580
atattagaca cagatctact aaaaacactt gttgttgctt atgggcaggg attaggaaca 32640
aacaccaatg gtgctcttgt tgcccaacta gcatacccac ttgtttttaa taccgctagc 32700
aaaattgccc ttaatttagg caatggacca ttaaaagtgg atgcaaatag actgaacatt 32760
aattgcaaaa gaggtatcta tgtcactacc acaaaagatg cactggagat taatatcagt 32820
tgggcaaatg ctatgacatt tataggaaat gccattggtg tcaatattga cacaaaaaaa 32880
ggcctacagt tcggcacttc aagcactgaa acagatgtta aaaatgcttt tccactccaa 32940
gtaaaacttg gagctggtct tacatttgac agcacaggtg ccattgttgc ttggaacaaa 33000

gaagatgaca aacttacact gtggaccaca gccgatccat ctccaaactg tcacatatat 33060
tctgcaaagg atgctaagct tacactctgc ttgacaaagt gtggtagtca gatactgggc 33120
actgtttctc tcatagctgt tgatactggt agcttaaatc caataacagg aaaagtaacc 33180
actgctcttg tttcacttaa attcgatgcc aatggagttt tgcaagccag ttcaacacta 33240
gataaagaat attggaattt cagaaaagga gatgtgacac ctgctgaccc ctacactaat 33300
gctataggct ttatgcccaa ccttaatgca tacccaaaaa acacaaacgc agctgcaaaa 33360
agtcacattg ttggaaaagt atacctacat ggggatgaaa gcaagccact agacttgata 33420
attacattta atgaaaccag tgatgaatcc tgtacttatt gcattaactt tcagtggcag 33480
tggggaactg accaatataa agatgaaaca cttgcagtca gttcattcac cttctcatac 33540
attgctaaag aataacatcc accctgcatg ccaacccatt tccctctatc tatacatgga 33600
aaactctgaa gcagaaaaaa taaagttcaa gtgttttatt gattcaacag tttttacaga 33660
attcgagtag ttatttttccc tccaccctcc caactcatgg aatacaccat cctctcccca 33720
cgcacagcct taaacatctg aatgccattg gtaatggaca tggttttggc ctccacattc 33780
cacacagttt cagagcgagc cagtctcggg tcggtcaggg agatgaaacc ctccgggcac 33840
tcctgcatct gcacctcaca gttcaacagc tgagggctgt cctcggtggt cgggatcaca 33900
gttatctgga agaagagcga tgagagtcat aatccgcgaa cgggatcggg cggttgtggc 33960
gcatcaggcc ccgcagcagt cgctgtctgc gccgctccgt caagctgctg ctcaaggggt 34020
ccgggtccag ggactccccg cgcatgatgc cgatggccct gagcatcagt cgcctggtgc 34080
ggcgggcgca gcagcggatg cggatctcac tcaggtcgga acagtacgtg cagcacagca 34140
ctaccaagtt gttcaacagt ccatagttca acgtgctcca gccaaaactc atctgtggaa 34200
ctatgctgcc cacatgtcca tcgtaccaga tcctgatgta aatcaggtgg cgccccctcc 34260
agaacacact gcccatgtac atgatctcct tgggcatgtg caggttcacc acctcccggt 34320
accacatcac ccgctggttg aacatgcagc cccggatgat cctgcggaac cacagggcca 34380
gcaccgcccc gcccgccatg cagcgcaggg accccgggtc ctggcaatgg cagtggatga 34440
tccaccgctc gtacccgtgg atcatctggg agctgaacaa gtctatgttg gcacagcaca 34500
ggcacacgct catgcatctc ttcagcactc tcagctcctc gggggtcaaa accatatccc 34560
agggtacggg gaactcttgc aggacagcga accccgcaga acagggcaaa cctcgcacag 34620
aacttacatt gtgcatggac agggtatcgc aatcaggcag caccgggtga tcctccacca 34680
gggaagcgcg ggtctcgatt tcctcacagc gtggtaaggg ggccggtcga tacgggtgat 34740
ggcgggacgc ggctgatcgt gttcgcgatc gtgtcatgat gcagttgctt tcggacattt 34800
tcgtacttgc tatagcagaa cctggtccgg gcgctgcaca ccgatcgccg gcggcggtct 34860
cggcgcttgg aacgctccgt gttgaaattg taaaacagcc actctctcag accgtgcagc 34920
agatctaggg cctcaggagt gatgaagatc ccatcatgcc tgatggctct gatcacatcg 34980
accaccgtgg aatgggccag acccagccag atgatgcaat tttgttgggt ttcggtgacg 35040
gcgggggagg gaagaacagg aagaaccatg attaacttta atccaaacgg tctcggagca 35100
cttcaaaatg aaggtcgcgg agatggcacc tctcgccccc gctgtgttgg tggaaaataa 35160
cagccaggtc aaaggtgata cggttctcga gatgttccac ggtggcttcc agcaaagcct 35220
ccacgcgcac atccagaaac aagacaatag cgaaagcggg agggttctct aattcctcaa 35280
tcatcatgtt acactcctgc accatcccca gataattttc atttttccag ccttgaatga 35340
ttcgaactag ttcctgaggt aaatccaagc cagccatgat aaagagctcg cgcagagcgc 35400
cctccaccgg cattcttaag cacaccctca taattccaag atattctgct cctggttcac 35460
ctgcagcaga ttgacaagcg gaatatcaaa ctctctgccg cgatccctaa gctcctccct 35520
cagcaataac tgtaagtact ctctcatatc ctctccgaaa tttttagcca taggaccgcc 35580
aggaataaga ttagggcaag ccacagtaca gataaaccga agtcctcccc agtgagcatt 35640
gccaaatgca agactgctat aagcatgctg gctagacccg gtgatatctt ccagataatt 35700
ggacagaaaa tcgcccaggc aatttttaag aaaatcaaca aaagaaaaat cctccaggtg 35760
cacgtttaga gcctcgggaa caacgatgga gtaaatgcaa gcggtgcgtt ccagcatggt 35820
tagttagctg atctgtagaa aaaacaaaaa tgaacattaa accatgctag cctggcgaac 35880
aggtgggtaa atcgttcttt ccagcaccag gcaggccacg gggtctccgg cgcgaccctc 35940
gtaaaaattg tcgctatgat tgaaaaccat cacagagaga cgttcccggt ggccggcgtg 36000
aatgattcga caagacgaat acacccccgg aacattggcg tccgcgagtg aaaaaaagcg 36060
cccgaggaag caataaggca ctacaatgct cagtctcaag tccagcaaag cgatgccatg 36120
cggatgaagc acaaaattct caggtgcgta caaaatgtaa ttactccct cctgcacagg 36180
cagcaaagcc cccgatccct ccaggtacac atacaaagcc tcagcgtcca tagcttaccg 36240
agcagcagcg gcacacaaca ggcgcaaaag tcagagaaag gctgagagct ctaacctgtc 36300

```
cacccgctct ctgctcaata tatagcccag atctacactg acgtaaaggc caaagtctaa 36360
aaatacccgc caaataatca cacacgccca gcacacgccc agaaaccggt gacacactca 36420
gaaaaatacg cgcacttcct caaacgccca aactgccgtc atttccgggt tcccacgcta 36480
cgtcatcaaa attcaacttt caaattccgt cgaccgttaa aaacgtcacc cgccccgccc 36540
ctaacggtcg ccgctcccgc agccaatcag cgccccgcat ccccaaattc aaacggctca 36600
tttgcatatt aacgcgcacc aaaagtttga ggtatattat tgatgatg              36648
```

<210> 3
<211> 36606
<212> DNA
<213> Chimpanzee Pan 6 (CV32) Genomic

<400> 3

```
catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg agctgtttga 60
atttggggag ggaggaaggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc tatgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
gtgacgaccc tccagagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgtatg 720
atctggaggt ggatgtgccc gagagcgacc ctaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgcg gtgaaccagg gagtgaaaac tgcgggcgag agctttagcc 1020
tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
acagtaagtg tgattaactt tagttgggaa ggcagagggt gactgggtgc tgactggttt 1200
atttatgtat atgttttttt atgtgtaggt cccgtctctg acgtagatga gacccccact 1260
tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
agaccagttg cagtgagagt caccgggcgg agagcagctg tggagagttt ggatgacttg 1380
ctacagggtg gggatgaacc tttggacttg tgtacccgga aacgccccag gcactaagtg 1440
ccacacatgt gtgtttactt aaggtgatgt cagtatttat agggtgtgga gtgcaataaa 1500
atccgtgttg actttaagtg cgtgttttat gactcagggg tggggactgt gggtatataa 1560
gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gactgtcttg 1620
gaagactttc accagactag acagttgcta gagaactcat cggagggagt ctcttacctg 1680
tggagattct gcttcggtgg gcctctagct aagctagtct atagggccaa acaggattat 1740
aaggaacaat ttgaggatat tttgagagag tgtcctggta tttttgactc tctcaacttg 1800
ggccatcagt ctcactttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc 1860
agaactaccg ccgcggtagc ctttttgcc tttattcttg acaaatggag tcaagaaacc 1920
catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg 1980
tgccagcgcc tgaatgcaat ctccggctac ttgccagtac agccggtaga cacgctgagg 2040
atcctgagtc tccagtcacc ccaggaacac caacgccgcc agcagccgca gcaggagcag 2100
cagcaagagg aggaccgaga agagaacccg agagccggtc tggaccctcc ggtggcggag 2160
gaggaggagt agctgacttg tttcccgagc tgcgccgggt gctgactagg tcttccagtg 2220
gacgggagag ggggattaag cgggagaggc atgaggagac tagccacaga actgaactga 2280
ctgtcagtct gatgagccgc aggcgcccag aatcggtgtg gtggcatgag gtgcagtcgc 2340
aggggataga tgaggtctcg gtgatgcatg agaaatattc cctagaacaa gtcaagactt 2400
gttggttgga gcccgaggat gattgggagg tagccatcag gaattatgcc aagctggctc 2460
tgaagccaga caagaagtac aagattacca aactgattaa tatcagaaat tcctgctaca 2520
```

tttcagggaa tggggccgag gtggagatca gtacccagga gagggtggcc ttcagatgtt 2580
gtatgatgaa tatgtacccg ggggtggtgg gcatggaggg agtcaccttt atgaacacga 2640
ggttcagggg tgatgggtat aatggggtgg tctttatggc caacaccaag ctgacagtgc 2700
acggatgctc cttctttggc ttcaataaca tgtgcatcga ggcctggggc agtgtttcag 2760
tgaggggatg cagcttttca gccaactgga tgggggtcgt gggcagaacc aagagcaagg 2820
tgtcagtgaa gaaatgcctg ttcgagaggt gccacctggg ggtgatgagc gagggcgaag 2880
ccaaagtcaa acactgcgcc tctaccgaga cgggctgctt tgtgctgatc aagggcaatg 2940
cccaagtcaa gcataacatg atctgtgggg cctcggatga gcgcggctac cagatgctga 3000
cctgcgccgg tgggaacagc catatgctgg ccaccgtgca tgtggcctcg caccccgca 3060
agacatggcc cgagttcgag cacaacgtca tgacccgctg caatgtgcac ctgggctccc 3120
gccgaggcat gttcatgccc taccagtgca acatgcaatt tgtgaaggtg ctgctggagc 3180
ccgatgccat gtccagagtg agcctgacgg gggtgtttga catgaatgtg gagctgtgga 3240
aaattctgag atatgatgaa tccaagacca ggtgccgggc ctgcgaatgc ggaggcaagc 3300
acgccaggct tcagcccgtg tgtgtggagg tgacggagga cctgcgaccc gatcatttgg 3360
tgttgtcctg caacgggacg gagttcggct ccagcgggga agaatctgac tagagtgagt 3420
agtgtttggg gctgggtgtg agcctgcatg aggggcagaa tgactaaaat ctgtggtttt 3480
ctgtgtgttg cagcagcatg agcggaagcg cctcctttga gggaggggta ttcagccctt 3540
atctgacggg gcgtctcccc tcctgggcgg gagtgtgtca gaatgttatg gnatccacgg 3600
tggacggccg gcccgtgcag cccgcgaact cttcaaccct gacctacgcg accctgagct 3660
cctcgtccgt ggacgcagct gccgccgcag ctgctgcttc cgccgccagc gccgtgcgcg 3720
gaatggccct gggcgccggc tactacagct ctctggtggc caactcgagt tccaccaata 3780
atcccgccag cctgaacgag gagaagctgc tgctgctgat ggcccagctc gaggccctga 3840
cccagcgcct gggcgagctg acccagcagg tggctcagct gcaggcggag acgcgggccg 3900
cggttgccac ggtgaaaacc aaataaaaaa tgaatcaata aataaacgga gacggttgtt 3960
gattttaaca cagagtcttg aatctttatt tgatttttcg cgcgcggtag gccctggacc 4020
accggtctcg atcattgagc acccggtgga tcttttccag gacccggtag aggtgggctt 4080
ggatgttgag gtacatgggc atgagcccgt cccgggggtg gaggtagctc cattgcaggg 4140
cctcgtgctc ggggatggtg ttgtaaatca cccagtcata gcaggggcgc agggcgtggt 4200
gctgcacgat gtccttgagg aggagactga tggccacggg cagccccttg gtgtaggtgt 4260
tgacgaacct gttgagctgg gagggatgca tgcgggggga gatgagatgc atcttggcct 4320
ggatcttgag attggcgatg ttcccgccca gatcccgccg ggggttcatg ttgtgcagga 4380
ccaccagcac ggtgtatccg gtgcacttgg ggaatttgtc atgcaacttg gaagggaagg 4440
cgtgaaagaa tttggagacg cccttgtgac cgcccaggtt ttccatgcac tcatccatga 4500
tgatggcgat gggcccgtgg gcggcggcct gggcaaagac gtttcggggg tcggacacat 4560
cgtagttgtg gtcctgggtg agctcgtcat aggccatttt aatgaatttg gggcggaggg 4620
tgcccgactg ggggacgaag gtgccctcga tcccgggggc gtagttgccc tcgcagatct 4680
gcatctccca ggccttgagc tcggaggggg ggatcatgtc cacctgcggg gcgatgaaaa 4740
aaacggtttc cggggcgggg gagatgagct gggccgaaag caggttccgg agcagctggg 4800
acttgccgca accggtgggg ccgtagatga ccccgatgac cggctgcagg tggtagttga 4860
gggagagaca gctgccgtcc tcgcggagga ggggggccac ctcgttcatc atctcgcgca 4920
catgcatgtt ctcgcgcacg agttccgcca ggaggcgctc gccccccagc gagaggagct 4980
cttgcagcga ggcgaagttt ttcagcggct tgagtccgtc ggccatgggc attttggaga 5040
gggtctgttg caagagttcc agacggtccc agagctcggt gatgtgctct agggcatctc 5100
gatccagcag acctcctcgt ttcgcgggtt ggggcgactg cgggagtagg gcaccaggcg 5160
atgggcgtcc agcgaggcca gggtccggtc cttccagggc cgcagggtcc gcgtcagcgt 5220
ggtctccgtc acggtgaagg ggtgcgcgcc gggctgggcg cttgcgaggg tgcgcttcag 5280
gctcatccgg ctggtcgaga accgctcccg gtcggcgccc tgcgcgtcgg ccaggtagca 5340
attgagcatg agttcgtagt tgagcgcctc ggccgcgtgg cccttggcgc ggagcttacc 5400
tttggaagtg tgtccgcaga cgggacagag gagggacttg agggcgtaga gcttgggggc 5460
gaggaagacg gactcggggg cgtaggcgtc cgcgccgcag ctggcgcaga cggtctcgca 5520
ctccacgagc caggtgaggt cggggcggtt ggggtcaaaa acgaggtttc ctccgtgctt 5580
tttgatgcgt ttcttacctc tggtctccat gagctcgtgt ccccgctggg tgacaaagag 5640
gctgtccgtg tccccgtaga ccgactttat gggccggtcc tcgagcgggg tgccgcggtc 5700
ctcgtcgtag aggaaccccg cccactccga gacgaaggcc cgggtccagg ccagcacgaa 5760
ggaggccacg tgggaggggt agcggtcgtt gtccaccagc gggtccacct tctccagggt 5820

atgcaagcac atgtcccccc cgtccacatc caggaaggtg attggcttgt aagtgtaggc 5880
cacgtgaccg ggggtcccgg ccgggggggt ataaaagggg gcgggcccct gctcgtcctc 5940
actgtcttcc ggatcgctgt ccaggagcgc cagctgttgg ggtaggtatt ccctctcgaa 6000
ggcgggcatg acctcggcac tcaggttgtc agtttctaga aacgaggagg atttgatatt 6060
gacggtgccg ttggagacgc ctttcatgag cccctcgtcc atttggtcag aaaagacgat 6120
ctttttgttg tcgagcttgg tggcgaagga gccgtagagg gcgttggaga gcagcttggc 6180
gatggagcgc atggtctggt tcttttcctt gtcggcgcgc tccttggcgg cgatgttgag 6240
ctgcacgtac tcgcgcgcca cgcacttcca ttcggggaag acggtggtga gctcgtcggg 6300
cacgattctg acccgccagc cgcggttgtg cagggtgatg aggtccacgc tggtggccac 6360
ctcgccgcgc aggggctcgt tggtccagca gaggcgcccg cccttgcgcg agcagaaggg 6420
gggcagcggg tccagcatga gctcgtcggg ggggtcggcg tccacggtga agatgccggg 6480
caggagctcg gggtcgaagt agctgatgca ggtgcccaga ttgtccagcg ccgcttgcca 6540
gtcgcgcacg gccagcgcgc gctcgtaggg gctgaggggc gtgccccagg gcatggggtg 6600
cgtgagcgcg gaggcgtaca tgccgcagat gtcgtagacg tagaggggct cctcgaggac 6660
gccgatgtag gtggggtagc agcgcccccc gcggatgctg gcgcgcacgt agtcgtacag 6720
ctcgtgcgag ggcgcgagga gccccgtgcc gaggttggag cgttgcggct tttcggcgcg 6780
gtagacgatc tggcggaaga tggcgtggga gttggaggag atggtgggcc tttggaagat 6840
gttgaagtgg gcgtggggca ggccgaccga gtccctgatg aagtgggcgt aggagtcctg 6900
cagcttggcg acgagctcgg cggtgacgag gacgtccagg gcgcagtagt cgagggtctc 6960
ttggatgatg tcatacttga gctggccctt ctgcttccac agctcgcggt tgagaaggaa 7020
ctcttcgcgg tccttccagt actcttcgag ggggaacccg tcctgatcgg cacggtaaga 7080
gcccaccatg tagaactggt tgacggcctt gtaggcgcag cagcccttct ccacggggag 7140
ggcgtaagct tgcgcggcct tgcgcaggga ggtgtgggtg agggcgaagg tgtcgcgcac 7200
catgaccttg aggaactggt gcttgaagtc gaggtcgtcg cagccgccct gctcccagag 7260
ttggaagtcc gtgcgcttct tgtaggcggg gttaggcaaa gcgaaagtaa catcgttgaa 7320
gaggatcttg cccgcgcggg gcatgaagtt gcgagtgatg cggaaaggct ggggcacctc 7380
ggcccggttg ttgatgacct gggcggcgag gacgatctcg tcgaagccgt tgatgttgtg 7440
cccgacgatg tagagttcca cgaatcgcgg gcggcccttg acgtggggca gcttcttgag 7500
ctcgtcgtag gtgagctcgg cggggtcgct gagcccgtgc tgctcgaggg cccagtcggc 7560
gacgtggggg ttggcgctga ggaaggaagt ccagagatcc acggccaggg cggtctgcaa 7620
gcggtcccgg tactgacgga actgttggcc cacggccatt ttttcggggg tgacgcagta 7680
gaaggtgcgg gggtcgccgt gccancggtc ccacttgagc tggagggcga ggtcgtgggc 7740
gagctcgacg agcggcgggt ccccggagag tttcatgacc agcatgaagg ggacgagctg 7800
cttgccgaag gaccccatcc aggtgtaggt ttccacatcg taggtgagga agagcctttc 7860
ggtgcgagga tgcgagccga tggggaagaa ctggatctcc tgccaccagt tggaggaatg 7920
gctgttgatg tgatggaagt agaaatgccg acggcgcgcc gagcactcgt gcttgtgttt 7980
atacaagcgt ccgcagtgct cgcaacgctg cacgggatgc acgtgctgca cgagctgtac 8040
ctgggttcct ttggcgagga atttcagtgg gcagtggagc gctggcggct gcatctcgtg 8100
ctgtactacg tcttggccat cggcgtggcc atcgtctgcc tcgatggtgg tcatgctgac 8160
gagcccgcgc gggaggcagg tccagacctc ggctcggacg ggtcggagag cgaggacgag 8220
ggcgcgcagg ccggagctgt ccagggtcct gagacgctgc ggagtcaggt cagtgggcag 8280
cggcggcgcg cggttgactt gcaggagctt ttccagggcg cgcgggaggt ccagatggta 8340
cttgatctcc acggcgccgt tggtggctac gtccacggct tgcagggtgc cgtgcccctg 8400
gggcgccacc accgtgcccc gtttcttctt gggcgctgct tccatgtcgg tcagaagcgg 8460
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg 8520
ggcacgtcgg cgccgcgcgc gggcaggttc tggtactgcg cccggagaag actggcgtga 8580
gcgacgacgc gacggttgac gtcctggatc tgacgcctct gggtgaaggc cacgggaccc 8640
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc 8700
tgccgcagga tctcttgcac gtcgcccgag ttgtcctggt aggcgatctc ggtcatgaac 8760
tgctcgatct cctcctcctg aaggtctccg cggccggcgc gctcgacggt ggccgcgagg 8820
tcgttggaga tgcggcccat gagctgcgag aaggcgttca tgccggcctc gttccagacg 8880
cggctgtaga ccacggctcc gtcggggtcg cgcgcgcgca tgaccacctg ggcgaggttg 8940
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc 9000
gtggtggcga tgtgctcggt gacgaagaag tacatgatcc agcggcggag cggcatctcg 9060
ctgacgtcgc ccagggcttc caagcgttcc atggcctcgt agaagtccac ggcgaagttg 9120

aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg 9180
gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc catctcctcc 9240
tcttcctcct ccactaacat ctcttctact tcctcctcag gaggcggtgg cgggggaggg 9300
gccctgcgtc gccggcggcg cacgggcaga cggtcgatga agcgctcgat ggtctccccg 9360
cgccggcgac gcatggtctc ggtgacggcg cgcccgtcct cgcggggccg cagcatgaag 9420
acgccgccgc gcatctccag gtggccgccg ggggggtctc cgttgggcag ggagagggcg 9480
ctgacgatgc atcttatcaa ttgacccgta gggactccgc gcaaggacct gagcgtctcg 9540
agatccacgg gatccgaaaa ccgctgaacg aaggcttcga gccagtcgca gtcgcaaggt 9600
aggctgagcc cggtttcttg ttcttcgggt atttggtcgg gaggcggcgg gcgatgctgc 9660
tggtgatgaa gttgaagtag gcggtcctga gacggcggat ggtggcgagg agcaccaggt 9720
ccttgggccc ggcttgctgg atgcgcagac ggtcggccat gccccaggcg tggtcctgac 9780
acctggcgag gtccttgtag tagtcctgca tgagccgctc cacgggcacc tcctcctcgc 9840
ccgcgcggcc gtgcatgcgc gtgagcccga acccgcgctg cggctggacg agcgccaggt 9900
cggcgacgac gcgctcggtg aggatggcct gctggatctg ggtgagggtg gtctggaagt 9960
cgtcgaagtc gacgaagcgg tggtaggctc cggtgttgat ggtgtaggag cagttggcca 10020
tgacggacca gttgacggtc tggtggccgg gtcgcacgag ctcgtggtac ttgaggcgcg 10080
agtaggcgcg cgtgtcgaag atgtagtcgt tgcaggcgcg cacgaggtac tggtatccga 10140
cgaggaagtg cggcggcggc tggcggtaga gcggccatcg ctcggtggcg ggggcgccgg 10200
gcgcgaggtc ctcgagcatg aggcggtggt agccgtagat gtacctggac atccaggtga 10260
tgccggcggc ggtggtggag gcgcgcggga actcgcggac gcggttccag atgttgcgca 10320
gcggcaggaa gtagttcatg gtggccgcgg tctggcccgt gaggcgcgcg cagtcgtgga 10380
tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactccgt ggcctggagg 10440
ctaagcgaac gggttgggct gcgcgtgtac cccggttcga atctcgaatc aggctggagc 10500
cgcagctaac gtggtactgg cactcccgtc tcgacccaag cctgctaacg aaacctccag 10560
gatacggagg cgggtcgttt tttggccttg gtcgctggtc atgaaaaact agtaagcgcg 10620
gaaagcggcc gcccgcgatg gctcgctgcc gtagtctgga gaaagaatcg ccagggttgc 10680
gttgcggtgt gccccggttc gagcctcagc gctcggcgcc ggccggattc cgcggctaac 10740
gtgggcgtgg ctgccccgtc gtttccaaga ccccttagcc agccgacttc tccagttacg 10800
gagcgagccc ctctttttttt ttcttgtgtt tttgccagat gcatcccgta ctgcggcaga 10860
tgcgcccccа ccctccacca caaccgcccc taccgcagca gcagcaacag ccggcgcttc 10920
tgcccccgcc ccagcagcag ccagccacta ccgcggcggc cgccgtgagc ggagccggcg 10980
ttcagtatga cctggccttg gaagagggcg aggggctggc gcggctgggg gcgtcgtcgc 11040
cggagcggca cccgcgcgtg cagatgaaaa gggacgctcg cgaggcctac gtgcccaagc 11100
agaacctgtt cagagacagg agcggcgagg agcccgagga gatgcgcgcc tcccgcttcc 11160
acgcggggcg ggagctgcgg cgcggcctga accgaaagcg ggtgctgagg gacgaggatt 11220
tcgaggcgga cgagctgacg gggatcagcc ccgtgcgcgc gcacgtggtc gnggncaacc 11280
tggtcacggc gtacgagcag accgtgaagg aggagagcaa cttccaaaaa tccttcaaca 11340
accacgtgcg caccttgatc gcgcgcgagg aggtgaccct gggcctgatg cacctgtggg 11400
acctgctgga ggccatcgtg cagaacccca cgagcaagcc gctgacggcg cagctgtttc 11460
tggtggtgca gcacagtcgg gacaacgaga cgttcaggga ggcgctgctg aatatcaccg 11520
agcccgaggg ccgctggctc ctggacctgg tgaacatttt gcagagcatc gtggtgcagg 11580
agcgcgggct gccgctgtcc gagaagctgg cggccatcaa cttctcggtg ctgagtctgg 11640
gcaagtacta cgctaggaag atctacaaga ccccgtacgt gcccatagac aaggaggtga 11700
agatcgacgg gttttacatg cgcatgaccc tgaaagtgct gaccctgagc gacgatctgg 11760
gggtgtaccg caacgacagg atgcaccgcg cggtgagcgc cagccgccgg cgcgagctga 11820
gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgaggggg 11880
agagctactt tgacatgggc gcggacctgc gctggcagcc cagccgccgg gccttggaag 11940
ctgccggcgg ttccccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000
tggaagactg atggcgcgac cgtatttttg ctagatgcag caacagccac cgccgccgcc 12060
tcctgatccc gcgatgcggg cggcgctgca gagccagccg tccggcatta actcctcgga 12120
cgattggacc caggccatgc aacgcatcat ggcgctgacg acccgcaatc ccgaagcctt 12180
tagacagcag cctcaggcca accggctctc ggccatcctg gaggccgtgg tgccctcgcg 12240
ctcgaacccc acgcacgaga aggtgctggc catcgtgaac gcgctggtgg agaacaaggc 12300
catccgcggt gacgaggccg ggctggtgta caacgcgctg ctggagcgcg tggcccgcta 12360
caacagcacc aacgtgcaga cgaacctgga ccgcatggtg accgacgtgc gcgaggcggt 12420

gtcgcagcgc gagcggttcc accgcgagtc gaacctgggc tccatggtgg cgctgaacgc 12480
cttcctgagc acgcagcccg ccaacgtgcc ccggggccag gaggactaca ccaacttcat 12540
cagcgcgctg cggctgatgg tggccgaggt gccccagagc gaggtgtacc agtcggggcc 12600
ggactacttc ttccagacca gtcgccaggg cttgcagacc gtgaacctga gccaggcttt 12660
caagaacttg cagggactgt ggggcgtgca ggccccggtc ggggaccgcg cgacggtgtc 12720
gagcctgctg acgccgaact cgcgcctgct gctgctgctg gtggcgccct tcacggacag 12780
cggcagcgtg agccgcgact cgtacctggg ctacctgctt aacctgtacc gcgaggccat 12840
cggacaggcg cacgtggacg agcagaccta ccaggagatc acccacgtga gccgcgcgct 12900
gggccaggag gacccgggca acctggaggc caccctgaac ttcctgctga ccaaccggtc 12960
gcagaagatc ccgccccagt acgcgctgag caccgaggag gagcgcatcc tgcgctacgt 13020
gcagcagagc gtggggctgt tcctgatgca ggaggggccc acgcccagcg cggcgctcga 13080
catgaccgcg cgcaacatgg agcccagcat gtacgcccgc aaccgcccgt tcatcaataa 13140
gctgatggac tacttgcatc gggcggccgc catgaactcg gactacttta ccaacgccat 13200
cttgaacccg cactggctcc cgccgcccgg gttctacacg ggcgagtacg acatgcccga 13260
ccccaacgac gggttcctgt gggacgacgt ggacagcagc gtgttctcgc cgcgtccagg 13320
aaccaatgcc gtgtggaaga aagagggcgg ggaccggcgg ccgtcctcgg cgctgtccgg 13380
tcgcgcgggt gctgccgcgg cggtgcccga ggccgccagc cccttcccga gcctgccctt 13440
ttcgctgaac agcgtgcgca gcagcgagct gggtcggctg acgcgaccgc gcctgctggg 13500
cgaggaggag tacctgaacg actccttgtt gaggcccgag cgcgagaaga acttccccaa 13560
taacgggata gagagcctgg tggacaagat gagccgctgg aagacgtacg cgcacgagca 13620
cagggacgag ccccgagcta gcagcgcagg cacccgtaga cgccagcggc acgacaggca 13680
gcggggactg gtgtgggacg atgaggattc cgccgacgac agcagcgtgt tggacttggg 13740
tgggagtggt ggtaacccgt tcgctcacct gcgcccccgt atcgggcgcc tgatgtaaga 13800
atctgaaaaa ataaaagacg gtactcacca aggccatggc gaccagcgtg cgttcttctc 13860
tgttgtttgt agtagtatga tgaggcgcgt gtacccggag ggtcctcctc cctcgtacga 13920
gagcgtgatg cagcaggcgg tggcggcggc gatgcagccc ccgctggagg cgccttacgt 13980
gcccccgcgg tacctggcgc ctacggaggg gcggaacagc attcgttact cggagctggc 14040
acccttgtac gataccaccc ggttgtacct ggtggacaac aagtcggcag acatcgcctc 14100
gctgaactac cagaacgacc acagcaactt cctgaccacc gtggtgcaga acaacgattt 14160
caccccccacg gaggccagca cccagaccat caactttgac gagcgctcgc ggtggggcgg 14220
ccagctgaaa accatcatgc acaccaacat gcccaacgtg aacgagttca tgtacagcaa 14280
caagttcaag gcgcgggtga tggtctcgcg caagacccccc aacggggtgg atgatgatta 14340
tgatggtagt caggacgagc tgacctacga gtgggtggag tttgagctgc ccgagggcaa 14400
cttctcggtg accatgacca tcgatctgat gaacaacgcc atcatcgaca actacttggc 14460
ggtggggcgg cagaacgggg tgctggagag cgacatcggc gtgaagttcg acacgcgcaa 14520
cttccggctg ggctgggacc ccgtgaccga gctggtgatg ccgggcgtgt acaccaacga 14580
ggccttccac cccgacatcg tcctgctgcc cggctgcggc gtggacttca ccgagagccg 14640
cctcagcaac ctgctgggca tccgcaagcg gcagcccttc caggagggct tccagatcct 14700
gtacgaggac ctggaggggg gcaacatccc cgcgctcttg gatgtcgaag cctacgagaa 14760
aagcaaggag gatagcaccg ccgcggcgac cgcagccgtg gccaccgcct ctaccgaggt 14820
gcggggcgat aattttgcta gcgctgcggc agcggccgag gcggctgaaa ccgaaagtaa 14880
gatagtcatc cagccggtgg agaaggacag caaggacagg agctacaacg tgctcgcgga 14940
caagaaaaac accgcctacc gcagctggta cctggcctac aactacggcg accccgagaa 15000
gggcgtgcgc tcctggacgc tgctcaccac ctcggacgtc acctgcggcg tggagcaagt 15060
ctactggtcg ctgcccgaca tgatgcaaga cccggtcacc ttccgctcca cgcgtcaagt 15120
tagcaactac ccggtggtgg gcgccgagct cctgcccgtc tactccaaga gcttcttcaa 15180
cgagcaggcc gtctactcgc agnagctgcg cgccttcacc tcgctcacgc acgtcttcaa 15240
ccgcttcccc gagaaccaga tcctcgtccg ccgccgcgcc caccattacc accgtcagtg 15300
aaaacgttcc tgctctcaca gatcacggga ccctgccgct gcgcagcagt atccggggag 15360
tccagcgcgt gaccgtcact gacgccagac gccgcacctg cccctacgtc tacaaggccc 15420
tgggcgtagt cgcgccgcgc gtcctctcga gccgcacctt ctaaaaaatg tccattctca 15480
tctcgcccag taataacacc ggttggggcc tgcgcgcgcc cagcaagatg tacggaggcg 15540
ctcgccaacg ctccacgcaa caccccgtgc gcgtgcgcgg gcacttccgc gctccctggg 15600
gcgccctcaa gggccgcgtg cgctcgcgca ccaccgtcga cgacgtgatc gaccaggtgg 15660
tggccgacgc gcgcaactac acgcccgccg ccgcgcccgt ctccaccgtg gacgccgtca 15720

tcgacagcgt ggtggccgac gcgcgccggt acgcccgcac caagagccgg cggcggcgca 15780
tcgcccggcg gcaccggagc acccccgcca tgcgcgcggc gcgagccttg ctgcgcaggg 15840
ccaggcgcac gggacgcagg gccatgctca gggcggccag acgcgcggcc tccggcagca 15900
gcagcgccgg caggacccgc agacgcgcgg ccacggcggc ggcggcggcc atcgccagca 15960
tgtcccgccc gcggcgcggc aacgtgtact gggtgcgcga cgccgccacc ggtgtgcgcg 16020
tgcccgtgcg cacccgcccc cctcgcactt gaagatgctg acttcgcgat gttgatgtgt 16080
cccagcggcg aggaggatgt ccaagcgcaa atacaaggaa gagatgctcc aggtcatcgc 16140
gcctgagatc tacggccccg cggcggcggt gaaggaggaa agaaagcccc gcaaactgaa 16200
gcgggtcaaa aaggacaaaa aggaggagga agatgacgga ctggtggagt ttgtgcgcga 16260
gttcgccccc cggcggcgcg tgcagtggcg cgggcggaaa gtgaaaccgg tgctgcggcc 16320
cggcaccacg gtggtcttca cgcccggcga gcgttccggc tccgcctcca agcgctccta 16380
cgacgaggtg tacggggacg aggacatcct cgagcaggcg gtcgagcgtc tgggcgagtt 16440
tgcgtacggc aagcgcagcc gccccgcgcc cttgaaagag gaggcggtgt ccatcccgct 16500
ggaccacggc aaccccacgc cgagcctgaa gccggtgacc ctgcagcagg tgctaccgag 16560
cgcggcgccg cgccggggct tcaagcgcga gggcggcgag gatctgtacc cgaccatgca 16620
gctgatggtg cccaagcgcc agaagctgga ggacgtgctg gagcacatga aggtggaccc 16680
cgaggtgcag cccgaggtca aggtgcggcc catcaagcag gtggccccgg gcctgggcgt 16740
gcagaccgtg gacatcaaga tccccacgga gcccatggaa acgcagaccg agcccgtgaa 16800
gcccagcacc agcaccatgg aggtgcagac ggatccctgg atgccagcac cagcttccac 16860
cagcactcgc cgaagacgca agtacggcgc ggccagcctg ctgatgccca actacgcggc 16920
tgcatccttc catcatcccc acgccgggct accgcggcac gcgcttctac cgcggctaca 16980
ccagcagccg ccgccgcaag accaccaccc gccgccgtcg tcgcagccgc cgcagcagca 17040
ccgcgacttc cgccttggtg cggagagtgt atcgcagcgg gcgcgagcct ctgaccctgc 17100
cgcgcgcgcg ctaccacccg agcatcgcca tttaactacc gcctcctact tgcagatatg 17160
gccctcacat gccgcctccg cgtccccatt acgggctacc gaggaagaaa gccgcgccgt 17220
agaaggctga cggggaacgg gctgcgtcgc catcaccacc ggcggcggcg cgccatcagc 17280
aagcggttgg ggggaggctt cctgcccgcg ctgatcccca tcatcgccgc ggcgatcggg 17340
gcgatccccg gcatagcttc cgtggcggtg caggcctctc agcgccactg agacacaaaa 17400
aagcatggat ttgtaataaa aaaaaaaatg gactgacgct cctggtcctg tgatgtgtgt 17460
ttttagatgg aagacatcaa tttttcgtcc ctggcaccgc gacacggcac gcggccgttt 17520
atgggcacct ggagcgacat cggcaacagc caactgaacg ggggcgcctt caattggagc 17580
agtctctgga gcgggcttaa gaatttcggg tccacgctca aaacctatgg caacaaggcg 17640
tggaacagca gcacagggca ggcgctgagg gaaaagctga aagaacagaa cttccagcag 17700
aaggtggttg atggcctggc ctcaggcatc aacggggtgg ttgacctggc caaccaggcc 17760
gtgcagaaac agatcaacag ccgcctggac gcggtcccgc ccgcggggtc cgtggagatg 17820
ccccaggtgg aggaggagct gcctcccctg gacaagcgcg gcgacaagcg accgcgtccc 17880
gacgcggagg agacgctgct gacgcacacg gacgagccgc ccccgtacga ggaggcggtg 17940
aaactgggcc tgcccaccac gcggcccgtg gcgcctctgg ccaccggagt gctgaaaccc 18000
agcagcagcc agcccgcgac cctggacttg cctccgcctc gcccctccac agtggctaag 18060
cccctgccgc cggtggccgt cgcgtcgcgc gccccccgag gccgccccca ggcgaactgg 18120
cagagcactc tgaacagcat cgtgggtctg ggagtgcaga gtgtgaagcg ccgccgctgc 18180
tattaaaaga cactgtagcg cttaacttgc ttgtctgtgt gtatatgtat gtccgccgac 18240
cagaaggagg agtgtgaaga ggcgcgtcgc cgagttgcaa gatggccacc ccatcgatgc 18300
tgccccagtg ggcgtacatg cacatcgccg gacaggacgc ttcggagtac ctgagtccgg 18360
gtctggtgca gttcgcccgc gccacagaca cctacttcag tctggggaac aagtttagga 18420
accccacggt ggcgcccacg cacaatgtga ccaccgaccg cagccagcgg ctgacggtgc 18480
gcttcgtgcc cgtggaccgc gaggacaaca cctactcgta caaagtgcgc tacacgctgg 18540
ccgtgggcga caaccgcgtg ctggacatgg ccagcaccta ctttgacatc cgcggcgtgc 18600
tggaccgggg ccctagcttc aaaccctact ctggcaccgc ctacaacagc ctagctccca 18660
agggagctcc caattccagc cagtgggagc aagcaaaaac aggcaatggg ggaactatgg 18720
aaacacacac atatggtgtg gccccaatgg gcggagagaa tattacaaaa gatggtcttc 18780
aaattggaac tgacgttaca gcgaatcaga ataaaccaat ttatgccgac aaaacatttc 18840
aaccagaacc gcaagtagga gaagaaaatt ggcaagaaac tgaaaacttt tatggcggta 18900
gagctcttaa aaaagacaca aacatgaaac cttgctatgg ctcctatgct agacccacca 18960
atgaaaaagg aggtcaagct aaacttaaag ttggagatga tggagttcca accaaagaat 19020

tcgacataga cctggctttc tttgatactc ccggtggcac cgtgaacggt caagacgagt 19080
ataaagcaga cattgtcatg tataccgaaa acacgtattt ggaaactcca gacacgcatg 19140
tggtatacaa accaggcaag gatgatgcaa gttctgaaat taacctggtt cagcagtcta 19200
tgcccaacag acccaactac attgggttca gggacaactt tatcggtctt atgtactaca 19260
acagcactgg caatatgggt gtgcttgctg gtcaggcctc ccagctgaat gctgtggttg 19320
atttgcaaga cagaaacacc gagctgtcct accagctctt gcttgactct ttgggtgaca 19380
gaacccggta tttcagtatg tggaaccagg cggtggacag ttatgacccc gatgtgcgca 19440
tcatcgaaaa ccatggtgtg gaggatgaat tgccaaacta ttgcttcccc ttggacggct 19500
ctggcactaa cgccgcatac caaggtgtga aagtaaaaga tggtcaagat ggtgatgttg 19560
agagtgaatg ggaaaatgac gatactgttg cagctcgaaa tcaattatgt aaaggtaaca 19620
ttttcgccat ggagattaat ctccaggcta acctgtggag aagtttcctc tactcgaacg 19680
tggccctgta cctgcccgac tcctacaagt acacgccgac caacgtcacg ctgccgacca 19740
acaccaacac ctacgattac atgaatggca gagtgacacc tccctcgctg gtagacgcct 19800
acctcaacat cggggcgcgc tggtcgctgg accccatgga caacgtcaac cccttcaacc 19860
accaccgcaa cgcgggcctg cgctaccgct ccatgctcct gggcaacggg cgctacgtgc 19920
ccttccacat ccaggtgccc caaaagtttt tcgccatcaa gagcctcctg ctcctgcccg 19980
ggtcctacac ctacgagtgg aacttccgca aggacgtcaa catgatcctg cagagctccc 20040
taggcaacga cctgcgcacg gacggggcct ccatcgcctt caccagcatc aacctctacg 20100
ccaccttctt ccccatggcg cacaacaccg cctccacgct cgaggccatg ctgcgcaacg 20160
acaccaacga ccagtccttc aacgactacc tctcggcggc caacatgctc taccccatcc 20220
cggccaacgc caccaacgtg cccatctcca tcccctcgcg caactgggcc gccttccgcg 20280
gatggtcctt cacgcgcctg aagacccgcg agacgccctc gctcggctcc gggttcgacc 20340
cctacttcgt ctactcgggc tccatccccc acctagacgg caccttctac ctcaaccaca 20400
ccttcaagaa ggtctccatc accttcgact cctccgtcag ctggcccggc aacgaccgcc 20460
tcctgacgcc caacgagttc gaaatcaagc gcaccgtcga cggagaggga tacaacgtgg 20520
cccagtgcaa catgaccaag gactggttcc tggtccagat gctggcccac tacaacatcg 20580
gctaccaggg cttctacgtg cccgagggct acaaggaccg catgtactcc ttcttccgca 20640
acttccagcc catgagccgc caggtcgtgg acgaggtcaa ctacaaggac taccaggccg 20700
tcaccctggc ctaccagcac aacaactcgg gcttcgtcgg ctacctcgcg cccaccatgc 20760
gccagggcca gccctacccc gccaactacc cctacccgct catcggcaag agcgccgtcg 20820
ccagcgtcac ccagaaaaag ttcctctgcg accgggtcat gtggcgcatc cccttctcca 20880
gcttcctcaa tgcccactcc gcctactttc gctcccaccg cgcgcgcatc gagaaggcca 20940
gcaacttcat gtccatgggc gcgctcaccg acctcggcca gaacatgctc tacgccaact 21000
ccgcccacgc gctagacatg aatttcgaag tcgaccccat ggatgagtcc acccttctct 21060
atgttgtctt cgaagtcttc gacgtcgtcc gagtgcacca gccccaccgc ggcgtcatcg 21120
aagccgtcta cctgcgcacg cccttctcgg ccggcaacgc caccacctaa gccgctcttg 21180
cttcttgcaa gatgacggcg ggctccggcg agcaggagct cagggccatc ctccgcgacc 21240
tgggctgcgg gccctgcttc ctgggcacct tcgacaagcg cttccctgga ttcatggccc 21300
cgcacaagct ggcctgcgcc atcgtgaaca cggccggccg cgagaccggg ggcgagcact 21360
ggctggcctt cgcctggaac ccgcgctccc acacatgcta cctcttcgac cccttcgggt 21420
tctcggacga gcgcctcaag cagatctacc agttcgagta cgagggcctg ctgcgtcgca 21480
gcgccctggc caccgaggac cgctgcgtca ccctggaaaa gtccacccag accgtgcagg 21540
gtccgcgctc ggccgcctgc gggctcttct gctgcatgtt cctgcacgcc ttcgtgcact 21600
ggcccgaccg ccccatggac aagaacccca ccatgaactt actgacgggg gtgcccaacg 21660
gcatgctcca gtcgccccag gtggaaccca ccctgcgccg caaccaggaa gcgctctacc 21720
ccgccttcga ccgcatgaat caagacatgt aaaaaaccgg tgtgtgtatg tgaatgcttt 21780
attcataata aacagcacat gtttatgcca ccttctctga ggctctgact ttatttagaa 21840
atcgaagggg ttctgccggc tctcggcatg gcccgcgggc agggatacgt tgcggaactg 21900
gtacttgggc agccacttga actcggggat cagcagcttg ggcacgggga ggtcggggaa 21960
cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg gcgcggagat 22020
cttgaaatcg cagttgggac ccgcgttctg cgcgcgagag ttgcggtaca cggggttgca 22080
gcactggaac accatcaggg ccgggtgctt cacgcttgcc agcaccgtcg cgtcggtgat 22140
gccctccacg tccagatcct cggcgttggc catcccgaag ggggtcatct tgcaggtctg 22200
ccgccccatg ctgggcacgc agccgggctt gtggttgcaa tcgcagtgca gggggatcag 22260
catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga aagccttcag 22320

ctggcggaag gcctgctgcg ccttgccgcc ctcggtgaag aagaccccgc aggacttgct 22380
agagaactgg ttggtggcgc agccggcgtc gtgcacgcag cagcgcgcgt cgttgttggc 22440
cagctgcacc acgctgcgcc cccagcggtt ctgggtgatc ttggcccggt tggggttctc 22500
cttcagcgcg cgctgcccgt tctcgctcgc cacatccatc tcgatagtgt gctccttctg 22560
gatcatcacg gtcccgtgca ggcaccgcag cttgccctcg gcttcggtgc agccgtgcag 22620
ccacagcgcg cagccggtgc actcccagtt cttgtgggcg atctgggagt gcgagtgcac 22680
gaagccctgc aggaagcggc ccatcatcgc ggtcagggtc ttgttgctgg tgaaggtcag 22740
cgggatgccg cggtgctcct cgttcacata caggtggcag atgcggcggt acacctcgcc 22800
ctgctcgggc atcagctgga aggcggactt caggtcgctc tccacgcggt accggtccat 22860
cagcagcgtc atcacttcca tgcccttctc ccaggccgaa acgatcggca ggctcagggg 22920
gttcttcacc gccattgtca tcttagtcgc cgccgccgag gtcaggggggt cgttctcgtc 22980
cagggtctca aacactcgct tgccgtcctt ctcgatgatg cgcacggggg gaaagctgaa 23040
gcccacggcc gccagctcct cctcggcctg cctttcgtcc tcgctgtcct ggctgatgtc 23100
ttgcaaaggc acatgcttgg tcttgcgggg tttcttttttg ggcggcagag gcggcggcga 23160
tgtgctggga gagcgcgagt tctcgttcac cacgactatt tcttcttctt ggccgtcgtc 23220
cgagaccacg cggcggtagg catgcctctt ctggggcaga ggcggaggcg acgggctctc 23280
gcggttcggc gggcggctgg cagagcccct tccgcgttcg ggggtgcgct cctggcggcg 23340
ctgctctgac tgacttcctc cgcggccggc cattgtgttc tcctagggag caacaacaag 23400
catggagact cagccatcgt cgccaacatc gccatctgcc cccgccgcca ccgccgacga 23460
gaaccagcag cagaatgaaa gcttaaccgc cccgccgccc agccccacct ccgacgccgc 23520
ggccccagac atgcaagaga tggaggaatc catcgagatt gacctgggct acgtgacgcc 23580
cgcggagcac gaggaggagc tggcagcgcg cttttcagcc ccggaagaga accaccaaga 23640
gcagccagag caggaagcag agaacgagca gaaccaggct gggcacgagc atggcgacta 23700
cctgagcggg gcagaggacg tgctcatcaa gcatctggcc cgccaatgca tcatcgtcaa 23760
ggacgcgctg ctcgaccgcg ccgaggtgcc cctcagcgtg gcggagctca gccgcgccta 23820
cgagcgcaac ctcttctcgc cgcgcgtgcc ccccaagcgc cagcccaacg gcacctgtga 23880
gcccaacccg cgcctcaact tctacccggt cttcgcggtg cccgaggccc tggccaccta 23940
ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc gcacccgcgc 24000
cgacgccctg ctcaacctgg gccccggcgc ccgcctacct gatatcacct ccttggaaga 24060
ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggccgcga acgctctgca 24120
aggaagcgga gaggagcatg agcaccacag cgccctggtg gagttggaag gcgacaacgc 24180
gcgcctggcg gtcctcaagc gcacggtcga gctgacccac ttcgcctacc cggcgctcaa 24240
cctgcccccc aaggtcatga gcgccgtcat ggaccaggtg ctcatcaagc gcgcctcgcc 24300
cctctcggag gaggagatgc aggaccccga gagttcggac gagggcaagc ccgtggtcag 24360
cgacgagcag ctggcgcgct ggctgggagc gagtancacc ccccagagcc tggaagagcg 24420
gcgcaagctc atgatggccg tggtcctggt gaccgtggag ctggagtgtc tgcgccgctt 24480
ctttgccgac gcggagaccc tgcgcaaggt cgaggagaac ctgcactacc tcttcaggca 24540
cgggttcgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc tggtctccta 24600
catgggcatc ctgcacgaga accgcctggg gcaaaacgtg ctgcacacca ccctgcgcgg 24660
ggaggcccgc cgcgactaca tccgcgactg cgtctacctg tacctctgcc acacctggca 24720
gacgggcatg ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag agctctgcaa 24780
gctcctgcag aagaacctca aggccctgtg gaccgggttc gacgagcgta ccaccgcctc 24840
ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca acgggctgcc 24900
cgactttatg agccaaagca tgttgcaaaa ctttcgctct ttcatcctcg aacgctccgg 24960
gatcctgccc gccacctgct ccgcgctgcc ctcggacttc gtgccgctga ccttccgcga 25020
gtgccccccg ccgctctgga gccactgcta cttgctgcgc ctggccaact acctggccta 25080
ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctggagtgcc actgccgctg 25140
caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga gcgagaccca 25200
gatcatcggc accttcgagt tgcaaggccc cggcgacggc gagggcaagg ggggtctgaa 25260
actcaccccg gggctgtgga cctcggccta cttgcgcaag ttcgtgcccg aggactacca 25320
tcccttcgag atcaggttct acgaggacca atcccagccg cccaaggccg agctgtcggc 25380
ctgcgtcatc acccaggggg ccatcctggc ccaattgcaa gccatccaga aatcccgcca 25440
agaatttctg ctgaaaaagg gccacggggt ctacttggac ccccagaccg gagaggagct 25500
caaccccagc ttcccccagg atgccccgag gaagcagcaa gaagctgaaa gtggagctgc 25560
cgccgccgga ggatttggag gaagactggg agagcagtca ggcagaggag gaggagatgg 25620

aagactggga cagcactcag gcagaggagg acagcctgca agacagtctg gaggaggaag 25680
acgaggtgga ggaggcagag gaagaagcag ccgccgccag accgtcgtcc tcggcggaga 25740
aagcaagcag cacggatacc atctccgctc cgggtcgggg tcgcggcggc cgggcccaca 25800
gtaggtggga cgagaccggg cgcttccgaa ccccaccacc cagaccggta agaaggagcg 25860
gcagggatac aagtcctggc gggggcacaa aaacgccatc gtctcctgct tgcaagcctg 25920
cgggggcaac atctccttca cccggcgcta cctgctcttt caccgcgggg tgaacttccc 25980
ccgcaacatc ttgcattact accgtcacct ccacagcccc tactactgtt tccaagaaga 26040
ggcagaaacc cagcagcagc agaaaaccag cggcagcagc agctagaaaa tccacagcgg 26100
cggcaggtgg actgaggatc gcggcgaacg agccggcgca gacccgggag ctgaggaacc 26160
ggatctttcc caccctctat gccatcttcc agcagagtcg ggggcaggag caggaactga 26220
aagtcaagaa ccgttctctg cgctcgctca cccgcagttg tctgtatcac aagagcgaag 26280
accaacttca gcgcactctc gaggacgccg aggctctctt caacaagtac tgcgcgctca 26340
ctcttaaaga gtagcccgcg cccgcccaca cacggaaaaa ggcgggaatt acgtcaccac 26400
ctgcgccctt cgcccgacca tcatgagcaa agagattccc acgccttaca tgtggagcta 26460
ccagccccag atgggcctgg ccgccggcgc cgcccaggac tactccaccc gcatgaactg 26520
gctcagtgcc gggcccgcga tgatctcacg ggtgaatgac atccgcgccc accgaaacca 26580
gatactccta gaacagtcag cgatcaccgc cacgccccgc catcacctta atccgcgtaa 26640
ttggcccgcc gccctggtgt accaggaaat tccccagccc acgaccgtac tacttccgcg 26700
agacgcccag gccgaagtcc agctgactaa ctcaggtgtc cagctggccg gcggcgccgc 26760
cctgtgtcgt caccgccccg ctcagggtat aaagcggctg gtgatccgag gcagaggcac 26820
acagctcaac gacgaggtgg tgagctcttc gctgggtctg cgacctgacg gagtcttcca 26880
actcgccgga tcggggagat cttccttcac gcctcgtcag gccgtcctga ctttggagag 26940
ttcgtcctcg cagccccgct cgggcggcat cggcactctc cagttcgtgg aggagttcac 27000
tccctcggtn tacttcaacc ccttctccgg ctcccccggc cactacccgg acgagttcat 27060
cccgaacttc gacgccatca gcgagtcggt ggacggctac gattgaatgt cccatggtgg 27120
cgcagctgac ctagctcggc ttcgacacct ggaccactgc cgccgcttcc gctgcttcgc 27180
tcgggatctc gccgagtttg cctactttga gctgcccgag gagcaccctc agggcccagc 27240
ccacggagtg cggatcatcg tcgaaggggg cctcgactcc cacctgcttc ggatcttcag 27300
ccagcgaccg atcctggtcg agcgcgaaca aggacagacc cttcttactt tgtactgcat 27360
ctgcaaccac cccggcctgc atgaaagtct ttgttgtctg ctgtgtactg agtataataa 27420
aagctgagat cagcgactac tccggactcg attgtggtgt tcctgctatc aaccggtccc 27480
tgttcttcac cgggaacgag accgagctcc agctccagtg taagccccac aagaagtacc 27540
tcacctggct gttccagggc tccccgatcg ccgttgtcaa ccactgcgac aacgacggag 27600
tcctgctgag cggccctgcc aaccttactt tttccacccg cagaagcaag ctccagctct 27660
tccaaccctt cctccccggg acctatcagt gcgtctcagg accctgccat cacaccttcc 27720
acctgatccc gaataccaca gcgccgctcc ccgctactaa caaccaaact acccaccaac 27780
gccaccgtcg cgacctttcc tctgaatcta ataccactac cggaggtgag ctccgaggtc 27840
gaccaacctc tgggatttac tacggccccc gggaggtggt ggggttaata gcgctaggcc 27900
tagttgcggg tgggcttttg gttctctgct acctatacct cccttgctgt tcgtacttag 27960
tggtgctgtg ttgctggttt aagaaatggg gaagatcacc ctagtgagct gcggtgcgct 28020
ggtggcggtg ttgctttcga ttgtgggact gggcggcgcg gctgtagtga aggagaaggc 28080
cgatccctgc ttgcatttca atcccaacaa atgccagctg agttttcagc ccgatggcaa 28140
tcggtgcgcg gtactgatca agtgcggatg ggaatgcgag aacgtgagaa tcgagtacaa 28200
taacaagact cggaacaata ctctcgcgtc cgtgtggcag cccggggacc ccgagtggta 28260
caccgtctct gtccccggtg ctgacggctc cccgcgcacc gtgaataata ctttcatttt 28320
tgcgcacatg tgcaacacgg tcatgtggat gagcaagcag tacgatatgt ggcccccccac 28380
gaaggagaac atcgtggtct tctccatcgc ttacagcctg tgcacggcgc taatcaccgc 28440
tatcgtgtgc ctgagcattc acatgctcat cgctattcgc cccagaaata atgccgagaa 28500
agagaaacag ccataacacg ttttttcaca caccttgttt ttacagacaa tgcgtctgtt 28560
aaatttttta aacattgtgc tcagtattgc ttatgcctct ggttatgcaa acatacagaa 28620
aaccctttat gtaggatctg atggtacact agagngtacc caatcacaag ccaaggttgc 28680
atggtatttt tatagaacca acactgatcc agttaaactt tgtaagggtg aattgccgcg 28740
tacacataaa actccactta catttagttg cagcaataat aatcttacac ttttttcaat 28800
tacaaaacaa tatactggta cttattacag tacaaacttt catacaggac aagataaata 28860
ttatactgtt aaggtagaaa atcctaccac tcctagaact accaccacca ccactactgc 28920

aaagcccact gtgaaaacta caactaggac caccacaact acagaaacca ccaccagcac 28980
aacacttgct gcaactacac acacacacac taagctaacc ttacagacca ctaatgattt 29040
gatcgccctg ctgcaaaagg gggataacag caccacttcc aatgaggaga tacccaaatc 29100
catgattggc attattgttg ctgtagtggt gtgcatgttg atcatcgcct tgtgcatggt 29160
gtactatgcc ttctgctaca gaaagcacag actgaacgac aagctggaac acttactaag 29220
tgttgaattt taatttttta gaaccatgaa gatcctaggc ctttttagtt tttctatcat 29280
tacctctgct ctttgtgaat cagtggatag agatgttact attaccactg gttctaatta 29340
tacactgaaa gggccaccct caggtatgct ttcgtggtat tgctattttg gaactgacac 29400
tgatcaaact gaattatgca attttcaaaa aggcaaaacc tcaaactcta aaatctctaa 29460
ttatcaatgc aatggcactg atctgatact actcaatgtc acgaaagcat atggtggcag 29520
ttattattgc cctggacaaa acactgaaga aatgattttt tacaaagtgg aagtggttga 29580
tcccactaca ccacccacca ccacaactat tcataccaca cacacagaac aaacaccaga 29640
ggcaacagaa gcagagttgg ccttccaggt tcacggagat tcctttgctg tcaatacccc 29700
tacacccgat cagcggtgtc cggggccgct agtcagcggc attgtcggtg tgctttcggg 29760
attagcagtc ataatcatct gcatgttcat ttttgcttgc tgctatagaa ggctttaccg 29820
acaaaaatca gacccactgc tgaacctcta tgtttaattt tttccagagc catgaaggca 29880
gttagcgctc tagtttttg ttctttgatt ggcattgttt ttaatagtaa aattaccaga 29940
gttagcttta ttaaacatgt taatgtaact gaaggagata acatcacact agcaggtgta 30000
gaaggtgctc aaaacaccac ctggacaaaa taccatctag gatggagaga tatttgcacc 30060
tggaatgtaa cttattattg cataggagtt aatcttacca ttgttaacgc taaccaatct 30120
cagaatgggt taattaaagg acagagtgtt agtgtgacca gtgatgggta ctatacccag 30180
catagtttta actacaacat tactgtcata ccactgccta cgcctagccc acctagcact 30240
accacacaga caaccacata cagtacatca aatcagccta ccaccactac agcagcagag 30300
gttgccagct cgtctggggt ccgagtggca tttttgatgt tggccccatc tagcagtccc 30360
actgctagta ccaatgagca gactactgaa tttttgtcca ctgtcgagag ccacaccaca 30420
gctacctcca gtgccttctc tagcaccgcc aatctctcct cgctttcctc tacaccaatc 30480
agccccgcta ctactcctag ccccgctcct cttcccactc ccctgaagca aacagacggc 30540
ggcatgcaat ggcagatcac cctgctcatt gtgatcgggt tggtcatcct ggccgtgttg 30600
ctctactaca tcttctgccg ccgcattccc aacgcgcacc gcaagccggc ctacaagccc 30660
atcgttatcg ggcagccgga gccgcttcag gtggaagggg gtctaaggaa tcttctcttc 30720
tcttttacag tatggtgatt gaantatgat tcctagacaa ttcttgatca ctattcttat 30780
ctgcctcctc caagtctgtg ccaccctcgc tctggtggcc aacgccagtc cagactgtat 30840
tgggcccttc gcctcctacg tgctctttgc cttcgtcacc tgcatctgct gctgtagcat 30900
agtctgcctg cttatcacct tcttccagtt cattgactgg atctttgtgc gcatcgccta 30960
cctgcgccac caccccagt accgcgacca gcgagtggcg cagctgctca ggctcctctg 31020
ataagcatgc gggctctgct acttntcgcg cttctgctgt tagtgctccc ccgtcccgtc 31080
gacccccggt cccccactca gtcccccgag gaggttcgca aatgcaaatt ccaagaaccc 31140
tggaaattcc tcaaatgcta ccgccaaaaa tcagacatgc atcccagctg gatcatgatc 31200
attgggatcg tgaacattct ggcctgcacc ctcatctcct ttgtgattta cccctgcttt 31260
gactttggtt ggaactcgcc agaggcgctc tatctcccgc ctgaacctga cacaccacca 31320
cagcatcaac ctcaggcaca cgcactacca ccaccacagc ctaggccaca atacatgccc 31380
atattagact atgaggccga gccacagcga cccatgctcc ccgctattag ttacttcaat 31440
ctaaccggcg gagatgactg acccactggc caataacaac gtcaacgacc ttctcctgga 31500
catggacggc cgcgcctcgg agcagcgact cgcccaactt cgcattcgtc agcagcagga 31560
gagagccgtc aaggagctgc aggacggcat agccatccac cagtgcaaga gaggcatctt 31620
ctgcctggtg aaacaggcca agatctccta cgaggtcacc cagaccgacc atcgcctctc 31680
ctacgagctc ctgcagcagc gccagaagtt cacctgcctg gtcggagtca accccatcgt 31740
catcacccag ccagcagtcg ggcgatacca aggggtgcat ccactgctcc tgcgactccc 31800
ccgactgcgt ccacactctg atcaagaccc tctgcggcct ccgcgacctc ctccccatga 31860
actaatcacc cccttatcca gtgaaataaa gatcatattg atgatgattt aaataaaaaa 31920
aataatcatt tgatttgaaa taaagataca atcatattga tgatttgagt ttaacaaaaa 31980
taaagaatca cttacttgaa atctgatacc aggtctctgt ccatgttttc tgccaacacc 32040
acctcactcc cctcttccca gctctggtac tgcaggcccc ggcgggctgc aaacttcctc 32100
cacacgctga aggggatgtc aaattcctcc tgtccctcaa tcttcatttt atcttctatc 32160
agatgtccaa aaagcgcgtc cgggtggatg atgacttcga ccccgtctac ccctacgatg 32220

cagacaacgc accgaccgtg cccttcatca acccccccct cgtctcttca gatggattcc 32280
aagagaagcc cctgggggtg ttgtccctgc gactggctga ccccgtcacc accaagaacg 32340
gggaaatcac cctcaagctg ggagaggggg tggacctcga ctcgtcggga aaactcatct 32400
ccaacacggc caccaaggcc gccgcccctc tcagtatttc aaacaacacc atttccctta 32460
aaactgctgc ccctttctac aacaacaatg gaactttaag cctcaatgtc tccacaccat 32520
tagcagtatt tcccacattt aacactttag gcataagtct tggaaacggt cttcagactt 32580
caaataagtt gttgactgta caactaactc atcctcttac attcagctca aatagcatca 32640
cagtaaaaac agacaaaggg ctatatatta actccagtgg aaacagagga cttgaggcta 32700
atataagcct aaaaagagga ctagtttttg acggtaatgc tattgcaaca tatattggaa 32760
atggcttaga ctatggatct tatgatagtg atggaaaaac aagacccgta attaccaaaa 32820
ttggagcagg attaaatttt gatgctaaca aagcaatagc tgtcaaacta ggcacaggtt 32880
taagttttga ctccgctggt gccttgacag ctggaaacaa acaggatgac aagctaacac 32940
tttggactac ccctgaccca agccctaatt gtcaattact ttcagacaga gatgccaaat 33000
ttactctctg tcttacaaaa tgcggtagtc aaatactagg cactgtggca gtggcggctg 33060
ttactgtagg atcagcacta aatccaatta atgacacagt caaaagcgcc atagttttcc 33120
ttagatttga ttccgatggt gtactcatgt caaactcatc aatggtaggt gattactgga 33180
actttaggga gggacagacc actcaaagtg tagcctatac aaatgctgtg ggattcatgc 33240
caaatatagg tgcatatcca aaaacccaaa gtaaaacacc taaaaatagc atagtcagtc 33300
aggtatattt aactggagaa actactatgc caatgacact aaccataact ttcaatggca 33360
ctgatgaaaa agacacaacc ccagttagca cctactctat gacttttaca tggcagtgga 33420
ctggagacta taaggacaaa aatattacct ttgctaccaa ctcattctct ttttcctaca 33480
tcgcccagga ataatcccac ccagcaagcc aacccctttt cccaccacct ttgtctatat 33540
ggaaactctg aaacagaaaa ataaagttca agtgttttat tgaatcaaca gttttacagg 33600
actcgagcag ttatttttcc tccaccctcc caggacatgg aatacaccac cctctccccc 33660
cgcacagcct tgaacatctg aatgccattg gtgatggaca tgcttttggt ctccacgttc 33720
cacacagttt cagagcgagc cagtctcgga tcggtcaggg agatgaaacc ctccgggcac 33780
tcccgcatct gcacctcaca gctcaacagc tgaggattgt cctcggtggt cgggatcacg 33840
gttatctgga agaagcagaa gagcggcggt gggaatcata gtccgcgaac gggatcggcc 33900
ggtggtgtcg catcaggccc cgcagcagtc gctgccgccg ccgctccgtc aagctgctgc 33960
tcaggggttt cgggtccagg gactccctca gcatgatgcc cacggccctc agcatcagtc 34020
gtctggtgcg gcgggcgcag cagcgcatgc gaatctcgct caggtcactg cagtacgtgc 34080
aacacaggac caccaggttg ttcaacagtc catagttcaa cacgctccag ccgaaactca 34140
tcgcgggaag gatgctaccc acgtggccgt cgtaccagat cctcaggtaa atcaagtggc 34200
gctccctcca gaagacgctg cccatgtaca tgatctcctt gggcatgtgg cggttcacca 34260
cctcccggta ccacatcacc ctctggttga acatgcagcc ccggatgatc ctgcggaacc 34320
acagggccag caccgccccg cccgccatgc agcgaagaga ccccggatcc cggcaatgac 34380
aatggaggac ccaccgctcg tacccgtgga tcatctggga gctgaacaag tctatgttgg 34440
cacagcacag gcatatgctc atgcatctct tcagcactct cagctcctcg ggggtcaaaa 34500
ccatatccca gggcacgggg aactcttgca ggacagcgaa ccccgcagaa cagggcaatc 34560
ctcgcacata acttacattg tgcatggaca gggtatcgca atcaggcagc accgggtgat 34620
cctccaccag agaagcgcgg gtctcggtct cctcacagcg tggtaagggg gccggccgat 34680
acgggtgatg gcgggacgcg gctgatcgtg ttctcgaccg tgtcatgatg cagttgcttt 34740
cggacatttt cgtacttgct gtagcagaac ctggtccggg cgctgcacac cgatcgccgg 34800
cggcggtctc ggcgcttgga acgctcggtg ttaaagttgt aaaacagcca ctctctcaga 34860
ccgtgcagca gatctagggc ctcaggagtg atgaagatcc catcatgcct gatagctctg 34920
atcacatcga ccaccgtgga atgggccagg cccagccaga tgatgcaatt ttgttgggtt 34980
tcggtgacgg cgggggaggg aagaacagga agaaccatga ttaactttta atccaaacgg 35040
tctcggagca cttcaaaatg aaggtcacgg agatggcacc tctcgccccc gctgtgttgg 35100
tggaaaataa cagccaggtc aaaggtgata cggttctcga gatgttccac ggtggcttcc 35160
agcaaagcct ccacgcgcac atcagaaaca agacaatagc gaaagcggga gggttctcta 35220
attcctcaac catcatgtta cactcctgca ccatccccag ataattttca tttttccagc 35280
cttgaatgat tcgaactagt tcctgaggta aatccaagcc agccatgata aaaagctcgc 35340
gcagagcacc ctccaccggc attcttaagc acaccctcat aattccaaga tattctgctc 35400
ctggttcacc tgcagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
ctcctccctc agcaataact gtaagtactc tttcatatcg tctccgaaat ttttagccat 35520

```
aggaccccca ggaataagag aagggcaagc cacattacag ataaaccgaa gtccccccca 35580
gtgagcattg ccaaatgtaa gattgaaata agcatgctgg ctagacccgg tgatatcttc 35640
cagataactg gacagaaaat cgggtaagca atttttaaga aaatcaacaa aagaaaaatc 35700
ttccaggtgc acgtttaggg cctcgggaac aacgatggag taagtgcaag gggtgcgttc 35760
cagcatggtt agttagctga tctgtaaaaa aacaaaaaat aaaacattaa accatgctag 35820
cctggcgaac aggtgggtaa atcgttctct ccagcaccag gcaggccacg gggtctccgg 35880
cgcgaccctc gtaaaaattg tcgctatgat tgaaaaccat cacagagaga cgttcccggt 35940
ggccggcgtg aatgattcga gaagaagcat acacccccgg gaacattgga gtccgtgagt 36000
gaaaaaaagc ggccgaggaa gcaatgaggc actacaacgc tcactctcaa gtccagcaaa 36060
gcgatgccat gcggatgaag cacaaaattt tcaggtgcgt aaaaaatgta attactcccc 36120
tcctgcacag gcagcgaagc tcccgatccc tccagataca catacaaagc ctcagcgtcc 36180
atagcttacc gagcggcagc agcagcggca cacaacaggc gcaagagtca gagaaaagac 36240
tgagctctaa cctgtccgcc cgctctctgc tcaatatata gccccagatc tacactgacg 36300
taaaggccaa agtctaaaaa tacccgccaa ataatcacac acgcccagca cacgcccaga 36360
aaccggtgac acactcagaa aaatacgcgc acttcctcaa acggccaaac tgccgtcatt 36420
tccgggttcc cacgctacgt catcaaaaca cgactttcaa attccgtcga ccgttaaaaa 36480
catcacccgc cccgccccta acggtcgccg ctcccgcagc caatcacctt cctccctccc 36540
caaattcaaa cagctcattt gcatattaac gcgcaccaaa agtttgaggt atattattga 36600
tgatgg                                                           36606
```

<210> 4
<211> 36535
<212> DNA
<213> Chimpanzee Pan 7 (CV33) Genomic

<400> 4

```
catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg agctgtttga 60
atttggggag ggaggaaggt gattggccga gagacgggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
gtggcgaccc tcctgagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgtatg 720
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgca tcgaaccagg gagtgaaagc tgcgggcgaa agctttagcc 1020
tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
acagtaagtg tgattaactt tagttgggaa ggcagagggt gactgggtgc tgactggttt 1200
atttatgtat atgtttttt atgtgtaggt cccgtctctg acgtagatga gacccccact 1260
tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
agaccagttg cagtgagagt caccgggcgg agagcagctg tggagagttt ggatgacttg 1380
ctacagggtg gggatgaacc tttggacttg tgtacccgga aacgccccag gcactaagtg 1440
ccacacatgt gtgtttactt aaggtgatgt cagtatttat agggtgtgga gtgcaataaa 1500
atccgtgttg actttaagtg cgtggtttat gactcagggg tggggactgt gggtatataa 1560
gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gacggtcttg 1620
gaagactttc accagactag acagctgcta gagaactcat cggaggggtc ctcttacctg 1680
tggagattct gcttcggtgg gcctctagct aagctagtct atagggccaa acaggattat 1740
```

aaggatcaat ttgaggatat tttgagagag tgtcctggta tttttgactc tctcaacttg 1800
ggccatcagt ctcactttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc 1860
agaactaccg ccgcggtagc ctttttttgcc tttatccttg acaaatggag tcaagaaacc 1920
catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg 1980
tgccagcgcc tgaatgcaat ctccggctac ttgccagtac agccggtaga cacgctgagg 2040
atcctgagtc tccagtcacc ccaggaacac caacgccgcc agcagccgca gcaggagcag 2100
cagcaagagg aggaggagga tcgagaagag aacccgagag ccggtctgga ccctccggtg 2160
gcggaggagg aggagtagct gacttgtttc ccgagctgcg ccgggtgctg actaggtctt 2220
ccagtggacg ggagaggggg attaagcggg agaggcatga ggagactagc cacagaactg 2280
aactgactgt cagtctgatg agccgcaggc gcccagaatc ggtgtggtgg catgaggttc 2340
agtcgcaggg gatagatgag gtctcggtga tgcatgagaa atattccctg gaacaagtca 2400
agacttgttg gttggagcct gaggatgatt gggaggtagc catcaggaat tatgccaagc 2460
tggctctgaa gccagacaag aagtacaaga ttaccaaact gattaatatc agaaattcct 2520
gctacatttc agggaatggg gccgaggtgg agatcagtac ccaggagagg gtggccttca 2580
gatgttgtat gatgaatatg tacccggggg tggtgggcat ggagggagtc acctttatga 2640
acgcgaggtt caggggtgat gggtataatg gggtggtctt tatggccaac accaagctga 2700
cagtgcacgg atgctccttc tttgggttca ataacatgtg catcgaggcc tggggcagtg 2760
tttcagtgag gggatgcagc ttttcagcca actggatggg ggtcgtgggc agaaccaaga 2820
gcaaggtgtc agtgaagaaa tgcctgttcg agaggtgcca cctgggggtg atgagcgagg 2880
gcgaagccaa agtcaaacac tgcgcctcta ctgagacggg ctgctttgtg ctgatcaagg 2940
gcaatgccca agtcaagcat aacatgatct gtggggcctc ggatgagcgc ggctaccaga 3000
tgctgacctg cgccggtggg aacagccata tgctggccac cgtgcatgtg acctcgcacc 3060
cccgcaagac atggcccgag ttcgagcaca acgtcatgac ccgatgcaat gtgcacctgg 3120
ggtcccgccg aggcatgttc atgccctacc agtgcaacat gcaatttgtg aaggtgctgc 3180
tggagcccga tgccatgtcc agagtgagcc tgacgggggt gtttgacatg aatgtggagc 3240
tgtggaaaat tctgagatat gatgaatcca agaccaggtg ccgggcctgc gaatgcggag 3300
gcaagcacgc caggcttcag cccgtgtgtg tggaggtgac ggaggacctg cgacccgatc 3360
atttggtgtt gtcctgcaac gggacggagt tcggctccag cggggaagaa tctgactaga 3420
gtgagtagtg tttgggggag gtggagggct tgtatgaggg gcagaatgac taaaatctgt 3480
gtttttctgt gtgttgcagc agcatgagcg gaagcgcctc ctttgaggga ggggtattca 3540
gcccttatct gacggggcgt ctcccctcct gggcgggagt gcgtcagaat gtgatgggat 3600
ccacggtgga cggccggccc gtgcagcccg cgaactcttc aaccctgacc tacgcgaccc 3660
tgagctcctc gtccgtggac gcagctgccg ccgcagctgc tgcttccgcc gccagcgccg 3720
tgcgcggaat ggccctgggc gccggctact acagctctct ggtggccaac tcgacttcca 3780
ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg 3840
ccctgaccca gcgcctgggc gagctgaccc agcaggtggc tcagctgcag gcggagacgc 3900
gggccgcggt tgccacggtg aaaaccaaat aaaaaatgaa tcaataaata aacggagacg 3960
gttgttgatt ttaacacaga gtcttgaatc tttatttgat ttttcgcgcg cggtaggccc 4020
tggaccaccg gtctcgatca ttgagcaccc ggtggatttt ttccaggacc cggtagaggt 4080
gggcttggat gttgaggtac atgggcatga gcccgtcccg ggggtggagg tagctccatt 4140
gcagggcctc gtgctcgggg gtggtgttgt aaatcaccca gtcatagcag gggcgcaggg 4200
cgtggtgctg cacgatgtcc ttgaggagga gactgatggc cacgggcagc cccttggtgt 4260
aggtgttgac gaacctgttg agctgggagg gatgcatgcg gggggagatg agatgcatct 4320
tggcctggat cttgagattg gcgatgttcc cgcccagatc ccgccggggg ttcatgttgt 4380
gcaggaccac cagcacggtg tatccggtgc acttggggaa tttgtcatgc aacttggaag 4440
ggaaggcgtg aaagaatttg gagacgccct tgtgaccgcc caggttttcc atgcactcat 4500
ccatgatgat ggcgatgggc ccgtgggcgg cggcctgggc aaagacgttt cgggggtcgg 4560
acacatcgta gttgtggtcc tgggtgagct cgtcataggc cattttaatg aatttggggc 4620
ggagggtgcc cgactggggg acgaaggtgc cctcgatccc gggggcgtag ttgccctcgc 4680
agatctgcat ctcccaggcc ttgagctcgg agggggggat catgtccacc tgcggggcga 4740
tgaaaaaaac ggtttccggg gcgggggaga tgagctgggc cgaaagcagg ttccggagca 4800
gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc tgcaggtggt 4860
agttgaggga gagacagctg ccgtcctcgc ggaggagggg ggccacctcg ttcatcatct 4920
cgcgcacatg catgttctcg cgcacgagtt ccgccaggag gcgctcgccc cccagcgaga 4980
ggagctcttg cagcgaggcg aagtttttca gcggcttgag cccgtcggcc atgggcattt 5040

```
tggagagggt ctgttgcaag agttccagac ggtcccagag ctcggtgatg tgctctaggg 5100
catctcgatc cagcagacct cctcgtttcg cgggttgggg cgactgcggg agtagggcac 5160
caggcgatgg gcgtccagcg aggccagggt ccggtccttc cagggtcgca gggtccgcgt 5220
cagcgtggtc tccgtcacgg tgaaggggtg cgcgccgggc tgggcgcttg cgagggtgcg 5280
cttcaggctc atccggctgg tcgagaaccg ctcccggtcg gcgccctgcg cgtcggccag 5340
gtagcaattg agcatgagtt cgtagttgag cgcctcggcc gcgtggccct tggcgcggag 5400
cttacctttg gaagtgtgtc cgcagacggg acagaggagg gacttgaggg cgtagagctt 5460
gggggcgagg aagacggact cgggggcgta ggcgtccgcg ccgcagctgg cgcagacggt 5520
ctcgcactcc acgagccagg tgaggtcggg ccggttgggg tcaaaaacga ggtttcctcc 5580
gtgctttttg atgcgtttct tacctctggt ctccatgagc tcgtgtcccc gctgggtgac 5640
aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cggtcctcga gcggggtgcc 5700
gcggtcctcg tcgtagagga accccgccca ctccgagacg aaggcccggg tccaggccag 5760
cacgaaggag gccacgtggg aggggtagcg gtcgttgtcc accagcgggt ccaccttctc 5820
cagggtatgc aagcacatgt ccccctcgtc cacatccagg aaggtgattg gcttgtaagt 5880
gtaggccacg tgaccggggg tcccggccgg gggggtataa aaggggcgg gcccctgctc 5940
gtcctcactg tcttccggat cgctgtccag gagcgccagc tgttggggta ggtattccct 6000
ctcgaaggct ggcataacct cggcactcag gttgtcagtt tctagaaacg aggaggattt 6060
gatattgacg gtgccgttgg agacgccttt catgagcccc tcgtccatct ggtcagaaaa 6120
gacgatcttt ttgttgtcga gcttggtggc gaaggagccg tagagggcgt tggagaggag 6180
cttggcgatg gagcgcatgg tctggttctt ttccttgtcg gcgcgctcct tggcggcgat 6240
gttgagctgc acgtactcgc gcgccacgca cttccattcg gggaagacgg tggtgagctc 6300
gtcgggcacg attctgaccc gccagccgcg gttgtgcagg gtgatgaggt ccacgctggt 6360
ggccacctcg ccgcgcaggg gctcgttggt ccagcagagg cgcccgccct tgcgcgagca 6420
gaaggggggc agcgggtcca gcatgagctc gtcgggggggg tcggcgtcca cggtgaagat 6480
gccgggcaga agctcggggt cgaagtagct gatgcaggtg tccagatcgt ccagcgccgc 6540
ttgccagtcg cgcacggcca gcgcgcgctc gtaggggctg aggggcgtgc cccagggcat 6600
ggggtgcgtg agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga ggggctcctc 6660
gaggacgccg atgtaggtgg ggtagcagcg ccccccgcgg atgctggcgc gcacgtagtc 6720
gtacagctcg tgcgagggcg cgaggagccc cgtgccgagg ttggagcgtt gcggcttttc 6780
ggcgcggtag acgatctggc ggaagatggc gtgggagttg gaggagatgg tgggcctctg 6840
gaagatgttg aagtgggcgt ggggcaggcc gaccgagtcc ctgatgaagt gggcgtagga 6900
gtcctgcagc ttggcgacga gctcggcggt gacgaggacg tccagggcgc agtagtcgag 6960
ggtctcttgg atgatgtcgt acttgagctg gcccttctgc ttccacagct cgcggttgag 7020
aaggaactct tcgcggtcct tccagtactc ttcgaggggg aacccgtcct gatcggcacg 7080
gtaagagccc accatgtaga actggttgac ggccttgtag gcgcagcagc ccttctccac 7140
ggggagggcg taagcttgtg cggccttgcg cagggaggtg tgggtgaggg cgaaggtgtc 7200
gcgcaccatg accttgagga actggtgctt gaagtcgagg tcgtcgcagc cgccctgctc 7260
ccagagctgg aagtccgtgc gcttcttgta ggcggggttg ggcaaagcga aagtaacatc 7320
gttgaagagg atcttgcccg cgcggggcat gaagttgcga gtgatgcgga aaggctgggg 7380
cacctcggcc cggttgttga tgacctgggc ggcgaggacg atctcgtcga agccgttgat 7440
gttgtgcccg acgatgtaga gttccacgaa tcgcgggcgg cccttaacgt ggggcagctt 7500
cttgagctcg tcgtaggtga gctcggcggg gtcgctgagc ccgtgctgct cgagggccca 7560
gtcggcgacg tgggggttgg cgctgaggaa ggaagtccag agatccacgg ccagggcggt 7620
ctgcaagcgg tcccggtact gacggaactg ctggcccacg gccatttttt cgggggtgac 7680
gcagtagaag gtgcgggggt cgccgtgcca gcggtcccac ttgagctgga gggcgaggtc 7740
gtgggcgagc tcgacgagcg gcgggtcccc ggagagtttc atgaccagca tgaaggggac 7800
gagctgcttg ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg tgaggaagag 7860
cctttcggtg cgaggatgcg agccgatggg gaagaactgg atctcctgcc accagttgga 7920
ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgagc actcgtgctt 7980
gtgtttatac aagcgtccgc agtgctcgca acgctgcacg ggatgcacgt gctgcacgag 8040
ctgtacctgg gttcctttga cgaggaattt cagtgggcag tggagcgctg gcggctgcat 8100
ctggtgctgt actacgtcct ggccatcggc gtggccatcg tctgcctcga tggtggtcat 8160
gctgacgagc ccgcgcggga ggcaggtcca gacttcggct cggacgggtc ggagagcgag 8220
gacgagggcg cgcaggccgg agctgtccag ggtcctgaga cgctgcggag tcaggtcagt 8280
gggcagcggc ggcgcgcggt tgacttgcag gagcttttcc agggcgcgcg ggaggtccag 8340
```

atggtacttg atctccacgg cgccgttggt ggcgacgtcc acggcttgca gggtcccgtg 8400
cccctggggc gccaccaccg tgccccgttt cttcttgggc gctgcttcca tgccggtcag 8460
aagcggcggc gaggacgcgc gccgggcggc aggggcggct cgggacccgg aggcaggggc 8520
ggcaggggca cgtcggcgcc gcgcgcgggc aggttctggt actgcgcccg gagaagactg 8580
gcgtgagcga cgacgcgacg gttgacgtcc tggatctgac gcctctgggt gaaggccacg 8640
ggacccgtga gtttgaacct gaaagagagt tcgacagaat caatctcggt atcgttgacg 8700
gcggcctgcc gcaggatctc ttgcacgtcg cccgagttgt cctggtaggc gatctcggtc 8760
atgaactgct cgatctcctc ctcctgaagg tctccgcggc cggcgcgctc gacggtggcc 8820
gcgaggtcgt tggagatgcg gcccatgagc tgcgagaagg cgttcatgcc ggcctcgttc 8880
cagacgcggc tgtagaccac ggctccgtcg gggtcgcgcg cgcgcatgac cacctgggcg 8940
aggttgagct cgacgtggcg cgtgaagacc gcgtagttgc agaggcgctg gtagaggtag 9000
ttgagcgtgg tggcgatgtg ctcggtgacg aagaagtaca tgatccagcg gcggagcggc 9060
atctcgctga cgtcgcccag ggcttccaag cgctccatgg cctcgtagaa gtccacggcg 9120
aagttgaaaa actgggagtt gcgcgccgag acggtcaact cctcctccag aagacggatg 9180
agctcagcga tggtggcgcg cacctcgcgc tcgaaggccc cgggggctc ctcttcttcc 9240
atctcttcct cctccactaa catctcttct acttcctcct caggaggcgg cggcggggga 9300
ggggccctgc gtcgccggcg gcgcacgggc agacggtcga tgaagcgctc gatggtctcc 9360
ccgcgccggc gacgcatggt ctcggtgacg gcgcgcccgt cctcgcgggg ccgcagcgtg 9420
aagacgccgc cgcgcatctc caggtggccg ccgggggggt ctccgttggg cagggagagg 9480
gcgctgacga tgcatcttat caattggccc gtagggactc cgcgcaagga cctgagcgtc 9540
tcgagatcca cgggatccga aaaccgctga acgaaggctt cgagccagtc gcagtcgcaa 9600
ggtaggctga gcccggtttc ttgttcttcg gggatttcgg gaggcgggcg ggcgatgctg 9660
ctggtgatga agttgaagta ggcggtcctg agacggcgga tggtggcgag gagcaccagg 9720
tccttgggcc cggcttgctg gatgcgcaga cggtcggcca tgccccaggc gtggtcctga 9780
cacctggcga ggtccttgta gtagtcctgc atgagccgct ccacgggcac ctcctcctcg 9840
cccgcgcggc cgtgcatgcg cgtgagcccg aacccgcgct ggggctggac gagcgccagg 9900
tcggcgacga cgcgctcggc gaggatggcc tgctgtatct gggtgagggt ggtctggaag 9960
tcgtcgaagt cgacgaagcg gtggtaggct ccggtgttga tggtatagga gcagttggcc 10020
atgacggacc agttgacggt ctggtggccg ggtcgcacga gctcgtggta cttgaggcgc 10080
gagtaggcgc gcgtgtcgaa gatgtagtcg ttgcaggtgc gcacgaggta ctggtatccg 10140
acgaggaagt gcggcggcgg ctggcggtag agcggccatc gctcggtggc gggggcgccg 10200
ggcgcgaggt cctcgagcat gaggcggtgg tagccgtaga tgtacctgga catccaggtg 10260
atgccggcgg cggtggtgga ggcgcgcggg aactcgcgga cgcggttcca gatgttgcgc 10320
agcggcagga agtagttcat ggtggccgcg gtctggcccg tgaggcgcgc gcagtcgtgg 10380
atgctctaga catacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag 10440
gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag 10500
ccgcagctaa cgtggtactg gcactcccgt ctcgacccaa gcctgctaac gaaacctcca 10560
ggatacggag gcgggtcgtt ttttggcctt ggtcgctggt catgaaaaac tagtaagcgc 10620
ggaaagcgac cgcccgcgat ggctcgctgc cgtagtctgg agaaagaatc gccagggttg 10680
cgttgcggtg tgccccggtt cgagcctcag cgctcggcgc cggccggatt ccgcggctaa 10740
cgtgggcgtg gctgccccgt cgtttccaag acccccttagc cagccgactt ctccagttac 10800
ggagcgagcc cctcttttttc ttgtgttttt gccagatgca tcccgtactg cggcagatgc 10860
gcccccaccc tccacctcaa ccgcccctac cgccgcagca gcagcaacag ccggcgcttc 10920
tgcccccgcc ccagcagcag ccagccacta ccgcggcggc cgccgtgagc ggagccggcg 10980
ttcagtatga cctggccttg gaagagggcg aggggctggc gcggctgggg gcgtcgtcgc 11040
cggagcggca cccgcgcgtg cagatgaaaa gggacgctcg cgaggcctac gtgcccaagc 11100
agaacctgtt cagagacagg agcggcgagg agcccgagga gatgcgcgcc tcccgcttcc 11160
acgcggggcg ggagctgcgg cgcggcctgg accgaaagcg ggtgctgagg gacgaggatt 11220
tcgaggcgga cgagctgacg gggatcagcc ccgcgcgcgc gcacgtggcc gcggccaacc 11280
tggtcacggc gtacgagcag accgtgaagg aggagagcaa cttccaaaaa tccttcaaca 11340
accacgtgcg cacgctgatc gcgcgcgagg aggtgaccct gggcctgatg cacctgtggg 11400
acctgctgga ggccatcgtg cagaacccca cgagcaagcc gctgacggcg cagctgtttc 11460
tggtggtgca gcacagtcgg gacaacgaga cgttcaggga ggcgctgctg aatatcaccg 11520
agcccgaggg ccgctggctc ctggacctgg tgaacattct gcagagcatc gtggtgcagg 11580
agcgcgggct gccgctgtcc gagaagctgg cggctatcaa cttctcggtg ctgagcctgg 11640

gcaagtacta cgctaggaag atctacaaga ccccgtacgt gcccatagac aaggaggtga 11700
agatcgacgg gttttacatg cgcatgaccc tgaaagtgct gaccctgagc gacgatctgg 11760
gggtgtaccg caacgacagg atgcaccgcg cggtgagcgc cagccgccgg cgcgagctga 11820
gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgaggggg 11880
agagctactt tgacatgggc gcggacctgc gctggcagcc cagccgccgg gccttggaag 11940
ctgccggcgg ttccccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000
tggaagactg atggcgcgac cgtatttttg ctagatgcag caacagccac cgcctcctga 12060
tcccgcgatg cgggcggcgc tgcagagcca gccgtccggc attaactcct cggacgattg 12120
gacccaggcc atgcaacgca tcatggcgct gacgacccgc aatcccgaag cctttagaca 12180
gcagcctcag gccaaccggc tctcggccat cctggaggcc gtggtgccct cgcgctcgaa 12240
ccccacgcac gagaaggtgc tggccatcgt gaacgcgctg gtggagaaca aggccatccg 12300
cggcgacgag gccgggctgg tgtacaacgc gctgctggag cgcgtggccc gctacaacag 12360
caccaacgtg cagacgaacc tggaccgcat ggtgaccgac gtgcgcgagg cggtgtcgca 12420
gcgcgagcgg ttccaccgcg agtcgaacct gggctccatg gtggcgctga acgccttcct 12480
gagcacgcag cccgccaacg tgccccgggg ccaggaggac tacaccaact tcatcagcgc 12540
gctgcggctg atggtggccg aggtgcccca gagcgaggtg taccagtcgg ggccggacta 12600
cttcttccag accagtcgcc agggcttgca gaccgtgaac ctgagccagg ctttcaagaa 12660
cttgcaggga ctgtggggcg tgcaggcccc ggtcggggac cgcgcgacgg tgtcgagcct 12720
gctgacgccg aactcgcgcc tgctgctgct gctggtggcg cccttcacgg acagcggcag 12780
cgtgagccgc gactcgtacc tgggctacct gcttaacctg taccgcgagg ccatcgggca 12840
ggcgcacgtg gacgagcaga cctaccagga gatcacccac gtgagccgcg cgctgggcca 12900
ggaggacccg ggcaacctgg aggccaccct gaacttcctg ctgaccaacc ggtcgcagaa 12960
gatcccgccc cagtacgcgc tgagcaccga ggaggagcgc atcctgcgct acgtgcagca 13020
gagcgtgggg ctgttcctga tgcaggaggg ggccacgccc agcgccgcgc tcgacatgac 13080
cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat 13140
ggactacttg catcgggcgg ccgccatgaa ctcggactac tttaccaacg ccatcttgaa 13200
cccgcactgg ctcccgccgc ccgggttcta cacgggcgag tacgacatgc ccgaccccaa 13260
cgacgggttc ctgtgggacg acgtggacag cagcgtgttc tcgccgcgcc ccgccaccac 13320
cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg gtcgcgcggg 13380
tgctgccgcg gcggtgcctg aggccgccag ccccttcccg agcctgccct tttcgctgaa 13440
cagcgtgcgc agcagcgagc tgggtcggct gacgcggccg cgcctgctgg gcgaggagga 13500
gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttcccca ataacgggat 13560
agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc acagggacga 13620
gccccgagct agcagcagcg caggcacccg tagacgccag cgacacgaca ggcagcgggg 13680
tctggtgtgg gacgatgagg attccgccga cgacagcagc gtgttggact tgggtgggag 13740
tggtggtggt aacccgttcg ctcacttgcg cccccgtatc gggcgcctga tgtaagaatc 13800
tgaaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttctctgt 13860
tgtttgtagt agtatgatga ggcgcgtgta cccggagggt cctcctccct cgtacgagag 13920
cgtgatgcag caggcggtgg cggcggcgat gcagcccccg ctggaggcgc cttacgtgcc 13980
cccgcggtac ctggcgccta cggaggggcg gaacagcatt cgttactcgg agctggcacc 14040
cttgtacgat accacccggt tgtacctggt ggacaacaag tcggcggaca tcgcctcgct 14100
gaactaccag aacgaccaca gcaacttcct gaccaccgtg gtgcagaaca acgatttcac 14160
ccccacggag gccagcaccc agaccatcaa ctttgacgag cgctcgcggt ggggcggcca 14220
gctgaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa 14280
gttcaaggcg cgggtgatgg tctcgcgcaa gacccccaat ggggtcgcgg tggatgagaa 14340
ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccgaggg 14400
caacttctcg gtgaccatga ccatcgatct gatgaacaac gccatcatcg acaactactt 14460
ggcggtgggg cgtcagaacg gggtgctgga gagcgacatc ggcgtgaagt tcgacacgcg 14520
caacttccgg ctgggctggg accccgtgac cgagctggtg atgccgggcg tgtacaccaa 14580
cgaggccttc caccccgaca tcgtcctgct gcccggctgc ggcgtggact tcaccgagag 14640
ccgcctcagc aacctgctgg gcatccgcaa gcggcagccc ttccaggagg gcttccagat 14700
cctgtacgag gacctggagg ggggcaacat ccccgcgctc ttggatgtcg aagcctatga 14760
gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccg cctctaccga 14820
ggtgcggggc gataattttg ctagcgccgc ggcagtggcc gaggcggctg aaaccgaaag 14880
taagatagtc atccagccgg tggagaagga cagcaaggac aggagctaca acgtgctcgc 14940

ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgaccccga 15000
gaagggcgtg cgctcctgga cgctgctcac cacctcggac gtcacctgcg gcgtggagca 15060
agtctactgg tcgctgcccg acatgatgca agacccggtc accttccgct ccacgcgtca 15120
agttagcaac tacccggtgg tgggcgccga gctcctgccc gtctactcca agagcttctt 15180
caacgagcag gccgtctact cgcagcagct gcgcgccttc acctcgctca cgcacgtctt 15240
caaccgcttc cccgagaacc agatcctcgt ccgcccgccc gcgcccacca ttaccaccgt 15300
cagtgaaaac gttcctgctc tcacagatca cgggaccctg ccgctgcgca gcagtatccg 15360
gggagtccag cgcgtgaccg tcactgacgc cagacgccgc acctgcccct acgtctacaa 15420
ggccctgggc gtagtcgcgc cgcgcgtcct ctcgagccgc accttctaaa aaatgtccat 15480
tctcatctcg cccagtaata acaccggttg gggcctgcgc gcgcccagca agatgtacgg 15540
aggcgctcgc caacgctcca cgcaacaccc cgtgcgcgtg cgcgggcact tccgcgctcc 15600
ctggggcgcc ctcaagggcc gcgtgcgctc gcgcaccacc gtcgacgacg tgatcgacca 15660
ggtggtggcc gacgcgcgca actacacgcc cgccgccgcg cccgcctcca ccgtggacgc 15720
cgtcatcgac agcgtggtgg ccgatgcgcg ccggtacgcc cgcgccaaga gccggcggcg 15780
gcgcatcgcc cggcggcacc ggagcacccc cgccatgcgc gcggcgcgag ccttgctgcg 15840
cagggccagg cgcacgggac gcagggccat gctcagggcg gccagacgcg cggcctccgg 15900
cagcagcagc gccggcagga cccgcagacg cgcggccacg gcggcggcgg cggccatcgc 15960
cagcatgtcc cgcccgcggc gcggcaacgt gtactgggtg cgcgacgccg ccaccggtgt 16020
gcgcgtgccc gtgcgcaccc gccccctcg cacttgaaga tgctgacttc gcgatgttga 16080
tgtgtcccag cggcgaggag gatgtccaag cgcaaataca aggaagagat gctccaggtc 16140
atcgcgcctg agatctacgg ccccgcggtg aaggaggaaa gaaagccccg caaactgaag 16200
cgggtcaaaa aggacaaaaa ggaggaggaa gatgtggacg gactggtgga gtttgtgcgc 16260
gagttcgccc cccggcggcg cgtgcagtgg cgcgggcgga aagtgaaacc ggtgctgcgg 16320
cccggcacca cggtggtctt cacgcccggc gagcgttccg gctccgcctc caagcgctcc 16380
tacgacgagg tgtacgggga cgaggacatc ctcgagcagg cggtcgagcg tctgggcgag 16440
tttgcttacg gcaagcgcag ccgccccgcg cccttgaaag aggaggcggt gtccatcccg 16500
ctggaccacg gcaaccccac gccgagcctg aagccggtga ccctgcagca ggtgctgccg 16560
agcgcggcgc cgcgccgggg cttcaagcgc gagggcggcg aggatctgta cccgaccatg 16620
cagctgatgg tgcccaagcg ccagaagctg gaggacgtgc tggagcacat gaaggtggac 16680
cccgaggtgc agcccgaggt caaggtgcgg cccatcaagc aggtggcccc gggcctgggc 16740
gtgcagaccg tggacatcaa gatccccacg gagcccatgg aaacgcagac cgagcccgtg 16800
aagcccagca ccagcaccat ggaggtgcag acggatccct ggatgccggc gccggcttcc 16860
accactcgcc gaagacgcaa gtacggcgcg gccagcctgc tgatgcccaa ctacgcgctg 16920
catccttcca tcatccccac gccgggctac cgcggcacgc gcttctaccg cggctacacc 16980
agcagccgcc gcaagaccac cacccgccgc cgccgtcgtc gcacccgccg cagcagcacc 17040
gcgacttccg ccgccgccct ggtgcggaga gtgtaccgca gcgggcgcga gcctctgacc 17100
ctgccgcgcg cgcgctacca cccgagcatc gccatttaac tctgccgtcg cctcctactt 17160
gcagatatgg ccctcacatg ccgcctccgc gtccccatta cgggctaccg aggaagaaag 17220
ccgcgccgta gaaggctgac ggggaacggg ctgcgtcgcc atcaccaccg gcggcggcgc 17280
gccatcagca agcggttggg gggaggcttc ctgcccgcgc tgatccccat catcgccgcg 17340
gcgatcgggg cgatccccgg catagcttcc gtggcggtgc aggcctctca gcgccactga 17400
gacacagctt ggaaaatttg taataaaaaa atggactgac gctcctggtc ctgtgatgtg 17460
tgtttttaga tggaagacat caatttttcg tccctggcac cgcgacacgg cacgcggccg 17520
tttatgggca cctggagcga catcggcaac agccaactga acgggggcgc cttcaattgg 17580
agcagtctct ggagcgggct taagaatttc gggtccacgc tcaaaacctа tggcaacaag 17640
gcgtggaaca gcagcacagg gcaggcgctg agggaaaagc tgaaagagca gaacttccag 17700
cagaaggtgg tcgatggcct ggcctcgggc atcaacgggg tggtggacct ggccaaccag 17760
gccgtgcaga aacagatcaa cagccgcctg gacgcggtcc cgcccgcggg gtccgtggag 17820
atgccccagg tggaggagga gctgcctccc ctggacaagc gcggcgacaa gcgaccgcgt 17880
cccgacgcgg aggagacgct gctgacgcac acggacgagc cgcccccgta cgaggaggcg 17940
gtgaaactgg gtctgcccac cacgcggccc gtggcgcctc tggccaccgg ggtgctgaaa 18000
cccagcagca gcagccagcc cgcgaccctg gacttgcctc cgcctgcttc ccgcccctcc 18060
acagtggcta agcccctgcc gccggtggcc gtcgcgtcgc gcgccccccg aggccgcccc 18120
caggcgaact ggcagagcac tctgaacagc atcgtgggtc tgggagtgca gagtgtgaag 18180
cgccgccgct gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt 18240

atgtccgccg accagaagga ggaagaggcg cgtcgccgag ttgcaagatg gccaccccat 18300
cgatgctgcc ccagtgggcg tacatgcaca tcgccggaca ggacgcttcg gagtacctga 18360
gtccgggtct ggtgcagttc gcccgcgcca cagacaccta cttcagtctg gggaacaagt 18420
ttaggaaccc cacggtggcg cccacgcacg atgtgaccac cgaccgcagc cagcggctga 18480
cgctgcgctt cgtgcccgtg gaccgcgagg acaacaccta ctcgtacaaa gtgcgctaca 18540
cgctggccgt gggcgacaac cgcgtgctgg acatggccag cacctacttt gacatccgcg 18600
gcgtgctgga tcgggggccc agcttcaaac cctactccgg caccgcctac aacagcctgg 18660
ctcccaaggg agcgcccaac acttgccagt ggacatataa agctggtgat actgatacag 18720
aaaaaaccta tacatatgga aatgcacctg tgcaaggcat tagcattaca aaggatggta 18780
ttcaacttgg aactgacagc gatggtcagg caatctatgc agacgaaact tatcaaccag 18840
agcctcaagt gggtgatgct gaatggcatg acatcactgg tactgatgaa aaatatggag 18900
gcagagctct taagcctgac accaaaatga agccttgcta tggttctttt gccaagccta 18960
ccaataaaga aggaggccag gcaaatgtga aaaccgaaac aggcggtacc aaagaatatg 19020
acattgacat ggcattcttc gataatcgaa gtgcagctgc cgccggccta gccccagaaa 19080
ttgttttgta tactgagaat gtggatctgg aaactccaga tacccatatt gtatacaagg 19140
caggtacaga tgacagtagc tcttctatca atttgggtca gcagtccatg cccaacagac 19200
ccaactacat tggcttcaga gacaacttta tcggtctgat gtactacaac agcactggca 19260
atatgggtgt actggctgga caggcctccc agctgaatgc tgtggtggac ttgcaggaca 19320
gaaacaccga actgtcctac cagctcttgc ttgactctct gggtgacaga accaggtatt 19380
tcagtatgtg gaatcaggcg gtggacagtt atgaccccga tgtgcgcatt attgaaaatc 19440
acggtgtgga ggatgaactt cctaactatt gcttcccccc ggatgctgtg ggtagaactg 19500
atacttacca gggaattaag gccaatggtg ataatcaaac cacctggacc aaagatgata 19560
ctgttaatga tgctaatgaa ttgggcaagg gcaatccttt cgccatggag atcaacatcc 19620
aggccaacct gtggcggaac ttcctctacg cgaacgtggc gctgtacctg cccgactcct 19680
acaagtacac gccggccaac atcacgctgc ccaccaacac caacacctac gattacatga 19740
acggccgcgt ggtggcgccc tcgctggtgg acgcctacat caacatcggg gcgcgctggt 19800
cgctggaccc catggacaac gtcaaccccct tcaaccacca ccgcaacgcg ggcctgcgat 19860
accgctccat gctcctgggc aacgggcgct acgtgccctt ccacatccag gtgccccaaa 19920
agtttttcgc catcaagagc ctcctgctcc tgcccgggtc ctacacctac gagtggaact 19980
tccgcaagga cgtcaacatg atcctgcaga gctccctcgg caacgacctg cgcacggacg 20040
gggcctccat cgccttcacc agcatcaacc tctacgccac cttcttcccc atggcgcaca 20100
acaccgcctc cacgctcgag gccatgctgc gcaacgacac caacgaccag tccttcaacg 20160
actacctctc ggcggccaac atgctctacc ccatcccggc caacgccacc aacgtgccca 20220
tctccatccc ctcgcgcaac tgggccgcct tccgcggctg gtccttcacg cgcctcaaga 20280
cccgcgagac gccctcgctc ggctccgggt tcgaccccta cttcgtctac tcgggctcca 20340
tcccctacct cgacggcacc ttctacctca accacacctt caagaaggtc tccatcacct 20400
tcgactcctc cgtcagctgg cccggcaacg accgcctcct gacgcccaac gagttcgaaa 20460
tcaagcgcac cgtcgacgga gaggggtaca acgtggccca gtgcaacatg accaaggact 20520
ggttcctggt ccagatgctg gcccactaca acatcggcta ccagggcttc tacgtgcccg 20580
agggctacaa ggaccgcatg tactccttct tccgcaactt ccagcccatg agccgccagg 20640
tcgtggacga ggtcaactac aaggactacc aggccgtcac cctggcctac cagcacaaca 20700
actcgggctt cgtcggctac ctcgcgccca ccatgcgcca gggccagccc taccccgcca 20760
actaccccta cccgctcatc ggcaagagcg ccgtcgccag cgtcacccag aaaaagttcc 20820
tctgcgaccg ggtcatgtgg cgcatccccct tctccagcaa cttcatgtcc atgggcgcgc 20880
tcaccgacct cggccagaac atgctctacg ccaactccgc ccacgcgcta gacatgaatt 20940
tcgaagtcga ccccatggat gagtccaccc ttctctatgt tgtcttcgaa gtcttcgacg 21000
tcgtccgagt gcaccagccc caccgcggcg tcatcgaggc cgtctacctg cgcacgccct 21060
tctcggccgg caacgccacc acctaagcct cttgcttctt gcaagatgac ggcctgcgcg 21120
ggctccggcg agcaggagct cagggccatc ctccgcgacc tgggctgcgg gccctgcttc 21180
ctgggcacct tcgacaagcg cttcccggga ttcatggccc cgcacaagct ggcctgcgcc 21240
atcgtcaaca cggccggccg cgagaccggg ggcgagcact ggctggcctt cgcctggaac 21300
ccgcgctccc acacctgcta cctcttcgac cccttcgggt tctcggacga gcgcctcaag 21360
cagatctacc agttcgagta cgagggcctg ctgcgtcgca gcgccctggc caccgaggac 21420
cgctgcgtca ccctggaaaa gtccacccag accgtgcagg gtccgcgctc ggccgcctgc 21480
gggctcttct gctgcatgtt cctgcacgcc ttcgtgcact ggcccgaccg ccccatggac 21540

aagaacccca ccatgaactt gctgacgggg gtgcccaacg gcatgctcca gtcgccccag 21600
gtggaaccca ccctgcgccg caaccaggag gcgctctacc gcttcctcaa cgcccactcc 21660
gcctactttc gctcccaccg cgcgcgcatc gagaaggcca ccgccttcga ccgcatgaat 21720
caagacatgt aatccggtgt gtgtatgtga atgctttatt catcataata aacagcacat 21780
gtttatgcca ccttctctga ggctctgact ttatttagaa atcgaagggg ttctgccggc 21840
tctcggcatg gcccgcgggc agggatacgt tgcggaactg gtacttgggc agccacttga 21900
actcggggat cagcagcttc ggcacgggga ggtcggggaa cgagtcgctc cacagcttgc 21960
gcgtgagttg cagggcgccc agcaggtcgg gcgcggagat cttgaaatcg cagttgggac 22020
ccgcgttctg cgcgcgagag ttacggtaca cggggttgca gcactggaac accatcaggg 22080
ccgggtgctt cacgctcgcc agcaccgtcg cgtcggtgat gccctccacg tccagatcct 22140
cggcgttggc catcccgaag ggggtcatct tgcaggtctg ccgccccatg ctgggcacgc 22200
agccgggctt gtggttgcaa tcgcagtgca gggggatcag catcatctgg gcctgctcgg 22260
agctcatgcc cgggtacatg gccttcatga aagcctccag ctggcggaag gcctgctgcg 22320
ccttgccgcc ctcggtgaag aagaccccgc aggacttgct agagaactgg ttggtggcgc 22380
agccagcgtc gtgcacgcag cagcgcgcgt cgttgttggc cagctgcacc acgctgcgcc 22440
cccagcggtt ctgggtgatc ttggcccggt cggggttctc cttcagcgcg cgctgcccgt 22500
tctcgctcgc cacatccatc tcgatcgtgt gctccttctg gatcatcacg gtcccgtgca 22560
ggcaccgcag cttgccctcg gcctcggtgc acccgtgcag ccacagcgcg cagccggtgc 22620
tctcccagtt cttgtgggcg atctgggagt gcgagtgcac gaagccctgc aggaagcggc 22680
ccatcatcgt ggtcagggtc ttgttgctgg tgaaggtcag cggaatgccg cggtgctcct 22740
cgttcacata caggtggcag atacggcggt acacctcgcc ctgctcgggc atcagctgga 22800
aggcggactt caggtcgctc tccacgcggt accggtccat cagcagcgtc atcacttcca 22860
tgcccttctc ccaggccgaa acgatcggca ggctcagggg gttcttcacc gttgtcatct 22920
tagtcgccgc cgccgaagtc agggggtcgt tctcgtccag ggtctcaaac actcgcttgc 22980
cgtccttctc ggtgatgcgc acgggggggaa agctgaagcc cacggccgcc agctcctcct 23040
cggcctgcct ttcgtcctcg ctgtcctggc tgatgtcttg caaaggcaca tgcttggtct 23100
tgcgggggttt ctttttgggc ggcagaggcg gcggcggaga cgtgctgggc gagcgcgagt 23160
tctcgctcac cacgactatt tcttctcctt ggccgtcgtc cgagaccacg cggcggtagg 23220
catgcctctt ctggggcaga ggcggaggcg acgggctctc gcggttcggc gggcggctgg 23280
cagagccgct tccgcgttcg ggggtgcgct cctggcggcg ctgctctgac tgacttcctc 23340
cgcggccggc cattgtgttc tcctagggag caagcatgga gactcagcca tcgtcgccaa 23400
catcgccatc tgcccccgcc gccgccgacg agaaccagca gcagcagaat gaaagcttaa 23460
ccgccccgcc gcccagcccc acctccgacg ccgcagcccc agacatgcaa gagatggagg 23520
aatccatcga gattgacctg ggctacgtga cgcccgcgga gcacgaggag gagctggcag 23580
cgcgcttttc agccccggaa gagaaccacc aagagcagcc agagcaggaa gcagagagcg 23640
agcagaacca ggctgggctc gagcatggcg actacctgag cggggcagag gacgtgctca 23700
tcaagcatct ggcccgccaa tgcatcatcg tcaaggacgc gctgctcgac cgcgccgagg 23760
tgcccctcag cgtggcggag ctcagccgcg cctacgagcg caacctcttc tcgccgcgcg 23820
tgcccccaa gcgccagccc aacggcacct gcgagcccaa cccgcgcctc aacttctacc 23880
cggtcttcgc ggtgcccgag gccctggcca cctaccacct ctttttcaag aaccaaagga 23940
tccccgtctc ctgccgcgcc aaccgcaccc gcgccgacgc cctgctcaac ctgggccccg 24000
gcgcccgcct acctgatatc gcctccttgg aagaggttcc caagatcttc gagggtctgg 24060
gcagcgacga gactcgggcc gcgaacgctc tgcaaggaag cggagaggag catgagcacc 24120
acagcgccct ggtggagttg gaaggcgaca acgcgcgcct ggcggtcctc aagcgcacgg 24180
tcgagctgac ccacttcgcc tacccggcgc tcaacctgcc ccccaaggtc atgagcgccg 24240
tcatggacca ggtgctcatc aagcgcgcct cgcccctctc ggaggaggag atgcaggacc 24300
ccgagagctc ggacgagggc aagcccgtgg tcagcgacga gcagctggcg cgctggctgg 24360
gagagcagtc aggcagagga ggaggagatg gaagactggg acagcactca ggcagaggag 24420
gagcgagtag cacccccag agcctggaag agcggcgcaa gctcatgatg gccgtggtcc 24480
tggtgaccgt ggagctggag tgtctgcgcc gcttcttcgc cgacgcggag accctgcgca 24540
aggtcgagga gaacctgcac tacctcttca gacacgggtt cgtgcgccag gcctgcaaga 24600
tctccaacgt ggagctgacc aacctggtct cctacatggg catcctgcac gagaaccgcc 24660
tggggcagaa cgtgctgcac accaccctgc gcggggaggc ccgccgcgac tacatccgcg 24720
actgcgtcta cctgtacctc tgccacacct ggcagacggg catgggcgtg tggcagcagt 24780
gcctggagga gcagaacctg aaagagctct gcaagctcct gcagaagaac ctcaaggccc 24840

```
tgtggaccgg gttcgacgag cgcaccaccg ccgcggacct ggccgacctc atcttccccg 24900
agcgcctgcg gctgacgctg cgcaacgggc tgcccgactt tatgagccaa agcatgttgc 24960
aaaactttcg ctctttcatc ctcgaacgct ccgggatcct gcccgccacc tgctccgcgc 25020
tgccctcgga cttcgtgccg ctgaccttcc gcgagtgccc cccgccgctc tggagccact 25080
gctacctgct gcgcctggcc aactacctgg cctaccactc ggacgtgatc gaggacgtca 25140
gcggcgaggg cctgctcgag tgccactgcc gctgcaacct ctgcacgccg caccgctccc 25200
tggcctgcaa cccccagctg ctgagcgaga cccagatcat cggcaccttc gagttgcaag 25260
gccccggcga gggcaagggg ggtctgaaac tcaccccggg gctgtggacc tcggcctact 25320
tgcgcaagtt cgtgcccgag gactaccatc ccttcgagat caggttctac gaggaccaat 25380
cccagccgcc caaggccgag ctgtcggcct gcgtcatcac ccaggggccc atcctggccc 25440
aattgcaagc catccagaaa tcccgccaag aatttctgct gaaaaagggc cacggggtct 25500
acttggaccc ccagaccgga gaggagctca accccagctt cccccaggat gccccgagga 25560
agcagcaaga agctgaaagt ggagctgccg ccgccgccgg aggatttgga ggaagactgg 25620
gacagcctgc aagacagtct ggaggaggaa gacgaggtgg aggaggcaga ggaagaagca 25680
gccgccgcca gaccgtcgtc ctcggcggag gaggagaaag caagcagcac ggataccatc 25740
tccgctccgg gtcggggtcg cggcggccgg gcccacagta gatgggacga gaccgggcgc 25800
ttcccgaacc ccaccaccca gaccggtaag aaggagcggc agggatacaa gtcctggcgg 25860
gggcacaaaa acgccatcgt ctcctgcttg caagcctgcg ggggcaacat ctccttcacc 25920
cggcgctacc tgctcttcca ccgcggggtg aacttccccc gcaacatctt gcattactac 25980
cgtcacctcc acagccccta ctactgtttc caagaagagg cagaaaccca gcagcagcag 26040
cagcagcaga aaaccagcgg cagcagctag aaaatccaca gcggcggcag gtggactgag 26100
gatcgcggcg aacgagccgg cgcagacccg ggagctgagg aaccggatct ttcccaccct 26160
ctatgccatc ttccagcaga gtcgggggca agagcaggaa ctgaaagtca agaaccgttc 26220
tctgcgctcg ctcacccgca gttgtctgta tcacaagagc gaagaccaac ttcagcgcac 26280
tctcgaggac gccgaggctc tcttcaacaa gtactgcgcg ctcactctta aagagtagcc 26340
cgcgcccgcc cacacacgga aaaaggcggg aattacgtca ccacctgcgc ccttcgcccg 26400
accatcatca tgagcaaaga gattcccacg ccttacatgt ggagctacca gccccagatg 26460
ggcctggccg ccggcgccgc ccaggactac tccacccgca tgaactggct cagtgccggg 26520
cccgcgatga tctcacgggt gaatgacatc cgcgcccacc gaaaccagat actcctagaa 26580
cagtcagcga tcaccgccac gccccgccat caccttaatc cgcgtaattg gcccgccgcc 26640
ctggtgtacc aggaaattcc ccagcccacg accgtactac ttccgcgaga cgcccaggcc 26700
gaagtccagc tgactaactc aggtgtccag ctggccggcg gcgccgccct gtgtcgtcac 26760
cgccccgctc agggtataaa gcggctggtg atccgaggca gaggcacaca gctcaacgac 26820
gaggtggtga gctcttcgct gggtctgcga cctgacggag tcttccaact cgccggatcg 26880
gggagatctt ccttcacgcc tcgtcaggcc gtcctgactt tggagagttc gtcctcgcag 26940
ccccgctcgg gtggcatcgg cactctccag ttcgtggagg agttcactcc ctcggtctac 27000
ttcaacccct tctccggctc ccccggccac tacccggacg agttcatccc gaacttcgac 27060
gccatcagcg agtcggtgga cggctacgat tgaatgtccc atggtggcgc ggctgaccta 27120
gctcggcttc gacacctgga ccactgccgc cgcttccgct gcttcgctcg ggatctcgcc 27180
gagtttgcct actttgagct gcccgaggag caccctcagg gcccggccca cggagtgcgg 27240
atcgtcgtcg aaggggggtct cgactcccac ctgcttcgga tcttcagcca gcgtccgatc 27300
ctggccgagc gcgagcaagg acagaccctt ctgaccctgt actgcatctg caaccacccc 27360
ggcctgcatg aaagtctttg ttgtctgctg tgtactgagt ataataaaag ctgagatcag 27420
cgactactcc ggacttccgt gtgttcctgc tatcaaccag tccctgttct tcaccgggaa 27480
cgagaccgag ctccagctcc agtgtaagcc ccacaagaag tacctcacct ggctgttcca 27540
gggctctccg atcgccgttg tcaaccactg cgacaacgac ggagtcctgc tgagcggccc 27600
tgccaacctt acttttcca cccgcagaag caagctccag ctcttccaac ccttcctccc 27660
cgggacctat cagtgcgtct cgggaccctg ccatcacacc ttccacctga tcccgaatac 27720
cacagcgtcg ctccccgcta ctaacaacca aactacccac caacgccacc gtcgcgacct 27780
ttcctctggg tctaatacca ctaccggagg tgagctccga ggtcgaccaa cctctgggat 27840
ttactacggc ccctgggagg tggtagggtt aatagcgcta ggcctagttg cgggtgggct 27900
tttggctctc tgctacctat acctcccttg ctgttcgtac ttagtggtgc tgtgttgctg 27960
gtttaagaaa tggggaagat caccctagtg agctgcggtg tgctggtggc ggtggtgctt 28020
tcgattgtgg gactgggcgg cgcggctgta gtgaaggaga aggccgatcc ctgcttgcat 28080
ttcaatcccg acaaatgcca gctgagtttt cagcccgatg gcaatcggtg cgcggtgctg 28140
```

atcaagtgcg gatgggaatg cgagaacgtg agaatcgagt acaataacaa gactcggaac 28200
aatactctcg cgtccgtgtg gcagcccggg gaccccgagt ggtacaccgt ctctgtcccc 28260
ggtgctgacg gctccccgcg caccgtgaat aatactttca tttttgcgca catgtgcgac 28320
acggtcatgt ggatgagcaa gcagtacgat atgtggcccc ccacgaagga gaacatcgtg 28380
gtcttctcca tcgcttacag cgtgtgcacg gcgctaatca ccgctatcgt gtgcctgagc 28440
attcacatgc tcatcgctat tcgccccaga aataatgccg aaaaagaaaa acagccataa 28500
cacgtttttt cacacacctt tttcagacca tggcctctgt taaatttttg cttttatttg 28560
ccagtctcat tgccgtcatt catggaatga gtaatgagaa aattactatt tacactggca 28620
ctaatcacac attgaaaggt ccagaaaaag ccacagaagt ttcatggtat tgttatttta 28680
atgaatcaga tgtatctact gaactctgtg gaaacaataa caaaaaaaat gagagcatta 28740
ctctcatcaa gtttcaatgt ggatctgact taaccctaat taacatcact agagactatg 28800
taggtatgta ttatggaact acagcaggca tttcggacat ggaatttat caagtttctg 28860
tgtctgaacc caccacgcct agaatgacca caaccacaaa aactacacct gttaccacta 28920
tacagctcac taccaatggc tttcttgcca tgcttcaagt ggctgaaaat agcaccagca 28980
ttcaacccac cccacccagt gaggaaattc ccagatccat gattggcatt attgttgctg 29040
tagtggtgtg catgttgatc atcgccttgt gcatggtgta ctatgccttc tgctacagaa 29100
agcacagact gaacgacaag ctggaacact tactaagtgt tgaatttaa ttttttagaa 29160
ccatgaagat cctaggcctt ttagtttttt ctatcattac ctctgctcta tgcaattctg 29220
acaatgagga cgttactgtc gttgtcggat caaattatac actaaaaggt ccagcaaaag 29280
gtatgctttc gtggtattgt tggttcggaa ctgacgagca acagacagaa ctttgcaatg 29340
ctcaaaaagg caaaacctca aattctaaaa tctctaatta tcaatgcaat ggcactgact 29400
tagtattgct caatgtcacg aaagcatatg ctggcagtta cacctgccct ggagatgatg 29460
ccgacaatat gatttttac aaagtggaag tggttgatcc cactactcca ccgcccacca 29520
ccacaactac tcataccaca cacacagaac aaacaccaga ggcagcagaa gcagagttgg 29580
ccttccaggt tcacggagat tcctttgctg tcaatacccc tacacccgat cagcggtgtc 29640
cggggctgct cgtcagcggc attgtcggtg tgctttcggg attagcagtc ataatcatct 29700
gcatgttcat ttttgcttgc tgctatagaa ggctttaccg acaaaaatca gacccactgc 29760
tgaacctcta tgtttaattt tttccagagc catgaaggca gttagcgctc tagtttttg 29820
ttctttgatt ggcattgttt ttagtgctgg gtttttgaaa aatcttacca tttatgaagg 29880
tgagaatgcc actctagtgg gcatcagtgg tcaaaatgtc agctggctaa aataccatct 29940
agatgggtgg aaagacattt gcgattggaa tgtcactgtg tatacatgta atggagttaa 30000
cctcaccatt actaatgcca cccaagatca gaatggtagg tttaagggcc agagtttcac 30060
tagaaataat gggtatgaat cccataacat gtttatctat gacgtcactg tcatcagaaa 30120
tgagactgcc accaccacac agatgcccac tacacacagt tctaccacta ctaccatgca 30180
aaccacacag acaaccacta catcaactca gcatatgacc accactacag cagcaaagcc 30240
aagtagtgca gcgcctcagc cccaggcttt ggctttgaaa gctgcacaac ctagtacaac 30300
tactaggacc aatgagcaga ctactgaatt tttgtccact gtcgagagcc acaccacagc 30360
tacctccagt gccttctcta gcaccgccaa tctctcctcg ctttcctcta caccaatcag 30420
tcccgctact actcccaccc cagctcttct ccccactccc ctgaagcaaa ctgaggacag 30480
cggcatgcaa tggcagatca ccctgctcat tgtgatcggg ttggtcatcc tggccgtgtt 30540
gctctactac atcttctgcc gccgcattcc caacgcgcac cgcaaaccgg cctacaagcc 30600
catcgttatc gggcagccgg agccgcttca ggtggaaggg ggtctaagga atcttctctt 30660
ctcttttaca gtatggtgat tgaactatga ttcctagaca attcttgatc actattctta 30720
tctgcctcct ccaagtctgt gccaccctcg ctctggtggc caacgccagt ccagactgta 30780
ttgggccctt cgcctcctac gtgctctttg ccttcatcac ctgcatctgc tgctgtagca 30840
tagtctgcct gcttatcacc ttcttccagt tcattgactg gatctttgtg cgcatcgcct 30900
acctgcgcca ccacccccag taccgcgacc agcgagtggc gcggctgctc aggctcctct 30960
gataagcatg cgggctctgc tacttctcgc gcttctgctg ttagtgctcc cccgccccgt 31020
cgacccccgg tcccccactc agtcccccga agaggtccgc aaatgcaaat tccaagaacc 31080
ctggaaattc ctcaaatgct accgccaaaa atcagacatg cttcccagct ggatcatgat 31140
cattgggatc gtgaacattc tggcctgcac cctcatctcc tttgtgattt acccctgctt 31200
tgactttggt tggaactcgc cagaggcgct ctatctcccg cctgaacctg acacaccacc 31260
acagcaacct caggcacacg cactaccacc accacagcct aggccacaat acatgcccat 31320
attagactat gaggccgagc cacagcgacc catgctcccc gctattagtt acttcaatct 31380
aaccggcgga gatgactgac ccactggcca acaacaacgt caacgacctt ctcctggaca 31440

tggacggccg cgcctcggag cagcgactcg cccaacttcg cattcgccag cagcaggaga 31500
gagccgtcaa ggagctgcag gacggcatag ccatccacca gtgcaagaaa ggcatcttct 31560
gcctggtgaa acaggccaag atctcctacg aggtcacccc gaccgaccat cgcctctcct 31620
acgagctcct gcagcagcgc cagaagttca cctgcctggt cggagtcaac cccatcgtca 31680
tcacccagca gtcgggcgat accaaggggt gcatccactg ctcctgcgac tcccccgact 31740
gcgtccacac tctgatcaag accctctgcg gcctccgcga cctcctcccc atgaactaat 31800
cacccccttа tccagtgaaa taaatatcat attgatgatg atttaaataa aaaataatca 31860
tttgatttga aataaagata caatcatatt gatgatttga gttttaaaaa ataaagaatc 31920
acttacttga aatctgatac caggtctctg tccatgtttt ctgccaacac cacctcactc 31980
ccctcttccc agctctggta ctgcagaccc cggcgggctg caaacttcct ccacacgctg 32040
aaggggatgt caaattcctc ctgtccctca atcttcattt tatcttctat cagatgtcca 32100
aaaagcgcgt ccgggtggat gatgacttcg accccgtcta cccctacgat gcagacaacg 32160
caccgaccgt gcccttcatc aacccccccc tcgtctcttc agatggattc caagagaagc 32220
ccctgggggt gctgtccctg cgactggctg accccgtcac caccaagaac ggggaaatca 32280
ccctcaagct gggagagggg gtggacctcg actcctcggg aaaactcatc tccaacacgg 32340
ccaccaaggc cgccgccccc ctcagttttt ccaacaacac catttccctt aacatggata 32400
cccctcttta taccaaagat ggaaaattat ccttacaagt ttctccaccg ttaaacatat 32460
taaaatcaac cattctgaac acattagctg tagcttatgg atcaggttta ggactgagtg 32520
gtggcactgc tcttgcagta cagttggcct ctccactcac ttttgatgaa aaaggaaata 32580
ttaaaattaa cctagccagt ggtccattaa cagttgatgc aagtcgactt agtatcaact 32640
gcaaaagagg ggtcactgtc actacctcag gagatgcaat tgaaagcaac ataagctggc 32700
ctaaaggtat aagatttgaa ggtaatggca tagctgcaaa cattggcaga ggattggaat 32760
ttggaaccac tagtacagag actgatgtca cagatgcata cccaattcaa gttaaattgg 32820
gtactggcct taccttttgac agtacaggcg ccattgttgc ttggaacaaa gaggatgata 32880
aacttacatt atggaccaca gccgacccct cgccaaattg caaaatatac tctgaaaaag 32940
atgccaaact cacactttgc ttgacaaagt gtggaagtca aattctgggt actgtgactg 33000
tattggcagt gaataatgga agtctcaacc caatcacaaa cacagtaagc actgcactcg 33060
tctccctcaa gtttgatgca agtggagttt tgctaagcag ctccacatta gacaaagaat 33120
attggaactt cagaaaggga gatgttacac ctgctgagcc ctatactaat gctataggtt 33180
ttatgcctaa cataaaggcc tatcctaaaa acacatctgc agcttcaaaa agccatattg 33240
tcagtcaagt ttatctcaat ggggatgagg ccaaaccact gatgctgatt attactttta 33300
atgaaactga ggatgcaact tgcacctaca gtatcacttt tcaatggaaa tgggatagta 33360
ctaagtacac aggtgaaaca cttgctacca gctccttcac cttctcctac atcgcccaag 33420
aatgaacact gtatcccacc ctgcatgcca acccttccca ccccactctg tctatggaaa 33480
aaactctgaa gcacaaaata aaataaagtt caagtgtttt attgattcaa cagttttaca 33540
ggattcgagc agttatttttt cctccaccct cccaggacat ggaatacacc accctctccc 33600
cccgcacagc cttgaacatc tgaatgccat tggtgatgga catgcttttg gtctccacgt 33660
tccacacagt ttcagagcga gccagtctcg ggtcggtcag ggagatgaaa ccctccgggc 33720
actcccgcat ctgcacctca cagctcaaca gctgaggatt gtcctcggtg gtcgggatca 33780
cggttatctg gaagaagcag aagagcggcg gtgggaatca tagtccgcga acgggatcgg 33840
ccggtggtgt cgcatcaggc cccgcagcag tcgctgccgc cgccgctccg tcaagctgct 33900
gctcaggggg tccgggtcca gggactccct cagcatgatg cccacggccc tcagcatcag 33960
tcgtctggtg cggcgggcgc agcagcgcat gcggatctcg ctcaggtcgc tgcagtacgt 34020
gcaacacagg accaccaggt tgttcaacag tccatagttc aacacgctcc agccgaaact 34080
catcgcggga aggatgctac ccacgtggcc gtcgtaccag atcctcaggt aaatcaagtg 34140
gcgctccctc cagaacacgc tgcccacgta catgatctcc ttgggcatgt ggcggttcac 34200
cacctcccgg taccacatca ccctctggtt gaacatgcag ccccggatga tcctgcggaa 34260
ccacagggcc agcaccgccc cgcccgccat gcagcgaaga gaccccgggt cccggcaatg 34320
gcaatggagg acccaccgct cgtacccgtg gatcatctgg gagctgaaca agtctatgtt 34380
ggcacagcac aggcatatgc tcatgcatct cttcagcact ctcagctcct cgggggtcaa 34440
aaccatatcc cagggcacgg ggaactcttg caggacagcg aaccccgcag aacagggcaa 34500
tcctcgcaca taacttacat tgtgcatgga cagggtatcg caatcaggca gcaccgggtg 34560
atcctccacc agagaagcgc gggtctcggt ctcctcacag cgtggtaagg gggccggccg 34620
atacgggtga tggcgggacg cggctgatcg tgttcgcgac cgtgtcatga tgcagttgct 34680
ttcggacatt ttcgtacttg ctgtagcaga acctggtccg ggcgctgcac accgatcgcc 34740

ggcggcggtc ccggcgcttg gaacgctcgg tgttgaaatt gtaaaacagc cactctctca 34800
gaccgtgcag cagatctagg gcctcaggag tgatgaagat cccatcatgc ctgatagctc 34860
tgatcacatc gaccaccgtg gaatgggcca gacccagcca gatgatgcaa ttttgttggg 34920
tttcggtgac ggcgggggag ggaagaacag gaagaaccat gattaacttt taatccaaac 34980
ggtctcggag cacttcaaaa tgaaggtcgc ggagatggca cctctcgccc ccgctgtgtt 35040
ggtggaaaat aacagccagg tcaaaggtga tacggttctc gagatgttcc acggtggctt 35100
ccagcaaagc ctccacgcgc acatccagaa acaagacaat agcgaaagcg ggagggttct 35160
ctaattcctc aatcatcatg ttacactcct gcaccatccc cagataattt tcatttttcc 35220
agccttgaat gattcgaact agttcctgag gtaaatccaa gccagccatg ataaagagct 35280
cgcgcagagc gccctccacc ggcattctta agcacaccct cataattcca agatattctg 35340
ctcctggttc acctgcagca gattgacaag cggaatatca aaatctctgc cgcgatccct 35400
aagctcctcc ctcagcaata actgtaagta ctctttcata tcctctccga aatttttagc 35460
cataggacca ccaggaataa gattagggca agccacagta cagataaacc gaagtcctcc 35520
ccagtgagca ttgccaaatg caagactgct ataagcatgc tggctagacc cggtgatatc 35580
ttccagataa ctggacagaa aatcacccag gcaatttta agaaaatcaa caaaagaaaa 35640
atcctccagg tgcacgttta gagcctcggg aacaacgatg aagtaaatgc aagcggtgcg 35700
ttccagcatg gttagttagc tgatctgtaa aaaacaaaaa ataaaacatt aaaccatgct 35760
agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggcca cggggtctcc 35820
ggcgcgaccc tcgtaaaaat tgtcgctatg attgaaaacc atcacagaga gacgttcccg 35880
gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg cgtccgcgag 35940
tgaaaaaaag cgcccgagga agcaataagg cactacaatg ctcagtctca agtccagcaa 36000
agcgatgcca tgcggatgaa gcacaaaatc ctcaggtgcg tacaaaatgt aattactccc 36060
ctcctgcaca ggcagcgaag cccccgatcc ctccagatac acatacaaag cctcagcgtc 36120
catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
acctgtccac ccgctctctg ctcaatatat agcccagatc tacactgacg taaaggccaa 36240
agtctaaaaa tacccgccaa ataatcacac acgcccagca cacgcccaga aaccggtgac 36300
acactcaaaa aaatacgcgc acttcctcaa acgcccaaac tgccgtcatt tccgggttcc 36360
cacgctacgt catcggaatt cgactttcaa attccgtcga ccgttaaaaa cgtcacccgc 36420
cccgccccta acggtcgccc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
cagctcattt gcatattaac gcgcaccaaa agtttgaggt atattattga tgatg 36535

<210> 5
<211> 32020
<212> DNA
<213> Chimpanzee Pan 5 (CV23) Genomic

<400> 5

catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg aggtatttga 60
atttggggat gcggggcggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggt gtcagctgat cgccagggta 480
ttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg 660
gtgacgaccc tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgtatg 720
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggctttagcc 1020
tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080

ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
acagtaagtg tgattaactt tagctgggga ggcagagggt gactgggtgc tgactggttt 1200
atttatgtat atgtttttta tgtgtaggtc ccgtctctga cgtagatgag acccccacta 1260
cagagtgcat ttcatcaccc ccagaaattg gcgaggaacc gcccgaagat attattcata 1320
gaccagttgc agtgagagtc accgggcgta gagcagctgt ggagagtttg gatgacttgc 1380
tacagggtgg ggatgaacct ttggacttgt gtacccggaa acgccccagg cactaagtgc 1440
cacacatgtg tgtttactta aggtgatgtc agtatttata gggtgtggag tgcaataaaa 1500
tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag 1560
caggtgcaga cctgtgtggt cagttcagag caggactcat ggagatctgg acagtcttgg 1620
aagactttca ccagactaga cagctgctag agaactcatc ggagggagtc tcttacctgt 1680
ggagattctg cttcggtggg cctctagcta agctagtcta tagggccaag caggattata 1740
aggatcaatt tgaggatatt ttgagagagt gtcctggtat ttttgactct ctcaacttgg 1800
gccatcagtc tcactttaac cagagtattc tgagagccct tgacttttct actcctggca 1860
gaactaccgc cgcggtagcc ttttttgcct ttatccttga caaatggagt caagaaaccc 1920
atttcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt 1980
gccagcgcct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga 2040
tcctgagtct ccagtcaccc caggaacacc aacgccgcca gcagccgcag caggagcagc 2100
agcaagagga ggaccgagaa gagaacctga gagccggtct ggaccctccg gtggcggagg 2160
aggaggagta gctgacttgt ttcccgagct gcgccgggtg ctgactaggt cttccagtgg 2220
acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac 2280
tgtcagtctg atgagtcgca ggcgcccaga atcggtgtgg tggcatgagg tgcagtcgca 2340
ggggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg 2400
ttggttggag cccgaggatg attgggaggt agccatcagg aattatgcca agctggctct 2460
gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat 2520
ttcagggaat ggggccgagg tggagatcag tacccaggag agggtggcct tcagatgctg 2580
catgatgaat atgtacccgg gggtggtggg catggaggga gtcaccttta tgaacgcgag 2640
gttcaggggt gatgggtata atggggtggt ctttatggcc aacaccaagc tgacagtgca 2700
cggatgctcc ttctttggct tcaataacat gtgcattgag gcctggggca gtgtttcagt 2760
gaggggatgc agtttttcag ccaactggat gggggtcgtg ggcagaacca agagcatggt 2820
gtcagtgaag aaatgcctgt tcgagaggtg ccacctgggg gtgatgagcg agggcgaagc 2880
caaagtcaaa cactgcgcct ctaccgagac gggctgcttt gtactgatca agggcaatgc 2940
caaagtcaag cataatatga tctgtggggc ctcggatgag cgcggctacc agatgctgac 3000
ctgcgccggt gggaacagcc atatgctagc caccgtgcat gtggcctcgc acccccgcaa 3060
gacatggccc gagttcgagc acaacgtcat gacccgctgc aatgtgcacc tggggtcccg 3120
ccgaggcatg ttcatgccct accagtgcaa catgcaattt gtgaaggtgc tgctggagcc 3180
cgatgccatg tccagagtga gcctgacggg ggtgtttgac atgaatgtgg agctgtggaa 3240
aattctgaga tatgatgaat ccaagaccag gtgccgggcc tgcgaatgcg gaggcaagca 3300
cgccaggctt cagcccgtgt gtgtggaggt gacggaggac ctgcgacccg atcatttggt 3360
gttgtcctgc aacgggacgg agttcggctc cagcggggaa gaatctgact agagtgagta 3420
gtgtttggga ctgggtggga gcctgcatga tgggcagaat gactaaaatc tgtgtttttc 3480
tgcgcagcag catgagcgga agcgcctcct ttgagggagg ggtattcagc ccttatctga 3540
cggggcgtct cccctcctgg gcgggagtgc gtcagaatgt gatgggatcc acggtggacg 3600
gccggcccgt gcagcccgcg aactcttcaa ccctgaccta cgcgaccctg agctcctcgt 3660
ccgtggacgc agctgccgcc gcagctgctg cttccgccgc cagcgccgtg cgcggaatgg 3720
ccctgggcgc cggctactac agctctctgg tggccaactc gagttccacc aataatcccg 3780
ccagcctgaa cgaggagaag ctgctgctgc tgatggccca gctcgaggcc ctgacccagc 3840
gcctgggcga gctgacccag caggtggctc agctgcaggc ggagacgcgg gccgcggttg 3900
ccacggtgaa aaccaaataa aaaatgaatc aataaataaa cggagacggt tgttgatttt 3960
aacacagagt cttgaatctt tatttgattt ttcgcgcgcg gtaggccctg gaccaccggt 4020
ctcgatcatt gagcacccgg tggatctttt ccaggacccg gtagaggtgg gcttggatgt 4080
tgaggtacat gggcatgagc ccgtcccggg ggtggaggta gctccattgc agggcctcgt 4140
gctcgggggt ggtgttgtaa atcacccagt catagcaggg gcgcagggcg tggtgctgca 4200
cgatgtcctt gaggaggaga ctgatggcca cgggcagccc cttggtgtag gtgttgacga 4260
acctgttgag ctgggaggga tgcatgcggg gggagatgag atgcatcttg gcctggatct 4320
tgagattggc gatgttcccg cccagatccc gccgggggtt catgttgtgc aggaccacca 4380

gcacggtgta tccggtgcac ttggggaatt tgtcatgcaa cttggaaggg aaggcgtgaa 4440
agaatttgga gacgcccttg tgaccgccca ggttttccat gcactcatcc atgatgatgg 4500
cgatgggccc gtgggcggcg gcttgggcaa agacgtttcg ggggtcggac acatcgtagt 4560
tgtggtcctg ggtgagctcg tcataggcca ttttaatgaa tttggggcgg agggtgcccg 4620
actgggggac gaaggtgccc tcgatcccgg gggcgtagtt gccctcgcag atctgcatct 4680
cccaggcctt gagctcggag gggggggatca tgtccacctg cggggcgatg aaaaaaacgg 4740
tttccggggc gggggagatg agctgggccg aaagcaggtt ccggagcagc tgggacttgc 4800
cgcagccggt ggggccgtag atgaccccga tgaccggctg caggtggtag ttgagggaga 4860
gacagctgcc gtcctcgcgg aggaggggggg ccacctcgtt catcatctcg cgcacatgca 4920
tgttctcgcg cacgagttcc gccaggaggc gctcgccccc aagcgagagg agctcttgca 4980
gcgaggcgaa gtttttcagc ggcttgagcc cgtcggccat gggcattttg gagagggtct 5040
gttgcaagag ttccagacgg tcccagagct cggtgatgtg ctctagggca tctcgatcca 5100
gcagacctcc tcgtttcgcg ggttggggcg actgcgggag tagggcacca ggcgatgggc 5160
gtccagcgag gccagggtcc ggtccttcca ggggcgcagg gtccgcgtca gcgtggtctc 5220
cgtcacggtg aaggggtgcg cgccgggctg ggcgcttgcg agggtgcgct tcaggctcat 5280
ccggctggtc gagaaccgct cccggtcggc gccctgcgcg tcggccaggt agcaattgag 5340
catgagttcg tagttgagcg cctcggccgc gtggcccttg gcgcggagct tacctttgga 5400
agtgtgtccg cagacgggac agaggaggga cttgagggcg tagagcttgg gggcgaggaa 5460
gacggactcg ggggcgtagg cgtccgcgcc gcagctggcg cagacggtct cgcactccac 5520
gagccaggtg aggtctggcc ggtcggggtc aaaaacgagg tttcctccgt gctttttgat 5580
gcgtttctta cctctggtct ccatgagctc gtgtccccgc tgggtgacaa agaggctgtc 5640
cgtgtccccg tagaccgact ttatgggccg gtcctcgagc ggggtgccgc ggtcctcgtc 5700
gtagaggaac cccgcccact ccgagacgaa ggcccgggtc caggccagca cgaaggaggc 5760
cacgtgggag gggtagcggt cgttgtccac cagcgggtcc accttctcca gggtatgcaa 5820
gcacatgtcc ccctcgtcca catccaggaa ggtgattggc ttgtaagtgt aggccacgtg 5880
accgggggtc ccggccgggg gggtataaaa gggggcgggc ccctgctcgt cctcactgtc 5940
ttccggatcg ctgtccagga gcgccagctg ttggggtagg tattccctct cgaaggcggg 6000
catgacctcg gcactcaggt tgtcagtttc tagaaacgag gaggatttga tattgacggt 6060
gccgttggag acgcctttca tgagcccctc gtccatctgg tcagaaaaga cgatcttttt 6120
gttgtcgagc ttggtggcga aggagccgta gagggcgttg gagagcagct tggcgatgga 6180
gcgcatggtc tggttctttt ccttgtcggc gcgctccttg gcggcgatgt tgagctgcac 6240
gtactcgcgc gccacgcact tccattcggg gaagacggtg gtgagcttgt cgggcacgat 6300
tctgacccgc cagccgcggt tgtgcagggt gatgaggtcc acgctggtgg ccacctcgcc 6360
gcgcaggggc tcgttggtcc agcagaggcg cccgcccttg cgcgagcaga agggggggcag 6420
cgggtccagc atgagctcgt cgggggggtc ggcgtccacg gtgaagatgc cgggcaggag 6480
ctcggggtcg aagtagctga tgcaggtgcc cagatcgtcc agcgccgctt gccagtcgcg 6540
cacggccagc gcgcgctcgt aggggctgag gggcgtgccc cagggcatgg ggtgcgtgag 6600
cgcggaggcg tacatgccgc agatgtcgta gacgtagagg ggctcctcga ggacgccgat 6660
gtaggtgggg tagcagcgcc ccccgcggat gctggcgcgc acgtagtcgt acagctcgtg 6720
cgagggcgcg aggagcccgg tgccgaggtt ggagcgctgc ggcttttcgg cgcggtagac 6780
gatctggcgg aagatggcgt gggagttgga ggagatggtg ggcctctgga agatgttgaa 6840
gtgggcgtgg ggcagtccga ccgagtccct gatgaagtgg gcgtaggagt cctgcagctt 6900
ggcgacgagc tcggcggtga cgaggacgtc cagggcgcag tagtcgaggg tctcttggat 6960
gatgtcgtac ttgagctggc ccttctgctt ccacagctcg cggttgagaa ggaactcttc 7020
gcggtccttc cagtactctt cgaggggaa cccgtcctga tcggcacggt aagagcccac 7080
catgtagaac tggttgacgg ccttgtaggc gcagcagccc ttctccacgg ggagggcgta 7140
agcttgcgcg gccttgcgca gggaggtgtg ggtgagggcg aaggtgtcgc gcaccatgac 7200
cttgaggaac tggtgcttga agtcgaggtc gtcgcagccg ccctgctccc agagctggaa 7260
gtccgtgcgc ttcttgtagg cggggttggg caaagcgaaa gtaacatcgt tgaagaggat 7320
cttgcccgcg cggggcatga agttgcgagt gatgcggaaa ggctggggca cctcggcccg 7380
gttgttgatg acctgggcgg cgaggacgat ctcgtcgaag ccgttgatgt tgtgcccgac 7440
gatgtagagt tccacgaatc gcgggcggcc cttgacgtgg ggcagcttct tgagctcgtc 7500
gtaggtgagc tcggcggggt cgctgaggcc gtgctgctcg agggcccagt cggcgaggtg 7560
ggggttggcg ccgaggaagg aagtccagag atccacggcc agggcggtct gcaagcggtc 7620
ccggtactga cggaactgct ggcccacggc catttttttcg ggggtgacgc agtagaaggt 7680

gcgggggtcg ccgtgccagc ggtcccactt gagctggagg gcgaggtcgt gggcgagctc 7740
gacgagcggc gggtccccgg agagtttcat gaccagcatg aaggggacga gctgcttgcc 7800
gaaggacccc atccaggtgt aggtttccac gtcgtaggtg aggaagagcc tttcggtgcg 7860
aggatgcgag ccgatgggga agaactggat ctcctgccac cagttggagg aatggctgtt 7920
gatgtgatgg aagtagaaat gccgacggcg cgccgagcac tcgtgcttgt gtttatacaa 7980
gcgtccgcag tgctcgcaac gctgcacggg atgcacgtgc tgcacgagct gtacctgggt 8040
tcctttgacg aggaatttca gtgggcagtg gagcgctggc ggctgcatct ggtgctgtac 8100
tacgtcctgg ccatcggcgt ggccatcgtc tgcctcgatg gtggtcatgc tgacgaggcc 8160
gcgcgggagg caggtccaga cctcggctcg gacgggtcgg agagcgagga cgagggcgcg 8220
caggccggag ctgtccaggg tcctgagacg ctgcggagtc aggtcagtgg gcagcggcgg 8280
cgcgcggttg acttgcagga gcttttccag ggcgcgcggg aggtccagat ggtacttgat 8340
ctccacggcg ccgttggtgg cgacgtccac ggcttgcagg gtcccgtgcc cctggggcgc 8400
caccaccgtg ccccgtttct tcttgggtgc tggcggcggc ggctccatgc ttagaagcgg 8460
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg 8520
ggcacgtcgg cgccgcgcgc gggcaggttc tggtactgcg cccggagaag actggcgtga 8580
gcgacgacgc gacggttgac gtcctggatc tgacgcctct gggtgaaggc cacgggaccc 8640
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc 8700
tgccgcagga tctcttgcac gtcgcccgag ttgtcctggt aggcgatctc ggtcatgaac 8760
tgctcgatct cctcctcctg aaggtctccg cgaccggcgc gctcgacggt ggccgcgagg 8820
tcgttggaga tgcggcccat gagctgcgag aaggcgttca tgccggcctc gttccagacg 8880
cggctgtaga ccacggctcc gtcggggtcg cgcgcgcgca tgaccacctg ggcgaggttg 8940
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc 9000
gtggtggcga tgtgctcggt gacgaagaag tacatgatcc agcggcggag cggcatctcg 9060
ctgacgtcgc ccagggcttc caagcgctcc atggcctcgt agaagtccac ggcgaagttg 9120
aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg 9180
gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc ttccatctcc 9240
tcctcctctt ccatctcctc cactaacatc tcttctactt cctcctcagg aggcggcggc 9300
gggggagggg ccctgcgtcg ccggcggcgc acgggcagac ggtcgatgaa gcgctcgatg 9360
gtctccccgc gccggcgacg catggtctcg gtgacggcgc gcccgtcctc gcggggccgc 9420
agcgtgaaga cgccgccgcg catctccagg tggccgccgg gggggtctcc gttgggcagg 9480
gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg 9540
agcgtctcga gatccacggg atccgaaaac cgctgaacga aggcttcgag ccagtcgcag 9600
tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg 9660
gcgatgctgc tggtgatgaa gttgaagtag gcggtcctga gacggcggat ggtggcgagg 9720
agcaccaggt ccttgggccc ggcttgctgg atgcgcagac ggtcggccat gccccaggcg 9780
tggtcctgac acctggcgag gtccttgtag tagtcctgca tgagccgctc cacgggcacc 9840
tcctcctcgc ccgcgcggcc gtgcatgcgc gtgagcccga acccgcgctg cggctggacg 9900
agcgccaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgagggtg 9960
gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cggtgttgat ggtgtaggag 10020
cagttggcca tgacggacca gttgacggtc tggtggccgg ggcgcacgag ctcgtggtac 10080
ttgaggcgcg agtaggcgcg cgtgtcgaag atgtagtcgt tgcaggtgcg cacgaggtac 10140
tggtatccga cgaggaagtg cggcggcggc tggcggtaga gcggccatcg ctcggtggcg 10200
ggggcgccgg gcgcgaggtc ctcgagcatg aggcggtggt agccgtagat gtacctggac 10260
atccaggtga tgccggcggc ggtggtggag gcgcgcggga actcgcggac gcggttccag 10320
atgttgcgca gcggcaggaa gtagttcatg gtggccgcgg tctggcccgt gaggcgcgcg 10380
cagtcgtgga tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactccgt 10440
ggcctggagg ctaagcgaac gggttgggct gcgcgtgtac cccggttcga gtccctgctc 10500
gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc 10560
taacgaaacc tccaggatac ggaggcgggt cgttttggcc attttcgtca ggccggaaat 10620
gaaactagta agcgcggaaa gcggccgtcc gcgatggctc gctgccgtag tctggagaaa 10680
gaatcgccag ggttgcgttg cggtgtgccc cggttcgagc ctcagcgctc ggcgccggcc 10740
ggattccgcg gctaacgtgg gcgtggctgc cccgtcgttt ccaagacccc ttagccagcc 10800
gacttctcca gttacggagc gagcccctct ttttcttgtg tttttgccag atgcatcccg 10860
tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac 10920
agccggcgct tctgcccccg ccccagcagc agcagccagc cactaccgcg gcggccgccg 10980

```
tgagcggagc cggcgttcag tatgacctgg ccttggaaga gggcgagggg ctggcgcggc 11040
tgggggcgtc gtcgccggag cggcacccgc gcgtgcagat gaaaagggac gctcgcgagg 11100
cctacgtgcc caagcagaac ctgttcagag acaggagcgg cgaggagccc gaggagatgc 11160
gcgcctcccg cttccacgcg gggcgggagc tgcggcgcgg cctggaccga aagcgggtgc 11220
tgagggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgcacg 11280
tggccgcggc caacctggtc acggcgtacg agcagaccgt gaaggaggag agcaacttcc 11340
aaaaatcctt caacaaccac gtgcgcacgc tgatcgcgcg cgaggaggtg accctgggcc 11400
tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa ccccacgagc aagccgctga 11460
cggcgcagct gtttctggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc 11520
tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga 11580
gcatcgtggt gcaggagcgc gggctgccgc tgtccgagaa gctggcggcc atcaacttct 11640
cggtgctgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgccca 11700
tagacaagga ggtgaagatc gacgggtttt acatgcgcat gaccctgaaa gtgctgaccc 11760
tgagcgacga tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgccagcc 11820
gccggcgcga gctgagcgac caggagctga tgcacagcct gcagcgggcc ctgaccgggg 11880
ccgggaccga gggggagagc tactttgaca tgggcgcgga cctgcgctgg cagcctagcc 11940
gccgggcctt ggaagctgcc ggcggttccc cctacgtgga ggaggtggac gatgaggagg 12000
aggagggcga gtacctggaa gactgatggc gcgaccgtat ttttgctaga tgcagcaaca 12060
gccaccgccg cctcctgatc ccgcgatgcg ggcggcgctg cagagccagc cgtccggcat 12120
taactcctcg gacgattgga cccaggccat gcaacgcatc atggcgctga cgacccgcaa 12180
tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt 12240
ggtgccctcg cgctcgaacc ccacgcacga gaaggtgctg gccatcgtga acgcgctggt 12300
ggagaacaag gccatccgcg gcgacgaggc cgggctggtg tacaacgcgc tgctggagcg 12360
cgtggcccgc tacaacagca ccaacgtgca gacgaacctg gaccgcatgg tgaccgacgt 12420
gcgcgaggcg gtgtcgcagc gcgagcggtt ccaccgcgag tcgaacctgg gctccatggt 12480
ggcgctgaac gccttcctga gcacgcagcc cgccaacgtg ccccggggcc aggaggacta 12540
caccaacttc atcagcgcgc tgcggctgat ggtggccgag gtgccccaga gcgaggtgta 12600
ccagtcgggg ccggactact tcttccagac cagtcgccag ggcttgcaga ccgtgaacct 12660
gagccaggct ttcaagaact tgcagggact gtggggcgtg caggccccgg tcggggaccg 12720
cgcgacggtg tcgagcctgc tgacgccgaa ctcgcgcctg ctgctgctgc tggtggcgcc 12780
cttcacggac agcggcagcg tgagccgcga ctcgtacctg ggctacctgc ttaacctgta 12840
ccgcgaggcc atcgggcagg cgcacgtgga cgagcagacc taccaggaga tcacccacgt 12900
gagccgcgcg ctgggccagg aggacccggg caacctggag gccaccctga acttcctgct 12960
gaccaaccgg tcgcagaaga tcccgcccca gtacgcgctg agcaccgagg aggagcgcat 13020
cctgcgctac gtgcagcaga gcgtggggct gttcctgatg caggaggggg ccacgcccag 13080
cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc 13140
gttcatcaat aagctgatgg actacttgca tcgggcggcc gccatgaact cggactactt 13200
taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta 13260
cgacatgccc gaccccaacg acgggttcct gtgggacgac gtggacagca gcgtgttctc 13320
gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggccgtcctc 13380
ggcgctgtcc ggtcgcgcgg gtgctgccgc ggcggtgccc gaggccgcca gcccttccc 13440
gagcctgccc ttttcgctga acagcgtgcg cagcagcgag ctgggtcggc tgacgcggcc 13500
gcgcctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccg agcgcgagaa 13560
gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta 13620
cgcgcacgag cacagggacg agccccgagc tagcagcagc accggcgcca cccgtagacg 13680
ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacag 13740
cagcgtgttg gacttgggtg ggagtggtgg tggtaacccg ttcgctcacc tgcgcccccg 13800
tatcgggcgc ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg 13860
cgaccagcgt gcgttcttct ctgttgtttg tagtagtatg atgaggcgcg tgtacccgga 13920
gggtcctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcggcgg cgatgcagcc 13980
cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag 14040
cattcgttac tcggagctgg cacccttgta cgataccacc cggttgtacc tggtggacaa 14100
caagtcggcg gacatcgcct cgctgaacta ccagaacgac cacagcaact tcctgaccac 14160
cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga 14220
cgagcgctcg cggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280
```

gaacgagttc atgtacagca acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc 14340
caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt gggtggagtt 14400
tgagctgccc gagggcaact tctcggtgac catgaccatc gatctgatga acaacgccat 14460
catcgacaac tacttggcgg tggggcggca gaacggggtg ctggagagcg acatcggcgt 14520
gaagttcgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tggtgatgcc 14580
gggcgtgtac accaacgagg ccttccaccc cgacatcgtc ctgctgcccg gctgcggcgt 14640
ggacttcacc gagagccgcc tcagcaacct gctgggcatc cgcaagcggc agcccttcca 14700
ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga 14760
cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcggcgaccg cagccgtggc 14820
caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcagcggc 14880
cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga 14940
gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc 15000
ctacaactac ggcgaccccg agaagggcgt gcgctcctgg acgctgctca ccacctcgga 15060
cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagacccggt 15120
caccttccgc tccacgcgtc aagttagcaa ctacccggtg gtgggcgccg agctcctgcc 15180
cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt 15240
cacctcgctc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgcccgcc 15300
cgcgcccacc attaccaccg tcagtgaaaa cgttcctgct ctcacagatc acgggaccct 15360
gccgctgcgc agcagtatcc ggggagtcca gcgcgtgacc gtcactgacg ccagacgccg 15420
cacctgcccc tacgtctaca aggccctggg cgtagtcgcg ccgcgcgtcc tctcgagccg 15480
caccttctaa aaaatgtcca ttctcatctc gcccagtaat aacaccggtt ggggcctgcg 15540
cgcgcccagc aagatgtacg gaggcgctcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600
gcgcgggcac ttccgcgctc cctggggcgc cctcaagggc cgcgtgcgct cgcgcaccac 15660
cgtcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgccgccgc 15720
gcccgtctcc accgtggacg ccgtcatcga cagcgtggtg gccgacgcgc gccggtacgc 15780
ccgcgccaag agccggcggc ggcgcatcgc ccggcggcac cggagcaccc ccgccatgcg 15840
cgcggcgcga gccttgctgc gcagggccag gcgcacggga cgcagggcca tgctcagggc 15900
ggccagacgc gcggcctccg gcagcagcag cgccggcagg acccgcagac gcgcggccac 15960
ggcggcggcg gcggccatcg ccagcatgtc ccgcccgcgg cgcggcaacg tgtactgggt 16020
gcgcgacgcc gccaccggtg tgcgcgtgcc cgtgcgcacc cgccccctc gcacttgaag 16080
atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaattc 16140
aaggaagaga tgctccaggt catcgcgcct gagatctacg gcccggcggc ggtgaaggag 16200
gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg 16260
gacggactgg tggagtttgt gcgcgagttc gccccccggc ggcgcgtgca gtggcgcggg 16320
cggaaagtga aaccggtgct gcgacccggc accacggtgg tcttcacgcc cggcgagcgt 16380
tccggctccg cctccaagcg ctcctacgac gaggtgtacg gggacgagga catcctcgag 16440
caggcggccg aacgtctggg cgagtttgct tacggcaagc gcagccgccc cgcgcccttg 16500
aaagaggagg cggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg 16560
gtgaccctgc agcaggtgct gcctggtgcg gcgccgcgcc ggggcttcaa gcgcgagggc 16620
ggcgaggatc tgtacccgac catgcagctg atggtgccca agcgccagaa gctggaggac 16680
gtgctggagc acatgaaggt ggaccccgag gtgcagcccg aggtcaaggt gcggcccatc 16740
aagcaggtgg ccccgggcct gggcgtgcag accgtggaca tcaagatccc cacggagccc 16800
atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat 16860
ccctggatgc cggcaccggc ttccaccacc cgccgaagac gcaagtacgg cgcggccagc 16920
ctgctgatgc ccaactacgc gctgcatcct tccatcatcc ccacgccggg ctaccgcggc 16980
acgcgcttct accgcggcta caccagcagc cgccgccgca agaccaccac ccgccgccgc 17040
cgtcgtcgca cccgccgcag cagcaccgcg acttccgccg ccgccctggt gcggagagtg 17100
taccgcagcg ggcgcgagcc tctgaccctg ccgcgcgcgc gctaccaccc gagcatcgcc 17160
atttaactac cgcctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtccccat 17220
tacgggctac cgaggaagaa agccgcgccg tagaaggctg acggggaacg ggctgcgtcg 17280
ccatcaccac cggcggcggc gcgccatcag caagcggttg gggggaggct tcctgcccgc 17340
gctgatgccc atcatcgccg cggcgatcgg ggcgatcccc ggcatagctt ccgtggcggt 17400
gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga 17460
cgctcctggt cctgtgatgt gtgtttttag atggaagaca tcaatttttc gtccctggca 17520
ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggcaa cagccaactg 17580

aacgggggcg ccttcaattg gagcagtctc tggagcgggc ttaagaattt cgggtccacg 17640
ctcaaaacct atggcaacaa ggcgtggaac agcagcacag ggcaggcgct gagggaaaag 17700
ctgaaagagc agaacttcca gcagaaggtg gtcgatggcc tggcctcggg catcaacggg 17760
gtggtggacc tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc 17820
ccgcccgcgg ggtccgtgga gatgccccag gtggaggagg agctgcctcc cctggacaag 17880
cgcggcgaca agcgaccgcg tcccgacgcg gaggagacgc tgctgacgca cacggacgag 17940
ccgcccccgt acgaggaggc ggtgaaactg ggtctgccca ccacgcggcc cgtggcgcct 18000
ctggccaccg gggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg 18060
cctccgcctg cttcccgccc ctccacagtg gctaagcccc tgccgccggt ggccgtcgcg 18120
tcgcgcgccc cccgaggccg cccccaggcg aactggcaga gcactctgaa cagcatcgtg 18180
ggtctgggag tgcagagtgt gaagcgccgc cgctgctatt aaaagacact gtagcgctta 18240
acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgcgt 18300
cgccgagttg caagatggcc accccatcga tgctgcccca gtgggcgtac atgcacatcg 18360
ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag 18420
acacctactt cagtctgggg aacaagttta ggaaccccac ggtggcgccc acgcacgatg 18480
tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccgtggac cgcgaggaca 18540
acacctactc gtacaaagtg cgctacacgc tggccgtggg cgacaaccgc gtgctggaca 18600
tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaaccct 18660
actccggcac cgcttacaac agcctggctc ccaagggagc gcccaacact tgccagtgga 18720
catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgcctg 18780
tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc 18840
ccatttatgc agataaaact tatcaaccag agcctcaagt gggtgatgct gaatggcatg 18900
acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga 18960
agccctgcta tggttctttt gccaagccta ccaataaaga aggaggtcag gcaaatgtga 19020
aaaccgaaac aggcggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa 19080
gtgcagctgc ggctggcctg gccccagaaa ttgttttgta tactgagaat gtggatctgg 19140
aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200
atttgggtca gcagtccatg cccaacagac ccaactacat tggctttaga gacaacttta 19260
tcgggctcat gtactacaac agcactggca acatgggcgt gctggctggt caggcctccc 19320
agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc 19380
ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct 19440
atgaccccga tgtgcgcatt attgaaaatc acggtgtgga ggatgaactc cctaactatt 19500
gcttcccctt ggatgctgtg ggtagaactg atacttacca gggaattaag gccaatggtg 19560
ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaatgaa ttgggcaagg 19620
gcaatccttt cgccatggag atcaacatcc aggccaacct gtggcggaac ttcctctacg 19680
cgaacgtggc gctgtacctg cccgactcct acaagtacac gccggccaac atcacgctgc 19740
cgaccaacac caacacctac gattacatga acggccgcgt ggtggcgccc tcgctggtgg 19800
acgcctacat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaacccct 19860
tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gctcctgggc aacgggcgct 19920
acgtgccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctcctgctcc 19980
tgcccgggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga 20040
gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc 20100
tctacgccac cttcttcccc atggcgcaca acaccgcctc cacgctcgag gccatgctgc 20160
gcaacgacac caacgaccag tccttcaacg actacctctc ggcggccaac atgctctacc 20220
ccatcccggc caacgccacc aacgtgccca tctccatccc ctcgcgcaac tgggccgcct 20280
tccgcggatg gtccttcacg cgcctcaaga cccgcgagac gccctcgctc ggctccgggt 20340
tcgaccccta cttcgtctac tcgggctcca tcccctacct cgacggcacc ttctacctca 20400
accacacctt caagaaggtc tccatcacct tcgactcctc cgtcagctgg cccggcaacg 20460
accgcctcct gacgcccaac gagttcgaaa tcaagcgcac cgtcgacgga gaggggtaca 20520
acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gcccactaca 20580
acatcggcta ccagggcttc tacgtgcccg agggctacaa ggaccgcatg tactccttct 20640
tccgcaactt ccagcccatg agccgccagg tcgtggacga ggtcaactac aaggactacc 20700
aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctcgcgccca 20760
ccatgcgcca gggacagccc taccccgcca actaccccta cccgctcatc ggcaagagcg 20820
ccgtcgccag cgtcacccag aaaaagttcc tctgcgaccg ggtcatgtgg cgcatccccc 20880

tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg 20940
ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccaccc 21000
ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg 21060
tcatcgaggc cgtctacctg cgcacgccct tctcggccgg caacgccacc acctaagccc 21120
cgctcttgct tcttgcaaga tgacggcctg tgcgggctcc ggcgagcagg agctcagggc 21180
catcctccgc gacctgggct gcgggccctg cttcctgggc accttcgaca agcgcttccc 21240
gggattcatg gccccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac 21300
cgggggcgag cactggctgg ccttcgcctg gaacccgcgc tcccacacct gctacctctt 21360
cgaccccttc gggttctcgg acgagcgcct caagcagatc taccagttcg agtacgaggg 21420
cctgctgcgc cgcagcgccc tggccaccga ggaccgctgc gtcaccctgg aaaagtccac 21480
ccagaccgtg cagggtccgc gctcggccgc ctgcgggctc ttctgctgca tgttcctgca 21540
cgccttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgctgac 21600
gggggtgccc aacggcatgc tccagtcgcc ccaggtggaa cccaccctgc gccgcaacca 21660
ggaggcgctc taccgcttcc tcaacgccca ctccgcctac tttcgctccc accgcgcgcg 21720
catcgagaag gccaccgcct tcgaccgcat gaatcaagac atgtaaaccg tgtgtgtatg 21780
tgaatgcttt attcataata aacagcacat gtttatgcca ccttttctga ggctctgact 21840
ttatttagaa atcgaagggg ttctgccggc tctcggcgtg ccccgcgggc agggatacgt 21900
tgcggaactg gtacttgggc agccacttga actcggggat cagcagcttc ggcacgggga 21960
ggtcggggaa cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg 22020
gcgcggagat cttgaaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca 22080
cggggttgca gcactggaac accatcaggg ccgggtgctt cacgctcgcc agcaccgtcg 22140
cgtcggtgat gccctccacg tccagatcct cggcgttggc catcccgaag ggggtcatct 22200
tgcaggtctg ccgccccatg ctgggcacgc agccgggctt gtggttgcaa tcgcagtgca 22260
gggggatcag catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga 22320
aagcctccag ctggcggaag gcctgctgcg ccttgccgcc ctcggtgaag aagaccccgc 22380
aggacttgct agagaactgg ttggtggcgc agccggcgtc gtgcacgcag cagcgcgcgt 22440
cgttgttggc cagctgcacc acgctgcgcc cccagcggtt ctgggtgatc ttggcccggt 22500
cggggttctc cttcagcgcg cgctgcccgt tctcgctcgc cacatccatc tcgatcgtgt 22560
gctccttctg gatcatcacg gtcccgtgca ggcatcgcag cttgccctcg gcctcggtgc 22620
acccgtgcag ccacagcgcg cagccggtgc actcccagtt cttgtgggcg atctgggagt 22680
gcgagtgcac gaagccctgc aggaagcggc ccatcatcgt ggtcagggtc ttgttgctgg 22740
tgaaggtcag cgggatgccg cggtgctcct cgttcacata caggtggcag atgcggcggt 22800
acacctcgcc ctgctcgggc atcagctgga aggcggactt caggtcgctc tccacgcggt 22860
accggtccat cagcagcgtc atgacttcca tgcccttctc ccaggccgag acgatcggca 22920
ggctcagggg gttcttcacc gccgttgtca tcttagtcgc cgccgctgag gtcaggggggt 22980
cgttctcgtc cagggtctca aacactcgct tgccgtcctt ctcggtgatg cgcacggggg 23040
gaaagctgaa gcccacggcc gccagctcct cctcggcctg cctttcgtcc tcgctgtcct 23100
ggctgatgtc ttgcaaaggc acatgcttgg tcttgcgggg tttcttttttg ggcggcagag 23160
gcggcggcgg agacgtgctg ggcgagcgcg agttctcgct caccacgact atttcttctt 23220
cttggccgtc gtccgagacc acgcggcggt aggcatgcct cttctggggc agaggcggag 23280
gcgacgggct ctcgcggttc ggcgggcggc tggcagagcc ccttccgcgt tcgggggtgc 23340
gctcctggcg gcgctgctct gactgacttc ctccgcggcc ggccattgtg ttctcctagg 23400
gagcaacaag catggagact cagccatcgt cgccaacatc gccatctgcc cccgccgccg 23460
ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gccgcccagc cccacctccg 23520
acgccgccgc ggccccagac atgcaagaga tggaggaatc catcgagatt gacctgggct 23580
acgtgacgcc cgcggagcac gaggaggagc tggcagcgcg cttttcagcc ccggaagaga 23640
accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcgagc 23700
atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcatctggcc cgccaatgca 23760
tcatcgtcaa ggacgcgctg ctcgaccgcg ccgaggtgcc cctcagcgtg gcggagctca 23820
gccgcgccta cgagcgcaac ctcttctcgc cgcgcgtgcc ccccaagcgc cagcccaacg 23880
gcacctgcga gcccaacccg cgcctcaact tctacccggt cttcgcggtg cccgaggccc 23940
tggccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc 24000
gcacccgcgc cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatcgcct 24060
ccttggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggccgcga 24120
acgctctgca aggaagcgga gaggagcatg agcaccacag cgccctggtg gagttggaag 24180

gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgacccac ttcgcctacc 24240
cggcgctcaa cctgcccccc aaggtcatga gcgccgtcat ggaccaggtg ctcatcaagc 24300
gcgcctcgcc cctctcggat gaggacatgc aggaccccga gagctcggac gagggcaagc 24360
ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gagtagcacc ccccagagct 24420
tggaagagcg gcgcaagctc atgatggccg tggtcctggt gaccgtggag ctggagtgtc 24480
tgcgccgctt cttcgccgac gcagagaccc tgcgcaaggt cgaggagaac ctgcactacc 24540
tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc 24600
tggtctccta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca 24660
ccctgcgcgg ggaggcccgc cgcgactaca tccgcgactg cgtctacctg tacctctgcc 24720
acacctggca gacgggcatg ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag 24780
agctctgcaa gctcctgcag aagaacctga aggccctgtg gaccgggttc gacgagcgca 24840
ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca 24900
acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct ttcatcctcg 24960
aacgctccgg gatcctgccc gccacctgct ccgcgctgcc ctcggacttc gtgccgctga 25020
ccttccgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact 25080
acctggccta ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140
actgccgctg caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga 25200
gcgagaccca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260
tgaaactcac cccggggctg tggacctcgg cctacttgcg caagttcgtg cccgaggact 25320
accatccctt cgagatcagg ttctacgagg accaatccca gccgcccaag gccgaactgt 25380
cggcctgcgt catcacccag ggggccatcc tggcccaatt gcaagccatc cagaaatccc 25440
gccaagaatt tctgctgaaa aagggccacg gggtctacct ggacccccag accggagagg 25500
agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag 25560
ctgccgccgc cggaggattt ggaggaagac tgggagagca gtcaggcaga ggaggaggag 25620
atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac 25680
gaggtggagg aggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctcggcggag 25740
aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcggcga ccgggcccac 25800
agtaggtggg acgagaccgg gcgcttcccg aaccccacca cccagaccgg taagaaggag 25860
cggcagggat acaagtcctg gcgggggcac aaaaacgcca tcgtctcctg cttgcaagcc 25920
tgcgggggca acatctcctt cacccgccgc tacctgctct tccaccgcgg ggtgaacttc 25980
ccccgcaaca tcttgcatta ctaccgtcac ctccacagcc cctactactg tttccaagaa 26040
gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc 26100
ggcggcaggt ggactgagga tcgcagcgaa cgagccggcg cagacccggg agctgaggaa 26160
ccggatcttt cccaccctct atgccatctt ccagcagagt cgggggcagg agcaggaact 26220
gaaagtcaag aaccgttctc tgcgctcgct cacccgcagt tgtctgtatc acaagagcga 26280
agaccaactt cagcgcactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgct 26340
cactcttaaa gagtagcccg cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc 26400
acctgcgccc ttcgcccgac catcatcatg agcaaagaga ttcccacgcc ttacatgtgg 26460
agctaccagc cccagatggg cctggccgcc ggcgccgccc aggactactc cacccgcatg 26520
aactggctca gcgccgggcc cgcgatgatc tcacgggtga atgacatccg cgcccgccga 26580
aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg 26640
cgtaattggc ccgccgccct ggtgtaccag gaaattcccc agcccacgac cgtactactt 26700
ccgcgagacg cccaggccga agtccagctg actaactcag gtgtccagct ggccggcggc 26760
gccgccctgt gtcgtcaccg ccccgctcag ggtataaagc ggctggtgat ccgaggcaga 26820
ggcacacagc tcaacgacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc 26880
ttccaactcg ccggatcggg gagatcttcc ttcacgcctc gtcaggccgt cctgactttg 26940
gagagttcgt cctcgcagcc ccgctcgggt ggcatcggca ctctccagtt cgtggaggag 27000
ttcactccct cggtctactt caaccccttc tccggctccc ccggccacta cccggacgag 27060
ttcatcccga acttcgacgc catcagcgag tcggtggacg gctacgattg aatgtcccat 27120
ggtggcgcag ctgacctagc tcggcttcga cacctggacc actgtcgcct ctcctacgag 27180
ctcctgcagc agcgccagaa gttcacctgc ctggtcggag tcaaccccat cgtcatcacc 27240
cagcagtcgg gcgataccaa ggggtgcatc cactgctcct gcgactcccc cgactgcgtc 27300
cacactctga tcaagaccct ctgcggcctc cgcgacctcc tccccatgaa ctaatcaccc 27360
ccttatccag tgaaataaag atcatattga tgatttgagt ttaataaaaa taaagaatca 27420
cttacttgaa atctgatacc aggtctctgt ccatgttttc tgccaacacc acttcactcc 27480

cctcttccca gctctggtac tgcaggcccc ggcgggctgc aaacttcctc cacaccctga 27540
aggggatgtc aaattcctcc tgtccctcaa tcttcatttt atcttctatc agatgtccaa 27600
aaagcgcgtc cgggtggatg atgacttcga ccccgtctac ccctacgatg cagacaacgc 27660
accgaccgtg cccttcatca acccccccct cgtctcttca gatggattcc aagagaagcc 27720
cctgggggtg ctgtccctgc gtctggccga tcccgtcacc accaagaacg gggaaatcac 27780
cctcaagctg ggagatgggg tggacctcga ctcctcggga aaactcatct ccaacacggc 27840
caccaaggcc gccgcccctc tcagtttttc caacaacacc atttcccttа acatggatac 27900
ccctttttac aacaacaatg gaaagttagg catgaaagtc actgctccac tgaagatact 27960
agacacagac ttgctaaaaa cacttgttgt agcttatgga caaggtttag gaacaaacac 28020
cactggtgcc cttgttgccc aactagcatc cccacttgct tttgatagca atagcaaaat 28080
tgcccttaat ttaggcaatg gaccattgaa agtggatgca aatagactga acatcaattg 28140
caatagagga ctctatgtta ctaccacaaa agatgcactg gaagccaata taagttgggc 28200
taatgctatg acatttatag gaaatgccat gggtgtcaat attgatacac aaaaaggctt 28260
gcaatttggc accactagta ccgtcgcaga tgttaaaaac gcttacccca tacaaatcaa 28320
acttggagct ggtctcacat ttgacagcac aggtgcaatt gttgcatgga acaaagatga 28380
tgacaagctt acactatgga ccacagccga cccctctcca aattgtcaca tatattctga 28440
aaaggatgct aagcttacac tttgcttgac aaagtgtggc agtcagattc tgggcactgt 28500
ttccctcata gctgttgata ctggcagttt aaatcccata acaggaacag taaccactgc 28560
tcttgtctca cttaaattcg atgcaaatgg agttttgcaa agcagctcaa cactagactc 28620
agactattgg aatttcagac agggagatgt tacacctgct gaagcctata ctaatgctat 28680
aggtttcatg cccaatctaa aagcataccc taaaaacaca agtggagctg caaaaagtca 28740
cattgttggg aaagtgtacc tacatgggga tacaggcaaa ccactggacc tcattattac 28800
tttcaatgaa acaagtgatg aatcttgcac ttactgtatt aactttcaat ggcagtgggg 28860
ggctgatcaa tataaaaatg aaacacttgc cgtcagttca ttcacctttt cctatattgc 28920
taaagaataa accccactct gtaccccatc tctgtctatg gaaaaaactc tgaaacacaa 28980
aataaaataa agttcaagtg ttttattgat tcaacagttt tacaggattc gagcagttat 29040
ttttcctcca ccctcccagg acatggaata caccaccctc tccccccgca cagccttgaa 29100
catctgaatg ccattggtga tggacatgct tttggtctcc acgttccaca cagtttcaga 29160
gcgagccagt ctcgggtcgg tcagggagat gaaaccctcc gggcactccc gcatctgcac 29220
ctcacagctc aacagctgag gattgtcctc ggtggtcggg atcacggtta tctggaagaa 29280
gcagaagagc ggcggtggga atcatagtcc gcgaacggga tcggccggtg gtgtcgcatc 29340
aggccccgca gcagtcgctg tcgccgccgc tccgtcaagc tgctgctcag ggggtccggg 29400
tccagggact ccctcagcat gatgcccacg gccctcagca tcagtcgtct ggtgcggcgg 29460
gcgcagcagc gcatgcggat ctcgctcagg tcgctgcagt acgtgcaaca caggaccacc 29520
aggttgttca acagtccata gttcaacacg ctccagccga aactcatcgc gggaaggatg 29580
ctacccacgt ggccgtcgta ccagatcctc aggtaaatca agtggcgccc cctccagaac 29640
acgctgccca tgtacatgat ctccttgggc atgtggcggt tcaccacctc ccggtaccac 29700
atcaccctct ggttgaacat gcagccccgg atgatcctgc ggaaccacag ggccagcacc 29760
gccccgcccg ccatgcagcg aagagacccc gggtcccgac aatggcaatg gaggacccac 29820
cgctcgtacc cgtggatcat ctgggagctg aacaagtcta tgttggcaca gcacaggcat 29880
atgctcatgc atctcttcag cactctcagc tcctcggggg tcaaaaccat atcccagggc 29940
acggggaact cttgcaggac agcgaacccc gcagaacagg gcaatcctcg cacataactt 30000
acattgtgca tggacagggt atcgcaatca ggcagcaccg ggtgatcctc caccagagaa 30060
gcgcgggtct cggtctcctc acagcgtggt aaggggggccg gccgatacgg gtgatggcgg 30120
gacgcggctg atcgtgttcg cgaccgtgtt atgatgcagt tgctttcgga cattttcgta 30180
cttgctgtag cagaacctgg tccgggcgct gcacaccgat cgccggcggc ggtcccggcg 30240
cttggaacgc tcggtgttga agttgtaaaa cagccactct ctcagaccgt gcagcagatc 30300
tagggcctca ggagtgatga agatcccatc atgcctgatg gctctaatca catcgaccac 30360
cgtggaatgg gccagaccca gccagatgat gcaattttgt tgggtttcgg tgacggcggg 30420
ggagggaaga acaggaagaa ccatgattaa cttttaatcc aaacggtctc ggagcacttc 30480
aaaatgaaga tcgcggagat ggcacctctc gcccccgctg tgttggtgga aaataacagc 30540
caggtcaaag gtgatacggt tctcgagatg ttccacggtg gcttccagca aagcctccac 30600
gcgcacatcc agaaacaaga caatagcgaa agcgggaggg ttctctaatt cctcaatcat 30660
catgttacac tcctgcacca tccccagata attttcattt ttccagcctt gaatgattcg 30720
aactagttcc tgaggtaaat ccaagccagc catgataaag agctcgcgca gagcgccctc 30780

caccggcatt cttaagcaca ccctcataat tccaagatat tctgctcctg gttcacctgc 30840
agcagattga caagcggaat atcaaaatct ctgccgcgat ccctaagctc ctccctcagc 30900
aataactgta agtactcttt catatcctct ccgaaatttt tagccatagg accaccagga 30960
ataagattag ggcaagccac agtacagata aaccgaagtc ctccccagtg agcattgcca 31020
aatgcaagac tgctataagc atgctggcta gacccggtga tatcttccag ataactggac 31080
agaaaatcgc ccaggcaatt tttaagaaaa tcaacaaaag aaaaatcctc caggtgcacg 31140
tttagagcct cgggaacaac gatggagtaa atgcaagcgg tgcgttccag catggttagt 31200
tagctgatct gtagaaaaaa acaaaaatga acattaaacc atgctagcct ggcgaacagg 31260
tgggtaaatc gttctctcca gcaccaggca ggccacgggg tctccggcac gaccctcgta 31320
aaaattgtcg ctatgattga aaaccatcac agagagacgt tcccggtggc cggcgtgaat 31380
gattcgacaa gatgaataca ccccccggaac attggcgtcc gcgagtgaaa aaaagcgccc 31440
aaggaagcaa taaggcacta caatgctcag tctcaagtcc agcaaagcga tgccatgcgg 31500
atgaagcaca aaattctcag gtgcgtacaa aatgtaatta ctcccctcct gcacaggcag 31560
caaagccccc gatccctcca ggtacacata caaagcctca gcgtccatag cttaccgagc 31620
agcagcacac aacaggcgca agagtcagag aaaggctgag ctctaacctg tccacccgct 31680
ctctgctcaa tatatagccc agatctacac tgacgtaaag gccaaagtct aaaaataccc 31740
gccaaataat cacacacgcc cagcacacgc ccagaaaccg gtgacacact caaaaaaata 31800
cgcgcacttc ctcaaacgcc caaactgccg tcatttccgg gttcccacgc tacgtcatca 31860
aaattcgact ttcaaattcc gtcgaccgtt aaaaacgtcg cccgccccgc ccctaacggt 31920
cgccgctccc gcagccaatc accgccccgc atccccaaat tcaaatacct catttgcata 31980
ttaacgcgca ccaaaagttt gaggtatatt attgatgatg 32020

<210> 6
<211> 1737
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 20 Fiber

<400> 6

atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtaccccta tgacacggaa 60
agcggccctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gaaagtcccc ccggggtcct gtctctgaac ctggccgagc ccctggtcac ttcccacggc 180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgctggcaa cctcacctct 240
caagatatca ccaccgctag ccctcccctc aaaaaaacca agaccaacct cagcctagaa 300
acctcatccc ccctaactgt gagcacctca ggcgccctca ccgtagcagc cgccgctccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg ccccctgac agtacaggat 420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg 480
caaacatcgg ccccgctgac ggccgctgac agcagcaccc tcacagtcag tgccacacca 540
ccccttagca caagcaatgg cagcttgggt attgacatgc aagcccccat ttacaccacc 600
aatggaaaac taggacttaa ctttggcgct cccctgcatg tggtagacag cctaaatgca 660
ctgactgtag ttactggcca aggtcttacg ataaacggaa cagccctaca aactagagtc 720
tcaggtgccc tcaactatga cacatcagga aacctagaat tgagagctgc agggggtatg 780
cgagttgatg caaatggtca acttatcctt gatgtagctt acccatttga tgcacaaaac 840
aatctcagcc ttaggcttgg acagggaccc ctgtttgtta actctgccca caacttggat 900
gttaactaca acagaggcct ctacctgttc acatctggaa ataccaaaaa gctagaagtt 960
aatatcaaaa cagccaaggg tctcatttat gatgacactg ctatagcaat caatgcgggt 1020
gatgggctac agtttgactc aggctcagat acaaatccat taaaaactaa acttggatta 1080
ggactggatt atgactccag cagagccata attgctaaac tgggaactgg cctaagcttt 1140
gacaacacag gtgccatcac agtaggcaac aaaaatgatg acaagctcac cttgtggacc 1200
acaccagacc catctcctaa ctgtagaatc tattcagaga aagatgctaa attcacactt 1260
gttttgacta aatgcggcag tcaggtgttg gccagcgttt ctgttttatc tgtaaaaggt 1320
agccttgcgc ccatcagtgg cacagtaact agtgctcaga ttgtcctcag atttgatgaa 1380
aatggagttc tactaagcaa ttcttccctt gaccctcaat actggaacta cagaaaaggt 1440
gaccttacag agggcactgc atataccaac gcagtgggat ttatgcccaa cctcacagca 1500
tacccaaaaa cacagagcca aactgctaaa agcaacattg taagtcaggt ttacttgaat 1560
ggggacaaat ccaaacccat gaccctcacc attaccctca atggaactaa tgaaacagga 1620

gatgccacag taagcactta ctccatgtca ttctcatgga actggaatgg aagtaattac 1680
attaatgaaa cgttccaaac caactccttc accttctcct acatcgccca agaataa 1737

<210> 7
<211> 1278
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 4 Fiber

<400> 7

atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgtt gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgatt cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggatcacc ccttttacac taaagatgga aaattatcct tacaagtttc tccaccatta 360
aatatactga gaacaagcat tctaaacaca ctagctttag gttttggatc aggtttagga 420
ctccgtggct ctgccttggc agtacagtta gtctctccac ttacatttga tactgatgga 480
aacataaagc ttaccttaga cagaggtttg catgttacaa caggagatgc aattgaaagc 540
aacataagct gggctaaagg tttaaaattt gaagatggag ccatagcaac caacattgga 600
aatgggttag agtttggaag cagtagtaca gaaacaggtg ttgatgatgc ttacccaatc 660
caagttaaac ttggatctgg ccttagcttt gacagtacag gagccataat ggctggtaac 720
aaagaagacg ataaactcac tttgtggaca acacctgatc catcaccaaa ctgtcaaata 780
ctcgcagaaa atgatgcaaa actaacactt tgcttgacta aatgtggtag tcaaatactg 840
gccactgtgt cagtcttagt tgtaggaagt ggaaacctaa accccattac tggcaccgta 900
agcagtgctc aggtgtttct acgttttgat gcaaacggtg ttcttttaac agaacattct 960
acactaaaaa aatactgggg gtataggcag ggagatagca tagatggcac tccatatacc 1020
aatgctgtag gattcatgcc caatttaaaa gcttatccaa agtcacaaag ttctactact 1080
aaaaataata tagtagggca agtatacatg aatggagatg tttcaaaacc tatgcttctc 1140
actataaccc tcaatggtac tgatgacagc aacagtacat attcaatgtc attttcatac 1200
acctggacta atggaagcta tgttggagca acatttgggg ctaactctta taccttctca 1260
tacatcgccc aagaatga 1278

<210> 8
<211> 1335
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 5 Fiber

<400> 8

atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgct gtccctgcgt ctggccgatc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agatggggtg gacctcgacg actcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggataccc ctctttacaa caacaatgga aagctaggta tgaaggtaac cgcaccatta 360
aagatattag acacagatct actaaaaaca cttgttgttg cttatgggca gggattagga 420
acaaacacca atggtgctct tgttgcccaa ctagcatacc cacttgtttt taataccgct 480
agcaaaattg cccttaattt aggcaatgga ccattaaaag tggatgcaaa tagactgaac 540
attaattgca aaagaggtat ctatgtcact accacaaaag atgcactgga gattaatatc 600
agttgggcaa atgctatgac atttatagga aatgccattg gtgtcaatat tgacacaaaa 660
aaaggcctac agttcggcac ttcaagcact gaaacagatg ttaaaaatgc tttttcactc 720
caagtaaaac ttggagctgg tcttacattt gacagcacag gtgccattgt tgcttggaac 780
aaagaagatg acaaacttac actgtggacc acagccgatc catctccaaa ctgtcacata 840
tattctgcaa aggatgctaa gcttacactc tgcttgacaa agtgtggtag tcaaatccta 900
ggcactgtct ccctattagc agtcagtggc agcttggctc ctatcacagg ggctgttaga 960
actgcacttg tatcactcaa attcaatgct aatggagccc ttttggacaa atcaactctg 1020

aacaaagaat actggaacta cagacaagga gatctaattc caggtacacc atatacacat 1080
gctgtgggtt tcatgcctaa caaaaaagcc taccctaaaa acacaactgc agcttccaag 1140
agccacattg tgggtgatgt gtatttagat ggagatgcag ataaaccttt atctcttatc 1200
atcactttca atgaaactga tgatgaaacc tgtgattact gcatcaactt tcaatggaaa 1260
tggggagctg atcaatataa ggataagaca ctcgcaacca gttcattcac cttctcatac 1320
atcgcccaag aataa 1335

<210> 9
<211> 1338
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 7 Fiber

<400> 9

atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgct gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgact cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggataccc ctttttacaa caataatgga aagttaggca tgaaagtcac tgctccactg 360
aagatactcg acacagactt gctaaaaaca cttgttgtag cttatggaca aggtttagga 420
acaaacacca ctggtgccct tgttgcccaa ctagcagccc cacttgcttt tgatagcaat 480
agcaaaattg cccttaattt aggcaatgga ccattgaaag tggatgcaaa tagactgaac 540
atcaattgca atagaggact ctatgttact accacaaaag atgcactgga aaccaacata 600
agttgggcta atgctatgac atttatagga aatgccatgg gtgtcaatat tgatacacaa 660
aaaggcttgc aatttggcac cactagtacc gtcgcagatg ttaaaaacgc ttaccccata 720
caagtcaaac tgggagctgg tctcacattt gacagcacag gtgcaattgt cgcttggaac 780
aaagaagatg acaaacttac actgtggacc acagccgatc catctccaaa ctgtcacata 840
tattctgaca aggatgctaa gcttacactc tgcttgacaa agtgtggcag tcagatactg 900
ggcactgttt ctctcatagc tgttgatact ggtagcttaa atccaataac aggacaagta 960
accactgctc ttgtttcact taaattcgat gccaatggag ttttgcaaac cagttcaaca 1020
ttggacaaag aatattggaa ttttagaaaa ggagatgtga cacctgctga gccatatact 1080
aatgctatag gtttcatgcc caatctaaag gcatacccta aaaacacaag tggagctgca 1140
aaaagtcaca ttgttgggaa agtgtaccta catggggata cagacaaacc actggacctg 1200
attattactt tcaatgaaac aagtgatgaa tcttgcactt actgtattaa ctttcaatgg 1260
aaatgggata gtactaagta cacaggtgaa acacttgcta caagctcctt caccttctcc 1320
tacattgccc aagaatga 1338

<210> 10
<211> 1278
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 9 Fiber

<400> 10

atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgtt gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgact cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc tgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggatcacc ccttttacac taaagatgga aaattagcct tacaagtttc tccaccatta 360
aatatactga gaacaagcat tctaaacaca ctagctttag gttttggatc aggtttagga 420
ctccgtggct ctgccttggc agtacagtta gtctctccac ttacatttga tactgatgga 480
aacataaagc ttaccttaga cagaggtttg catgttacaa caggagatgc aattgaaagc 540
aacataagct gggctaaagg tttaaaattt gaagatggag ccatagcaac caacattgga 600
aatgggttag agtttggaag cagtagtaca gaaacaggtg tcgatgatgc ttacccaatc 660
caagttaaac ttggatctgg ccttagcttt gacagtacag gagccataat ggctggtaac 720

aaagaagacg ataaactcac tttgtggaca acacctgatc catcaccaaa ctgtcaaata 780
ctcgcagaaa atgatgcaaa actaacactt tgcttgacta aatgtggtag tcaaatactg 840
gccactgtgt cagtcttagt tgtaggaagt ggagacctaa accccattac tggcaccgta 900
agcagtgctc aggtgtttct acgttttgat gcaaacggtg ttcttttaac agaacattct 960
acactaaaaa aatactgggg gtataggcag ggagatagca tagatggcac tccatatgcc 1020
aatgctgtag gattcatgcc caatttaaaa gcttatccaa agtcacaaag ttctactact 1080
aaaaataata tagtagggca agtatacatg aatggagatg tttcaaaacc tatgcttctc 1140
actataaccc tcaatggtac tgatgacagc aacagtacat attcaatgtc attttcatac 1200
acctggacta atggaagcta tgttggagca acatttggag ctaactctta taccttctcc 1260
tacatcgccc aagaatga 1278

<210> 11
<211> 1278
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 10 Fiber

<400> 11

atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgct gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agaggggtg gacctcgact cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggatcacc ccttttacac taaagatgga aaattatcct tacaagtttc tccaccatta 360
aatatactga gaacaagcat tctaaacaca ctagctttag gttttggatc aggtttagga 420
ctccgtggct ctgccttggc agtacagtta gtctctccac ttacatttga tactgatgga 480
aacataaagc ttaccttaga cagaggtttg catgttacaa caggagatgc aattgaaagc 540
aacataagct gggctaaagg tttaaaattt gaagatggag ccatagcaac caacattgga 600
aatgggttag agtttggaag cagtagtaca gaaacaggtg ttgatgatgc ttacccaatc 660
caagttaaac ttggatctgg ccttagcttt gacagtacag gagccataat ggctggtaac 720
aaagaagacg ataaactcac tttgtggaca acacctgatc catcgccaaa ctgtcaaata 780
ctcgcagaaa atgatgcaaa actaacactt tgcttgacta aatgtggtag tcaaatactg 840
gccactgtgt cagtcttagt tgtaggaagt ggaaacctaa accccattac tggcaccgta 900
agcagtgctc aggtgtttct acgttttgat gcaaacggtg ttcttttaac agaacattct 960
acactaaaaa aatactgggg gtataggcag ggagatagca tagatggcac tccatatacc 1020
aatgctgtag gattcatgcc caatttaaaa gcttatccaa agtcacaaag ttctactact 1080
aaaaataata tagtagggca agtatacatg aatggagatg tttcaaaacc tatgcttctc 1140
actataaccc tcaatggtac tgatgacagc aacagtacat attcaatgtc attttcatac 1200
acctggacta atggaagcta tgttggagca acatttgggg ctaactctta taccttctca 1260
tacatcgccc aagaatga 1278

<210> 12
<211> 1737
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 11 Fiber

<400> 12

atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtaccccta tgacacggaa 60
aacggtcctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gagagccccc ccggggtcct gtctctgaac ctggccgagc ccctggtcac ttcccacggc 180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgccggcaa cctcacctct 240
caagatgtca ccaccactac ccctcccctg aaaaaaacca agaccaacct cagcctagaa 300
acctcagccc ccctgactgt gagcacctca ggcgccctca ccctagcagc cgccgttccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg ccccctgac agtccaagat 420
gcaaaactca ccctggccac caagggcccc ctgaccgtgt ctgaaggcaa actagccttg 480
cagacctcgg ccccgctgac ggccgctgac agcagcaccc tcacaatcag cgccacaccg 540

```
ccccttagca caagcaatgg cagcttgggt attgacatgc aagcccccat ttacactact 600
aacggaaaac tgggacttaa ctttggtgct cccctgcatg tggtagacag cctaaatgca 660
ctgactgtag tgactggcca aggtcttacg ataaacggta cagccctaca aactagagtc 720
tcaggtgccc tcaactatga ctcatcagga aacctagaat tgagagctgc agggggtatg 780
cgagttgatg caaatggcaa acttatcctt gacgtagctt acccatttga tgctcaaaac 840
aacctcagcc ttagacttgg acagggaccc ctgtttgtta actctgccca caacttggat 900
gttaactaca acagaggcct ctacctgttc acatctggaa ataccaaaaa gctagaagtt 960
aatatcaaaa cagccaaagg cctcatttat gatgacactg ctatagcaat caatccaggc 1020
gatgggctag agtttggctc aggctcagat acaaatccat taaaaactaa acttggattg 1080
ggactagagt atgactccag cagagccata attgctaagc tgggaaccgg cctaagcttt 1140
gacaacacag gtgccatcac agtgggcaac aaaaatgatg acaagcttac cttgtggacc 1200
acaccagacc cctctcccaa ctgtagaatt tattcagaaa aagatgctaa atttacacta 1260
gttttaacta aatgcggcag tcaggtgttg gccagcgttt ctgttttatc tgtaaaaggc 1320
agccttgcgc ccatcagtgg cacagtaact agcgctcaga ttattctcag atttgatgaa 1380
aatggagttc tactaagcaa ttcttctctt gacccccaat actggaacta cagaaaaggt 1440
gaccttacag agggcactgc atataccaac gcagtgggat ttatgcccaa cctcacagca 1500
tacccaaaaa cacagagtca aactgctaaa agcaacattg taagccaggt ttacttgaat 1560
ggggacaaat ccaaacccat gatcctcacc attaccctca atggaactaa tgaaacaggg 1620
gatgctacag ttagcactta ctccatgtca ttctcatgga attggaatgg aagtaattac 1680
attaatgaaa cgttccaaac caactctttc accttctcct acatcgccca agaataa    1737
```

<210> 13
<211> 1632
<212> DNA
<213> Chimpanzee Adnovirus- ChAd 16 Fiber

<400> 13

```
atgaagcgca ccaaaacgtc tgacaagagc ttcaaccccg tgtaccccta tgacacggaa 60
aacggtcctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gagagccccc ccggggtcct gtctctgaac ctggccgagc ccctggtcac ttcccacggc 180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgccggcaa cctcacctct 240
caagatgtca ccaccactac ccctcccctg aaaaaaacca agaccaacct cagcctagaa 300
acctcagccc ccctgactgt gagcacctca ggcgccctca ccctagcagc cgccgccccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg ccccccctgac agtccaagat 420
gcaaaactca ccctggccac caagggcccc ctgaccgtgt ctgaaggcaa actggccttg 480
cagacctcgg ccccgctgac ggccgctgac agcagcaccc tcaccgttag cgccacacca 540
cccatcagtg taagcagtgg aagtttgggc ttagacatgg aagacccaat gtatactcat 600
gatggaaaac tgggaataag aattgggggc ccactgagag tagtagacag cctgcacaca 660
ctgactgtag ttaccggaaa tggaatagct gtagataaca atgccctcca aactagagtt 720
acgggcgccc tgggttatga cacatcagga aacctacaac tgagagccgc ggggggtatg 780
cgaattgatg caaatggcca acttatcctt gatgtggcat acccatttga tgctcaaaac 840
aatctcagcc ttagacttgg tcagggaccc ctgtatgtaa acacagacca caacctagat 900
ttgaattgca acagaggtct gaccacaact accaccaaca acacaaaaaa acttgaaact 960
aaaattggct caggcttaga ctatgatacc aatggtgctg ttattattaa acttggcact 1020
ggtgtcagct ttgacagcac aggtgcccta agtgtgggaa acactggcga tgataaactg 1080
actctgtgga caaccccaga cccatctcca aattgcagaa ttcactcaga caaagactgc 1140
aagtttactc tagtcctaac taagtgtgga agtcaaatcc tggcttctgt cgccgcccta 1200
gcggtgtcag gaaatctggc ttcaataaca ggcaccgttt ccagcgttac catctttctc 1260
agatttgatc agaatggagt gcttatggaa aactcctcgc tagacaagca gtactggaac 1320
ttcagaaatg gtaattcaac caatgccacc ccctacacca atgcagttgg gtttatgcca 1380
aacctcgcag catacccccaa gacacagagc cagactgcaa aaaacaacat tgtaagtcag 1440
gtttacttga atggggacaa atccaaaccc atgaccctta ccattaccct caatggaact 1500
aatgaatcca gtgaaactag ccaggtgagt cactactcca tgtcatttac gtgggcttgg 1560
gagagtgggc aatatgccac cgaaaccttt gccaccaatt cctttacctt ctcttacatt 1620
gctgaacaat aa                                                    1632
```

<210> 14

<211> 1632
<212> DNA
<213> Chimpanzee Adenovirus-ChAd 17 Fiber

<400> 14

atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtaccccta tgacacggaa 60
agcggccctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gaaagccccc ccggggtcct gtctctgaac ctggccgagc ccctggtcac ttcccacggc 180
atgcttgccc tgaaaatggg aagtggcctc tccctggacg acgctggcaa ccttacctct 240
caagatatta cctccactac ccctcccctc aaaaaaacca agaccaacct cagcctagaa 300
acctcatccc ccctaactgt aagcacctca ggcgccctca ccgtagcagc cgccgctccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg ccccccctggc agtacaggat 420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg 480
caaacatcgg ccccgctgac ggccgctgac agcagcaccc tcaccgttag ctccactcca 540
ccaattagtg taagcagtgg aagtttgggc ttggacatgg aagaccccat gtatactcac 600
gatggaaaac tgggaataag aattgggggt ccactaagag tagtagacag cttgcacaca 660
ctcactgtag ttaccggaaa tggactaact gtagataaca atgccctcca aactagagtt 720
acgggcgccc taggttatga cacatcagga aatctacaac tgagagccgc agggggtatg 780
cgaattgatg caaatggcca acttatcctt gatgtggcat acccatttga tgctcaaaac 840
aatctcagcc ttagacttgg tcagggaccc ctgtatgtaa atacagacca caacctggat 900
ttaaattgca acagaggtct aaccacaact accaccaaca acacaaaaaa acttgagact 960
aaaattagct caggcttaga ctatgacacc aatggtgctg tcattattaa acttggcact 1020
ggtctaagct tcgacaacac aggcgcccta actgtgggaa acactggtga tgataaactg 1080
actctgtgga cgaccccaga cccatctcca aattgcagaa ttcactcaga caaagactgc 1140
aagtttactc tcgtcctaac taagtgtgga agccaaatcc tggcctctgt cgccgcccta 1200
gcggtatcag gaaatctggc ttcgataaca ggcaccgttg ccagcgttac catctttctt 1260
agatttgatc agaatggagt gcttatggaa aactcctcac tagacaagca gtactggaac 1320
ttcagaaatg gcaattcaac taatgctgcc ccctacacca acgcagttgg gttcatgcca 1380
aacctcgcag cgtaccccaa aacgcaaagc cagactgcta aaaacaacat tgtaagtcag 1440
gtttacttga atggagacaa atccaaaccc atgaccctta ccatcaccct caatggaact 1500
aatgaatcca gtgaaactag tcaggtgagt cactactcca tgtcatttac atgggcttgg 1560
gaaagcgggc aatatgccac tgaaaccttt gccaccaact ccttcacctt ttcttacatt 1620
gctgaacaat aa 1632

<210> 15
<211> 1632
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 19 Fiber

<400> 15

atgaagcgca ccaaaacgtc tgacaagagc ttcaaccccg tgtaccccta tgacacggaa 60
aacggtcctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gagagccccc ccggggtcct gtctctgaac ctggccgagc ccctggtcac ttcccacggc 180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgccggcaa cctcacctct 240
caagatgtca ccaccactac ccctcccctg aaaaaaacca agaccaacct cagcctagaa 300
acctcagccc ccctgactgt gagcacctca ggcgccctca ccctagcagc cgccgccccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg ccccccctgac agtccaagat 420
gcaaaactca ccctggccac caagggcccc ctgaccgtgt ctgaaggcaa actggccttg 480
cagacctcgg ccccgctgac ggccgctgac agcagcaccc tcaccgttag cgccacacca 540
cccatcagtg taagcagtgg aagtttgggc ttagacatgg aagacccaat gtatactcat 600
gatggaaaac tgggaataag aattgggggc ccactgagag tagtagacag cctgcacaca 660
ctgactgtag ttaccggaaa tggaatagct gtagataaca atgccctcca aactagagtt 720
acgggcgccc tgggttatga cacatcagga aacctacaac tgagagccgc ggggggtatg 780

```
cgaattgatg caaatggcca acttatcctt gatgtggcat acccatttga tgctcaaaac 840
aatctcagcc ttagacttgg tcagggaccc ctgtatgtaa acacagacca caacctagat 900
ttgaattgca acagaggtct gaccacaact accaccaaca acacaaaaaa acttgaaact 960
aaaattggct caggcttaga ctatgatacc aatggtgctg ttattattaa acttggcact 1020
ggtgtcagct ttgacagcac aggtgcccta agtgtgggaa acactggcga tgataaactg 1080
actctgtgga caaccccaga cccatctcca aattgcagaa ttcactcaga caaagactgc 1140
aagtttactc tagtcctaac taagtgtgga agtcaaatcc tggcttctgt cgccgcccta 1200
gcggtgtcag gaaatctggc ttcaataaca ggcaccgttt ccagcgttac catctttctc 1260
agatttgatc agaatggagt gcttatggaa aactcctcgc tagacaagca gtactggaac 1320
ttcagaaatg gtaattcaac caatgccacc ccctacacca atgcagttgg gtttatgcca 1380
aacctcgcag cataccccaa gacacagagc cagactgcaa aaaacaacat tgtaagtcag 1440
gtttacttga atggggacaa atccaaaccc atgaccctta ccattaccct caatggaact 1500
aatgaatcca gtgaaactag ccaggtgagt cactactcca tgtcatttac gtgggcttgg 1560
gagagtgggc aatatgccac cgaaaccttt gccaccaatt cctttacctt ctcttacatt 1620
gctgaacaat aa                                                     1632
```

<210> 16
<211> 2865
<212> DNA
<213> Chimp0anzee Adenovirus- ChAd 20 Hexon

<400> 16

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tctcagcgcc tgacgctgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggggtgct ggaccggggt cccactttca agccctactc tggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aacccatgcg agtgggatga ggctgctact 420
gcccttgaca ttgatttgaa cgcagaagac gatgaagaaa gcgacgaagc tcaaggggaa 480
gcagatcagc agaaaactca tgtatttggc caggcgccct actccggaca gaacattaca 540
aaagaaggca tacagatagg catagatgct gccagtcaag cccagacacc tgtatatgcc 600
gataaaacat tccaaccaga acctcaagtt ggagaatcac agtggaatga gacagagatt 660
agttatggag cgggacgggt gcttaaaaaa accactctca tgaaaccttg ctatgggtcg 720
tatgcaaggc ctactaatga gaacggaggt cagggcatcc tcttggaaca agatggaaag 780
aaagaaagtc aagtggaaat gcaatttttc tctactactc aggcagccgc gggtaattca 840
gataatccta ccccaaaggt tgttttgtac agcgaggatg ttaacctgga aacaccagat 900
acacacattt catacatgcc caccaacaac gagacaaatt caagagagct tttgggacaa 960
caggccatgc ccaacaggcc taattacatt ggcttcagag acaactttat cggtctcatg 1020
tattacaaca gcactggcaa catgggagtg cttgcaggtc aggcctctca gttgaacgca 1080
gtggtggact tgcaagacag aaacacagaa ctgtcatacc agctcttgct tgattccatg 1140
ggtgacagaa ccagatactt ttccatgtgg aatcaggcag tggacagtta tgacccagat 1200
gtcagaatta ttgaaaatca tggaactgaa gacgagctcc ccaactattg tttccctctg 1260
ggcggcgtaa tcaatacgga aactttcaca aaagtaaaac ctaaagctgc acaggacgct 1320
cagtgggaaa aagattcaga attttcagat aaaaatgaaa taagggtggg aaacaacttc 1380
gccatggaaa ttaacctcaa tgccaatctg tggaggaact ttttgtactc caacgtagcc 1440
ctctacttgc ctgacaagct taagtatact ccatccaatg tgcaaatttc caacaatccc 1500
aactcctacg attacatgaa caagcgagtg gtggccccgg ggctggtgga ctgctacatc 1560
aacctgggcg cgcgctggtc gctggactac atggacaacg tcaacccctt caaccaccac 1620
cgcaatgcgg gcctgcgcta ccgctccatg ctcctgggca acgggcgcta cgtgcccttc 1680
cacatccagg tgccccagaa gttctttgcc atcaagaacc tcctcctcct gccgggctcc 1740
tacacctacg agtggaactt caggaaggat gtcaacatgg tcctccagag ctctctgggt 1800
aacgatctca gggtggacgg ggccagcatc aagttcgaga gcatctgcct ctacgccacc 1860
ttcttcccca tggcccacaa cacggcctcc acgctcgagg ccatgctcag gaacgacacc 1920
aacgaccagt ccttcaatga ctacctttcc gccgccaaca tgctctaccc catacccgcc 1980
```

aacgccacca acgtccccat ctccatcccc tcgcgcaact gggcggcctt ccgcggctgg 2040
gccttcaccc gcctcaagac caaggagacc ccctccctgg gctcgggatt cgacccctac 2100
tacacctact cgggctccat tccctacctg gacggcacct tctacctcaa ccacactttc 2160
aagaaggtct cggtcacctt cgactcctcg gtcagctggc cgggcaacga ccgtctgctc 2220
acccccaacg agttcgagat caagcgctcg gtcgacgggg agggctacaa cgtggcccag 2280
tgcaacatga ccaaggactg gttcctggtc cagatgctgg ccaactacaa catcggctac 2340
cagggcttct acatcccaga gagctacaag gacaggatgt actccttctt caggaacttc 2400
cagcccatga gccggcaggt ggtggaccag accaagtaca aggactacca ggaggtgggc 2460
atcatccacc agcacaacaa ctcgggcttc gtgggctacc tcgcccccac catgcgcgag 2520
ggacaggcct accccgccaa cttcccctac ccgctcatag gcaagaccgc ggtcgacagc 2580
atcacccaga aaaagttcct ctgcgaccgc accctctggc gcatcccctt ctccagcaac 2640
ttcatgtcca tgggtgcgct ctcggacctg ggccagaact tgctctacgc caactccgcc 2700
cacgccctcg acatgacctt cgaggtcgac cccatggacg agcccaccct tctctatgtt 2760
ctgttcgaag tctttgacgt ggtccgggtc caccagccgc accgcggcgt catcgagacc 2820
gtgtacctgc gtacgccctt ctcggccggc aacgccacca cctaa 2865

<210> 17
<211> 2823
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 4 Hexon

<400> 17

atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggatcggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggctcccaa gggagcgccc aattccagcc agtgggagca aaaaaagact 420
ggcaataatg ccaatggaga tacggagaat gtcacttatg gtgtagctgc catgggagga 480
attgacatcg ataaaaatgg ccttcaaatt ggaaccgatg acaccaaaga tgacgataat 540
gaaatttatg cagacaaaac atatcagcct gagccgcaaa taggagagga aaactggcaa 600
gaaacatatt cctactatgg aggtagagct cttaaaaaag ataccaaaat gaagccatgc 660
tatggctcat ttgccagacc taccaatgtg aaaggaggac aggcaaaaat aaaaacagat 720
ggagatgtta agtcatttga catagaccta gccttctttg atattcccaa ttctggcgcg 780
ggaaatggca caaatgttaa cgatgatcca gatatggtta tgtatacaga aaatgtaaat 840
ctggaaaccc cagatactca tattgtgtac aaaccaggaa cttcagatga cagctcaaag 900
gtcaacttgt gtcagcaatc catgcctaac agacccaatt atattggctt cagagacaat 960
tttattgggc ttatgtacta caacagcact ggcaatatgg gtgtgctggc tggtcaggcc 1020
tctcaactga atgccgtggt ggacttgcaa gacagaaaca cagagctgtc ctaccagctc 1080
ttgcttgact ctctgggtga cagaaccagg tatttcagta tgtggaatca ggcggtggac 1140
agttatgatc ctgatgtgcg cattattgaa aaccatggtg tggaggatga attgccaaac 1200
tattgcttcc ccttggatgg agcaggcacc aattcggttt accaaggtgt taaaccaaaa 1260
actgacaatg gcaacgatca gtgggaaaca gattccacag tttcaagtca caatcagata 1320
tgcaaaggca atatctatgc catggagatc aacctccagg ccaacctgtg gagaagtttt 1380
ctctactcga acgtggccct gtacctgccc gattcttaca agtacacgcc ggccaacatc 1440
accctgccca ccaacaccaa cacctacgat tacatgaacg ggagagtggt gcctccctcg 1500
ctggtggacg cctacatcaa catcggggcg cgctggtcgc tggaccccat ggacaacgtg 1560
aatcccttca accaccaccg caacgcgggc ctgcgctacc gctccatgct cctgggcaac 1620
gggcgctacg tgccccttcca catccaggtg ccccagaaat ttttcgccat caagagcctc 1680
ctgctcctgc ccgggtccta cacctacgag tggaacttcc gcaaggacgt caacatgatc 1740
ctgcagagct ccctcggcaa cgacctgcgc acggacgggg cctccatctc cttcaccagc 1800
atcaacctct acgccacctt cttccccatg gcgcacaaca cggcctccac gctcgaggcc 1860
atgctgcgca acgacaccaa cgaccagtcc ttcaacgact acctctcggc ggccaacatg 1920
ctctacccca tcccggccaa cgccaccaac gtgcccatct ccatcccctc gcgcaactgg 1980

gccgccttcc gcggctggtc cttcacgcgc ctcaagaccc gcgagacgcc ctcgctgggc 2040
tccgggttcg acccctactt cgtctactcg ggctccatcc cctacctcga cggcaccttc 2100
tacctcaacc acaccttcaa gaaggtctcc atcaccttcg actcctccgt cagctggccc 2160
ggcaacgacc gcctcctgac gcccaacgag ttcgaaatca agcgcaccgt cgacggagag 2220
ggatacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2280
cactacaaca tcggctacca gggcttctac gtgcccgagg gctacaagga ccgcatgtac 2340
tccttcttcc gcaacttcca gcccatgagc cgccaggtgg tggacgaggt caactacaag 2400
gactaccagg ccgtcaccct ggcctaccag cacaacaact cgggcttcgt cggctacctc 2460
gcgcccacca tgcgccaggg ccagccctac cccgccaact acccgtaccc gctcatcgga 2520
aagagcgccg tcaccagcgt cacccagaaa aagttcctct gcgacagggt catgtggcgc 2580
atccccttct ccagcaactt catgtccatg ggcgcgctca ccgacctcgg ccagaacatg 2640
ctctatgcca actccgccca cgcgctagac atgaatttcg aagtcgaccc catggatgag 2700
tccacccttc tctatgttgt cttcgaagtc ttcgacgtcg tccgagtgca ccagccccac 2760
cgcggcgtca tcgaggccgt ctacctgcgc acccccttct cggccggtaa cgccaccacc 2820
taa 2823

<210> 18
<211> 2823
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 5 Hexon

<400> 18

atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggatcggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggctcccaa gggagcgccc aattccagcc agtgggagca aaaaaagact 420
ggcaataatg ccaatggaga tacggagaat gtcacttatg gtgtagctgc catgggagga 480
attgacatcg ataaaaatgg ccttcaaatt ggaaccgatg acaccaaaga tgacgataat 540
gaaatttatg cagacaaaac atatcagcct gagccgcaaa taggagagga aaactggcaa 600
gaaacatatt cctactatgg aggtagagct cttaaaaaag ataccaaaat gaagccatgc 660
tatggctcat ttgccagacc taccaatgtg aaaggaggac aggcaaaaat aaaaacagat 720
ggagatgtta agtcatttga catagaccta gccttctttg atattcccaa ttctggcgcg 780
ggaaatggca caaatgttaa cgatgatcca gatatggtta tgtatacaga aaatgtaaat 840
ctggaaaccc cagatactca tattgtgtac aaaccaggaa cttcagatga cagctcaaag 900
gtcaacttgt gtcagcaatc catgcctaac agacccaatt atattggctt cagagacaat 960
tttattgggc ttatgtacta caacagcact ggcaatatgg gtgtgctggc tggtcaggcc 1020
tctcaactga atgccgtggt ggacttgcaa gacagaaaca cagagctgtc ctaccagctc 1080
ttgcttgact ctctgggtga cagaaccagg tatttcagta tgtggaatca ggcggtggac 1140
agttatgatc ctgatgtgcg cattattgaa aaccatggtg tggaggatga attgccaaac 1200
tattgcttcc ccttggatgg agcaggcacc aattcggttt accaaggtgt taaaccaaaa 1260
actgacaatg gcaacgatca gtgggaaaca gattccacag tttcaagtca caatcagata 1320
tgcaaaggca atatctatgc catggagatc aacctccagg ccaacctgtg gagaagtttt 1380
ctctactcga acgtggccct gtacctgccc gattcttaca agtacacgcc ggccaacatc 1440
accctgccca ccaacaccaa cacctacgat tacatgaacg ggagagtggt gcctccctcg 1500
ctggtggacg cctacatcaa catcggggcg cgctggtcgc tggaccccat ggacaacgtg 1560
aatcccttca accaccaccg caacgcgggc ctgcgctacc gctccatgct cctgggcaac 1620
gggcgctacg tgccccttcca catccaggtg ccccagaaat tttttgccat caagagcctc 1680
ctgctcctgc ccgggtccta cacctacgag tggaacttcc gcaaggacgt caacatgatc 1740
ctgcagagct ccctcggcaa cgacctgcgc acggacgggg cctccatctc cttcaccagc 1800
atcaacctct acgccacctt cttccccatg gcgcacaaca cggcctccac gctcgaggcc 1860
atgctgcgca acgacaccaa cgaccagtcc ttcaacgact acctctcggc ggccaacatg 1920
ctctacccca tcccggccaa cgccaccaac gtgcccatct ccatcccctc gcgcaactgg 1980

```
gccgccttcc gcggctggtc cttcacgcgc ctcaagaccc gcgagacgcc ctcgctgggc 2040
tccgggttcg acccctactt cgtctactcg ggctccatcc cctacctcga cggcaccttc 2100
tacctcaacc acaccttcaa gaaggtctcc atcaccttcg actcctccgt cagctggccc 2160
ggcaacgacc gcctcctgac gcccaacgag ttcgaaatca agcgcaccgt cgacggagag 2220
ggatacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2280
cactacaaca tcggctacca gggcttctac gtgcccgagg gctacaagga ccgcatgtac 2340
tccttcttcc gcaacttcca gcccatgagc cgccaggtcg tggacgaggt caactacaag 2400
gactaccagg ccgtcaccct ggcctaccag cacaacaact cgggcttcgt cggctacctc 2460
gcgcccacca tgcgccaggg ccagccctac cccgccaact acccctaccc gctcatcggc 2520
aagagcgccg tcgccagcgt cacccagaaa aagttcctct gcgaccgggt catgtggcgc 2580
atccccttct ccagcaactt catgtccatg ggcgcgctca ccgacctcgg ccagaacatg 2640
ctctacgcca actccgccca cgcgctagac atgaatttcg aagtcgaccc catggatgag 2700
tccacccttc tctatgttgt cttcgaagtc ttcgacgtcg tccgagtgca ccagccccac 2760
cgcggcgtca tcgaggccgt ctacctgcgc acccccttct cggccggtaa agccaccacc 2820
taa                                                               2823
```

<210> 19
<211> 2823
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 7 Hexon

<400> 19

```
atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccttactc cggcaccgct 360
tacaacagcc tggcccccaa gggagcaccc aattccagcc agtgggagca aaaaaagact 420
ggcaaaaatg ccaatggaga tacggagaat gtcacttatg gtgtagctgc catgggagga 480
attgacatcg ataaaaatgg ccttcaaatt ggaaccgatg acaccaaaga tggcgataat 540
gaaatttatg cagacaaaac atatcagcct gagccgcaaa taggagagga aaactggcaa 600
gaaacatatt cctactatgg aggtagagct cttaaaaaag ataccaaaat gaagccatgc 660
tatggctcat ttgctagacc taccaatgtg aaaggaggac aggcaaaaat aaaaacagat 720
ggagatgtta agtcatttga catagaccta gccttctttg atattccaaa ttctggcgcg 780
ggaaatggca caaatgttaa cgatgatcca gatatggtta tgtatacaga aaatgtaaat 840
ctggaaaccc cagatactca tattgtgtac aaaccaggaa cttcagatga cagctccgag 900
gtcaacttgt gtcagcaatc catgcctaac agacccaatt atattggctt cagagacaat 960
tttattgggc ttatgtacta caacagcact ggcaatatgg gtgtgctggc tggtcaggcc 1020
tctcaactga atgccgtggt ggacttgcaa gacagaaaca cagagctgtc ctaccagctc 1080
ttgcttgact ctctgggtga cagaaccagg tatttcagta tgtggaatca ggcggtggac 1140
agttatgatc ctgatgtgcg cattattgaa aaccatggtg tggaggatga attgccaaac 1200
tattgcttcc ccttggatgg agcaggcacc aattcggttt accaaggtgt taaaccaaaa 1260
actgacaatg gcaacgatca gtgggaaaca gattccacag tttcaagtca caatcagata 1320
tgcaaaggca atatctatgc catggagatc aatctccagg ccaacctgtg gagaagtttc 1380
ctctactcga acgtggccct gtacctgccc gattcttaca agtacacgcc ggccaacatc 1440
accctgccca ccaacaccaa cacctacgat tacatgaacg ggagagtggt gcctccctcg 1500
ctggtggatg cctacatcaa catcggagcg cgctggtcgc tggaccccat ggacaacgtc 1560
aatcccttca accaccaccg caatgcgggg ctgcgctacc gctccatgct cctgggcaac 1620
gggcgctacg tgccccttcca catccaggtg ccccagaaat ttttcgccat caagagcctt 1680
ctgctcctgc ccgggtccta cacctacgag tggaacttcc gcaaggacgt caacatgatc 1740
ctgcagagct ccctcggcaa cgacctgcgc acggacgggg cctccatctc cttcaccagc 1800
atcaacctct acgccacctt cttccccatg gcgcacaaca cggcctccac gctcgaggcc 1860
atgctgcgca acgacaccaa cgaccagtcc ttcaacgact acctctcggc ggccaacatg 1920
ctctacccca tcccggccaa cgccaccaac gtgcccatct ccatcccctc gcgcaactgg 1980
```

gccgccttcc gcggctggtc cttcacgcgc ctcaagacca aggagacgcc ctcgctgggc 2040
tccgggttcg acccatactt cgtctactcg ggctccatcc cctacctcga cggcaccttc 2100
tacctcaacc acaccttcaa gaaggtctcc atcaccttcg attcctccgt cagctggccc 2160
ggcaacgacc ggctcctgac gcccaacgag ttcgaaatca agcgcaccgt cgacggcgag 2220
ggatacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2280
cactacaaca tcggctacca gggcttctac gtgcccgagg gctacaagga ccgcatgtac 2340
tccttcttcc gcaacttcca gcccatgagc cgccaggtgg tggacgaggt caactacaag 2400
gactaccagg ccgtcaccct ggcctaccag cacaacaact cgggcttcgt cggctacctc 2460
gcgcccacca tgcgccaggg ccagccctac cccgccaact acccgtaccc gctcatcggc 2520
aagagcgccg tcaccagcgt cacccagaaa aagttcctct gcgacagggt catgtggcgc 2580
atccccttct ccagcaactt catgtccatg ggcgcgctca ccgacctcgg gcagaacatg 2640
ctctatgcca actccgccca cgcgctagac atgaatttcg aagtcgaccc catggatgag 2700
tccacccttc tctatgttgt cttcgaagtc ttcgacgtcg tccgagtgca ccagccccac 2760
cgcggcgtca tcgaggccgt ctacctgcgc acccccttct cggccggtaa cgccaccacc 2820
taa 2823

<210> 20
<211> 2793
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 9 Hexon

<400> 20

atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcacccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggatcggggc cctagcttca aaccctactc cggcaccgct 360
tacaacagcc tggctcccaa gggagcgccc aacacttgcc agtggacata tactgataac 420
caaactgaga aaacagccac atatggaaat gcacccgtag agggcattaa cattacaaaa 480
gatggcattc aacttggaac tgacagcgat ggtcaggcaa tctatgcaga cgaaacttat 540
cagcccgaac ctcaggtggg agatcctgaa tggcatgata ccacaggtac agaagaaaaa 600
tatggaggca gagcgcttaa acctgccacc gacatgaaac cttgctatgg ctcttttgcc 660
aagccaacta atgttaaggg aggtcaggcc aaaagcagaa caaaaactga tggaacaact 720
gagcctgata ttgacatggc cttttttgat ggcagaaatg caacaacagc tggtttgact 780
ccagaaattg ttttgtatac tgaaaatgtg gatctggaaa ctccagatac ccatattgta 840
tacaaggcag gcacagatga cagcagctct tctatcaatt tgggtcagca gtccatgccc 900
aacagaccca actacattgg cttcagagac aactttatcg ggctcatgta ctacaacagc 960
actggcaata tgggtgtact ggctggacag gcctcccagc tgaatgctgt ggtggacttg 1020
caggacagaa acactgaact gtcctaccag ctcttgcttg actctctggg tgacagaacc 1080
aggtatttca gtatgtggaa tcaggcggtg gacagttatg accccgatgt gcgcattatt 1140
gaaaatcacg gtgtggagga tgaactcccc aactattgct tccccctgaa tgctgtgggt 1200
agaacaaata gttatcaggg aattaaaccc aatggaggcg atccagctac atgggccaaa 1260
gatgaaagcg tcaatgattc taatgaattg ggcaagggca atcctttcgc catggagatc 1320
aacatccagg ccaacctgtg gcggaacttc ctctacgcga acgtggcgct gtacctgccc 1380
gactcctaca agtacacgcc ggccaacatc acgctgcccg ccaacaccaa cacctacgat 1440
tacatgaacg gccgcgtggt ggcgccctcg ctggtggacg cctacatcaa catcggggcg 1500
cgctggtcgc tggaccccat ggacaacgtc aacccttca accaccaccg caacgcgggc 1560
ctgcgctacc gctccatgct cctgggcaac gggcgctacg tgccttcca catccaggtg 1620
ccccaaaagt ttttcgccat caagagcctc ctgctcctgc ccgggtccta cacctacgag 1680
tggaacttcc gcaaggacgt caacatgatc ctgcagagct ccctcggcaa cgacctgcgc 1740
acggacgggg cctccatcgc cttcaccagc atcaacctct acgccacctt cttccccatg 1800
gcgcacaaca ccgcctccac gctcgaggcc atgctgcgca acgacaccaa cgaccagtcc 1860
ttcaacgact acctctcggc ggccaacatg ctctacccca tcccggccaa cgccaccaac 1920
gtgcccatct ccatcccctc gcgcaactgg gccgccttcc gcggatggtc cttcacgcgc 1980

ctcaagaccc gcgagacgcc ctcgctaggc tccgggttcg acccctactt cgtctactcg 2040
ggctccatcc cctacctcga cggcaccttc tacctcaacc acaccttcaa gaaggtctcc 2100
atcaccttcg actcctccgt cagctggccc ggcaacgacc gcctcctgac gcccaacgag 2160
ttcgaaatca agcgcaccgt cgacggagag ggatacaacg tggcccagtg caacatgacc 2220
aaggactggt tcctggtcca gatgctggcc cactacaaca tcggctacca gggcttctac 2280
gtgcccgagg gctacaagga ccgcatgtac tccttcttcc gcaacttcca gcccatgagc 2340
cgccaggtcg tggacgaggt caactacaag gactaccagg ccgtcaccct ggcctaccag 2400
cacaacaact cgggcttcgt cggctacctc gcgcccacca tgcgccaggg ccagccctac 2460
cccgccaact acccctaccc gctcatcggc aagagcgccg tcgccagcgt cacccagaaa 2520
aagttcctct gcgaccgggt catgtggcgc atccccttct ccagcaactt catgtccatg 2580
ggcgcgctca ccgacctcgg ccagaacatg ctctacgcca actccgccca cgcgctagac 2640
atgaatttcg aagtcgaccc catggatgag tccacccttc tctatgttgt cttcgaagtc 2700
ttcgacgtcg tccgagtgca ccagccccac cgcggcgtca tcgaggccgt ctacctgcgc 2760
acgcccttct cggccggcaa cgccaccacc taa 2793

<210> 21
<211> 2793
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 10 Hexon

<400> 21

atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggatcggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggctcccaa gggagcgccc aacacttgcc agtggacata tactgataac 420
caaactgaga aaacagccac atatggaaat gcgcctgtgc aaggcattag tattacaaaa 480
gatggtattc aacttggaac tgacactgat gatcagccca tttatgcaga taaaacttat 540
caaccagagc ctcaagtggg tgatgctgaa tggcatgaca tcactggtac tgatgaaaaa 600
tatggaggca gagctctcaa gcctgacacc aaaatgaagc cctgctatgg ttcttttgcc 660
aagcctacca ataaagaagg aggtcaggca aatgtgaaaa ccgaaacagg cggtaccaaa 720
gaatatgaca ttgacatggc attcttcgat aatcgaagtg cagctgcggc tggcctggcc 780
ccagaaattg ttttgtatac tgagaatgtg gatctggaaa ctccagatac tcatattgta 840
tacaaggcag gcacagatga cagcagctct tctatcaatt tgggtcagca gtccatgccc 900
aacagaccca actacattgg cttcagagac aactttatcg gtctcatgta ctacaacagc 960
actggcaata tgggtgtact ggctggtcag gcctcccagc tgaatgctgt ggtggacttg 1020
caggacagaa acactgaact gtcctaccag ctcttgcttg actctctggg tgacagaacc 1080
aggtatttta gtatgtggaa tcaggcggtg gacagttatg accccgatgt gcgcattatt 1140
gaaaatcacg gtgtggagga tgaactccct aattattgct tccccccttaa tgctgtgggt 1200
agaactgata cttaccaggg aattaaggcc aatggtgctg atcaaaccac atggaccaaa 1260
gatgatactg ttaatgatgc taatgaattg ggcaagggca atcctttcgc catggagatc 1320
aacatccagg ccaacctgtg gcggaacttc ctctacgcga acgtggccct gtacctgccc 1380
gactcctaca agtacacgcc ggccaacatc acgctgccca ccaacaccaa cacctacgat 1440
tacatgaacg gccgcgtggt ggcgccctcg ctggtggacg cctacatcaa catcggggcg 1500
cgctggtcgc tggaccccat ggacaacgtc aacccttca accaccaccg caacgcgggc 1560
ctgcgctacc gctccatgct cctgggcaac gggcgctacg tgcccttcca catccaggtg 1620
ccccaaaagt tcttcgccat caagagcctc ctgctcctgc ccgggtccta cacctacgag 1680
tggaacttcc gcaaggacgt caacatgatc ctgcagagct ccctcggcaa cgacctgcgc 1740
acggacgggg cctccatcgc cttcaccagc atcaacctct acgccacctt cttccccatg 1800
gcgcacaaca ccgcctccac gctcgaggcc atgctgcgca acgacaccaa cgaccagtcc 1860
ttcaacgact acctctcggc ggccaacatg ctctacccca tcccggccaa tgccaccaac 1920
gtgcccatct ccatcccctc gcgcaactgg gccgccttcc gcggatggtc cttcacgcgc 1980
ctcaagaccc gcgagacgcc ctcgctaggc tccgggttcg acccctactt cgtctactcg 2040

```
ggctccatcc cctacctcga cggcaccttc tacctcaacc acaccttcaa gaaggtctcc 2100
atcaccttcg actcctccgt cagctggccc ggcaacgacc gcctcctgac gcccaacgag 2160
ttcgaaatca agcgcaccgt cgacggagag gggtacaacg tggcccagtg caacatgacc 2220
aaggactggt tcctggtcca gatgctggcc cactacaaca tcggctacca gggcttctac 2280
gtgcccgagg gctacaagga ccgcatgtac tccttcttcc gcaacttcca gcccatgagc 2340
cgccaggtcg tggacgaggt caactacaag gactaccagg ccgtcaccct ggcctaccag 2400
cacaacaact cgggcttcgt cggctacctc gcgcccacca tgcgccaggg ccagccctac 2460
cccgccaact acccctaccc gctcatcggc aagagcgccg tcgccagcgt cacccagaaa 2520
aagttcctct gcgaccgggt catgtggcgc atccccttct ccagcaactt catgtccatg 2580
ggcgcgctca ccgacctcgg ccagaacatg ctctacgcca actccgccca cgcgctagac 2640
atgaatttcg aagtcgaccc catggatgag tccacccttc tctatgttgt cttcgaagtc 2700
ttcgacgtcg tccgagtgca ccagccccac cgcggcgtca tcgaggccgt ctacctgcgc 2760
acgcccttct cggccggcaa cgccaccacc taa                               2793
```

<210> 22
<211> 2883
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 11 Hexon

<400> 22

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagtcccgg gctggtgcag ttcgctcgcg ccaccgagag ctacttcagt 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tcccagcgcc tgacgctgcg gttcatcccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggcgtgct ggaccgcggc cccaccttca agccctactc cggcaccgcy 360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact 420
caggcagttg aagaagcagc agaagaggag gaagaagatg ctgacggtca agctgaggaa 480
gagcaagcag ctaccaaaaa gactcatgta tatgctcagg ctcccctttc cggcgaaaaa 540
attagcaaag acggtctgca gataggaacg gacgctacag caaccgaaca aaaacctatt 600
tatgcagacc ctacattcca gcccgaaccc caaatcgggg agtcccagtg gaatgaggca 660
gatgctacag tcgctggtgg tagagtgctc aagaaaacca ctcccatgaa accatgctat 720
ggttcctatg caagacccac gaatgctaat ggaggtcagg gtgtactagc ggcaaatgcc 780
caaggacagc tagaatctca ggttgaaatg caattctttt caacttctga aaacgcccgt 840
aacgaggcta acaacattca gcccaaattg gtgctgtata gcgaggatgt gcacatggag 900
accccggata cacacctctc ttacaagccc acaaaaagcg atgacaattc taaagttatg 960
ctgggccaac aggccatgcc caacaggcct aattacattg gcttcagaga caactttatc 1020
ggtctcatgt actacaacag cactggcaac atgggagtgc ttgcaggtca ggcctctcag 1080
ttgaatgcag tggtggactt gcaagacaga aacacagaac tgtcctacca gctcttgctt 1140
gattccatgg gtgacagaac cagatatttc tccatgtgga atcaggcagt ggacagttat 1200
gacccagatg tcagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt 1260
ttccctctgg gcggcatagg ggtaactgac acttaccagg ctgttaagac caacaatggc 1320
aataatgggg gtcaggtgac ttggacaaaa gatgaaactt ttgcagagcg caatgagata 1380
ggggtgggaa acaatttcgc catggagatc aacctcaatg ccaacctgtg gaggaacttc 1440
ctgtactcca acgtggccct gtacctgcca gacaagctta agtacaaccc ctccaacgtg 1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggt ggccccgggg 1560
ctggtggact gctacatcaa cctgggcgcg cgctggtcgc tggactacat ggacaacgtc 1620
aacctttca accaccaccg caacgcgggc ctgcgctacc gctccatgct cctgggcaac 1680
gggcgctacg tgcccttcca catccaggtg ccccagaagt tctttgccat caagaacctc 1740
ctcctcctgc cgggctccta cacctacgag tggaacttca ggaaggatgt caacatggtc 1800
ctccagagct ctctgggcaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc 1860
atctgcctct acgccacctt cttccccatg gcccacaaca ccgcctccac gctcgaggcc 1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctctccgc cgccaacatg 1980
ctctacccca tccccgccaa cgccaccaac gtccccatct ccatcccctc gcgcaactgg 2040
gcggccttcc gcggctgggc cttcacccgc ctcaagacca aggagacccc ctccctgggc 2100
```

```
tcgggattcg acccctacta cacctactcg ggatccattc cctacctgga cggcaccttc 2160
tacctcaacc acactttcaa gaaggtctcg gtcaccttcg actcctcggt cagctggccg 2220
ggcaacgacc gcctgctcac ccccaacgag ttcgagatca agcgctcggt cgacggggag 2280
ggctacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2340
aactacaaca tcggctacca gggcttctac atcccagaga gctacaagga caggatgtac 2400
tccttcttca ggaacttcca gcccatgagc cggcaggtgg tggaccagac caagtacaag 2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggcttcgt gggctacctc 2520
gcccccacca tgcgcgaggg acaggcctac cccgccaact tcccctaccc gctcataggc 2580
aagaccgcgg tcgacagcat cacccagaaa aagttcctct gcgaccgcac cctctggcgc 2640
atccccttct ccagcaactt catgtccatg ggtgcgctca cggacctggg ccagaacctg 2700
ctctatgcca actccgccca cgcgctcgac atgaccttcg aggtcgaccc catggacgag 2760
cccacccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac 2820
cgcggcgtca tcgagaccgt gtacctgcgc acgcccttct cggccggcaa cgccaccacc 2880
taa                                                               2883
```

<210> 23
<211> 2835
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 16 Hexon

<400> 23

```
atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggcccccaa gggagctccc aattccagtc agtgggagca gacggagaac 420
gggggcggac aggctacgac taaaacacac acctatggag ttgccccaat gggtggaact 480
aatattacag tcgacggact acaaattgga actgacgcta cagctgatac ggaaaaacca 540
atttatgctg ataaaacatt ccaacctgag cctcagatag gagaggaaaa ctggcaagaa 600
actgaaagct tttatggcgg tagggctctt aagaaagaca caaacatgaa gccttgttat 660
ggctcatttg ccagacctac caatgaaaag ggaggtcaag ctaaacttaa agttggagct 720
gatgggctgc cgaccaaaga atttgacata gacctagcat tctttgatac tcctggtggc 780
actgtgaccg gaggtacaga ggagtataaa gcagatattg ttatgtatac cgaaaacacg 840
tatctggaaa ctccagacac acatgtggtg tataaaccag gcaaggataa cacaagttct 900
aaaattaacc tggtccagca gtctatgccc aacaggccca actacattgg gtttagggac 960
aactttattg ggctcatgta ttacaacagc actggcaata tgggtgtgct ggccggtcag 1020
gcttctcagt tgaatgctgt ggttgacttg caagacagaa acactgaact gtcttaccag 1080
ctcttgcttg actctttggg tgacagaacc aggtatttca gtatgtggaa tcaggcggtg 1140
gacagttatg atcctgatgt gcgcattatt gaaaaccatg gtgtggaaga tgaacttccc 1200
aactattgct tccccctgga tgggtctggc actaacgccg cttaccaagg tgtgaaagta 1260
aaaaatggtc aagatggtga tgttgagagc gaatgggaaa aagatgatac tgtcgcagct 1320
cgaaatcaat tatgcaaggg caacattttt gccatggaga tcaatctcca ggccaacctg 1380
tggagaagtt ttctctactc gaacgtggcc ctgtacctgc ccgattctta caagtacacg 1440
ccggccaaca tcaccctgcc caccaacacc aacacctacg attacatgaa cgggagagtg 1500
gtgcctccct cgctggtgga cgcctacatc aacatcgggg cgcgctggtc gctggacccc 1560
atggacaacg tcaatccctt caaccaccat cgcaacgcgg ggctgcgcta ccgctccatg 1620
ctcctgggca acgggcgcta cgtgcccttc cacatccagg tgccccagaa atttttcgcc 1680
attaagagcc tcctgctcct gcccgggtcc tacacctacg agtggaactt ccgcaaggac 1740
gtcaacatga tcctgcagag ctccctcggc aacgacctgc gcacggacgg ggcctccatc 1800
tccttcacca gcatcaacct ctacgccacc ttcttcccca tggcgcacaa caccgcctcc 1860
acgctcgagg ccatgctgcg caacgacacc aacgaccagt ccttcaacga ctacctctcg 1920
gcggccaaca tgctctaccc catcccggcc aacgccacca acgtgcccat ctccatcccc 1980
tcgcgcaact gggccgcctt ccgcggctgg tccttcacgc gcctcaagac caaggagacg 2040
```

```
ccctcgctgg gctccgggtt cgacccctac ttcgtctact cgggctccat cccctacctc 2100
gacggcacct tctacctcaa ccacaccttc aagaaggtct ccatcacctt cgactcctcc 2160
gtcagctggc ccggcaacga ccggctcctg acgcccaacg agttcgaaat caagcgcacc 2220
gtcgacggcg agggctacaa cgtggcccag tgcaacatga ccaaggactg gttcctggtc 2280
cagatgctgg cccactacaa catcggctac cagggcttct acgtgcccga gggctacaag 2340
gaccgcatgt actccttctt ccgcaacttc cagcccatga gccgccaggt cgtggacgag 2400
gtcaactaca aggactacca ggccgtcacc ctggcctacc agcacaacaa ctcgggcttc 2460
gtcggctacc tcgcgcccac catgcgccag ggccagccct accccgccaa ctacccctac 2520
ccgctcatcg gcaagagcgc cgtcgccagc gtcacccaga aaaagttcct ctgcgaccgg 2580
gtcatgtggc gcatcccctt ctccagcaac ttcatgtcca tgggcgcgct caccgacctc 2640
ggccagaaca tgctctacgc caactccgcc cacgcgctag acatgaattt cgaagtcgac 2700
cccatggatg agtccaccct tctctatgtt gtcttcgaag tcttcgacgt cgtccgagtg 2760
caccagcccc accgcggcgt catcgaggcc gtctacctgc gcaccccctt ctcggccggt 2820
aacgccacca cctaa                                                  2835
```

<210> 24
<211> 2883
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 17 Hexon

<400> 24

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcngagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tctcagcgcc tgacgctgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggggtgct ggaccggggt cccactttca agccctactc tggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact 420
caggcagttg aagaagcagc agaagaggaa gaagaagatg ctgacggtca agctgaggaa 480
gagcaagcag ctaccaaaaa gactcatgta tatgctcagg ctcccctttc tggcgaaaaa 540
attagtaaag atggtctgca aataggaacg gacgctacag ctacagaaca aaaacctatt 600
tatgcagacc ctacattcca gcccgaaccc caaatcgggg agtcacagtg gaatgaggca 660
gatgctacag tcgccggcgg tagagtgcta aagaaatcta ctcccatgaa accatgctat 720
ggttcctatg caagacccac aaatgctaat ggaggtcagg gtgtactaac ggcaaatgcc 780
cagggacagc tagaatctca ggttgaaatg caattctttt caacttctga aaacgcccgt 840
aacgagacta acaacattca gcccaaattg gtgctgtata gtgaggatgt gcacatggag 900
accccggata cgcacctttc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg 960
ctgggtcagc agtccatgcc caacagacct aattacatcg gcttcagaga taactttatc 1020
ggcctcatgt attacaatag cactggcaac atgggagtgc ttgcaggtca ggcctctcag 1080
ttgaatgcag tggtggactt gcaagacaga aacacagaac tgtcctacca gctcttgctt 1140
gattccatgg gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagttat 1200
gacccagatg ttagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt 1260
ttccctctgg gtggcatagg ggtaactgac acttaccagg ctgttaaaac caacaatggc 1320
aataacgggg gccaggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata 1380
ggggtgggaa acaatttcgc tatggagata aacctcagtg ccaacctgtg gagaaacttc 1440
ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaatgtg 1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggt ggccccgggg 1560
ctggtggact gctacatcaa cctgggcgcg cgctggtcgc tggactacat ggacaacgtc 1620
aacccttca accaccaccg caatgcgggc ctgcgctacc gctccatgct cctgggcaac 1680
gggcgctacg tgcccttcca catccaggtg ccccagaagt tctttgccat caagaacctc 1740
ctcctcctgc cgggctccta cacctacgag tggaacttca ggaaggatgt caacatggtc 1800
ctccagagct ctctgggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc 1860
atctgcctct acgccacctt cttccccatg gcccacaaca cggcctccac gctcgaggcc 1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctctccgc cgccaacatg 1980
ctctacccca tacccgccaa cgccaccaac gtccccatct ccatcccctc gcgcaactgg 2040
```

```
gcggccttcc gcggctgggc cttcacccgc ctcaagacca aggagacccc ctccctgggc 2100
tcgggattcg acccctacta cacctactcg ggctccattc cctacctgga cggcaccttc 2160
tacctcaacc acactttcaa gaaggtctcg gtcaccttcg actcctcggt cagctggccg 2220
ggcaacgacc gtctgctcac ccccaacgag ttcgagatca agcgctcggt cgacggggag 2280
ggctacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2340
aactacaaca tcggctacca gggcttctac atcccagaga gctacaagga caggatgtac 2400
tccttcttca ggaacttcca gcccatgagc cggcaggtgg tggaccagac caagtacaag 2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggcttcgt gggctacctc 2520
gcccccacca tgcgcgaggg acaggcctac cccgccaact tcccctatcc gctcataggc 2580
aagaccgcgg tcgacagcat cacccagaaa aagttcctct gcgaccgcac cctctggcgc 2640
atccccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg 2700
ctctacgcca actccgccca cgccctcgac atgaccttcg aggtcgaccc catggacgag 2760
cccacccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac 2820
cgcggcgtca tcgagaccgt gtacctgcgt acgcccttct cggccggcaa cgccaccacc 2880
taa                                                               2883
```

<210> 25
<211> 2877
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 19 Hexon

<400> 25

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagt 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tcccagcgcc tgacgctgcg gttcatcccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggcgtgct ggaccgcggc cccaccttca agccctactc cggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aactcttgtg agtgggagca attagaagaa 420
gcccaggccg ctttggaaga cgaagaatta gaagatgaag acgaggaacc acaggatgag 480
gcgcctgtga aaaagaccca tgtatacgct caggctcccc tttctggaga agaaattact 540
aaagacggtt tgcaaatagg gtcagataac acagaagctc agtctaagcc tatatatgca 600
gaccctacat tccagcccga accccaaatc ggggagtccc agtggaacga ggcagatgct 660
acagtcgctg gtggtagagt gctcaagaaa accactccca tgaaaccatg ctatggttcc 720
tatgcaagac ccacgaatgc taatggaggt cagggtgtgc tggtggctga tgataagggg 780
gtccttcaat ctaaagttga attgcaattt ttttcaaata ctactactct taatcagcgg 840
gagggtaatg atacaaaacc aaaagtagtg ctgtatagcg aggatgtgca catggaaaca 900
ccagacaccc acatttctta caagcccaca aaaagcgatg acaattctaa agttatgctg 960
ggccaacagt ccatgcccaa caggcctaat tacatcggct tcagagacaa ctttatcggt 1020
ctcatgtact acaacagcac tggcaacatg ggagtgcttg caggtcaggc ctctcagttg 1080
aatgcagtgg tggacttgca agacagaaac acagaactgt cctaccagct cttgcttgat 1140
tccatgggtg acagaaccag atatttctcc atgtggaatc aggcagtgga cagttatgac 1200
ccggatgtca gaattattga aaatcatgga accgaagacg agctccccaa ctattgtttt 1260
cctctgggtg gcataggggt aactgacact taccaggtca ttaaaactaa tggcaatggt 1320
caagcagacc caacctggga aaaagataca gagtttgcag accgcaatga aatagggggtg 1380
ggaaacaatt tcgccatgga gatcaacctc aatgccaacc tgtggaggaa cttcctgtac 1440
tccaacgtgg ccctgtacct gccagacaag cttaagtaca acccctccaa cgtggacatc 1500
tctgacaacc ccaacaccta cgattacatg aacaagcgag tggtggcccc ggggctggtg 1560
gactgctaca tcaacctggg cgcgcgctgg tcgctggact acatggacaa cgtcaacccc 1620
ttcaaccacc accgcaacgc gggcctgcgc taccgctcca tgctcctggg caacgggcgc 1680
tacgtgccct tccacatcca ggtgccccag aagttctttg ccatcaagaa cctcctcctc 1740
ctgccgggct cctacaccta cgagtggaac ttcaggaagg atgtcaacat ggtcctccag 1800
agctctttgg gcaacgatct cagggtggac ggggccagca tcaagttcga gagcatctgc 1860
ctctacgcca ccttcttccc catggcccac aacaccgcct ccacgctcga ggccatgctc 1920
aggaacgaca ccaacgacca gtccttcaat gactacctct ccgccgccaa catgctctac 1980
```

cccatccccg ccaacgccac caacgtccct atctccatcc cctcgcgcaa ctgggcggcc 2040
ttccgcggct gggccttcac ccgcctcaag accaaggaga caccctccct gggctcggga 2100
ttcgacccct actacaccta ctcgggatcc attccctacc tggacggcac cttctacctc 2160
aaccacactt tcaagaaggt ctcggtcacc ttcgactcct cggtcagctg gccgggcaac 2220
gaccgcctgc tcacccccaa cgagttcgag atcaagcgct cggtcgacgg ggagggctac 2280
aacgtggccc agtgcaacat gaccaaggac tggttcctgg tccagatgct ggccaactac 2340
aacatcggct accagggctt ctacatccca gagagctaca aggacaggat gtactccttc 2400
ttcaggaact tccagcccat gagccggcag gtggtggacc aaaccaagta caaggactac 2460
caggaggtgg gcatcatcca ccagcacaac aactcgggct tcgtgggcta cctcgccccc 2520
accatgcgcg agggacaggc ctaccccgcc aacttccccc acccgctcat aggcaagacc 2580
gcggtcgaca gcatcaccca gaaaaagttc ctctgcgacc gcaccctctg gcgcatcccc 2640
ttctccagca acttcatgtc catgggtgcg ctcacggacc tgggccagaa cctgctctat 2700
gccaactccg cccacgcgct cgacatgacc ttcgaggtcg accccatgga cgagcccacc 2760
cttctctatg ttctgttcga agtctttgac gtggtccggg tccaccagcc gcaccgcggc 2820
gtcatcgaga ccgtgtacct gcgcacgccc ttctcggccg gcaacgccac cacctaa 2877

<210> 26
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> polA

<400> 26
aataaaagat ctttattttc attagatctg tgtgttggtt ttttgtgtg 49

<210> 27
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 27
atggaattcg tttaaaccat catcaataat atacctc 37

<210> 28
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 28
cgctggcact caagagtggc ctc 23

<210> 29
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 29

atgaagcttg tttaaaccca tcatcaataa tatacct 37

<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 30
atctagacag cgtccatagc ttaccg 26

<210> 31
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 31
atgctacgta gcgatcgcgt gagtagtgtt tgggggtggg tggg 44

<210> 32
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 32
taggcgcgcc gcttctcctc gttcaggctg gcg 33

<210> 33
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 33
gatctagtta gtttaaacga attcggatct gcgacgcg 38

<210> 34
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 34
ttcgatcatg tttaaacgaa attaagaatt cggatcc 37

<210> 35

<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 35
tattctgcga tcgctgaggt gggtgagtgg gcg 33

<210> 36.
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 36
taggcgcgcc cttaaacggc atttgtggga g 31

<210> 37
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 37
cgtctagaag acccgagtct taccagt 27

<210> 38
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 38
cgggatccgt ttaaaccatc atcaataata taccttatt 39

<210> 39
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 39
atggaattcg tttaaaccat catcaataat atacctt 37

<210> 40
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligomer

<400> 40
atgacgcgat cgctgatatc ctataataat aaaacgcaga ctttg 45

<210> 41
<211> 2880
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 3 Hexon

<400> 41

atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tctcagcgcc tgacgctgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggggtgct ggaccggggt cccactttca agccctactc tggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact 420
caggcagttg aagaagcagc agaagaggaa gaagaagatg ctgacggtca agctgaggaa 480
gagcaagcag ctaccaaaaa gactcatgta tatgctcagg ctccccttc tggcgaaaaa 540
attagtaaag atggtctgca aataggaacg gacgctacag ctacagaaca aaaacctatt 600
tatgcagacc ctacattcca gcccgaaccc caaatcgggg agtcccagtg gaatgaggca 660
gatgctacag tcgccggcgg tagagtgcta aagaaatcta ctcccatgaa accatgctat 720
ggttcctatg caagacccac aaatgctaat ggaggtcagg gtgtactaac ggcaaatgcc 780
cagggacagc tagaatctca ggttgaaatg caattctttt caacttctga aaacgcccgt 840
aacgaggcta acaacattca gcccaaattg gtgctgtata gtgaggatgt gcacatggag 900
accccggata cgcacctttc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg 960
ctgggtcagc agtccatgcc caacagacct aattacatcg gcttcagaga caactttatc 1020
ggcctcatgt attacaatag cactggcaac atgggagtgc ttgcaggtca ggcctctcag 1080
ttgaatgcag tggtggactt gcaagacaga aacacagaac tgtcctacca gctcttgctt 1140
gattccatgg gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagttat 1200
gacccagatg ttagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt 1260
ttccctctgg gtggcatagg ggtaactgac acttaccagg ctgttaaaac caacaatggc 1320
aataacgggg gccaggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata 1380
ggggtgggaa acaatttcgc tatggagatc aacctcagtg ccaacctgtg gagaaacttc 1440
ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaatgtg 1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggt ggccccgggg 1560
ctggtggact gctacatcaa cctgggcgcg cgctggtcgc tggactacat ggacaacgtc 1620
aaccccttca accaccaccg caatgcgggc ctgcgctacc gctccatgct cctgggcaac 1680
gggcgctacg tgcccttcca catccaggtg ccccagaagt tctttgccat caagaacctc 1740
ctcctcctgc cgggctccta cacctacgag tggaacttca ggaaggatgt caacatggtc 1800
ctccagagct ctctgggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc 1860
atctgcctct acgccacctt cttccccatg gcccacaaca cggcctccac gctcgaggcc 1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctttccgc cgccaacatg 1980

ctctacccca tacccgccaa cgccaccaac gtccccatct ccatcccctc gcgcaactgg 2040
gcggccttcc gcggctgggc cttcacccgc ctcaagacca aggagacccc ctccctgggc 2100
tcgggattcg acccctacta cacctactcg ggctctattc cctacctgga cggcaccttc 2160
tacctcaacc acactttcaa gaaggtctcg gtcaccttcg actcctcggt cagctggccg 2220
ggcaacgacc gtctgctcac ccccaacgag ttcgagatca agcgctcggt cgacggggaa 2280
ggctacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc 2340
aactacaaca tcggctacca gggcttctac atcccagaga gctacaagga caggatgtac 2400
tccttcttca ggaacttcca gcccatgagc cggcaggtgg tggaccagac caagtacaag 2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggcttcgt gggctacctc 2520
gcccccacca tgcgcgaggg acaggcctac cccgccaact tcccctaccc gctcataggc 2580
aagaccgcgg tcgacagcat cacccagaaa aagttcctct gcgaccgcac cctctggcgc 2640
atccccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg 2700
ctctacgcca actccgccca cgccctcgac atgaccttcg aggtcgaccc catggacgag 2760
cccacccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac 2820
cgcggcgtca tcgagaccgt gtacctgcgt acgcccttct cggccggcaa cgccaccacc 2880

<210> 42
<211> 1683
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 3 Fiber

<400> 42

atgtcagatt cttgctcctg tccctccgca cccactatct tcatgttgtt gcagatgaag 60
cgcaccaaaa cgtctgacga gagcttcaac cccgtgtacc cctatgacac ggaaagcggc 120
cctccctccg tccctttcct cacccctccc ttcgtgtctc ccgatggatt ccaagaaagc 180
ccccccgggg tcctgtctct gaacctggcc gagcccctgg tcacttccca cggcatgctc 240
gccctgaaaa tgggaagtgg cctctccctg gacgacgctg gcaacctcac ctctcaagat 300
atcaccaccg ctagccctcc cctcaaaaaa accaagacca acctcagcct agaaacctca 360
tccccctaa ctgtaagcac ctcaggcgcc ctcaccgtag cagccgccgc tcccctggca 420
gtggccggca cctccctcac catgcaatca gaggcccccc tgacagtaca ggatgcaaaa 480
ctcaccctgg ccaccaaagg ccccctgacc gtgtctgaag gcaaactggc cttgcaaaca 540
tcggccccgc tgacggccgc tgacagcagc accctcaccg ttagcgccac accaccaatt 600
aatgtaagca gtggaagttt aggcttagac atggaagacc ctatgtatac tcacgatgga 660
aaactgggaa taagaattgg gggtccacta agagtagtag acagcttgca cacactcact 720
gtagttaccg gaaatggact aactgtagat aacaatgccc tccaaactag agttacgggc 780
gccctaggtt atgacacatc aggaaatcta caattgagag ctgcaggagg tatgcgaatt 840
gatgcaaatg gccaacttat ccttaatgtg gcatacccat ttgatgctca gaacaatctc 900
agccttagac ttggtcaggg acccctgtat ataaacacag accacaacct ggatttgaat 960
tgcaacagag gtctaaccac aactaccacc aacaacacaa aaaaacttga gactaaaatt 1020
agctcaggct tagactatga caccaatggt gctgtcatta ttaaacttgg cactggtcta 1080
agcttcgaca acacaggcgc cctaactgtg ggaaacactg gtgatgataa actgactctg 1140
tggacgaccc cagacccatc tccaaattgc agaattcact cagacaaaga ctgcaagttt 1200
actctagtcc taactaagtg tggaagccaa atcctggcct ctgtcgccgc cctagcggta 1260
tcaggaaatc tggcttcgat aacaggcacc gttgccagcg ttaccatctt tctcagattt 1320
gatcagaatg gagtgcttat ggaaaactcc tcgctagaca ggcagtactg gaacttcaga 1380
aatggcaact caactaacgc tgcccctac accaatgcag ttgggttcat gccaaacctc 1440
gcagcatacc ccaaaacgca aagccagact gctaaaaaca acattgtaag tcaggtttac 1500
ttgaatggag acaaatccaa acccatgacc cttaccatca ccctcaatgg aactaatgaa 1560
tccagtgaaa ctagccaggt gagtcactac tccatgtcat ttacatgggc ttgggaaagt 1620
gggcaatatg ccactgaaac ctttgccacc aactccttca ccttttctta cattgctgaa 1680
caa 1683

<210> 43
<211> 2859
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 6 Hexon

<400> 43

```
atgtatgtcc gccgaccaga aggaagaggc gcgtcgccga gttgcaagat ggccacccca 60
tcgatgctgc cccagtgggc gtacatgcac atcgccggac aggacgcttc ggagtacctg 120
agtccgggtc tggtgcagtt cgcccgcgcc acagacacct acttcagtct ggggaacaag 180
tttaggaacc ccacggtggc gcccacgcac gatgtgacca ccgaccgcag ccagcggctg 240
acgctgcgct tcgtgcccgt ggaccgcgag gacaacacct actcgtacaa agtgcgctac 300
acgctggccg tgggcgacaa ccgcgtgctg gacatggcca gcacctactt tgacatccgc 360
ggcgtgctgg atcggggccc cagcttcaaa ccctactccg gcaccgccta caacagcctg 420
gctcccaagg gagcgcccaa cacctcacag tggataacca aagacaatgg aactgataag 480
acatacagtt ttggaaatgc tccagtcaga ggattggaca ttacagaaga gggtctccaa 540
ataggacccg atgagtcagg gggtgaaagc aagaaaattt ttgcagacaa aacctatcag 600
cctgaacctc agcttggaga tgaggaatgg catgatacta ttggagctga agacaagtat 660
ggaggcagag cgcttaaacc tgccaccaac atgaaaccct gctatgggtc tttcgccaag 720
ccaactaatg ctaagggagg tcaggctaaa agcagaacca aggacgatgg cactactgag 780
cctgatattg acatggcctt ctttgacgat cgcagtcagc aagctagttt cagtccagaa 840
cttgttttgt atactgagaa tgtcgatctg gacaccccgg atacccacat tatttacaaa 900
cctggcactg atgaaacaag ttcttctttc aacttgggtc agcagtccat gcccaacaga 960
cccaactaca tcggcttcag agacaacttt atcggtctca tgtactacaa cagtactggc 1020
aatatgggtg tactagctgg acaggcctcc cagctgaatg ctgtggtgga cttgcaggac 1080
agaaacactg aactgtccta ccagctcttg cttgactctc tgggtgacag aaccaggtat 1140
ttcagtatgt ggaaccaggc ggtggacagc tacgaccccg atgtgcgcat tattgaaaat 1200
cacggtgtgg aggatgaact acccaactat tgcttccctt tgaatggtgt gggctttaca 1260
gatacattcc agggaattaa ggttaaaact accaataacg gaacagcaaa tgctacagag 1320
tgggaatctg atacctctgt caataatgct aatgagattg ccaagggcaa tcctttcgcc 1380
atggagatca acatccaggc caacctgtgg cggaacttcc tctacgcgaa cgtggcgctg 1440
tacctgcccg actcctacaa gtacacgccg gccaacatca cgctgcccgc caacaccaac 1500
acctacgatt acatgaacgg ccgcgtggta gcgccctcgc tggtggacgc ctacatcaac 1560
atcggggcgc gctggtcgct ggaccccatg gacaacgtca acccccttcaa ccaccaccgc 1620
aacgcgggcc tgcgctaccg ctccatgctc ctgggcaacg ggcgctacgt gcccttccac 1680
atccaggtgc cccaaaagtt tttcgccatc aagagcctcc tgctcctgcc cgggtcctac 1740
acctacgagt ggaacttccg caaggacgtc aacatgatcc tgcagagctc cctcggcaac 1800
gacctgcgca cggacggggc ctccatcgcc ttcaccagca tcaacctcta cgccaccttc 1860
ttccccatgg cgcacaacac cgcctccacg ctcgaggcca tgctgcgcaa cgacaccaac 1920
gaccagtcct tcaacgacta cctctcggcg gccaacatgc tctaccccat cccggccaac 1980
gccaccaacg tgcccatctc catcccctcg cgcaactggg ccgccttccg cggctggtcc 2040
ttcacgcgcc tcaagacccg cgagacgccc tcgctcggct ccgggttcga cccctacttc 2100
gtctactcgg gctccatccc ctacctcgac ggcaccttct acctcaacca caccttcaag 2160
aaggtctcca tcaccttcga ctcctccgtc agctggcccg gcaacgaccg cctcctgacg 2220
cccaacgagt tcgaaatcaa gcgcaccgtc gacggagagg ggtacaacgt ggcccagtgc 2280
aacatgacca aggactggtt cctggttcag atgctggccc actacaacat cggctaccag 2340
ggcttctacg tgcccgaggg ctacaaggac cgcatgtact ccttcttccg caacttccag 2400
cccatgagcc gccaggtcgt ggacgaggtc aactacaagg actaccaggc cgtcaccctg 2460
gcctaccagc acaacaactc gggcttcgtc ggctacctcg cgcccaccat gcgccaggga 2520
cagccctacc ccgccaacta cccctacccg ctcatcggca agagcgccgt cgccagcgtc 2580
acccagaaaa agttcctctg cgaccgggtc atgtggcgca tcccccttctc cagcaacttc 2640
atgtccatgg gcgcgctcac cgacctcggc cagaacatgc tctacgccaa ctccgcccac 2700
gcgctagaca tgaatttcga agtcgacccc atggatgagt ccacccttct ctatgttgtc 2760
ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc gcggcgtcat cgaggccgtc 2820
tacctgcgca cgcccttctc ggccggtaac gccaccacc                        2859
```

<210> 44
<211> 1335
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 6 Fiber

<400> 44

```
atgtccaaaa agcgcgcgcg ggtggatgat gacttcgacc ccgtgtaccc ctacgatgca 60
gacaacgcac cgactgtgcc cttcatcaac cctcccttcg tctcttcaga tggattccaa 120
gaaaagcccc tgggggtgtt gtccctgcga ctggccgatc ccgtcaccac caagaacggg 180
gctgtcaccc tcaagctggg ggaggggggtg gacctcgacg actcgggaaa actcatctcc 240
aaaaatgcca ccaaggccac tgcccctctc agtatttcca acaacaccat ttcccttaac 300
atggataccc ctctttacaa caacaatgga aagctaggta tgaaggtaac cgcaccatta 360
aagatattag acacagatct actaaaaaca cttgttgttg cttatgggca gggattagga 420
acaaacacca atggtgctct tgttgcccaa ctagcatacc cacttgtttt taataccgct 480
agcaaaattg cccttaattt aggcaatgga ccattaaaag tggatgcaaa tagactgaac 540
attaattgca aaagaggtat ctatgtcact accacaaaag atgcactgga gattaatatc 600
agttgggcaa atgctatgac atttatagga aatgccattg gtgtcaatat tgacacaaaa 660
aaaggcctac agttcggcac ttcaagcact gaaacagatg ttaaaaatgc ttttccactc 720
caagtaaaac ttggagctgg tcttacattt gacagcacag gtgccattgt tgcttggaac 780
aaagaagatg acaaacttac actgtggacc acagccgatc catctccaaa ctgtcacata 840
tattctgcaa aggatgctaa gcttacactc tgcttgacaa agtgtggtag tcagatactg 900
ggcactgttt ctctcatagc tgttgatact ggtagcttaa atccaataac aggaaaagta 960
accactgctc ttgtttcact taaattcgat gccaatggag ttttgcaagc cagttcaaca 1020
ctagataaag aatattggaa tttcagaaaa ggagatgtga cacctgctga cccctacact 1080
aatgctatag gctttatgcc caaccttaat gcatacccaa aaaacacaaa cgcagctgca 1140
aaaagtcaca ttgttggaaa agtataccta catggggatg aaagcaagcc actagacttg 1200
ataattacat ttaatgaaac cagtgatgaa tcctgtactt attgcattaa ctttcagtgg 1260
cagtggggaa ctgaccaata taaagatgaa acacttgcag tcagttcatt caccttctca 1320
tacattgcta aagaa                                                  1335
```

<210> 45
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 45
tgtcctacca rctcttgctt ga 22

<210> 46
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 46
gtggaarggc acgtagcg 18

<210> 47
<211> 9
<212> PRT
<213> HIV gag CD8 Epitope

<220>
<223> Primer

<400> 47

Ala Met Gln Met Leu Lys Glu Thr Ile
1 5

<210> 48
<211> 578
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 20 Fiber

<400> 48

```
Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
                165                 170                 175
Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
            180                 185                 190
Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
        195                 200                 205
Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
    210                 215                 220
Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Ser Gly Ala Leu Asn Tyr Asp Thr Ser Gly Asn Leu Glu Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Val Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
            260                 265                 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
    290                 295                 300
Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305                 310                 315                 320
```

```
Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
                325                 330                 335
Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
            340                 345                 350
Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
        355                 360                 365
Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
    370                 375                 380
Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385                 390                 395                 400
Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
                405                 410                 415
Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
            420                 425                 430
Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
        435                 440                 445
Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
    450                 455                 460
Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465                 470                 475                 480
Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
                485                 490                 495
Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
            500                 505                 510
Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
        515                 520                 525
Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
    530                 535                 540
Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545                 550                 555                 560
Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
                565                 570                 575
Gln Glu
```

<210> 49
<211> 425
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 4 Fiber

<400> 49

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp His Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
```

```
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Arg Thr Ser Ile Leu
        115                 120                 125
Asn Thr Leu Ala Leu Gly Phe Gly Ser Gly Leu Gly Leu Arg Gly Ser
    130                 135                 140
Ala Leu Ala Val Gln Leu Val Ser Pro Leu Thr Phe Asp Thr Asp Gly
145                 150                 155                 160
Asn Ile Lys Leu Thr Leu Asp Arg Gly Leu His Val Thr Thr Gly Asp
                165                 170                 175
Ala Ile Glu Ser Asn Ile Ser Trp Ala Lys Gly Leu Lys Phe Glu Asp
            180                 185                 190
Gly Ala Ile Ala Thr Asn Ile Gly Asn Gly Leu Glu Phe Gly Ser Ser
        195                 200                 205
Ser Thr Glu Thr Gly Val Asp Asp Ala Tyr Pro Ile Gln Val Lys Leu
    210                 215                 220
Gly Ser Gly Leu Ser Phe Asp Ser Thr Gly Ala Ile Met Ala Gly Asn
225                 230                 235                 240
Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro
                245                 250                 255
Asn Cys Gln Ile Leu Ala Glu Asn Asp Ala Lys Leu Thr Leu Cys Leu
            260                 265                 270
Thr Lys Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Val Val
        275                 280                 285
Gly Ser Gly Asn Leu Asn Pro Ile Thr Gly Thr Val Ser Ser Ala Gln
    290                 295                 300
Val Phe Leu Arg Phe Asp Ala Asn Gly Val Leu Leu Thr Glu His Ser
305                 310                 315                 320
Thr Leu Lys Lys Tyr Trp Gly Tyr Arg Gln Gly Asp Ser Ile Asp Gly
                325                 330                 335
Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Lys Ala Tyr
            340                 345                 350
Pro Lys Ser Gln Ser Ser Thr Thr Lys Asn Asn Ile Val Gly Gln Val
        355                 360                 365
Tyr Met Asn Gly Asp Val Ser Lys Pro Met Leu Leu Thr Ile Thr Leu
    370                 375                 380
Asn Gly Thr Asp Asp Ser Asn Ser Thr Tyr Ser Met Ser Phe Ser Tyr
385                 390                 395                 400
Thr Trp Thr Asn Gly Ser Tyr Val Gly Ala Thr Phe Gly Ala Asn Ser
                405                 410                 415
Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
            420                 425
```

<210> 50
<211> 444
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 5 Fiber

<400> 50

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
```

```
Lys Leu Gly Asp Gly Val Asp Leu Asp Asp Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Leu Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Asn
    130                 135                 140
Gly Ala Leu Val Ala Gln Leu Ala Tyr Pro Leu Val Phe Asn Thr Ala
145                 150                 155                 160
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Lys Arg Gly Ile Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Ile Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Ile Gly Val Asn Ile Asp Thr Lys Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Ser Ser Thr Glu Thr Asp Val Lys Asn Ala Phe Ser Leu
225                 230                 235                 240
Gln Val Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Ala Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Leu Ala Val Ser Gly Ser Leu Ala Pro Ile Thr Gly Ala Val Arg
305                 310                 315                 320
Thr Ala Leu Val Ser Leu Lys Phe Asn Ala Asn Gly Ala Leu Leu Asp
                325                 330                 335
Lys Ser Thr Leu Asn Lys Glu Tyr Trp Asn Tyr Arg Gln Gly Asp Leu
            340                 345                 350
Ile Pro Gly Thr Pro Tyr Thr His Ala Val Gly Phe Met Pro Asn Lys
        355                 360                 365
Lys Ala Tyr Pro Lys Asn Thr Thr Ala Ala Ser Lys Ser His Ile Val
    370                 375                 380
Gly Asp Val Tyr Leu Asp Gly Asp Ala Asp Lys Pro Leu Ser Leu Ile
385                 390                 395                 400
Ile Thr Phe Asn Glu Thr Asp Asp Glu Thr Cys Asp Tyr Cys Ile Asn
                405                 410                 415
Phe Gln Trp Lys Trp Gly Ala Asp Gln Tyr Lys Asp Lys Thr Leu Ala
            420                 425                 430
Thr Ser Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 51
<211> 445
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 7 Fiber

<400> 51

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Phe Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Thr
    130                 135                 140
Gly Ala Leu Val Ala Gln Leu Ala Ala Pro Leu Ala Phe Asp Ser Asn
145                 150                 155                 160
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Asn Arg Gly Leu Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Thr Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Met Gly Val Asn Ile Asp Thr Gln Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Thr Ser Thr Val Ala Asp Val Lys Asn Ala Tyr Pro Ile
225                 230                 235                 240
Gln Val Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Asp Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Ile Ala Val Asp Thr Gly Ser Leu Asn Pro Ile Thr Gly Gln Val
305                 310                 315                 320
Thr Thr Ala Leu Val Ser Leu Lys Phe Asp Ala Asn Gly Val Leu Gln
                325                 330                 335
Thr Ser Ser Thr Leu Asp Lys Glu Tyr Trp Asn Phe Arg Lys Gly Asp
            340                 345                 350
Val Thr Pro Ala Glu Pro Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn
        355                 360                 365
Leu Lys Ala Tyr Pro Lys Asn Thr Ser Gly Ala Ala Lys Ser His Ile
    370                 375                 380
Val Gly Lys Val Tyr Leu His Gly Asp Thr Asp Lys Pro Leu Asp Leu
385                 390                 395                 400
Ile Ile Thr Phe Asn Glu Thr Ser Asp Glu Ser Cys Thr Tyr Cys Ile
                405                 410                 415
Asn Phe Gln Trp Lys Trp Asp Ser Thr Lys Tyr Thr Gly Glu Thr Leu
            420                 425                 430

        Ala Thr Ser Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
                435                 440                 445
```

<210> 52

<211> 425
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 9 Fiber

<400> 52

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp His Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ala Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Arg Thr Ser Ile Leu
        115                 120                 125
Asn Thr Leu Ala Leu Gly Phe Gly Ser Gly Leu Gly Leu Arg Gly Ser
    130                 135                 140
Ala Leu Ala Val Gln Leu Val Ser Pro Leu Thr Phe Asp Thr Asp Gly
145                 150                 155                 160
Asn Ile Lys Leu Thr Leu Asp Arg Gly Leu His Val Thr Thr Gly Asp
                165                 170                 175
Ala Ile Glu Ser Asn Ile Ser Trp Ala Lys Gly Leu Lys Phe Glu Asp
            180                 185                 190
Gly Ala Ile Ala Thr Asn Ile Gly Asn Gly Leu Glu Phe Gly Ser Ser
        195                 200                 205
Ser Thr Glu Thr Gly Val Asp Asp Ala Tyr Pro Ile Gln Val Lys Leu
    210                 215                 220
Gly Ser Gly Leu Ser Phe Asp Ser Thr Gly Ala Ile Met Ala Gly Asn
225                 230                 235                 240
Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro
                245                 250                 255
Asn Cys Gln Ile Leu Ala Glu Asn Asp Ala Lys Leu Thr Leu Cys Leu
            260                 265                 270
Thr Lys Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Val Val
        275                 280                 285
Gly Ser Gly Asp Leu Asn Pro Ile Thr Gly Thr Val Ser Ser Ala Gln
    290                 295                 300
Val Phe Leu Arg Phe Asp Ala Asn Gly Val Leu Leu Thr Glu His Ser
305                 310                 315                 320
Thr Leu Lys Lys Tyr Trp Gly Tyr Arg Gln Gly Asp Ser Ile Asp Gly
                325                 330                 335
Thr Pro Tyr Ala Asn Ala Val Gly Phe Met Pro Asn Leu Lys Ala Tyr
            340                 345                 350
Pro Lys Ser Gln Ser Ser Thr Thr Lys Asn Asn Ile Val Gly Gln Val
        355                 360                 365
```

```
Tyr Met Asn Gly Asp Val Ser Lys Pro Met Leu Leu Thr Ile Thr Leu
    370                 375                 380
Asn Gly Thr Asp Asp Ser Asn Ser Thr Tyr Ser Met Ser Phe Ser Tyr
385                 390                 395                 400
Thr Trp Thr Asn Gly Ser Tyr Val Gly Ala Thr Phe Gly Ala Asn Ser
                405                 410                 415
Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
            420                 425
```

<210> 53
<211> 425
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 10 Fiber

<400> 53

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp His Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Arg Thr Ser Ile Leu
        115                 120                 125
Asn Thr Leu Ala Leu Gly Phe Gly Ser Gly Leu Gly Leu Arg Gly Ser
    130                 135                 140
Ala Leu Ala Val Gln Leu Val Ser Pro Leu Thr Phe Asp Thr Asp Gly
145                 150                 155                 160
Asn Ile Lys Leu Thr Leu Asp Arg Gly Leu His Val Thr Thr Gly Asp
                165                 170                 175
Ala Ile Glu Ser Asn Ile Ser Trp Ala Lys Gly Leu Lys Phe Glu Asp
            180                 185                 190
Gly Ala Ile Ala Thr Asn Ile Gly Asn Gly Leu Glu Phe Gly Ser Ser
        195                 200                 205
Ser Thr Glu Thr Gly Val Asp Asp Ala Tyr Pro Ile Gln Val Lys Leu
    210                 215                 220
Gly Ser Gly Leu Ser Phe Asp Ser Thr Gly Ala Ile Met Ala Gly Asn
225                 230                 235                 240
Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro
                245                 250                 255
Asn Cys Gln Ile Leu Ala Glu Asn Asp Ala Lys Leu Thr Leu Cys Leu
            260                 265                 270
Thr Lys Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Val Val
        275                 280                 285
Gly Ser Gly Asn Leu Asn Pro Ile Thr Gly Thr Val Ser Ser Ala Gln
    290                 295                 300
Val Phe Leu Arg Phe Asp Ala Asn Gly Val Leu Leu Thr Glu His Ser
305                 310                 315                 320
```

```
Thr Leu Lys Lys Tyr Trp Gly Tyr Arg Gln Gly Asp Ser Ile Asp Gly
                325                 330                 335
Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Lys Ala Tyr
            340                 345                 350
Pro Lys Ser Gln Ser Ser Thr Thr Lys Asn Asn Ile Val Gly Gln Val
        355                 360                 365
Tyr Met Asn Gly Asp Val Ser Lys Pro Met Leu Leu Thr Ile Thr Leu
    370                 375                 380
Asn Gly Thr Asp Asp Ser Asn Ser Thr Tyr Ser Met Ser Phe Ser Tyr
385                 390                 395                 400
Thr Trp Thr Asn Gly Ser Tyr Val Gly Ala Thr Phe Gly Ala Asn Ser
                405                 410                 415
Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
            420                 425
```

<210> 54
<211> 578
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 11 Fiber

<400> 54

```
Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Leu Ala Ala Ala Val Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Ile
                165                 170                 175
Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
            180                 185                 190
Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
        195                 200                 205
Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
    210                 215                 220
Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Ser Gly Ala Leu Asn Tyr Asp Ser Ser Gly Asn Leu Glu Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Val Asp Ala Asn Gly Lys Leu Ile Leu Asp Val
            260                 265                 270
```

```
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
    290                 295                 300
Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305                 310                 315                 320
Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
                325                 330                 335
Ile Asn Pro Gly Asp Gly Leu Glu Phe Gly Ser Gly Ser Asp Thr Asn
            340                 345                 350
Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Ser Arg
        355                 360                 365
Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
    370                 375                 380
Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385                 390                 395                 400
Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
                405                 410                 415
Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
            420                 425                 430
Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
        435                 440                 445
Val Thr Ser Ala Gln Ile Ile Leu Arg Phe Asp Glu Asn Gly Val Leu
    450                 455                 460
Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465                 470                 475                 480
Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
                485                 490                 495
Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
            500                 505                 510
Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile
        515                 520                 525
Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
    530                 535                 540
Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545                 550                 555                 560
Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
                565                 570                 575
Gln Glu
```

<210> 55
<211> 442
<212> PRT
<213> Chimpansee Adenovirus- ChAd 16 Fiber

<400> 55

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
```

```
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Thr Met Gln Val Thr Ala Pro Leu Lys Leu Ala Asn Thr Ala Ile Leu
        115                 120                 125
Asn Thr Leu Ala Met Ala Tyr Gly Asn Gly Leu Gly Leu Ser Asn Asn
    130                 135                 140
Ala Leu Thr Val Gln Leu Gln Ser Pro Leu Thr Phe Asn Asn Ser Lys
145                 150                 155                 160
Val Ala Ile Asn Leu Gly Asn Gly Pro Leu Asn Val Thr Ser Asn Arg
                165                 170                 175
Leu Ser Ile Asn Cys Lys Arg Gly Val Tyr Val Thr Thr Thr Gly Asp
            180                 185                 190
Ala Ile Glu Thr Asn Ile Ser Trp Ser Asn Ala Ile Lys Phe Ile Gly
        195                 200                 205
Asn Ala Met Gly Val Asn Ile Asp Thr Asn Lys Gly Leu Gln Phe Gly
    210                 215                 220
Thr Thr Ser Thr Val Thr Asp Val Thr Asn Ala Phe Pro Ile Gln Val
225                 230                 235                 240
Lys Leu Gly Ala Gly Leu Ala Phe Asp Ser Thr Gly Ala Ile Val Ala
                245                 250                 255
Trp Asn Lys Glu Asp Asp Ser Leu Thr Leu Trp Thr Thr Pro Asp Pro
            260                 265                 270
Ser Pro Asn Cys Lys Ile Ala Ser Asp Lys Asp Ala Lys Leu Thr Leu
        275                 280                 285
Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser Leu Leu
    290                 295                 300
Ala Val Ser Gly Ser Leu Ala Pro Ile Thr Gly Ala Val Ser Thr Ala
305                 310                 315                 320
Leu Val Ser Leu Lys Phe Asp Ala Asn Gly Ala Leu Leu Glu Lys Ser
                325                 330                 335
Thr Leu Asn Arg Glu Tyr Trp Asn Tyr Arg Gln Gly Asp Leu Ile Pro
            340                 345                 350
Gly Thr Pro Tyr Thr His Ala Val Gly Phe Met Pro Asn Lys Lys Ala
        355                 360                 365
Tyr Pro Lys Asn Thr Thr Ala Ala Ser Lys Ser His Ile Val Gly Glu
    370                 375                 380
Val Tyr Leu Asp Gly Asp Ala Asp Lys Pro Leu Ser Leu Ile Ile Thr
385                 390                 395                 400
Phe Asn Glu Thr Asp Asp Glu Ser Cys Asp Tyr Cys Met Asn Phe Gln
                405                 410                 415
Trp Lys Trp Gly Ala Asp Gln Tyr Lys Asp Lys Thr Leu Ala Thr Ser
            420                 425                 430
Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 56
<211> 543
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 17 Fiber

<400> 56

```
Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Ile Thr Ser Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Ala Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
                165                 170                 175
Ser Ser Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
            180                 185                 190
Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
        195                 200                 205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
    210                 215                 220
Thr Gly Asn Gly Leu Thr Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
            260                 265                 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
    290                 295                 300
Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305                 310                 315                 320
Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
                325                 330                 335
Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
            340                 345                 350
Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
        355                 360                 365
Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
    370                 375                 380
Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385                 390                 395                 400
Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
                405                 410                 415
Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
            420                 425                 430
```

```
Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
        435                 440                 445
Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
    450                 455                 460
Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465                 470                 475                 480
Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
                485                 490                 495
Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
            500                 505                 510
Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
        515                 520                 525
Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
    530                 535                 540
```

<210> 57
<211> 543
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 19 Fiber

<400> 57

```
Met Lys Arg Thr Lys Thr Ser Asp Lys Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
                165                 170                 175
Ser Ala Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
            180                 185                 190
Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
        195                 200                 205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
    210                 215                 220
Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
            260                 265                 270
```

```
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
    290                 295                 300
Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305                 310                 315                 320
Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
                325                 330                 335
Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val
            340                 345                 350
Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
        355                 360                 365
Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
    370                 375                 380
Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385                 390                 395                 400
Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
                405                 410                 415
Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
            420                 425                 430
Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
        435                 440                 445
Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
    450                 455                 460
Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465                 470                 475                 480
Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
                485                 490                 495
Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
            500                 505                 510
Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
        515                 520                 525
Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
    530                 535                 540
```

<210> 58
<211> 963
<212> DNA
<213> Chimapnzee Adenovirus- ChAd 8 Fiber

<400> 58

```
atgaccaaac gagttcgact aagcagctcc ttcaatccgg tctaccccta tgaagatgaa 60
agcagctccc aacacccctt tataaaccct ggtttcattt cctcaaatgg atttacacaa 120
agcccagatg gggttcttac acttaaatgc ttatcgccgc tcaccaccac aggcggctcc 180
cttcaactta aagttggagg aggattatca gtggatgaca ctgacggttc attagaagaa 240
aacataagca ttacagcacc acttaataaa acaagtcact caataggttt atccatagga 300
gatgggttgg aaacaaaaaa caaccaacta tgtgctaagc tgggagacgg tcttacattt 360
aatacaggca gcatatgcat agatactgac attaatacat tatggacagg agcaacacca 420
gacgctaatt gcttagtcct tggaactgaa tctaatgatt gtaaacttac actggcactt 480
gtaaagtcag gagccttagt aaatgcttac gtagcacttg ttggagcctc agacgccgtt 540
aatgatttaa ccacagaaac aagtgctcaa ataattgcag acatatattt tgatgcgcaa 600
ggaaaacttc ttcctgattt atcagcactc aaaacagagc taaaacacaa atctggacaa 660
ggcacttcga cagcagatcc caataactgt aaaagcttta tgccaagtct aaatgcatat 720
ccactgcgcc ccaatggagg caacggaaac tatatttatg gaaccaccta ctacagggcc 780
agagatgaaa ccctttatga acttaaaacc tctgtaatgc ttaactacaa aattaccagt 840
```

ggactatgtg catatgccat gcattttcag tggtcttgga atagtgggac taaaccagaa 900
gacactcccg ccactttcat tgcctccccc tttgtctttt cctacattag agaagatgac 960
tga 963

<210> 59
<211> 320
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 8 Fiber

<400> 59

Met Thr Lys Arg Val Arg Leu Ser Ser Ser Phe Asn Pro Val Tyr Pro
1 5 10 15
Tyr Glu Asp Glu Ser Ser Ser Gln His Pro Phe Ile Asn Pro Gly Phe
20 25 30
Ile Ser Ser Asn Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu
35 40 45
Lys Cys Leu Ser Pro Leu Thr Thr Thr Gly Gly Ser Leu Gln Leu Lys
50 55 60
Val Gly Gly Gly Leu Ser Val Asp Asp Thr Asp Gly Ser Leu Glu Glu
65 70 75 80
Asn Ile Ser Ile Thr Ala Pro Leu Asn Lys Thr Ser His Ser Ile Gly
85 90 95
Leu Ser Ile Gly Asp Gly Leu Glu Thr Lys Asn Asn Gln Leu Cys Ala
100 105 110
Lys Leu Gly Asp Gly Leu Thr Phe Asn Thr Gly Ser Ile Cys Ile Asp
115 120 125
Thr Asp Ile Asn Thr Leu Trp Thr Gly Ala Thr Pro Asp Ala Asn Cys
130 135 140
Leu Val Leu Gly Thr Glu Ser Asn Asp Cys Lys Leu Thr Leu Ala Leu
145 150 155 160
Val Lys Ser Gly Ala Leu Val Asn Ala Tyr Val Ala Leu Val Gly Ala
165 170 175
Ser Asp Ala Val Asn Asp Leu Thr Thr Glu Thr Ser Ala Gln Ile Ile
180 185 190
Ala Asp Ile Tyr Phe Asp Ala Gln Gly Lys Leu Leu Pro Asp Leu Ser
195 200 205
Ala Leu Lys Thr Glu Leu Lys His Lys Ser Gly Gln Gly Thr Ser Thr
210 215 220
Ala Asp Pro Asn Asn Cys Lys Ser Phe Met Pro Ser Leu Asn Ala Tyr
225 230 235 240
Pro Leu Arg Pro Asn Gly Gly Asn Gly Asn Tyr Ile Tyr Gly Thr Thr
245 250 255
Tyr Tyr Arg Ala Arg Asp Glu Thr Leu Tyr Glu Leu Lys Thr Ser Val
260 265 270
Met Leu Asn Tyr Lys Ile Thr Ser Gly Leu Cys Ala Tyr Ala Met His
275 280 285
Phe Gln Trp Ser Trp Asn Ser Gly Thr Lys Pro Glu Asp Thr Pro Ala
290 295 300
Thr Phe Ile Ala Ser Pro Phe Val Phe Ser Tyr Ile Arg Glu Asp Asp
305 310 315 320

<210> 60
<211> 1062
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 22 Fiber

<400> 60

```
atggccaaac gagctcggct aagcagctcc ttcaatccgg tctaccccta tgaagatgaa 60
agcagctcac aacacccctt tataaaccct ggtttcattt cctcaaatgg ttttgcacaa 120
agcccagatg gagttctaac tcttaaatgt gttaatccgc tcactaccgc cagcggaccc 180
ctccaactta aagttggaag cagtcttaca gtagataata tcgatgggtc tttggaggaa 240
aatataactg ccgcagcgcc actcactaaa actaaccact ccataggttt atcaatagga 300
tctggcttgc aaacaaagga tgataaactt tgtttatcgc tgggagatgg gttggtaaca 360
aaggatgata aactatgttt atcgctggga gatgggttaa taacaaaaga tgatacacta 420
tgtgccaaac taggacatgg ccttgtgttt gactcttcca atgctatcac catagaaaac 480
aacaccttgt ggacaggtgc aaaaccaagc gccaactgtg taattaaaga gggagaagat 540
tccccagact gtaagctcac tttagttcta gtgaagaatg gaggactgat aaatggatac 600
ataacattaa tgggagcctc agaatatact aacaccttgt ttaaaaacaa acaagttaca 660
atcgatgtaa acctcgcatt tgataatact ggccaaatta tcacttacct atcatccctt 720
aaaagtaacc tgaactttaa agacaaccaa aacatggcta ctggaaccat aaccagtgcc 780
aaaggcttca tgcccagcac caccgcctat ccatttataa catacgccac tcagtcccta 840
aatgaagatt acatttatgg agagtgttac tacaaatcta ccaatggaac tctctttcca 900
ctaaaagtta ctgtcacact aaacagacgt atgtcagctt ctggaatggc ctatgctatg 960
aatttttcat ggtctctaaa tgcagaggaa gccccggaaa ctaccgaagt cactctcatt 1020
acctcccccc tctttttttc ttatatcaga gaagatgact ga                    1062
```

<210> 61
<211> 353
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 22 Fiber

<400> 61

```
Met Ala Lys Arg Ala Arg Leu Ser Ser Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Glu Asp Glu Ser Ser Ser Gln His Pro Phe Ile Asn Pro Gly Phe
            20                  25                  30
Ile Ser Ser Asn Gly Phe Ala Gln Ser Pro Asp Gly Val Leu Thr Leu
        35                  40                  45
Lys Cys Val Asn Pro Leu Thr Thr Ala Ser Gly Pro Leu Gln Leu Lys
    50                  55                  60
Val Gly Ser Ser Leu Thr Val Asp Asn Ile Asp Gly Ser Leu Glu Glu
65                  70                  75                  80
Asn Ile Thr Ala Ala Ala Pro Leu Thr Lys Thr Asn His Ser Ile Gly
                85                  90                  95
Leu Ser Ile Gly Ser Gly Leu Gln Thr Lys Asp Asp Lys Leu Cys Leu
            100                 105                 110
Ser Leu Gly Asp Gly Leu Val Thr Lys Asp Asp Lys Leu Cys Leu Ser
        115                 120                 125
Leu Gly Asp Gly Leu Ile Thr Lys Asp Asp Thr Leu Cys Ala Lys Leu
    130                 135                 140
Gly His Gly Leu Val Phe Asp Ser Ser Asn Ala Ile Thr Ile Glu Asn
145                 150                 155                 160
Asn Thr Leu Trp Thr Gly Ala Lys Pro Ser Ala Asn Cys Val Ile Lys
                165                 170                 175
Glu Gly Glu Asp Ser Pro Asp Cys Lys Leu Thr Leu Val Leu Val Lys
            180                 185                 190
Asn Gly Gly Leu Ile Asn Gly Tyr Ile Thr Leu Met Gly Ala Ser Glu
        195                 200                 205
Tyr Thr Asn Thr Leu Phe Lys Asn Lys Gln Val Thr Ile Asp Val Asn
    210                 215                 220
```

```
Leu Ala Phe Asp Asn Thr Gly Gln Ile Ile Thr Tyr Leu Ser Ser Leu
225                 230                 235                 240
Lys Ser Asn Leu Asn Phe Lys Asp Asn Gln Asn Met Ala Thr Gly Thr
                245                 250                 255
Ile Thr Ser Ala Lys Gly Phe Met Pro Ser Thr Thr Ala Tyr Pro Phe
            260                 265                 270
Ile Thr Tyr Ala Thr Gln Ser Leu Asn Glu Asp Tyr Ile Tyr Gly Glu
        275                 280                 285
Cys Tyr Tyr Lys Ser Thr Asn Gly Thr Leu Phe Pro Leu Lys Val Thr
    290                 295                 300
Val Thr Leu Asn Arg Arg Met Ser Ala Ser Gly Met Ala Tyr Ala Met
305                 310                 315                 320
Asn Phe Ser Trp Ser Leu Asn Ala Glu Glu Ala Pro Glu Thr Thr Glu
                325                 330                 335
Val Thr Leu Ile Thr Ser Pro Phe Phe Phe Ser Tyr Ile Arg Glu Asp
            340                 345                 350
Asp
```

<210> 62
<211> 1686
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 24 Fiber

<400> 62

```
atgtcagatt cttgctcctg tccttccgca cccactatct tcatgttgtt gcagatgaag 60
cgcaccaaaa cgtctgacga gagcttcaac cccgtgtacc cctatgacac ggaaaacggt 120
cctccctccg tccctttcct cacccctccc ttcgtgtctc ccgatggatt ccaagagagc 180
ccccccgggg tcctgtctct gaacctggcc gagcccctgg tcacttccca cggcatgctc 240
gccctgaaaa tgggaagtgg cctctccctg gacgacgccg gcaacctcac ctctcaagat 300
gtcaccacca ctacccctcc cctgaaaaaa accaagacca acctcagcct agaaacctca 360
gccccctga ctgtgagcac ctcaggcgcc ctcaccctag cggccgccgc tcccctggcg 420
gtggccggca cctccctcac catgcaatca gaggcccccc tgacagtaca ggatgcaaaa 480
ctcaccctgg ccaccaaggg ccccctgacc gtgtctgaag gcaaactggc cttgcagacc 540
tcggccccac tgacggccgc tgacagcagc accctcaccg ttagcgccac accacccatc 600
aatgtaagca gtggaagttt gggcttagac atggaaaatc ccatgtatac tcatgacgga 660
aaactgggaa taagaattgg gggcccactg agagtagtag acagcctgca cacactgact 720
gtagttaccg gaaatggaat agctgtagat aacaatgccc tccaaactag agttacgggc 780
gccctgggtt atgacacatc aggaaaccta caactgagag ccgcgggggg tatgcgaatt 840
gatgcaaatg gccaacttat ccttgatgtg gcatacccat ttgatgctca aaacaatctc 900
agccttagac ttggtcaggg acccctgtat gtaaacacag accacaacct agatttgaat 960
tgcaacagag gtctgaccac aactaccacc aacaacacaa aaaaacttga aactaaaatt 1020
ggctcaggct tagactatga taccaatggt gctgtcatta ttaaacttgg cactggtgtc 1080
agctttgaca gcacaggcgc cctaagtgtg ggaaacactg gcgatgataa actgactctg 1140
tggacaaccc cagacccatc tccaaattgc agaattcact cagacaaaga ctgcaagttt 1200
actctagtcc taactaagtg tggaagtcaa atcctggctt ctgtcgccgc cctagcggtg 1260
tcaggaaatc tggcttcaat aacaggcacc gtttccagcg ttaccatctt tctcagattt 1320
gatcagaatg gagtgcttat ggaaaactcc tcgctagaca agcagtactg gaactttaga 1380
aatggtaatt caaccaatgc cacccccctac accaatgcag ttggtttcat gccaaacctc 1440
gcagcatacc ccaagacaca gagtcagact gctaaaaaca acattgtaag tcaggtttac 1500
ttgaatgggg acaaatccaa acccatgatc cttaccatta ccctcaatgg aactaatgaa 1560
tccagtgaaa ctagccaggt gagtcactac tccatgtcat ttacgtgggc ttgggagagt 1620
gggcaatatg ccaccgaaac ctttgccacc aattccttta ccttctctta cattgctgaa 1680
caataa                                                          1686
```

<210> 63
<211> 543
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 24 Fiber

<400> 63

```
Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
                165                 170                 175
Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp
            180                 185                 190
Met Glu Asn Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
        195                 200                 205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
    210                 215                 220
Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
            260                 265                 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
    290                 295                 300
Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305                 310                 315                 320
Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
                325                 330                 335
Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val
            340                 345                 350
Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
        355                 360                 365
Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
    370                 375                 380
Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385                 390                 395                 400
```

```
Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
                405                 410                 415
Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
            420                 425                 430
Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
        435                 440                 445
Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
    450                 455                 460
Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465                 470                 475                 480
Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr
                485                 490                 495
Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
            500                 505                 510
Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
        515                 520                 525
Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
    530                 535                 540
```

<210> 64
<211> 1335
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 26 Fiber

<400> 64

```
atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgct gtccctgcgt ctggccgatc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agatggggtg gacctcgacg actcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggataccc ctctttacaa caacaatgga aagctaggta tgaaggtaac cgcaccatta 360
aagatattag acacagatct actaaaaaca cttgttgttg cttatgggca gggattagga 420
acaaacacca atggtgctct tgttgcccaa ctagcatacc cacttgtttt taataccgct 480
agcaaaattg cccttaattt aggcaatgga ccattaaaag tggatgcaaa tagactgaac 540
attaattgca aaagaggtat ctatgtcact accacaaaag atgcactgga gattaatatc 600
agttgggcaa atgctatgac atttatagga aatgccattg gtgtcaatat tgacacaaaa 660
aaaggcctac agttcggcac ttcaagcact gaaacagatg ttaaaaatgc ttttccactc 720
caagtaaaac ttggagctgg tcttacattt gacagcacag gtgccattgt tgcttggaac 780
aaagaagatg acaaacttac actgtggacc acagccgatc catctccaaa ctgtcacata 840
tattctgcaa aggatgctaa gcttacactc tgcttgacaa agtgtggtag tcaaatccta 900
ggcactgtct ccctattagc agtcagtggc agcttggctc ctatcacagg ggctgttaga 960
actgcacttg tatcactcaa attcaatgct aatggagccc ttttggacaa atcaactctg 1020
aacaaagaat actggaacta cagacaagga gatctaattc caggtacacc atatacacat 1080
gctgtgggtt tcatgcctaa caaaaaagcc taccctaaaa acacaactgc agcttccaag 1140
agccacattg tgggtgatgt gtatttagat ggagatgcag ataaaccttt atctcttatc 1200
atcactttca atgaaactga tgatgaaacc tgtgattact gcatcaactt tcaatggaaa 1260
tggggagctg atcaatataa ggataagaca ctcgcaacca gttcattcac cttctcatac 1320
atcgcccaag aataa                                                  1335
```

<210> 65
<211> 444
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 26 Fiber

<400> 65

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Asp Gly Val Asp Leu Asp Asp Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Leu Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Asn
    130                 135                 140
Gly Ala Leu Val Ala Gln Leu Ala Tyr Pro Leu Val Phe Asn Thr Ala
145                 150                 155                 160
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Lys Arg Gly Ile Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Ile Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Ile Gly Val Asn Ile Asp Thr Lys Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Ser Ser Thr Glu Thr Asp Val Lys Asn Ala Phe Pro Leu
225                 230                 235                 240
Gln Val Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Ala Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Leu Ala Val Ser Gly Ser Leu Ala Pro Ile Thr Gly Ala Val Arg
305                 310                 315                 320
Thr Ala Leu Val Ser Leu Lys Phe Asn Ala Asn Gly Ala Leu Leu Asp
                325                 330                 335
Lys Ser Thr Leu Asn Lys Glu Tyr Trp Asn Tyr Arg Gln Gly Asp Leu
            340                 345                 350
Ile Pro Gly Thr Pro Tyr Thr His Ala Val Gly Phe Met Pro Asn Lys
        355                 360                 365
Lys Ala Tyr Pro Lys Asn Thr Thr Ala Ala Ser Lys Ser His Ile Val
    370                 375                 380
Gly Asp Val Tyr Leu Asp Gly Asp Ala Asp Lys Pro Leu Ser Leu Ile
385                 390                 395                 400
Ile Thr Phe Asn Glu Thr Asp Asp Glu Thr Cys Asp Tyr Cys Ile Asn
                405                 410                 415
Phe Gln Trp Lys Trp Gly Ala Asp Gln Tyr Lys Asp Lys Thr Leu Ala
            420                 425                 430

        Thr Ser Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
                435                 440
```

<210> 66

<211> 1062
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 30 Fiber

<400> 66

```
atggccaaac gagctcggct aagcagctcc ttcaatccgg tctaccccta tgaagatgaa 60
agcagctcac aacacccctt tataaaccct ggtttcattt cctcaaatgg ttttgcacaa 120
agcccagatg gagttctaac tcttaaatgt gttaatccgc tcactaccgc cagcggaccc 180
ctccaactta aagttggaag cagtcttaca gtagatacta tcgatgggtc tttggaggaa 240
aatataactg ccgcagcgcc actcactaaa actaaccact ccataggttt atcaatagga 300
tctggcttgc aaacaaagga tgataaactt tgtttatcgc tgggagatgg gttggtaaca 360
aaggatgata aactatgttt atcgctggga gatgggttaa taacaaaaga tgatacacta 420
tgtgccaaac taggacatgg ccttgtgttt gactcttcca atgctatcac catagaaaac 480
aacaccttgt ggacaggtgc aaaaccaagc gccaactgtg taattaaaga gggagaagat 540
tccccagact gtaagctcac tttagttcta gtgaagaatg gaggactgat aaatggatac 600
ataacattaa tgggagcctc agaatatact aacaccttgt ttaaaaacaa acaagttaca 660
atcgatgtaa acctcgcatt tgataatact ggccaaatta tcacttacct atcatccctt 720
aaaagtaacc tgaactttaa agacaaccaa aacatggcta ctggaaccat aaccagtgcc 780
aaaggcttca tgcccagcac caccgcctat ccatttataa catacgccac tcagtcccta 840
aatgaagatt acatttatgg agagtgttac tacaaatcta ccaatggaac tctctttcca 900
ctaaaagtta ctgtcacact aaacagacgt atgtcagctt ctggaatggc ctatgctatg 960
aatttttcat ggtctctaaa tgcagaggaa gccccggaaa ctaccgaagt cactctcatt 1020
acctcccccc tctttttttc ttatatcaga gaagatgact ga . .                 1062
```

<210> 67
<211> 353
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 30 Fiber

<400> 67

```
Met Ala Lys Arg Ala Arg Leu Ser Ser Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Glu Asp Glu Ser Ser Ser Gln His Pro Phe Ile Asn Pro Gly Phe
            20                  25                  30
Ile Ser Ser Asn Gly Phe Ala Gln Ser Pro Asp Gly Val Leu Thr Leu
        35                  40                  45
Lys Cys Val Asn Pro Leu Thr Thr Ala Ser Gly Pro Leu Gln Leu Lys
    50                  55                  60
Val Gly Ser Ser Leu Thr Val Asp Thr Ile Asp Gly Ser Leu Glu Glu
65                  70                  75                  80
Asn Ile Thr Ala Ala Ala Pro Leu Thr Lys Thr Asn His Ser Ile Gly
                85                  90                  95
Leu Ser Ile Gly Ser Gly Leu Gln Thr Lys Asp Asp Lys Leu Cys Leu
            100                 105                 110
Ser Leu Gly Asp Gly Leu Val Thr Lys Asp Asp Lys Leu Cys Leu Ser
        115                 120                 125
Leu Gly Asp Gly Leu Ile Thr Lys Asp Asp Thr Leu Cys Ala Lys Leu
    130                 135                 140
Gly His Gly Leu Val Phe Asp Ser Ser Asn Ala Ile Thr Ile Glu Asn
145                 150                 155                 160
```

```
Asn Thr Leu Trp Thr Gly Ala Lys Pro Ser Ala Asn Cys Val Ile Lys
                165                 170                 175
Glu Gly Glu Asp Ser Pro Asp Cys Lys Leu Thr Leu Val Leu Val Lys
            180                 185                 190
Asn Gly Gly Leu Ile Asn Gly Tyr Ile Thr Leu Met Gly Ala Ser Glu
        195                 200                 205
Tyr Thr Asn Thr Leu Phe Lys Asn Lys Gln Val Thr Ile Asp Val Asn
    210                 215                 220
Leu Ala Phe Asp Asn Thr Gly Gln Ile Ile Thr Tyr Leu Ser Ser Leu
225                 230                 235                 240
Lys Ser Asn Leu Asn Phe Lys Asp Asn Gln Asn Met Ala Thr Gly Thr
                245                 250                 255
Ile Thr Ser Ala Lys Gly Phe Met Pro Ser Thr Thr Ala Tyr Pro Phe
            260                 265                 270
Ile Thr Tyr Ala Thr Gln Ser Leu Asn Glu Asp Tyr Ile Tyr Gly Glu
        275                 280                 285
Cys Tyr Tyr Lys Ser Thr Asn Gly Thr Leu Phe Pro Leu Lys Val Thr
    290                 295                 300
Val Thr Leu Asn Arg Arg Met Ser Ala Ser Gly Met Ala Tyr Ala Met
305                 310                 315                 320
Asn Phe Ser Trp Ser Leu Asn Ala Glu Glu Ala Pro Glu Thr Thr Glu
                325                 330                 335
Val Thr Leu Ile Thr Ser Pro Phe Phe Phe Ser Tyr Ile Arg Glu Asp
            340                 345                 350
Asp
```

<210> 68
<211> 1791
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 31 Fiber

<400> 68

```
atgtcagatt cttgctcctg tccctccgca cccactatct tcatgttgtt gcagatgaag 60
cgcaccaaaa cgtctgacga gagcttcaac cccgtgtacc cctatgacac ggaaagcggc 120
cctccctccg tccctttcct cacccctccc ttcgtgtctc ccgatggatt ccaagaaagt 180
ccccccgggg tcctgtctct gaacctggcc gagcccctgg tcacttccca cggcatgctc 240
gccctgaaaa tgggaagtgg cctctccctg gacgacgctg gcaacctcac ctctcaagat 300
atcaccaccg ctagccctcc cctcaaaaaa accaagacca acctcagcct agaaacctca 360
tccccctaa ctgtgagcac ctcaggcgcc ctcaccgtag cagccgccgc tcccctggcg 420
gtggccggca cctccctcac catgcaatca gaggcccccc tgacagtaca ggatgcaaaa 480
ctcaccctgg ccaccaaagg ccccctgacc gtgtctgaag gcaaactggc cttgcaaaca 540
tcggccccgc tgacggccgc tgacagcagc accctcacag tcagtgccac accacccctt 600
agcacaagca atggcagctt gggtattgac atgcaagccc ccatttacac caccaatgga 660
aaactaggac ttaactttgg cgctcccctg catgtggtag acagcctaaa tgcactgact 720
gtagttactg gccaaggtct tacgataaac ggaacagccc tacaaactag agtctcaggt 780
gccctcaact atgacacatc aggaaaccta gaattgagag ctgcaggggg tatgcgagtt 840
gatgcaaatg gtcaacttat ccttgatgta gcttacccat ttgatgcaca aaacaatctc 900
agccttaggc ttggacaggg acccctgttt gttaactctg cccacaactt ggatgttaac 960
tacaacagag gcctctacct gttcacatct ggaaatacca aaaagctaga agttaatatc 1020
aaaacagcca agggtctcat ttatgatgac actgctatag caatcaatgc gggtgatggg 1080
ctacagtttg actcaggctc agatacaaat ccattaaaaa ctaaacttgg attaggactg 1140
gattatgact ccagcagagc cataattgct aaactgggaa ctggcctaag ctttgacaac 1200
acaggtgcca tcacagtagg caacaaaaat gatgacaagc ttaccttgtg gaccacacca 1260
gacccatccc ctaactgtag aatctattca gagaaagatg ctaaattcac acttgttttg 1320
actaaatgcg gcagtcaggt gttggccagc gtttctgttt tatctgtaaa aggtagcctt 1380
```

gcgcccatca gtggcacagt aactagtgct cagattgtcc tcagatttga tgaaaatgga 1440
gttctactaa gcaattcttc ccttgaccct caatactgga actacagaaa aggtgacctt 1500
acagagggca ctgcatatac caacgcagtg ggatttatgc ccaacctcac agcataccca 1560
aaaacacaga gccaaactgc taaaagcaac attgtaagtc aggtttactt gaatggggac 1620
aaatccaaac ccatgaccct caccattacc ctcaatggaa ctaatgaaac aggagatgcc 1680
acagtaagca cttactccat gtcattctca tggaactgga atggaagtaa ttacattaat 1740
gaaacgttcc aaaccaactc cttcaccttc tcctacatcg cccaagaata a 1791

<210> 69
<211> 578
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 31 Fiber

<400> 69

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15
Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80
Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95
Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110
Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145 150 155 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175
Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
180 185 190
Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
195 200 205
Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
210 215 220
Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225 230 235 240
Ser Gly Ala Leu Asn Tyr Asp Thr Ser Gly Asn Leu Glu Leu Arg Ala
245 250 255
Ala Gly Gly Met Arg Val Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285
Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300
Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320

```
Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
                325                 330                 335
Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
            340                 345                 350
Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
        355                 360                 365
Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
    370                 375                 380
Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385                 390                 395                 400
Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
                405                 410                 415
Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
            420                 425                 430
Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
        435                 440                 445
Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
    450                 455                 460
Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465                 470                 475                 480
Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
                485                 490                 495
Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
            500                 505                 510
Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
        515                 520                 525
Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
    530                 535                 540
Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545                 550                 555                 560
Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
                565                 570                 575
Gln Glu
```

<210> 70
<211> 978
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 37 Fiber

<400> 70

```
atggccaaac gggctcgtct aagcagctcc ttcaacccgg tgtaccccta tgaagacgag 60
agcagctcac aacacccatt tataaacccc ggcttcattt cccctgatgg ctttacacaa 120
agcccagacg gagttctaac actgaaatgt gtttcccctc ttactaccac cagtggcgct 180
ctagacatta aagtgggaag agggcttaaa gtagatagca ctgatggttc cctggaagaa 240
aatatagaca ttacagctcc cctcactaaa tttaaccact cagtaggatt agcatttggc 300
gacggtctag aaacaaaaga aaacaagctt tatgtaaaac ttggagatgg acttaaattt 360
agctctggca gtatatacat tgaccatgat gttaacactt tatggacagg agtcaatcca 420
agtgctaact gtataattac agacaatgga gaaaccaatg acagcaagct taccctaata 480
cttgttaagt caggtggatt aataaatgct tatgtctcat taatgggtga ctcagacaca 540
gtcaataaat taaccacaga aaaaagtgct caaattaccg ttgacatata ctttgataat 600
caaggaaaag ttcttactga actatcggcc cttaaaacag atcttaaaca taaatttggt 660
caaaacatgg cttctagcga agtatcaaac tgcaaaggct ttatgccaag cttaaatgca 720
tacccattca gaaatccaac taaacctacc aaaggaagag aagactacat ttatggaata 780
acttactatc aagccacaga tggtaatctc tatgagctaa aaactactat tactctaaac 840
```

cacagtgtca ttagttctct atgtgcatat gcaatgcaca tttcatggtc atgggacacc 900
gtaacagagc cagagacaac acccactact cttattacct ccccttctc cttttcctat 960
atcagagaag atgactga 978

<210> 71
<211> 325
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 37 Fiber

<400> 71

Met Ala Lys Arg Ala Arg Leu Ser Ser Ser Phe Asn Pro Val Tyr Pro
1 5 10 15
Tyr Glu Asp Glu Ser Ser Ser Gln His Pro Phe Ile Asn Pro Gly Phe
20 25 30
Ile Ser Pro Asp Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu
35 40 45
Lys Cys Val Ser Pro Leu Thr Thr Thr Ser Gly Ala Leu Asp Ile Lys
50 55 60
Val Gly Arg Gly Leu Lys Val Asp Ser Thr Asp Gly Ser Leu Glu Glu
65 70 75 80
Asn Ile Asp Ile Thr Ala Pro Leu Thr Lys Phe Asn His Ser Val Gly
85 90 95
Leu Ala Phe Gly Asp Gly Leu Glu Thr Lys Glu Asn Lys Leu Tyr Val
100 105 110
Lys Leu Gly Asp Gly Leu Lys Phe Ser Ser Gly Ser Ile Tyr Ile Asp
115 120 125
His Asp Val Asn Thr Leu Trp Thr Gly Val Asn Pro Ser Ala Asn Cys
130 135 140
Ile Ile Thr Asp Asn Gly Glu Thr Asn Asp Ser Lys Leu Thr Leu Ile
145 150 155 160
Leu Val Lys Ser Gly Gly Leu Ile Asn Ala Tyr Val Ser Leu Met Gly
165 170 175
Asp Ser Asp Thr Val Asn Lys Leu Thr Thr Glu Lys Ser Ala Gln Ile
180 185 190
Thr Val Asp Ile Tyr Phe Asp Asn Gln Gly Lys Val Leu Thr Glu Leu
195 200 205
Ser Ala Leu Lys Thr Asp Leu Lys His Lys Phe Gly Gln Asn Met Ala
210 215 220
Ser Ser Glu Val Ser Asn Cys Lys Gly Phe Met Pro Ser Leu Asn Ala
225 230 235 240
Tyr Pro Phe Arg Asn Pro Thr Lys Pro Thr Lys Gly Arg Glu Asp Tyr
245 250 255
Ile Tyr Gly Ile Thr Tyr Tyr Gln Ala Thr Asp Gly Asn Leu Tyr Glu
260 265 270
Leu Lys Thr Thr Ile Thr Leu Asn His Ser Val Ile Ser Ser Leu Cys
275 280 285
Ala Tyr Ala Met His Ile Ser Trp Ser Trp Asp Thr Val Thr Glu Pro
290 295 300
Glu Thr Thr Pro Thr Thr Leu Ile Thr Ser Pro Phe Ser Phe Ser Tyr
305 310 315 320
Ile Arg Glu Asp Asp
325

<210> 72
<211> 1332
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 38 Fiber

<400> 72

```
atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgtt gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgact cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtttttcca acaacaccat ttcccttaac 300
atggataccc ctttttatac caaagatgga aaattatcct tacaagtttc tccaccatta 360
aacatattaa aatcaaccat tctgaacaca ttagctgtag cttatggatc aggtttagga 420
ctcagtggtg gcactgctct tgcagtacag ttggcctctc cactcacctt tgatgaaaaa 480
ggaaatatta aaattaacct agccagtggt ccattaacag ttgatgcaag tcgacttagt 540
atcaactgca aaagaggggt cactgtcact accgcaggag atgcaattaa aagcaacata 600
agctggccta aaggtataag atttgaaggt gatgccatag ctgcaaacat tggcagagga 660
ttggaatttg gaaccactag tacagagact gatgtcacag atgcataccc aattcaagtt 720
aaattgggta ctggtctcac ctttgacagt acaggcgcca ttgttgcatg gaacaaagag 780
gatgataaac ttacattatg gaccacagcc gacccctcgc caaattgcaa aatatactct 840
gaaaaagatg ctaaactcac actttgcttg acaaaatgtg gaagccaaat tctgggcact 900
gtgactgtat tggcagtgaa taatggaagt ctcaacccaa tcacaaacac agtaagcact 960
gcacttgtct ccctcaagtt tgatgcaagt ggagttttgc taagcagctc cacattagac 1020
aaagaatatt ggaacttccg aaagggagat gttacacctg ctgaacccta tactaatgct 1080
ataggtttta tgcctaacat aaaggcctat cctaaaaaca catctgcagc ttcaaaaagc 1140
catattgtca gtcaagttta tctcaatggg gatgaaacca aacctctgat gctgattatt 1200
acttttaatg aaactgagga tgcaacttgc acctatagta tcacttttca atggaaatgg 1260
gatagtacta agtacacagg taaaacactt gctaccagct ccttcacctt ctcctacatt 1320
gctcaagaat ga                                                    1332
```

<210> 73
<211> 443
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 38 Fiber

<400> 73

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Lys Ser Thr Ile Leu
        115                 120                 125
Asn Thr Leu Ala Val Ala Tyr Gly Ser Gly Leu Gly Leu Ser Gly Gly
    130                 135                 140
Thr Ala Leu Ala Val Gln Leu Ala Ser Pro Leu Thr Phe Asp Glu Lys
145                 150                 155                 160
```

```
Gly Asn Ile Lys Ile Asn Leu Ala Ser Gly Pro Leu Thr Val Asp Ala
                165                 170                 175
Ser Arg Leu Ser Ile Asn Cys Lys Arg Gly Val Thr Val Thr Thr Ala
            180                 185                 190
Gly Asp Ala Ile Lys Ser Asn Ile Ser Trp Pro Lys Gly Ile Arg Phe
        195                 200                 205
Glu Gly Asp Ala Ile Ala Ala Asn Ile Gly Arg Gly Leu Glu Phe Gly
    210                 215                 220
Thr Thr Ser Thr Glu Thr Asp Val Thr Asp Ala Tyr Pro Ile Gln Val
225                 230                 235                 240
Lys Leu Gly Thr Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile Val Ala
                245                 250                 255
Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala Asp Pro
            260                 265                 270
Ser Pro Asn Cys Lys Ile Tyr Ser Glu Lys Asp Ala Lys Leu Thr Leu
        275                 280                 285
Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Thr Val Leu
    290                 295                 300
Ala Val Asn Asn Gly Ser Leu Asn Pro Ile Thr Asn Thr Val Ser Thr
305                 310                 315                 320
Ala Leu Val Ser Leu Lys Phe Asp Ala Ser Gly Val Leu Leu Ser Ser
                325                 330                 335
Ser Thr Leu Asp Lys Glu Tyr Trp Asn Phe Arg Lys Gly Asp Val Thr
            340                 345                 350
Pro Ala Glu Pro Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn Ile Lys
        355                 360                 365
Ala Tyr Pro Lys Asn Thr Ser Ala Ala Ser Lys Ser His Ile Val Ser
    370                 375                 380
Gln Val Tyr Leu Asn Gly Asp Glu Thr Lys Pro Leu Met Leu Ile Ile
385                 390                 395                 400
Thr Phe Asn Glu Thr Glu Asp Ala Thr Cys Thr Tyr Ser Ile Thr Phe
                405                 410                 415
Gln Trp Lys Trp Asp Ser Thr Lys Tyr Thr Gly Lys Thr Leu Ala Thr
            420                 425                 430
Ser Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 74
<211> 1332
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 44 Fiber

<400> 74

```
atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgtt gtccctgcga ctggctgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgact cgtcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agtatttcaa acaacaccat ttcccttaaa 300
actgctgccc ctttctacaa caacaatgga actttaagcc tcaatgtctc cacaccatta 360
gcagtatttc ccacatttaa cactttaggc ataagtcttg gaaacggtct tcagacttca 420
aataagttgt tgactgtaca actaactcat cctcttacat tcagctcaaa tagcatcaca 480
gtaaaaacag acaaagggct atatattaac tccagtggaa acagaggact tgaggctaat 540
ataagcctaa aaagaggact agtttttgac ggtaatgcta ttgcaacata tattggaaat 600
ggcttagact atggatctta tgatagtgat ggaaaaacaa gacccgtaat taccaaaatt 660
ggagcaggat taaattttga tgctaacaaa gcaatagctg tcaaactagg cacaggttta 720
```

```
agttttgact ccgctggtgc cttgacagct ggaaacaaac aggatgacaa gctaacactt 780
tggactaccc ctgacccaag ccctaattgt caattacttt cagacagaga tgccaaattt 840
actctctgtc ttacaaaatg cggtagtcaa atactaggca ctgtggcagt ggcggctgtt 900
actgtaggat cagcactaaa tccaattaat gacacagtca aaagcgccat agttttcctt 960
agatttgatt ccgatggtgt actcatgtca aactcatcaa tggtaggtga ttactggaac 1020
tttagggagg gacagaccac tcaaagtgta gcctatacaa atgctgtggg attcatgcca 1080
aatataggtg catatccaaa aacccaaagt aaaacaccta aaaatagcat agtcagtcag 1140
gtatatttaa ctggagaaac tactatgcca atgacactaa ccataacttt caatggcact 1200
gatgaaaaag acacaacccc agttagcacc tactctatga cttttacatg gcagtggact 1260
ggagactata aggacaaaaa tattaccttt gctaccaact cattctcttt ttcctacatc 1320
gcccaggaat aa                                                     1332
```

<210> 75
<211> 443
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 44 Fiber

<400> 75

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Ile Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Lys Thr Ala Ala Pro Phe Tyr Asn Asn Asn Gly Thr Leu
            100                 105                 110
Ser Leu Asn Val Ser Thr Pro Leu Ala Val Phe Pro Thr Phe Asn Thr
        115                 120                 125
Leu Gly Ile Ser Leu Gly Asn Gly Leu Gln Thr Ser Asn Lys Leu Leu
    130                 135                 140
Thr Val Gln Leu Thr His Pro Leu Thr Phe Ser Ser Asn Ser Ile Thr
145                 150                 155                 160
Val Lys Thr Asp Lys Gly Leu Tyr Ile Asn Ser Ser Gly Asn Arg Gly
                165                 170                 175
Leu Glu Ala Asn Ile Ser Leu Lys Arg Gly Leu Val Phe Asp Gly Asn
            180                 185                 190
Ala Ile Ala Thr Tyr Ile Gly Asn Gly Leu Asp Tyr Gly Ser Tyr Asp
        195                 200                 205
Ser Asp Gly Lys Thr Arg Pro Val Ile Thr Lys Ile Gly Ala Gly Leu
    210                 215                 220
Asn Phe Asp Ala Asn Lys Ala Ile Ala Val Lys Leu Gly Thr Gly Leu
225                 230                 235                 240
Ser Phe Asp Ser Ala Gly Ala Leu Thr Ala Gly Asn Lys Gln Asp Asp
                245                 250                 255
Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro Asn Cys Gln Leu
            260                 265                 270
Leu Ser Asp Arg Asp Ala Lys Phe Thr Leu Cys Leu Thr Lys Cys Gly
        275                 280                 285
```

```
Ser Gln Ile Leu Gly Thr Val Ala Val Ala Ala Val Thr Val Gly Ser
    290                 295                 300
Ala Leu Asn Pro Ile Asn Asp Thr Val Lys Ser Ala Ile Val Phe Leu
305                 310                 315                 320
Arg Phe Asp Ser Asp Gly Val Leu Met Ser Asn Ser Ser Met Val Gly
                325                 330                 335
Asp Tyr Trp Asn Phe Arg Glu Gly Gln Thr Thr Gln Ser Val Ala Tyr
            340                 345                 350
Thr Asn Ala Val Gly Phe Met Pro Asn Ile Gly Ala Tyr Pro Lys Thr
        355                 360                 365
Gln Ser Lys Thr Pro Lys Asn Ser Ile Val Ser Gln Val Tyr Leu Thr
    370                 375                 380
Gly Glu Thr Thr Met Pro Met Thr Leu Thr Ile Thr Phe Asn Gly Thr
385                 390                 395                 400
Asp Glu Lys Asp Thr Thr Pro Val Ser Thr Tyr Ser Met Thr Phe Thr
                405                 410                 415
Trp Gln Trp Thr Gly Asp Tyr Lys Asp Lys Asn Ile Thr Phe Ala Thr
            420                 425                 430
Asn Ser Phe Ser Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 76
<211> 1278
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 63 Fiber

<400> 76

```
atgtccaaaa agcgcgtccg ggtggatgat gacttcgacc ccgtctaccc ctacgatgca 60
gacaacgcac cgaccgtgcc cttcatcaac ccccccttcg tctcttcaga tggattccaa 120
gagaagcccc tgggggtgct gtccctgcga ctggccgacc ccgtcaccac caagaacggg 180
gaaatcaccc tcaagctggg agagggggtg gacctcgact cctcgggaaa actcatctcc 240
aacacggcca ccaaggccgc cgcccctctc agttttttcca acaacaccat ttcccttaac 300
atggatcacc ccttttacac taaagatgga aaattatcct tacaagtttc tccaccatta 360
aatatactga gaacaagcat tctaaacaca ctagctttag gttttggatc aggtttagga 420
ctccgtggct ctgccttggc agtacagtta gtctctccac ttacatttga tactgatgga 480
aacataaagc ttaccttaga cagaggtttg catgttacaa caggagatgc aattgaaagc 540
aacataagct gggctaaagg tttaaaattt gaagatggag ccatagcaac caacattgga 600
aatgggttag agtttggaag cagtagtaca gaaacaggtg ttgatgatgc ttacccaatc 660
caagttaaac ttggatctgg ccttagcttt gacagtacag gagccataat ggctggtaac 720
aaagaagacg ataaactcac tttgtggaca acacctgatc catcgccaaa ctgtcaaata 780
ctcgcagaaa atgatgcaaa actaacactt tgcttgacta aatgtggtag tcaaatactg 840
gccactgtgt cagtcttagt tgtaggaagt ggaaacctaa accccattac tggcaccgta 900
agcagtgctc aggtgtttct acgttttgat gcaaacggtg ttcttttaac agaacattct 960
acactaaaaa aatactgggg gtataggcag ggagatagca tagatggcac tccatatacc 1020
aatgctgtag gattcatgcc caatttaaaa gcttatccaa agtcacaaag ttctactact 1080
aaaaataata tagtagggca agtatacatg aatggagatg tttcaaaacc tatgcttctc 1140
actataaccc tcaatggtac tgatgacagc aacagtacat attcaatgtc attttcatac 1200
acctggacta atggaagcta tgttggagca acatttgggg ctaactctta taccttctca 1260
tacatcgccc aagaatga                                                1278
```

<210> 77
<211> 425
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 63 Fiber

<400> 77

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp His Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Arg Thr Ser Ile Leu
        115                 120                 125
Asn Thr Leu Ala Leu Gly Phe Gly Ser Gly Leu Gly Leu Arg Gly Ser
    130                 135                 140
Ala Leu Ala Val Gln Leu Val Ser Pro Leu Thr Phe Asp Thr Asp Gly
145                 150                 155                 160
Asn Ile Lys Leu Thr Leu Asp Arg Gly Leu His Val Thr Thr Gly Asp
                165                 170                 175
Ala Ile Glu Ser Asn Ile Ser Trp Ala Lys Gly Leu Lys Phe Glu Asp
            180                 185                 190
Gly Ala Ile Ala Thr Asn Ile Gly Asn Gly Leu Glu Phe Gly Ser Ser
        195                 200                 205
Ser Thr Glu Thr Gly Val Asp Asp Ala Tyr Pro Ile Gln Val Lys Leu
    210                 215                 220
Gly Ser Gly Leu Ser Phe Asp Ser Thr Gly Ala Ile Met Ala Gly Asn
225                 230                 235                 240
Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro
                245                 250                 255
Asn Cys Gln Ile Leu Ala Glu Asn Asp Ala Lys Leu Thr Leu Cys Leu
            260                 265                 270
Thr Lys Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Val Val
        275                 280                 285
Gly Ser Gly Asn Leu Asn Pro Ile Thr Gly Thr Val Ser Ser Ala Gln
    290                 295                 300
Val Phe Leu Arg Phe Asp Ala Asn Gly Val Leu Leu Thr Glu His Ser
305                 310                 315                 320
Thr Leu Lys Lys Tyr Trp Gly Tyr Arg Gln Gly Asp Ser Ile Asp Gly
                325                 330                 335
Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Lys Ala Tyr
            340                 345                 350
Pro Lys Ser Gln Ser Ser Thr Thr Lys Asn Asn Ile Val Gly Gln Val
        355                 360                 365
Tyr Met Asn Gly Asp Val Ser Lys Pro Met Leu Leu Thr Ile Thr Leu
    370                 375                 380
Asn Gly Thr Asp Asp Ser Asn Ser Thr Tyr Ser Met Ser Phe Ser Tyr
385                 390                 395                 400
Thr Trp Thr Asn Gly Ser Tyr Val Gly Ala Thr Phe Gly Ala Asn Ser
                405                 410                 415
Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
            420                 425
```

<210> 78
<211> 1338
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 82 Fiber

<400> 78

```
atgtccaaaa agcgcgcgcg ggtggatgat gacttcgacc ccgtgtaccc ctacgatgca 60
gacaacgcac cgactgtgcc cttcatcaac cctcccttcg tctcttcaga tggattccaa 120
gaaaagcccc tgggggtgtt gtccctgcga ctggccgatc ccgtcaccac caagaacggg 180
gctgtcaccc tcaagctggg ggaggggggtg gacctcgacg actcgggaaa actcatctcc 240
aaaaatgcca ccaaggccac tgcccctctc agtatttcca acaacaccat ttcccttaac 300
atggataccc ctctttacaa caacaatgga aagctaggta tgaaggtaac cgcaccatta 360
aagatattag acacagatct actaaaaaca cttgttgttg cttatgggca gggattagga 420
acaaacacca atggtgctct tgttgcccaa ctagcatacc cacttgtttt taataccgct 480
agcaaaattg cccttaattt aggcaatgga ccattaaaag tggatgcaaa tagactgaac 540
attaattgca aaagaggtat ctatgtcact accacaaaag atgcactgga gattaatatc 600
agttgggcaa atgctatgac atttatagga aatgccattg gtgtcaatat tgacacaaaa 660
aaaggcctac agttcggcac ttcaagcact gaaacagatg ttaaaaatgc ttttccactc 720
caagtaaaac ttggagctgg tcttacattt gacagcacag gtgccattgt tgcttggaac 780
aaagaagatg acaaacttac actgtggacc acagccgatc catctccaaa ctgtcacata 840
tattctgcaa aggatgctaa gcttacactc tgcttgacaa agtgtggtag tcagatactg 900
ggcactgttt ctctcatagc tgttgatact ggtagcttaa atccaataac aggaaaagta 960
accactgctc ttgtttcact taaattcgat gccaatggag ttttgcaagc cagttcaaca 1020
ctagataaag aatattggaa tttcagaaaa ggagatgtga cacctgctga cccctacact 1080
aatgctatag gctttatgcc caaccttaat gcatacccaa aaaacacaaa cgcagctgca 1140
aaaagtcaca ttgttggaaa agtataccta catggggatg taagcaagcc actagacttg 1200
ataattacat ttaatgaaac cagtgatgaa tcctgtactt attgcattaa ctttcagtgg 1260
cggtggggaa ctgaccaata taaagatgaa acacttgcag tcagttcatt caccttctca 1320
tacattgcta aagaataa                                              1338
```

<210> 79
<211> 445
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 82 Fiber

<400> 79

```
Met Ser Lys Lys Arg Ala Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Ala Val Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Asp Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Lys Asn Ala Thr Lys Ala Thr Ala Pro Leu Ser Ile Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Leu Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Asn
    130                 135                 140
```

```
Gly Ala Leu Val Ala Gln Leu Ala Tyr Pro Leu Val Phe Asn Thr Ala
145                 150                 155                 160
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Lys Arg Gly Ile Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Ile Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Ile Gly Val Asn Ile Asp Thr Lys Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Ser Ser Thr Glu Thr Asp Val Lys Asn Ala Phe Pro Leu
225                 230                 235                 240
Gln Val Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Ala Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Ile Ala Val Asp Thr Gly Ser Leu Asn Pro Ile Thr Gly Lys Val
305                 310                 315                 320
Thr Thr Ala Leu Val Ser Leu Lys Phe Asp Ala Asn Gly Val Leu Gln
                325                 330                 335
Ala Ser Ser Thr Leu Asp Lys Glu Tyr Trp Asn Phe Arg Lys Gly Asp
            340                 345                 350
Val Thr Pro Ala Asp Pro Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn
        355                 360                 365
Leu Asn Ala Tyr Pro Lys Asn Thr Asn Ala Ala Ala Lys Ser His Ile
    370                 375                 380
Val Gly Lys Val Tyr Leu His Gly Asp Val Ser Lys Pro Leu Asp Leu
385                 390                 395                 400
Ile Ile Thr Phe Asn Glu Thr Ser Asp Glu Ser Cys Thr Tyr Cys Ile
                405                 410                 415
Asn Phe Gln Trp Arg Trp Gly Thr Asp Gln Tyr Lys Asp Glu Thr Leu
            420                 425                 430
Ala Val Ser Ser Phe Thr Phe Ser Tyr Ile Ala Lys Glu
        435                 440                 445
```

<210> 80
<211> 445
<212> PRT
<213> Chimpanzee Adenovirus- CV23/Pan5 Fiber

<400> 80

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Asp Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
```

```
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Phe Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Thr
    130                 135                 140
Gly Ala Leu Val Ala Gln Leu Ala Ser Pro Leu Ala Phe Asp Ser Asn
145                 150                 155                 160
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Asn Arg Gly Leu Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Ala Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Met Gly Val Asn Ile Asp Thr Gln Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Thr Ser Thr Val Ala Asp Val Lys Asn Ala Tyr Pro Ile
225                 230                 235                 240
Gln Ile Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Asp Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Glu Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Ile Ala Val Asp Thr Gly Ser Leu Asn Pro Ile Thr Gly Thr Val
305                 310                 315                 320
Thr Thr Ala Leu Val Ser Leu Lys Phe Asp Ala Asn Gly Val Leu Gln
                325                 330                 335
Ser Ser Ser Thr Leu Asp Ser Asp Tyr Trp Asn Phe Arg Gln Gly Asp
            340                 345                 350
Val Thr Pro Ala Glu Ala Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn
        355                 360                 365
Leu Lys Ala Tyr Pro Lys Asn Thr Ser Gly Ala Ala Lys Ser His Ile
    370                 375                 380
Val Gly Lys Val Tyr Leu His Gly Asp Thr Gly Lys Pro Leu Asp Leu
385                 390                 395                 400
Ile Ile Thr Phe Asn Glu Thr Ser Asp Glu Ser Cys Thr Tyr Cys Ile
                405                 410                 415
Asn Phe Gln Trp Gln Trp Gly Ala Asp Gln Tyr Lys Asn Glu Thr Leu
            420                 425                 430
Ala Val Ser Ser Phe Thr Phe Ser Tyr Ile Ala Lys Glu
        435                 440                 445
```

<210> 81
<211> 443
<212> PRT
<213> Chimpanzee Adenovirus- CV32/Pan6 Fiber

<400> 81

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
```

```
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
                20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
            35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
        50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Ile Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Lys Thr Ala Ala Pro Phe Tyr Asn Asn Asn Gly Thr Leu
            100                 105                 110
Ser Leu Asn Val Ser Thr Pro Leu Ala Val Phe Pro Thr Phe Asn Thr
        115                 120                 125
Leu Gly Ile Ser Leu Gly Asn Gly Leu Gln Thr Ser Asn Lys Leu Leu
    130                 135                 140
Thr Val Gln Leu Thr His Pro Leu Thr Phe Ser Ser Asn Ser Ile Thr
145                 150                 155                 160
Val Lys Thr Asp Lys Gly Leu Tyr Ile Asn Ser Ser Gly Asn Arg Gly
                165                 170                 175
Leu Glu Ala Asn Ile Ser Leu Lys Arg Gly Leu Val Phe Asp Gly Asn
            180                 185                 190
Ala Ile Ala Thr Tyr Ile Gly Asn Gly Leu Asp Tyr Gly Ser Tyr Asp
        195                 200                 205
Ser Asp Gly Lys Thr Arg Pro Val Ile Thr Lys Ile Gly Ala Gly Leu
    210                 215                 220
Asn Phe Asp Ala Asn Lys Ala Ile Ala Val Lys Leu Gly Thr Gly Leu
225                 230                 235                 240
Ser Phe Asp Ser Ala Gly Ala Leu Thr Ala Gly Asn Lys Gln Asp Asp
                245                 250                 255
Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro Asn Cys Gln Leu
            260                 265                 270
Leu Ser Asp Arg Asp Ala Lys Phe Thr Leu Cys Leu Thr Lys Cys Gly
        275                 280                 285
Ser Gln Ile Leu Gly Thr Val Ala Val Ala Ala Val Thr Val Gly Ser
    290                 295                 300
Ala Leu Asn Pro Ile Asn Asp Thr Val Lys Ser Ala Ile Val Phe Leu
305                 310                 315                 320
Arg Phe Asp Ser Asp Gly Val Leu Met Ser Asn Ser Ser Met Val Gly
                325                 330                 335
Asp Tyr Trp Asn Phe Arg Glu Gly Gln Thr Thr Gln Ser Val Ala Tyr
            340                 345                 350
Thr Asn Ala Val Gly Phe Met Pro Asn Ile Gly Ala Tyr Pro Lys Thr
        355                 360                 365
Gln Ser Lys Thr Pro Lys Asn Ser Ile Val Ser Gln Val Tyr Leu Thr
    370                 375                 380
Gly Glu Thr Thr Met Pro Met Thr Leu Thr Ile Thr Phe Asn Gly Thr
385                 390                 395                 400
Asp Glu Lys Asp Thr Thr Pro Val Ser Thr Tyr Ser Met Thr Phe Thr
                405                 410                 415
Trp Gln Trp Thr Gly Asp Tyr Lys Asp Lys Asn Ile Thr Phe Ala Thr
            420                 425                 430
Asn Ser Phe Ser Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 82
<211> 443
<212> PRT
<213> Chimpanzee Adenovirus- CV33/Pan7 Fiber

<400> 82

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Leu Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Lys Ser Thr Ile Leu
        115                 120                 125
Asn Thr Leu Ala Val Ala Tyr Gly Ser Gly Leu Gly Leu Ser Gly Gly
    130                 135                 140
Thr Ala Leu Ala Val Gln Leu Ala Ser Pro Leu Thr Phe Asp Glu Lys
145                 150                 155                 160
Gly Asn Ile Lys Ile Asn Leu Ala Ser Gly Pro Leu Thr Val Asp Ala
                165                 170                 175
Ser Arg Leu Ser Ile Asn Cys Lys Arg Gly Val Thr Val Thr Thr Ser
            180                 185                 190
Gly Asp Ala Ile Glu Ser Asn Ile Ser Trp Pro Lys Gly Ile Arg Phe
        195                 200                 205
Glu Gly Asn Gly Ile Ala Ala Asn Ile Gly Arg Gly Leu Glu Phe Gly
    210                 215                 220
Thr Thr Ser Thr Glu Thr Asp Val Thr Asp Ala Tyr Pro Ile Gln Val
225                 230                 235                 240
Lys Leu Gly Thr Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile Val Ala
                245                 250                 255
Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala Asp Pro
            260                 265                 270
Ser Pro Asn Cys Lys Ile Tyr Ser Glu Lys Asp Ala Lys Leu Thr Leu
        275                 280                 285
Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Thr Val Leu
    290                 295                 300
Ala Val Asn Asn Gly Ser Leu Asn Pro Ile Thr Asn Thr Val Ser Thr
305                 310                 315                 320
Ala Leu Val Ser Leu Lys Phe Asp Ala Ser Gly Val Leu Leu Ser Ser
                325                 330                 335
Ser Thr Leu Asp Lys Glu Tyr Trp Asn Phe Arg Lys Gly Asp Val Thr
            340                 345                 350
Pro Ala Glu Pro Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn Ile Lys
        355                 360                 365
Ala Tyr Pro Lys Asn Thr Ser Ala Ala Ser Lys Ser His Ile Val Ser
    370                 375                 380

Gln Val Tyr Leu Asn Gly Asp Glu Ala Lys Pro Leu Met Leu Ile Ile
385                 390                 395                 400
Thr Phe Asn Glu Thr Glu Asp Ala Thr Cys Thr Tyr Ser Ile Thr Phe
                405                 410                 415
Gln Trp Lys Trp Asp Ser Thr Lys Tyr Thr Gly Glu Thr Leu Ala Thr
            420                 425                 430
Ser Ser Phe Thr Phe Ser Tyr Ile Ala Gln Glu
        435                 440
```

<210> 83

<211> 543
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 3 Fiber

<400> 83

```
Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
            20                  25                  30
Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
        35                  40                  45
Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
    50                  55                  60
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65                  70                  75                  80
Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
                85                  90                  95
Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
            100                 105                 110
Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
        115                 120                 125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
    130                 135                 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145                 150                 155                 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
                165                 170                 175
Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp
            180                 185                 190
Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
        195                 200                 205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
    210                 215                 220
Thr Gly Asn Gly Leu Thr Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225                 230                 235                 240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
                245                 250                 255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asn Val
            260                 265                 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
        275                 280                 285
Gly Pro Leu Tyr Ile Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
    290                 295                 300
Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305                 310                 315                 320
```

```
Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
                325                 330                 335
Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
            340                 345                 350
Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
        355                 360                 365
Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
    370                 375                 380
Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385                 390                 395                 400
Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
                405                 410                 415
Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
            420                 425                 430
Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
        435                 440                 445
Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
    450                 455                 460
Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465                 470                 475                 480
Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
                485                 490                 495
Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
            500                 505                 510
Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
        515                 520                 525
Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
    530                 535                 540
```

<210> 84
<211> 445
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 6 Fiber

<400> 84

```
Met Ser Lys Lys Arg Ala Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Ala Val Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Asp Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Lys Asn Ala Thr Lys Ala Thr Ala Pro Leu Ser Ile Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp Thr Pro Leu Tyr Asn Asn Asn Gly Lys Leu
            100                 105                 110
Gly Met Lys Val Thr Ala Pro Leu Lys Ile Leu Asp Thr Asp Leu Leu
        115                 120                 125
Lys Thr Leu Val Val Ala Tyr Gly Gln Gly Leu Gly Thr Asn Thr Asn
    130                 135                 140
Gly Ala Leu Val Ala Gln Leu Ala Tyr Pro Leu Val Phe Asn Thr Ala
145                 150                 155                 160
```

```
Ser Lys Ile Ala Leu Asn Leu Gly Asn Gly Pro Leu Lys Val Asp Ala
                165                 170                 175
Asn Arg Leu Asn Ile Asn Cys Lys Arg Gly Ile Tyr Val Thr Thr Thr
            180                 185                 190
Lys Asp Ala Leu Glu Ile Asn Ile Ser Trp Ala Asn Ala Met Thr Phe
        195                 200                 205
Ile Gly Asn Ala Ile Gly Val Asn Ile Asp Thr Lys Lys Gly Leu Gln
    210                 215                 220
Phe Gly Thr Ser Ser Thr Glu Thr Asp Val Lys Asn Ala Phe Pro Leu
225                 230                 235                 240
Gln Val Lys Leu Gly Ala Gly Leu Thr Phe Asp Ser Thr Gly Ala Ile
                245                 250                 255
Val Ala Trp Asn Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Ala
            260                 265                 270
Asp Pro Ser Pro Asn Cys His Ile Tyr Ser Ala Lys Asp Ala Lys Leu
        275                 280                 285
Thr Leu Cys Leu Thr Lys Cys Gly Ser Gln Ile Leu Gly Thr Val Ser
    290                 295                 300
Leu Ile Ala Val Asp Thr Gly Ser Leu Asn Pro Ile Thr Gly Lys Val
305                 310                 315                 320
Thr Thr Ala Leu Val Ser Leu Lys Phe Asp Ala Asn Gly Val Leu Gln
                325                 330                 335
Ala Ser Ser Thr Leu Asp Lys Glu Tyr Trp Asn Phe Arg Lys Gly Asp
            340                 345                 350
Val Thr Pro Ala Asp Pro Tyr Thr Asn Ala Ile Gly Phe Met Pro Asn
        355                 360                 365
Leu Asn Ala Tyr Pro Lys Asn Thr Asn Ala Ala Ala Lys Ser His Ile
    370                 375                 380
Val Gly Lys Val Tyr Leu His Gly Asp Glu Ser Lys Pro Leu Asp Leu
385                 390                 395                 400
Ile Ile Thr Phe Asn Glu Thr Ser Asp Glu Ser Cys Thr Tyr Cys Ile
                405                 410                 415
Asn Phe Gln Trp Gln Trp Gly Thr Asp Gln Tyr Lys Asp Glu Thr Leu
            420                 425                 430
Ala Val Ser Ser Phe Thr Phe Ser Tyr Ile Ala Lys Glu
        435                 440                 445
```

<210> 85
<211> 322
<212> PRT
<213> Chimpanzee Adenovirus- C1 Fiber

<400> 85

```
Met Ala Lys Arg Thr Arg Leu Ser Ser Ser Phe Asn Pro Val Tyr Pro
 1               5                  10                  15
Tyr Glu Asp Glu Asn Ser Ser His Pro Phe Ile Asn Pro Gly Phe Ile
            20                  25                  30
Ser Pro Asn Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu Asn
        35                  40                  45
Cys Val Ala Pro Leu Thr Thr Ala Asn Gly Ala Leu Asp Ile Lys Val
    50                  55                  60
Gly Gly Gly Leu Lys Val Asn Ser Thr Asp Gly Phe Leu Glu Glu Asn
65                  70                  75                  80
Ile Asn Ile Thr Ser Pro Leu Thr Lys Ser Asn His Ser Ile Gly Leu
                85                  90                  95
```

```
Glu Trp Ser Asp Gly Leu Gln Thr Asn Glu Ala Lys Leu Cys Val Lys
            100                 105                 110
Leu Gly Lys Gly Leu Val Phe Asp Ser Ser Ser Ala Ile Ala Met Glu
        115                 120                 125
Asn Asn Thr Leu Trp Thr Gly Ala Lys Pro Ser Ala Asn Cys Val Ile
    130                 135                 140
Lys Glu Gly Glu Asp Ser Pro Asp Cys Lys Leu Thr Leu Val Leu Val
145                 150                 155                 160
Lys Asn Gly Gly Leu Val Asn Gly Tyr Ile Thr Leu Met Gly Asp Ser
                165                 170                 175
Glu Tyr Thr Asn Thr Leu Phe Lys Asn Lys Gln Val Thr Ile Asp Val
            180                 185                 190
Asn Leu Ala Phe Asp Asn Thr Gly Gln Ile Ile Thr Tyr Leu Ser Ser
        195                 200                 205
Leu Lys Ser Asn Leu Asn Phe Lys Asp Asn Gln Asn Met Ala Thr Gly
    210                 215                 220
Thr Ile Thr Ser Ala Lys Gly Phe Met Pro Ser Thr Thr Ala Tyr Pro
225                 230                 235                 240
Phe Ile Thr Tyr Ala Thr Gln Ser Leu Asn Glu Asp Tyr Ile Tyr Gly
                245                 250                 255
Glu Cys Tyr Tyr Lys Ser Thr Asn Gly Thr Leu Phe Pro Leu Lys Val
            260                 265                 270
Thr Val Thr Leu Asn Arg Arg Met Ser Ala Ser Gly Met Ala Tyr Ala
        275                 280                 285
Met Asn Phe Ser Trp Ser Leu Asn Ala Glu Glu Ala Pro Glu Thr Thr
    290                 295                 300
Glu Val Thr Leu Ile Thr Ser Pro Phe Phe Phe Ser Tyr Ile Arg Glu
305                 310                 315                 320
Asp Asp
```

<210> 86
<211> 425
<212> PRT
<213> Chimpanzee Adenovirus- CV68 Fiber

<400> 86

```
Met Ser Lys Lys Arg Val Arg Val Asp Asp Asp Phe Asp Pro Val Tyr
 1               5                  10                  15
Pro Tyr Asp Ala Asp Asn Ala Pro Thr Val Pro Phe Ile Asn Pro Pro
            20                  25                  30
Phe Val Ser Ser Asp Gly Phe Gln Glu Lys Pro Leu Gly Val Leu Ser
        35                  40                  45
Leu Arg Leu Ala Asp Pro Val Thr Thr Lys Asn Gly Glu Ile Thr Leu
    50                  55                  60
Lys Leu Gly Glu Gly Val Asp Leu Asp Ser Ser Gly Lys Leu Ile Ser
65                  70                  75                  80
Asn Thr Ala Thr Lys Ala Ala Ala Pro Leu Ser Phe Ser Asn Asn Thr
                85                  90                  95
Ile Ser Leu Asn Met Asp His Pro Phe Tyr Thr Lys Asp Gly Lys Leu
            100                 105                 110
Ser Leu Gln Val Ser Pro Pro Leu Asn Ile Leu Arg Thr Ser Ile Leu
        115                 120                 125
Asn Thr Leu Ala Leu Gly Phe Gly Ser Gly Leu Gly Leu Arg Gly Ser
    130                 135                 140
```

```
Ala Leu Ala Val Gln Leu Val Ser Pro Leu Thr Phe Asp Thr Asp Gly
145                 150                 155                 160
Asn Ile Lys Leu Thr Leu Asp Arg Gly Leu His Val Thr Thr Gly Asp
                165                 170                 175
Ala Ile Glu Ser Asn Ile Ser Trp Ala Lys Gly Leu Lys Phe Glu Asp
            180                 185                 190
Gly Ala Ile Ala Thr Asn Ile Gly Asn Gly Leu Glu Phe Gly Ser Ser
        195                 200                 205
Ser Thr Glu Thr Gly Val Asp Asp Ala Tyr Pro Ile Gln Val Lys Leu
    210                 215                 220
Gly Ser Gly Leu Ser Phe Asp Ser Thr Gly Ala Ile Met Ala Gly Asn
225                 230                 235                 240
Lys Glu Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser Pro
                245                 250                 255
Asn Cys Gln Ile Leu Ala Glu Asn Asp Ala Lys Leu Thr Leu Cys Leu
            260                 265                 270
Thr Lys Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Val Val
        275                 280                 285
Gly Ser Gly Asn Leu Asn Pro Ile Thr Gly Thr Val Ser Ser Ala Gln
    290                 295                 300
Val Phe Leu Arg Phe Asp Ala Asn Gly Val Leu Leu Thr Glu His Ser
305                 310                 315                 320
Thr Leu Lys Lys Tyr Trp Gly Tyr Arg Gln Gly Asp Ser Ile Asp Gly
                325                 330                 335
Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Lys Ala Tyr
            340                 345                 350
Pro Lys Ser Gln Ser Ser Thr Thr Lys Asn Asn Ile Val Gly Gln Val
        355                 360                 365
Tyr Met Asn Gly Asp Val Ser Lys Pro Met Leu Leu Thr Ile Thr Leu
    370                 375                 380
Asn Gly Thr Asp Asp Ser Asn Ser Thr Tyr Ser Met Ser Phe Ser Tyr
385                 390                 395                 400
Thr Trp Thr Asn Gly Ser Tyr Val Gly Ala Thr Phe Gly Ala Asn Ser
                405                 410                 415
Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
            420                 425
```

<210> 87
<211> 954
<212> PRT
<213> Chimpanzee Adenovirus- ChAd20 Hexon

<400> 87

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
```

```
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
                100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Pro Cys Glu Trp Asp Glu Ala Ala Thr Ala Leu Asp Ile
    130                 135                 140
Asp Leu Asn Ala Glu Asp Asp Glu Glu Ser Asp Glu Ala Gln Gly Glu
145                 150                 155                 160
Ala Asp Gln Gln Lys Thr His Val Phe Gly Gln Ala Pro Tyr Ser Gly
                165                 170                 175
Gln Asn Ile Thr Lys Glu Gly Ile Gln Ile Gly Ile Asp Ala Ala Ser
            180                 185                 190
Gln Ala Gln Thr Pro Val Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro
        195                 200                 205
Gln Val Gly Glu Ser Gln Trp Asn Glu Thr Glu Ile Ser Tyr Gly Ala
    210                 215                 220
Gly Arg Val Leu Lys Lys Thr Thr Leu Met Lys Pro Cys Tyr Gly Ser
225                 230                 235                 240
Tyr Ala Arg Pro Thr Asn Glu Asn Gly Gly Gln Gly Ile Leu Leu Glu
                245                 250                 255
Gln Asp Gly Lys Lys Glu Ser Gln Val Glu Met Gln Phe Phe Ser Thr
            260                 265                 270
Thr Gln Ala Ala Ala Gly Asn Ser Asp Asn Pro Thr Pro Lys Val Val
        275                 280                 285
Leu Tyr Ser Glu Asp Val Asn Leu Glu Thr Pro Asp Thr His Ile Ser
    290                 295                 300
Tyr Met Pro Thr Asn Asn Glu Thr Asn Ser Arg Glu Leu Leu Gly Gln
305                 310                 315                 320
Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe
                325                 330                 335
Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala
            340                 345                 350
Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn
        355                 360                 365
Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly Asp Arg Thr
    370                 375                 380
Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp
385                 390                 395                 400
Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu Pro Asn Tyr
                405                 410                 415
Cys Phe Pro Leu Gly Gly Val Ile Asn Thr Glu Thr Phe Thr Lys Val
            420                 425                 430
Lys Pro Lys Ala Ala Gln Asp Ala Gln Trp Glu Lys Asp Ser Glu Phe
        435                 440                 445
Ser Asp Lys Asn Glu Ile Arg Val Gly Asn Asn Phe Ala Met Glu Ile
    450                 455                 460
Asn Leu Asn Ala Asn Leu Trp Arg Asn Phe Leu Tyr Ser Asn Val Ala
465                 470                 475                 480
Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Thr Pro Ser Asn Val Gln Ile
                485                 490                 495
Ser Asn Asn Pro Asn Ser Tyr Asp Tyr Met Asn Lys Arg Val Val Ala
            500                 505                 510
Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala Arg Trp Ser Leu
        515                 520                 525
```

```
Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly
    530                 535                 540
Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe
545                 550                 555                 560
His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu
                565                 570                 575
Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn
            580                 585                 590
Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala
        595                 600                 605
Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr Phe Phe Pro Met
    610                 615                 620
Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr
625                 630                 635                 640
Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr
                645                 650                 655
Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg
            660                 665                 670
Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg Leu Lys Thr Lys
        675                 680                 685
Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Tyr Thr Tyr Ser
    690                 695                 700
Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe
705                 710                 715                 720
Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn
                725                 730                 735
Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Ser Val Asp
            740                 745                 750
Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe
        755                 760                 765
Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr
    770                 775                 780
Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe
785                 790                 795                 800
Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys Tyr Lys Asp Tyr
                805                 810                 815
Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser Gly Phe Val Gly
            820                 825                 830
Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Phe
        835                 840                 845
Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser Ile Thr Gln Lys
    850                 855                 860
Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro Phe Ser Ser Asn
865                 870                 875                 880
Phe Met Ser Met Gly Ala Leu Ser Asp Leu Gly Gln Asn Leu Leu Tyr
                885                 890                 895
Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met
            900                 905                 910
Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val
        915                 920                 925
Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr Val Tyr Leu Arg
    930                 935                 940

Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950
```

<210> 88
<211> 940
<212> PRT

<213> Chimpanzee Adenovirus- ChAd 4 Hexon

<400> 88

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Lys Lys Thr Gly Asn Asn Ala
    130                 135                 140
Asn Gly Asp Thr Glu Asn Val Thr Tyr Gly Val Ala Ala Met Gly Gly
145                 150                 155                 160
Ile Asp Ile Asp Lys Asn Gly Leu Gln Ile Gly Thr Asp Asp Thr Lys
                165                 170                 175
Asp Asp Asp Asn Glu Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro
            180                 185                 190
Gln Ile Gly Glu Glu Asn Trp Gln Glu Thr Tyr Ser Tyr Tyr Gly Gly
        195                 200                 205
Arg Ala Leu Lys Lys Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe
    210                 215                 220
Ala Arg Pro Thr Asn Val Lys Gly Gly Gln Ala Lys Ile Lys Thr Asp
225                 230                 235                 240
Gly Asp Val Lys Ser Phe Asp Ile Asp Leu Ala Phe Phe Asp Ile Pro
                245                 250                 255
Asn Ser Gly Ala Gly Asn Gly Thr Asn Val Asn Asp Asp Pro Asp Met
            260                 265                 270
Val Met Tyr Thr Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile
        275                 280                 285
Val Tyr Lys Pro Gly Thr Ser Asp Asp Ser Ser Lys Val Asn Leu Cys
    290                 295                 300
Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn
305                 310                 315                 320
Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu
                325                 330                 335
Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg
            340                 345                 350
Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg
        355                 360                 365
```

```
Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro
    370                 375                 380
Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn
385                 390                 395                 400
Tyr Cys Phe Pro Leu Asp Gly Ala Gly Thr Asn Ser Val Tyr Gln Gly
                405                 410                 415
Val Lys Pro Lys Thr Asp Asn Gly Asn Asp Gln Trp Glu Thr Asp Ser
            420                 425                 430
Thr Val Ser Ser His Asn Gln Ile Cys Lys Gly Asn Ile Tyr Ala Met
        435                 440                 445
Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn
    450                 455                 460
Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Ile
465                 470                 475                 480
Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val
                485                 490                 495
Val Pro Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp
            500                 505                 510
Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn
        515                 520                 525
Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val
    530                 535                 540
Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu
545                 550                 555                 560
Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp
                565                 570                 575
Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp
            580                 585                 590
Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe
        595                 600                 605
Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn
    610                 615                 620
Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met
625                 630                 635                 640
Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro
                645                 650                 655
Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys
            660                 665                 670
Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val
        675                 680                 685
Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His
    690                 695                 700
Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro
705                 710                 715                 720
Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr
                725                 730                 735
Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp
            740                 745                 750
Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly
        755                 760                 765
Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg
    770                 775                 780
Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys
785                 790                 795                 800
```

```
Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe
                805                 810                 815
Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala
            820                 825                 830
Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Thr Ser Val Thr
        835                 840                 845
Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser
    850                 855                 860
Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met
865                 870                 875                 880
Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp
                885                 890                 895
Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp
            900                 905                 910
Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr
        915                 920                 925
Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935                 940
```

<210> 89
<211> 940
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 5 Hexon

<400> 89

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Lys Lys Thr Gly Asn Asn Ala
    130                 135                 140
Asn Gly Asp Thr Glu Asn Val Thr Tyr Gly Val Ala Ala Met Gly Gly
145                 150                 155                 160
Ile Asp Ile Asp Lys Asn Gly Leu Gln Ile Gly Thr Asp Asp Thr Lys
                165                 170                 175
Asp Asp Asp Asn Glu Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro
            180                 185                 190
Gln Ile Gly Glu Glu Asn Trp Gln Glu Thr Tyr Ser Tyr Tyr Gly Gly
        195                 200                 205
Arg Ala Leu Lys Lys Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe
    210                 215                 220
Ala Arg Pro Thr Asn Val Lys Gly Gly Gln Ala Lys Ile Lys Thr Asp
225                 230                 235                 240
```

```
Gly Asp Val Lys Ser Phe Asp Ile Asp Leu Ala Phe Phe Asp Ile Pro
                245                 250                 255
Asn Ser Gly Ala Gly Asn Gly Thr Asn Val Asn Asp Asp Pro Asp Met
            260                 265                 270
Val Met Tyr Thr Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile
        275                 280                 285
Val Tyr Lys Pro Gly Thr Ser Asp Asp Ser Ser Lys Val Asn Leu Cys
    290                 295                 300
Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn
305                 310                 315                 320
Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu
                325                 330                 335
Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg
            340                 345                 350
Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg
        355                 360                 365
Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro
    370                 375                 380
Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn
385                 390                 395                 400
Tyr Cys Phe Pro Leu Asp Gly Ala Gly Thr Asn Ser Val Tyr Gln Gly
                405                 410                 415
Val Lys Pro Lys Thr Asp Asn Gly Asn Asp Gln Trp Glu Thr Asp Ser
            420                 425                 430
Thr Val Ser Ser His Asn Gln Ile Cys Lys Gly Asn Ile Tyr Ala Met
        435                 440                 445
Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn
    450                 455                 460
Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Ile
465                 470                 475                 480
Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val
                485                 490                 495
Val Pro Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp
            500                 505                 510
Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn
        515                 520                 525
Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val
    530                 535                 540
Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu
545                 550                 555                 560
Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp
                565                 570                 575
Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp
            580                 585                 590
Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe
        595                 600                 605
Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn
    610                 615                 620
Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met
625                 630                 635                 640
Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro
                645                 650                 655
Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys
            660                 665                 670
```

```
Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val
        675                 680                 685
Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His
    690                 695                 700
Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro
705                 710                 715                 720
Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr
                725                 730                 735
Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp
            740                 745                 750
Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly
        755                 760                 765
Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg
    770                 775                 780
Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys
785                 790                 795                 800
Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe
                805                 810                 815
Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala
            820                 825                 830
Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Ala Ser Val Thr
        835                 840                 845
Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser
    850                 855                 860
Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met
865                 870                 875                 880
Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp
                885                 890                 895
Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp
            900                 905                 910
Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr
        915                 920                 925
Leu Arg Thr Pro Phe Ser Ala Gly Lys Ala Thr Thr
    930                 935                 940
```

<210> 90
<211> 940
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 7 Hexon

<400> 90

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
```

```
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Lys Lys Thr Gly Lys Asn Ala
    130                 135                 140
Asn Gly Asp Thr Glu Asn Val Thr Tyr Gly Val Ala Ala Met Gly Gly
145                 150                 155                 160
Ile Asp Ile Asp Lys Asn Gly Leu Gln Ile Gly Thr Asp Asp Thr Lys
                165                 170                 175
Asp Gly Asp Asn Glu Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro
            180                 185                 190
Gln Ile Gly Glu Glu Asn Trp Gln Glu Thr Tyr Ser Tyr Tyr Gly Gly
        195                 200                 205
Arg Ala Leu Lys Lys Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe
    210                 215                 220
Ala Arg Pro Thr Asn Val Lys Gly Gly Gln Ala Lys Ile Lys Thr Asp
225                 230                 235                 240
Gly Asp Val Lys Ser Phe Asp Ile Asp Leu Ala Phe Phe Asp Ile Pro
                245                 250                 255
Asn Ser Gly Ala Gly Asn Gly Thr Asn Val Asn Asp Asp Pro Asp Met
            260                 265                 270
Val Met Tyr Thr Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile
        275                 280                 285
Val Tyr Lys Pro Gly Thr Ser Asp Asp Ser Ser Glu Val Asn Leu Cys
    290                 295                 300
Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn
305                 310                 315                 320
Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu
                325                 330                 335
Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg
            340                 345                 350
Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg
        355                 360                 365
Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro
    370                 375                 380
Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn
385                 390                 395                 400
Tyr Cys Phe Pro Leu Asp Gly Ala Gly Thr Asn Ser Val Tyr Gln Gly
                405                 410                 415
Val Lys Pro Lys Thr Asp Asn Gly Asn Asp Gln Trp Glu Thr Asp Ser
            420                 425                 430
Thr Val Ser Ser His Asn Gln Ile Cys Lys Gly Asn Ile Tyr Ala Met
        435                 440                 445
Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn
    450                 455                 460
Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Ile
465                 470                 475                 480
Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val
                485                 490                 495
Val Pro Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp
            500                 505                 510
Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn
        515                 520                 525
Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val
    530                 535                 540
```

```
Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu
545                 550                 555                 560
Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp
                565                 570                 575
Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp
            580                 585                 590
Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe
        595                 600                 605
Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn
    610                 615                 620
Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met
625                 630                 635                 640
Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro
                645                 650                 655
Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys
            660                 665                 670
Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val
        675                 680                 685
Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His
    690                 695                 700
Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro
705                 710                 715                 720
Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr
                725                 730                 735
Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp
            740                 745                 750
Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly
        755                 760                 765
Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg
    770                 775                 780
Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys
785                 790                 795                 800
Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe
                805                 810                 815
Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala
            820                 825                 830
Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Thr Ser Val Thr
        835                 840                 845
Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser
    850                 855                 860
Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met
865                 870                 875                 880
Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp
                885                 890                 895
Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp
            900                 905                 910
Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr
        915                 920                 925
Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
930                 935                 940
```

<210> 91
<211> 930
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 9 Hexon

<400> 91

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Thr Tyr Thr Asp Asn Gln Thr Glu Lys
    130                 135                 140
Thr Ala Thr Tyr Gly Asn Ala Pro Val Glu Gly Ile Asn Ile Thr Lys
145                 150                 155                 160
Asp Gly Ile Gln Leu Gly Thr Asp Ser Asp Gly Gln Ala Ile Tyr Ala
                165                 170                 175
Asp Glu Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Pro Glu Trp His
            180                 185                 190
Asp Thr Thr Gly Thr Glu Glu Lys Tyr Gly Gly Arg Ala Leu Lys Pro
        195                 200                 205
Ala Thr Asp Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys Pro Thr Asn
    210                 215                 220
Val Lys Gly Gly Gln Ala Lys Ser Arg Thr Lys Thr Asp Gly Thr Thr
225                 230                 235                 240
Glu Pro Asp Ile Asp Met Ala Phe Phe Asp Gly Arg Asn Ala Thr Thr
                245                 250                 255
Ala Gly Leu Thr Pro Glu Ile Val Leu Tyr Thr Glu Asn Val Asp Leu
            260                 265                 270
Glu Thr Pro Asp Thr His Ile Val Tyr Lys Ala Gly Thr Asp Asp Ser
        275                 280                 285
Ser Ser Ser Ile Asn Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn
    290                 295                 300
Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser
305                 310                 315                 320
Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala
                325                 330                 335
Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu
            340                 345                 350
Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln
        355                 360                 365
Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly
    370                 375                 380
Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asn Ala Val Gly
385                 390                 395                 400
```

```
Arg Thr Asn Ser Tyr Gln Gly Ile Lys Pro Asn Gly Gly Asp Pro Ala
                405                 410                 415
Thr Trp Ala Lys Asp Glu Ser Val Asn Asp Ser Asn Glu Leu Gly Lys
            420                 425                 430
Gly Asn Pro Phe Ala Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg
        435                 440                 445
Asn Phe Leu Tyr Ala Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys
    450                 455                 460
Tyr Thr Pro Ala Asn Ile Thr Leu Pro Ala Asn Thr Asn Thr Tyr Asp
465                 470                 475                 480
Tyr Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp Ala Tyr Ile
                485                 490                 495
Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro
            500                 505                 510
Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu
        515                 520                 525
Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe
    530                 535                 540
Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu
545                 550                 555                 560
Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly
                565                 570                 575
Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ala Phe Thr Ser Ile Asn
            580                 585                 590
Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu
        595                 600                 605
Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr
    610                 615                 620
Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn
625                 630                 635                 640
Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp
                645                 650                 655
Ser Phe Thr Arg Leu Lys Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly
            660                 665                 670
Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly
        675                 680                 685
Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp
    690                 695                 700
Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu
705                 710                 715                 720
Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln
                725                 730                 735
Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr
            740                 745                 750
Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg
        755                 760                 765
Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val
    770                 775                 780
Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln
785                 790                 795                 800
His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln
                805                 810                 815
Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser
            820                 825                 830
```

```
Ala Val Ala Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met
        835                 840                 845
Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr
    850                 855                 860
Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp
865                 870                 875                 880
Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val
                885                 890                 895
Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly
            900                 905                 910
Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala
        915                 920                 925
Thr Thr
    930
```

<210> 92
<211> 930
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 10 Hexon

<400> 92

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Thr Tyr Thr Asp Asn Gln Thr Glu Lys
    130                 135                 140
Thr Ala Thr Tyr Gly Asn Ala Pro Val Gln Gly Ile Ser Ile Thr Lys
145                 150                 155                 160
Asp Gly Ile Gln Leu Gly Thr Asp Thr Asp Asp Gln Pro Ile Tyr Ala
                165                 170                 175
Asp Lys Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Ala Glu Trp His
            180                 185                 190
Asp Ile Thr Gly Thr Asp Glu Lys Tyr Gly Gly Arg Ala Leu Lys Pro
        195                 200                 205
Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys Pro Thr Asn
    210                 215                 220
Lys Glu Gly Gly Gln Ala Asn Val Lys Thr Glu Thr Gly Gly Thr Lys
225                 230                 235                 240
Glu Tyr Asp Ile Asp Met Ala Phe Phe Asp Asn Arg Ser Ala Ala Ala
                245                 250                 255
Ala Gly Leu Ala Pro Glu Ile Val Leu Tyr Thr Glu Asn Val Asp Leu
            260                 265                 270
```

```
Glu Thr Pro Asp Thr His Ile Val Tyr Lys Ala Gly Thr Asp Asp Ser
        275                 280                 285
Ser Ser Ser Ile Asn Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn
    290                 295                 300
Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser
305                 310                 315                 320
Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala
                325                 330                 335
Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu
            340                 345                 350
Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln
        355                 360                 365
Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly
    370                 375                 380
Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asn Ala Val Gly
385                 390                 395                 400
Arg Thr Asp Thr Tyr Gln Gly Ile Lys Ala Asn Gly Ala Asp Gln Thr
                405                 410                 415
Thr Trp Thr Lys Asp Asp Thr Val Asn Asp Ala Asn Glu Leu Gly Lys
            420                 425                 430
Gly Asn Pro Phe Ala Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg
        435                 440                 445
Asn Phe Leu Tyr Ala Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys
    450                 455                 460
Tyr Thr Pro Ala Asn Ile Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp
465                 470                 475                 480
Tyr Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp Ala Tyr Ile
                485                 490                 495
Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro
            500                 505                 510
Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu
        515                 520                 525
Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe
    530                 535                 540
Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu
545                 550                 555                 560
Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly
                565                 570                 575
Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ala Phe Thr Ser Ile Asn
            580                 585                 590
Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu
        595                 600                 605
Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr
    610                 615                 620
Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn
625                 630                 635                 640
Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp
                645                 650                 655
Ser Phe Thr Arg Leu Lys Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly
            660                 665                 670
Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly
        675                 680                 685
Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp
    690                 695                 700
```

```
Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu
705                 710                 715                 720
Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln
                725                 730                 735
Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr
            740                 745                 750
Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg
        755                 760                 765
Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val
    770                 775                 780
Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln
785                 790                 795                 800
His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln
                805                 810                 815
Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser
            820                 825                 830
Ala Val Ala Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met
        835                 840                 845
Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr
    850                 855                 860
Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp
865                 870                 875                 880
Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val
                885                 890                 895
Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly
            900                 905                 910
Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala
        915                 920                 925
Thr Thr
    930
```

<210> 93
<211> 960
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 11 Hexon

<400> 93

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Cys Glu Trp Glu Gln Glu Glu Thr Gln Ala Val Glu
    130                 135                 140
```

```
Glu Ala Ala Glu Glu Glu Glu Glu Asp Ala Asp Gly Gln Ala Glu Glu
145                 150                 155                 160
Glu Gln Ala Ala Thr Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu
                165                 170                 175
Ser Gly Glu Lys Ile Ser Lys Asp Gly Leu Gln Ile Gly Thr Asp Ala
            180                 185                 190
Thr Ala Thr Glu Gln Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro
        195                 200                 205
Glu Pro Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val
    210                 215                 220
Ala Gly Gly Arg Val Leu Lys Lys Thr Thr Pro Met Lys Pro Cys Tyr
225                 230                 235                 240
Gly Ser Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln Gly Val Leu
                245                 250                 255
Ala Ala Asn Ala Gln Gly Gln Leu Glu Ser Gln Val Glu Met Gln Phe
            260                 265                 270
Phe Ser Thr Ser Glu Asn Ala Arg Asn Glu Ala Asn Asn Ile Gln Pro
        275                 280                 285
Lys Leu Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr
    290                 295                 300
His Leu Ser Tyr Lys Pro Thr Lys Ser Asp Asp Asn Ser Lys Val Met
305                 310                 315                 320
Leu Gly Gln Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
                325                 330                 335
Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
            340                 345                 350
Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln
        355                 360                 365
Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly
    370                 375                 380
Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
385                 390                 395                 400
Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu
                405                 410                 415
Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr
            420                 425                 430
Gln Ala Val Lys Thr Asn Asn Gly Asn Asn Gly Gly Gln Val Thr Trp
        435                 440                 445
Thr Lys Asp Glu Thr Phe Ala Glu Arg Asn Glu Ile Gly Val Gly Asn
    450                 455                 460
Asn Phe Ala Met Glu Ile Asn Leu Asn Ala Asn Leu Trp Arg Asn Phe
465                 470                 475                 480
Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn
                485                 490                 495
Pro Ser Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met
            500                 505                 510
Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu
        515                 520                 525
Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn
    530                 535                 540
His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
545                 550                 555                 560
Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
                565                 570                 575
```

```
Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
            580                 585                 590
Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp
        595                 600                 605
Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr
    610                 615                 620
Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
625                 630                 635                 640
Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser
                645                 650                 655
Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro
            660                 665                 670
Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe
        675                 680                 685
Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp
    690                 695                 700
Pro Tyr Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe
705                 710                 715                 720
Tyr Leu Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser
                725                 730                 735
Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu
            740                 745                 750
Ile Lys Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn
        755                 760                 765
Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile
    770                 775                 780
Gly Tyr Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr
785                 790                 795                 800
Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln
                805                 810                 815
Thr Lys Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn
            820                 825                 830
Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln
        835                 840                 845
Ala Tyr Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val
    850                 855                 860
Asp Ser Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg
865                 870                 875                 880
Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu
                885                 890                 895
Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr
            900                 905                 910
Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe
        915                 920                 925
Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile
    930                 935                 940
Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955                 960
```

<210> 94
<211> 944
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 16 Hexon

<400> 94

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Thr Glu Asn Gly Gly Gly Gln
    130                 135                 140
Ala Thr Thr Lys Thr His Thr Tyr Gly Val Ala Pro Met Gly Gly Thr
145                 150                 155                 160
Asn Ile Thr Val Asp Gly Leu Gln Ile Gly Thr Asp Ala Thr Ala Asp
                165                 170                 175
Thr Glu Lys Pro Ile Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro Gln
            180                 185                 190
Ile Gly Glu Glu Asn Trp Gln Glu Thr Glu Ser Phe Tyr Gly Gly Arg
        195                 200                 205
Ala Leu Lys Lys Asp Thr Asn Met Lys Pro Cys Tyr Gly Ser Phe Ala
    210                 215                 220
Arg Pro Thr Asn Glu Lys Gly Gly Gln Ala Lys Leu Lys Val Gly Ala
225                 230                 235                 240
Asp Gly Leu Pro Thr Lys Glu Phe Asp Ile Asp Leu Ala Phe Phe Asp
                245                 250                 255
Thr Pro Gly Gly Thr Val Thr Gly Gly Thr Glu Glu Tyr Lys Ala Asp
            260                 265                 270
Ile Val Met Tyr Thr Glu Asn Thr Tyr Leu Glu Thr Pro Asp Thr His
        275                 280                 285
Val Val Tyr Lys Pro Gly Lys Asp Asn Thr Ser Ser Lys Ile Asn Leu
    290                 295                 300
Val Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp
305                 310                 315                 320
Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val
                325                 330                 335
Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp
            340                 345                 350
Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp
        355                 360                 365
Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp
    370                 375                 380
Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro
385                 390                 395                 400
Asn Tyr Cys Phe Pro Leu Asp Gly Ser Gly Thr Asn Ala Ala Tyr Gln
                405                 410                 415
Gly Val Lys Val Lys Asn Gly Gln Asp Gly Asp Val Glu Ser Glu Trp
            420                 425                 430
```

```
Glu Lys Asp Asp Thr Val Ala Ala Arg Asn Gln Leu Cys Lys Gly Asn
        435                 440                 445
Ile Phe Ala Met Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe
    450                 455                 460
Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr
465                 470                 475                 480
Pro Ala Asn Ile Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met
                485                 490                 495
Asn Gly Arg Val Val Pro Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile
            500                 505                 510
Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn
        515                 520                 525
His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
    530                 535                 540
Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
545                 550                 555                 560
Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
                565                 570                 575
Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp
            580                 585                 590
Leu Arg Thr Asp Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr
        595                 600                 605
Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
    610                 615                 620
Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser
625                 630                 635                 640
Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro
                645                 650                 655
Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe
            660                 665                 670
Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp
        675                 680                 685
Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe
    690                 695                 700
Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser
705                 710                 715                 720
Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu
                725                 730                 735
Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn
            740                 745                 750
Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile
        755                 760                 765
Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr
    770                 775                 780
Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu
785                 790                 795                 800
Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn
                805                 810                 815
Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln
            820                 825                 830
Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val
        835                 840                 845
Ala Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg
    850                 855                 860
```

```
Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu
865                 870                 875                 880
Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn
                885                 890                 895
Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe
            900                 905                 910
Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile
        915                 920                 925
Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935                 940
```

<210> 95
<211> 960
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 17 Hexon

<400> 95

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Cys Glu Trp Glu Gln Glu Glu Thr Gln Ala Val Glu
    130                 135                 140
Glu Ala Ala Glu Glu Glu Glu Glu Asp Ala Asp Gly Gln Ala Glu Glu
145                 150                 155                 160
Glu Gln Ala Ala Thr Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu
                165                 170                 175
Ser Gly Glu Lys Ile Ser Lys Asp Gly Leu Gln Ile Gly Thr Asp Ala
            180                 185                 190
Thr Ala Thr Glu Gln Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro
        195                 200                 205
Glu Pro Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val
    210                 215                 220
Ala Gly Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr
225                 230                 235                 240
Gly Ser Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln Gly Val Leu
                245                 250                 255
Thr Ala Asn Ala Gln Gly Gln Leu Glu Ser Gln Val Glu Met Gln Phe
            260                 265                 270
Phe Ser Thr Ser Glu Asn Ala Arg Asn Glu Thr Asn Asn Ile Gln Pro
        275                 280                 285
Lys Leu Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr
    290                 295                 300
```

```
His Leu Ser Tyr Lys Pro Ala Lys Ser Asp Asp Asn Ser Lys Ile Met
305                 310                 315                 320
Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
                325                 330                 335
Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
            340                 345                 350
Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln
        355                 360                 365
Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly
    370                 375                 380
Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
385                 390                 395                 400
Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu
                405                 410                 415
Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr
            420                 425                 430
Gln Ala Val Lys Thr Asn Asn Gly Asn Asn Gly Gly Gln Val Thr Trp
        435                 440                 445
Thr Lys Asp Glu Thr Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn
    450                 455                 460
Asn Phe Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe
465                 470                 475                 480
Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn
                485                 490                 495
Pro Ser Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met
            500                 505                 510
Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu
        515                 520                 525
Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn
    530                 535                 540
His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
545                 550                 555                 560
Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
                565                 570                 575
Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
            580                 585                 590
Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp
        595                 600                 605
Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr
    610                 615                 620
Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
625                 630                 635                 640
Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser
                645                 650                 655
Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro
            660                 665                 670
Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe
        675                 680                 685
Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp
    690                 695                 700
Pro Tyr Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe
705                 710                 715                 720
Tyr Leu Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser
                725                 730                 735
```

```
Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu
            740                 745                 750
Ile Lys Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn
        755                 760                 765
Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile
    770                 775                 780
Gly Tyr Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr
785                 790                 795                 800
Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln
                805                 810                 815
Thr Lys Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn
            820                 825                 830
Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln
        835                 840                 845
Ala Tyr Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val
    850                 855                 860
Asp Ser Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg
865                 870                 875                 880
Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu
                885                 890                 895
Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr
            900                 905                 910
Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe
        915                 920                 925
Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile
    930                 935                 940
Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955                 960
```

<210> 96
<211> 958
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 19 Hexon

<400> 96

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Cys Glu Trp Glu Gln Leu Glu Glu Ala Gln Ala Ala
    130                 135                 140
Leu Glu Asp Glu Glu Leu Glu Asp Glu Asp Glu Glu Pro Gln Asp Glu
145                 150                 155                 160
```

```
Ala Pro Val Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu Ser Gly
                165                 170                 175
Glu Glu Ile Thr Lys Asp Gly Leu Gln Ile Gly Ser Asp Asn Thr Glu
            180                 185                 190
Ala Gln Ser Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro Glu Pro
        195                 200                 205
Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val Ala Gly
    210                 215                 220
Gly Arg Val Leu Lys Lys Thr Thr Pro Met Lys Pro Cys Tyr Gly Ser
225                 230                 235                 240
Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln Gly Val Leu Val Ala
                245                 250                 255
Asp Asp Lys Gly Val Leu Gln Ser Lys Val Glu Leu Gln Phe Phe Ser
            260                 265                 270
Asn Thr Thr Thr Leu Asn Gln Arg Glu Gly Asn Asp Thr Lys Pro Lys
        275                 280                 285
Val Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr His
    290                 295                 300
Ile Ser Tyr Lys Pro Thr Lys Ser Asp Asp Asn Ser Lys Val Met Leu
305                 310                 315                 320
Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp
                325                 330                 335
Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val
            340                 345                 350
Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp
        355                 360                 365
Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly Asp
    370                 375                 380
Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp
385                 390                 395                 400
Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu Pro
                405                 410                 415
Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr Gln
            420                 425                 430
Val Ile Lys Thr Asn Gly Asn Gly Gln Ala Asp Pro Thr Trp Glu Lys
        435                 440                 445
Asp Thr Glu Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn Asn Phe
    450                 455                 460
Ala Met Glu Ile Asn Leu Asn Ala Asn Leu Trp Arg Asn Phe Leu Tyr
465                 470                 475                 480
Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn Pro Ser
                485                 490                 495
Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met Asn Lys
            500                 505                 510
Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala
        515                 520                 525
Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His
    530                 535                 540
Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
545                 550                 555                 560
Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
                565                 570                 575
Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
            580                 585                 590
```

```
Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
        595                 600                 605
Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr
    610                 615                 620
Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu
625                 630                 635                 640
Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
                645                 650                 655
Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
            660                 665                 670
Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg
        675                 680                 685
Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr
    690                 695                 700
Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
705                 710                 715                 720
Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser
                725                 730                 735
Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
            740                 745                 750
Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr
        755                 760                 765
Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr
    770                 775                 780
Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe
785                 790                 795                 800
Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys
                805                 810                 815
Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser
            820                 825                 830
Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr
        835                 840                 845
Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser
    850                 855                 860
Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro
865                 870                 875                 880
Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln
                885                 890                 895
Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu
            900                 905                 910
Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val
        915                 920                 925
Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr
    930                 935                 940
Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955
```

<210> 97
<211> 2865
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 8 Hexon

<400> 97

```
atggccaccc catcgatgtt gccccagtgg gcatacatgc acatcgccgg acaggatgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgtg ccacagacac ctacttcaat 120
```

```
ctggggaaca agtttaggaa ccccaccgtg gcccccaccc acgatgtgac caccgaccga 180
agccagcggc tgatgctgcg ctttgtgccc gttgatcggg aggacaatac ctactcatac 240
aaagttcgct acacactggc tgtgggcgac aacagagtgc tggatatggc cagcaccttc 300
tttgacatcc ggggggtgct tgacagaggt cccagtttca agccatactc tggcacggct 360
tacaactcct tggctcctaa gggtgccccc aatacatgcc agtggatagc taaagggtcg 420
cccgttcaag atgatgctga acaagctcag gaacaaaaag atgttaccta tacttttggc 480
aatgcgccag taaaagcaga agatgacatt acaaaagacg gattagaagt aggcatacaa 540
attattggtg atgaggagaa tcccatttat gcagataaaa catatcaacc agagccacag 600
gttggtgacg agcaatggca tgacacaact ggaaccactg agcagtatgg aggcagagct 660
cttaagccag ctacaaacat gaggccatgc tatggctctt ttgccagacc tacaaacaaa 720
aaaggagggc aagctaaaac cagaaaagta gaaaaaactg aaggtgacaa aaagactgaa 780
gttgaagaac ttgacattga tatggatttt tatgatgcaa gatctaaaaa acaaggctat 840
gatcctcaaa tagtgctata ttcagaaaat gtaaatctgg aaacgcccga cactcatatt 900
gtgtacaaac cgggaactga tgaaaccagt tcctccacta atttgggcca gcaagctatg 960
cccaacagac ccaattacat tggtttcagg gacaacttca ttggacttat gtattataac 1020
agtactggta acatgggagt gctggccggt caagcttctc agttgaatgc tgtggtcgac 1080
ttgcaggaca ggaacacaga actgtcctac cagctgctgc ttgactctct gggtgacaga 1140
accagatact ttagcatgtg gaatcaggcc gtggatagct atgacccaga tgtgcgcatt 1200
attgaaaatc atggtgtgga agatgaactt cccaattatt gtttcccatt ggatggtgtt 1260
ggtccaatta cagaaaccta tcaaggtatc aagccaaaaa cagcagacaa tgcaaatgac 1320
caatgggaaa aaaataccga agtaaatgga gctaatgaaa taggaaaggg aaacaattat 1380
gcaatggaaa ttaatctaca agctaacctc tggagaagtt ttctttactc caacgtggct 1440
ctgtatcttc cagacggtta caaatatacc ccagccaatg ttacgctgcc agacaacaaa 1500
aatacctatg ggtacataaa cggacgagta gtgtctccat ctttggtgga ttcatacatc 1560
aacattggag ccagatggtc tttggatctt atggacaatg tcaacccatt taatcaccac 1620
cgcaatgctg ggctgcgcta ccgttccatg cttcttggca atggacgcta tgtgcccttc 1680
cacatccaag tgccgcagaa attctttgct atcaagaact tgctgcttct gccaggctcc 1740
tacacctatg agtggaactt cagaaaggat gtgaacatgg tcctacaaag ttcccttggt 1800
aatgatctca gaactgatgg agccagcatc agttttacca gcatcaacct ctatgccacc 1860
tttttcccaa tggctcacaa cactgcttcc acccttgaag ccatgctgcg caatgacacc 1920
aatgaccagt cattcaacga ctacctctct gcagccaaca tgctctaccc catccctgcc 1980
aatgccacta acattcccat ctccattccc tctcgcaact gggctgcctt caggggctgg 2040
tccttcacca gactcaaaac caaggagact ccctctttgg gatcagggtt cgatccctac 2100
tttgtctatt caggctccat tccctacctg gatggtacct tctacctcaa ccacactttc 2160
aagaaggtct ccatcatgtt tgactcctca gtcagctggc cgggcaatga cagactgttg 2220
tgtcccaatg agtttgaaat caagcgcact gtggatgggg aagggtacaa cgtggctcaa 2280
tgcaacatga ccaaggattg gttcttggtc cagatgcttg ccaactacaa cattggctac 2340
cagggcttct acatcccaga ggggtacaag gatcgcatgt actccttctt cagaaacttc 2400
cagcccatga gcagacaggt agttgatgaa gtcaattaca aggagtacca agctgtcaca 2460
cttgcttacc agcacaacaa ctctggtttt gtgggttacc atgcacccac tcttcgtcag 2520
ggccaaccat acccagctaa ctaccctac ccgctcattg gaaccactgc tgtcaccagc 2580
gtcacccaga aaaagttctt gtgcgacagg accatgtggc gcatcccatt ctccagcaac 2640
ttcatgtcca tgggtgccct taccgacctg gggcagaaca tgctttatgc caactcagcc 2700
catgcgctgg acatgacttt tgaggtggat cccatggatg agcccacact gctttatctt 2760
ctttttgaag tcttcgacgt ggtcagagtg caccagccac accgcggcgt catcgaggct 2820
gtctacctgc gtaccccatt ctcagctggt aacgccacca cataa                 2865
```

<210> 98
<211> 954
<212> PRT
<213> Chimpanzee Adenovirus ChAd 8

<400> 98

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
```

```
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
                20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Asn Leu Gly Asn Lys Phe Arg Asn Pro
            35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
        50                  55                  60
Met Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Phe Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Ile Ala Lys Gly Ser Pro Val Gln Asp
    130                 135                 140
Asp Ala Glu Gln Ala Gln Glu Gln Lys Asp Val Thr Tyr Thr Phe Gly
145                 150                 155                 160
Asn Ala Pro Val Lys Ala Glu Asp Asp Ile Thr Lys Asp Gly Leu Glu
                165                 170                 175
Val Gly Ile Gln Ile Ile Gly Asp Glu Glu Asn Pro Ile Tyr Ala Asp
            180                 185                 190
Lys Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Glu Gln Trp His Asp
        195                 200                 205
Thr Thr Gly Thr Thr Glu Gln Tyr Gly Gly Arg Ala Leu Lys Pro Ala
    210                 215                 220
Thr Asn Met Arg Pro Cys Tyr Gly Ser Phe Ala Arg Pro Thr Asn Lys
225                 230                 235                 240
Lys Gly Gly Gln Ala Lys Thr Arg Lys Val Glu Lys Thr Glu Gly Asp
                245                 250                 255
Lys Lys Thr Glu Val Glu Glu Leu Asp Ile Asp Met Asp Phe Tyr Asp
            260                 265                 270
Ala Arg Ser Lys Lys Gln Gly Tyr Asp Pro Gln Ile Val Leu Tyr Ser
        275                 280                 285
Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile Val Tyr Lys Pro
    290                 295                 300
Gly Thr Asp Glu Thr Ser Ser Ser Thr Asn Leu Gly Gln Gln Ala Met
305                 310                 315                 320
Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu
                325                 330                 335
Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala
            340                 345                 350
Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu
        355                 360                 365
Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe
    370                 375                 380
Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile
385                 390                 395                 400
Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro
                405                 410                 415
Leu Asp Gly Val Gly Pro Ile Thr Glu Thr Tyr Gln Gly Ile Lys Pro
            420                 425                 430
Lys Thr Ala Asp Asn Ala Asn Asp Gln Trp Glu Lys Asn Thr Glu Val
        435                 440                 445
```

```
Asn Gly Ala Asn Glu Ile Gly Lys Gly Asn Asn Tyr Ala Met Glu Ile
    450                 455                 460
Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala
465                 470                 475                 480
Leu Tyr Leu Pro Asp Gly Tyr Lys Tyr Thr Pro Ala Asn Val Thr Leu
                485                 490                 495
Pro Asp Asn Lys Asn Thr Tyr Gly Tyr Ile Asn Gly Arg Val Val Ser
            500                 505                 510
Pro Ser Leu Val Asp Ser Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu
        515                 520                 525
Asp Leu Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly
    530                 535                 540
Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe
545                 550                 555                 560
His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu
                565                 570                 575
Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn
            580                 585                 590
Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala
        595                 600                 605
Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met
    610                 615                 620
Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr
625                 630                 635                 640
Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr
                645                 650                 655
Pro Ile Pro Ala Asn Ala Thr Asn Ile Pro Ile Ser Ile Pro Ser Arg
            660                 665                 670
Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys
        675                 680                 685
Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser
    690                 695                 700
Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe
705                 710                 715                 720
Lys Lys Val Ser Ile Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn
                725                 730                 735
Asp Arg Leu Leu Cys Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp
            740                 745                 750
Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe
        755                 760                 765
Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr
    770                 775                 780
Ile Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe
785                 790                 795                 800
Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys Glu Tyr
                805                 810                 815
Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly
            820                 825                 830
Tyr His Ala Pro Thr Leu Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr
        835                 840                 845
Pro Tyr Pro Leu Ile Gly Thr Thr Ala Val Thr Ser Val Thr Gln Lys
    850                 855                 860
Lys Phe Leu Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn
865                 870                 875                 880
```

```
Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr
                885                 890                 895
Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met
            900                 905                 910
Asp Glu Pro Thr Leu Leu Tyr Leu Leu Phe Glu Val Phe Asp Val Val
        915                 920                 925
Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg
    930                 935                 940
Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950
```

<210> 99
<211> 2871
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 22 Hexon

<400> 99

```
atggccaccc catcgatgct gccccagtgg gcatacatgc acatcgccgg acaggatgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcaat 120
ctggggaaca agtttaggaa ccctaccgtg gcgcccaccc atgatgtgac caccgaccgc 180
agtcaacggc tgatgctccg ctttgtgccc gttgaccggg aggacaatac ctactcatac 240
aaagttcgat acaccttggc tgtgggcgac aacagagtgc tggatatggc cagtactttc 300
tttgacattc ggggtgtgtt ggatagaggc cctagcttca agccatattc tggcactgct 360
tacaactcat tggcccctaa gggcgctccc aatacatctc agtggattgc tgaaggcgta 420
aaaaaagaaa atggggaagc tgacaatgaa gcagctgtcg aagaggaaga ggaagagaaa 480
aatcttacca cttacacttt tggaaatgcc ccagtgaaag cagaaggtgg tgatatcact 540
aaagacaaag gtcttccaat tggttcagaa attacagacg gcgaagccaa accaatttat 600
gcagataaac tataccaacc agaacctcag gtgggagagg aaacttggac tgacacagat 660
ggaacaactg agaagtatgg tggtagagct ctaaagccag aaactaaaat gaaaccctgc 720
tatgggtctt ttgctaaacc cactaacgtc aaaggcggac aggcaaaaca aaaaactact 780
gaacaactgc aaaaccagca ggttgaatat gatattgaca tgaacttttt tgatcaagcg 840
tcacagaaag caaacttcag tccaaaaatt gtgatgtatg cagaaaatgt agacttggaa 900
accccagaca ctcacgtggt gtacaaacct ggtacttcag aagaaagttc tcatgctaat 960
ctcggtcaac aatctatgcc caacagaccc aactacattg gctttagaga taactttatt 1020
ggacttatgt actacaacag tactggcaac atgggagtgc tggcaggtca agcatcccaa 1080
ttgaatgcgg tggttgactt gcaggacaga aacacagaac tatcatatca actactgctt 1140
gattctctgg gtgacagaac cagatacttc agcatgtgga atcaagcagt cgatagctat 1200
gatcctgatg tgcgcattat tgaaaatcat ggggtggaag atgagcttcc caactactgc 1260
tttccattgg atggagtagg ggtaccaaca actagttaca aaataattga accaaatgga 1320
gagggtgcag attggaaaga gcctgacata aatggaacaa gtgaaattgg acaaggaaat 1380
ctctttgcca tggaaattaa cctccaagct aatctctgga gaagttttct ttattccaat 1440
gtggctctgt atctcccaga ctcctacaaa tacaccccag ccaatgtcac tcttccaact 1500
aacaccaaca cttatgacta catgaatggg cgggtggttc ccccatccct agtggatacc 1560
tacgtaaaca ttggcgccag atggtctttg gatgccatgg acaatgtcaa ccccctttaac 1620
catcaccgca acgctggcct gcgataccgg tccatgcttt tgggcaatgg tcgctacgtg 1680
cctttccaca ttcaagtgcc tcagaaattc tttgctgtga agaacctgct gcttctaccc 1740
ggttcttaca cctacgagtg gaacttcaga aaggatgtga acatggtcct gcagagttcc 1800
cttggtaatg atctccgggt cgatggtgcc agcatcagtt ttaccagcat caatctctat 1860
gccaccttct tccccatggc ccacaacact gcctccaccc ttgaagccat gctgcgcaat 1920
gacaccaatg atcaatcatt caatgactac ctttctgcag ccaacatgct ctaccccatc 1980
ccggccaacg ctaccaacgt tcccatctcc attccctctc gtaactgggc cgccttcaga 2040
ggctggtcct tcaccagact caaaaccaaa gagactccct ctttgggatc agggttcgat 2100
ccctactttg tttactctgg ttctatacct tacctggatg gtaccttcta ccttaaccac 2160
acttttaaga aagtctctat catgtttgac tcttcagtca gctggcctgg taatgacaga 2220
ttgctaactc caaatgagtt cgaaatcaag cgcacagttg atggggaagg ctacaatgtg 2280
```

gcccaatgta acatgaccaa agactggttc ctggtccaga tgcttgccaa ctacaacatt 2340
ggataccagg gcttctacgt tcctgagggt tacaaggatc gcatgtactc cttcttcaga 2400
aacttccagc ccatgagtag acaggtggtt gatgagatta actacaaaga ctataaagct 2460
gtcgccgtac cctaccagca taataactct ggctttgtgg gttacatggc tcctaccatg 2520
cgtcagggtc aagcgtaccc tgctaactac ccataccccc taattggaac cactgcagta 2580
accagtgtca cccagaaaaa attcctgtgc gacaggacca tgtggcgcat cccattctct 2640
agcaacttca tgtccatggg tgcccttaca gacctgggac agaacttgct gtatgccaac 2700
tcagcccatg cgctggacat gacttttgag gtggatccca tggatgagcc caccctgctt 2760
tatcttcttt ttgaagtatt cgacgtggtc agagtgcacc aaccacaccg cggcgtcatc 2820
gaggccgtct acctgcgcac accgttctcg gctggtaacg ccaccacata a 2871

<210> 100
<211> 956
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 22 Hexon

<400> 100

Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
1 5 10 15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
20 25 30
Arg Ala Thr Asp Thr Tyr Phe Asn Leu Gly Asn Lys Phe Arg Asn Pro
35 40 45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
50 55 60
Met Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65 70 75 80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95
Ala Ser Thr Phe Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
100 105 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125
Ala Pro Asn Thr Ser Gln Trp Ile Ala Glu Gly Val Lys Lys Glu Asn
130 135 140
Gly Glu Ala Asp Asn Glu Ala Ala Val Glu Glu Glu Glu Glu Glu Lys
145 150 155 160
Asn Leu Thr Thr Tyr Thr Phe Gly Asn Ala Pro Val Lys Ala Glu Gly
165 170 175
Gly Asp Ile Thr Lys Asp Lys Gly Leu Pro Ile Gly Ser Glu Ile Thr
180 185 190
Asp Gly Glu Ala Lys Pro Ile Tyr Ala Asp Lys Leu Tyr Gln Pro Glu
195 200 205
Pro Gln Val Gly Glu Glu Thr Trp Thr Asp Thr Asp Gly Thr Thr Glu
210 215 220
Lys Tyr Gly Gly Arg Ala Leu Lys Pro Glu Thr Lys Met Lys Pro Cys
225 230 235 240
Tyr Gly Ser Phe Ala Lys Pro Thr Asn Val Lys Gly Gly Gln Ala Lys
245 250 255
Gln Lys Thr Thr Glu Gln Leu Gln Asn Gln Gln Val Glu Tyr Asp Ile
260 265 270
Asp Met Asn Phe Phe Asp Gln Ala Ser Gln Lys Ala Asn Phe Ser Pro
275 280 285
Lys Ile Val Met Tyr Ala Glu Asn Val Asp Leu Glu Thr Pro Asp Thr
290 295 300

```
His Val Val Tyr Lys Pro Gly Thr Ser Glu Glu Ser Ser His Ala Asn
305                 310                 315                 320
Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
                325                 330                 335
Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
            340                 345                 350
Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln
        355                 360                 365
Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly
    370                 375                 380
Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
385                 390                 395                 400
Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu
                405                 410                 415
Pro Asn Tyr Cys Phe Pro Leu Asp Gly Val Gly Val Pro Thr Thr Ser
            420                 425                 430
Tyr Lys Ile Ile Glu Pro Asn Gly Glu Gly Ala Asp Trp Lys Glu Pro
        435                 440                 445
Asp Ile Asn Gly Thr Ser Glu Ile Gly Gln Gly Asn Leu Phe Ala Met
    450                 455                 460
Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn
465                 470                 475                 480
Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Val
                485                 490                 495
Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val
            500                 505                 510
Val Pro Pro Ser Leu Val Asp Thr Tyr Val Asn Ile Gly Ala Arg Trp
        515                 520                 525
Ser Leu Asp Ala Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn
    530                 535                 540
Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val
545                 550                 555                 560
Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Val Lys Asn Leu
                565                 570                 575
Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp
            580                 585                 590
Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Val Asp
        595                 600                 605
Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe
    610                 615                 620
Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn
625                 630                 635                 640
Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met
                645                 650                 655
Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro
            660                 665                 670
Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys
        675                 680                 685
Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val
    690                 695                 700
Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His
705                 710                 715                 720
Thr Phe Lys Lys Val Ser Ile Met Phe Asp Ser Ser Val Ser Trp Pro
                725                 730                 735
```

```
Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr
                740                 745                 750
Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp
            755                 760                 765
Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly
        770                 775                 780
Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg
785                 790                 795                 800
Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Ile Asn Tyr Lys
                805                 810                 815
Asp Tyr Lys Ala Val Ala Val Pro Tyr Gln His Asn Asn Ser Gly Phe
                820                 825                 830
Val Gly Tyr Met Ala Pro Thr Met Arg Gln Gly Gln Ala Tyr Pro Ala
            835                 840                 845
Asn Tyr Pro Tyr Pro Leu Ile Gly Thr Thr Ala Val Thr Ser Val Thr
        850                 855                 860
Gln Lys Lys Phe Leu Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser
865                 870                 875                 880
Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Leu
                885                 890                 895
Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp
                900                 905                 910
Pro Met Asp Glu Pro Thr Leu Leu Tyr Leu Leu Phe Glu Val Phe Asp
            915                 920                 925
Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr
        930                 935                 940
Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955
```

<210> 101
<211> 2865
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 24 Hexon

<400> 101

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tcccagcgcc tgacgctgcg gttcatcccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggcgtgct ggaccgcggc cccaccttca agccctactc cggcaccgcc 360
tacaactccc tggcccccaa gggcgcgccc aacccatgcg agtgggatga ggctgctact 420
gcccttgaca ttgatttgaa cgcagaagaa gatgaagaag gcgatgaagc ccaaggggaa 480
gcagatcagc agaaaactca tgtatttggc caggcgccat actccggaca gaacattaca 540
aaagaaggca tacagatagg cattgatgct accagtcaag cccaaacacc tctatatgcc 600
gacaaaacat tccaacccga acctcaggtt ggagaatcac aatggaatga gacagagatt 660
agccatggag cgggacgggt gctaaaaaag accactctca tgaaaccttg ctatgggtca 720
tatgcaaggc ctactaatga gaacggaggt cagggcatcc tcctggaaca agatggaaag 780
aaagaaagtc aagtggaaat gcaatttttc tccactactc aggcagctgc gggtaattca 840
gataatccta ctccaaagct tgttttgtac agcgaggatg ttaacctgga aacaccagat 900
acacacattt catacatgcc cactaacaac gaaaccaatt caagagaact gttgggacaa 960
caggccatgc ccaacaggcc taattacatc ggcttcagag acaactttat cggtctcatg 1020
tactacaaca gcactggcaa catgggagtg cttgcaggtc aggcctctca gttgaatgca 1080
gtggtggact tgcaagacag aaacacagaa ctgtcctacc agctcttgct tgattccatg 1140
ggtgacagaa ccagatattt ctccatgtgg aatcaggcag tggacagtta tgacccagat 1200
```

```
gtcagaatta ttgaaaatca tggaaccgaa gacgagctcc ccaactactg ttttcctctg 1260
ggtggcataa tcaatacgga aacttttaca aaagtcaagc ctaaagctgg acaggacgct 1320
cagtgggaaa aagattcaga attttcagat aaaaatgaaa taagagtggg aaacaacttc 1380
gctatggaaa tcaacatcaa tgccaacctg tggaggaact tcctgtactc caacgtggcc 1440
ctgtacctgc cagacaagct taagtatact ccatccaatg tgcaaatttc caacaacccc 1500
aactcctacg attacatgaa caagcgagtg gtggccccgg ggctggtgga ctgctacatc 1560
aacctgggcg cgcggtggtc gctggactac atggacaacg tcaacccctt caaccaccac 1620
cgcaacgcgg gcctgcgcta ccgctccatg ctcctgggca acgggcgcta cgtgcccttc 1680
cacatccagg tgccccagaa gttctttgcc atcaagaacc tcctcctcct gccgggctcc 1740
tacacctacg agtggaactt caggaaggat gtcaacatgg tcctccagag ctctctgggc 1800
aacgatctca gggtggacgg ggccagcatc aagttcgaga gcatctgcct ctacgccacc 1860
ttcttcccca tggcccacaa cacggcctcc acgctcgagg ccatgctcag gaacgacacc 1920
aacgaccagt ccttcaatga ctacctctcc gccgccaaca tgctctaccc catccccgcc 1980
aacgccacca acgtccccat ctccatcccc tcgcgcaact gggcggcctt ccgcggctgg 2040
gccttcaccc gcctcaagac caaggagacc ccctccctgg gctcgggatt cgacccctac 2100
tacacctact cgggatccat tccctacctg gacggcacct tctacctcaa ccacactttc 2160
aagaaggtct cggtcacctt cgactcctcg gtcagctggc cgggcaacga ccgcctgctc 2220
acccccaacg agttcgagat caagcgctcg gtcgacgggg agggctacaa cgtggcccag 2280
tgcaacatga ccaaggactg gttcctggtc cagatgctgg ccaactacaa catcggctac 2340
cagggcttct acatcccaga gagctacaag gacaggatgt actccttctt caggaacttc 2400
cagcccatga gccggcaggt ggtggaccag accaagtaca aggactacca ggaggtgggc 2460
atcatccacc agcacaacaa ctcgggcttc gtgggctacc tcgcccccac catgcgcgag 2520
ggacaggcct accccgccaa cttcccctac ccgctcatag gcaagaccgc ggtcgacagc 2580
atcacccaga aaaagttcct ctgcgaccgc accctctggc gcatcccctt ctccagcaac 2640
ttcatgtcca tgggtgcgct cacggacctg ggccagaacc tgctctatgc caactccgcc 2700
cacgcgctcg acatgacctt cgaggtcgac cccatggacg agcccaccct tctctatgtt 2760
ctgttcgaag tctttgacgt ggtccgggtc caccagccgc accgcggcgt catcgagacc 2820
gtgtacctgc gcacgccctt ctcggccggc aacgccacca cctaa                  2865
```

<210> 102
<211> 954
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 24 Hexon

<400> 102

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Pro Cys Glu Trp Asp Glu Ala Ala Thr Ala Leu Asp Ile
    130                 135                 140
Asp Leu Asn Ala Glu Glu Asp Glu Glu Gly Asp Glu Ala Gln Gly Glu
145                 150                 155                 160
```

```
Ala Asp Gln Gln Lys Thr His Val Phe Gly Gln Ala Pro Tyr Ser Gly
                165                 170                 175
Gln Asn Ile Thr Lys Glu Gly Ile Gln Ile Gly Ile Asp Ala Thr Ser
            180                 185                 190
Gln Ala Gln Thr Pro Leu Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro
        195                 200                 205
Gln Val Gly Glu Ser Gln Trp Asn Glu Thr Glu Ile Ser His Gly Ala
    210                 215                 220
Gly Arg Val Leu Lys Lys Thr Thr Leu Met Lys Pro Cys Tyr Gly Ser
225                 230                 235                 240
Tyr Ala Arg Pro Thr Asn Glu Asn Gly Gly Gln Gly Ile Leu Leu Glu
                245                 250                 255
Gln Asp Gly Lys Lys Glu Ser Gln Val Glu Met Gln Phe Phe Ser Thr
            260                 265                 270
Thr Gln Ala Ala Ala Gly Asn Ser Asp Asn Pro Thr Pro Lys Leu Val
        275                 280                 285
Leu Tyr Ser Glu Asp Val Asn Leu Glu Thr Pro Asp Thr His Ile Ser
    290                 295                 300
Tyr Met Pro Thr Asn Asn Glu Thr Asn Ser Arg Glu Leu Leu Gly Gln
305                 310                 315                 320
Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe
                325                 330                 335
Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala
            340                 345                 350
Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn
        355                 360                 365
Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly Asp Arg Thr
    370                 375                 380
Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp
385                 390                 395                 400
Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu Pro Asn Tyr
                405                 410                 415
Cys Phe Pro Leu Gly Gly Ile Ile Asn Thr Glu Thr Phe Thr Lys Val
            420                 425                 430
Lys Pro Lys Ala Gly Gln Asp Ala Gln Trp Glu Lys Asp Ser Glu Phe
        435                 440                 445
Ser Asp Lys Asn Glu Ile Arg Val Gly Asn Asn Phe Ala Met Glu Ile
    450                 455                 460
Asn Ile Asn Ala Asn Leu Trp Arg Asn Phe Leu Tyr Ser Asn Val Ala
465                 470                 475                 480
Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Thr Pro Ser Asn Val Gln Ile
                485                 490                 495
Ser Asn Asn Pro Asn Ser Tyr Asp Tyr Met Asn Lys Arg Val Val Ala
            500                 505                 510
Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala Arg Trp Ser Leu
        515                 520                 525
Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly
    530                 535                 540
Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe
545                 550                 555                 560
His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu
                565                 570                 575
Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn
            580                 585                 590
```

Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala
595 600 605
Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr Phe Phe Pro Met
610 615 620
Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr
625 630 635 640
Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr
645 650 655
Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg
660 665 670
Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg Leu Lys Thr Lys
675 680 685
Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Tyr Thr Tyr Ser
690 695 700
Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe
705 710 715 720
Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn
725 730 735
Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Ser Val Asp
740 745 750
Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe
755 760 765
Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr
770 775 780
Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe
785 790 795 800
Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys Tyr Lys Asp Tyr
805 810 815
Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser Gly Phe Val Gly
820 825 830
Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Phe
835 840 845
Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser Ile Thr Gln Lys
850 855 860
Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro Phe Ser Ser Asn
865 870 875 880
Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Leu Leu Tyr
885 890 895
Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met
900 905 910
Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val
915 920 925
Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr Val Tyr Leu Arg
930 935 940
Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945 950

<210> 103
<211> 2841
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 26 Hexon

<400> 103

atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120

```
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggctcccaa gggagcaccc aattccagcc agtgggagca aaaaaagaca 420
ggcaataaca atggaaatgg cggcactgaa tctgttacct ttggtgtagc cgccatgggc 480
ggagagaata ttacaaaaga gggtcttcaa attggaagtg atgaaactaa aaccgataac 540
aaagaaattt atgcagacaa gacctaccaa cctgagcctc aaataggaga ggagaactgg 600
caagaaacat tctcttttta tggtggtaga gctcttaaaa aagatactaa gatgaaacca 660
tgctatggct cctttgctag gcccacaaat gaaaagggag gtcaggccaa atttaaagtt 720
caagacggtg tgcagactac agaatatgac atcgacctgg ctttctttga tattccaagc 780
accggcacag ggggcaatgg tacaaatgta aatgataagc cagacatggt catgtacact 840
gaaaatgtga atttggagac gccagatact catattgtgt acaaacctgg aacttcagat 900
gacagctcta aagccaactt gtgccagcag gccatgccaa acagacccaa ctacattggt 960
ttcagagaca actttattgg gctcatgtat tacaacagta ctggcaatat gggggtgctg 1020
gctggtcagg cctctcagct gaatgctgtg gtcgacttgc aagacagaaa caccgaactg 1080
tcttaccagc tcttgctcga ctctctgggt gacagaacca ggtatttcag tatgtggaac 1140
caggcggtgg acagttatga ccctgatgtg cgcattattg aaaaccatgg tgtggaggat 1200
gaattgccaa actattgctt ccccttggat ggagctggca ctaatgctgt ataccagggt 1260
gttaaagcaa aagataacgg aaacgcagcc aatggaaact gggaacaaga cacaggcgtt 1320
tcaagtatta accagatatg caaggggaac atctatgcca tggaaatcaa cattcaagcc 1380
aacctgtgga gaagtttcct ttactcgaac gtggccctgt acctgcccga ctcttacaag 1440
tacacgccgg ccaacatcac cctgcccacc aacaccaaca cctacgatta catgaacggt 1500
cgggtggtgc ctccctcgct ggtggacgcc tacatcaaca tcggggcgcg ctggtcgctg 1560
gaccccatgg acaacgtcaa tcccttcaac caccaccgca acgcgggcct gcgctaccgc 1620
tccatgctcc tgggcaacgg gcgctacgtg cccttccaca tccaggtgcc ccagaaattt 1680
ttcgccatca agagcctcct gctcctgccc gggtcctaca cctacgagtg gaacttccgc 1740
aaggacgtca acatgatcct gcagagctcc ctcggcaacg acctgcgcac ggacggggcc 1800
tccatctcct tcaccagcat caacctctac gccaccttct tccccatggc gcacaacacg 1860
gcttccacgc tcgaggccat gctgcgcaac gacaccaacg accagtcctt caacgactac 1920
ctctcggcgg ccaacatgct ctaccccatc ccggccaacg ccaccaacgt gcccatctcc 1980
atcccctcgc gcaactgggc cgccttccgc ggctggtcct tcacgcgtct caagaccaag 2040
gaaacgccct cgctgggctc cgggttcgac ccctacttcg tctactcggg ctccatcccc 2100
tacctcgacg gcaccttcta cctcaaccac accttcaaga aggtctccat caccttcgac 2160
tcctccgtca gctggcccgg caacgaccgg ctcctgacgc ccaacgagtt cgaaatcaag 2220
cgcaccgtcg acggcgaggg ctacaacgtg gcccagtgca acatgaccaa ggactggttc 2280
ctggtccaga tgctggccca ctacaacatc ggctaccagg gcttctacgt gcccgagggc 2340
tacaaggacc gcatgtactc cttcttccgc aacttccagc ccatgagccg ccaggtggtg 2400
gacgaggtca actacaagga ctaccaggcc gtcaccctgg cctaccagca caacaactcg 2460
ggcttcgtcg gctacctcgc gcccaccatg cgccagggcc agccctaccc cgccaactac 2520
ccctacccgc tcatcggcaa gagcgccgtc accagcgtca cccagaaaaa gttcctctgc 2580
gacagggtca tgtggcgcat ccccttctcc agcaacttca tgtccatggg cgcgctcacc 2640
gacctcggcc agaacatgct ctatgccaac tccgcccacg cgctagacat gaatttcgaa 2700
gtcgacccca tggatgagtc caccttctc tatgttgtct tcgaagtctt cgacgtcgtc 2760
cgagtgcacc agccccaccg cggcgtcatc gaggccgtct acctgcgcac ccccttctcg 2820
gccggcaacg ccaccaccta a                                          2841
```

<210> 104
<211> 946
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 26 Hexon

<400> 104

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
```

```
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
                20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
            35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
        50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Lys Lys Thr Gly Asn Asn Asn
    130                 135                 140
Gly Asn Gly Gly Thr Glu Ser Val Thr Phe Gly Val Ala Ala Met Gly
145                 150                 155                 160
Gly Glu Asn Ile Thr Lys Glu Gly Leu Gln Ile Gly Ser Asp Glu Thr
                165                 170                 175
Lys Thr Asp Asn Lys Glu Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu
            180                 185                 190
Pro Gln Ile Gly Glu Glu Asn Trp Gln Glu Thr Phe Ser Phe Tyr Gly
        195                 200                 205
Gly Arg Ala Leu Lys Lys Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser
    210                 215                 220
Phe Ala Arg Pro Thr Asn Glu Lys Gly Gly Gln Ala Lys Phe Lys Val
225                 230                 235                 240
Gln Asp Gly Val Gln Thr Thr Glu Tyr Asp Ile Asp Leu Ala Phe Phe
                245                 250                 255
Asp Ile Pro Ser Thr Gly Thr Gly Gly Asn Gly Thr Asn Val Asn Asp
            260                 265                 270
Lys Pro Asp Met Val Met Tyr Thr Glu Asn Val Asn Leu Glu Thr Pro
        275                 280                 285
Asp Thr His Ile Val Tyr Lys Pro Gly Thr Ser Asp Asp Ser Ser Lys
    290                 295                 300
Ala Asn Leu Cys Gln Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly
305                 310                 315                 320
Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn
                325                 330                 335
Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp
            340                 345                 350
Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser
        355                 360                 365
Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp
    370                 375                 380
Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp
385                 390                 395                 400
Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asp Gly Ala Gly Thr Asn Ala
                405                 410                 415
Val Tyr Gln Gly Val Lys Ala Lys Asp Asn Gly Asn Ala Ala Asn Gly
            420                 425                 430
Asn Trp Glu Gln Asp Thr Gly Val Ser Ser Ile Asn Gln Ile Cys Lys
        435                 440                 445
```

```
Gly Asn Ile Tyr Ala Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg
    450                 455                 460
Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys
465                 470                 475                 480
Tyr Thr Pro Ala Asn Ile Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp
                485                 490                 495
Tyr Met Asn Gly Arg Val Val Pro Pro Ser Leu Val Asp Ala Tyr Ile
            500                 505                 510
Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro
        515                 520                 525
Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu
    530                 535                 540
Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe
545                 550                 555                 560
Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu
                565                 570                 575
Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly
            580                 585                 590
Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn
        595                 600                 605
Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu
    610                 615                 620
Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr
625                 630                 635                 640
Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn
                645                 650                 655
Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp
            660                 665                 670
Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly
        675                 680                 685
Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly
    690                 695                 700
Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp
705                 710                 715                 720
Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu
                725                 730                 735
Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln
            740                 745                 750
Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr
        755                 760                 765
Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg
    770                 775                 780
Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val
785                 790                 795                 800
Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln
                805                 810                 815
His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln
            820                 825                 830
Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser
        835                 840                 845
Ala Val Thr Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met
    850                 855                 860
Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr
865                 870                 875                 880
```

```
Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp
                885                 890                 895
Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val
            900                 905                 910
Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly
        915                 920                 925
Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala
    930                 935                 940
Thr Thr
945
```

<210> 105
<211> 2838
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 30 Hexon

<400> 105

```
atggccaccc catcgatgct gccccagtgg gcatacatgc acatcgccgg acaggatgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgtg ccacagacac ctacttcaat 120
ctggggaaca agtttaggaa ccccaccgtg gcacccaccc acgatgtgac caccgaccga 180
agccagcggc tgatgctgcg ctttgtgccc gttgatcgtg aggacaatac ttactcgtac 240
aaagttcgtt acacactggc tgtgggcgac aacagagtgc tagatatggc cagcaccttc 300
tttgacatca gaggggtgct tgacagaggt cccagcttca agccctactc tggcacagct 360
tacaactctc tggctcctaa gggagctcct aatcctagcc aatggctaga acaatcaacg 420
acagagggag aagacgatcc aactaacact acacacacat ttggaatagc ttctatgaaa 480
ggagaaaaca tcaccaaaga aggtttgcaa attggaaaag aagtaaccac tactggagat 540
aaaccaattt atgcagataa aacctttcag ccagaacccc aagtgggaga agagacttgg 600
actgatactg acgggacaaa tgaaaagttt ggcggtagaa ctcttaaaag tgctactaat 660
atgaaaccat gctatgggtc ttttgctagg cccacaaaca aacaaggcgg acaagctaaa 720
accagaaaag tagcggcagt tgatgggggg gaggaaactg aagaaccaga catcgacatg 780
gtgttttatg atgatagagg tgctacagaa gccatgatgg ctcctgaagt tgtactttat 840
gcagaaaatg taaatctgga aacaccagac acccatgtgg tgtacaaacc aggaacctct 900
gatattaatt ctcatgaaaa tttgggtcag caggctatgc caaacaggcc caattacatt 960
ggattcagag ataactttgt tggactcatg tactacaata gtactggcaa tatgggtgtg 1020
ttggcagggc aggcatcaca gctaaatgca gtagttgact tgcaagacag aaacactgag 1080
ctatcgtacc agctcttact tgattccctg ggcgacagaa ctcgatattt cagcatgtgg 1140
aaccaggctg tggatagcta tgaccctgat gtgcgcatta ttgaaaatca tggtatcgaa 1200
gatgaactac caaattactg cttccctctt gatggaatag gaccaggtaa aacataccaa 1260
ggtattaaag aaaaacaagg tgatgaggcc aacaaatggg aacaagacaa aacctatgcc 1320
acctctaatg aaatagccat aggtaataac ctggctatgg aaattaatat ccaggctaac 1380
ctttggagaa gttttctgta ctccaacgtg gctctgtacc ttccagacgc ttacaagtac 1440
acgccggcca acattacttt acctgccaat accaacacct atgaatacat gaacgggcga 1500
gtggtggcac catctttggt tgattcctac atcaacattg gtgccaggtg gtctcttgac 1560
ccaatggaca atgtgaaccc cttcaatcac caccgcaacg ctgggctgcg ttacagatcc 1620
atgcttctgg gcaatggtcg ctatgtgcct ttccacatcc aagtgcctca aaaattcttt 1680
gctatcaaaa acctgcttct cctccccgga tcctacacct atgagtggaa cttcagaaag 1740
gacgtaaaca tggtcctgca gagttccctt ggtaatgatc tcagaactga tggtgctagc 1800
attagtttta ccagcatcaa cctctatgcc acctttttcc caatggctca caacactgct 1860
tccacacttg aagccatgct gcgcaatgac accaatgacc agtcattcaa tgactacctt 1920
tctgcagcta acatgctcta cccaattcca gcaaatgcta ccaacattcc catttccatt 1980
ccctctcgca actgggctgc cttcaggggc tggtcattca ccagactcaa aaccaaggag 2040
actccctctt tgggatcagg ctttgatccc tactttgttt actctggctc cattccctac 2100
ctggatggta ccttctacct caaccacact ttcaaaaagg tttccattat gtttgattcc 2160
tcagtcagct ggccgggcaa tgacagattg ctaactccaa atgagttcga aatcaagcgc 2220
actgtggatg gggaagggta caatgtagct caatgcaaca tgaccaagga ctggttcctg 2280
```

```
gttcagatgc ttgccaacta taacattggc taccagggct tctacatccc agaggggtac 2340
aaggatcgca tgtattcctt cttcagaaac ttccagccca tgagcagaca ggtggttgat 2400
gaagttaatt acaaggagta ccaagccgtc acacttgctt accaacacaa caactctggc 2460
tttgtgggtt accttgcacc cactatgagg cagggagaac cttaccccgc taactaccca 2520
taccccctaa tcggaaccac tgctgttaag agtgttaccc acaaaaagtt cctgtgcgac 2580
aggaccatgt ggcgcatccc cttctccagc aacttcatgt ccatgggtgc ccttaccgac 2640
ctgggacaga acatgcttta tgccaactca tcccatgcgc tggacatgac ttttgaggtg 2700
gatcccatgg atgagcccac cctgctttat cttcttttcg aagttttcga cgtggtcaga 2760
gcgcaccagc cacaccgcgg cgtcatcgag gctgtctacc tgcgtactcc attctcagct 2820
ggtaacgcca ccacataa                                              2838
```

<210> 106
<211> 945
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 30 Hexon

<400> 106

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Asn Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Met Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Phe Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Pro Ser Gln Trp Leu Glu Gln Ser Thr Thr Glu Gly Glu
    130                 135                 140
Asp Asp Pro Thr Asn Thr Thr His Thr Phe Gly Ile Ala Ser Met Lys
145                 150                 155                 160
Gly Glu Asn Ile Thr Lys Glu Gly Leu Gln Ile Gly Lys Glu Val Thr
                165                 170                 175
Thr Thr Gly Asp Lys Pro Ile Tyr Ala Asp Lys Thr Phe Gln Pro Glu
            180                 185                 190
Pro Gln Val Gly Glu Glu Thr Trp Thr Asp Thr Asp Gly Thr Asn Glu
        195                 200                 205
Lys Phe Gly Gly Arg Thr Leu Lys Ser Ala Thr Asn Met Lys Pro Cys
    210                 215                 220
Tyr Gly Ser Phe Ala Arg Pro Thr Asn Lys Gln Gly Gly Gln Ala Lys
225                 230                 235                 240
Thr Arg Lys Val Ala Ala Val Asp Gly Gly Glu Glu Thr Glu Glu Pro
                245                 250                 255
Asp Ile Asp Met Val Phe Tyr Asp Asp Arg Gly Ala Thr Glu Ala Met
            260                 265                 270
Met Ala Pro Glu Val Val Leu Tyr Ala Glu Asn Val Asn Leu Glu Thr
        275                 280                 285
Pro Asp Thr His Val Val Tyr Lys Pro Gly Thr Ser Asp Ile Asn Ser
    290                 295                 300
```

```
His Glu Asn Leu Gly Gln Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile
305                 310                 315                 320
Gly Phe Arg Asp Asn Phe Val Gly Leu Met Tyr Tyr Asn Ser Thr Gly
                325                 330                 335
Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val
            340                 345                 350
Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp
        355                 360                 365
Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val
    370                 375                 380
Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Ile Glu
385                 390                 395                 400
Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asp Gly Ile Gly Pro Gly
                405                 410                 415
Lys Thr Tyr Gln Gly Ile Lys Glu Lys Gln Gly Asp Glu Ala Asn Lys
            420                 425                 430
Trp Glu Gln Asp Lys Thr Tyr Ala Thr Ser Asn Glu Ile Ala Ile Gly
        435                 440                 445
Asn Asn Leu Ala Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg Ser
    450                 455                 460
Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Ala Tyr Lys Tyr
465                 470                 475                 480
Thr Pro Ala Asn Ile Thr Leu Pro Ala Asn Thr Asn Thr Tyr Glu Tyr
                485                 490                 495
Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp Ser Tyr Ile Asn
            500                 505                 510
Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe
        515                 520                 525
Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly
    530                 535                 540
Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe
545                 550                 555                 560
Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp
                565                 570                 575
Asn Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn
            580                 585                 590
Asp Leu Arg Thr Asp Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu
        595                 600                 605
Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu
    610                 615                 620
Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu
625                 630                 635                 640
Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Ile
                645                 650                 655
Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser
            660                 665                 670
Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe
        675                 680                 685
Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr
    690                 695                 700
Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Met Phe Asp Ser
705                 710                 715                 720
Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe
                725                 730                 735
```

```
Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys
            740                 745                 750
Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn
        755                 760                 765
Ile Gly Tyr Gln Gly Phe Tyr Ile Pro Glu Gly Tyr Lys Asp Arg Met
    770                 775                 780
Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp
785                 790                 795                 800
Glu Val Asn Tyr Lys Glu Tyr Gln Ala Val Thr Leu Ala Tyr Gln His
                805                 810                 815
Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly
            820                 825                 830
Glu Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Thr Thr Ala
        835                 840                 845
Val Lys Ser Val Thr His Lys Lys Phe Leu Cys Asp Arg Thr Met Trp
    850                 855                 860
Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp
865                 870                 875                 880
Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ser His Ala Leu Asp Met
                885                 890                 895
Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Leu Leu
            900                 905                 910
Phe Glu Val Phe Asp Val Val Arg Ala His Gln Pro His Arg Gly Val
        915                 920                 925
Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr
    930                 935                 940
Thr
945
```

<210> 107
<211> 2877
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 31 Hexon

<400> 107

```
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgg 180
tctcagcgcc tgacgctgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcggggtgct ggaccggggc cccactttca agccttactc tggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca attagaagaa 420
gcccaggccg ctgtggaaga cgaagaatta gaagatgaag acgaggaacc acaggatgag 480
gcacctgtga aaaaaaccca tgtatacgct caggctcccc tttctggaga agaaattact 540
aaaaacggtt tgcaaatagg gtcagataac acagaagccc agtctaagcc catatatgca 600
gatcctacat tccagcctga accccaaatc ggggaatccc agtggaatga ggcagatgct 660
acagttgccg gcggtagagt gctaaagaaa tccactccca tgaagccatg ctatggttcc 720
tatgcaagac ccacaaactc caatggaggt caaggtgtgc tggtggctga tgataagggg 780
gttcttcaat ctaaagttga attgcaattt ttttcaaata ctactactct taatcagcgg 840
gagggtaacg atacaaaacc aaaagtggtg ctgtatagcg aagatgtgca catggaaact 900
ccagacaccc acatttctta caagcccaca aaaagcgatg acaattcaaa aatcatgctg 960
ggtcagcagt ccatgcccaa cagacctaat tacatcggct tcagagacaa ctttatcggc 1020
ctcatgtatt acaatagcac tggcaacatg ggagtgcttg caggtcaggc ctctcagttg 1080
aatgcagtgg tggacttgca agacagaaac acagaactgt cctaccagct cttgcttgat 1140
tccatgggtg acagaaccag atacttttcc atgtggaatc aggcagtgga cagttatgac 1200
```

```
ccagatgtca gaattattga aaatcatgga actgaagacg agctccccaa ctattgtttc 1260
cctctgggcg gcataggggt aactgacact taccaggcca ttaaaaccaa tggcaatggt 1320
caagaaaacc caacctggga aaaagataca gagtttgcag accgcaatga aataggggtg 1380
ggaaacaatt ttgctatgga gatcaacctc agtgccaacc tgtggagaaa cttcctgtac 1440
tccaacgtgg cgctgtacct accagacaag cttaagtaca acccctccaa tgtggacatc 1500
tctgacaacc ccaacaccta cgattacatg aacaagcgag tggtggcccc ggggctggtg 1560
gactgctaca tcaacctggg cgcgcgctgg tcgctggact acatggacaa cgtcaacccc 1620
ttcaaccacc accgcaatgc gggcctgcgc taccgctcca tgctcctggg caacgggcgc 1680
tatgtgccct tccacatcca ggtgccccag aagttctttg ccatcaagaa cctcctcctc 1740
ctgccgggct cctacaccta cgagtggaac ttcaggaagg atgtcaacat ggtcctccag 1800
agctctctgg gtaacgatct cagggtggac ggggccagca tcaagttcga gagcatctgc 1860
ctctacgcca ccttcttccc catggcccac aacacggcct ccacgctcga ggccatgctc 1920
aggaacgaca ccaacgacca gtccttcaat gactaccttt ccgccgccaa catgctctac 1980
cccatacccg ccaacgccac caacgtcccc atctccatcc cctcgcgcaa ctgggcggcc 2040
ttccgcggct gggccttcac ccgcctcaag accaaggaga ccccctccct gggctcggga 2100
ttcgacccct actacaccta ctcgggctct attccctacc tggacggcac cttctacctc 2160
aaccacactt tcaagaaggt ctcggtcacc ttcgactcct cggtcagctg gccgggcaac 2220
gaccgtctgc tcacccccaa cgagttcgag atcaagcgct cggtcgacgg ggagggctac 2280
aacgtggccc agtgcaacat gaccaaggac tggttcctgg tccagatgct ggccaactac 2340
aacatcggct accagggctt ctacatccca gagagctaca aggacaggat gtactccttc 2400
ttcaggaact tccagcccat gagccggcag gtggtggacc agaccaagta caaggactac 2460
caggaggtgg gcatcatcca ccagcacaac aactcgggct tcgtgggcta cctcgccccc 2520
accatgcgcg agggacaggc ctaccccgcc aacttccct acccgctcat aggcaagacc 2580
gcggtcgaca gcatcaccca gaaaaagttc ctctgcgacc gcaccctctg gcgcatcccc 2640
ttctccagca acttcatgtc catgggtgcg ctctcggacc tgggccagaa cttgctctac 2700
gccaactccg cccacgccct cgacatgacc ttcgaggtcg accccatgga cgagcccacc 2760
cttctctatg ttctgttcga agtctttgac gtggtccggg tccaccagcc gcaccgcggc 2820
gtcatcgaga ccgtgtacct gcgtacgccc ttctcggccg gcaacgccac cacctaa    2877
```

<210> 108

<211> 958

<212> PRT

<213> Chimpanzee Adenovirus- ChAd 31 Hexon

<400> 108

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Cys Glu Trp Glu Gln Leu Glu Glu Ala Gln Ala Ala
    130                 135                 140
Val Glu Asp Glu Glu Leu Glu Asp Glu Asp Glu Glu Pro Gln Asp Glu
145                 150                 155                 160
```

```
Ala Pro Val Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu Ser Gly
                165                 170                 175
Glu Glu Ile Thr Lys Asn Gly Leu Gln Ile Gly Ser Asp Asn Thr Glu
            180                 185                 190
Ala Gln Ser Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro Glu Pro
        195                 200                 205
Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val Ala Gly
    210                 215                 220
Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr Gly Ser
225                 230                 235                 240
Tyr Ala Arg Pro Thr Asn Ser Asn Gly Gly Gln Gly Val Leu Val Ala
                245                 250                 255
Asp Asp Lys Gly Val Leu Gln Ser Lys Val Glu Leu Gln Phe Phe Ser
            260                 265                 270
Asn Thr Thr Thr Leu Asn Gln Arg Glu Gly Asn Asp Thr Lys Pro Lys
        275                 280                 285
Val Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr His
    290                 295                 300
Ile Ser Tyr Lys Pro Thr Lys Ser Asp Asp Asn Ser Lys Ile Met Leu
305                 310                 315                 320
Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp
                325                 330                 335
Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val
            340                 345                 350
Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp
        355                 360                 365
Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly Asp
    370                 375                 380
Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp
385                 390                 395                 400
Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu Pro
                405                 410                 415
Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr Gln
            420                 425                 430
Ala Ile Lys Thr Asn Gly Asn Gly Gln Glu Asn Pro Thr Trp Glu Lys
        435                 440                 445
Asp Thr Glu Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn Asn Phe
    450                 455                 460
Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe Leu Tyr
465                 470                 475                 480
Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn Pro Ser
                485                 490                 495
Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met Asn Lys
            500                 505                 510
Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala
        515                 520                 525
Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His
    530                 535                 540
Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
545                 550                 555                 560
Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
                565                 570                 575
Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
            580                 585                 590
```

```
Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
        595                 600                 605
Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr
    610                 615                 620
Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu
625                 630                 635                 640
Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
                645                 650                 655
Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
            660                 665                 670
Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg
        675                 680                 685
Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr
    690                 695                 700
Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
705                 710                 715                 720
Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser
                725                 730                 735
Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
            740                 745                 750
Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr
        755                 760                 765
Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr
    770                 775                 780
Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe
785                 790                 795                 800
Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys
                805                 810                 815
Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser
            820                 825                 830
Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr
        835                 840                 845
Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser
    850                 855                 860
Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro
865                 870                 875                 880
Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu Gly Gln
                885                 890                 895
Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu
            900                 905                 910
Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val
        915                 920                 925
Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr
    930                 935                 940
Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955
```

<210> 109
<211> 2856
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 37 Hexon

<400> 109

```
atggccaccc catcgatgct gccccagtgg gcatacatgc acatcgccgg acaggatgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgtg ccacagacac ctacttcaat 120
```

```
ctggggaaca agtttaggaa ccccaccgtg gcccctaccc acgatgtgac caccgaccgt 180
agccagcggc tgatgctgcg ctttgtgccc gttgatcggg aggacaatac ctactcttac 240
aaagttcgct acacactggc tgtgggcgac aacagagtgc tggacatggc cagcaccttc 300
tttgacatca ggggggtgct tgacagaggt cccagtttca agccatactc tggcacagct 360
tacaattccc tggcgcctaa gggcgcgccc aatacatgcc agtggattgc caaggggcg 420
cctgttaccg atcaagacaa tgaagaacag gaattaacag atgttactta cgcttttggc 480
aatgctccag tacaagcaga agccaaaatt acaaaagatg gtctgccagt aggtttggaa 540
attacagaag atgaacaaaa gtcaatttat gcagacaaat tgtatcagcc agagccccaa 600
attggcgatg aacaatggca tgacaccact ggcactaatg aacaatacgg cggcagagct 660
ctaaaaccgg ccaccaacat gaaaccatgt tatggctcat ttgccagacc cacaaataaa 720
aaaggcggtc aggctaaaac tagaaaaata gaaaaggaag agaatggagt taaaaccgta 780
actgaagaag ctgacattga tatggacttt tatgacttaa gatcacaaag agcaaatttt 840
gatcctaaaa ttgttcttta ttctgaaaat gtaaatttgg aaactccaga tacacatatt 900
gtgtataaac caggaacaga tgaaactagt tcctctgtta acttgggaca gcaggcaatg 960
cccaacagac ccaactacat tggttttagg gacaacttca ttggacttat gttttacaac 1020
agtaccggca acatgggcgt gctggccggg caagcttctc agttaaatgc tgtggttgac 1080
ttgcaggaca ggaacacaga actgtcctac cagctgctgc ttgactctct gggtgacaga 1140
accagatact ttagcatgtg gaatcaggcc gtggatagct atgacccaga cgtgcgcatt 1200
attgaaaacc acggtgtgga agacgaactt cctaactatt gttttccatt agatggagtg 1260
ggaccaatta cgggcactta tcaggggtt gagcctgatg gaaacaatgg aaactggaag 1320
aaaaacacaa acataaatgg agcaaatgaa attggcaagg gaaataacta tgctatggaa 1380
attaatctac aagctaacct ctggagaagt tttctatatt ccaatgtggc tctgtattta 1440
ccagacggtt acaaatatac cccagccaat gttacactgc cagaaaacaa aaacacctat 1500
ggctatataa acggacgagt agtatcccca tctttggtgg attcatacat caacattgga 1560
gccagatggt ctttggatct tatggacaat gtaaacccat tcaatcacca ccgcaatgca 1620
ggcctgcgtt accgttccat gcttttagga aatggtcgct atgtgccttt ccacatccaa 1680
gtgcctcaga aaatctttgc tgtcaagaac ctgttgcttc ttcccggctc ctacacctat 1740
gagtggaact tcagaaagga cgtaaacatg gtcctgcaaa gttcccttgg taatgatctc 1800
agaactgatg gtgctagcat cagttttacc agcatcaatc tatatgctac ctttttcccc 1860
atggcccaca acactgcttc cacccttgaa gccatgctgc gcaatgacac caatgaccag 1920
tcatttaatg actacctttc tgcagctaac atgctctacc ctattccagc caatgcaacc 1980
aacatcccca tttccattcc ctctcgcaat tgggccgcct tcaggggctg gtccttcacc 2040
agactcaaaa ccaaggagac cccatctctg ggatcagggt tcgatcccta ctttgtctat 2100
tctggttcta ttccctacct tgatggcacc ttctacctta accacacttt caagaaggtc 2160
tccatcatgt ttgactcctc agtcagctgg ccaggcaatg acaggcttct aactccaaat 2220
gagtttgaaa tcaaacgcac tgtggatggg gaagggtaca atgtggctca atgcaacatg 2280
accaaggact ggttcctggt tcaaatgctc gccaactaca acattggcta ccagggcttc 2340
tacatcccag aggggtacaa ggatcgcatg tactccttct tcagaaactt ccagcccatg 2400
agtaggcagg tggttgatga gatcaactac aaggagtacc aagctgtcac acttgcttac 2460
cagcacaaca actctggctt tgtgggttac catgcaccca ctctccgtca gggtcaacca 2520
tacccagcta actacccata cccgcttatt ggaaccactg ctgtcaccag cgtcacccag 2580
aaaaagttct tgtgcgacag gaccatgtgg cgcatcccct tctccagcaa cttcatgtcc 2640
atgggtgccc ttaccgacct ggggcagaac atgctttatg ctaactcagc tcatgcgctg 2700
gacatgactt ttgaggtgga tcccatggat gagcccacac tgctttatct tcttttcgaa 2760
gtcttcgacg tggtcagagt gcaccagcca caccgcggcg tcatcgaggc cgtctacctg 2820
cgcacaccgt tctcggccgg caacgccacc acataa                          2856
```

<210> 110
<211> 951
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 37 Hexon

<400> 110

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
```

```
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
                20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Asn Leu Gly Asn Lys Phe Arg Asn Pro
            35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
        50                  55                  60
Met Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Phe Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Ile Ala Lys Gly Ala Pro Val Thr Asp
    130                 135                 140
Gln Asp Asn Glu Glu Gln Glu Leu Thr Asp Val Thr Tyr Ala Phe Gly
145                 150                 155                 160
Asn Ala Pro Val Gln Ala Glu Ala Lys Ile Thr Lys Asp Gly Leu Pro
                165                 170                 175
Val Gly Leu Glu Ile Thr Glu Asp Glu Gln Lys Ser Ile Tyr Ala Asp
            180                 185                 190
Lys Leu Tyr Gln Pro Glu Pro Gln Ile Gly Asp Glu Gln Trp His Asp
        195                 200                 205
Thr Thr Gly Thr Asn Glu Gln Tyr Gly Gly Arg Ala Leu Lys Pro Ala
    210                 215                 220
Thr Asn Met Lys Pro Cys Tyr Gly Ser Phe Ala Arg Pro Thr Asn Lys
225                 230                 235                 240
Lys Gly Gly Gln Ala Lys Thr Arg Lys Ile Glu Lys Glu Glu Asn Gly
                245                 250                 255
Val Lys Thr Val Thr Glu Glu Ala Asp Ile Asp Met Asp Phe Tyr Asp
            260                 265                 270
Leu Arg Ser Gln Arg Ala Asn Phe Asp Pro Lys Ile Val Leu Tyr Ser
        275                 280                 285
Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile Val Tyr Lys Pro
    290                 295                 300
Gly Thr Asp Glu Thr Ser Ser Ser Val Asn Leu Gly Gln Gln Ala Met
305                 310                 315                 320
Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu
                325                 330                 335
Met Phe Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala
            340                 345                 350
Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu
        355                 360                 365
Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe
    370                 375                 380
Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile
385                 390                 395                 400
Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro
                405                 410                 415
Leu Asp Gly Val Gly Pro Ile Thr Gly Thr Tyr Gln Gly Val Glu Pro
            420                 425                 430
Asp Gly Asn Asn Gly Asn Trp Lys Lys Asn Thr Asn Ile Asn Gly Ala
        435                 440                 445
```

```
Asn Glu Ile Gly Lys Gly Asn Asn Tyr Ala Met Glu Ile Asn Leu Gln
    450                 455                 460
Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu
465                 470                 475                 480
Pro Asp Gly Tyr Lys Tyr Thr Pro Ala Asn Val Thr Leu Pro Glu Asn
                485                 490                 495
Lys Asn Thr Tyr Gly Tyr Ile Asn Gly Arg Val Val Ser Pro Ser Leu
            500                 505                 510
Val Asp Ser Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp Leu Met
        515                 520                 525
Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr
    530                 535                 540
Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln
545                 550                 555                 560
Val Pro Gln Lys Ile Phe Ala Val Lys Asn Leu Leu Leu Leu Pro Gly
                565                 570                 575
Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Val Leu
            580                 585                 590
Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ser
        595                 600                 605
Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn
    610                 615                 620
Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln
625                 630                 635                 640
Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro
                645                 650                 655
Ala Asn Ala Thr Asn Ile Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala
            660                 665                 670
Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro
        675                 680                 685
Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile
    690                 695                 700
Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val
705                 710                 715                 720
Ser Ile Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu
                725                 730                 735
Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly
            740                 745                 750
Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln
        755                 760                 765
Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Ile Pro Glu
    770                 775                 780
Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met
785                 790                 795                 800
Ser Arg Gln Val Val Asp Glu Ile Asn Tyr Lys Glu Tyr Gln Ala Val
                805                 810                 815
Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr His Ala
            820                 825                 830
Pro Thr Leu Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro
        835                 840                 845
Leu Ile Gly Thr Thr Ala Val Thr Ser Val Thr Gln Lys Lys Phe Leu
    850                 855                 860
Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser
865                 870                 875                 880
```

```
Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser
                885                 890                 895
Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro
            900                 905                 910
Thr Leu Leu Tyr Leu Leu Phe Glu Val Phe Asp Val Val Arg Val His
        915                 920                 925
Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe
    930                 935                 940
Ser Ala Gly Asn Ala Thr Thr
945                 950
```

<210> 111
<211> 2817
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 38 Hexon

<400> 111

```
atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacgccc tggcccctaa agccgctcct aatcctagcc agtgggagga aaccactact 420
ggaacggatg gaaatgctgc tactacaact acacatagtt ttggcctggc tgctatgaaa 480
ggggacaata ttacatctga cggtttgcaa ataggaacag atgctacttc tggagaggaa 540
aaacccatct atgcagataa actgtaccag ccagaacccc aaatagggga agagtcatgg 600
actgatactg atggaaccaa tgaaaaattc ggaggaagag tccttaaaaa ggacacaagc 660
atgaaaccct gctacgggtc atttgccaaa ccaaccaata acaaaggtgg tcaagcaaaa 720
caaaaggcaa ctgaaggaac cgctgtagaa tatgatgtag acatgaactt ttttgatggt 780
agagatgcag ctgctaactt tactccagaa gtagtgttgt atgctgaaaa tgtggatttg 840
gaaactccag atacacatat tgtatacaaa ccaggaacct cagatgtgag ttcccatgtt 900
aatttgggac agcaagccat gcccaacaga cccaactaca ttggtttcag agacaacttt 960
attggtctca tgtattacaa cagcactggc aatatggggg tgctggccgg tcaggcttct 1020
cagctgaatg ctgtggttga cttgcaagac agaaacactg aactgtccta ccagctgttg 1080
cttgactctt tgggtgacag aactcggtat ttcagtatgt ggaatcaggc ggtggacagc 1140
tatgatcctg atgtgcgcat cattgaaaat catggtatag aagatgagct tcccaactat 1200
tgctttccta ttgacgctgt gggtattact agaacttatc aaggtatcaa aacgcaaaat 1260
ggtcaaacta caacttggga aaaggacacc agtgttagta cggccaatga aataggcatt 1320
ggcaacaatc ttgccatgga aatcaacatc caagccaacc tgtggaggaa cttcctctac 1380
gccaacgtgg ccctgtacct gcccgattct tacaagtaca caccggccaa cgtcaccctg 1440
cccaccaata ccaacaccta cgattacatg aacggccggg tggtggcgcc ctcgctggtg 1500
gacgcctaca tcaacatcgg ggcgcgctgg tcgctggacc ccatggataa cgtgaatccc 1560
ttcaaccacc accgcaacgc ggggctgcgc taccgctcca tgctcctggg caacgggcgc 1620
tacgtgccct tccacatcca ggtgccccag aaatttttcg ccatcaagag cctcctgctc 1680
ctgcccgggt cctacaccta cgagtggaac ttccgcaagg acgtcaacat gatcctgcag 1740
agctccctcg gcaacgacct gcgcacggac ggggcctcca tctccttcac cagcatcaac 1800
ctctacgcca ccttcttccc catggcgcac aacacggcct ccacgctcga ggccatgctg 1860
cgcaacgaca ccaacgacca gtccttcaac gactacctct cggcggccaa catgctctac 1920
cccatcccgg ccaacgccac caacgtgccc atctccatcc cctcgcgcaa ctgggccgcc 1980
ttccgcggct ggtccttcac gcgcctcaag accaaggaga cgccctcgct gggctccggg 2040
ttcgaccccc acttcgtcta ctcgggctcc atcccctacc tcgacggcac cttctacctc 2100
aaccacacct tcaagaaggt ctccatcacc ttcgactcct ccgtcagctg gcccggcaac 2160
gaccggctcc tgacgcccaa cgagttcgaa atcaagcgca ccgtcgacgg cgagggctac 2220
aacgtggccc agtgcaacat gaccaaggac tggttcctgg tccagatgct ggcccactac 2280
```

```
aacatcggct accagggctt ctacgtgccc gagggctaca aggaccgcat gtactccttc 2340
ttccgcaact tccaacccat gagccgccag gtggtggacg aggtcaacta caaggactac 2400
caggccgtca ccctggccta ccagcacaac aactcgggct tcgtcggcta cctcgcgccc 2460
accatgcgcc agggccagcc ctaccccgcc aactaccctt acccgctcat cggcaagagc 2520
gccgtcacca gcgtcaccca gaaaaagttc ctctgcgaca gggtcatgtg gcgcatcccc 2580
ttctccagca acttcatgtc catgggcgcg ctcaccgacc tcggccagaa catgctctat 2640
gccaactccg cccacgcgct agacatgaat ttcgaagtcg accccatgga tgagtccacc 2700
cttctctatg ttgtcttcga agtcttcgac gtcgtccgag tgcaccagcc ccaccgcggc 2760
gtcatcgagg ccgtctacct gcgcaccccc ttctcggccg gtaacgccac cacctaa    2817
```

<210> 112
<211> 938
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 38 Hexon

<400> 112

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ala Leu Ala Pro Lys Ala
        115                 120                 125
Ala Pro Asn Pro Ser Gln Trp Glu Glu Thr Thr Thr Gly Thr Asp Gly
    130                 135                 140
Asn Ala Ala Thr Thr Thr Thr His Ser Phe Gly Leu Ala Ala Met Lys
145                 150                 155                 160
Gly Asp Asn Ile Thr Ser Asp Gly Leu Gln Ile Gly Thr Asp Ala Thr
                165                 170                 175
Ser Gly Glu Glu Lys Pro Ile Tyr Ala Asp Lys Leu Tyr Gln Pro Glu
            180                 185                 190
Pro Gln Ile Gly Glu Glu Ser Trp Thr Asp Thr Asp Gly Thr Asn Glu
        195                 200                 205
Lys Phe Gly Gly Arg Val Leu Lys Lys Asp Thr Ser Met Lys Pro Cys
    210                 215                 220
Tyr Gly Ser Phe Ala Lys Pro Thr Asn Asn Lys Gly Gly Gln Ala Lys
225                 230                 235                 240
Gln Lys Ala Thr Glu Gly Thr Ala Val Glu Tyr Asp Val Asp Met Asn
                245                 250                 255
Phe Phe Asp Gly Arg Asp Ala Ala Ala Asn Phe Thr Pro Glu Val Val
            260                 265                 270
Leu Tyr Ala Glu Asn Val Asp Leu Glu Thr Pro Asp Thr His Ile Val
        275                 280                 285
Tyr Lys Pro Gly Thr Ser Asp Val Ser Ser His Val Asn Leu Gly Gln
    290                 295                 300
```

```
Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe
305                 310                 315                 320
Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala
                325                 330                 335
Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn
            340                 345                 350
Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr
        355                 360                 365
Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp
    370                 375                 380
Val Arg Ile Ile Glu Asn His Gly Ile Glu Asp Glu Leu Pro Asn Tyr
385                 390                 395                 400
Cys Phe Pro Ile Asp Ala Val Gly Ile Thr Arg Thr Tyr Gln Gly Ile
                405                 410                 415
Lys Thr Gln Asn Gly Gln Thr Thr Thr Trp Glu Lys Asp Thr Ser Val
            420                 425                 430
Ser Thr Ala Asn Glu Ile Gly Ile Gly Asn Asn Leu Ala Met Glu Ile
        435                 440                 445
Asn Ile Gln Ala Asn Leu Trp Arg Asn Phe Leu Tyr Ala Asn Val Ala
    450                 455                 460
Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Val Thr Leu
465                 470                 475                 480
Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val Val Ala
                485                 490                 495
Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu
            500                 505                 510
Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly
        515                 520                 525
Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe
    530                 535                 540
His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu
545                 550                 555                 560
Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn
                565                 570                 575
Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala
            580                 585                 590
Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met
        595                 600                 605
Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr
    610                 615                 620
Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr
625                 630                 635                 640
Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg
                645                 650                 655
Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys
            660                 665                 670
Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser
        675                 680                 685
Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe
    690                 695                 700
Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn
705                 710                 715                 720
Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp
                725                 730                 735
```

```
Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe
                740                 745                 750
Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr
            755                 760                 765
Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe
        770                 775                 780
Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys Asp Tyr
785                 790                 795                 800
Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly
                805                 810                 815
Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr
                820                 825                 830
Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Thr Ser Val Thr Gln Lys
            835                 840                 845
Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser Ser Asn
        850                 855                 860
Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr
865                 870                 875                 880
Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp Pro Met
                885                 890                 895
Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp Val Val
                900                 905                 910
Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg
            915                 920                 925
Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
        930                 935
```

<210> 113
<211> 2781
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 44 Hexon

<400> 113

```
atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacctac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggcccccaa gggagcgccc aatcccagcc agtgggaaca aactgaaacc 420
aatgttaata aaacacacac cttcggaatg gcagccatga aaggagaggc tattgacaaa 480
aatggtctgc aaattggaac tgacgctgcg gatcaggata aaccaattta tgcagataaa 540
acattccagc ctgaacctca agtaggagag gaagactgga ttgacaaagc agatttttat 600
ggcggaagag ctcttaaaaa agataccaag atgaaaccat gctatggctc atttgccaaa 660
cctacaaatg tcaagggagg acaggcaacg cccaggacta aagcagatgg aactactgag 720
cctgatattg acatgaactt ctttgaccca accactatta acacaccaga tgtagtgttg 780
tatgctgaaa atgttgattt gcaaactcca gacacccata tagtttacaa agcaggaact 840
tcagatgaca gttctgaggt caatttggct cagcaagcta tgcctaacag gcccaactac 900
attggtttca gagacaactt tatcggactt atgtattaca atagcactgg caatatgggt 960
gtgctcgctg gtcaggcttc ccagctaaat gctgtggtgg acttgcaaga cagaaacacc 1020
gagctgtcct accagctctt gcttgactct ctgggtgaca gaaccaggta tttcagtatg 1080
tggaatcagg cggtggacag ctatgatcct gatgtgcgca ttattgaaaa ccatggtgtg 1140
gaggatgaac tgccaaacta ttgctttcct ttggatggtg tgggcactaa taccgcatac 1200
caaggcgtta aagttaagac aactaatgga aacgacacgt gggaaaaaga tgaaactgtt 1260
tatgagttta atcaaattgg aaaggggat atctatgcta tggaaatcaa cattcaagcc 1320
```

```
aacctgtgga gaagttttct ctactcgaac gtggccctgt acctgcccga ttcttacaag 1380
tacacgccgg ccaacgtcac cctgcccacc aacaccaaca cctacgatta catgaacggg 1440
agagtggtgc ctccctcgct ggtggacgcc tacatcaaca tcggggcgcg ctggtcgctg 1500
gaccccatgg acaacgtgaa ccccttcaac caccaccgca acgcggggct gcgctaccgc 1560
tccatgctcc tgggcaacgg gcgcttcgtg cccttccaca tccaggtgcc ccagaaattt 1620
ttcgccatca agagcctcct gctcctgcca gggtcctaca cctacgagtg gaacttccgc 1680
aaggacgtca acatgatcct gcagagctcc ctcggcaacg acctgcgcac ggacggggcc 1740
tccatctcct tcaccagtat caacctctac gccaccttct tccccatggc gcacaacacg 1800
gcctccacgc tcgaggccat gctgcgcaac gacaccaacg accagtcctt caacgactac 1860
ctctcggcgg ccaacatgct ctaccccatc ccggccaacg ccaccaacgt gcccatctcc 1920
atcccctcgc gcaactgggc cgccttccgc ggctggtcct tcacgcgcct caagaccaag 1980
gagacgccct cgctgggctc cgggttcgac ccctacttcg tctactcggg ctccatcccc 2040
tacctcgacg gcaccttcta cctcaaccac accttcaaga aggtctccat caccttcgac 2100
tcctccgtca gctggcccgg caacgaccgc ctcctgacgc ccaacgagtt cgaaatcaag 2160
cgcaccgtcg acggagaggg gtacaacgtg gcccagtgca acatgaccaa ggactggttc 2220
ctggtccaga tgctggccca ctacaacatc ggctaccagg gcttctacgt gcccgagggc 2280
tacaaggacc gcatgtactc cttcttccgc aacttccagc ccatgagccg ccaggtggtg 2340
gacgaggtca actacaagga ctaccaggcc gtcaccctgg cctaccagca caacaactcg 2400
ggcttcgtcg gctacctcgc gcccaccatg cgccagggcc agccctaccc cgccaactac 2460
ccgtacccgc tcatcggcaa gagcgccgtc accagcgtca cccagaaaaa gttcctctgc 2520
gacagggtca tgtggcgcat ccccttctcc agcaacttca tgtccatggg cgcgctcacc 2580
gacctcggcc agaacatgct ctatgccaac tccgcccacg cgctagacat gaatttcgaa 2640
gtcgacccca tggatgagtc caccttctc tatgttgtct tcgaagtctt cgacgtcgtc 2700
cgagtgcacc agccccaccg cggcgtcatc gaggccgtct acctgcgcac ccccttctcg 2760
gccggtaacg ccaccaccta a                                          2781
```

<210> 114
<211> 926
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 44 Hexon

<400> 114

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Pro Ser Gln Trp Glu Gln Thr Glu Thr Asn Val Asn Lys
    130                 135                 140
Thr His Thr Phe Gly Met Ala Ala Met Lys Gly Glu Ala Ile Asp Lys
145                 150                 155                 160
Asn Gly Leu Gln Ile Gly Thr Asp Ala Ala Asp Gln Asp Lys Pro Ile
                165                 170                 175
```

```
Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro Gln Val Gly Glu Glu Asp
            180                 185                 190
Trp Ile Asp Lys Ala Asp Phe Tyr Gly Gly Arg Ala Leu Lys Lys Asp
        195                 200                 205
Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys Pro Thr Asn Val
    210                 215                 220
Lys Gly Gly Gln Ala Thr Pro Arg Thr Lys Ala Asp Gly Thr Thr Glu
225                 230                 235                 240
Pro Asp Ile Asp Met Asn Phe Phe Asp Pro Thr Thr Ile Asn Thr Pro
                245                 250                 255
Asp Val Val Leu Tyr Ala Glu Asn Val Asp Leu Gln Thr Pro Asp Thr
            260                 265                 270
His Ile Val Tyr Lys Ala Gly Thr Ser Asp Asp Ser Ser Glu Val Asn
        275                 280                 285
Leu Ala Gln Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
    290                 295                 300
Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
305                 310                 315                 320
Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln
                325                 330                 335
Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly
            340                 345                 350
Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
        355                 360                 365
Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu
    370                 375                 380
Pro Asn Tyr Cys Phe Pro Leu Asp Gly Val Gly Thr Asn Thr Ala Tyr
385                 390                 395                 400
Gln Gly Val Lys Val Lys Thr Thr Asn Gly Asn Asp Thr Trp Glu Lys
                405                 410                 415
Asp Glu Thr Val Tyr Glu Phe Asn Gln Ile Gly Lys Gly Asp Ile Tyr
            420                 425                 430
Ala Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr
        435                 440                 445
Ser Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala
    450                 455                 460
Asn Val Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly
465                 470                 475                 480
Arg Val Val Pro Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala
                485                 490                 495
Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His
            500                 505                 510
Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
        515                 520                 525
Phe Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
    530                 535                 540
Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
545                 550                 555                 560
Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
                565                 570                 575
Thr Asp Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr
            580                 585                 590
Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu
        595                 600                 605
```

```
Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
    610                 615                 620
Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
625                 630                 635                 640
Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg
                645                 650                 655
Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr
            660                 665                 670
Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
        675                 680                 685
Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser
    690                 695                 700
Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
705                 710                 715                 720
Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr
                725                 730                 735
Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr
            740                 745                 750
Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe
        755                 760                 765
Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn
    770                 775                 780
Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser
785                 790                 795                 800
Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr
                805                 810                 815
Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Thr Ser
            820                 825                 830
Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro
        835                 840                 845
Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln
    850                 855                 860
Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu
865                 870                 875                 880
Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val
                885                 890                 895
Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala
            900                 905                 910
Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
        915                 920                 925
```

<210> 115
<211> 2877
<212> DNA
<213> Chimpanzee Adenovirus- ChAd 63 Hexon

<400> 115

```
atgtatgtcc gccgaccaga aggaggaaga ggcgcgtcgc cgagttgcaa gatggccacc 60
ccatcgatgc tgccccagtg ggcgtacatg cacatcgccg gacaggacgc ttcggagtac 120
ctgagtccgg gtctggtgca gttcgcccgc gccacagaca cctacttcag tctggggaac 180
aagtttagga accccacggt ggcgcccacg cacgatgtga ccaccgaccg cagccagcgg 240
ctgacgctgc gcttcgtgcc cgtggaccgc gaggacaaca cctactcgta caaagtgcgc 300
tacacgctgg ccgtgggcga caaccgcgtg ctggacatgg ccagcaccta ctttgacatc 360
cgcggcgtgc tggatcgggg ccccagcttc aaaccctact ccggcaccgc ctacaacagc 420
ctagctccca agggagcgcc caacacctca cagtggaagg attccgacag caaaatgcat 480
```

```
actttggag ttgctgccat gcccggtgtt gttggtaaaa aaatagaagc cgatggtctg 540
cctattggaa tagattcatc ctctggaact gacaccataa tttatgctga taaaactttc 600
caaccagagc cacaggttgg aagtgacagt tgggtcgaca ccaatggtgc agaggaaaaa 660
tatggaggta gagctcttaa ggacactaca aacatgaagc cctgctacgg ttcttttgcc 720
aggcctacca acaaagaagg tggacaggct aacataaaag attctgaaac tgccagcact 780
actcctaact atgatataga tttggcattc tttgacagca aaaatattgc agctaactac 840
gatccagata ttgtaatgta cacagaaaat gttgagttgc aaactccaga tactcatatt 900
gtgtttaagc caggaacttc agatgaaagt tcagaagcca atttgggcca gcaggccatg 960
cccaacagac ccaactacat cgggttcaga gacaacttta tcgggctcat gtactacaac 1020
agcactggca atatgggtgt actggctggt caggcctccc agctaaatgc tgtggtggac 1080
ttgcaggaca gaaacaccga actgtcctac cagctcttgc ttgactctct gggtgacaga 1140
accaggtatt tcagtatgtg gaatcaggcg gtggacagct atgaccccga tgtgcgcatt 1200
attgaaaatc acggtgtgga ggatgaactc cccaattatt gcttcccttt gaatggtgta 1260
ggctttacag atacttacca gggtgttaaa gttaagacag atacagccgc tactggtacc 1320
aatggaacgc agtgggacaa agatgatacc acagtcagca ctgccaatga gatccactca 1380
ggcaatcctt tcgccatgga gatcaacatc caggccaacc tgtggcggaa cttcctctac 1440
gcgaacgtgg cgctgtacct gcccgactcc tacaagtaca cgccggccaa catcacgctg 1500
ccgaccaaca ccaacaccta cgattacatg aacggccgcg tggtggcgcc ctcgctggtg 1560
gacgcctaca tcaacatcgg ggcgcgctgg tcgctggacc ccatggacaa cgtcaacccc 1620
ttcaaccacc accgcaacgc gggcctgcgc taccgctcca tgctcctggg caacgggcgc 1680
tacgtgccct tccacatcca ggtgccccaa aagtttttcg ccatcaagag cctcctgctc 1740
ctgcccgggt cctacaccta cgagtggaac ttccgcaagg acgtcaacat gatcctgcag 1800
agctccctcg gcaacgacct gcgcacggac ggggcctcca tcgccttcac cagcatcaac 1860
ctctacgcca ccttcttccc catggcgcac aacaccgcct ccacgctcga ggccatgctg 1920
cgcaacgaca ccaacgacca gtccttcaac gactacctct cggcggccaa catgctctac 1980
cccatcccgg ccaacgccac caacgtgccc atctccatcc cctcgcgcaa ctgggccgcc 2040
ttccgcggat ggtccttcac gcgcctcaag acccgcgaga cgccctcgct cggctccggg 2100
ttcgacccct acttcgtcta ctcgggctcc atcccctacc tcgacggcac cttctacctc 2160
aaccacacct tcaagaaggt ctccatcacc ttcgactcct ccgtcagctg gcccggcaac 2220
gaccgcctcc tgacgcccaa cgagttcgaa atcaagcgca ccgtcgacgg agagggatac 2280
aacgtggccc agtgcaacat gaccaaggac tggttcctgg tccagatgct ggcccactac 2340
aacatcggct accagggctt ctacgtgccc gagggctaca aggaccgcat gtactccttc 2400
ttccgcaact tccagcccat gagccgccag gtcgtggacg aggtcaacta caaggactac 2460
caggccgtca ccctggccta ccagcacaac aactcgggct tcgtcggcta cctcgcgccc 2520
accatgcgcc agggccagcc ctaccccgcc aactaccctt acccgctcat cggcaagagc 2580
gccgtcgcca gcgtcaccca gaaaaagttc ctctgcgacc gggtcatgtg gcgcatcccc 2640
ttctccagca acttcatgtc catgggcgcg ctcaccgacc tcggccagaa catgctctac 2700
gccaactccg cccacgcgct agacatgaat ttcgaagtcg accccatgga tgagtccacc 2760
cttctctatg ttgtcttcga agtcttcgac gtcgtccgag tgcaccagcc ccaccgcggc 2820
gtcatcgagg ccgtctacct gcgcacgccc ttctcggccg gcaacgccac cacctaa    2877
```

<210> 116
<211> 941
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 63 Hexon

<400> 116

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
```

```
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Ser Gln Trp Lys Asp Ser Asp Ser Lys Met His Thr
    130                 135                 140
Phe Gly Val Ala Ala Met Pro Gly Val Val Gly Lys Lys Ile Glu Ala
145                 150                 155                 160
Asp Gly Leu Pro Ile Gly Ile Asp Ser Ser Ser Gly Thr Asp Thr Ile
                165                 170                 175
Ile Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro Gln Val Gly Ser Asp
            180                 185                 190
Ser Trp Val Asp Thr Asn Gly Ala Glu Glu Lys Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Asp Thr Thr Asn Met Lys Pro Cys Tyr Gly Ser Phe Ala Arg
    210                 215                 220
Pro Thr Asn Lys Glu Gly Gly Gln Ala Asn Ile Lys Asp Ser Glu Thr
225                 230                 235                 240
Ala Ser Thr Thr Pro Asn Tyr Asp Ile Asp Leu Ala Phe Phe Asp Ser
                245                 250                 255
Lys Asn Ile Ala Ala Asn Tyr Asp Pro Asp Ile Val Met Tyr Thr Glu
            260                 265                 270
Asn Val Glu Leu Gln Thr Pro Asp Thr His Ile Val Phe Lys Pro Gly
        275                 280                 285
Thr Ser Asp Glu Ser Ser Glu Ala Asn Leu Gly Gln Gln Ala Met Pro
    290                 295                 300
Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met
305                 310                 315                 320
Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser
                325                 330                 335
Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser
            340                 345                 350
Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser
        355                 360                 365
Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile
    370                 375                 380
Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu
385                 390                 395                 400
Asn Gly Val Gly Phe Thr Asp Thr Tyr Gln Gly Val Lys Val Lys Thr
                405                 410                 415
Asp Thr Ala Ala Thr Gly Thr Asn Gly Thr Gln Trp Asp Lys Asp Asp
            420                 425                 430
Thr Thr Val Ser Thr Ala Asn Glu Ile His Ser Gly Asn Pro Phe Ala
        435                 440                 445
Met Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg Asn Phe Leu Tyr Ala
    450                 455                 460
Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn
465                 470                 475                 480
Ile Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg
                485                 490                 495
```

```
Val Val Ala Pro Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg
                500                 505                 510
Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg
            515                 520                 525
Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr
        530                 535                 540
Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser
545                 550                 555                 560
Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys
                565                 570                 575
Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr
            580                 585                 590
Asp Gly Ala Ser Ile Ala Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe
        595                 600                 605
Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg
    610                 615                 620
Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn
625                 630                 635                 640
Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile
                645                 650                 655
Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu
            660                 665                 670
Lys Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe
        675                 680                 685
Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn
    690                 695                 700
His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp
705                 710                 715                 720
Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg
                725                 730                 735
Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys
            740                 745                 750
Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln
        755                 760                 765
Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe
    770                 775                 780
Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr
785                 790                 795                 800
Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly
                805                 810                 815
Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro
            820                 825                 830
Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Ala Ser Val
        835                 840                 845
Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe
    850                 855                 860
Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn
865                 870                 875                 880
Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val
                885                 890                 895
Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe
            900                 905                 910
Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val
        915                 920                 925

Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935                 940
```

<210> 117
<211> 2811
<212> DNA

<213> Chimpanzee Adenovirus- ChAd 82 Hexon

<400> 117

```
atggccaccc catcgatgct gccccagtgg gcgtacatgc acatcgccgg acaggacgct 60
tcggagtacc tgagtccggg tctggtgcag ttcgcccgcg ccacagacac ctacttcagt 120
ctggggaaca agtttaggaa ccccacggtg gcgcccacgc acgatgtgac caccgaccgc 180
agccagcggc tgacgctgcg cttcgtgccc gtggaccgcg aggacaacac ctactcgtac 240
aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc tggacatggc cagcacttac 300
tttgacatcc gcggcgtgct ggaccggggc cctagcttca aaccctactc cggcaccgcc 360
tacaacagcc tggctcccaa gggagcgccc aattccagcc agtgggagca aaatgaaaac 420
aatggtcaag gtcaagctaa gacacacacc tatggtgttg ctgctatggg cggacttgat 480
attacaaaag agggtcttaa aattgtaact gatgctagta aggaagatga caatgaaatt 540
tatgcagata aaacatatca gcccgagcct caaataggag aggaaaattg gcaagacact 600
aaaaactttt atggaggcag agctcttaaa aaagatacca agatgaagcc atgctatggc 660
tcatttgcca gacctaccaa tgtgaaggga gggcaagcca aagtgaaaac agaagaaaat 720
gttcagtcat ttgacataga tctggctttc tttgatattc caagcaccgg cacaggggac 780
aatggtacaa atgtaaatga taagccagac atggttatgt acactgaaaa tgtgaatctg 840
gagacgccag atactcatat tgtgtacaaa cctggaactt cagatgacag ctctgaagcc 900
aacttgtgcc agcaggccat gccaaacaga cccaactaca ttggtttcag agacaacttt 960
attgggctca tgtattacaa cagtactggc aatatggggg tgctggctgg tcaggcctcc 1020
cagctgaatg ctgtggttga cttgcaagac agaaacaccg agctgtcata ccagctcttg 1080
cttgactctc tgggtgacag aacccggtat ttcagcatgt ggaaccaggc ggtggacagt 1140
tatgaccctg atgtgcgcat tattgaaaac catggtgtgg aggatgaatt gccaaactat 1200
tgcttccgct tggatggagc tggcactaat gctgtatacc ggggtgttaa agcaaaagat 1260
aacggaaact gggaacaaga cacaggcgtt tcaagtatta accagatatg caaggggaac 1320
atctatgcca tggaaatcaa cattcaagcc aacctgtgga gaagtttcct ttactcgaac 1380
gtggccctgt acctgcccga ctcttacaag tacacgccgg ccaacatcac cctgcccacc 1440
aacaccaaca cctacgatta catgaacggt cgggtggtgc ctccctcgct ggtggacgcc 1500
tacatcaaca tcggggcgcg ctggtcgctg gaccccatgg acaacgtcaa tcccttcaac 1560
caccaccgca acgcgggcct gcgctaccgc tccatgctcc tgggcaacgg gcgctacgtg 1620
cccttccaca tccaggtgcc ccagaaattt ttcgccatca agagcctcct gctcctgccc 1680
gggtcctaca cctacgagtg gaacttccgc aaggacgtca acatgatcct gcagagctcc 1740
ctcggcaacg acctgcgcac ggacggggcc tccatctcct tcaccagcat caacctctac 1800
gccaccttct tccccatggc gcacaacacg gcctccacgc tcgaggccat gctgcgcaac 1860
gacaccaacg accagtcctt caacgactac ctctcggcgg ccaacatgct ctaccccatc 1920
ccggccaacg ccaccaacgt gcccatctcc atcccctcgc gcaactgggc cgccttccgc 1980
ggctggtcct tcacgcgtct caagaccaag gagacgccct cgctgggctc cgggttcgac 2040
ccctacttcg tctactcggg ctccatcccc tacctcgacg gcaccttcta cctcaaccac 2100
accttcaaga aggtctccat caccttcgac tcctccgtca gctggcccgg caacgaccgg 2160
ctcctgacgc ccaacgagtt cgaaatcaag cgcaccgtcg acggcgaggg ctacaacgtg 2220
gcccagtgca acatgaccaa ggactggttc ctggtccaga tgctggccca ctacaacatc 2280
ggctaccagg gcttctacgt gcccgagggc tacaaggacc gcatgtactc cttcttccgc 2340
aacttccagc ccatgagccg ccaggtggtg gacgaggtca actacaagga ctaccaggcc 2400
gtcaccctgg cctaccagca caacaactcg ggcttcgtcg gctacctcgc gcccaccatg 2460
cgccaggggc agccctaccc cgccaactac ccgtacccgc tcatcggcaa gagcgccgtc 2520
accagcgtca cccagaaaaa gttcctctgc gacagggtca tgtggcgcat ccccttctcc 2580
agcaacttca tgtccatggg cgcgctcacc gacctcggcc agaacatgct ctatgccaac 2640
tccgcccacg cgctagacat gaatttcgaa gtcgacccca tggatgagtc caccttctc 2700
tatgttgtct tcgaagtctt cgacgtcgtc cgagtgcacc agccccaccg cggcgtcatc 2760

gaggccgtct acctgcgcac acccttctcg gccggtaacg ccaccaccta a          2811
```

<210> 118
<211> 936
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 82 Hexon

<400> 118

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Asn Glu Asn Asn Gly Gln Gly
    130                 135                 140
Gln Ala Lys Thr His Thr Tyr Gly Val Ala Ala Met Gly Gly Leu Asp
145                 150                 155                 160
Ile Thr Lys Glu Gly Leu Lys Ile Val Thr Asp Ala Ser Lys Glu Asp
                165                 170                 175
Asp Asn Glu Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro Gln Ile
            180                 185                 190
Gly Glu Glu Asn Trp Gln Asp Thr Lys Asn Phe Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Lys Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Arg
    210                 215                 220
Pro Thr Asn Val Lys Gly Gly Gln Ala Lys Val Lys Thr Glu Glu Asn
225                 230                 235                 240
Val Gln Ser Phe Asp Ile Asp Leu Ala Phe Phe Asp Ile Pro Ser Thr
                245                 250                 255
Gly Thr Gly Gly Asn Gly Thr Asn Val Asn Asp Lys Pro Asp Met Val
            260                 265                 270
Met Tyr Thr Glu Asn Val Asn Leu Glu Thr Pro Asp Thr His Ile Val
        275                 280                 285
Tyr Lys Pro Gly Thr Ser Asp Asp Ser Ser Glu Ala Asn Leu Cys Gln
    290                 295                 300
Gln Ala Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe
305                 310                 315                 320
Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala
                325                 330                 335
Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn
            340                 345                 350
Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr
        355                 360                 365
```

```
Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp
    370                 375                 380
Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr
385                 390                 395                 400
Cys Phe Pro Leu Asp Gly Ala Gly Thr Asn Ala Val Tyr Arg Gly Val
                405                 410                 415
Lys Ala Lys Asp Asn Gly Asn Trp Glu Gln Asp Thr Gly Val Ser Ser
            420                 425                 430
Ile Asn Gln Ile Cys Lys Gly Asn Ile Tyr Ala Met Glu Ile Asn Ile
        435                 440                 445
Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr
    450                 455                 460
Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Ile Thr Leu Pro Thr
465                 470                 475                 480
Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val Val Pro Pro Ser
                485                 490                 495
Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro
            500                 505                 510
Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg
        515                 520                 525
Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile
    530                 535                 540
Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro
545                 550                 555                 560
Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile
                565                 570                 575
Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile
            580                 585                 590
Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala His
        595                 600                 605
Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp
    610                 615                 620
Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile
625                 630                 635                 640
Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg Asn Trp
                645                 650                 655
Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr
            660                 665                 670
Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser
        675                 680                 685
Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys
    690                 695                 700
Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg
705                 710                 715                 720
Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu
                725                 730                 735
Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Val
            740                 745                 750
Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro
        755                 760                 765
Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro
    770                 775                 780
Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala
785                 790                 795                 800
```

```
Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr Leu
                805                 810                 815
Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr
            820                 825                 830
Pro Leu Ile Gly Lys Ser Ala Val Thr Ser Val Thr Gln Lys Lys Phe
        835                 840                 845
Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met
    850                 855                 860
Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn
865                 870                 875                 880
Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp Pro Met Asp Glu
                885                 890                 895
Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp Val Val Arg Val
            900                 905                 910
His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro
        915                 920                 925
Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935
```

<210> 119
<211> 933
<212> PRT
<213> Chimpanzee Adenovirus- CV23 Pan5 Hexon

<400> 119

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Thr Tyr Lys Ala Asp Gly Asp Thr Gly
    130                 135                 140
Thr Glu Lys Thr Tyr Thr Tyr Gly Asn Ala Pro Val Gln Gly Ile Ser
145                 150                 155                 160
Ile Thr Lys Asp Gly Ile Gln Leu Gly Thr Asp Thr Asp Asp Gln Pro
                165                 170                 175
Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Ala
            180                 185                 190
Glu Trp His Asp Ile Thr Gly Thr Asp Glu Lys Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Pro Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys
    210                 215                 220
Pro Thr Asn Lys Glu Gly Gly Gln Ala Asn Val Lys Thr Glu Thr Gly
225                 230                 235                 240
```

Gly Thr Lys Glu Tyr Asp Ile Asp Met Ala Phe Phe Asp Asn Arg Ser
245 250 255
Ala Ala Ala Ala Gly Leu Ala Pro Glu Ile Val Leu Tyr Thr Glu Asn
260 265 270
Val Asp Leu Glu Thr Pro Asp Thr His Ile Val Tyr Lys Ala Gly Thr
275 280 285
Asp Asp Ser Ser Ser Ser Ile Asn Leu Gly Gln Gln Ser Met Pro Asn
290 295 300
Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr
305 310 315 320
Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln
325 330 335
Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr
340 345 350
Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met
355 360 365
Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu
370 375 380
Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asp
385 390 395 400
Ala Val Gly Arg Thr Asp Thr Tyr Gln Gly Ile Lys Ala Asn Gly Ala
405 410 415
Asp Gln Thr Thr Trp Thr Lys Asp Asp Thr Val Asn Asp Ala Asn Glu
420 425 430
Leu Gly Lys Gly Asn Pro Phe Ala Met Glu Ile Asn Ile Gln Ala Asn
435 440 445
Leu Trp Arg Asn Phe Leu Tyr Ala Asn Val Ala Leu Tyr Leu Pro Asp
450 455 460
Ser Tyr Lys Tyr Thr Pro Ala Asn Ile Thr Leu Pro Thr Asn Thr Asn
465 470 475 480
Thr Tyr Asp Tyr Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp
485 490 495
Ala Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn
500 505 510
Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser
515 520 525
Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro
530 535 540
Gln Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr
545 550 555 560
Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser
565 570 575
Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ala Phe Thr
580 585 590
Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala
595 600 605
Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
610 615 620
Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
625 630 635 640
Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
645 650 655
Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Arg Glu Thr Pro Ser Leu
660 665 670

```
Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr
        675                 680                 685
Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
    690                 695                 700
Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
705                 710                 715                 720
Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
                725                 730                 735
Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu
            740                 745                 750
Ala His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
        755                 760                 765
Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
    770                 775                 780
Gln Val Val Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu
785                 790                 795                 800
Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr
                805                 810                 815
Met Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
            820                 825                 830
Gly Lys Ser Ala Val Ala Ser Val Thr Gln Lys Lys Phe Leu Cys Asp
        835                 840                 845
Arg Val Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
    850                 855                 860
Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
865                 870                 875                 880
Ala Leu Asp Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu
                885                 890                 895
Leu Tyr Val Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro
            900                 905                 910
His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
        915                 920                 925
Gly Asn Ala Thr Thr
    930
```

<210> 120
<211> 942
<212> PRT
<213> Chimpanzee Adenovirus- CV32 Pan6 Hexon

<400> 120

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asn Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Val Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
```

```
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Ser Gln Trp Glu Gln Ala Lys Thr Gly Asn Gly Gly
    130                 135                 140
Thr Met Glu Thr His Thr Tyr Gly Val Ala Pro Met Gly Gly Glu Asn
145                 150                 155                 160
Ile Thr Lys Asp Gly Leu Gln Ile Gly Thr Asp Val Thr Ala Asn Gln
                165                 170                 175
Asn Lys Pro Ile Tyr Ala Asp Lys Thr Phe Gln Pro Glu Pro Gln Val
            180                 185                 190
Gly Glu Glu Asn Trp Gln Glu Thr Glu Asn Phe Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Lys Asp Thr Asn Met Lys Pro Cys Tyr Gly Ser Tyr Ala Arg
    210                 215                 220
Pro Thr Asn Glu Lys Gly Gly Gln Ala Lys Leu Lys Val Gly Asp Asp
225                 230                 235                 240
Gly Val Pro Thr Lys Glu Phe Asp Ile Asp Leu Ala Phe Phe Asp Thr
                245                 250                 255
Pro Gly Gly Thr Val Asn Gly Gln Asp Glu Tyr Lys Ala Asp Ile Val
            260                 265                 270
Met Tyr Thr Glu Asn Thr Tyr Leu Glu Thr Pro Asp Thr His Val Val
        275                 280                 285
Tyr Lys Pro Gly Lys Asp Asp Ala Ser Ser Glu Ile Asn Leu Val Gln
    290                 295                 300
Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe
305                 310                 315                 320
Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala
                325                 330                 335
Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn
            340                 345                 350
Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr
        355                 360                 365
Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp
    370                 375                 380
Val Arg Ile Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr
385                 390                 395                 400
Cys Phe Pro Leu Asp Gly Ser Gly Thr Asn Ala Ala Tyr Gln Gly Val
                405                 410                 415
Lys Val Lys Asp Gly Gln Asp Gly Asp Val Glu Ser Glu Trp Glu Asn
            420                 425                 430
Asp Asp Thr Val Ala Ala Arg Asn Gln Leu Cys Lys Gly Asn Ile Phe
        435                 440                 445
Ala Met Glu Ile Asn Leu Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr
    450                 455                 460
Ser Asn Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Thr
465                 470                 475                 480
Asn Val Thr Leu Pro Thr Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly
                485                 490                 495
Arg Val Thr Pro Pro Ser Leu Val Asp Ala Tyr Leu Asn Ile Gly Ala
            500                 505                 510
Arg Trp Ser Leu Asp Pro Met Asp Asn Val Asn Pro Phe Asn His His
        515                 520                 525
Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
    530                 535                 540
```

```
Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
545                 550                 555                 560
Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
                565                 570                 575
Lys Asp Val Asn Met Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
            580                 585                 590
Thr Asp Gly Ala Ser Ile Ala Phe Thr Ser Ile Asn Leu Tyr Ala Thr
        595                 600                 605
Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu
    610                 615                 620
Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
625                 630                 635                 640
Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
                645                 650                 655
Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg
            660                 665                 670
Leu Lys Thr Arg Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr
        675                 680                 685
Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
    690                 695                 700
Asn His Thr Phe Lys Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser
705                 710                 715                 720
Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
                725                 730                 735
Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr
            740                 745                 750
Lys Asp Trp Phe Leu Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr
        755                 760                 765
Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe
    770                 775                 780
Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Val Asn
785                 790                 795                 800
Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser
                805                 810                 815
Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr
            820                 825                 830
Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly Lys Ser Ala Val Ala Ser
        835                 840                 845
Val Thr Gln Lys Lys Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro
    850                 855                 860
Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln
865                 870                 875                 880
Asn Met Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu
                885                 890                 895
Val Asp Pro Met Asp Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val
            900                 905                 910
Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala
        915                 920                 925
Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935                 940
```

<210> 121
<211> 932
<212> PRT
<213> Chimpanzee Adenovirus- CV33 Pan7 Hexon

<400> 121

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Thr Tyr Lys Ala Gly Asp Thr Asp Thr
    130                 135                 140
Glu Lys Thr Tyr Thr Tyr Gly Asn Ala Pro Val Gln Gly Ile Ser Ile
145                 150                 155                 160
Thr Lys Asp Gly Ile Gln Leu Gly Thr Asp Ser Asp Gly Gln Ala Ile
                165                 170                 175
Tyr Ala Asp Glu Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Ala Glu
            180                 185                 190
Trp His Asp Ile Thr Gly Thr Asp Glu Lys Tyr Gly Gly Arg Ala Leu
        195                 200                 205
Lys Pro Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys Pro
    210                 215                 220
Thr Asn Lys Glu Gly Gly Gln Ala Asn Val Lys Thr Glu Thr Gly Gly
225                 230                 235                 240
Thr Lys Glu Tyr Asp Ile Asp Met Ala Phe Phe Asp Asn Arg Ser Ala
                245                 250                 255
Ala Ala Ala Gly Leu Ala Pro Glu Ile Val Leu Tyr Thr Glu Asn Val
            260                 265                 270
Asp Leu Glu Thr Pro Asp Thr His Ile Val Tyr Lys Ala Gly Thr Asp
        275                 280                 285
Asp Ser Ser Ser Ser Ile Asn Leu Gly Gln Gln Ser Met Pro Asn Arg
    290                 295                 300
Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr
305                 310                 315                 320
Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
                325                 330                 335
Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln
            340                 345                 350
Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp
        355                 360                 365
Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
    370                 375                 380
His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asp Ala
385                 390                 395                 400
Val Gly Arg Thr Asp Thr Tyr Gln Gly Ile Lys Ala Asn Gly Asp Asn
                405                 410                 415
```

```
Gln Thr Thr Trp Thr Lys Asp Asp Thr Val Asn Asp Ala Asn Glu Leu
                420                 425                 430
Gly Lys Gly Asn Pro Phe Ala Met Glu Ile Asn Ile Gln Ala Asn Leu
            435                 440                 445
Trp Arg Asn Phe Leu Tyr Ala Asn Val Ala Leu Tyr Leu Pro Asp Ser
        450                 455                 460
Tyr Lys Tyr Thr Pro Ala Asn Ile Thr Leu Pro Thr Asn Thr Asn Thr
465                 470                 475                 480
Tyr Asp Tyr Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp Ala
                485                 490                 495
Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn Val
            500                 505                 510
Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met
        515                 520                 525
Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln
    530                 535                 540
Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr Thr
545                 550                 555                 560
Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser Ser
                565                 570                 575
Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ala Phe Thr Ser
            580                 585                 590
Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser
        595                 600                 605
Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn
    610                 615                 620
Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala
625                 630                 635                 640
Thr Asn Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg
                645                 650                 655
Gly Trp Ser Phe Thr Arg Leu Lys Thr Arg Glu Thr Pro Ser Leu Gly
            660                 665                 670
Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr Leu
        675                 680                 685
Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile Thr
    690                 695                 700
Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro
705                 710                 715                 720
Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val
                725                 730                 735
Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala
            740                 745                 750
His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys
        755                 760                 765
Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln
    770                 775                 780
Val Val Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu Ala
785                 790                 795                 800
Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met
                805                 810                 815
Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile Gly
            820                 825                 830
Lys Ser Ala Val Ala Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg
        835                 840                 845
```

```
Val Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala
    850                 855                 860
Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His Ala
865                 870                 875                 880
Leu Asp Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu Leu
                885                 890                 895
Tyr Val Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro His
            900                 905                 910
Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly
        915                 920                 925
Asn Ala Thr Thr
    930
```

<210> 122
<211> 960
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 3 Hexon

<400> 122

```
Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65                  70                  75                  80
Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Ser Cys Glu Trp Glu Gln Glu Glu Thr Gln Ala Val Glu
    130                 135                 140
Glu Ala Ala Glu Glu Glu Glu Glu Asp Ala Asp Gly Gln Ala Glu Glu
145                 150                 155                 160
Glu Gln Ala Ala Thr Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu
                165                 170                 175
Ser Gly Glu Lys Ile Ser Lys Asp Gly Leu Gln Ile Gly Thr Asp Ala
            180                 185                 190
Thr Ala Thr Glu Gln Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro
        195                 200                 205
Glu Pro Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val
    210                 215                 220
Ala Gly Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr
225                 230                 235                 240
Gly Ser Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln Gly Val Leu
                245                 250                 255
Thr Ala Asn Ala Gln Gly Gln Leu Glu Ser Gln Val Glu Met Gln Phe
            260                 265                 270
Phe Ser Thr Ser Glu Asn Ala Arg Asn Glu Ala Asn Asn Ile Gln Pro
        275                 280                 285
```

```
Lys Leu Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr
    290                 295                 300
His Leu Ser Tyr Lys Pro Ala Lys Ser Asp Asp Asn Ser Lys Ile Met
305                 310                 315                 320
Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
                325                 330                 335
Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
            340                 345                 350
Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln
        355                 360                 365
Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly
    370                 375                 380
Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
385                 390                 395                 400
Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu
                405                 410                 415
Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr
            420                 425                 430
Gln Ala Val Lys Thr Asn Asn Gly Asn Asn Gly Gly Gln Val Thr Trp
        435                 440                 445
Thr Lys Asp Glu Thr Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn
    450                 455                 460
Asn Phe Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe
465                 470                 475                 480
Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn
                485                 490                 495
Pro Ser Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met
            500                 505                 510
Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu
        515                 520                 525
Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn
    530                 535                 540
His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
545                 550                 555                 560
Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
                565                 570                 575
Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
            580                 585                 590
Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp
        595                 600                 605
Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr
    610                 615                 620
Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
625                 630                 635                 640
Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser
                645                 650                 655
Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro
            660                 665                 670
Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe
        675                 680                 685
Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp
    690                 695                 700
Pro Tyr Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe
705                 710                 715                 720
```

```
Tyr Leu Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser
                725                 730                 735
Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu
            740                 745                 750
Ile Lys Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn
        755                 760                 765
Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile
    770                 775                 780
Gly Tyr Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr
785                 790                 795                 800
Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln
                805                 810                 815
Thr Lys Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn
            820                 825                 830
Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln
        835                 840                 845
Ala Tyr Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val
    850                 855                 860
Asp Ser Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg
865                 870                 875                 880
Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu
                885                 890                 895
Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr
            900                 905                 910
Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe
        915                 920                 925
Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile
    930                 935                 940
Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955                 960
```

<210> 123
<211> 937
<212> PRT
<213> Chimpanzee Adenovirus- ChAd 6 Hexon

<400> 123

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Ser Gln Trp Ile Thr Lys Asp Asn Gly Thr Asp Lys
    130                 135                 140
```

```
Thr Tyr Ser Phe Gly Asn Ala Pro Val Arg Gly Leu Asp Ile Thr Glu
145                 150                 155                 160
Glu Gly Leu Gln Ile Gly Pro Asp Glu Ser Gly Gly Glu Ser Lys Lys
                165                 170                 175
Ile Phe Ala Asp Lys Thr Tyr Gln Pro Glu Pro Gln Leu Gly Asp Glu
            180                 185                 190
Glu Trp His Asp Thr Ile Gly Ala Glu Asp Lys Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Pro Ala Thr Asn Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys
    210                 215                 220
Pro Thr Asn Ala Lys Gly Gly Gln Ala Lys Ser Arg Thr Lys Asp Asp
225                 230                 235                 240
Gly Thr Thr Glu Pro Asp Ile Asp Met Ala Phe Phe Asp Asp Arg Ser
                245                 250                 255
Gln Gln Ala Ser Phe Ser Pro Glu Leu Val Leu Tyr Thr Glu Asn Val
            260                 265                 270
Asp Leu Asp Thr Pro Asp Thr His Ile Ile Tyr Lys Pro Gly Thr Asp
        275                 280                 285
Glu Thr Ser Ser Ser Phe Asn Leu Gly Gln Gln Ser Met Pro Asn Arg
    290                 295                 300
Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr
305                 310                 315                 320
Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
                325                 330                 335
Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln
            340                 345                 350
Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp
        355                 360                 365
Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
    370                 375                 380
His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asn Gly
385                 390                 395                 400
Val Gly Phe Thr Asp Thr Phe Gln Gly Ile Lys Val Lys Thr Thr Asn
                405                 410                 415
Asn Gly Thr Ala Asn Ala Thr Glu Trp Glu Ser Asp Thr Ser Val Asn
            420                 425                 430
Asn Ala Asn Glu Ile Ala Lys Gly Asn Pro Phe Ala Met Glu Ile Asn
        435                 440                 445
Ile Gln Ala Asn Leu Trp Arg Asn Phe Leu Tyr Ala Asn Val Ala Leu
    450                 455                 460
Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Ala Asn Ile Thr Leu Pro
465                 470                 475                 480
Ala Asn Thr Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val Val Ala Pro
                485                 490                 495
Ser Leu Val Asp Ala Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp
            500                 505                 510
Pro Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu
        515                 520                 525
Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His
    530                 535                 540
Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu Leu
545                 550                 555                 560
Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met
                565                 570                 575
```

```
Ile Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser
                580                 585                 590
Ile Ala Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala
            595                 600                 605
His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn
        610                 615                 620
Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro
625                 630                 635                 640
Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg Asn
                645                 650                 655
Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Arg Glu
            660                 665                 670
Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly
        675                 680                 685
Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys
    690                 695                 700
Lys Val Ser Ile Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp
705                 710                 715                 720
Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly
                725                 730                 735
Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu
            740                 745                 750
Val Gln Met Leu Ala His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val
        755                 760                 765
Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln
    770                 775                 780
Pro Met Ser Arg Gln Val Val Asp Glu Val Asn Tyr Lys Asp Tyr Gln
785                 790                 795                 800
Ala Val Thr Leu Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr
                805                 810                 815
Leu Ala Pro Thr Met Arg Gln Gly Gln Pro Tyr Pro Ala Asn Tyr Pro
            820                 825                 830
Tyr Pro Leu Ile Gly Lys Ser Ala Val Ala Ser Val Thr Gln Lys Lys
        835                 840                 845
Phe Leu Cys Asp Arg Val Met Trp Arg Ile Pro Phe Ser Ser Asn Phe
    850                 855                 860
Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala
865                 870                 875                 880
Asn Ser Ala His Ala Leu Asp Met Asn Phe Glu Val Asp Pro Met Asp
                885                 890                 895
Glu Ser Thr Leu Leu Tyr Val Val Phe Glu Val Phe Asp Val Val Arg
            900                 905                 910
Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr
        915                 920                 925
Pro Phe Ser Ala Gly Asn Ala Thr Thr
    930                 935
```

<210> 124
<211> 956
<212> PRT
<213> Chimpanzee Adenovirus- C1 Hexon

<400> 124

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
```

```
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
                20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Asn Leu Gly Asn Lys Phe Arg Asn Pro
            35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
        50                  55                  60
Met Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Phe Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Ser Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Ser Gln Trp Leu Asp Lys Gly Val Thr Thr Thr Asp
    130                 135                 140
Asn Asn Thr Glu Asn Gly Asp Glu Glu Asp Glu Val Ala Glu Glu Gly
145                 150                 155                 160
Glu Glu Glu Lys Gln Ala Thr Tyr Thr Phe Gly Asn Ala Pro Val Lys
                165                 170                 175
Ala Glu Ala Glu Ile Thr Lys Glu Gly Leu Pro Ile Gly Leu Glu Val
            180                 185                 190
Pro Ser Glu Gly Asp Pro Lys Pro Ile Tyr Ala Asp Lys Leu Tyr Gln
        195                 200                 205
Pro Glu Pro Gln Val Gly Glu Glu Ser Trp Thr Asp Thr Asp Gly Thr
    210                 215                 220
Asp Glu Lys Tyr Gly Gly Arg Ala Leu Lys Pro Glu Thr Lys Met Lys
225                 230                 235                 240
Pro Cys Tyr Gly Ser Phe Ala Lys Pro Thr Asn Val Lys Gly Gly Gln
                245                 250                 255
Ala Lys Val Lys Lys Val Glu Glu Gly Lys Val Glu Tyr Asp Ile Asp
            260                 265                 270
Met Asn Phe Phe Asp Leu Arg Ser Gln Lys Thr Gly Leu Lys Pro Lys
        275                 280                 285
Ile Val Met Tyr Ala Glu Asn Val Asp Leu Glu Thr Pro Asp Thr His
    290                 295                 300
Val Val Tyr Lys Pro Gly Ala Ser Asp Ala Ser Ser His Ala Asn Leu
305                 310                 315                 320
Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp
                325                 330                 335
Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val
            340                 345                 350
Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp
        355                 360                 365
Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Leu Gly Asp
    370                 375                 380
Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp
385                 390                 395                 400
Pro Asp Val Arg Val Ile Glu Asn His Gly Val Glu Asp Glu Leu Pro
                405                 410                 415
Asn Tyr Cys Phe Pro Leu Asp Gly Val Gly Pro Arg Thr Asp Ser Tyr
            420                 425                 430
Lys Gly Ile Glu Thr Asn Gly Asp Glu Asn Thr Thr Trp Lys Asp Leu
        435                 440                 445
```

```
Asp Pro Asn Gly Ile Ser Glu Leu Ala Lys Gly Asn Pro Phe Ala Met
    450                 455                 460
Glu Ile Asn Ile Gln Ala Asn Leu Trp Arg Ser Phe Leu Tyr Ser Asn
465                 470                 475                 480
Val Ala Leu Tyr Leu Pro Asp Ser Tyr Lys Tyr Thr Pro Thr Asn Val
                485                 490                 495
Thr Leu Pro Glu Asn Lys Asn Thr Tyr Asp Tyr Met Asn Gly Arg Val
            500                 505                 510
Val Pro Pro Ser Leu Val Asp Thr Tyr Val Asn Ile Gly Ala Arg Trp
        515                 520                 525
Ser Leu Asp Ala Met Asp Asn Val Asn Pro Phe Asn His His Arg Asn
    530                 535                 540
Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg Tyr Val
545                 550                 555                 560
Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Val Lys Asn Leu
                565                 570                 575
Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg Lys Asp
            580                 585                 590
Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg Val Asp
        595                 600                 605
Gly Ala Ser Ile Ser Phe Thr Ser Ile Asn Leu Tyr Ala Thr Phe Phe
    610                 615                 620
Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu Arg Asn
625                 630                 635                 640
Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala Asn Met
                645                 650                 655
Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser Ile Pro
            660                 665                 670
Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ser Phe Thr Arg Leu Lys
        675                 680                 685
Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Val
    690                 695                 700
Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn His
705                 710                 715                 720
Thr Phe Lys Lys Val Ser Ile Met Phe Asp Ser Ser Val Ser Trp Pro
                725                 730                 735
Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys Arg Thr
            740                 745                 750
Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr Lys Asp
        755                 760                 765
Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr Gln Gly
    770                 775                 780
Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg Met Tyr Ser Phe Phe Arg
785                 790                 795                 800
Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Glu Ile Asn Tyr Lys
                805                 810                 815
Asp Tyr Lys Ala Val Ala Val Pro Tyr Gln His Asn Asn Ser Gly Phe
            820                 825                 830
Val Gly Tyr Met Ala Pro Thr Met Arg Gln Gly Gln Ala Tyr Pro Ala
        835                 840                 845
Asn Tyr Pro Tyr Pro Leu Ile Gly Thr Thr Ala Val Thr Ser Val Thr
    850                 855                 860
Gln Lys Lys Phe Leu Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser
865                 870                 875                 880
```

```
Ser Asn Phe Met Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Leu
                885                 890                 895
Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp
            900                 905                 910
Pro Met Asp Glu Pro Thr Leu Leu Tyr Leu Leu Phe Glu Val Phe Asp
        915                 920                 925
Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr
    930                 935                 940
Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945                 950                 955
```

<210> 125
<211> 933
<212> PRT
<213> Chimpanzee Adenovirus- CV68 Hexon

<400> 125

```
Met Ala Thr Pro Ser Met Leu Pro Gln Trp Ala Tyr Met His Ile Ala
 1               5                  10                  15
Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
            20                  25                  30
Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
        35                  40                  45
Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
    50                  55                  60
Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Asn Thr Tyr Ser Tyr
65                  70                  75                  80
Lys Val Arg Tyr Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
                85                  90                  95
Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
            100                 105                 110
Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
        115                 120                 125
Ala Pro Asn Thr Cys Gln Trp Thr Tyr Lys Ala Asp Gly Glu Thr Ala
    130                 135                 140
Thr Glu Lys Thr Tyr Thr Tyr Gly Asn Ala Pro Val Gln Gly Ile Asn
145                 150                 155                 160
Ile Thr Lys Asp Gly Ile Gln Leu Gly Thr Asp Thr Asp Asp Gln Pro
                165                 170                 175
Ile Tyr Ala Asp Lys Thr Tyr Gln Pro Glu Pro Gln Val Gly Asp Ala
            180                 185                 190
Glu Trp His Asp Ile Thr Gly Thr Asp Glu Lys Tyr Gly Gly Arg Ala
        195                 200                 205
Leu Lys Pro Asp Thr Lys Met Lys Pro Cys Tyr Gly Ser Phe Ala Lys
    210                 215                 220
Pro Thr Asn Lys Glu Gly Gly Gln Ala Asn Val Lys Thr Gly Thr Gly
225                 230                 235                 240
Thr Thr Lys Glu Tyr Asp Ile Asp Met Ala Phe Phe Asp Asn Arg Ser
                245                 250                 255
Ala Ala Ala Ala Gly Leu Ala Pro Glu Ile Val Leu Tyr Thr Glu Asn
            260                 265                 270
Val Asp Leu Glu Thr Pro Asp Thr His Ile Val Tyr Lys Ala Gly Thr
        275                 280                 285
Asp Asp Ser Ser Ser Ser Ile Asn Leu Gly Gln Gln Ala Met Pro Asn
    290                 295                 300
```

```
Arg Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr
305                 310                 315                 320
Tyr Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln
                325                 330                 335
Leu Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr
            340                 345                 350
Gln Leu Leu Leu Asp Ser Leu Gly Asp Arg Thr Arg Tyr Phe Ser Met
        355                 360                 365
Trp Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu
    370                 375                 380
Asn His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Asp
385                 390                 395                 400
Ala Val Gly Arg Thr Asp Thr Tyr Gln Gly Ile Lys Ala Asn Gly Thr
                405                 410                 415
Asp Gln Thr Thr Trp Thr Lys Asp Asp Ser Val Asn Asp Ala Asn Glu
            420                 425                 430
Ile Gly Lys Gly Asn Pro Phe Ala Met Glu Ile Asn Ile Gln Ala Asn
        435                 440                 445
Leu Trp Arg Asn Phe Leu Tyr Ala Asn Val Ala Leu Tyr Leu Pro Asp
    450                 455                 460
Ser Tyr Lys Tyr Thr Pro Ala Asn Val Thr Leu Pro Thr Asn Thr Asn
465                 470                 475                 480
Thr Tyr Asp Tyr Met Asn Gly Arg Val Val Ala Pro Ser Leu Val Asp
                485                 490                 495
Ser Tyr Ile Asn Ile Gly Ala Arg Trp Ser Leu Asp Pro Met Asp Asn
            500                 505                 510
Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser
        515                 520                 525
Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro
    530                 535                 540
Gln Lys Phe Phe Ala Ile Lys Ser Leu Leu Leu Leu Pro Gly Ser Tyr
545                 550                 555                 560
Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser
                565                 570                 575
Ser Leu Gly Asn Asp Leu Arg Thr Asp Gly Ala Ser Ile Ser Phe Thr
            580                 585                 590
Ser Ile Asn Leu Tyr Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala
        595                 600                 605
Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
    610                 615                 620
Asn Asp Tyr Leu Ser Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
625                 630                 635                 640
Ala Thr Asn Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
                645                 650                 655
Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu
            660                 665                 670
Gly Ser Gly Phe Asp Pro Tyr Phe Val Tyr Ser Gly Ser Ile Pro Tyr
        675                 680                 685
Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
    690                 695                 700
Thr Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
705                 710                 715                 720
Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
                725                 730                 735
```

```
Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu
            740                 745                 750
Ala His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
        755                 760                 765
Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
    770                 775                 780
Gln Val Val Asp Glu Val Asn Tyr Lys Asp Tyr Gln Ala Val Thr Leu
785                 790                 795                 800
Ala Tyr Gln His Asn Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr
                805                 810                 815
Met Arg Gln Gly Gln Pro Tyr Pro Ala Xaa Tyr Pro Tyr Pro Leu Ile
            820                 825                 830
Gly Lys Ser Ala Val Thr Ser Val Thr Gln Lys Lys Phe Leu Cys Asp
        835                 840                 845
Arg Val Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
    850                 855                 860
Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
865                 870                 875                 880
Ala Leu Asp Met Asn Phe Glu Val Asp Pro Met Asp Glu Ser Thr Leu
                885                 890                 895
Leu Tyr Val Val Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro
            900                 905                 910
His Arg Gly Val Ile Glu Ala Val Tyr Xaa Arg Thr Pro Phe Ser Ala
        915                 920                 925
Gly Asn Ala Thr Thr
    930
```

Claims

1. A replication defective chimpanzee adenoviral vector comprising a recombinant adenoviral genome comprising a sequence of nucleotides that encodes a peptide as set forth in SEQ ID 116 and a transgene which encodes at least one immunogen operatively linked to regulatory sequences which direct expression of said transgene in mammalian cells, wherein said vector is modified in that the gene E1, is functionally deleted.

2. The replication defective chimpanzee adenoviral vector according to claim 1, wherein the recombinant adenoviral genome comprises a sequence of nucleotides as set forth in SEQ ID NO:115.

3. The replication defective chimpanzee adenoviral vector according to claim 1 or 2, wherein the recombinant adenoviral genome comprises a sequence of nucleotides as set forth in SEQ ID NO:76.

4. A host cell comprising a replication-defective chimpanzee adenoviral vector according to any one of claims 1-3 wherein said host cell expresses one or more adenoviral regions selected from the group consisting of E1a, E1b, E2a, E2b, E4 orfs 1, 2, 3, 4, 5, 6, 6/7, pIX, IVa2, regions L1, L2, L3, L4, L5.

5. An isolated recombinant adenovirus genome comprising:

   a) a hexon gene sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:115, and
   b) a fiber gene sequence comprising a sequence of nucleotides as set forth in SEQ ID NO:76.

6. A replication-defective chimpanzee adenoviral vector according to any one of claims 1-3 for use in a method of boosting an antigen-specific immune response in a mammal comprising administering to said mammal a sufficient amount of said recombinant ChAd vector, wherein administration of said ChAd vector elicits a boosted response.

7. The vector of claim 6 wherein the chimpanzee adenoviral vector comprises a complete deletion of its E1 genes and further wherein the vector optionally comprises a deletion of its E3 genes.

8. The vector of claim 6 wherein the immune response is a boosted immune response that is specific for a TAA.

9. The vector of claim 6 or claim 8 wherein the boosted immune response comprises the production of antigen-specific CD8+ T cells.

10. A replication-defective chimpanzee adenoviral vector according to any one of claims 1-3 for use in a method of eliciting an immune response in a naïve mammal comprising administering to said mammal a sufficient amount of said adenoviral vector, wherein administration of the adenoviral vector elicits a primary immune response.

11. The vector of claim 6 or claim 10 wherein the boosted or primary immune response, respectively, is specific for an antigen derived from an infectious agent such as, but not limited to HIV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae,* Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, *Mycobacterium tuberculosis,* CMV and RSV.

12. The vector of claim 6 wherein the immune response is a primary immune response that is specific for a TAA against which the mammal is tolerant.

13. A replication-defective chimpanzee adenoviral vector according to any one of claims 1-3 for use in a method of inducing an immune response against an antigen derived from an infectious agent comprising the steps of:

(a) priming a host to respond to an infectious agent-antigen by administering a first vaccine composition comprising a nucleotide sequence encoding an infectious agent-antigen against which an antigen-specific immune response is desired; and
(b) boosting the immune response of step (a) by administering a second vaccine composition comprising a recombinant ChAd vector of claim 1 containing at least a functional deletion of its E1 gene, and in the site of the E1 gene deletion, a sequence comprising a promoter capable of directing expression of DNA encoding the same infectious agent-antigen delivered in the priming step;

wherein administration of the boosting composition elicits an immune response which has the effect of conferring protective immunity wherein the antigen derived from an infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, HSV1, HSV2, SARS, CoV, *Leishmania major, Plasmodium malaria,* Ebola virus, West Nile virus, Dengue virus, Influenza A, Influenza B, RSV, CMV and *Mycobacterium tuberculosis.*

14. The vector of claim 13 wherein the immune response comprises the production of antigen-specific CD8+ T cells.

15. A replication-defective chimpanzee adenoviral vector according to any one of claims 1-3 for use in a method of breaking host tolerance to a self-antigen comprising:

(a) priming a host to respond to a self-antigen by administering a first vaccine composition comprising a nucleotide sequence encoding a self-antigen against which an antigen-specific immune response is desired, thereby eliciting a primed response; and
(b) boosting the primed immune response of step (a) by administering a second vaccine composition comprising a recombinant ChAd vector of claim 1 containing at least a functional deletion of its E1 gene, and in the site of the E1 gene deletion, a sequence comprising a promoter capable of directing expression of DNA encoding the same self-antigen delivered in the priming step, wherein administration of the boosting composition elicits an immune response which has the effect of breaking host tolerance to the self-antigen.

16. The replication defective chimpanzee adenoviral vector for use in a method according to claim 15 wherein the self antigen is selected from the group consisting of: HER2 NEU, CEA, HEPCAM, PSA, PSMA, TELOMERASE, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELYSIN 3, TYROSINASE, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 and VEGFR2.

17. The replication defective chimpanzee adenoviral vector of claim 15 wherein the immune response comprises the production of antigen-specific CD8+ T cells.

18. The replication defective chimpanzee adenoviral vector of claim 15 or 16 wherein the first and second vaccine compositions are both adenoviral vectors **characterized by** different serotypes.

Patentansprüche

1. Ein replikationsdefizienter Schimpansen-Adenovirus-Vektor, umfassend ein rekombinantes Adenovirus-Genom, das eine Nukleotidsequenz umfasst, die ein Peptid kodiert, wie in SEQ ID 116 angegeben, und ein Transgen, das wenigstens ein Immunogen kodiert, operativ verknüpft mit regulatorischen Sequenzen, die die Expression des Transgens in Säugerzellen lenken, wobei der Vektor derart modifiziert ist, dass das Gen E1 funktional deletiert ist.

2. Der replikationsdefiziente Schimpansen-Adenovirus-Vektor gemäß Anspruch 1, wobei das rekombinante Adenovirus-Genom eine wie in SEQ ID NO: 115 angegebene Nukleotidsequenz umfasst.

3. Der replikationsdefiziente Schimpansen-Adenovirus-Vektor gemäß Anspruch 1 oder 2, wobei das rekombinante Adenovirus-Genom eine wie in SEQ ID NO: 76 angegebene Nukleotidsequenz umfasst.

4. Eine Wirtszelle, die einen replikationsdefizienten Schimpansen-Adenovirus-Vektor gemäß einem der Ansprüche 1-3 umfasst, wobei die Wirtszelle eine oder mehrere Adenovirus-Regionen, ausgewählt aus der Gruppe bestehend aus E1a, E1 b, E2a, E2b, E4 orfs 1, 2, 3, 4, 5, 6, 6/7, pIX, IVa2, Regionen L1, L2, L3, L4, L5, exprimiert.

5. Ein isoliertes rekombinantes Adenovirus-Genom, umfassend:

   a) eine Hexon-Gensequenz, umfassend eine wie in SEQ ID NO: 115 angegebene Nukleotidsequenz, und
   b) eine Faser-Gensequenz, umfassend eine wie in SEQ ID NO: 76 angegebene Nukleotidsequenz.

6. Ein replikationsdefizienter Schimpansen-Adenovirus-Vektor gemäß einem der Ansprüche 1-3 zur Verwendung bei einem Verfahren zur Verstärkung einer antigenspezifischen Immunreaktion bei einem Säuger, umfassend die Verabreichung einer ausreichenden Menge des rekombinanten ChAd-Vektors an den Säuger, wobei die Verabreichung des ChAd-Vektors eine verstärkte Reaktion hervorruft.

7. Der Vektor nach Anspruch 6, wobei der Schimpansen-Adenovirus-Vektor eine vollständige Deletion seiner E1-Gene umfasst, und wobei ferner der Vektor gegebenenfalls eine Deletion seiner E3-Gene umfasst.

8. Der Vektor nach Anspruch 6, wobei die Immunreaktion eine verstärkte Immunreaktion ist, die für ein TAA spezifisch ist.

9. Der Vektor nach Anspruch 6 oder Anspruch 8, wobei die verstärkte Immunreaktion die Produktion von antigenspezifischen CD8+-T-Zellen umfasst.

10. Ein replikationsdefizienter Schimpansen-Adenovirus-Vektor gemäß einem der Ansprüche 1-3 zur Verwendung bei einem Verfahren zum Hervorrufen einer Immunreaktion bei einem naiven Säuger, umfassend die Verabreichung einer ausreichenden Menge des Adenovirus-Vektors an den Säuger, wobei die Verabreichung des Adenovirus-Vektors eine primäre Immunreaktion hervorruft.

11. Der Vektor nach Anspruch 6 oder Anspruch 10, wobei die verstärkte bzw. primäre Immunreaktion spezifisch ist für ein Antigen, das von einem Infektionserreger hergeleitet ist, wie z.B., ohne jedoch darauf beschränkt zu sein, HIV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae,* Ebola-Virus, West-Nil-Virus, Dengue-Virus, Influenza A, Influenza B, *Mycobacterium tuberculosis,* CMV und RSV.

12. Der Vektor nach Anspruch 6, wobei die Immunreaktion eine primäre Immunreaktion ist, die spezifisch ist für ein TAA, gegenüber dem der Säuger tolerant ist.

13. Ein replikationsdefizienter Schimpansen-Adenovirus-Vektor gemäß einem der Ansprüche 1-3 zur Verwendung bei einem Verfahren zum Hervorrufen einer Immunreaktion gegen ein Antigen, das von einem Infektionserreger hergeleitet ist, umfassend die Schritte:

   (a) Induzieren eines Wirts, so dass er auf ein Infektionserreger-Antigen reagiert, durch Verabreichen einer ersten Impfzusammensetzung, die eine Nukleotidsequenz umfasst, welche ein Infektionserreger-Antigen, gegen das eine antigenspezifische Immunreaktion erwünscht ist, kodiert, und
   (b) Verstärken der Immunreaktion von Schritt (a) durch Verabreichen einer zweiten Impfzusammensetzung, die einen rekombinanten ChAd-Vektor nach Anspruch 1 umfasst, welcher wenigstens eine funktionale Deletion

seines E1-Gens enthält, und an der Stelle der E1-Gen-Deletion eine Sequenz, die einen Promotor umfasst, der in der Lage ist, die Expression von DNA zu lenken, die das gleiche Infektionserreger-Antigen kodiert, das im Induzierungsschritt zugeführt wird,

wobei die Verabreichung der verstärkenden Zusammensetzung eine Immunreaktion hervorruft, welche eine Schutzimmunität verleihende Wirkung besitzt, wobei das Antigen von einem Infektionserreger hergeleitet ist, ausgewählt aus der Gruppe bestehend aus: HIV, HBV, HCV, HPV, HSV1, HSV2, SARS, CoV, *Leishmania major, Plasmodium malariae,* Ebola-Virus, West-Nil-Virus, Dengue-Virus, Influenza A, Influenza B, RSV, CMV und *Mycobacterium tuberculosis.*

14. Der Vektor nach Anspruch 13, wobei die Immunreaktion die Produktion von antigenspezifischen CD8+-T-Zellen umfasst.

15. Ein replikationsdefizienter Schimpansen-Adenovirus-Vektor gemäß einem der Ansprüche 1-3 zur Verwendung bei einem Verfahren zum Brechen der Toleranz gegenüber einem Selbstantigen, umfassend:

(a) Induzieren eines Wirts, so dass er auf ein Selbstantigen reagiert, durch Verabreichen einer ersten Impfzusammensetzung, die eine Nukleotidsequenz umfasst, welche ein Selbstantigen, gegen das eine antigenspezifische Immunreaktion erwünscht ist, kodiert, wodurch eine induzierte Reaktion hervorgerufen wird, und
(b) Verstärken der induzierten Immunreaktion von Schritt (a) durch Verabreichen einer zweiten Impfzusammensetzung, die einen rekombinanten ChAd-Vektor nach Anspruch 1 umfasst, welcher wenigstens eine funktionale Deletion seines E1-Gens enthält, und an der Stelle der E1-Gen-Deletion eine Sequenz, die einen Promotor umfasst, der in der Lage ist, die Expression von DNA zu lenken, die das gleiche Selbstantigen kodiert, das im Induzierungsschritt zugeführt wird, wobei die Verabreichung der verstärkenden Zusammensetzung eine Immunreaktion hervorruft, welche bewirkt, dass die Toleranz des Wirts gegenüber dem Selbstantigen gebrochen wird.

16. Der replikationsdefiziente Schimpansen-Adenovirus-Vektor zur Verwendung bei einem Verfahren gemäß Anspruch 15, wobei das Selbstantigen ausgewählt ist aus der Gruppe bestehend aus: HER2 NEU, CEA, HEPCAM, PSA, PSMA, TELOMERASE, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELYSIN 3, TYROSINASE, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 und VEGFR2.

17. Der replikationsdefiziente Schimpansen-Adenovirus-Vektor nach Anspruch 15, wobei die Immunreaktion die Produktion von antigenspezifischen CD8+-T-Zellen umfasst.

18. Der replikationsdefiziente Schimpansen-Adenovirus-Vektor nach Anspruch 15 oder 16, wobei die ersten und zweiten Impfzusammensetzungen beides Adenovirus-Vektoren sind, die durch unterschiedliche Serotypen gekennzeichnet sind.

Revendications

1. Vecteur adénoviral de chimpanzé défectueux à la réplication comprenant un génome adénoviral recombinant comprenant une séquence de nucléotides qui code un peptide telle que présentée dans la SEQ ID NO: 116 et un transgène qui code au moins un immunogène lié d'une manière opérationnelle à des séquences régulatrices qui dirigent l'expression dudit transgène dans des cellules mammaliennes, où ledit vecteur est modifié en ce que le gène E1 est fonctionnellement supprimé.

2. Vecteur adénoviral de chimpanzé défectueux à la réplication selon la revendication 1, dans lequel le génome adénoviral recombinant comprend une séquence de nucléotides telle que présentée dans la SEQ ID NO: 115.

3. Vecteur adénoviral de chimpanzé défectueux à la réplication selon la revendication 1 ou 2, dans lequel le génome adénoviral recombinant comprend une séquence de nucléotides telle que présentée dans la SEQ ID NO: 76.

4. Cellule hôte comprenant un vecteur adénoviral de chimpanzé défectueux à la réplication selon l'une quelconque des revendications 1-3, où ladite cellule hôte exprime une ou plusieurs régions adénovirales sélectionnées parmi le groupe consistant en E1a, E1b, E2a, E2b, E4 cadres de lecture ouverts 1, 2, 3, 4, 5, 6, 6/7, pIX, IVa2, régions

L1, L2, L3, L4, L5.

5. Génome d'adénovirus recombinant isolé, comprenant:

a) une séquence de gène de l'hexon comprenant une séquence de nucléotides telle que présentée dans la SEQ ID NO: 115; et
b) une séquence de gène de la fibre comprenant une séquence de nucléotides telle que présentée dans la SEQ ID NO: 76.

6. Vecteur adénoviral de chimpanzé défectueux à la réplication selon l'une quelconque des revendications 1-3 à utiliser dans un procédé destiné à potentialiser une réponse immunitaire antigène-spécifique chez un mammifère, comprenant administrer audit mammifère une quantité suffisante dudit vecteur ChAd recombinant, où l'administration dudit vecteur ChAd provoque une réponse potentialisée.

7. Vecteur selon la revendication 6, où le vecteur adénoviral de chimpanzé comprend une délétion complète de ses gènes E1 et en outre où le vecteur comprend optionnellement une délétion de ses gènes E3.

8. Vecteur selon la revendication 6, où la réponse immunitaire est une réponse immunitaire potentialisée qui est spécifique pour un TAA.

9. Vecteur selon la revendication 6 ou la revendication 8, où la réponse immunitaire potentialisée comprend la production de cellules T CD8+ antigène-spécifiques.

10. Vecteur adénoviral de chimpanzé défectueux à la réplication selon l'une quelconque des revendications 1-3 à utiliser dans un procédé destiné à provoquer une réponse immunitaire chez un mammifère naïf, comprenant administrer audit mammifère une quantité suffisante dudit vecteur adénoviral, où l'administration du vecteur adénoviral provoque une réponse immunitaire primaire.

11. Vecteur selon la revendication 6 ou la revendication 10, où la réponse immunitaire potentialisée ou primaire, respectivement, est spécifique pour un antigène dérivé d'un agent infectieux tel que, sans toutefois s'y limiter: VIH, VHC, VPH, VHS1, VHS2, SARS-CoV, *Plasmodium malarias,* virus Ebola, virus du Nil occidental, virus de la dengue, influenza A, influenza B, *Mycobacterium tuberculosis,* CMV et VRS.

12. Vecteur selon la revendication 6, où la réponse immunitaire est une réponse immunitaire primaire qui est spécifique pour un TAA contre lequel le mammifère est tolérant.

13. Vecteur adénoviral de chimpanzé défectueux à la réplication selon l'une quelconque des revendications 1-3, à utiliser dans un procédé d'induction d'une réponse immunitaire contre un antigène dérivé d'un agent infectieux, comprenant les étapes consistant à:

(a) sensibiliser un hôte à répondre à un agent infectieux-antigène en administrant une première composition de vaccin comprenant une séquence de nucléotides codant un agent infectieux-antigène contre lequel une réponse immunitaire antigène-spécifique est désirée; et
(b) potentialiser la réponse immunitaire de l'étape (a) en administrant une deuxième composition de vaccin comprenant un vecteur ChAd recombinant selon la revendication 1 contenant au moins une délétion fonctionnelle de son gène E1, et dans le site de délétion du gène E1, une séquence comprenant un promoteur capable de diriger l'expression d'ADN codant le même agent infectieux-antigène qu'administré à l'étape de sensibilisation;

où l'administration de la composition de potentialisation provoque une réponse immunitaire qui a l'effet de conférer une immunité protectrice, où l'antigène dérivé d'un agent infectieux est sélectionné parmi le groupe consistant en: VIH, VHB, VHC, VPH, VHS1, VHS2, SARS-CoV, *Leishmania major, Plasmodium malarias,* virus Ebola, virus du Nil occidental, virus de la dengue, influenza A, influenza B, VRS, CMV et *Mycobacterium tuberculosis.*

14. Vecteur selon la revendication 13, où la réponse immunitaire comprend la production de cellules T CD8+ antigène-spécifiques.

15. Vecteur adénoviral de chimpanzé défectueux à la réplication selon l'une quelconque des revendications 1-3, à utiliser dans un procédé destiné à casser la tolérance de l'hôte à un auto-antigène, comprenant:

(a) sensibiliser un hôte à répondre à un auto-antigène en administrant une première composition de vaccin comprenant une séquence de nucléotides codant un auto-antigène contre lequel une réponse immunitaire antigène-spécifique est désirée, provoquant ainsi une réponse sensibilisée; et

(b) potentialiser la réponse immunitaire sensibilisée de l'étape (a) en administrant une deuxième composition de vaccin comprenant un vecteur ChAd recombinant selon la revendication 1 contenant au moins une délétion fonctionnelle de son gène E1, et dans le site de délétion du gène E1, une séquence comprenant un promoteur capable de diriger l'expression d'ADN codant le même auto-antigène qu'administré à l'étape de sensibilisation, où l'administration de la composition de potentialisation provoque une réponse immunitaire qui a l'effet de casser la tolérance de l'hôte à l'auto-antigène.

16. Vecteur adénoviral de chimpanzé défectueux à la réplication, à utiliser dans un procédé selon la revendication 15, où l'auto-antigène est sélectionné parmi le groupe consistant en: HER2 NEU, CEA, HEPCAM, PSA, PSMA, TELOMERASE, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELYSIN 3, TYROSINASE, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 et VEGFR2.

17. Vecteur adénoviral de chimpanzé défectueux à la réplication selon la revendication 15, où la réponse immunitaire comprend la production de cellules T CD8+ antigène-spécifiques.

18. Vecteur adénoviral de chimpanzé défectueux à la réplication selon la revendication 15 ou 16, où la première et la deuxième composition de vaccin sont toutes les deux des vecteurs adénoviraux **caractérisés par** des sérotypes différents.

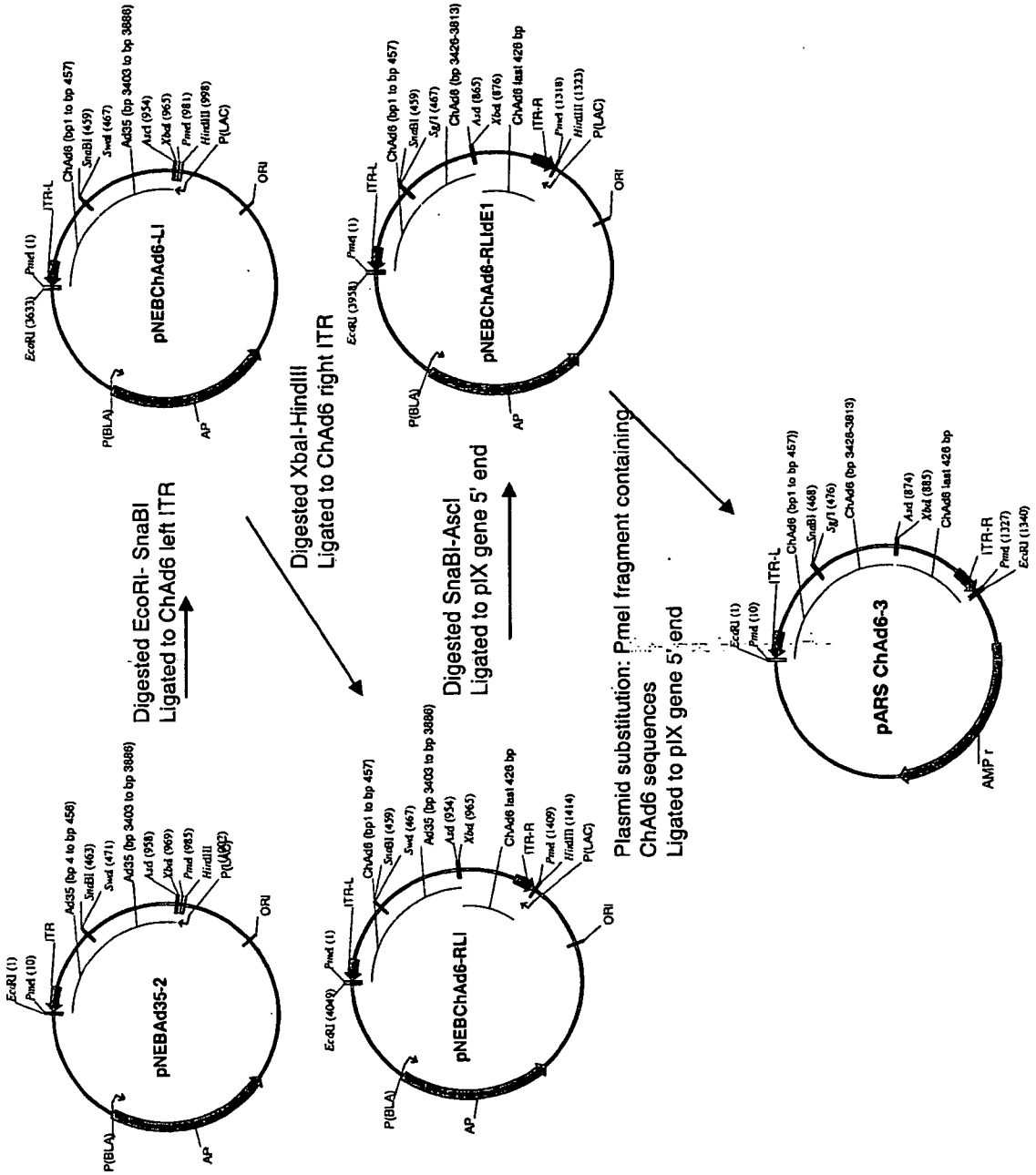
pNEBAd35-2
pNEBChAd6-LI
pNEBChAd6-RLI
pNEBChAd6-RLIdE1
pARS ChAd6-3
Digested EcoRI- SnaBI
Ligated to ChAd6 left ITR
Digested XbaI-HindIII
Ligated to ChAd6 right ITR
Digested SnaBI-AscI
Ligated to pIX gene 5' end
Plasmid substitution: PmeI fragment containing
ChAd6 sequences
Ligated to pIX gene 5' end
ITR
ITR-L
ITR-R
ORI
P(BLA)
AP
AMP r

FIG. 1

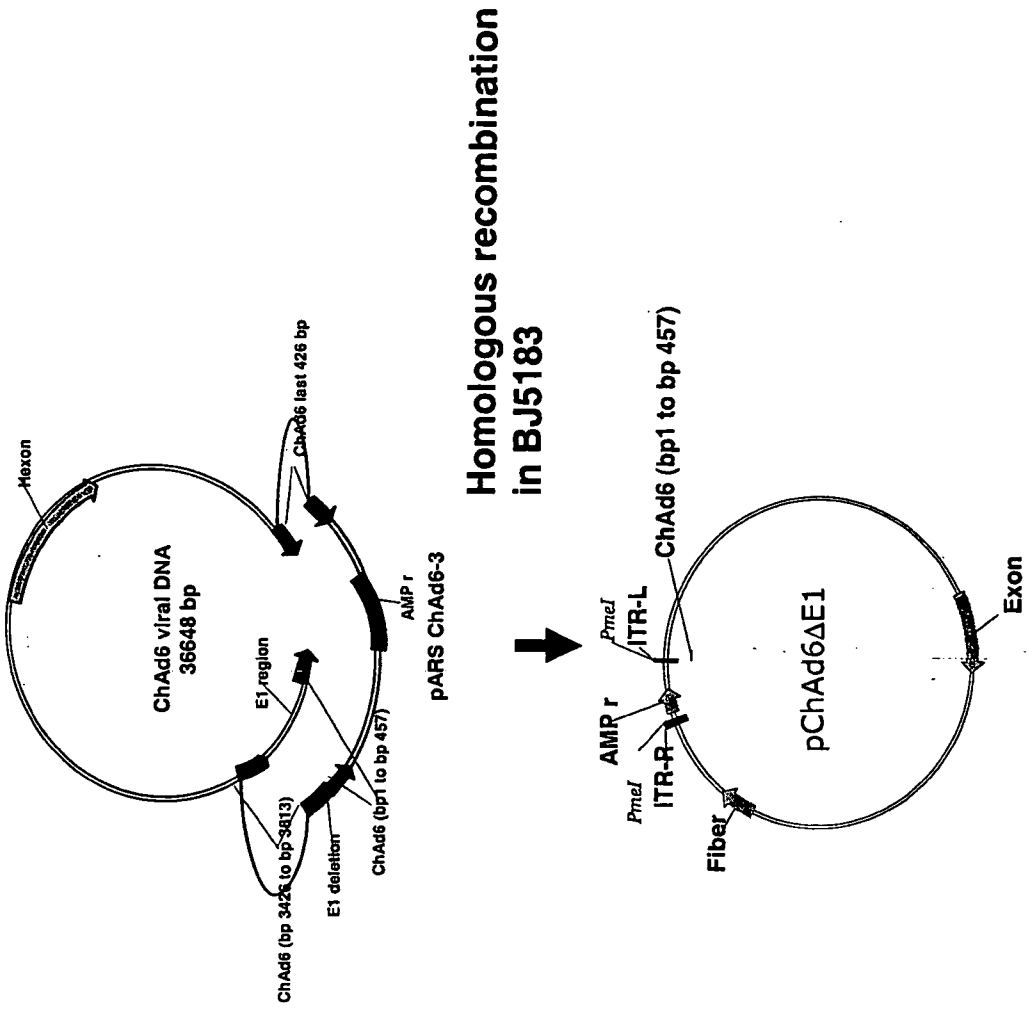
Hexon
ChAd6 viral DNA
36648 bp
ChAd6 last 426 bp
AMP r
pARS ChAd6-3
E1 region
ChAd6 (bp1 to bp 457)
E1 deletion
ChAd6 (bp 3426 to bp 3813)
Homologous recombination
in BJ5183
ChAd6 (bp1 to bp 457)
PmeI
ITR-L
AMP r
PmeI
ITR-R
pChAd6ΔE1
Exon
Fiber

FIG. 2

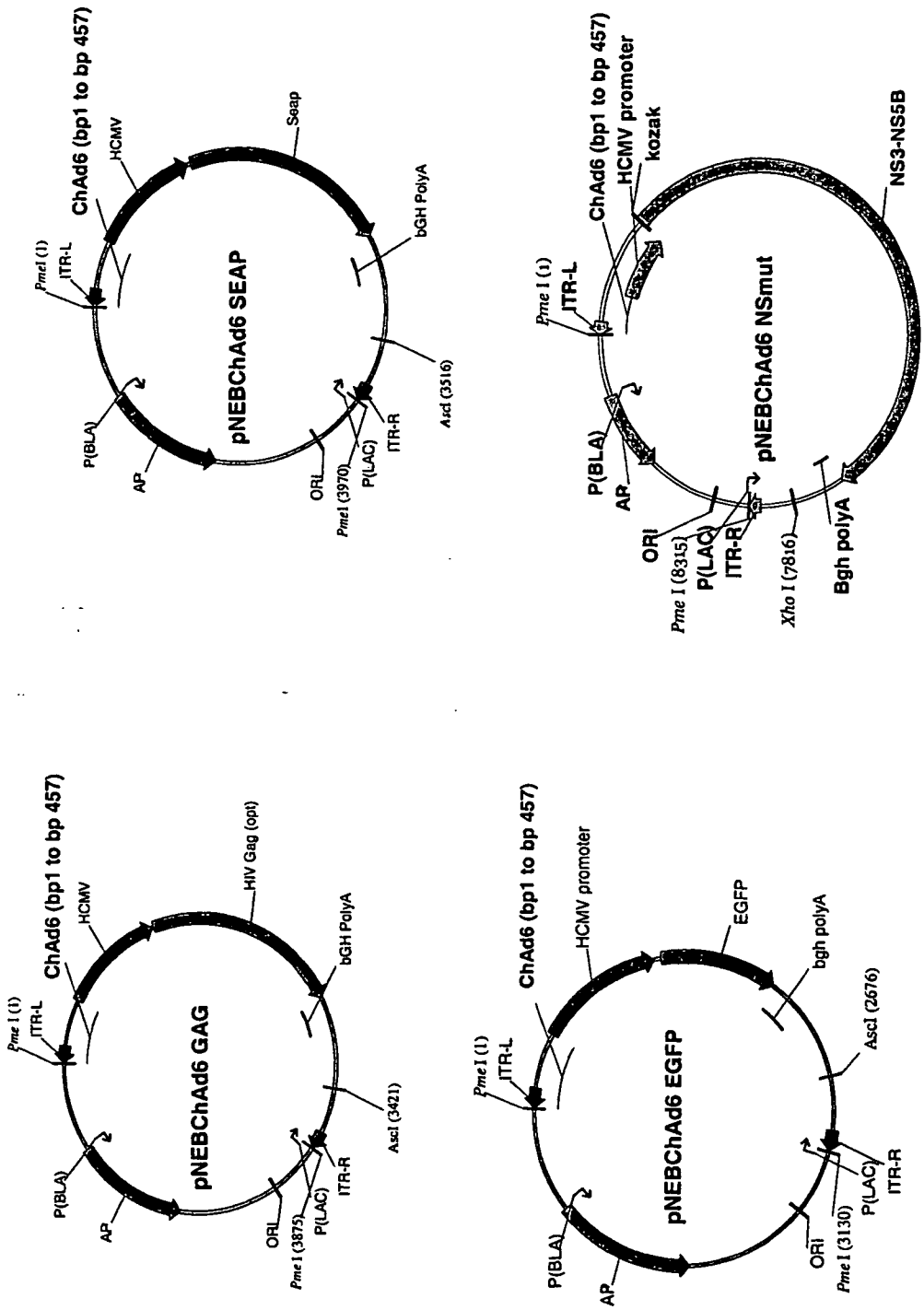
pNEBChAd6 SEAP
ChAd6 (bp1 to bp 457)
HCMV
Seap
bGH PolyA
Pme I (1)
ITR-L
P(BLA)
AP
ORI
Pme I (3970)
P(LAC)
ITR-R
Ascl (3516)
pNEBChAd6 NSmut
ChAd6 (bp1 to bp 457)
HCMV promoter
kozak
NS3-NS5B
Pme I (1)
ITR-L
P(BLA)
AP
ORI
Pme I (8315)
P(LAC)
ITR-R
Xho I (7816)
Bgh polyA
pNEBChAd6 GAG
ChAd6 (bp1 to bp 457)
HCMV
HIV Gag (opt)
bGH PolyA
Pme I (1)
ITR-L
P(BLA)
AP
ORI
Pme I (3875)
P(LAC)
ITR-R
Ascl (3421)
pNEBChAd6 EGFP
ChAd6 (bp1 to bp 457)
HCMV promoter
EGFP
bgh polyA
AscI (2676)
Pme I (1)
ITR-L
P(BLA)
AP
ORI
Pme I (3130)
P(LAC)
ITR-R

FIG. 3

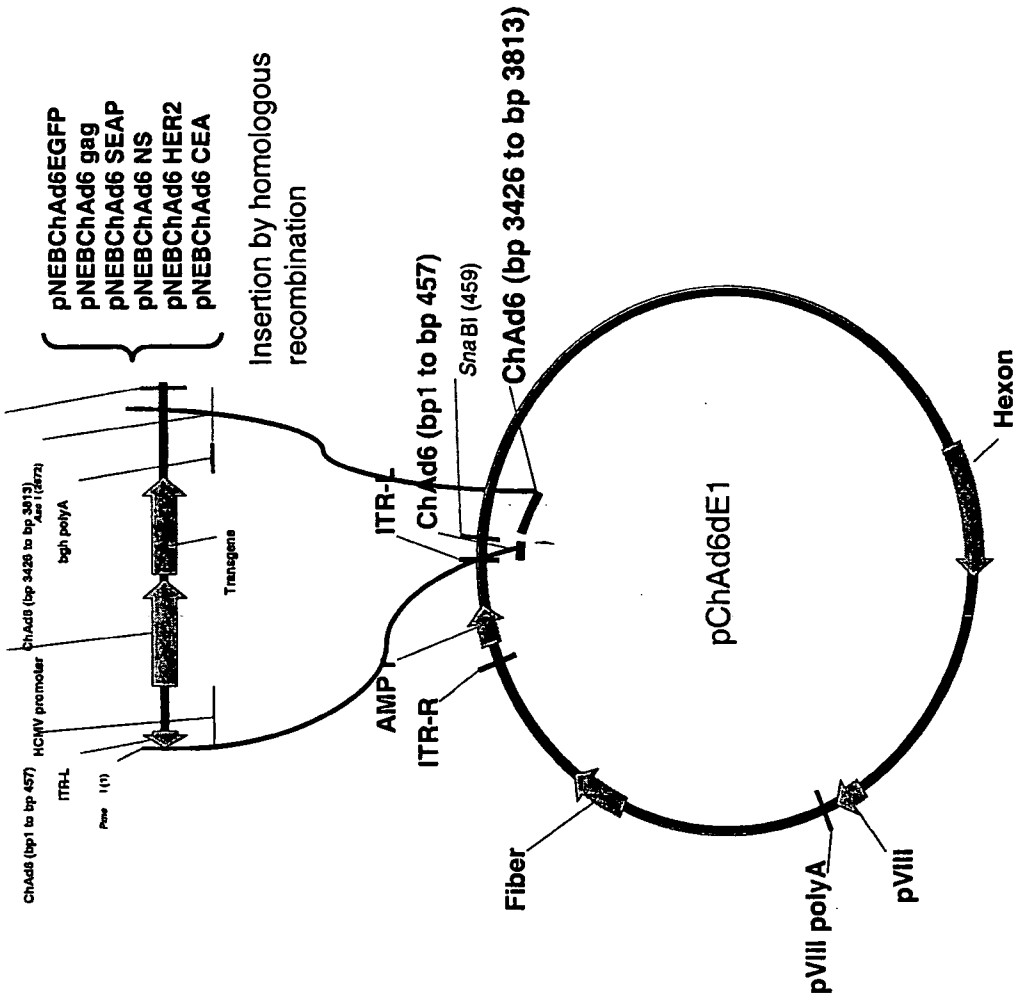
pNEBChAd6EGFP
pNEBChAd6 gag
pNEBChAd6 SEAP
pNEBChAd6 NS
pNEBChAd6 HER2
pNEBChAd6 CEA
Insertion by homologous recombination
ChAd6 (bp 3426 to bp 3813)
Sna BI (459)
ChAd6 (bp1 to bp 457)
ITR-L
AMP
ITR-R
pChAd6dE1
Hexon
Fiber
pVIII polyA
pVIII
HCMV promoter
bgh polyA
Transgene

FIG. 4

```
   1  CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG ATGGGCGGCG
  61  CGAGGCGGGG CGCGGGGCGG GAGGCGGGTT TGGGGGCGGG CCGGCGGGCG GGGCGGTGTG
 121  GCGGAAGTGG ACTTTGTAAG TGTGGCGGAT GTGACTTGCT AGTGCCGGGC GCGGTAAAAG
 181  TGACGTTTTC CGTGCGCGAC AACGCCCCCG GGAAGTGACA TTTTTCCCGC GGTTTTTACC
 241  GGATGTTGTA GTGAATTTGG GCGTAACCAA GTAAGATTTG GCCATTTTCG CGGGAAAACT
 301  GAAACGGGGA AGTGAAATCT GATTAATTTT GCGTTAGTCA TACCGCGTAA TATTTGTCTA
 361  GGGCCGAGGG ACTTTGGCCG ATTACGTGGA GGACTCGCCC AGGTGTTTTT TGAGGTGAAT
 421  TTCCGCGTTC CGGGTCAAAG TCTCCGTTTT ATTATTATAG TCAGCTGACG CGGAGTGTAT
 481  TTATACCCTC TGATCTCGTC AAGAGGCCAC TCTTGAGTGC CAGCGAGTAG AGTTTTCTCC
 541  TCTGCCGCTC TCCGCTCCGC TCCGCTCGGC TCTGACACCG GGGAAAAAAT GAGACATTTC
 601  ACCTACGATG GCGGTGTGCT CACCGGCCAG CTGGCTGCTG AGGTCCTGGA CACCCTGATC
 661  GAGGAGGTAT TGGCCGATAA TTATCCTCCC TCGACTCCTT TTGAGCCACC TACACTTCAC
 721  GAACTATACG ATCTGGATGT GGTGGGGCCC AGCGATCCGA ACGAGCAGGC GGTTTCCAGT
 781  TTTTTTCCAG AGTCCATGTT GTTGGCCAGC CAGGAGGGGG TCGAACTTGA GACCCCTCCT
 841  CCGATCGTGG ATTCCCCCGA TCCGCCGCAG CTGACTAGGC AGCCCGAGCG CTGTGCGGGA
 901  CCTGAGACTA TGCCCCAGCT GCTACCTGAG GTGATCGATC TCACCTGTAA TGAGTCTGGT
 961  TTTCCACCCA GCGAGGATGA GGACGAAGAG GGTGAGCAGT TTGTGTTAGA TTCTGTGGAA
1021  CAACCCGGGC GAGGATGCAG GTCTTGTCAA TATCACCGGA AAAACACAGG AGACTCCCAG
1081  ATTATGTGTT CTCTGTGTTA TATGAAGATG ACCTGTATGT TTATTTACAG TAAGTTTATC
1141  ATCGGTGGGC AGGTGGGCTA TAGTGTGGGT GGTGGTCTTT GGGGGGTTTT TTAATATATG
1201  TCAGGGGTTA TGCTGAAGAC TTTTTTATTG TGATTTTTAA AGGTCCAGTG TCTGAGCCCG
1261  AGCAAGAACC TGAACCGGAG CCTGAGCCTT CTCGCCCCAG GAGAAAGCCT GTAATCTTAA
1321  CTAGACCCAG CGCACCGGTA GCGAGAGGCC TCAGCAGCGC GGAGACCACC GACTCCGGTG
1381  CTTCCTCATC ACCCCCGGAG ATTCACCCCC TGGTGCCCCT ATGTCCCGTT AAGCCCGTTG
1441  CCGTGAGAGT CAGTGGGCGG CGGTCTGCTG TGGAGTGCAT TGAGGACTTG CTTTTTGATT
1501  CACAGGAACC TTTGGACTTG AGCTTGAAAC GCCCCAGGCA TTAAACCTGG TCACCTGGAC
1561  TGAATGAGTT GACGCCTATG TTTGCTTTTG AATGACTTAA TGTGTATAGA TAATAAAGAG
1621  TGAGATAATG TTTTAATTGC ATGGTGTGTT TAACTTGGGC GGAGTCTGCT GGGTATATAA
1681  GCTTCCCTGG GCTAAACTTG GTTACACTTG ACCTCATGGA GGCCTGGGAG TGTTTGGAGA
1741  ACTTTGCCGG AGTTCGTGCC TTGCTGGACG AGAGCTCTAA CAATACCTCT TGGTGGTGGA
1801  GGTATTTGTG GGGCTCTCCC CAGGGCAAGT TAGTTTGTAG AATCAAGGAG GATTACAAGT
1861  GGGAATTTGA AGAGCTTTTG AAATCCTGTG GTGAGCTATT GGATTCTTTG AATCTAGGCC
1921  ACCAGGCTCT CTTCCAGGAG AAGGTCATCA GGACTTTGGA TTTTTCCACA CCGGGGCGCA
1981  TTGCAGCCGC GGTTGCTTTT CTAGCTTTTT TGAAGGATAG ATGGAGCGAA GAGACCCACT
2041  TGAGTTCGGG CTACGTCCTG GATTTTCTGG CCATGCAACT GTGGAGAGCA TGGATCAGAC
2101  ACAAGAACAG GCTGCAACTG TTGTCTTCCG TCCGCCCGTT GCTGATTCCG GCGGAGGAGC
2161  AACAGGCCGG GTCAGAGGAC CGGGCCCGTC GGGATCCGGA GGAGAGGGCA CCGAGGCCGG
2221  GCGAGAGGAG CGCGCTGAAC CTGGGAACCG GGCTGAGCGG CCATCCACAT CGGGAGTGAA
2281  TGTCGGGCAG GTGGTGGATC TTTTTCCAGA ACTGCGGCGG ATTTTGACTA TTAGGGAGGA
2341  TGGGCAATTT GTTAAGGGTC TTAAGAGGGA GAGGGGGGCT TCTGAGCATA ACGAGGAGGC
2401  CAGTAATTTA GCTTTTAGCT TGATGACCAG ACACCGTCCA GAGTGCATCA CTTTTCAGCA
2461  GATTAAGGAC AATTGTGCCA ATGAGTTGGA TCTGTTGGGT CAGAAGTATA GCATAGAGCA
2521  GCTGACCACT TACTGGCTGC AGCCGGGTGA TGATCTGGAG GAAGCTATTA GGGTGTATGC
2581  TAAGGTGGCC CTGCGGCCCG ATTGCAAGTA CAAGCTCAAG GGGCTGGTGA ATATCAGGAA
2641  TTGTTGCTAC ATTTCTGGCA ACGGGGCGGA GGTGGAGATA GAGACCGAAG ACAGGGTGGC
2701  TTTCAGATGC AGCATGATGA ATATGTGGCC GGGGGTGCTG GGCATGGACG GGGTGGTGAT
2761  TATGAATGTG AGGTTCACGG GGCCCAACTT TAACGGCACG GTGTTTTTGG GGAACACCAA
2821  CCTGGTCCTG CACGGGGTGA GCTTCTATGG GTTTAACAAC ACCTGTGTGG AGGCCTGGAC
2881  CGATGTGAAG GTCCGCGGTT GCGCCTTTTA TGGATGTTGG AAGGCCATAG TGAGCCGCCC
2941  TAAGAGCAGG AGTTCCATTA AGAAATGCTT GTTTGAGAGG TGCACCTTGG GGATCCTGGC
3001  CGAGGGCAAC TGCAGGGTGC GCCACAATGT GGCCTCCGAG TGCGGTTGCT TCATGCTAGT
3061  CAAGAGCGTG GCGGTAATCA AGCATAATAT GGTGTGCGGC AACAGCGAGG ACAAGGCCTC
3121  ACAGATGCTG ACCTGCACGG ATGGCAACTG CCACTTGCTG AAGACCATCC ATGTAACCAG
3181  CCACAGCCGG AAGGCCTGGC CCGTGTTCGA GCACAACTTG CTGACCCGCT GCTCCTTGCA
3241  TCTGGGCAAC AGGCGGGGGG TGTTCCTGCC CTATCAATGC AACTTTAGTC ACACCAAGAT
3301  CTTGCTAGAG CCCGAGAGCA TGTCCAAGGT GAACTTGAAC GGGGTGTTTG ACATGACCAT
3361  GAAGATCTGG AAGGTGCTGA GGTACGACGA GACCAGGTCC CGGTGCAGAC CCTGCGAGTG
3421  CGGGGGCAAG CATATGAGGA ACCAGCCCGT GATGCTGGAT GTGACCGAGG AGCTGAGGAC
```

FIG. 5A

```
3481  AGACCACTTG GTTCTGGCCT GCACCAGGGC CGAGTTTGGT TCTAGCGATG AAGACACAGA
3541  TTGAGGTGGG TGAGTGGGCG TGGCCTGGGG TGGTCATGAA AATATATAAG TTGGGGGTCT
3601  TAGGGTCTCT TTATTTGTGT TGCAGAGACC GCCGGAGCCA TGAGCGGGAG CAGCAGCAGC
3661  AGCAGTAGCA GCAGCGCCTT GGATGGCAGC ATCGTGAGCC CTTATTTGAC GACGCGGATG
3721  CCCCACTGGG CCGGGGTGCG TCAGAATGTG ATGGGCTCCA GCATCGACGG CCGACCCGTC
3781  CTGCCCGCAA ATTCCGCCAC GCTGACCTAT GCGACCGTCG CGGGGACGCC GTTGGACGCC
3841  ACCGCCGCCG CCGCCGCCAC CGCAGCCGCC TCGGCCGTGC GCAGCCTGGC CACGGACTTT
3901  GCATTCCTGG GACCACTGGC GACAGGGGCT ACTTCTCGGG CCGCTGCTGC CGCCGTTCGC
3961  GATGACAAGC TGACCGCCCT GCTGGCGCAG TTGGATGCGC TTACTCGGGA ACTGGGTGAC
4021  CTTTCTCAGC AGGTCATGGC CCTGCGCCAG CAGGTCTCCT CCCTGCAAGC TGGCGGGAAT
4081  GCTTCTCCCA CAAATGCCGT TTAAGATAAA TAAAACCAGA CTCTGTTTGG ATTAAAGAAA
4141  AGTAGCAAGT GCATTGCTCT CTTTATTTCA TAATTTTCCG CGCGCGATAG GCCCTAGACC
4201  AGCGTTCTCG GTCGTTGAGG GTGCGGTGTA TCTTCTCCAG GACGTGGTAG AGGTGGCTCT
4261  GGACGTTGAG ATACATGGGC ATGAGCCCGT CCCGGGGGTG GAGGTAGCAC CACTGCAGAG
4321  CTTCATGCTC CGGGGTGGTG TTGTAGATGA TCCAGTCGTA GCAGGAGCGC TGGGCATGGT
4381  GCCTAAAAAT GTCCTTCAGC AGCAGGCCGA TGGCCAGGGG GAGGCCCTTG GTGTAAGTGT
4441  TTACAAAACG GTTAAGTTGG GAAGGGTGCA TTCGGGGAGA GATGATGTGC ATCTTGGACT
4501  GTATTTTTAG ATTGGCGATG TTTCCGCCCA GATCCCTTCT GGGATTCATG TTGTGCAGGA
4561  CCACCAGTAC AGTGTATCCG GTGCACTTGG GGAATTTGTC ATGCAGCTTA GAGGGAAAAG
4621  CGTGGAAGAA CTTGGAGACG CCCTTGTGGC CTCCCAGATT TTCCATGCAT TCGTCCATGA
4681  TGATGGCAAT GGGCCCGCGG GAGGCAGCTT GGGCAAAGAT ATTTCTGGGG TCGCTGACGT
4741  CGTAGTTGTG TTCCAGGGTG AGGTCGTCAT AGGCCATTTT TACAAAGCGC GGGCGGAGGG
4801  TGCCCGACTG GGGGATGATG GTCCCCTCTG GCCCTGGGGC GTAGTTGCCC TCGCAGATCT
4861  GCATTTCCCA GGCCTTAATC TCGGAGGGGG GAATCATATC CACCTGCGGG GCGATGAAGA
4921  AAACGGTTTC CGGAGCCGGG GAGATTAACT GGGATGAGAG CAGGTTTCTA AGCAGCTGTG
4981  ATTTTCCACA ACCGGTGGGC CCATAAATAA CACCTATAAC CGGTTGCAGC TGGTAGTTTA
5041  GAGAGCTGCA GCTGCCGTCG TCCCGGAGGA GGGGGGCCAC CTCGTTGAGC ATGTCCCTGA
5101  CGCGCATGTT CTCCCCGACC AGATCCGCCA GAAGGCGCTC GCCGCCCAGG GACAGCAGCT
5161  CTTGCAAGGA AGCAAAGTTT TTCAGCGGCT TGAGGCCGTC CGCCGTGGGC ATGTTTTTCA
5221  GGGTCTGGCT CAGCAGCTCC AGGCGGTCCC AGAGCTCGGT GACGTGCTCT ACGGCATCTC
5281  TATCCAGCAT ATCTCCTCGT TTCGCGGGTT GGGGCGACTT TCGCTGTAGG GCACCAAGCG
5341  GTGGTCGTCC AGCGGGGCCA AAGTCATGTC CTTCCATGGG CGCAGGGTCC TCGTCAGGGT
5401  GGTCTGGGTC ACGGTGAAGG GGTGCGCTCC GGGCTGAGCG CTTGCCAAGG TGCGCTTGAG
5461  GCTGGTTCTG CTGGTGCTGA AGCGCTGCCG GTCTTCGCCC TGCGCGTCGG CCAGGTAGCA
5521  TTTGACCATG GTGTCATAGT CCAGCCCCTC CGCGGCGTGT CCCTTGGCGC GCAGCTTGCC
5581  CTTGGAGGTG GCGCCGCACG AGGGGCAGAG CAGGCTCTTG AGCGCGTAGA GCTTGGGGGC
5641  GAGGAAGACC GATTCGGGGG AGTAGGCGTC CGCGCCGCAG ACCCCGCACA CGGTCTCGCA
5701  CTCCACCAGC CAGGTGAGCT CGGGGCGCGC CGGGTCAAAA ACCAGGTTTC CCCCATGCTT
5761  TTTGATGCGT TTCTTACCTC GGGTCTCCAT GAGGTGGTGT CCCCGCTCGG TGACGAAGAG
5821  GCTGTCCGTG TCTCCGTAGA CCGACTTGAG GGGTCTTTTC TCCAGGGGGG TCCCTCGGTC
5881  TTCCTCGTAG AGGAACTCGG ACCACTCTGA GACGAAGGCC CGCGTCCAGG CCAGGACGAA
5941  GGAGGCTATG TGGGAGGGGT AGCGGTCGTT GTCCACTAGG GGGTCCACCT TCTCCAAGGT
6001  GTGAAGACAC ATGTCGCCTT CCTCGGCGTC CAGGAAGGTG ATTGGCTTGT AGGTGTAGGC
6061  CACGTGACCG GGGGTTCCTG ACGGGGGGGT ATAAAAGGGG GTGGGGGCGC GCTCGTCGTC
6121  ACTCTCTTCC GCATCGCTGT CTGCGAGGGC CAGCTGCTGG GGTGAGTATT CCCTCTCGAA
6181  GGCGGGCATG ACCTCCGCGC TGAGGTTGTC AGTTTCCAAA AACGAGGAGG ATTTGATGTT
6241  CACCTGTCCC GAGGTGATAC CTTTGAGGGT ACCCGCGTCC ATCTGGTCAG AAAACACGAT
6301  CTTTTTATTG TCCAGCTTGG TGGCGAACGA CCCGTAGAGG GCGTTGGAGA GCAGCTTGGC
6361  GATGGAGCGC AGGGTCTGGT TCTTGTCCCT GTCGGCGCGC TCCTTGGCCG CGATGTTGAG
6421  CTGCACGTAC TCGCGCGCGA CGCAGCGCCA CTCGGGGAAG ACGGTGGTGC GCTCGTCGGG
6481  CACCAGGCGC ACGCGCCAGC CGCGGTTGTG CAGGGTGACC AGGTCCACGC TGGTGGCGAC
6541  CTCGCCGCGC AGGCGCTCGT TGGTCCAGCA GAGACGGCCG CCCTTGCGCG AGCAGAAGGG
6601  GGGCAGGGGG TCGAGCTGGG TCTCGTCCGG GGGGTCCGCG TCCACGGTGA AAACCCCGGG
6661  GCGCAGGCGC GCGTCGAAGT AGTCTATCTT GCAACCTTGC ATGTCCAGCG CCTGCTGCCA
6721  GTCGCGGGCG GCGAGCGCGC GCTCGTAGGG GTTGAGCGGC GGGCCCCAGG GCATGGGGTG
6781  GGTGAGTGCG GAGGCGTACA TGCCGCAGAT GTCATAGACG TAGAGGGGCT CCCGCAGGAC
6841  CCCGATGTAG GTGGGGTAGC AGCGGCCGCC GCGGATGCTG GCGCGCACGT AGTCATACAG
6901  CTCGTGCGAG GGGGCGAGGA GGTCGGGGCC CAGGTTGGTG CGGGCGGGGC GCTCCGCGCG
```

FIG. 5B

```
6961  GAAGACGATC TGCCTGAAGA TGGCATGCGA GTTGGAAGAG ATGGTGGGGC GCTGGAAGAC
7021  GTTGAAGCTG GCGTCCTGCA GGCCGACGGC GTCGCGCACG AAGGAGGCGT AGGAGTCGCG
7081  CAGCTTGTGT ACCAGCTCGG CGGTGACCTG CACGTCGAGC GCGCAGTAGT CGAGGGTCTC
7141  GCGGATGATG TCATATTTAG CCTGCCCCTT CTTTTTCCAC AGCTCGCGGT TGAGGACAAA
7201  CTCTTCGCGG TCTTTCCAGT ACTCTTGGAT CGGGAAACCG TCCGGTTCCG AACGGTAAGA
7261  GCCTAGCATG TAGAACTGGT TGACGGCCTG GTAGGCGCAG CAGCCCTTCT CCACGGGGAG
7321  GGCGTAGGCC TGCGCGGCCT TGCGGAGCGA GGTGTGGGTC AGGGCGAAGG TGTCCCTGAC
7381  CATGACTTTG AGGTACTGGT GCTTGAAGTC GGAGTCGTCG CAGCCGCCCC GCTCCCAGAG
7441  CGAGAAGTCG GTGCGCTTCT TGGAGCGGGG GTTGGGCAGA GCGAAGGTGA CATCGTTGAA
7501  GAGGATTTTG CCCGCGCGGG GCATGAAGTT GCGGGTGATG CGGAAGGGCC CCGGCACTTC
7561  AGAGCGGTTG TTGATGACCT GGGCGGCGAG CACGATCTCG TCGAAGCCGT TGATGTTGTG
7621  GCCCACGATG TAGAGTTCCA GGAAGCGGGG CCGGCCCTTT ACGGTGGGCA GCTTCTTTAG
7681  CTCTTCGTAG GTGAGCTCCT CGGGCGAGGC GAGGCCGTGC TCGGCCAGGG CCCAGTCCGC
7741  GAGGTGCGGG TTGTCTCTGA GGAAGGACTC CCAGAGGTCG CGGGCCAGGA GGGTCTGCAG
7801  GCGGTCCCTG AAGGTCCTGA ACTGGCGGCC CACGGCCATT TTTTCGGGGG TGATGCAGTA
7861  GAAGGTGAGG GGGTCTTGCT GCCAGCGGTC CCAGTCGAGC TGCAGGGCGA GGTCGCGCGC
7921  GGCGGTGACC AGGCGCTCGT CGCCCCCGAA TTTCATGACC AGCATGAAGG GCACGAGCTG
7981  CTTTCCGAAG GCCCCCATCC AAGTGTAGGT CTCTACATCG TAGGTGACAA AGAGGCGCTC
8041  CGTGCGAGGA TGCGAGCCGA TCGGGAAGAA CTGGATCTCC CGCCACCAGT TGGAGGAGTG
8101  GCTGTTGATG TGGTGGAAGT AGAAGTCCCG TCGCCGGGCC GAACACTCGT GCTGGCTTTT
8161  GTAAAAGCGA GCGCAGTACT GGCAGCGCTG CACGGGCTGT ACCTCCTGCA CGAGATGCAC
8221  CTTTCGCCCG CGCACGAGGA AGCCGAGGGG AAATCTGAGC CCCCCGCCTG GCTCGCGGCA
8281  TGGCTGGTGC TCTTCTACTT TGGATGCGTG TCCGTCTCCG TCTGGCTCCT CGAGGGGTGT
8341  TACGGTGGAG CGGACCACCA CGCCGCGCGA GCCGCAGGTC CAGATATCGG CGCGCGGCGG
8401  TCGGAGTTTG ATGACGACAT CGCGCAGCTG GGAGCTGTCC ATGGTCTGGA GCTCCCGCGG
8461  CGGCGGCAGG TCAGCCGGGA GTTCTTGCAG GTTCACCTCG CAGAGTCGGG CCAGGGCGCG
8521  GGGCAGGTCT AGGTGGTACC TGATCTCTAG GGGCGTGTTG GTGGCGGCGT CGATGGCTTG
8581  CAGGAGCCCG CATCCCCGGG GGGCGACGAC GGTGCCCCGC GGGGTGGTGG TGGTGGTGGT
8641  GGTGGTGGTG GTGGCGGTGC AGCTCAGAAG CGGTGCCGCG GGCGGGCCCC CGGAGGTAGG
8701  GGGGGCTCCG GTCCCGCCGG CAGGGGCGGC AGCGGCACGT CGGCGTGGAG CGCGGGCAGG
8761  AGTTGGTGCT GTGCCCGGAG GTTGCTGGCG AAGGCGACGA CGCGGCGGTT GATCTCCTGG
8821  ATCTGGCGCC TCTGCGTGAA GACGACGGGC CCGGTGAGCT TGAACCTGAA AGAGAGTTCG
8881  ACAGAATCAA TCTCGGTGTC ATTGACCGCG GCCTGGCGCA GGATCTCCTG CACGTCTCCC
8941  GAGTTGTCTT GGTAGGCGAT CTCGGCCATG AACTGCTCGA TCTCTTCCTC CTGGAGGTCT
9001  CCGCGTCCGG CGCGTTCCAC GGTGGCCGCC AGGTCGTTGG AGATGCGCCC CATGAGCTGC
9061  GAGAAGGCGT TGAGTCCGCC CTCGTTCCAG ACTCGGCTGT AGACCACGCC CCCCTGGTCA
9121  TCGCGGGCGC GCATGACCAC CTGCGCGAGG TTGAGCTCCA CGTGCCGCGC GAAGACGGCG
9181  TAGTTGCGCA GACGCTGGAA GAGGTAGTTG AGGGTGGTGG CGGTGTGCTC GGCCACGAAG
9241  AAGTTCATGA CCCAGCGGCG CAACGTGGAT TCGTTGATGT CCCCCAAGGC CTCCAGCCGT
9301  TCCATGGCCT CGTAGAAGTC CACGGCGAAG TTGAAAAACT GGGAGTTGCG CGCCGACACG
9361  GTCAACTCCT CCTCCAGAAG ACGGATGAGC TCGGCGACGG TGTCGCGCAC CTCGCGCTCG
9421  AAGGCTATGG GGATCTCTTC CTCCGCTAGC ATCACCACCT CCTCCTCTTC CTCCTCTTCT
9481  GGCACTTCCA TGATGGCTTC CTCCTCTTCG GGGGGCGGCG GCGGCGGCGG TGGGGGAGGG
9541  GGCGCTCTGC GCCGGCGGCG GCGCACCGGG AGGCGGTCCA CGAAGCGCGC GATCATCTCC
9601  CCGCGGCGGC GGCGCATGGT CTCGGTGACG GCGCGGCCGT TCTCCCGGGG GCGCAGTTGG
9661  AAGACGCCGC CGGACATCTG GTGCTGGGGC GGGTGGCCGT GAGGCAGCGA AACGGCGCTG
9721  ACGATGCATC TCAACAATTG CTGCGTAGGT ACGCCGCCGA GGGACCTGAG GGAGTCCATA
9781  TCCACCGGAT CCGAAAACCT TTCGAGGAAG GCGTCTAACC AGTCGCAGTC GCAAGGTAGG
9841  CTGAGCACCG TGGCGGGCGG CGGGGGGTGG GGGGAGTGTC TGGCGGAGGT GCTGCTGATG
9901  ATGTAATTGA AGTAGGCGGA CTTGACACGG CGGATGGTCG ACAGGAGCAC CATGTCCTTG
9961  GGTCCGGCCT GCTGGATGCG GAGGCGGTCG GCTATGCCCC AGGCTTCGTT CTGGCATCGG
10021 CGCAGGTCCT TGTAGTAGTC TTGCATGAGC CTTTCCACCG GCACCTCTTC TCCTTCCTCT
10081 TCTGCTTCTT CCATGTCTGC TTCGGCCCTG GGGCGGCGCC GCGCCCCCCT GCCCCCCATG
10141 CGCGTGACCC CGAACCCCCT GAGCGGTTGG AGCAGGGCCA GGTCGGCGAC GACGCGCTCG
10201 GCCAGGATGG CCTGCTGCAC CTGCGTGAGG GTGGTTTGGA AGTCATCCAA GTCCACGAAG
10261 CGGTGGTAGG CGCCCGTGTT GATGGTGTAG GTGCAGTTGG CCATGACGGA CCAGTTGACG
10321 GTCTGGTGGC CCGGTTGCGA CATCTCGGTG TACCTGAGTC GCGAGTAGGC GCGGGAGTCG
10381 AAGACGTAGT CGTTGCAAGT CCGCACCAGG TACTGGTAGC CCACCAGGAA GTGCGGCGGC
```

FIG. 5C

```
10441  GGCTGGCGGT AGAGGGGCCA GCGCAGGGTG GCGGGGGCTC CGGGGGCCAG GTCTTCCAGC
10501  ATGAGGCGGT GGTAGGCGTA GATGTACCTG GACATCCAGG TGATACCCGC GGCGGTGGTG
10561  GAGGCGCGCG GGAAGTCGCG CACCCGGTTC CAGATGTTGC GCAGGGGCAG AAAGTGCTCC
10621  ATGGTAGGCG TGCTCTGTCC AGTCAGACGC GCGCAGTCGT TGATACTCTA GACCAGGGAA
10681  AACGAAAGCC GGTCAGCGGG CACTCTTCCG TGGTCTGGTG AATAGATCGC AAGGGTATCA
10741  TGGCGGAGGG CCTCGGTTCG AGCCCCGGGT CCGGGCCGGA CGGTCCGCCA TGATCCACGC
10801  GGTTACCGCC CGCGTGTCGA ACCCAGGTGT GCGACGTCAG ACAACGGTGG AGTGTTCCTT
10861  TTGGCGTTTT TCTGGCCGGG CGCCGGCGTC GCGTAAGAGA CTAAGCCGCG AAAGCGAAAG
10921  CAGTAAGTGG CTCGCTCCCC GTAGCCGGAG GGATCCTTGC TAAGGGTTGC GTTGCGGCGA
10981  ACCCCGGTTC GAATCCCGTA CTCGGGCCGG CCGGACCCGC GGCTAAGGTG TTGGATTGGC
11041  CTCCCCCTCG TATAAAGACC CCGCTTGCGG ATTGACTCCG GACACGGGGA CGAGCCCCTT
11101  TTATTTTTGC TTTCCCCAGA TGCATCCGGT GCTGCGGCAG ATGCGCCCCC CGCCCCAGCA
11161  GCAGCAACAA CACCAGCAAG AGCGGCAGCA ACAGCAGCGG GAGTCATGCA GGGCCCCCTC
11221  ACCCACCCTC GGCGGGCCGG CCACCTCGGC GTCCGCGGCC GTGTCTGGCG CCTGCGGCGG
11281  CGGCGGGGGG CCGGCTGACG ACCCCGAGGA GCCCCCGCGG CGCAGGGCCA GACACTACCT
11341  GGACCTGGAG GAGGGCGAGG GCCTGGCGCG GCTGGGGGCG CCGTCTCCCG AGCGCCACCC
11401  GCGGGTGCAG CTGAAGCGCG ACTCGCGCGA GGCGTACGTG CCTCGGCAGA ACCTGTTCAG
11461  GGACCGCGCG GGCGAGGAGC CCGAGGAGAT GCGGGACAGG AGGTTCAGCG CAGGGCGGGA
11521  GCTGCGGCAG GGGCTGAACC GCGAGCGGCT GCTGCGCGAG GAGGACTTTG AGCCCGACGC
11581  GCGGACGGGG ATCAGCCCCG CGCGCGCGCA CGTGGCGGCC GCCGACCTGG TGACGGCGTA
11641  CGAGCAGACG GTGAACCAGG AGATCAACTT CCAAAAGAGT TTCAACAACC ACGTGCGCAC
11701  GCTGGTGGCG CGCGAGGAGG TGACCATCGG GCTGATGCAC CTGTGGGACT TTGTAAGCGC
11761  GCTGGTGCAG AACCCCAACA GCAAGCCTCT GACGGCGCAG CTGTTCCTGA TAGTGCAGCA
11821  CAGCAGGGAC AACGAGGCGT TTAGGGACGC GCTGCTGAAC ATCACCGAGC CCGAGGGTCG
11881  GTGGCTGCTG GACCTGATTA ACATCCTGCA GAGCATAGTG GTGCAGGAGC GCAGCCTGAG
11941  CCTGGCCGAC AAGGTGGCGG CCATCAACTA CTCGATGCTG AGCCTGGGCA AGTTTTACGC
12001  GCGCAAGATC TACCAGACGC CGTACGTGCC CATAGACAAG GAGGTGAAGA TCGACGGTTT
12061  TTACATGCGC ATGGCGCTGA AGGTGCTCAC CCTGAGCGAC GACCTGGGCG TGTACCGCAA
12121  CGAGCGCATC CACAAGGCCG TGAGCGTGAG CCGGCGGCGC GAGCTGAGCG ACCGCGAGCT
12181  GATGCACAGC CTGCAGCGGG CGCTGGCGGG CGCCGGCAGC GGCGACAGGG AGGCGGAGTC
12241  CTACTTCGAT GCGGGGGCGG ACCTGCGCTG GGCGCCCAGC CGGCGGGCCC TGGAGGCCGC
12301  GGGGGTCCGC GAGGACTATG ACGAGGACGG CGAGGAGGAT GAGGAGTACG AGCTAGAGGA
12361  GGGCGAGTAC CTGGACTAAA CCGCGGGTGG TGTTTCCGGT AGATGCAAGA CCCGAACGTG
12421  GTGGACCCGG CGCTGCGGGC GGCTCTGCAG AGCCAGCCGT CCGGCCTTAA CTCCTCAGAC
12481  GACTGGCGAC AGGTCATGGA CCGCATCATG TCGCTGACGG CGCGTAACCC GGACGCGTTC
12541  CGGCAGCAGC CGCAGGCCAA CAGGCTCTCC GCCATCCTGG AGGCGGTGGT GCCTGCGCGC
12601  TCGAACCCCA CGCACGAGAA GGTGCTGGCC ATAGTGAACG CGCTGGCCGA GAACAGGGCC
12661  ATCCGCCCGG ACGAGGCCGG GCTGGTGTAC GACGCGCTGC TGCAGCGCGT GGCCCGCTAC
12721  AACAGCGGCA ACGTGCAGAC CAACCTGGAC CGGCTGGTGG GGGACGTGCG CGAGGCGGTG
12781  GCGCAGCGCG AGCGCGCGGA TCGGCAGGGC AACCTGGGCT CCATGGTGGC GCTGAATGCC
12841  TTCCTGAGCA CGCAGCCGGC CAACGTGCCG CGGGGGCAGG AAGACTACAC CAACTTTGTG
12901  AGCGCGCTGC GGCTGATGGT GACCGAGACC CCCCAGAGCG AGGTGTACCA GTCGGGCCCG
12961  GACTACTTCT TCCAGACCAG CAGACAGGGC CTGCAGACGG TGAACCTGAG CCAGGCTTTC
13021  AAGAACCTGC GGGGGCTGTG GGGCGTGAAG GCGCCCACCG GCGACCGGGC GACGGTGTCC
13081  AGCCTGCTGA CGCCCAACTC GCGCCTGCTG CTGCTGCTGA TCGCGCCGTT CACGGACAGC
13141  GGCAGCGTGT CCCGGGACAC CTACCTGGGG CACCTGCTGA CCCTGTACCG CGAGGCCATC
13201  GGGCAGGCGC AGGTGGACGA GCACACCTTC CAGGAGATCA CCAGCGTGAG CCGCGCGCTG
13261  GGGCAGGAGG ACACGAGCAG CCTGGAGGCG ACTCTGAACT ACCTGCTGAC CAACCGGCGG
13321  CAGAAGATTC CCTCGCTGCA CAGCCTGACC TCCGAGGAGG AGCGCATCTT GCGCTACGTG
13381  CAGCAGAGCG TGAGCCTGAA CCTGATGCGC GACGGGGTGA CGCCCAGCGT GGCGCTGGAC
13441  ATGACCGCGC GCAACATGGA ACCGGGCATG TACGCCGCGC ACCGGCCTTA CATCAACCGC
13501  CTGATGGACT ACCTGCATCG CGCGGCGGCC GTGAACCCCG AGTACTTTAC CAACGCCATC
13561  CTGAACCCGC ACTGGCTCCC GCCGCCCGGG TTCTACAGCG GGGGCTTCGA GGTCCCGGAG
13621  GCCAACGATG GCTTCCTGTG GGACGACATG GACGACAGCG TGTTCTCCCC GCGGCCGCAG
13681  GCGCTGGCGG AAGCGTCCCT GCTGCGTCCC AAGAAGGAGG AGGAGGAGGC GAGTCGCCGC
13741  CGCGGCAGCA GCGGCGTGGC TTCTCTGTCC GAGCTGGGGG CGGCAGCCGC CGCGCGCCCC
13801  GGGTCCCTGG GCGGCAGCCC CTTTCCGAGC CTGGTGGGGT CTCTGCACAG CGAGCGCACC
13861  ACCCGCCCTC GGCTGCTGGG CGAGGACGAG TACCTGAATA ACTCCCTGCT GCAGCCGGTG
```

FIG. 5D

```
13921  CGGGAGAAAA ACCTGCCCCC CGCCTTCCCC AACAACGGGA TAGAGAGCCT GGTGGACAAG
13981  ATGAGCAGAT GGAAGACCTA TGCGCAGGAG CACAGGGACG CGCCCGCGCT CCGGCCGCCC
14041  ACGCGGCGCC AGCGCCACGA CCGGCAGCGG GGGCTGGTGT GGGATGACGA GGACTCCGCG
14101  GACGATAGCA GCGTGCTGGA CCTGGGAGGG AGCGGCAACC CGTTCGCGCA CCTGCGCCCC
14161  CGCCTGGGGA GGATGTTTTA AAAAAAAAAA AAGCAAGAAG CATGATGCAA AATTAAATAA
14221  AACTCACCAA GGCCATGGCG ACCGAGCGTT GGTTTCTTGT GTTCCCTTCA GTATGCGGCG
14281  CGCGGCGATG TACCAGGAGG GACCTCCTCC CTCTTACGAG AGCGTGGTGG GCGCGGCGGC
14341  GGCGGCGCCC TCTTCTCCCT TTGCGTCGCA GCTGCTGGAG CCGCCGTACG TGCCTCCGCG
14401  CTACCTGCGG CCTACGGGGG GGAGAAACAG CATCCGTTAC TCGGAGCTGG CGCCCCTGTT
14461  CGACACCACC CGGGTGTACC TGGTGGACAA CAAGTCGGCG GACGTGGCCT CCCTGAACTA
14521  CCAGAACGAC CACAGCAATT TTTTGACCAC GGTCATCCAG AACAATGACT ACAGCCCGAG
14581  CGAGGCCAGC ACCCAGACCA TCAATCTGGA TGACCGGTCG CACTGGGGCG GCGACCTGAA
14641  AACCATCCTG CACACCAACA TGCCCAACGT GAACGAGTTC ATGTTCACCA ATAAGTTCAA
14701  GGCGCGGGTG ATGGTGTCGC GCTCGCACAC CAAGGAAGAC CGGGTGGAGC TGAAGTACGA
14761  GTGGGTGGAG TTCGAGCTGC CAGAGGGCAA CTACTCCGAG ACCATGACCA TTGACCTGAT
14821  GAACAACGCG ATCGTGGAGC ACTATCTGAA AGTGGGCAGG CAAAACGGGG TCCTGGAGAG
14881  CGACATCGGG GTCAAGTTCG ACACCAGGAA CTTCCGCCTG GGGCTGGACC CCGTGACCGG
14941  GCTGGTTATG CCCGGGGTGT ACACCAACGA GGCCTTCCAT CCCGACATCA TCCTGCTGCC
15001  CGGCTGCGGG GTGGACTTCA CTTACAGCCG CCTGAGCAAC CTCCTGGGCA TCCGCAAGCG
15061  GCAGCCCTTC CAGGAGGGCT TCAGGATCAC CTACGAGGAC CTGGAGGGGG GCAACATCCC
15121  CGCGCTCCTC GATGTGGAGG CCTACCAGGA TAGCTTGAAG GAAAATGAGG CGGGACAGGA
15181  GGATACCACC CCCGCCGCCT CCGCCGCCGC CGAGCAGGGC GAGGATGCTG CTGACACCGC
15241  GGCCGCGGAC GGGGCAGAGG CCGACCCCGC TATGGTGGTG GAGGCTCCCG AGCAGGAGGA
15301  GGATATGAAT GACAGTGCGG TGCGCGGAGA CACCTTCGTC ACCCGGGGGG AGGAAAAGCA
15361  AGCGGAGGCC GAGGCCGCGG CCGAGGAAAA GCAACTGGCG GCAGCAGCGG CGGCGGCGGC
15421  GTTGGCCGCG GCGGAGGCTG AGTCTGAGGG GACCAAGCCC GCCAAGGAGC CCGTGATTAA
15481  GCCCCTGACC GAAGATAGCA AGAAGCGCAG TTACAACCTG CTCAAGGACA GCACCAACAC
15541  CGCGTACCGC AGCTGGTACC TGGCCTACAA CTACGGCGAC CCGTCGACGG GGGTGCGCTC
15601  CTGGACCCTG CTGTGCACGC CGGACGTGAC CTGCGGCTCG GAGCAGGTGT ACTGGTCGCT
15661  GCCCGACATG ATGCAAGACC CCGTGACCTT CCGCTCCACG CGGCAGGTCA GCAACTTCCC
15721  GGTGGTGGGC GCCGAGCTGC TGCCCGTGCA CTCCAAGAGC TTCTACAACG ACCAGGCCGT
15781  CTACTCCCAG CTCATCCGCC AGTTCACCTC TCTGACCCAC GTGTTCAATC GCTTTCCTGA
15841  GAACCAGATT CTGGCGCGCC CGCCCGCCCC CACCATCACC ACCGTCAGTG AAAACGTTCC
15901  TGCTCTCACA GATCACGGGA CGCTACCGCT GCGCAACAGC ATCGGAGGAG TCCAGCGAGT
15961  GACCGTTACT GACGCCAGAC GCCGCACCTG CCCCTACGTT TACAAGGCCT TGGGCATAGT
16021  CTCGCCGCGC GTCCTTTCCA GCCGCACTTT TTGAGCAACA CCACCATCAT GTCCATCCTG
16081  ATCTCACCCA GCAATAACTC CGGCTGGGGA CTGCTGCGCG CGCCCAGCAA GATGTTCGGA
16141  GGGGCGAGGA AGCGTTCCGA GCAGCACCCC GTGCGCGTGC GCGGGCACTT CCGCGCCCCC
16201  TGGGGAGCGC ACAAACGCGG CCGCGCGGGG CGCACCACCG TGGACGACGC CATCGACTCG
16261  GTGGTGGAGC AGGCGCGCAA CTACAGGCCC GCGGTCTCTA CCGTGGACGC GGCCATCCAG
16321  ACCGTGGTGC GGGGCGCGCG GCGGTACGCC AAGCTGAAGA GCCGCCGGAA GCGCGTGGCC
16381  CGCCGCCACC GCCGCCGACC CGGGGCCGCC GCCAAACGCG CCGCCGCGGC CCTGCTTCGC
16441  CGGGCCAAGC GCACGGGCCG CCGCGCCGCC ATGAGGGCCG CGCGCCGCTT GGCCGCCGGC
16501  ATCACCGCCG CCACCATGGC CCCCCGTACC CGAAGACGCG CGGCCGCCGC CGCCGCCGCC
16561  GCCATCAGTG ACATGGCCAG CAGGCGCCGG GGCAACGTGT ACTGGGTGCG CGACTCGGTG
16621  ACCGGCACGC GCGTGCCCGT GCGCTTCCGC CCCCCGCGGA CTTGAGATGA TGTGAAAAAA
16681  CAACACTGAG TCTCCTGCTG TTGTGTGTAT CCCAGCGGCG GCGGCGCGCG CAGCGTCATG
16741  TCCAAGCGCA AAATCAAAGA AGAGATGCTC CAGGTCGTCG CGCCGGAGAT CTATGGGCCC
16801  CCGAAGAAGG AAGAGCAGGA TTCGAAGCCC CGCAAGATAA AGCGGGTCAA AAAGAAAAAG
16861  AAAGATGATG ACGATGCCGA TGGGGAGGTG GAGTTCCTGC GCGCCACGGC GCCCAGGCGC
16921  CCGGTGCAGT GGAAGGGCCG GCGCGTAAAG CGCGTCCTGC GCCCCGGCAC CGCGGTGGTC
16981  TTCACGCCCG GCGAGCGCTC CACCCGGACT TTCAAGCGCG TCTATGACGA GGTGTACGGC
17041  GACGAAGACC TGCTGGAGCA GGCCAACGAG CGCTTCGGAG AGTTTGCTTA CGGGAAGCGT
17101  CAGCGGGCGC TGGGGAAGGA GGACCTGCTG GCGCTGCCGC TGGACCAGGG CAACCCCACC
17161  CCCAGTCTGA AGCCCGTGAC CCTGCAGCAG GTGCTGCCGA GCAGCGCACC CTCCGAGGCG
17221  AAGCGGGGTC TGAAGCGCGA GGGCGGCGAC CTGGCGCCCA CCGTGCAGCT CATGGTGCCC
17281  AAGCGGCAGA GGCTGGAGGA TGTGCTGGAG AAAATGAAAG TAGACCCCGG TCTGCAGCCG
17341  GACATCAGGG TCCGTCCCAT CAAGCAGGTG GCGCCGGGCC TCGGCGTGCA GACCGTGGAC
```

FIG. 5E

```
17401  GTGGTCATCC CCACCGGCAA CTCCCCCGCC GCCACCACCA CTACCGCTGC CTCCACGGAC
17461  ATGGAGACAC AGACCGATCC CGCCGCAGCC GCAGCCGCCG CCGCAGCCGC GACCTCCTCG
17521  GCGGAGGTGC AGACGGACCC CTGGCTGCCG CCGGCGATGT CAGCTCCCCG CGCGCGCCGC
17581  GGACGCAGAA AGTACGGCGC CGCCAACGCG CTCCTGCCCG AGTACGCCTT GCATCCTTCC
17641  ATCGCGCCCA CCCCCGGCTA CCGAGGCTAT ACCTACCGCC CGCGAAGAGC CAAGGGTTCC
17701  ACCCGCCGTC CCCGCCGACG CGCCGCCGCC ACCACCCGCC GCCGCCGCCG CAGACGCCAG
17761  CCCGCACTGG CTCCAGTCTC CGTGAGGAGA GTGGCGCGCG ACGGACACAC CCTGGTGCTG
17821  CCCAGGGCGC GCTACCACCC CAGCATCGTT TAAAAGCCTG TTGTGGTTCT TGCAGATATG
17881  GCCCTCACTT GCCGCCTCCG TTTCCCGGTG CCGGGATACC GAGGAGGAAG ATCGCGCCGC
17941  AGGAGGGGTC TGGCCGGCCG CGGCCTGAGC GGAGGCAGCC GCCGCGCGCA CCGGCGGCGA
18001  CGCGCCACCA GCCGACGCAT GCGCGGCGGG GTGCTGCCCC TGTTAATCCC CCTGATCGCC
18061  GCGGCGATCG GCGCCGTGCC CGGGATCGCC TCCGTGGCCT TGCAAGCGTC CCAGAGGCAT
18121  TGACAGACTT GCAAACTTGC AAATATGGAA AAAAAAAAAA AACCCCAATA AAAAGTCTAG
18181  ACTCTCACGC TCGCTTGGTC CTGTGACTAT TTTGTAGAAT GGAAGACATC AACTTTGCGT
18241  CGCTGGCCCC GCGTCACGGC TCGCGCCCGT TCCTGGGACA CTGGAACGAT ATCGGCACCA
18301  GCAACATGAG CGGTGGCGCC TTCAGTTGGG GCTCTCTGTG GAGCGGCATT AAAAGTATCG
18361  GGTCTGCCGT TAAAAATTAC GGCTCCCGGG CCTGGAACAG CAGCACGGGC CAGATGTTGA
18421  GAGACAAGTT GAAAGAGCAG AACTTCCAGC AGAAGGTGGT GGAGGGCCTG GCCTCCGGCA
18481  TCAACGGGGT GGTGGACCTG GCCAACCAGG CCGTGCAGAA TAAAATCAAC AGCAGACTGG
18541  ACCCCCGGCC GCCGGTGGAG GAGGTGCCGC CGGCGCTGGA GACGGTGTCC CCCGATGGGC
18601  GTGGCGAGAA GCGCCCGCGG CCCGATAGGG AAGAGACCAC TCTGGTCACG CAGACCGATG
18661  AGCCGCCCCC GTATGAGGAG GCCCTAAAGC AAGGTCTGCC CACCACGCGG CCCATCGCGC
18721  CCATGGCCAC CGGGGTGGTG GGCCGCCACA CCCCCGCCAC GCTGGACTTG CCTCCGCCCG
18781  CCGATGTGCC GCAGCAGCAG AAGGCGGCAC AGCCGGGCCC GCCCGCGACC GCCTCCCGTT
18841  CCTCCGCCGG TCCTCTGCGC CGCGCGGCCA GCGGCCCCCG CGGGGGGGTC GCGAGGCACG
18901  GCAACTGGCA GAGCACGCTG AACAGCATCG TGGGTCTGGG GGTGCGGTCC GTGAAGCGCC
18961  GCCGATGCTA CTGAATAGCT TAGCTAACGT GTTGTATGTG TGTATGCGCC CTATGTCGCC
19021  GCCAGAGGAG CTGCTGAGTC GCCGCCGTTC GCGCGCCCAC CACCACCGCC ACTCCGCCCC
19081  TCAAGATGGC GACCCCATCG ATGATGCCGC AGTGGTCGTA CATGCACATC TCGGGCCAGG
19141  ACGCCTCGGA GTACCTGAGC CCCGGGCTGG TGCAGTTCGC CCGCGCCACC GAGAGCTACT
19201  TCAGCCTGAG TAACAAGTTT AGGAACCCCA CGGTGGCGCC CACGCACGAT GTGACCACCG
19261  ACCGGTCTCA GCGCCTGACG CTGCGGTTCA TTCCCGTGGA CCGCGAGGAC ACCGCGTACT
19321  CGTACAAGGC GCGGTTCACC CTGGCCGTGG GCGACAACCG CGTGCTGGAC ATGGCCTCCA
19381  CCTACTTTGA CATCCGCGGG GTGCTGGACC GGGGTCCCAC TTTCAAGCCC TACTCTGGCA
19441  CCGCCTACAA CTCCCTGGCC CCCAAGGGCG CTCCCAACTC CTGCGAGTGG GAGCAAGAGG
19501  AAACTCAGGC AGTTGAAGAA GCAGCAGAAG AGGAAGAAGA AGATGCTGAC GGTCAAGCTG
19561  AGGAAGAGCA AGCAGCTACC AAAAAGACTC ATGTATATGC TCAGGCTCCC CTTTCTGGCG
19621  AAAAAATTAG TAAAGATGGT CTGCAAATAG GAACGGACGC TACAGCTACA GAACAAAAAC
19681  CTATTTATGC AGACCCTACA TTCCAGCCCG AACCCCAAAT CGGGGAGTCC CAGTGGAATG
19741  AGGCAGATGC TACAGTCGCC GGCGGTAGAG TGCTAAAGAA ATCTACTCCC ATGAAACCAT
19801  GCTATGGTTC CTATGCAAGA CCCACAAATG CTAATGGAGG TCAGGGTGTA CTAACGGCAA
19861  ATGCCCAGGG ACAGCTAGAA TCTCAGGTTG AAATGCAATT CTTTTCAACT TCTGAAAACG
19921  CCCGTAACGA GGCTAACAAC ATTCAGCCCA AATTGGTGCT GTATAGTGAG GATGTGCACA
19981  TGGAGACCCC GGATACGCAC CTTTCTTACA AGCCCGCAAA AAGCGATGAC AATTCAAAAA
20041  TCATGCTGGG TCAGCAGTCC ATGCCCAACA GACCTAATTA CATCGGCTTC AGAGACAACT
20101  TTATCGGCCT CATGTATTAC AATAGCACTG GCAACATGGG AGTGCTTGCA GGTCAGGCCT
20161  CTCAGTTGAA TGCAGTGGTG GACTTGCAAG ACAGAAACAC AGAACTGTCC TACCAGCTCT
20221  TGCTTGATTC CATGGGTGAC AGAACCAGAT ACTTTTCCAT GTGGAATCAG GCAGTGGACA
20281  GTTATGACCC AGATGTTAGA ATTATTGAAA ATCATGGAAC TGAAGACGAG CTCCCCAACT
20341  ATTGTTTCCC TCTGGGTGGC ATAGGGGTAA CTGACACTTA CCAGGCTGTT AAAACCAACA
20401  ATGGCAATAA CGGGGGCCAG GTGACTTGGA CAAAAGATGA AACTTTTGCA GATCGCAATG
20461  AAATAGGGGT GGGAAACAAT TTCGCTATGG AGATCAACCT CAGTGCCAAC CTGTGGAGAA
20521  ACTTCCTGTA CTCCAACGTG GCGCTGTACC TACCAGACAA GCTTAAGTAC AACCCCTCCA
20581  ATGTGGACAT CTCTGACAAC CCCAACACCT ACGATTACAT GAACAAGCGA GTGGTGGCCC
20641  CGGGGCTGGT GGACTGCTAC ATCAACCTGG GCGCGCGCTG GTCGCTGGAC TACATGGACA
20701  ACGTCAACCC CTTCAACCAC CACCGCAATG CGGGCCTGCG CTACCGCTCC ATGCTCCTGG
20761  GCAACGGGCG CTACGTGCCC TTCCACATCC AGGTGCCCCA GAAGTTCTTT GCCATCAAGA
20821  ACCTCCTCCT CCTGCCGGGC TCCTACACCT ACGAGTGGAA CTTCAGGAAG GATGTCAACA
```

FIG. 5F

```
20881  TGGTCCTCCA GAGCTCTCTG GGTAACGATC TCAGGGTGGA CGGGGCCAGC ATCAAGTTCG
20941  AGAGCATCTG CCTCTACGCC ACCTTCTTCC CCATGGCCCA CAACACGGCC TCCACGCTCG
21001  AGGCCATGCT CAGGAACGAC ACCAACGACC AGTCCTTCAA TGACTACCTT TCCGCCGCCA
21061  ACATGCTCTA CCCCATACCC GCCAACGCCA CCAACGTCCC CATCTCCATC CCCTCGCGCA
21121  ACTGGGCGGC CTTCCGCGGC TGGGCCTTCA CCCGCCTCAA GACCAAGGAG ACCCCCTCCC
21181  TGGGCTCGGG ATTCGACCCC TACTACACCT ACTCGGGCTC TATTCCCTAC CTGGACGGCA
21241  CCTTCTACCT CAACCACACT TTCAAGAAGG TCTCGGTCAC CTTCGACTCC TCGGTCAGCT
21301  GGCCGGGCAA CGACCGTCTG CTCACCCCCA ACGAGTTCGA GATCAAGCGC TCGGTCGACG
21361  GGGAAGGCTA CAACGTGGCC CAGTGCAACA TGACCAAGGA CTGGTTCCTG GTCCAGATGC
21421  TGGCCAACTA CAACATCGGC TACCAGGGCT TCTACATCCC AGAGAGCTAC AAGGACAGGA
21481  TGTACTCCTT CTTCAGGAAC TTCCAGCCCA TGAGCCGGCA GGTGGTGGAC CAGACCAAGT
21541  ACAAGGACTA CCAGGAGGTG GGCATCATCC ACCAGCACAA CAACTCGGGC TTCGTGGGCT
21601  ACCTCGCCCC CACCATGCGC GAGGGACAGG CCTACCCCGC CAACTTCCCC TACCCGCTCA
21661  TAGGCAAGAC CGCGGTCGAC AGCATCACCC AGAAAAAGTT CCTCTGCGAC CGCACCCTCT
21721  GGCGCATCCC CTTCTCCAGC AACTTCATGT CCATGGGTGC GCTCTCGGAC CTGGGCCAGA
21781  ACTTGCTCTA CGCCAACTCC GCCCACGCCC TCGACATGAC CTTCGAGGTC GACCCCATGG
21841  ACGAGCCCAC CCTTCTCTAT GTTCTGTTCG AAGTCTTTGA CGTGGTCCGG GTCCACCAGC
21901  CGCACCGCGG CGTCATCGAG ACCGTGTACC TGCGTACGCC CTTCTCGGCC GGCAACGCCA
21961  CCACCTAAAG AAGCAAGCCG CAGTCATCGC CGCCTGCATG CCGTCGGGTT CCACCGAGCA
22021  AGAGCTCAGG GCCATCGTCA GAGACCTGGG ATGCGGGCCC TATTTTTTGG GCACCTTCGA
22081  CAAGCGCTTC CCTGGCTTTG TCTCCCCACA CAAGCTGGCC TGCGCCATCG TCAACACGGC
22141  CGGCCGCGAG ACCGGGGGCG TGCACTGGCT GGCCTTTGCC TGGAACCCGC GCTCCAAAAC
22201  ATGCTTCCTC TTTGACCCCT TCGGCTTTTC GGACCAGCGG CTCAAGCAAA TCTACGAGTT
22261  CGAGTACGAG GGCTTGCTGC GTCGCAGCGC CATCGCCTCC TCGCCCGACC GCTGCGTCAC
22321  CCTCGAAAAG TCCACCCAGA CCGTGCAGGG GCCCGACTCG GCCGCCTGCG GTCTCTTCTG
22381  CTGCATGTTT CTGCACGCCT TTGTGCACTG GCCTCAGAGT CCCATGGACC GCAACCCCAC
22441  CATGAACTTG CTGACGGGGG TGCCCAACTC CATGCTCCAA AGCCCCCAGG TCGAGCCCAC
22501  CCTGCGCCGC AACCAGGAGC AGCTCTACAG CTTCCTGGAG CGCCACTCGC CCTACTTCCG
22561  CCGCCACAGC GCACAGATCA GGAGGGCCAC CTCCTTCTGC CACTTGCAAG AGATGCAAGA
22621  AGGGTAATAA CGATGTACAC ACTTTTTTCT CAATAAATGG CATTTTTTTT TTATTTATAC
22681  AAGCTCTCTG GGGTATTCAT TTCCCACCAC CACCACCCGC CGTTGTCGCC ATCTGGCTCT
22741  ATTTAGAAAT CGAAAGGGTT CTGCCGGGAG TCGCCGTGCG CCACGGGCAG GGACACGTTG
22801  CGATACTGGT AGCGGGTGCC CCACTTGAAC TCGGGCACCA CCAGGCGAGG CAGCTCGGGG
22861  AAGTTTTCGC TCCACAGGCT GCGGGTCAGC ACCAGCGCGT TCATCAGGTC GGGCGCCGAG
22921  ATCTTGAAGT CGCAGTTGGG GCCGCCGCCC TGCGCGCGCG AGTTGCGGTA CACCGGGTTG
22981  CAGCACTGGA ACACCAACAG CGCCGGGTGC TTCACGCTGG CCAGCACGCT GCGGTCGGAG
23041  ATCAGCTCGG CGTCCAGGTC CTCCGCGTTG CTCAGCGCGA ACGGGGTCAT CTTGGGCACT
23101  TGCCGCCCCA GGAAGGGCGC GTGCCCCGGT TTCGAGTTGC AGTCGCAGCG CAGCGGGATC
23161  AGCAGGTGCC CGTGCCCGGA CTCGGCGTTG GGGTACAGCG CGCGCATGAA GGCCTGCATC
23221  TGGCGGAAGG CCATCTGGGC CTTGGCGCCC TCCGAGAAGA ACATGCCGCA GGACTTGCCC
23281  GAGAACTGGT TTGCGGGGCA GCTGGCGTCG TGCAGGCAGC AGCGCGCGTC GGTGTTGGCG
23341  ATCTGCACCA CGTTGCGCCC CCACCGGTTC TTCACGATCT TGGCCTTGGA CGATTGCTCC
23401  TTCAGCGCGC GCTGCCCGTT CTCGCTGGTC ACATCCATCT CGATCACATG TTCCTTGTTC
23461  ACCATGCTGC TGCCGTGCAG ACACTTCAGC TCGCCCTCCG TCTCGGTGCA GCGGTGCTGC
23521  CACAGCGCGC AGCCCGTGGG CTCGAAAGAC TTGTAGGTCA CCTCCGCGAA GGACTGCAGG
23581  TACCCCTGCA AAAAGCGGCC CATCATGGTC ACGAAGGTCT TGTTGCTGCT GAAGGTCAGC
23641  TGCAGCCCGC GGTGCTCCTC GTTCAGCCAG GTCTTGCACA CGGCCGCCAG CGCCTCCACC
23701  TGGTCGGGCA GCATCTTGAA GTTCACCTTC AGCTCATTCT CCACGTGGTA CTTGTCCATC
23761  AGCGTGCGCG CCGCCTCCAT GCCCTTCTCC CAGGCCGACA CCAGCGGCAG GCTCACGGGG
23821  TTCTTCACCA TCACCGTGGC CGCCGCCTCC GCCGCGCTTT CGCTTTCCGC CCCGCTGTTC
23881  TCTTCCTCTT CCTCCTCTTC CTCGCCGCCG CCCACTCGCA GCCCCCGCAC CACGGGGTCG
23941  TCTTCCTGCA GGCGCTGCAC CTTGCGCTTG CCGTTGCGCC CCTGCTTGAT GCGCACGGGC
24001  GGGTTGCTGA AGCCCACCAT CACCAGCGCG GCCTCTTCTT GCTCGTCCTC GCTGTCCAGA
24061  ATGACCTCCG GGGAGGGGGG GTTGGTCATC CTCAGTACCG AGGCACGCTT CTTTTTCTTC
24121  CTGGGGGCGT TCGCCAGCTC CGCGGCTGCG GCCGCTGCCG AGGTCGAAGG CCGAGGGCTG
24181  GGCGTGCGCG GCACCAGCGC GTCTTGCGAG CCGTCCTCGT CCTCCTCGGA CTCGAGACGG
24241  AGGCGGGCCC GCTTCTTCGG GGGCGCGCGG GGCGGCGGAG GCGGCGGCGG CGACGGAGAC
24301  GGGGACGAGA CATCGTCCAG GGTGGGTGGA CGGCGGGCCG CGCCGCGTCC GCGCTCGGGG
```

FIG. 5G

```
24361  GTGGTTTCGC GCTGGTCCTC TTCCCGACTG GCCATCTCCC ACTGCTCCTT CTCCTATAGG
24421  CAGAAAGAGA TCATGGAGTC TCTCATGCGA GTCGAGAAGG AGGAGGACAG CCTAACCGCC
24481  CCCTCTGAGC CCTCCACCAC CGCCGCCACC ACCGCCAATG CCGCCGCGGA CGACGCGCCC
24541  ACCGAGACCA CCGCCAGTAC CACCCTCCCC AGCGACGCAC CCCCGCTCGA GAATGAAGTG
24601  CTGATCGAGC AGGACCCGGG TTTTGTGAGC GGAGAGGAGG ATGAGGTGGA TGAGAAGGAG
24661  AAGGAGGAGG TCGCCGCCTC AGTGCCAAAA GAGGATAAAA AGCAAGACCA GGACGACGCA
24721  GATAAGGATG AGACAGCAGT CGGGCGGGGG AACGGAAGCC ATGATGCTGA TGACGGCTAC
24781  CTAGACGTGG GAGACGACGT GCTGCTTAAG CACCTGCACC GCCAGTGCGT CATCGTCTGC
24841  GACGCGCTGC AGGAGCGCTG CGAAGTGCCC CTGGACGTGG CGGAGGTCAG CCGCGCCTAC
24901  GAGCGGCACC TCTTCGCGCC GCACGTGCCC CCCAAGCGCC GGGAGAACGG CACCTGCGAG
24961  CCCAACCCGC GTCTCAACTT CTACCCGGTC TTCGCGGTAC CCGAGGTGCT GGCCACCTAC
25021  CACATCTTCT TCCAAAACTG CAAGATCCCC CTCTCCTGCC GCGCTAACCG CACCCGCGCC
25081  GACAAAACCC TGACCCTGCG GCAGGGCGCC CACATACCTG ATATTGCCTC TCTGGAGGAA
25141  GTGCCCAAGA TCTTCGAGGG TCTCGGTCGC GACGAGAAAC GGGCGGCGAA CGCTCTGCAC
25201  GGAGACAGCG AAAACGAGAG TCACTCGGGG GTGCTGGTGG AGCTCGAGGG CGACAACGCG
25261  CGCCTGGCCG TACTCAAGCG CAGCATAGAG GTCACCCACT TTGCCTACCC GGCGCTCAAC
25321  CTGCCCCCCA AGGTCATGAG TGTGGTCATG GGCGAGCTCA TCATGCGCCG CGCTCAGCCC
25381  CTGGCCGCGG ATGCAAACTT GCAAGAGTCC TCCGAGGAAG GCCTGCCCGC GGTCAGCGAC
25441  GAGCAGCTAG CGCGCTGGCT GGAGACCCGC GACCCCGCGC AGCTGGAGGA GCGGCGCAAG
25501  CTCATGATGG CCGCGGTGCT GGTCACCGTG GAGCTCGAGT GTCTGCAGCG CTTCTTCGCG
25561  GACCCCGAGA TGCAGCGCAA GCTCGAGGAG ACCCTGCACT ACACCTTCCG CCAGGGCTAC
25621  GTGCGCCAGG CCTGCAAGAT CTCCAACGTG GAGCTCTGCA ACCTGGTCTC CTACCTGGGC
25681  ATCCTGCACG AGAACCGCCT CGGGCAGAAC GTCCTGCACT CCACCCTCAA AGGGGAGGCG
25741  CGCCGCGACT ACATCCGCGA CTGCGCCTAC CTCTTCCTCT GCTACACCTG GCAGACGGCC
25801  ATGGGGGTCT GGCAGCAGTG CCTGGAGGAG CGCAACCTCA AGGAGCTGGA AAAGCTACTC
25861  AAGCGCACCC TCAGGGACCT CTGGACGGGC TTCAACGAGC GCTCGGTGGC CGCCGCGCTG
25921  GCGGACATCA TCTTCCCCGA GCGCCTGCTC AAGACCCTGC AGCAGGGCCT GCCCGACTTC
25981  ACCAGCCAGA GCATGCTGCA GAACTTTAGG ACTTTCATCC TGGAGCGCTC GGGCATCCTG
26041  CCTGCCACTT GCTGCGCGCT GCCCAGCGAC TTCGTGCCCA TCAAGTACAG GGAGTGCCCG
26101  CCGCCGCTCT GGGGCCACTG CTACCTCTTC CAGCTGGCCA ACTACCTCGC CTACCACTCG
26161  GACCTCATGG AAGACGTGAG CGGCGAGGGC CTGCTCGAGT GCCACTGCCG CTGCAACCTC
26221  TGCACGCCCC ACCGCTCTCT AGTCTGCAAC CCGCAGCTGC TCAGCGAGAG TCAGATTATC
26281  GGTACCTTCG AGCTGCAGGG TCCCTCGCCT GACGAGAAGT CCGCGGCTCC GGGGCTGAAA
26341  CTCACTCCGG GGCTGTGGAC TTCCGCCTAC CTACGCAAAT TTGTACCTGA GGACTACCAC
26401  GCCCACGAGA TCAGGTTCTA CGAAGACCAA TCCCGCCCGC CCAAGGCGGA GCTCACCGCC
26461  TGCGTCATCA CCCAGGGGCA CATCCTGGGC CAATTGCAAG CCATCAACAA AGCCCGCCGA
26521  GAGTTCTTGC TGAAAAAGGG TCGGGGGGTG TACCTGGACC CCCAGTCCGG CGAGGAGCTA
26581  AACCCGCTAC CCCCGCCGCC GCCCCAGCAG CGGGACCTTG CTTCCCAGGA TGGCACCCAG
26641  AAAGAAGCAG CAGCCGCCGC CGCCGCAGCC ATACATGCTT CTGGAGGAAG AGGAGGAGGA
26701  CTGGGACAGT CAGGCAGAGG AGGTTTCGGA CGAGGAGCAG GAGGAGATGA TGGAAGACTG
26761  GGAGGAGGAC AGCAGCCTAG ACGAGGAAGC TTCAGAGGCC GAAGAGGTGG CAGACGCAAC
26821  ACCATCACCC TCGGTCGCAG CCCCCTCGCC GGGGCCCCTG AAATCCTCCG AACCCAGCAC
26881  CAGCGCTATA ACCTCCGCTC CTCCGGCGCC GGCGCCACCC GCCCGCAGAC CCAACCGTAG
26941  ATGGGACACC ACAGGAACCG GGGTCGGTAA GTCCAAGTGC CCGCCGCCGC CACCGCAGCA
27001  GCAGCAGCAG CGCCAGGGCT ACCGCTCGTG GCGCGGGCAC AAGAACGCCA TAGTCGCCTG
27061  CTTGCAAGAC TGCGGGGGCA ACATCTCTTT CGCCCGGCGC TTCCTGCTAT TCCACCACGG
27121  GGTCGCCTTT CCCCGCAATG TCCTGCATTA CTACCGTCAT CTCTACAGCC CCTACTGCAG
27181  CGGCGACCCA GAGGCGGCAG CGGCAGCCAC AGCGGCGACC ACCACCTAGG AAGATATCCT
27241  CCGCGGGCAA GACAGCGGCA GCAGCGGCCA GGAGACCCGC GGCAGCAGCG GCGGGAGCGG
27301  TGGGCGCACT GCGCCTCTCG CCCAACGAAC CCCTCTCGAC CCGGGAGCTC AGACACAGGA
27361  TCTTCCCCAC TTTGTATGCC ATCTTCCAAC AGAGCAGAGG CCAGGAGCAG GAGCTGAAAA
27421  TAAAAAACAG ATCTCTGCGC TCCCTCACCC GCAGCTGTCT GTATCACAAA AGCGAAGATC
27481  AGCTTCGGCG CACGCTGGAG GACGCGGAGG CACTCTTCAG CAAATACTGC GCGCTCACTC
27541  TTAAAGACTA GCTCCGCGCC CTTCTCGAAT TTAGGCGGGA GAAAACTACG TCATCGCCGG
27601  CCGCCGCCCA GCCCGCCCAG CCGAGATGAG CAAAGAGATT CCCACGCCAT ACATGTGGAG
27661  CTACCAGCCG CAGATGGGAC TCGCGGCGGG AGCGGCCCAG GACTACTCCA CCCGCATGAA
27721  CTACATGAGC GCGGGACCCC ACATGATCTC ACAGGTCAAC GGGATCCGCG CCCAGCGAAA
27781  CCAAATACTG CTGGAACAGG CGGCCATCAC CGCCACGCCC CGCCATAATC TCAACCCCCG
```

FIG. 5H

```
27841  AAATTGGCCC GCCGCCCTCG TGTACCAGGA AACCCCCTCC GCCACCACCG TACTACTTCC
27901  GCGTGACGCC CAGGCCGAAG TCCAGATGAC TAACTCAGGG GCGCAGCTCG CGGGCGGCTT
27961  TCGTCACGGG GCGCGGCCGC TCCGACCAGG TATAAGACAC CTGATGATCA GAGGCCGAGG
28021  TATCCAGCTC AACGACGAGT CGGTGAGCTC TTCGCTCGGT CTCCGTCCGG ACGGAACTTT
28081  CCAGCTCGCC GGATCCGGCC GCTCTTCGTT CACGCCCCGC CAGGCGTACC TGACTCTGCA
28141  GACCTCGTCC TCGGAGCCCC GCTCCGGAGG CATCGGAACC CTCCAGTTCG TGGAGGAGTT
28201  CGTGCCCTCG GTCTACTTCA ACCCCTTCTC GGGACCTCCC GGACGCTACC CCGACCAGTT
28261  CATTCCGAAC TTTGACGCGG TGAAGGACTC GGCGGACGGC TACGACTGAA TGTCAGGTGC
28321  CGAGGCAGAG CAGCTTCGCC TGAGACACCT CGAGCACTGC CGCCGCCACA AGTGCTTCGC
28381  CCGCGGTTCC GGTGAGTTCT GCTACTTTCA GCTACCCGAG GAGCATACCG AGGGGCCGGC
28441  GCACGGCGTC CGCCTGACCA CCCAGGGCGA GGTTACCTGT TCCCTCATCC GGGAGTTCAC
28501  CCTCCGTCCC CTGCTAGTGG AGCGGGAGCG GGGTCCCTGT GTCCTAACTA TCGCCTGCAA
28561  CTGCCCTAAC CCTGGATTAC ATCAAGATCT TTGCTGTCAT CTCTGTGCTG AGTTTAATAA
28621  ACGCTGAGAT CAGAATCTAC TGGGGCTCCT GTCGCCATCC TGTGAACGCC ACCGTCTTCA
28681  CCCACCCCGA CCAGGCCCAG GCGAACCTCA CCTGCGGTCT GCATCGGAGG GCCAAGAAGT
28741  ACCTCACCTG GTACTTCAAC GGCACCCCCT TTGTGGTTTA CAACAGCTTC GACGGGGACG
28801  GAGTCTCCCT GAAAGACCAG CTCTCCGGTC TCAGCTACTC CATCCACAAG AACACCACCC
28861  TCCAACTCTT CCCTCCCTAC CTGCCGGGAA CCTACGAGTG CGTCACCGGC CGCTGCACCC
28921  ACCTCACCCG CCTGATCGTA AACCAGAGCT TTCCGGGAAC AGATAACTCC CTCTTCCCCA
28981  GAACAGGAGG TGAGCTCAGG AAACTCCCCG GGGACCAGGG CGGAGACGTA CCTTCGACCC
29041  TTGTGGGGTT AGGATTTTTT ATTACCGGGT TGCTGGCTCT TTTAATCAAA GCTTCCTTGA
29101  GATTTGTTCT TTCCTTCTAC GTGTATGAAC ACCTCAGCCT CCAATAACTC TACCCTTTCT
29161  TCGGAATCAG GTGACTTCTC TGAAATCGGG CTTGGTGTGC TGCTTACTCT GTTGATTTTT
29221  TTCCTTATCA TACTCAGCCT TCTGTGCCTC AGGCTCGCCG CCTGCTGCGC ACACATCTAT
29281  ATCTACTGCT GGTTGCTCAA GTGCAGGGGT CGCCACCCAA GATGAACAGG TACATGGTCC
29341  TATCGATCCT AGGCCTGCTG GCCCTGGCGG CCTGCAGCGC CGCCAAAAAA GAGATTACCT
29401  TTGAGGAGCC CGCTTGCAAT GTAACTTTCA AGCCCGAGGG TGACCAATGC ACCACCCTCG
29461  TCAAATGCGT TACCAATCAT GAGAGGCTGC GCATCGACTA CAAAAACAAA ACTGGCCAGT
29521  TTGCGGTCTA TAGTGTGTTT ACGCCCGGAG ACCCCTCTAA CTACTCTGTC ACCGTCTTCC
29581  AGGGCGGACA GTCTAAGATA TTCAATTACA CTTTCCCTTT TTATGAGTTA TGCGATGCGG
29641  TCATGTACAT GTCAAAACAG TACAACCTGT GGCCTCCCTC TCCCCAGGCG TGTGTGGAAA
29701  ATACTGGGTC TTACTGCTGT ATGGCTTTGG CAATCACTAC GCTCGCTCTA ATCTGCACGG
29761  TGCTATACAT AAAATTCAGG CAGAGGCGAA TCTTTATCGA TGAAAAGAAA ATGCCTTGAT
29821  CGCTAACACC GGCTTTCTAT CTGCAGAATG AATGCAATCA CCTCCCTACT AATCACCACC
29881  ACCCTCCTTG CGATTGCCCA TGGGTTGACA CGAATCGAAG TGCCAGTGGG GTCCAATGTC
29941  ACCATGGTGG GCCCCGCCGG CAATTCCACC CTCATGTGGG AAAAATTTGT CCGCAATCAA
30001  TGGGTTCATT TCTGCTCTAA CCGAATCAGT ATCAAGCCCA GAGCCATCTG CGATGGGCAA
30061  AATCTAACTC TGATCAATGT GCAAATGATG GATGCTGGGT ACTATTACGG GCAGCGGGGA
30121  GAAATCATTA ATTACTGGCG ACCCCACAAG GACTACATGC TGCATGTAGT CGAGGCACTT
30181  CCCACTACCA CCCCCACTAC CACCTCTCCC ACCACCACTA CCACCACTAC TACTACTACT
30241  ACTACCACTA CCGCTGCCCG CCATACCCGC AAAAGCACCA TGATTAGCAC AAAGCCCCCT
30301  CGTGCTCACT CCCACGCCGG CGGGCCCATC GGTGCGACCT CAGAAACCAC CGAGCTTTGC
30361  TTCTGCCAAT GCACTAACGC CAGCGCTCAT GAACTGTTCG ACCTGGAGAA TGAGGATGCC
30421  CAGCAGAGCT CCGCTTGCCT GACCCAGGAG GCTGTGGAGC CCGTTGCCCT GAAGCAGATC
30481  GGTGATTCAA TAATTGACTC TTCTTCTTTT GCCACTCCCG AATACCCTCC CGATTCTACT
30541  TTCCACATCA CGGGTACCAA AGACCCTAAC CTCTCTTTCT ACCTGATGCT GCTGCTCTGT
30601  ATCTCTGTGG TCTCTTCCGC GCTGATGTTA CTGGGGATGT TCTGCTGCCT GATCTGCCGC
30661  AGAAAGAGAA AAGCTCGCTC TCAGGGCCAA CCACTGATGC CCTTCCCCTA CCCCCCGGAT
30721  TTTGCAGATA ACAAGATATG AGCTCGCTGC TGACACTAAC CGCTTTACTA GCCTGCGCTC
30781  TAACCCTTGT CGCTTGCGAC TCGAGATTCC ACAATGTCAC AGCTGTGGCA GGAGAAAATG
30841  TTACTTTCAA CTCCACGGCC GATACCCAGT GGTCGTGGAG TGGCTCAGGT AGCTACTTAA
30901  CTATCTGCAA TAGCTCCACT TCCCCCAGCA TATCCCCAAC CAAGTACCAA TGCAATGCCA
30961  GCCTGTTCAC CCTCATCAAC GCTTCCACCC TGGACAATGG ACTCTATGTA GGCTATGTAC
31021  CCTTTGGTGG GCAAGGAAAG ACCCACGCTT ACAACCTGGA AGTTCGCCAG CCCAGAACCA
31081  CTACCCAAGC TTCTCCCACC ACCACCACCA CCACCACCAC CACCATCACC AGCAGCAGCA
31141  GCAGCCACAG CAGCAGCAGC AGATTATTGA CTTTGGTTTT GGCCAGCTCA TCTGCCGCTA
31201  CCCAGGCCAT CTACAGCTCT GTGCCCGAAA CCACTCAGAT CCACCGCCCA GAAACGACCA
31261  CCGCCACCAC CCTACACACC TCCAGCGATC AGATGCCGAC CAACATCACC CCCTTGGCTC
```

FIG. 5I

```
31321  TTCAAATGGG ACTTACAAGC CCCACTCCAA AACCAGTGGA TGCGGCCGAG GTCTCCGCCC
31381  TCGTCAATGA CTGGGCGGGG CTGGGAATGT GGTGGTTCGC CATAGGCATG ATGGCGCTCT
31441  GCCTGCTTCT GCTCTGGCTC ATCTGCTGCC TCCACCGCAG GCGAGCCAGA CCCCCCATCT
31501  ATAGACCCAT CATTGTCCTG AACCCCGATA ATGATGGGAT CCATAGATTG GATGGCCTGA
31561  AAAACCTACT TTTTTCTTTT ACAGTATGAT AAATTGAGAC ATGCCTCGCA TTTTCTTGTA
31621  CATGTTCCTT CTCCCACCTT TTCTGGGGTG TTCTACGCTG GCCGCTGTGT CTCACCTGGA
31681  GGTAGACTGC CTCTCACCCT TCACTGTCTA CCTGCTTTAC GGATTGGTCA CCCTCACTCT
31741  CATCTGCAGC CTAATCACAG TAATCATCGC CTTCATCCAG TGCATTGATT ACATCTGTGT
31801  GCGCCTCGCA TACTTCAGAC ACCACCCGCA GTACCGAGAC AGGAACATTG CCCAACTTCT
31861  AAGACTGCTC TAATCATGCA TAAGACTGTG ATCTGCCTTC TGATCCTCTG CATCCTGCCC
31921  ACCCTCACCT CCTGCCAGTA CACCACAAAA TCTCCGCGCA AAAGACATGC CTCCTGCCGC
31981  TTCACCCAAC TGTGGAATAT ACCCAAATGC TACAACGAAA AGAGCGAGCT CTCCGAAGCT
32041  TGGCTGTATG GGGTCATCTG TGTCTTAGTT TTCTGCAGCA CTGTCTTTGC CCTCATGATC
32101  TACCCCTACT TTGATTTGGG ATGGAACGCG ATCGATGCCA TGAATTACCC CACCTTTCCC
32161  GCACCCGAGA TAATTCCACT GCGACAAGTT GTACCCGTTG TCGTTAATCA ACGCCCCCCA
32221  TCCCCTACGC CCACTGAAAT CAGCTACTTT AACCTAACAG GCGGAGATGA CTGACGCCCT
32281  AGATCTAGAA ATGGACGGCA TCAGTACCGA GCAGCGTCTC CTAGAGAGGC GCAGGCAGGC
32341  GGCTGAGCAA GAGCGCCTCA ATCAGGAGCT CCGAGATCTC GTTAACCTGC ACCAGTGCAA
32401  AAGAGGCATC TTTTGTCTGG TAAAGCAGGC CAAAGTCACC TACGAGAAGA CCGGCAACAG
32461  CCACCGCCTC AGTTACAAAT TGCCCACCCA GCGCCAGAAG CTGGTGCTCA TGGTGGGTGA
32521  GAATCCCATC ACCGTCACCC AGCACTCGGT AGAGACCGAG GGGTGTCTGC ACTCTCCCTG
32581  TCGGGGTCCA GAAGACCTCT GCACCCTGGT AAAGACCCTG TGCGGTCTCA GAGATTTAGT
32641  CCCCTTTAAC TAATCAAACA CTGGAATCAA TAAAAAGAAT CACTTACTTA AAATCAGACA
32701  GCAGGTCTCT GTCCAGTTTA TTCAGCAGCA CCTCCTTCCC CTCCTCCCAA CTCTGGTACT
32761  CCAAACGCCT TCTGGCGGCA AACTTCCTCC ACACCCTGAA GGGAATGTCA GATTCTTGCT
32821  CCTGTCCCTC CGCACCCACT ATCTTCATGT TGTTGCAGAT GAAGCGCACC AAAACGTCTG
32881  ACGAGAGCTT CAACCCCGTG TACCCCTATG ACACGGAAAG CGGCCCTCCC TCCGTCCCTT
32941  TCCTCACCCC TCCCTTCGTG TCTCCCGATG GATTCCAAGA AAGCCCCCCC GGGGTCCTGT
33001  CTCTGAACCT GGCCGAGCCC CTGGTCACTT CCCACGGCAT GCTCGCCCTG AAAATGGGAA
33061  GTGGCCTCTC CCTGGACGAC GCTGGCAACC TCACCTCTCA AGATATCACC ACCGCTAGCC
33121  CTCCCCTCAA AAAAACCAAG ACCAACCTCA GCCTAGAAAC CTCATCCCCC CTAACTGTAA
33181  GCACCTCAGG CGCCCTCACC GTAGCAGCCG CCGCTCCCCT GGCAGTGGCC GGCACCTCCC
33241  TCACCATGCA ATCAGAGGCC CCCCTGACAG TACAGGATGC AAAACTCACC CTGGCCACCA
33301  AAGGCCCCCT GACCGTGTCT GAAGGCAAAC TGGCCTTGCA AACATCGGCC CCGCTGACGG
33361  CCGCTGACAG CAGCACCCTC ACCGTTAGCG CCACACCACC AATTAATGTA AGCAGTGGAA
33421  GTTTAGGCTT AGACATGGAA GACCCTATGT ATACTCACGA TGGAAAACTG GGAATAAGAA
33481  TTGGGGGTCC ACTAAGAGTA GTAGACAGCT TGCACACACT CACTGTAGTT ACCGGAAATG
33541  GACTAACTGT AGATAACAAT GCCCTCCAAA CTAGAGTTAC GGGCGCCCTA GGTTATGACA
33601  CATCAGGAAA TCTACAATTG AGAGCTGCAG GAGGTATGCG AATTGATGCA AATGGCCAAC
33661  TTATCCTTAA TGTGGCATAC CCATTTGATG CTCAGAACAA TCTCAGCCTT AGACTTGGTC
33721  AGGGACCCCT GTATATAAAC ACAGACCACA ACCTGGATTT GAATTGCAAC AGAGGTCTAA
33781  CCACAACTAC CACCAACAAC ACAAAAAAAC TTGAGACTAA AATTAGCTCA GGCTTAGACT
33841  ATGACACCAA TGGTGCTGTC ATTATTAAAC TTGGCACTGG TCTAAGCTTC GACAACACAG
33901  GCGCCCTAAC TGTGGGAAAC ACTGGTGATG ATAAACTGAC TCTGTGGACG ACCCCAGACC
33961  CATCTCCAAA TTGCAGAATT CACTCAGACA AAGACTGCAA GTTTACTCTA GTCCTAACTA
34021  AGTGTGGAAG CCAAATCCTG GCCTCTGTCG CCGCCCTAGC GGTATCAGGA AATCTGGCTT
34081  CGATAACAGG CACCGTTGCC AGCGTTACCA TCTTTCTCAG ATTTGATCAG AATGGAGTGC
34141  TTATGGAAAA CTCCTCGCTA GACAGGCAGT ACTGGAACTT CAGAAATGGC AACTCAACTA
34201  ACGCTGCCCC CTACACCAAT GCAGTTGGGT TCATGCCAAA CCTCGCAGCA TACCCCAAAA
34261  CGCAAAGCCA GACTGCTAAA AACAACATTG TAAGTCAGGT TTACTTGAAT GGAGACAAAT
34321  CCAAACCCAT GACCCTTACC ATCACCCTCA ATGGAACTAA TGAATCCAGT GAAACTAGCC
34381  AGGTGAGTCA CTACTCCATG TCATTTACAT GGGCTTGGGA AAGTGGGCAA TATGCCACTG
34441  AAACCTTTGC CACCAACTCC TTCACCTTTT CTTACATTGC TGAACAATAA AAAGCATGAC
34501  ACTGATGTTC ATTTCTGATT CTTATTTTAT TATTTTCAAA CACAACAAAA TCATTCAAGT
34561  CATTCTTCCA TCTTAGCTTA ATAGACACAG TAGCTTAATA GACCCAGTAG TGCAAAGCCC
34621  CATTCTAGCT TATAGATCAG ACAGTGATAA TTAACCACCA CCACCACCAT ACCTTTTGAT
34681  TCAGGAAATC ATGATCATCA CAGGATCCTA GTCTTCAGGC CGCCCCCTCC CTCCCAAGAC
34741  ACAGAATACA CAGTCCTCTC CCCCCGACTG GCTTTAAATA ACACCATCTG GTTGGTCACA
```

FIG. 5J

```
34801  GACATGTTCT TAGGGGTTAT ATTCCACACG GTCTCCTGCC GCGCCAGGCG CTCGTCGGTG
34861  ATGTTGATAA ACTCTCCCGG CAGCTCGCTC AAGTTCACGT CGCTGTCCAG CGGCTGAACC
34921  TCCGGCTGAC GCGATAACTG TGCGACCGGC TGCTGGACAA ACGGAGGCCG CGCCTACAAG
34981  GGGGTAGAGT CATAATCCTC GGTCAGGATA GGGCGGTGAT GCAGCAGCAG CGAGCGAAAC
35041  ATCTGCTGCC GCCGCCGCTC CGTCCGGCAG GAAAACAACA AGCCGGTGGT CTCCTCCGCG
35101  ATAATCCGCA CCGCCCGCAG CATCAGCTTC CTCGTTCTCC GCGCGCAGCA CCTCACCCTG
35161  ATCTCGCTCA AGTCGGCGCA GTAGGTACAG CACAGCACCA CGATGTTATT CATGATCCCA
35221  CAGTGCAGGG CGCTGTATCC AAAGCTCATG CCGGGAACCA CCGCCCCCAC GTGGCCATCG
35281  TACCACAAGC GCACGTAAAT TAAGTGTCGA CCCCTCATGA ACGTGCTGGA CACAAACATT
35341  ACTTCCTTGG GCATGTTGTA ATTCACCACC TCCCGGTACC AGATAAACCT CTGGTTAAAC
35401  AGGGCACCTT CCACCACCAT CCTGAACCAA GAGGCCAGAA CCTGCCCACC GGCTATGCAC
35461  TGCAGGGAAC CCGGGTTGGA ACAATGACAA TGCAGACTCC AAGGCTCGTA ACCGTGGATC
35521  ATCCGGCTGC TGAAGGCATC GATGTTGGCA CAACACAGAC ACACGTGCAT GCACTTTCTC
35581  ATGATTAGCA GCTCTTCCCT CGTCAGGATC ATATCCCAAG GAATAACCCA TTCTTGAATC
35641  AACGTAAAAC CCACACAGCA GGGAAGGCCT CGCACATAAC TCACGTTGTG CATGGTCAGC
35701  GTGTTGCATT CTGGAAACAG CGGATGATCC TCCAGTATCG AGGCGCGGGT CTCCTTCTCA
35761  CAGGGAGGTA AAGGGTCCCT GCTGTACGGA CTGCGCCGGG ACGACCGAGA TCGTGTTGAG
35821  CGTAGTGTCA TGGAAAAGGG AACGCCGGAC GTGGTCATAC TTCTTGAAGC AGAACCAGGT
35881  TCGCGCGTGG CAGGCCTCCT TGCGTCTGCG GTCTCGCCGT CTAGCTCGCT CCGTGTGATA
35941  GTTGTAGTAC AGCCACTCCC GCAGAGCGTC GAGGCGCACC CTGGCTTCCG GATCTATGTA
36001  GACTCCGTCT TGCACCGCGG CCCTGATAAT ATCCACCACC GTAGAATAAG CAACACCCAG
36061  CCAAGCAATA CACTCGCTCT GCGAGCGGCA GACAGGAGGA GCGGGCAGAG ATGGGAGAAC
36121  CATGATAAAA AACTTTTTTT AAAGAATATT TTCCAATTCT TCGAAAGTAA GATCTATCAA
36181  GTGGCAGCGC TCCCCTCCAC TGGCGCGGTC AAACTCTACG GCCAAAGCAC AGACAACGGC
36241  ATTTCTAAGA TGTTCCTTAA TGGCGTCCAA AAGACACACC GCTCTCAAGT TGCAGTAAAC
36301  TATGAATGAA AACCCATCCG GCTGATTTTC CAATATAGAC GCGCCGGCGG CGTCCACCAA
36361  ACCCAGATAA TTTTCTTCTC TCCAGCGGTT TAGAATCTGT CTAAGCAAAT CCCTTATATC
36421  AAGTCCGGCC ATGCCAAAAA TCTGCTCAAG AGCGCCCTCC ACCTTCATGA CCAAGCAGCG
36481  CATCATGATT GCAAAAATTC AGGTTCTTCA GAGACCTGTA TAAGATTCAA AATGGGAACA
36541  TTAACAAAAA TTCCTCTGTC GCGCAGATCC CTTCGCAGGG CAAGCTGAAC ATAATCAGAC
36601  AGGTCTGAAC GGACCAGTGA GGCCAAATCC CCACCAGGAA CCAGATCCAG AGACCCTATA
36661  CTGATTATGA CGCGCATACT CGGGGCTATG CTGACCAGCG TAGCGCCGAT GTAGGCGTGC
36721  TGCATGGGCG GCGAGATAAA ATGCAAAGTG CTGGTTAAAA AATCAGGCAA AGCCTCGCGC
36781  AAAAAAGCTA ACACATCATA ATCATGCTCA TGCAGGTAGT TGCAGGTAAG CTCAGGAACC
36841  AAAACGGAAT AACACACGAT TTTCCTCTCA AACATGACTT CGCGGATACT GCGTAAAACA
36901  AAAATTATAA ATAAAAAATT AATTAACTTA AACATTGGAA GCCTGTCTCA CAACAGGAAA
36961  AACCACTTTA ATCAACATAA GACGGGCCAC GGGCATGCCG GCATAGCCGT AAAAAAATTG
37021  GTCCCCGTGA TTAACAAGTA CCACAGACAG CTCCCCGGTC ATGTCGGGGG TCATCATGTG
37081  AGACTCTGTA TACACGTCTG GATTGTGAAC ATCAGACAAA CAAAGAAATC GAGCCACGTA
37141  GCCCGGAGGT ATAATCACCC GCAGGCGGAG GTACAGCAAA ACGACCCCCA TAGGAGGAAT
37201  CACAAAATTA GTAGGAGAAA AAAATACATA AACACCAGAA AAACCCTGTT GCTGAGGCAA
37261  AATAGCGCCC TCCCGATCCA AAACAACATA AAGCGCTTCC ACAGGAGCAG CCATAACAAA
37321  GACCCGAGTC TTACCAGTAA AAGAAAAAAG ATCTCTCAAC GCAGCACCAG CACCAACACT
37381  TCGCAGTGTA AAAGGCCAAG TGCCGAGAGA GTATATATAG GAATAAAAAG TGACGTAAAC
37441  GGGCAAAGTC CAAAAAACGC CCAGAAAAAC CGCACGCGAA CCTACGCCCC GAAACGAAAG
37501  CCAAAAAACA CTAGACACTC CCTTCCGGCG TCAACTTCCG CTTTCCCACG CTACGTCACT
37561  TGCCCCAGTC AAACAAACTA CATATCCCGA ACTTCCAAGT CGCCACGCCC AAAACACCGC
37621  CTACACCTCC CCGCCCGCCG GCCCGCCCCC AAACCCGCCT CCCGCCCCGC GCCCCGCCTC
37681  GCGCCGCCCA TCTCATTATC ATATTGGCTT CAATCCAAAA TAAGGTATAT TATTGATGAT
37741  G  (SEQ ID NO: 1)
```

FIG. 5K

```
   1  CATCATCAAT AATATACCTC AAACTTTTGG TGCGCGTTAA TATGCAAATG AGCCGTTTGA
  61  ATTTGGGGAT GCGGGGCGCT GATTGGCTGC GGGAGCGGCG ACCGTTAGGG GCGGGGCGGG
 121  TGACGTTTTG ATGACGTGTT TGTGAGGCGG AGCCGGTTTG CAAGTTCTCG TGGGAAAAGT
 181  GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCTCTGACA
 241  GGAAATGAGG TGTTTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTCGC GCGAAAACTG
 301  AATGAGGAAG TGAAAATCTG AGTAATTTCG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
 361  GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTTCGATT ACCGTATTTT TCACCTAAAT
 421  TTCCGCGTAC GGTGTCAAAG TCCGGTGTTT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
 481  TTTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTTCT
 541  CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
 601  GTAATGTTTT CCTGGCTACT GGGAACGAGA TTCTGGAACT GGTGGTGGAC GCCATGATGG
 661  GTGACGACCC TCCTGAGCCC CCTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTTGTATG
 721  ATCTGGAGGT GGATGTGCCC GAGAACGACC CCAACGAGGA GGCGGTGAAT GATTTGTTTA
 781  GCGATGCCGC GCTGCTGGCC GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
 841  CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCCGAG CTTAAAGGGG
 901  AAGAGCTCGA CCTGCGCTGC TATGAGGAAT GCTTGCCTCC GAGCGATGAT GAGGAGGACG
 961  AGGAGGCGAT TCGAGCTGCA GCGAGCGAGG GAGTGAAAGC TGCGGGCGAG AGCTTTAGCC
1021  TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTCAT CGCATGAATA
1081  CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTGTGTTT
1141  ACAGTAAGTG TGATTAACTT TAGCTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201  ATTTATGTAT ATGTTTTTTA TGTGTAGGTC CCGTCTCTGA CGTAGATGAG ACCCCCACTT
1261  CAGAGTGCAT TTCATCACCC CCAGAAATTG GCGAGGAACC GCCCGAAGAT ATTATTCATA
1321  GACCAGTTGC AGTGAGAGTC ACCGGGCGGA GAGCAGCTGT GGAGAGTTTG GATGACTTGC
1381  TACAGGGTGG GGATGAACCT TTGGACTTGT GTACCCGGAA ACGCCCCAGG CACTAAGTGC
1441  CACACATGTG TGTTTACTTA AGGTGATGTC AGTATTTATA GGGTGTGGAG TGCAATAAAA
1501  TCCGTGTTGA CTTTAAGTGT GTGGTTTATG ACTCAGGGGT GGGGACTGTG GGTATATAAG
1561  CAGGTGCAGA CCTGTGTGGT CAGTTCAGAG CAGGACTCAT GGAGATCTGG ACGGTCTTGG
1621  AAGACTTTCA CCAGACTAGA CAGCTGCTAG AGAACTCATC GGAGGAAGTC TCTTACCTGT
1681  GGAGATTTTG CTTCGGTGGG GCTCTAGCTA AGCTAGTCTA TAGGGCCAAA CAGGATTATA
1741  AGGATCAATT TGAGGATATT TTGAGAGAGT GTCCTAGTAT TTTTGACTCT CTCAACTTGG
1801  GCCATCAGTC TCACTTTAAC CAGAGTATTC TGAGAGCCCT TGACTTTTCT ACTCCTGGCA
1861  GAACTACCGC CGCGGTAGCC TTTTTTGCCT TTATTCTTGA CAAATGGAGT CAAGAAACCC
1921  ATTTCAGCAG GGATTACCGT CTGGACTGCT TAGCAGTAGC TTTGTGGAGA ACATGGAGGT
1981  GCCAGCGCCT GAATGCAATC TCCGGCTACT TGCCAGTACA GCCGGTAGAC ACGCTGAGGA
2041  TCCTGAGTCT CCAGTCACCC CAGGAACACC AACGCCGCCA GCAGCCGCAG CAGGAGCAGC
2101  AGCAAGAGGA GGAGGAGGAG GAGGACCGAG AAGAGAACCC GAGAGCCGGT CTGGACCCTC
2161  CGGTGGCGGA GGAGGAGGAG TAGCTGACTT GTTTCCCGAG CTGCGCCGGG TGCTGACTAG
2221  GTCTTCCAGT GGACGGGAGA GGGGGATTAA GCGGGAGAGG CATGAGGAGA CTAGTCACAG
2281  AACTGAACTG ACTGTCAGTC TGATGAGCCG CAGGCGCCCA GAATCGGTGT GGTGGCATGA
2341  GGTTCAGTCG CAGGGGATAG ATGAGGTCTC GGTAATGCAT GAGAAATATT CCCTAGAACA
2401  AGTCAAGACT TGTTGGTTGG AGCCCGAGGA TGATTGGGAG GTAGCCATCA GGAATTATGC
2461  CAAGCTGGCT CTGAGGCCAG ACAAGAAGTA CAAGATTACC AAACTGATTA ATATCAGAAA
2521  TTCCTGCTAC ATTTCGGGGA ATGGGGCCGA GGTGGAGATC AGTACCCAGG AGAGGGTGGC
2581  CTTCAGATGT TGTATGATGA ATATGTACCC GGGGGTGGTG GGCATGGAGG GAGTCACCTT
2641  TATGAACGCG AGGTTTAGGG GTGATGGGTA TAATGGGGTG GTCTTTATGG CCAACACCAA
2701  GCTGACAGTG CACGGATGCT CCTTCTTTGG CTTCAATAAC ATGTGCATCG AGGCCTGGGG
2761  CAGTGTTTCA GTGAGGGGAT GCAGCTTTTC AGCCAACTGG ATGGGGGTCG TGGGCAGAAC
2821  CAAGAGCGTG GTGTCAGTGA AGAAATGCCT GTTCGAGAGG TGCCACCTGG GGGTGATGAG
2881  CGAGGGCGAA GCCAAAGTCA AACACTGCGC CTCTACCGAG ACGGGCTGCT TTGTGATGAT
2941  CAAGGGCAAT GCCAAAGTCA AGCATAACAT GATTTGTGGG GCCTCGGATG AGCGCGGCTA
3001  CCAGATGCTG ACCTGTGCCG GTGGGAACAG CCATATGCTG GCCACCGTGC ATGTGGCCTC
3061  GCACCCCCGC AAGACATGGC CCGAGTTCGA GCACAACGTC ATGACCCGCT GCAATGTGCA
3121  CCTGGGGTCC CGCCGAGGCA TGTTCATGCC CTACCAGTGC AACATGCAAT TTGTGAAGGT
3181  GCTGCTGGAG CCCGATGCCA TGTCCAGAGT GAGCCTGGTG GGGGTGTTTG ACATGAATGT
3241  GGAGGTGTGG AAAATTCTGA GATATGATGA ATCCAAGACC AGGTGCCGGG CCTGCGAATG
3301  CGGAGGCAAG CACGCCAGGC TTCAGCCCGT GTGTGTGGAG GTGACGGAGG ACCTGCGACC
3361  CGATCATTTG GTGTTGTCCT GCAACGGGAC GGAGTTCGGC TCCAGCGGGG AAGAATCTGA
```

FIG. 6A

```
3421  CTAGAGTGAG TAGTGTTTGG GGGTGGGTGG GAGTCTGCAT GATGGGCAGA ATGACTAAAA
3481  TCTGTGTTTT TCTGCGCAGC AGCATGAGCG GAAGCGCCTC CTTTGAGGGA GGGGTATTCA
3541  GCCCTTATCT GACGGGGCGT CTCCCCTCCT GGGCGGGAGT GCGTCAGAAT GTGATGGGAT
3601  CCACGGTGGA CGGCCGGCCC GTGCAGCCCG CGAACTCTTC AACCCTGACC TACGCGACCC
3661  TGAGCTCCTC GTCCGTGGAC GCAGCTGCCG CCGCAGCTGC TGCTTCCGCC GCCAGCGCCG
3721  TGCGCGGAAT GGCCTTGGGC GCCGGCTACT ACAGCTCTCT GGTGGCCAAC TCGAGTTCCA
3781  CCAATAATCC CGCCAGCCTG AACGAGGAGA AGCTGCTGCT GCTGATGGCC CAGCTCGAGG
3841  CCCTGACCCA GCGCCTGGGC GAGCTGACCC AGCAGGTGGC TCAGCTGCAG GCGGAGACGC
3901  GGGCCGCGGT TGCCACGGTG AAAACCAAAT AAAAAATGAA TCAATAAATA AACGGAGACG
3961  GTTGTTGATT TTAACACAGA GTCTTGATCT TTATTTGATT TTTCGCGCGC GGTAGGCCCT
4021  GGACCACCGG TCTCGATCAT TGAGCACCCG GTGGATTTTT TCCAGGACCC GGTAGAGGTG
4081  GGCTTGGATG TTGAGGTACA TGGGCATGAG CCCGTCCCGG GGGTGGAGGT AGCTCCATTG
4141  CAGGGCCTCG TGCTCGGGGG TGGTGTTGTA AATCACCCAG TCATAGCAGG GGCGCAGGGC
4201  GTGGTGCTGC ACGATGTCCT TGAGGAGGAG ACTGATGGCC ACGGGCAGCC CCTTGGTGTA
4261  GGTGTTGACG AACCTGTTGA GCTGGGAGGG ATGCATGCGG GGGGAGATGA GATGCATCTT
4321  GGCCTGGATC TTGAGATTGG CGATGTTCCC GCCCAGATCC CGCCGGGGGT TCATGTTGTG
4381  CAGGACCACC AGCACGGTGT ATCCGGTGCA CTTGGGGAAT TTGTCATGCA ACTTGGAAGG
4441  GAAGGCGTGA AAGAATTTGG AGACGCCCTT GTGACCGCCC AGGTTTTCCA TGCACTCATC
4501  CATGATGATG GCGATGGGCC CGTGGGCGGC GGCCTGGGCA AAGACGTTTC GGGGGTCGGA
4561  CACATCGTAG TTGTGGTCCT GGGTGAGCTC GTCATAGGCC ATTTTAATGA ATTTGGGGCG
4621  GAGAGTGCCC GACTGGGGGA CGAAGGTGCC CTCGATCCCG GGGGCGTAGT TCCCCTCGCA
4681  GATCTGCATC TCCCAGGCCT TGAGCTCGGA GGGGGGGATC ATGTCCACCT GCGGGGCGAT
4741  GAAAAAAACG GTTTCCGGGG CGGGGGAGAT GAGCTGGGCC GAAAGCAGGT TCCGGAGCAG
4801  CTGGGACTTG CCGCAGCCGG TGGGACCGTA GATGACCCCG ATGACCGGCT GCAGGTGGTA
4861  GTTGAGGGAG AGACAGCTGC CATCCTCGCG GAGGAGGGGG GCCACCTCGT TCATCATCTC
4921  GCGCACATGC ATGTTCTCGC GCACGAGTTC CGCCAGGAGG CGCTCGCCCC CCAGCGAGAG
4981  GAGCTCTTGC AGCGAGGCGA AGTTTTTCAG CGGCTTGAGC CCGTCGGCCA TGGGCATTTT
5041  GGAGAGGGTC TGTTGCAAGA GTTCCAGACG GTCCCAGAGC TCGGTGATGT GCTCTAGGGC
5101  ATCTCGATCC AGCAGACCTC CTCGTTTCGC GGGTTGGGGC GACTGCGGGA GTAGGGCACC
5161  AGGCGATGGG CGTCCAGCGA GGCCAGGGTC CGGTCCTTCC AGGGTCGCAG GGTCCGCGTC
5221  AGCGTGGTCT CCGTCACGGT GAAGGGGTGC GCGCCGGGCT GGGCGCTTGC GAGGGTGCGC
5281  TTCAGGCTCA TCCGGCTGGT CGAGAACCGC TCCCGGTCGG CGCCCTGCGC GTCGGCCAGG
5341  TAGCAATTGA GCATGAGTTC GTAGTTGAGC GCCTCGGCCG CGTGGCCCTT GGCGCGGAGC
5401  TTACCTTTGG AAGTGTGTCC GCAGACGGGA CAGAGGAGGG ACTTGAGGGC GTAGAGCTTG
5461  GGGGCGAGGA AGACGGACTC GGGGGCGTAG GCGTCCGCGC CGCAGCTGGC GCAGACGGTC
5521  TCGCACTCCA CGAGCCAGGT GAGGTCGGGG CGGTCGGGGT CAAAAACGAG GTTTCCTCCG
5581  TGCTTTTTGA TGCGTTTCTT ACCTCTGGTC TCCATGAGCT CGTGTCCCCG CTGGGTGACA
5641  AAGAGGCTGT CCGTGTCCCC GTAGACCGAC TTTATGGGCC GGTCCTCGAG CGGGGTGCCG
5701  CGGTCCTCGT CGTAGAGGAA CCCCGCCCAC TCCGAGACGA AGGCCCGGGT CCAGGCCAGC
5761  ACGAAGGAGG CCACGTGGGA GGGGTAGCGG TCGTTGTCCA CCAGCGGGTC CACCTTCTCC
5821  AGGGTATGCA AGCACATGTC CCCCTCGTCC ACATCCAGGA AGGTGATTGG CTTGTAAGTG
5881  TAGGCCACGT GACCGGGGGT CCCGGCCGGG GGGGTATAAA AGGGGGCGGG CCCCTGCTCG
5941  TCCTCACTGT CTTCCGGATC GCTGTCCAGG AGCGCCAGCT GTTGGGGTAG GTATTCCCTC
6001  TCGAAGGCGG GCATGACCTC GGCACTCAGG TTGTCAGTTT CTAGAAACGA GGAGGATTTG
6061  ATATTGACGG TGCCGTTGGA GACGCCTTTC ATGAGCCCCT CGTCCATCTG GTCAGAAAAG
6121  ACGATCTTTT TGTTGTCGAG CTTGGTGGCG AAGGAGCCGT AGAGGGCGTT GGAGAGGAGC
6181  TTGGCGATGG AGCGCATGGT CTGGTTCTTT TCCTTGTCGG CGCGCTCCTT GGCGGCGATG
6241  TTGAGCTGCA CGTACTCGCG CGCCACGCAC TTCCATTCGG GGAAGACGGT GGTGAGCTCG
6301  TCGGGCACGA TTCTGACCCG CCAGCCGCGG TTGTGCAGGG TGATGAGGTC CACGCTGGTG
6361  GCCACCTCGC CGCGCAGGGG CTCGTTGGTC CAGCAGAGGC GCCCGCCCTT GCGCGAGCAG
6421  AAGGGGGGCA GCGGGTCCAG CATGAGCTCG TCTGGGGGGT CGGCGTCCAC GGTGAAGATG
6481  CCGGGCAGGA GCTCGGGGTC GAAGTAGCTG ATGGAAGTGG CCAGATCGTC CAGGGAAGCT
6541  TGCCAGTCGC GCACGGCCAG CGCGCGCTCG TAGGGGCTGA GGGGCGTGCC CCAGGGCATG
6601  GGGTGCGTGA GCGCGGAGGC GTACATGCCG CAGATGTCGT AGACGTAGAG GGGCTCCTCG
6661  AGGATGCCGA TGTAGGTGGG GTAGCAGCGC CCCCCGCGGA TGCTGGCGCG CACGTAGTCG
6721  TACAGCTCGT GCGAGGGCGC GAGGAGCCCC GTGCCGAGAT TGGAGCGCTG CGGCTTTTCG
6781  GCGCGGTAGA CGATCTGGCG GAAGATGGCG TGGGAGTTGG AGGAGATGGT GGGCCTCTGG
```

FIG. 6B

```
 6841  AAGATGTTGA AGTGGGCGTG GGGCAGGCCG ACCGAGTCCC TGATGAAGTG GGCGTAGGAG
 6901  TCCTGCAGCT TGGCGACGAG CTCGGCGGTG ACGAGGACGT CCAGGGCGCA GTAGTCGAGG
 6961  GTCTCTTGGA TGATGTCATA CTTGAGCTGG CCCTTCTGCT TCCACAGCTC GCGGTTGAGA
 7021  AGGAACTCTT CGCGGTCCTT CCAGTACTCT TCGAGGGGGA ACCCGTCCTG ATCGGCACGG
 7081  TAAGAGCCCA CCATGTAGAA CTGGTTGACG GCCTTGTAGG CGCAGCAGCC CTTCTCCACG
 7141  GGGAGGGCGT AAGCTTGCGC GGCCTTGCGC AGGGAGGTGT GGGTGAGGGC GAAGGTGTCG
 7201  CGCACCATGA CCTTGAGGAA CTGGTGCTTG AAGTCGAGGT CGTCGCAGCC GCCCTGCTCC
 7261  CAGAGTTGGA AGTCCGTGCG CTTCTTGTAG GCGGGGTTGG GCAAAGCGAA AGTAACATCG
 7321  TTGAAGAGGA TCTTGCCCGC GCGGGGCATG AAGTTGCGAG TGATGCGGAA AGGCTGGGGC
 7381  ACCTCGGCCC GGTTGTTGAT GACCTGGGCG GCGAGGACGA TCTCGTCGAA GCCGTTGATG
 7441  TTGTGCCCGA CGATGTAGAG TTCCACGAAT CGCGGGCGGC CCTTGACGTG GGGCAGCTTC
 7501  TTGAGCTCGT CGTAGGTGAG CTCGGCGGGG TCGCTGAGCC CGTGCTGTTC GAGGGCCCAG
 7561  TCGGCGACGT GGGGGTTGGC GCTGAGGAAG GAAGTCCAGA GATCCACGGC CAGGGCGGTC
 7621  TGCAAGCGGT CCCGGTACTG ACGGAACTGC TGGCCCACGG CCATTTTTTC GGGGGTGACG
 7681  CAGTAGAAGG TGCGGGGGTC GCCGTGCCAG CGGTCCCACT TGAGCTGGAG GGCGAGGTCG
 7741  TGGGCGAGCT CGACGAGCGG TGGGTCCCCG GAGAGTTTCA TGACCAGCAT GAAGGGGACG
 7801  AGCTGCTTGC CGAAGGACCC CATCCAGGTG TAGGTTTCCA CATCGTAGGT GAGGAAGAGC
 7861  CTTTCGGTGC GAGGATGCGA GCCGATGGGG AAGAACTGGA TCTCCTGCCA CCAGTTGGAG
 7921  GAATGGCTGT TGATGTGATG GAAGTAGAAA TGCCGACGGC GCGCCGAGCA CTCGTGCTTG
 7981  TGTTTATACA AGCGTCCGCA GTGCTCGCAA CGCTGCACGG GATGCACGTG CTGCACGAGC
 8041  TGTACCTGAG TTCCTTTGAC GAGGAATTTC AGTGGGCAGT GGAGCGCTGG CGGCTGCATC
 8101  TGGTGCTGTA CTACGTCCTG GCCATCGGCG TGGCCATCGT CTGCCTCGAT GGTGGTCATG
 8161  CTGACGAGGC CGCGCGGGAG GCAGGTCCAG ACCTCGGCTC GGACGGGTCG GAGAGCGAGG
 8221  ACGAGGGCGC GCAGGCCGGA GCTGTCCAGG GTCCTGAGAC GCTGCGGAGT CAGGTCAGTG
 8281  GGCAGCGGCG GCGCGCGGTT GACTTGCAGG AGCTTTTCCA GGGCGCGCGG GAGGTCCAGA
 8341  TGGTACTTGA TCTCCACGGC GCCGTTGGTG GCGACGTCCA CGGCTTGCAG GGTCCCGTGC
 8401  CCCTGGGGCG CCACCACCGT GCCCCGTTTC TTCTTGGGCG CTGGTTCCAT GCCGGTCAGA
 8461  AGCGGCGGCG AGGACGCGCG CCGGGCGGCA GGGGCGGCTC GGGGCCCGGA GGCAGGGGCG
 8521  GCAGGGGCAC GTCGGCGCCG CGCGCGGGCA GGTTCTGGTA CTGCGCCCGG AGAAGACTGG
 8581  CGTGAGCGAC GACGCGACGG TTGACGTCCT GGATCTGACG CCTCTGGGTG AAGGCCACGG
 8641  GACCCGTGAG TTTGAACCTG AAAGAGAGTT CGACAGAATC AATCTCGGTA TCGTTGACGG
 8701  CGGCCTGCCG CAGGATCTCT TGCACGTCGC CCGAGTTGTC CTGGTAGGCG ATCTCGGTCA
 8761  TGAACTGCTC GATCTCCTCC TCCTGAAGGT CTCCGCGGCC GGCGCGCTCG ACGGTGGCCG
 8821  CGAGGTCGTT GGAGATGCGG CCCATGAGCT GCGAGAAGGC GTTCATGCCG GCCTCGTTCC
 8881  AGACGCGGCT GTAGACCACG GCTCCGTCGG GGTCGCGCGC GCGCATGACC ACCTGGGCGA
 8941  GGTTGAGCTC GACGTGGCGC GTGAAGACCG CGTAGTTGCA GAGGCGCTGG TAGAGGTAGT
 9001  TGAGCGTGGT GGCGATGTGC TCGGTGACGA AGAAGTACAT GATCCAGCGG CGGAGCGGCA
 9061  TCTCGCTGAC GTCGCCCAGG GCTTCCAAGC GCTCCATGGC CTCGTAGAAG TCCACGGCGA
 9121  AGTTGAAAAA CTGGGAGTTG CGCGCCGAGA CGGTCAACTC CTCCTCCAGA AGACGGATGA
 9181  GCTCGGCGAT GGTGGCGCGC ACCTCGCGCT CGAAGGCCCC GGGGGGCTCC TCTTCTTCCA
 9241  TCTCCTCCTC CTCTTCCTCC TCCACTAACA TCTCTTCTAC TTCCTCCTCA GGAGGCGGTG
 9301  GCGGGGGAGG GGCCCTGCGT CGCCGGCGGC GCACGGGCAG ACGGTCGATG AAGCGCTCGA
 9361  TGGTCTCCCC GCGCCGGCGA CGCATGGTCT CGGTGACGGC GCGCCCGTCC TCGCGGGGCC
 9421  GCAGCGTGAA GACGCCGCCG CGCATCTCCA GGTGGCCGCC GGGGGGGTCT CCGTTGGGCA
 9481  GGGAGAGGGC GCTGACGATG CATCTTATCA ATTGGCCCGT AGGGACTCCG CGCAAGGACC
 9541  TGAGCGTCTC GAGATCCACG GGATCCGAAA ACCGCTGAAC GAAGGCTTCG AGCCAGTCGC
 9601  AGTCGCAAGG TAGGCTGAGC CCGGTTTCTT CGGGTATTTG GTCGGGAGGC GGGCGGGCGA
 9661  TGCTGCTGGT GATGAAGTTG AAGTAGGCGG TCCTGAGACG GCGGATGGTG GCGAGGAGCA
 9721  CCAGGTCCTT GGGCCCGGCT TGCTGGATGC GCAGACGGTC GGCCATGCCC CAGGCGTGGT
 9781  CCTGACACCT GGCGAGGTCC TTGTAGTAGT CCTGCATGAG CCGCTCCACG GGCACCTCCT
 9841  CCTCGCCCGC GCGGCCGTGC ATGCGCGTGA GCCCGAACCC GCGCTGCGGC TGGACGAGCG
 9901  CCAGGTCGGC GACGACGCGC TCGGCGAGGA TGGCCTGCTG GATCTGGGTG AGGGTGGTCT
 9961  GGAAGTCGTC GAAGTCGACG AAGCGGTGGT AGGCTCCGGT GTTGATGGTG TATGAGCAGT
10021  TGGCCATGAC GGACCAGTTG ACGGTCTGGT GGCCGGGGCG CACGAGCTCG TGGTACTTGA
10081  GGCGCGAGTA GGCGCGCGTG TCGAAGATGT AGTCGTTGCA GGTGCGCACG AGGTACTGGT
10141  ATCCGACGAG GAAGTGCGGC GGCGGCTGGC GGTAGAGCGG CCATCGCTCG GTGGCGGGGG
10201  CGCCGGGCGC GAGGTCCTCG AGCATGAGGC GGTGGTAGCC GTAGATGTAC CTGGACATCC
```

FIG. 6C

```
10261  AGGTGATGCC GGCGGCGGTG GTGGAGGCGC GCGGGAACTC GCGGACGCGG TTCCAGATGT
10321  TGCGCAGCGG CAGGAAGTAG TTCATGGTGG CCGCGGTCTG GCCCGTGAGG CGCGCGCAGT
10381  CGTGGATGCT CTATACGGGC AAAAACGAAA GCGGTCAGCG GCTCGACTCC GTGGCCTGGA
10441  GGCTAAGCGA ACGGGTTGGG CTGCGCGTGT ACCCCGGTTC GAATCTCGAA TCAGGCTGGA
10501  GCCGCAGCTA ACGTGGTACT GGCACTCCCG TCTCGACCCA AGCCTGCACA AAACCTCCAG
10561  GATACGGAGG CGGGTCGTTT TGCAACTTTT TGAGGCCGGA AATGAAACTA GTAAGCGCGA
10621  AAAGCGGCCG ACCGCGATGG CTCGCTGCCG TAGTCTGGAG AAGAATCGCC AGGGTTGCGT
10681  TGCGGTGTGC CCCGGTTCGA GGCCGGCCGG ATTCCGCGGC TAACGAGGGC GTGGCTGCCC
10741  CGTCGTTTCC AAGACCCCTA GCCAGCCGAC TTCTCCAGTT ACGGAGCGAG CCCCTCTTTT
10801  GTTTTGTTTG TTTTTGCCAG ATGCATCCCG TACTGCGGCA GATGCGCCCC CACCACCCTC
10861  CACCGCAACA ACAGCCCACT CCACAGCCGG CGCTTCTGCC CCCGCCCCAG CAGCAGCAAC
10921  TTCCAGCCAC GACCGCCGCG GCCGCCGTGA GCGGGGCTGG ACAGACTTCT CAGTATGACC
10981  ACCTGGCCTT GGAAGAGGGC GAGGGGCTGG CGCGCCTGGG GGCGTCGTCG CCGGAGCGGC
11041  ACCCGCGCGT GCAGATGAAA CGGGACGCTC GCGAGGCCTA CGTGCCCAAG CAGAACCTGT
11101  TCAGAGACAG GAGCGGCGAG GAGCCCGAGG AGATGCGCGC GGCCCGGTTC CACGCGGGGC
11161  GGGAGCTGCG GCGCGGCCTG GACCGAAAGA GGGTGCTGAG GGACGAGGAT TTCGAGGCGG
11221  ACGAGCTGAC GGGGATCAGC CCCGCGCGCG CGCACGTGGC CGCGGCCAAC CTGGTCACGG
11281  CGTACGAGCA GACCGTGAAG GAGGAGAGCA ACTTCCAAAA ATCCTTCAAC AACCACGTGC
11341  GCACCCTGAT CGCGCGCGAG GAGGTGACCC TGGGCCTGAT GCACCTGTGG GACCTGCTGG
11401  AGGCCATCGT GCAGAACCCC ACCAGCAAGC CGCTGACGGC GCAGCTGTTC CTGGTGGTGC
11461  AGCACAGTCG GGACAACGAG GCGTTCAGGG AGGCGCTGCT GAATATCACC GAGCCCGAGG
11521  GCCGCTGGCT CCTGGACCTG GTGAACATTC TGCAGAGCAT CGTGGTGCAG GAGCGCGGGC
11581  TGCCGCTGTC CGAGAAGCTG GCGGCCATCA ACTTCTCGGT GCTGAGTCTG GGCAAGTACT
11641  ACGCTAGGAA GATCTACAAG ACCCCGTACG TGCCCATAGA CAAGGAGGTG AAGATCGACG
11701  GGTTTTACAT GCGCATGACC CTGAAAGTGC TGACCCTGAG CGACGATCTG GGGGTGTACC
11761  GCAACGACAG GATGCACCGC GCGGTGAGCG CCAGCCGCCG GCGCGAGCTG AGCGACCAGG
11821  AGCTGATGCA CAGCCTGCAG CGGGCCCTGA CCGGGGCCGG GACCGAGGGG GAGAGCTACT
11881  TTGACATGGG CGCGGACCTG CGCTGGCAGC CCAGCCGCCG GGCCTTGGAA GCTGCCGGCG
11941  GTTCCCCCTA CGTGGAGGAG GTGGACGATG AGGAGGAGGA GGGCGAGTAC CTGGAAGACT
12001  GATGGCGCGA CCGTATTTTT GCTAGATGCA GCAACAGCCA CCGCCGCCTC CTGATCCCGC
12061  GATGCGGGCG GCGCTGCAGA GCCAGCCGTC CGGCATTAAC TCCTCGGACG ATTGGACCCA
12121  GGCCATGCAA CGCATCATGG CGCTGACGAC CCGCAATCCC GAAGCCTTTA GACAGCAGCC
12181  TCAGGCCAAC CGGCTCTCGG CCATCCTGGA GGCCGTGGTG CCCTCGCGCT CGAACCCCAC
12241  GCACGAGAAG GTGCTGGCCA TCGTGAACGC GCTGGTGGAG AACAAGGCCA TCCGCGGCGA
12301  CGAGGCCGGG CTGGTGTACA ACGCGCTGCT GGAGCGCGTG GCCCGCTACA ACAGCACCAA
12361  CGTGCAGACG AACCTGGACC GCATGGTGAC CGACGTGCGC GAGGCGGTGT CGCAGCGCGA
12421  GCGGTTCCAC CGCGAGTCGA ACCTGGGCTC CATGGTGGCG CTGAACGCCT TCCTGAGCAC
12481  GCAGCCCGCC AACGTGCCCC GGGGCCAGGA GGACTACACC AACTTCATCA GCGCGCTGCG
12541  GCTGATGGTG GCCGAGGTGC CCCAGAGCGA GGTGTACCAG TCGGGGCCGG ACTACTTCTT
12601  CCAGACCAGT CGCCAGGGCT TGCAGACCGT GAACCTGAGC CAGGCTTTCA AGAACTTGCA
12661  GGGACTGTGG GGCGTGCAGG CCCCGGTCGG GGACCGCGCG ACGGTGTCGA GCCTGCTGAC
12721  GCCGAACTCG CGCCTGCTGC TGCTGCTGGT GGCGCCCTTC ACGGACAGCG GCAGCGTGAG
12781  CCGCGACTCG TACCTGGGCT ACCTGCTTAA CCTGTACCGC GAGGCCATCG GGCAGGCGCA
12841  CGTGGACGAG CAGACCTACC AGGAGATCAC CCACGTGAGC CGCGCGCTGG GCCAGGAGGA
12901  CCCGGGCAAC CTGGAGGCCA CCCTGAACTT CCTGCTGACC AACCGGTCGC AGAAGATCCC
12961  GCCCCAGTAC GCGCTGAGCA CCGAGGAGGA GCGCATCCTG CGCTACGTGC AGCAGAGCGT
13021  GGGGCTGTTC CTGATGCAGG AGGGGGCCAC GCCCAGCGCC GCGCTCGACA TGACCGCGCG
13081  CAACATGGAG CCCAGCATGT ACGCCCGCAA CCGCCCGTTC ATCAATAAGC TGATGGACTA
13141  CTTGCATCGG GCGGCCGCCA TGAACTCGGA CTACTTTACC AACGCCATCT TGAACCCGCA
13201  CTGGCTCCCG CCGCCCGGGT TCTACACGGG CGAGTATGAC ATGCCCGACC CCAACGACGG
13261  GTTCCTGTGG GATGACGTGG ACAGCAGCGT GTTCTCGCCG CGCCCCGCCA CCACCGTGTG
13321  GAAGAAAGAG GGCGGGGACC GGCGGCCGTC CTCGGCGCTG TCCGGTCGCG CGGGTGCTGC
13381  CGCGGCGGTG CCCGAGGCCG CCAGCCCCTT CCCGAGCCTG CCCTTTTCGC TGAACAGCGT
13441  GCGCAGCAGC GAGCTGGGAC GGCTGACGCG GCCGCGCCTG CTGGGCGAGG AGGAGTACCT
13501  GAACGACTCC TTGTTGAGGC CCGAGCGCGA GAAGAACTTC CCCAATAACG GGATAGAGAG
13561  CCTGGTGGAC AAGATGAGCC GCTGGAAGAC GTACGCGCAC GAGCACAGGG ACGAGCCGCG
13621  AGCTAGCAGC AGCACCGGCG CCCGTAGACG CCAGCGGCAC GACAGGCAGC GGGGACTGGT
```

FIG. 6D

```
13681  GTGGGACGAT GAGGATTCCG CCGACGACAG CAGCGTGTTG GACTTGGGTG GGAGTGGTGG
13741  TGGTAACCCG TTCGCTCACC TGCGCCCCCG TATCGGGCGC CTGATGTAAG AATCTGAAAA
13801  AATAAAAAAC GGTACTCACC AAGGCCATGG CGACCAGCGT GCGTTCTTCT CTGTTGTTTG
13861  TAGTAGTATG ATGAGGCGCG TGTACCCGGA GGGTCCTCCT CCCTCGTACG AGAGCGTGAT
13921  GCAGCAGGCG GTGGCGGCGG CGATGCAGCC CCCGCTGGAG GCGCCTTACG TGCCCCCGCG
13981  GTACCTGGCG CCTACGGAGG GGCGGAACAG CATTCGTTAC TCGGAGCTGG CACCCTTGTA
14041  CGATACCACC CGGTTGTACC TGGTGGACAA CAAGTCGGCG GACATCGCCT CGCTGAACTA
14101  CCAGAACGAC CACAGCAACT TCCTGACCAC CGTGGTGCAG AACAACGATT TCACCCCCAC
14161  GGAGGCCAGC ACCCAGACCA TCAACTTTGA CGAGCGCTCG CGGTGGGGCG GCCAGCTGAA
14221  AACCATCATG CACACCAACA TGCCCAACGT GAACGAGTTC ATGTACAGCA ACAAGTTCAA
14281  GGCGCGGGTC ATGGTCTCGC GCAAGACCCC CAACGGGGTC GCGGTAGGGG ATGATTATGA
14341  TGGTAGTCAG GACGAGCTGA CCTACGAGTG GGTGGAGTTT GAGCTGCCCG AGGGCAACTT
14401  CTCGGTGACC ATGACCATCG ATCTGATGAA CAACGCCATC ATCGACAATT ACTTGGCGGT
14461  GGGACGGCAG AACGGGGTGC TGGAGAGCGA CATCGGCGTG AAGTTCGACA CGCGCAACTT
14521  CCGGCTGGGC TGGGACCCCG TGACCGAGCT GGTGATGCCG GGCGTGTACA CCAACGAGGC
14581  CTTCCACCCC GACATCGTCC TGCTGCCCGG CTGCGGCGTG GACTTCACCG AGAGCCGCCT
14641  CAGCAACCTG CTGGGCATCC GCAAGCGGCA GCCCTTCCAG GAGGGCTTCC AGATCCTGTA
14701  CGAGGACCTG GAGGGGGGCA ACATCCCCGC GCTCTTGGAT GTCGAAGCCT ATGAAGAAAG
14761  TAAGGAAAAA GCAGAGGCTG AGGCAACTGC AGCCGTGGCT ACTGCCGCTG TCACCGATGC
14821  AGATGCAGCT ACTACCAGGG GCGATACATT CGCCACTGTG GCTGAAGAAG CAGCCGCCGT
14881  AGCGGCGACC GATGATAGTG AAAGTAAGAT AGTCATCAAG CCGGTGGAGA AGGACAGCAA
14941  GAACAGGAGC TACAACGTTC TATCGGATGG AAAGAACACC GCCTACCGCA GCTGGTACCT
15001  GGCCTACAAC TACGGCGACC CCGAGAAGGG CGTGCGCTCC TGGACGCTGC TCACCACCTC
15061  GGACGTCACC TGCGGCGTGG AGCAAGTCTA CTGGTCGCTG CCCGACATGA TGCAAGACCC
15121  GGTCACCTTC CGCTCCACGC GTCAAGTTAG CAACTACCCG GTGGTGGGCG CCGAGCTCCT
15181  GCCCGTCTAC TCCAAGAGCT TCTTCAACGA GCAGGCCGTC TACTCGCAGC AGCTGCGCGC
15241  CTTCACCTCG CTCACGCACG TCTTCAACCG CTTCCCCGAG AACCAGATCC TCGTCCGCCC
15301  GCCCGCGCCC ACCATTACCA CCGTCAGTGA AAACGTTCCT GCTCTCACAG ATCACGGGAC
15361  CCTGCCGCTG CGCAGCAGTA TCCGGGGAGT CCAGCGCGTG ACCGTCACTG ACGCCAGACG
15421  CCGCACCTGC CCCTACGTCT ACAAGGCCCT GGGCGTAGTC GCGCCGCGCG TCCTCTCGAG
15481  CCGCACCTTC TAAAAAATGT CCATTCTCAT CTCGCCCAGT AATAACACCG GTTGGGGCCT
15541  GCGCGCGCCC AGCAAGATGT ACGGAGGCGC TCGCCAACGC TCCACGCAAC ACCCCGTGCG
15601  CGTGCGCGGG CACTTCCGCG CTCCCTGGGG CGCCCTCAAG GGCCGCGTGC GCTCGCGCAC
15661  CACCGTCGAC GACGTGATCG ACCAGGTGGT GGCCGACGCG CGCAACTACA CGCCCGCCGC
15721  CGCGCCCGTC TCCACCGTGG ACGCCGTCAT CGACAGCGTG GTGGCCGACG CGCGCCGGTA
15781  CGCCCGCGCC AAGAGCCGGC GGCGGCGCAT CGCCCGGCGG CACCGGAGCA CCCCCGCCAT
15841  GCGCGCGGCG CGAGCCTTGC TGCGCAGGGC CAGGCGCACG GGACGCAGGG CCATGCTCAG
15901  GGCGGCCAGA CGCGCGGCCT CTGGCAGCAG CAGCGCCGGC AGGACCCGCA GACGCGCGGC
15961  CACGGCGGCG GCGGCGGCCA TCGCCAGCAT GTCCCGCCCG CGGCGCGGCA ACGTGTACTG
16021  GGTGCGCGAC GCCGCCACCG GTGTGCGCGT GCCCGTGCGC ACCCGCCCCC CTCGCACTTG
16081  AAGATGCTGA CTTCGCGATG TTGATGTGTC CCAGCGGCGA GGAGGATGTC CAAGCGCAAA
16141  TTCAAGGAAG AGATGCTCCA GGTCATCGCG CCTGAGATCT ACGGCCCCGC GGCGGCGGTG
16201  AAGGAGGAAA GAAAGCCCCG CAAACTGAAG CGGGTCAAAA AGGACAAAAA GGAGGAGGAA
16261  GATGTGGACG GACTGGTGGA GTTTGTGCGC GAGTTCGCCC CCCGGCGGCG CGTGCAGTGG
16321  CGCGGGCGGA AAGTGAAACC GGTGCTGCGA CCCGGCACCA CCGTGGTCTT CACGCCCGGC
16381  GAGCGTTCCG GCTCCGCCTC CAAGCGCTCC TACGACGAGG TGTACGGGGA CGAGGACATC
16441  CTCGAGCAGG CGGCCGAGCG TCTGGGCGAG TTTGCTTACG GCAAGCGCAG CCGCCCCGCG
16501  CCCTTGAAAG AGGAGGCGGT GTCCATCCCG CTGGACCACG GCAACCCCAC GCCGAGTCTG
16561  AAGCCGGTGA CCCTGCAGCA GGTGCTGCCG AGCGCGGCGC CGCGCCGGGG CTTCAAGCGC
16621  GAGGGCGGCG AGGATCTGTA CCCGACCATG CAGCTGATGG TGCCCAAGCG CCAGAAGCTG
16681  GAGGACGTGC TGGAGCACAT GAAGGTGGAC CCCGAGGTGC AGCCCGAGGT CAAGGTGCGG
16741  CCCATCAAGC AGGTGGCCCC GGGCCTGGGC GTGCAGACCG TGGACATCAA GATCCCCACG
16801  GAGCCCATGG AAACGCAGAC CGAGCCCGTG AAGCCCAGCA CCAGCACCAT GGAGGTGCAG
16861  ACGGATCCCT GGATGCCGGC GCCGGCTTCC ACCACCACCA CTCGCCGAAG ACGCAAGTAC
16921  GGCGCGGCCA GCCTGCTGAT GCCCAACTAC GCGCTGCATC CTTCCATCAT CCCCACGCCG
16981  GGCTACCGCG GCACGCGCTT CTACCGCGGC TACAGCAGCC GCCGCAAGAC CACCACCCGC
17041  CGCCGCCGTC GTCGCACCCG CCGCAGCAGC ACCGCGACTT CCGCCGCCTT GGTGCGGAGA
```

FIG. 6E

```
17101  GTGTACCGCA GCGGGCGCGA GCCTCTGACC CTGCCGCGCG CGCGCTACCA CCCGAGCATC
17161  GCCATTTAAC TCTGCCGTCG CCTCCTACTT GCAGATATGG CCCTCACATG CCGCCTCCGC
17221  GTCCCCATTA CGGGCTACCG AGGAAGAAAG CCGCGCCGTA GAAGGCTGAC GGGGAACGGG
17281  CTGCGTCGCC ATCACCACCG GCGGCGGCGC GCCATCAGCA AGCGGTTGGG GGGAGGCTTC
17341  CTGCCCGCGC TGATCCCCAT CATCGCCGCG GCGATCGGGG CGATCCCCGG CATAGCTTCC
17401  GTGGCGGTGC AGGCCTCTCA GCGCCACTGA GACACAGCTT GGAAAATTTG TAATAAAAAA
17461  TGGACTGACG CTCCTGGTCC TGTGATGTGT GTTTTTAGAT GGAAGACATC AATTTTTCGT
17521  CCCTGGCACC GCGACACGGC ACGCGGCCGT TTATGGGCAC CTGGAGCGAC ATCGGCAACA
17581  GCCAACTGAA CGGGGGCGCC TTCAATTGGA GCAGTCTCTG GAGCGGGCTT AAGAATTTCG
17641  GGTCCACGCT CAAAACCTAT GGCAACAAGG CGTGGAACAG CAGCACAGGG CAGGCGCTGA
17701  GGGAAAAGCT GAAAGAGCAG AACTTCCAGC AGAAGGTGGT CGATGGCCTG GCCTCGGGCA
17761  TCAACGGGGT GGTGGACCTG GCCAACCAGG CCGTGCAGAA ACAGATCAAC AGCCGCCTGG
17821  ACGCGGTCCC GCCCGCGGGG TCCGTGGAGA TGCCCCAGGT GGAGGAGGAG CTGCCTCCCC
17881  TGGACAAGCG CGGCGACAAG CGACCGCGTC CCGATGCAGA GGAGACGCTG CTGACGCACA
17941  CGGACGAGCC GCCCCCGTAC GAGGAGGCGG TGAAACTGGG TCTGCCCACC ACGCGGCCCG
18001  TGGCGCCTCT GGCCACCGGG GTGCTGAAAC CCAGCAGCAG CAGCCAGCCC GCGACCCTGG
18061  ACTTGCCTCC GCCTGCTTCC CGCCCCTCCA CAGTGGCTAA GCCCCTGCCG CCGGTGGCCG
18121  TCGCGTCGCG CGCCCCCCGA GGCCGCCCCC AGGCGAACTG GCAGAGCACT CTGAACAGCA
18181  TCGTGGGTCT GGGAGTGCAG AGTGTGAAGC GCCGCCGCTG CTATTAAAAG ACACTGTAGC
18241  GCTTAACTTG CTTGTCTGTG TGTATATGTA TGTCCGCCGA CCAGAAGGAA GAGGCGCGTC
18301  GCCGAGTTGC AAGATGGCCA CCCCATCGAT GCTGCCCCAG TGGGCGTACA TGCACATCGC
18361  CGGACAGGAC GCTTCGGAGT ACCTGAGTCC GGGTCTGGTG CAGTTCGCCC GCGCCACAGA
18421  CACCTACTTC AGTCTGGGGA ACAAGTTTAG GAACCCCACG GTGGCGCCCA CGCACGATGT
18481  GACCACCGAC CGCAGCCAGC GGCTGACGCT GCGCTTCGTG CCCGTGGACC GCGAGGACAA
18541  CACCTACTCG TACAAAGTGC GCTACACGCT GGCCGTGGGC GACAACCGCG TGCTGGACAT
18601  GGCCAGCACC TACTTTGACA TCCGCGGCGT GCTGGATCGG GGCCCCAGCT TCAAACCCTA
18661  CTCCGGCACC GCCTACAACA GCCTGGCTCC CAAGGGAGCG CCCAACACCT CACAGTGGAT
18721  AACCAAAGAC AATGGAACTG ATAAGACATA CAGTTTTGGA AATGCTCCAG TCAGAGGATT
18781  GGACATTACA GAAGAGGGTC TCCAAATAGG ACCCGATGAG TCAGGGGGTG AAAGCAAGAA
18841  AATTTTTGCA GACAAAACCT ATCAGCCTGA ACCTCAGCTT GGAGATGAGG AATGGCATGA
18901  TACTATTGGA GCTGAAGACA AGTATGGAGG CAGAGCGCTT AAACCTGCCA CCAACATGAA
18961  ACCCTGCTAT GGGTCTTTCG CCAAGCCAAC TAATGCTAAG GGAGGTCAGG CTAAAAGCAG
19021  AACCAAGGAC GATGGCACTA CTGAGCCTGA TATTGACATG GCCTTCTTTG ACGATCGCAG
19081  TCAGCAAGCT AGTTTCAGTC CAGAACTTGT TTTGTATACT GAGAATGTCG ATCTGGACAC
19141  CCCGGATACC CACATTATTT ACAAACCTGG CACTGATGAA ACAAGTTCTT CTTTCAACTT
19201  GGGTCAGCAG TCCATGCCCA ACAGACCCAA CTACATCGGC TTCAGAGACA ACTTTATCGG
19261  TCTCATGTAC TACAACAGTA CTGGCAATAT GGGTGTACTA GCTGGACAGG CCTCCCAGCT
19321  GAATGCTGTG GTGGACTTGC AGGACAGAAA CACTGAACTG TCCTACCAGC TCTTGCTTGA
19381  CTCTCTGGGT GACAGAACCA GGTATTTCAG TATGTGGAAC CAGGCGGTGG ACAGCTACGA
19441  CCCCGATGTG CGCATTATTG AAAATCACGG TGTGGAGGAT GAACTACCCA ACTATTGCTT
19501  CCCTTTGAAT GGTGTGGGCT TTACAGATAC ATTCCAGGGA ATTAAGGTTA AAACTACCAA
19561  TAACGGAACA GCAAATGCTA CAGAGTGGGA ATCTGATACC TCTGTCAATA ATGCTAATGA
19621  GATTGCCAAG GGCAATCCTT TCGCCATGGA GATCAACATC CAGGCCAACC TGTGGCGGAA
19681  CTTCCTCTAC GCGAACGTGG CGCTGTACCT GCCCGACTCC TACAAGTACA CGCCGGCCAA
19741  CATCACGCTG CCCGCCAACA CCAACACCTA CGATTACATG AACGGCCGCG TGGTAGCGCC
19801  CTCGCTGGTG GACGCCTACA TCAACATCGG GGCGCGCTGG TCGCTGGACC CCATGGACAA
19861  CGTCAACCCC TTCAACCACC ACCGCAACGC GGGCCTGCGC TACCGCTCCA TGCTCCTGGG
19921  CAACGGGCGC TACGTGCCCT TCCACATCCA GGTGCCCCAA AAGTTTTTCG CCATCAAGAG
19981  CCTCCTGCTC CTGCCCGGGT CCTACACCTA CGAGTGGAAC TTCCGCAAGG ACGTCAACAT
20041  GATCCTGCAG AGCTCCCTCG GCAACGACCT GCGCACGGAC GGGGCCTCCA TCGCCTTCAC
20101  CAGCATCAAC CTCTACGCCA CCTTCTTCCC CATGGCGCAC AACACCGCCT CCACGCTCGA
20161  GGCCATGCTG CGCAACGACA CCAACGACCA GTCCTTCAAC GACTACCTCT CGGCGGCCAA
20221  CATGCTCTAC CCCATCCCGG CCAACGCCAC CAACGTGCCC ATCTCCATCC CCTCGCGCAA
20281  CTGGGCCGCC TTCCGCGGCT GGTCCTTCAC GCGCCTCAAG ACCCGCGAGA CGCCCTCGCT
20341  CGGCTCCGGG TTCGACCCCT ACTTCGTCTA CTCGGGCTCC ATCCCCTACC TCGACGGCAC
20401  CTTCTACCTC AACCACACCT TCAAGAAGGT CTCCATCACC TTCGACTCCT CCGTCAGCTG
20461  GCCCGGCAAC GACCGCCTCC TGACGCCCAA CGAGTTCGAA ATCAAGCGCA CCGTCGACGG
```

FIG. 6F

```
20521  AGAGGGGTAC AACGTGGCCC AGTGCAACAT GACCAAGGAC TGGTTCCTGG TTCAGATGCT
20581  GGCCCACTAC AACATCGGCT ACCAGGGCTT CTACGTGCCC GAGGGCTACA AGGACCGCAT
20641  GTACTCCTTC TTCCGCAACT TCCAGCCCAT GAGCCGCCAG GTCGTGGACG AGGTCAACTA
20701  CAAGGACTAC CAGGCCGTCA CCCTGGCCTA CCAGCACAAC AACTCGGGCT TCGTCGGCTA
20761  CCTCGCGCCC ACCATGCGCC AGGGACAGCC CTACCCCGCC AACTACCCCT ACCCGCTCAT
20821  CGGCAAGAGC GCCGTCGCCA GCGTCACCCA GAAAAAGTTC CTCTGCGACC GGGTCATGTG
20881  GCGCATCCCC TTCTCCAGCA ACTTCATGTC CATGGGCGCG CTCACCGACC TCGGCCAGAA
20941  CATGCTCTAC GCCAACTCCG CCCACGCGCT AGACATGAAT TTCGAAGTCG ACCCCATGGA
21001  TGAGTCCACC CTTCTCTATG TTGTCTTCGA AGTCTTCGAC GTCGTCCGAG TGCACCAGCC
21061  CCACCGCGGC GTCATCGAGG CCGTCTACCT GCGCACGCCC TTCTCGGCCG GTAACGCCAC
21121  CACCTAAGCC CCGCTCTTGC TTCTTGCAAG ATGACGGCCT GTGCGGGCTC CGGCGAGCAG
21181  GAGCTCAGGG CCATCCTCCG CGACCTGGGC TGCGGGCCCT GCTTCCTGGG CACCTTCGAC
21241  AAGCGCTTCC CGGGATTCAT GGCCCCGCAC AAGCTGGCCT GCGCCATCGT CAACACGGCC
21301  GGCCGCGAGA CCGGGGGCGA GCACTGGCTG GCCTTCGCCT GGAACCCGCG CTCCCACACC
21361  TGCTACCTCT TCGACCCCTT CGGGTTCTCG AACGAGCGCC TCAAGCAGAT CTACCAGTTC
21421  GAGTACGAGG GCCTGCTGCG CCGCAGCGCC CTGGCCACCG AGGACCGCTG CGTCACCCTG
21481  GAAAAGTCCA CCCAGACCGT GCAGGGTCCG CGCTCGGCCG CCTGCGGGCT CTTCTGCTGC
21541  ATGTTCCTGC ACGCCTTCGT GCACTGGCCC GACCGCCCCA TGGACAAGAA CCCCACCATG
21601  AACTTGCTGA CGGGGGTGCC CAACGGCATG CTCCAGTCGC CCCAGGTGGA ACCCACCCTG
21661  CGCCGCAACC AGGAAGCGCT CTACCGCTTC CTCAACGCCC ACTCCGCCTA CTTTCGCTCC
21721  CACCGCGCGC GCATCGAGAA GGCCACCGCC TTCGACCGCA TGAATCAAGA CATGTAAACC
21781  GTGTGTGTAT GTGAATGCTT TATTCATAAT AAACAGCACA TGTTTATGCC ACCTTCTCTG
21841  AGGCTCTGAC TTTATTTAGA AATCGAAGGG GTTCTGCCGG CTCTCGGCAT GCCCCGCGGG
21901  CAGGGATACG TTGCGGAACT GGTACTTGGG CAGCCACTTG AACTCGGGGA TCAGCAGCTT
21961  GGGCACGGGG AGGTCGGGGA ACGAGTCGCT CCACAGCTTG CGCGTGAGTT GCAGGGCGCC
22021  CAGCAGGTCG GGCGCGGAGA TCTTGAAATC GCAGTTGGGA CCCGCGTTCT GCGCGCGAGA
22081  GTTGCGGTAC ACGGGGTTGC AGCACTGGAA CACCATCAGG GCCGGGTGCT TCACGCTCGC
22141  CAGCACCGTC GCGTCGGTGA TGCCCTCCAC GTCCAGATCC TCGGCGTTGG CCATCCCGAA
22201  GGGGGTCATC TTGCAGGTCT GCCGCCCCAT GCTGGGCACG CAGCCGGGCT TGTGGTTGCA
22261  ATCGCAGTGC AGGGGGATCA GCATCATCTG GGCCTGCTCG GAGCTCATGC CCGGGTACAT
22321  GGCCTTCATG AAAGCCTCCA GCTGGCGGAA GGCCTGCTGC GCCTTGCCGC CCTCGGTGAA
22381  GAAGACCCCG CAGGACTTGC TAGAGAACTG GTTGGTGGCG CAGCCCGCGT CGTGCACGCA
22441  GCAGCGCGCG TCGTTGTTGG CCAGCTGCAC CACGCTGCGC CCCCAGCGGT TCTGGGTGAT
22501  CTTGGCCCGG TCGGGGTTCT CCTTCAGCGC GCGCTGTCCG TTCTCGCTCG CCACATCCAT
22561  CTCGATCGTG TGCTCCTTCT GGATCATCAC GGTCCCGTGC AGGCACCGCA GCTTGCTCTC
22621  GGCCTCGGTG CACCCGTGCA GCCACAGCGC GCAGCCGGTG CTCTCCCAGT TCTTGTGGGC
22681  GATCTGGGAG TGCGAGTGCA CGAAGCCCTG CAGGAAGCGG CCCATCATCG CGGTCAGGGT
22741  CTTGTTGCTG GTGAAGGTCA GCGGGATGCC GCGGTGCTCC TCGTTCACAT ACAGGTGGCA
22801  GATGCGGCGG TACACCTCGC CCTGCTCGGG CATCAGCTGG AAGGCGGACT TCAGGTCGCT
22861  CTCCACGCGG TACCGGTCCA TCAGCAGCGT CATGACTTCC ATGCCCTTCT CCCAGGCCGA
22921  AACGATCGGC AGGCTCAGGG GGTTCTTCAC CGTTGTCATC TTAGTCGCCG CCGCCGAGGT
22981  CAGGGGGTCG TTCTCGTCCA GGGTCTCAAA CACTCGCTTG CCGTCCTTCT CGATGATGCG
23041  CACGGGGGGG AAGCTGAAGC CCACGGCCGC CAGCTCCTCC TCGGCCTGCC TTTCGTCCTC
23101  GCTGTCCTGG CTGATGTCTT GCAAAGGCAC ATGCTTGGTC TTGCGGGGTT TCTTTTTGGG
23161  CGGCAGAGGC GGCGGCGGAG ACGTGCTGGG CGAGCGCGAG TTCTCGCTCA CCACGACTAT
23221  TTCTTCTTCT TGGCCGTCGT CCGAGACCAC GCGGCGGTAG GCATGCCTCT TCTGGGGCAG
23281  AGGCGGAGGC GACGGGCTCT CGCGGTTCGG CGGGCGGCTG GCAGAGCCCC TTCCGCGTTC
23341  GGGGGTGCGC TCCTGGCGGC GCTGCTCTGA CTGACTTCCT CCGCGGCCGG CCATTGTGTT
23401  CTCCTAGGGA GCAACAACAA GCATGGAGAC TCAGCCATCG TCGCCAACAT CGCCATCTGC
23461  CCCCGCCGCC GACGAGAACC AGCAGAATGA AGCTTAACC GCCCCGCCGC CCAGCCCCAC
23521  CTCCGACGCC GCGGCCCCAG ACATGCAAGA GATGGAGGAA TCCATCGAGA TTGACCTGGG
23581  CTACGTGACG CCCGCGGAGC ACGAGGAGGA GCTGGCAGCG CGCTTTTCAG CCCCGGAAGA
23641  GAACCACCAA GAGCAGCCAG AGCAGGAAGC AGAGAGCGAG CAGAACCAGG CTGGGCTCGA
23701  GCATGGCGAC TACCTGAGCG GGGCAGAGGA CGTGCTCATC AAGCATCTGA CCCGCCAATG
23761  CATCATCGTC AAGGACGCGC TGCTCGACCG CGCCGAGGTG CCCCTCAGCG TGGCGGAGCT
23821  CAGCCGCGCC TACGAGCGCA ACCTCTTCTC GCCGCGCGTG CCCCCCAAGC GCCAGCCCAA
23881  CGGCACCTGC GAGCCCAACC CGCGCCTCAA CTTCTACCCG GTCTTCGCGG TGCCCGAGGC
```

FIG. 6G

```
23941  CCTGGCCACC TACCACCTCT TTTTCAAGAA CCAAAGGATC CCCGTCTCCT GCCGCGCCAA
24001  CCGCACCCGC GCCGACGCCC TGCTCAACCT GGGCCCCGGC GCCCGCCTAC CTGATATCAC
24061  CTCCTTGGAA GAGGTTCCCA AGATCTTCGA GGGTCTGGGC AGCGACGAGA CTCGGGCCGC
24121  GAACGCTCTG CAAGGAAGCG GAGAGGAACA TGAGCACCAC AGCGCCCTGG TGGAGTTGGA
24181  AGGCGACAAC GCGCGCCTGG CGGTGCTCAA GCGCACGGTC GAGCTGACCC ACTTCGCCTA
24241  CCCGGCGCTC AACCTGCCCC CCAAGGTCAT GAGCGCCGTC ATGGACCAGG TGCTCATCAA
24301  GCGCGCCTCG CCCATTGAGG ACATGCAGGA CCCCGAGAGC TCGGACGAGG GCAAGCCCGT
24361  GGTCAGCGAC GAGCAGCTGG CGCGCTGGCT GGGAGCGAGT AGCACCCCCC AGAGCCTGGA
24421  AGAGCGGCGC AAGCTCATGA TGGCCGTGGT CCTGGTGACC GTGGAGCTGG AGTGTCTGCG
24481  CCGCTTCTTC GCCGACGCAG AGACCCTGCG CAAGGTCGAG GAGAACCTGC ACTACCTCTT
24541  CAGGCACGGG TTCGTGCGCC AGGCCTGCAA GATCTCCAAC GTGGAGCTGA CCAACCTGGT
24601  CTCCTACATG GGCATCCTGC ACGAGAACCG CCTGGGGCAG AACGTGCTGC ACACCACCCT
24661  GCGCGGGGAG GCCCGCCGCG ACTACATCCG CGACTGCGTC TACCTGTACC TCTGCCACAC
24721  CTGGCAGACG GGCATGGGCG TGTGGCAGCA GTGCCTGGAG GAGCAGAACC TGAAAGAGCT
24781  CTGCAAGCTC CTGCAGAAGA ACCTCAAGGC CCTGTGGACC GGGTTCGACG AGCGCACCAC
24841  CGCCTCGGAC CTGGCCGACC TCATCTTCCC CGAGCGCCTG CGGCTGACGC TGCGCAACGG
24901  GCTGCCCGAC TTTATGAGCC AAAGCATGTT GCAAAACTTT CGCTCTTTCA TCCTCGAACG
24961  CTCCGGGATC CTGCCCGCCA CCTGCTCCGC GCTGCCCTCG GACTTCGTGC CGCTGACCTT
25021  CCGCGAGTGC CCCCCGCCGC TCTGGAGCCA CTGCTACTTG CTGCGCCTGG CCAACTACCT
25081  GGCCTACCAC TCGGACGTGA TCGAGGACGT CAGCGGCGAG GGTCTGCTGG AGTGCCACTG
25141  CCGCTGCAAC CTCTGCACGC CGCACCGCTC CCTGGCCTGC AACCCCCAGC TGCTGAGCGA
25201  GACCCAGATC ATCGGCACCT TCGAGTTGCA AGGCCCCGGC GAGGAGGGCA AGGGGGGTCT
25261  GAAACTCACC CCGGGGCTGT GGACCTCGGC CTACTTGCGC AAGTTCGTGC CCGAGGACTA
25321  CCATCCCTTC GAGATCAGGT TCTACGAGGA CCAATCCCAG CCGCCCAAGG CCGAGCTGTC
25381  GGCCTGCGTC ATCACCCAGG GGGCCATCCT GGCCCAATTG CAAGCCATCC AGAAATCCCG
25441  CCAAGAATTT CTGCTGAAAA AGGGCCACGG GGTCTACTTG GACCCCCAGA CCGGAGAGGA
25501  GCTCAACCGC AGCTTCCCCC AGGATGCCCC GAGGAAGCAG CAAGAAGCTG AAAGTGGAGC
25561  TGCCGCCGCC GGAGGATTTG GAGGAAGACT GGGAGAGCAG TCAGGCAGAG GAGGAGATGG
25621  AAGACTGGGA CAGCACTCAG GCAGAGGAGG ACAGCCTGCA AGACAGTCTG GAGGAGGAAG
25681  ACGAGGTGGA GGAGGAGGAG GCAGAGGAAG AAGCAGCCGC CGCCAGACCG TCGTCCTCGG
25741  CGGAGAAAGC AAGCAGCACG GATACCATCT CCGCTCCGGG TCGGGGTCGC GGCGGCCGGG
25801  CCCACAGTAG GTGGGACGAG ACCGGGCGCT TCCCGAACCC CACCACCCAG ACCGGTAAGA
25861  AGGAGCGGCA GGGATACAAG TCCTGGCGGG GGCACAAAAA CGCCATCGTC TCCTGCTTGC
25921  AAGCCTGCGG GGGCAACATC TCCTTCACCC GGCGCTACCT GCTCTTCCAC CGCGGGGTGA
25981  ACTTCCCCCG CAACATCTTG CATTACTACC GTCACCTCCA CAGCCCCTAC TACTGTTTCC
26041  AAGAAGAGGC AGAAACCCAG CAGCAGCAGA AAACCAGCGA CAGCGGCAGC AGCTAGAAAA
26101  TCCACAGCGG CAGGTGGACT GAGGATCGCG GCGAACGAGC CGGCGCAGAC CCGGGAGCTG
26161  AGGAACCGGA TCTTTCCCAC CCTCTATGCC ATCTTCCAGC AGAGTCGGGG GCAGGAGCAG
26221  GAACTGAAAG TCAAGAACCG TTCTCTGCGC TCGCTCACCC GCAGTTGTCT GTATCACAAG
26281  AGCGAAGACC AACTTCAGCG CACTCTCGAG GACGCCGAGG CTCTCTTCAA CAAGTACTGC
26341  GCGCTCACTC TTAAAGAGTA GCCCGCGCCC GCCCACACAC GGAAAAAGGC GGGAATTACG
26401  TCACCACCTG CGCCCTTCGC CCGACCATCA TCATGAGCAA AGAGATTCCC ACGCCTTACA
26461  TGTGGAGCTA CCAGCCCCAG ATGGGTCTGG CCGCCGGCGC CGCCCAGGAC TACTCCACCC
26521  GCATGAACTG GCTCAGTGCC GGGCCCGCGA TGATCTCACG GGTGAATGAC ATCCGCGCCC
26581  ATCGAAACCA GATACTCCTA GAACAGTCAG CGATCACCGC CACGCCCCGC CATCACCTTA
26641  ATCCGCGTAA TTGGCCCGCC GCCCTGGTGT ACCAGGAAAT TCCCCAGCCC ACGACCGTAC
26701  TACTTCCGCG AGACGCCCAG GCCGAAGTCC AGCTGACTAA CTCAGGTGTC CAGCTGGCCG
26761  GCGGCGCCGC CCTGTGTCGT CACCGCCCCG CTCAGGGTAT AAAGCGGCTG GTGATCCGAG
26821  GCAGAGGCAC ACAGCTCAAC GACGAGGTGG TGAGCTCTTC GCTGGGTCTG CGACCTGACG
26881  GAGTCTTCCA ACTCGCCGGA TCGGGGAGAT CTTCCTTCAC GCCTCGTCAG GCCGTCCTGA
26941  CTTTGGAGAG TTCGTCCTCG CAGCCCCGCT CGGGCGGCAT CGGCACTCTC CAGTTCGTGG
27001  AGGAGTTCAC TCCCTCGGTC TACTTCAACC CCTTCTCCGG CTCCCCCGGC CACTACCCGG
27061  ACGAGTTCAT CCCGAACTTC GACGCCATCA GCGAGTCGGT GGACGGCTAC GATTGAATGT
27121  CCCATGGTGG CGCAGCTGAC CTAGCTCGGC TTCGACACCT GGACCACTGC CGCCGCTTCC
27181  GCTGCTTCGC TCGGGATCTC GCCGAGTTTG CCTACTTTGA GCTGCCCGAG GAGCACCCTC
27241  AGGGCCCGGC CCACGGAGTG CGGATCATCA TCGAAGGGGG CCTCGACTCC CACCTGCTTC
27301  GGATCTTCAG CCAGCGACCG ATCCTGGTCG AGCGCGAGCA AGGACAGACC CGTCTGACCC
```

FIG. 6H

```
27361  TGTACTGCAT CTGCAACCAC CCCGGCCTGC ATGAAAGTCT TTGTTGTCTG CTGTGTACTG
27421  AGTATAATAA AAGCTGAGAT CAGCGACTAC TCCGGACTCG ATTGTGGTGT TCCTGCTATC
27481  AACCGGTCCC TGTTCTTCAC CGGGAACGAG ACCGAGCTCC AGCTCCAGTG TAAGCCCCAC
27541  AAGAAGTACC TCACCTGGCT GTTCCAGGGC TCTCCGATCG CCGTTGTCAA CCACTGCGAC
27601  AACGACGGAG TCCTGCTGAG CGGCCCTGCC AACCTTACTT TTTCCACCCG CAGAAGCAAG
27661  CTCCAGCTCT TCCAACCCTT CCTCCCCGGG ACCTATCAGT GCGTCTCGGG ACCCTGCCAT
27721  CACACCTTCC ACCTGATCCC GAATACCACA GCGCCGCTCC CCGCTACTAA CAACCAAACT
27781  ACCCACCAAC GCCACCGTCG CGACCTTTCC TCTGAATCTA ATACTACCAC CCACACCGGA
27841  GGTGAGCTCC GAGGTCGACC AACCTCTGGG ATTTACTACG GCCCCTGGGA GGTGGTGGGG
27901  TTAATAGCGC TAGGCCTAGT TGTGGGTGGG CTTTTGGCTC TCTGCTACCT ATACCTCCCT
27961  TGCTGTTCGT ACTTAGTGGT GCTGTGTTGC TGGTTTAAGA AATGGGGAAG ATCACCCTAG
28021  TGAGCTGCGG TGTGCTGGTG GCGGTGTTGC TTTCGATTGT GGGACTGGGC GGCGCGGCTG
28081  TAGTGAAGGA GGAGAAGGCC GATCCCTGCT TGCATTTCAA TCCCGACAAA TGCCAGCTGA
28141  GTTTTCAGCC CGATGGCAAT CGGTGCACGG TGCTGATCAA GTGCGGATGG GAATGTGAGA
28201  ACGTGAGAAT CGAGTACAAT AACAAGACTC GGAACAATAC TCTCGCGTCC GTGTGGCAAC
28261  CCGGGGACCC CGAGTGGTAC ACCGTCTCTG TCCCCGGTGC TGACGGCTCC CCGCGCACCG
28321  TGAATAATAC TTTCATTTTT GCGCACATGT GCGACACGGT CATGTGGATG AGCAAGCAGT
28381  ACGATATGTG GCCCCCCACG AAGGAGAACA TCGTGGTCTT CTCCATCGCT TACAGCGTGT
28441  GCACGGCGCT AATCACCGCT ATCGTGTGCC TGAGCATTCA CATGCTCATC GCTATTCGCC
28501  CCAGAAATAA TGCCGAAAAA GAGAAACAGC CATAACACGT TTTTTCACAC ACCTTTTTCA
28561  GACCATGGCC TCTGTTACTG CCCTAATTAT TTTTTTGGGT CTCGTGGGCA CTAGCAGCAC
28621  TTTTCAGCAT ATAAACAAAA CTGTTTATGC TGGTTCTAAT TCTGTATTAC CTGGGCATCA
28681  ATCACACCAG AAAGTTTCAT GGTACTGGTA TGATAAAAAT AACACGCCAG TCACACTCTG
28741  CAAGGGTCAT CAAACACCCA TAAACCGTAG TGGAATTTTT TTTAAATGTA ATCATAATAA
28801  TATTACACTA CTTTCAATTA CAAAGCACTA TTCTGGTACT TACTATGGAA CCAATTTTAA
28861  CATAAAACAG GACACTTACT ATAGTGTCAC AGTATTGGAT CCAACTACTC CTAGAACAAC
28921  TACAAAACCC ACAACTACTA AGAGGCACAC TAAACCTAAA ACTACCAAGA AAACCACTGT
28981  CAAAACAACA ACTAGGACCA CCACAACTAC AGAGGCTACC ACCAGCACAA CACTTGCTGC
29041  AACTACACAC ACACACACTG AGCTAACCTT ACAGACCACT AATGATTTGA TAGCCCTGTT
29101  GCAAAAGGGG GATAACAGCA CCACTTCCGA TGAGGAAATA CCCAAATCCA TGATTGGCAT
29161  TATTGTTGCT GTAGTGGTGT GCATGTTGAT CATCGCCTTG TGCATGGTGT ACTATGCCTT
29221  CTGCTACAGA AAGCACAGAC TGAACGACAA GCTGGAACAC TTACTAAGTG TTGAATTTTA
29281  ATTTTTTAGA ACCATGAAGA TCCTAGGCCT TTTAGTTTTT TCTATCATTA CCTCTGCTCT
29341  TTGTGAATCA GTGAATAAAG ATGTTACTAT TACCACTGGT TCTAATTATA CACTGAAAGG
29401  GCCACCCTCA GGTATGCTTT CGTGGTATTG CTATTTTGGA ACTGACACTG ATCAAACTGA
29461  ATTATGCAAT TTTCAAAAAG GCAAAACCTC AAACTCTAAA ATCTCTAATT ATCAATGCAA
29521  TGGCACTGAT CTGATACTAC TCAATGTCAC GAAAGCATAT GGTGGCAGTT ATTCTTGCCC
29581  TGGACAAAAC ACTGAAGAAA TGATTTTTTA CAAAGTGGAA GTGGTTGATC CCACTACTCC
29641  ACCCACCACC ACAACTACTC ACACCACACA CACAGAACAA ACCACAGCAG AGGAGGCAGC
29701  AAAGTTAGCC TTGCAGGTCC AAGACAGTTC ATTTGTTGGC ATTACCCCTA CACCTGATCA
29761  GCGGTGTCCG GGGCTGCTAG TCAGCGGCAT TGTCGGTGTG CTTTCGGGAT TAGCAGTCAT
29821  AATCATCTGC ATGTTCATTT TTGCTTGCTG CTATAGAAGG CTTTACCGAC AAAAATCAGA
29881  CCCACTGCTG AACCTCTATG TTTAATTTTT TCCAGAGCCA TGAAGGCAGT TAGCACTCTA
29941  GTTTTTTGTT CTTTGATTGG CATTGTTTTT AGTGCTGGGT TTTTGAAAAA TCTTACCATT
30001  TATGAAGGTG AGAATGCCAC TCTAGTGGGC ATCAGTGGTC AAAATGTCAG CTGGCTAAAA
30061  TACCATCTAG ATGGGTGGAA AGACATTTGC GATTGGAATG TCACTGTGTA TACATGTAAT
30121  GGAGTTAACC TCACCATTAC TAATGCCACC CAAGATCAGA ATGGTAGGTT TAAGGGTCAG
30181  AGTTTCACTA GAAATAATGG GTATGAATCC CATAACATGT TTATCTATGA CGTCACTGTC
30241  ATCAGAAATG AGACCGCCAC CACCACACAG ATGCCCACTA CACACAGTTC TACCACTACT
30301  ACCAAGCAAA CCACACAGAC AACCACTTTT TATACATCAA CTCAGCATAT GACCACCACT
30361  ACAGCAGCAA AGCCAAGTAG CGCAGCGCCT CAGCCACAGG CTTTGGCTTT GAAAGCTGCA
30421  CAACCTAGTA CAACTACTAA GACCAATGAG CAGACTACTG ATTTTTTGTC CACTGTCGAG
30481  AGCCACACCA CAGCTACCTC CAGTGCCTTC TCTAGCACCG CCAATCTCTC CTCGCTTTCC
30541  TCTACACCAA TCAGTCCCGC TACTACTCCT AGCCCCGCTC CTCTTCCCAC TCCCCTGAAG
30601  CAAACAGACG GCGGCATGCA ATGGCAGATC ACCCTGCTCA TTGTGATCGG GTTGGTCATC
30661  CTGGCCGTGT TGCTCTACTA CATCTTCTGC CGCCGCATTC CCAACGCGCA CCGCAAGCCG
30721  GTCTACAAGC CCATCGTTGT CGGGCAGCCG GAGCCGCTTC AGGTGGAAGG GGGTCTAAGG
```

FIG. 6I

30781 AATCTTCTCT TCTCTTTTAC AGTATGGTGA TTGAACTATG ATTCCTAGAC AATTCTTGAT
30841 CACTATTCTT ATCTGCCTCC TCCAAGTCTG TGCCACCCTC GCTCTGGTGG CCAACGCCAG
30901 TCCAGACTGT ATTGGGCCCT TCGCCTCCTA CGTGCTCTTT GCCTTCATCA CCTGCATCTG
30961 CTGTTGTAGC ATAGTCTGCC TGCTTATCAC CTTCTTCCAG TTCATTGACT GGATCTTTGT
31021 GCGCATCGCC TACCTGCGCC ACCACCCCCA GTACCGCGAC CAGCGAGTGG CGCGACTGCT
31081 CAGGCTCCTC TGATAAGCAT GCGGGCTCTG CTACTTCTCG CGCTTCTGCT GTTAGTGCTC
31141 CCCCGTCCCG TCGACCCCCG GTCCCCCGAG GAGGTCCGCA AATGCAAATT CCAAGAACCC
31201 TGGAAATTCC TCAAATGCTA CCGCCAAAAA TCAGACATGC ATCCCAGCTG GATCATGATC
31261 ATTGGGATCG TGAACATTCT GGCCTGCACC CTCATCTCCT TTGTGATTTA CCCCTGCTTT
31321 GACTTTGGTT GGAACTCGCC AGAGGCACTC TATCTCCCGC CTGAGCCTGA CACACCACCA
31381 CAGCAGCAAC CTCAGGCACA CGCACTACCA CCACCACAGC CTAGGCCACA ATACATGCCC
31441 ATATTAGACT ATGAGGCCGA GCCACAGCGA CCCATGCTCC CCGCTATTAG TTACTTCAAT
31501 CTAACCGGCG GAGATGACTG ACCCACTGGC CAACAACAAC GTCAACGACC TTCTCCTGGA
31561 CATGGACGGC CGCGCCTCGG AGCAGCGACT CGCCCAACTC CGCATCCGCC AGCAGCAGGA
31621 GAGAGCCGTC AAGGAGCTGC AGGATGCGGT GGCCATCCAC CAGTGCAAGA AAGGCATCTT
31681 CTGCCTGGTG AAGCAGGCCA AGATCTCCTA CGAGGTCACC CAGACCGACC ATCGCCTCTC
31741 CTACGAGCTC CTGCAGCAGC GCCAGAAGTT CACCTGCCTG GTCGGAGTCA ACCCCATCGT
31801 CATCACCCAG CAGTCGGGCG ATACCAAGGG GTGCATCCAC TGCTCCTGCG ACTCCCCCGA
31861 GTGCGTTCAC ACCATGATCA AGACCCTCTG CGGCCTCCGC GACCTCCTCC CCATGAACTA
31921 ATCACCCCCT TATCCAGTGA AATAAAGATC ATATTGATGA TGATTTAAAT AAAAAAATAA
31981 TCATTTGATT TGAAATAAAG ATACAATCAT ATTGATGATT TGAGTTTAAC AAAAATAAAG
32041 AATCACTTAC TTGAAATCTG ATACCAGGTC TCTGTCCATG TTTTCTGCCA ACACCACCTC
32101 ACTCCCCTCT TCCCAGCTCT GGTACTGCAG GCCCCGGCGG GCTGCAAACT TCCTCCACAC
32161 GCTGAAGGGG ATGTCAAATT CCTCCTGTCC CTCAATCTTC ATTTTCTCTT CTATCAGATG
32221 TCCAAAAAGC GCGCGCGGGT GGATGATGAC TTCGACCCCG TGTACCCCTA CGATGCAGAC
32281 AACGCACCGA CTGTGCCCTT CATCAACCCT CCCTTCGTCT CTTCAGATGG ATTCCAAGAA
32341 AAGCCCCTGG GGGTGTTGTC CCTGCGACTG GCCGATCCCG TCACCACCAA GAACGGGGCT
32401 GTCACCCTCA AGCTGGGGGA GGGGGTGGAC CTCGACGACT CGGGAAAACT CATCTCCAAA
32461 AATGCCACCA AGGCCACTGC CCCTCTCAGT ATTTCCAACA ACACCATTTC CCTTAACATG
32521 GATACCCCTC TTTACAACAA CAATGGAAAG CTAGGTATGA AGGTAACCGC ACCATTAAAG
32581 ATATTAGACA CAGATCTACT AAAAACACTT GTTGTTGCTT ATGGGCAGGG ATTAGGAACA
32641 AACACCAATG GTGCTCTTGT TGCCCAACTA GCATACCCAC TTGTTTTTAA TACCGCTAGC
32701 AAAATTGCCC TTAATTTAGG CAATGGACCA TTAAAAGTGG ATGCAAATAG ACTGAACATT
32761 AATTGCAAAA GAGGTATCTA TGTCACTACC ACAAAAGATG CACTGGAGAT TAATATCAGT
32821 TGGGCAAATG CTATGACATT TATAGGAAAT GCCATTGGTG TCAATATTGA CACAAAAAAA
32881 GGCCTACAGT TCGGCACTTC AAGCACTGAA ACAGATGTTA AAAATGCTTT TCCACTCCAA
32941 GTAAAACTTG GAGCTGGTCT TACATTTGAC AGCACAGGTG CCATTGTTGC TTGGAACAAA
33001 GAAGATGACA AACTTACACT GTGGACCACA GCCGATCCAT CTCCAAACTG TCACATATAT
33061 TCTGCAAAGG ATGCTAAGCT TACACTCTGC TTGACAAAGT GTGGTAGTCA GATACTGGGC
33121 ACTGTTTCTC TCATAGCTGT TGATACTGGT AGCTTAAATC CAATAACAGG AAAAGTAACC
33181 ACTGCTCTTG TTTCACTTAA ATTCGATGCC AATGGAGTTT TGCAAGCCAG TTCAACACTA
33241 GATAAAGAAT ATTGGAATTT CAGAAAAGGA GATGTGACAC CTGCTGACCC CTACACTAAT
33301 GCTATAGGCT TTATGCCCAA CCTTAATGCA TACCCAAAAA ACACAAACGC AGCTGCAAAA
33361 AGTCACATTG TTGGAAAAGT ATACCTACAT GGGGATGAAA GCAAGCCACT AGACTTGATA
33421 ATTACATTTA ATGAAACCAG TGATGAATCC TGTACTTATT GCATTAACTT TCAGTGGCAG
33481 TGGGGAACTG ACCAATATAA AGATGAAACA CTTGCAGTCA GTTCATTCAC CTTCTCATAC
33541 ATTGCTAAAG AATAACATCC ACCCTGCATG CCAACCCATT TCCCTCTATC TATACATGGA
33601 AAACTCTGAA GCAGAAAAAA TAAAGTTCAA GTGTTTTATT GATTCAACAG TTTTTACAGA
33661 ATTCGAGTAG TTATTTTCCC TCCACCCTCC CAACTCATGG AATACACCAT CCTCTCCCCA
33721 CGCACAGCCT TAAACATCTG AATGCCATTG GTAATGGACA TGGTTTTGGC CTCCACATTC
33781 CACACAGTTT CAGAGCGAGC CAGTCTCGGG TCGGTCAGGG AGATGAAACC CTCCGGGCAC
33841 TCCTGCATCT GCACCTCACA GTTCAACAGC TGAGGGCTGT CCTCGGTGGT CGGGATCACA
33901 GTTATCTGGA AGAAGAGCGA TGAGAGTCAT AATCCGCGAA CGGGATCGGG CGGTTGTGGC
33961 GCATCAGGCC CCGCAGCAGT CGCTGTCTGC GCCGCTCCGT CAAGCTGCTG CTCAAGGGGT
34021 CCGGGTCCAG GGACTCCCCG CGCATGATGC CGATGGCCCT GAGCATCAGT CGCCTGGTGC
34081 GGCGGGCGCA GCAGCGGATG CGGATCTCAC TCAGGTCGGA ACAGTACGTG CAGCACAGCA
34141 CTACCAAGTT GTTCAACAGT CCATAGTTCA ACGTGCTCCA GCCAAAACTC ATCTGTGGAA

FIG. 6J

```
34201  CTATGCTGCC CACATGTCCA TCGTACCAGA TCCTGATGTA AATCAGGTGG CGCCCCCTCC
34261  AGAACACACT GCCCATGTAC ATGATCTCCT TGGGCATGTG CAGGTTCACC ACCTCCCGGT
34321  ACCACATCAC CCGCTGGTTG AACATGCAGC CCCGGATGAT CCTGCGGAAC CACAGGGCCA
34381  GCACCGCCCC GCCCGCCATG CAGCGCAGGG ACCCCGGGTC CTGGCAATGG CAGTGGATGA
34441  TCCACCGCTC GTACCCGTGG ATCATCTGGG AGCTGAACAA GTCTATGTTG GCACAGCACA
34501  GGCACACGCT CATGCATCTC TTCAGCACTC TCAGCTCCTC GGGGGTCAAA ACCATATCCC
34561  AGGGTACGGG GAACTCTTGC AGGACAGCGA ACCCCGCAGA ACAGGGCAAA CCTCGCACAG
34621  AACTTACATT GTGCATGGAC AGGGTATCGC AATCAGGCAG CACCGGGTGA TCCTCCACCA
34681  GGGAAGCGCG GGTCTCGATT TCCTCACAGC GTGGTAAGGG GGCCGGTCGA TACGGGTGAT
34741  GGCGGGACGC GGCTGATCGT GTTCGCGATC GTGTCATGAT GCAGTTGCTT TCGGACATTT
34801  TCGTACTTGC TATAGCAGAA CCTGGTCCGG GCGCTGCACA CCGATCGCCG GCGGCGGTCT
34861  CGGCGCTTGG AACGCTCCGT GTTGAAATTG TAAAACAGCC ACTCTCTCAG ACCGTGCAGC
34921  AGATCTAGGG CCTCAGGAGT GATGAAGATC CCATCATGCC TGATGGCTCT GATCACATCG
34981  ACCACCGTGG AATGGGCCAG ACCCAGCCAG ATGATGCAAT TTTGTTGGGT TTCGGTGACG
35041  GCGGGGGAGG GAAGAACAGG AAGAACCATG ATTAACTTTA ATCCAAACGG TCTCGGAGCA
35101  CTTCAAAATG AAGGTCGCGG AGATGGCACC TCTCGCCCCC GCTGTGTTGG TGGAAAATAA
35161  CAGCCAGGTC AAAGGTGATA CGGTTCTCGA GATGTTCCAC GGTGGCTTCC AGCAAAGCCT
35221  CCACGCGCAC ATCCAGAAAC AAGACAATAG CGAAAGCGGG AGGGTTCTCT AATTCCTCAA
35281  TCATCATGTT ACACTCCTGC ACCATCCCCA GATAATTTTC ATTTTTCCAG CCTTGAATGA
35341  TTCGAACTAG TTCCTGAGGT AAATCCAAGC CAGCCATGAT AAAGAGCTCG CGCAGAGCGC
35401  CCTCCACCGG CATTCTTAAG CACACCCTCA TAATTCCAAG ATATTCTGCT CCTGGTTCAC
35461  CTGCAGCAGA TTGACAAGCG GAATATCAAA CTCTCTGCCG CGATCCCTAA GCTCCTCCCT
35521  CAGCAATAAC TGTAAGTACT CTCTCATATC CTCTCCGAAA TTTTTAGCCA TAGGACCGCC
35581  AGGAATAAGA TTAGGGCAAG CCACAGTACA GATAAACCGA AGTCCTCCCC AGTGAGCATT
35641  GCCAAATGCA AGACTGCTAT AAGCATGCTG GCTAGACCCG GTGATATCTT CCAGATAATT
35701  GGACAGAAAA TCGCCCAGGC AATTTTTAAG AAAATCAACA AAAGAAAAAT CCTCCAGGTG
35761  CACGTTTAGA GCCTCGGGAA CAACGATGGA GTAAATGCAA GCGGTGCGTT CCAGCATGGT
35821  TAGTTAGCTG ATCTGTAGAA AAAACAAAAA TGAACATTAA ACCATGCTAG CCTGGCGAAC
35881  AGGTGGGTAA ATCGTTCTTT CCAGCACCAG GCAGGCCACG GGGTCTCCGG CGCGACCCTC
35941  GTAAAAATTG TCGCTATGAT TGAAAACCAT CACAGAGAGA CGTTCCCGGT GGCCGGCGTG
36001  AATGATTCGA CAAGACGAAT ACACCCCCGG AACATTGGCG TCCGCGAGTG AAAAAAAGCG
36061  CCCGAGGAAG CAATAAGGCA CTACAATGCT CAGTCTCAAG TCCAGCAAAG CGATGCCATG
36121  CGGATGAAGC ACAAAATTCT CAGGTGCGTA CAAAATGTAA TTACTCCCCT CCTGCACAGG
36181  CAGCAAAGCC CCCGATCCCT CCAGGTACAC ATACAAAGCC TCAGCGTCCA TAGCTTACCG
36241  AGCAGCAGCG GCACACAACA GGCGCAAAAG TCAGAGAAAG GCTGAGAGCT CTAACCTGTC
36301  CACCCGCTCT CTGCTCAATA TATAGCCCAG ATCTACACTG ACGTAAAGGC CAAAGTCTAA
36361  AAATACCCGC CAAATAATCA CACACGCCCA GCACACGCCC AGAAACCGGT GACACACTCA
36421  GAAAAATACG CGCACTTCCT CAAACGCCCA AACTGCCGTC ATTTCCGGGT TCCCACGCTA
36481  CGTCATCAAA ATTCAACTTT CAAATTCCGT CGACCGTTAA AAACGTCACC CGCCCCGCCC
36541  CTAACGGTCG CCGCTCCCGC AGCCAATCAG CGCCCCGCAT CCCCAAATTC AAACGGCTCA
36601  TTTGCATATT AACGCGCACC AAAAGTTTGA GGTATATTAT TGATGATG  (SEQ ID NO: 2)
```

FIG. 6K

```
   1 CATCATCAAT AATATACCTC AAACTTTTGG TGCGCGTTAA TATGCAAATG AGCTGTTTGA
  61 ATTTGGGGAG GGAGGAAGGT GATTGGCTGC GGGAGCGGCG ACCGTTAGGG GCGGGGCGGG
 121 TGACGTTTTG ATGACGTGGC TATGAGGCGG AGCCGGTTTG CAAGTTCTCG TGGGAAAAGT
 181 GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCTCTGACA
 241 GGAAATGAGG TGTTTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTCGC GCGAAAACTG
 301 AATGAGGAAG TGAAAATCTG AGTAATTTCG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
 361 GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTTCGATT ACCGTATTTT TCACCTAAAT
 421 TTCCGCGTAC GGTGTCAAAG TCCGGTGTTT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
 481 TTTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTTCT
 541 CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
 601 GTAATGTTTT CCTGGCTACT GGGAACGAGA TTCTGGAATT GGTGGTGGAC GCCATGATGG
 661 GTGACGACCC TCCAGAGCCC CCTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTTGTATG
 721 ATCTGGAGGT GGATGTGCCC GAGAGCGACC CTAACGAGGA GGCGGTGAAT GATTTGTTTA
 781 GCGATGCCGC GCTGCTGGCT GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
 841 CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCCGAG CTTAAAGGGG
 901 AAGAGCTCGA CCTGCGCTGC TATGAGGAAT GCTTGCCTCC GAGCGATGAT GAGGAGGACG
 961 AGGAGGCGAT TCGAGCTGCG GTGAACCAGG GAGTGAAAAC TGCGGGCGAG AGCTTTAGCC
1021 TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTCAT CGCATGAATA
1081 CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTGTGTTT
1141 ACAGTAAGTG TGATTAACTT TAGTTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201 ATTTATGTAT ATGTTTTTTT ATGTGTAGGT CCCGTCTCTG ACGTAGATGA GACCCCCACT
1261 TCAGAGTGCA TTTCATCACC CCCAGAAATT GGCGAGGAAC CGCCCGAAGA TATTATTCAT
1321 AGACCAGTTG CAGTGAGAGT CACCGGGCGG AGAGCAGCTG TGGAGAGTTT GGATGACTTG
1381 CTACAGGGTG GGGATGAACC TTTGGACTTG TGTACCCGGA AACGCCCCAG GCACTAAGTG
1441 CCACACATGT GTGTTTACTT AAGGTGATGT CAGTATTTAT AGGGTGTGGA GTGCAATAAA
1501 ATCCGTGTTG ACTTTAAGTG CGTGTTTTAT GACTCAGGGG TGGGGACTGT GGGTATATAA
1561 GCAGGTGCAG ACCTGTGTGG TCAGTTCAGA GCAGGACTCA TGGAGATCTG GACTGTCTTG
1621 GAAGACTTTC ACCAGACTAG ACAGTTGCTA GAGAACTCAT CGGAGGGAGT CTCTTACCTG
1681 TGGAGATTCT GCTTCGGTGG GCCTCTAGCT AAGCTAGTCT ATAGGGCCAA ACAGGATTAT
1741 AAGGAACAAT TTGAGGATAT TTTGAGAGAG TGTCCTGGTA TTTTTGACTC TCTCAACTTG
1801 GGCCATCAGT CTCACTTTAA CCAGAGTATT CTGAGAGCCC TTGACTTTTC TACTCCTGGC
1861 AGAACTACCG CCGCGGTAGC CTTTTTTGCC TTTATTCTTG ACAAATGGAG TCAAGAAACC
1921 CATTTCAGCA GGGATTACCG TCTGGACTGC TTAGCAGTAG CTTTGTGGAG AACATGGAGG
1981 TGCCAGCGCC TGAATGCAAT CTCCGGCTAC TTGCCAGTAC AGCCGGTAGA CACGCTGAGG
2041 ATCCTGAGTC TCCAGTCACC CCAGGAACAC CAACGCCGCC AGCAGCCGCA GCAGGAGCAG
2101 CAGCAAGAGG AGGACCGAGA AGAGAACCCG AGAGCCGGTC TGGACCCTCC GGTGGCGGAG
2161 GAGGAGGAGT AGCTGACTTG TTTCCCGAGC TGCGCCGGGT GCTGACTAGG TCTTCCAGTG
2221 GACGGGAGAG GGGGATTAAG CGGGAGAGGC ATGAGGAGAC TAGCCACAGA ACTGAACTGA
2281 CTGTCAGTCT GATGAGCCGC AGGCGCCCAG AATCGGTGTG GTGGCATGAG GTGCAGTCGC
2341 AGGGGATAGA TGAGGTCTCG GTGATGCATG AGAAATATTC CCTAGAACAA GTCAAGACTT
2401 GTTGGTTGGA GCCCGAGGAT GATTGGGAGG TAGCCATCAG GAATTATGCC AAGCTGGCTC
2461 TGAAGCCAGA CAAGAAGTAC AAGATTACCA AACTGATTAA TATCAGAAAT TCCTGCTACA
2521 TTTCAGGGAA TGGGGCCGAG GTGGAGATCA GTACCCAGGA GAGGGTGGCC TTCAGATGTT
2581 GTATGATGAA TATGTACCCG GGGGTGGTGG GCATGGAGGG AGTCACCTTT ATGAACACGA
2641 GGTTCAGGGG TGATGGGTAT AATGGGGTGG TCTTTATGGC CAACACCAAG CTGACAGTGC
2701 ACGGATGCTC CTTCTTTGGC TTCAATAACA TGTGCATCGA GGCCTGGGGC AGTGTTTCAG
2761 TGAGGGGATG CAGCTTTTCA GCCAACTGGA TGGGGGTCGT GGGCAGAACC AAGAGCAAGG
2821 TGTCAGTGAA GAAATGCCTG TTCGAGAGGT GCCACCTGGG GGTGATGAGC GAGGGCGAAG
2881 CCAAAGTCAA ACACTGCGCC TCTACCGAGA CGGGCTGCTT TGTGCTGATC AAGGGCAATG
2941 CCCAAGTCAA GCATAACATG ATCTGTGGGG CCTCGGATGA GCGCGGCTAC CAGATGCTGA
3001 CCTGCGCCGG TGGGAACAGC CATATGCTGG CCACCGTGCA TGTGGCCTCG CACCCCCGCA
3061 AGACATGGCC CGAGTTCGAG CACAACGTCA TGACCCGCTG CAATGTGCAC CTGGGCTCCC
3121 GCCGAGGCAT GTTCATGCCC TACCAGTGCA ACATGCAATT TGTGAAGGTG CTGCTGGAGC
3181 CCGATGCCAT GTCCAGAGTG AGCCTGACGG GGGTGTTTGA CATGAATGTG GAGCTGTGGA
3241 AAATTCTGAG ATATGATGAA TCCAAGACCA GGTGCCGGGC CTGCGAATGC GGAGGCAAGC
3301 ACGCCAGGCT TCAGCCCGTG TGTGTGGAGG TGACGGAGGA CCTGCGACCC GATCATTTGG
3361 TGTTGTCCTG CAACGGGACG GAGTTCGGCT CCAGCGGGGA AGAATCTGAC TAGAGTGAGT
3421 AGTGTTTGGG GCTGGGTGTG AGCCTGCATG AGGGGCAGAA TGACTAAAAT CTGTGGTTTT
```

FIG. 7A

```
3481 CTGTGTGTTG CAGCAGCATG AGCGGAAGCG CCTCCTTTGA GGGAGGGGTA TTCAGCCCTT
3541 ATCTGACGGG GCGTCTCCCC TCCTGGGCGG GAGTGTGTCA GAATGTTATG GNATCCACGG
3601 TGGACGGCCG GCCCGTGCAG CCCGCGAACT CTTCAACCCT GACCTACGCG ACCCTGAGCT
3661 CCTCGTCCGT GGACGCAGCT GCCGCCGCAG CTGCTGCTTC CGCCGCCAGC GCCGTGCGCG
3721 GAATGGCCCT GGGCGCCGGC TACTACAGCT CTCTGGTGGC CAACTCGAGT TCCACCAATA
3781 ATCCCGCCAG CCTGAACGAG GAGAAGCTGC TGCTGCTGAT GGCCCAGCTC GAGGCCCTGA
3841 CCCAGCGCCT GGGCGAGCTG ACCCAGCAGG TGGCTCAGCT GCAGGCGGAG ACGCGGGCCG
3901 CGGTTGCCAC GGTGAAAACC AAATAAAAAA TGAATCAATA AATAAACGGA GACGGTTGTT
3961 GATTTTAACA CAGAGTCTTG AATCTTTATT TGATTTTTCG CGCGCGGTAG GCCCTGGACC
4021 ACCGGTCTCG ATCATTGAGC ACCCGGTGGA TCTTTTCCAG GACCCGGTAG AGGTGGGCTT
4081 GGATGTTGAG GTACATGGGC ATGAGCCCGT CCCGGGGGTG GAGGTAGCTC CATTGCAGGG
4141 CCTCGTGCTC GGGGATGGTG TTGTAAATCA CCCAGTCATA GCAGGGGCGC AGGGCGTGGT
4201 GCTGCACGAT GTCCTTGAGG AGGAGACTGA TGGCCACGGG CAGCCCCTTG GTGTAGGTGT
4261 TGACGAACCT GTTGAGCTGG GAGGGATGCA TGCGGGGGGA GATGAGATGC ATCTTGGCCT
4321 GGATCTTGAG ATTGGCGATG TTCCCGCCCA GATCCCGCCG GGGGTTCATG TTGTGCAGGA
4381 CCACCAGCAC GGTGTATCCG GTGCACTTGG GGAATTTGTC ATGCAACTTG GAAGGGAAGG
4441 CGTGAAAGAA TTTGGAGACG CCCTTGTGAC CGCCCAGGTT TTCCATGCAC TCATCCATGA
4501 TGATGGCGAT GGGCCCGTGG GCGGCGGCCT GGGCAAAGAC GTTTCGGGGG TCGGACACAT
4561 CGTAGTTGTG GTCCTGGGTG AGCTCGTCAT AGGCCATTTT AATGAATTTG GGGCGGAGGG
4621 TGCCCGACTG GGGGACGAAG GTGCCCTCGA TCCCGGGGGC GTAGTTGCCC TCGCAGATCT
4681 GCATCTCCCA GGCCTTGAGC TCGGAGGGGG GGATCATGTC CACCTGCGGG GCGATGAAAA
4741 AAACGGTTTC CGGGGCGGGG GAGATGAGCT GGGCCGAAAG CAGGTTCCGG AGCAGCTGGG
4801 ACTTGCCGCA ACCGGTGGGG CCGTAGATGA CCCCGATGAC CGGCTGCAGG TGGTAGTTGA
4861 GGGAGAGACA GCTGCCGTCC TCGCGGAGGA GGGGGGCCAC CTCGTTCATC ATCTCGCGCA
4921 CATGCATGTT CTCGCGCACG AGTTCCGCCA GGAGGCGCTC GCCCCCCAGC GAGAGGAGCT
4981 CTTGCAGCGA GGCGAAGTTT TTCAGCGGCT TGAGTCCGTC GGCCATGGGC ATTTTGGAGA
5041 GGGTCTGTTG CAAGAGTTCC AGACGGTCCC AGAGCTCGGT GATGTGCTCT AGGGCATCTC
5101 GATCCAGCAG ACCTCCTCGT TTCGCGGGTT GGGGCGACTG CGGGAGTAGG GCACCAGGCG
5161 ATGGGCGTCC AGCGAGGCCA GGGTCCGGTC CTTCCAGGGC CGCAGGGTCC GCGTCAGCGT
5221 GGTCTCCGTC ACGGTGAAGG GGTGCGCGCC GGGCTGGGCG CTTGCGAGGG TGCGCTTCAG
5281 GCTCATCCGG CTGGTCGAGA ACCGCTCCCG GTCGGCGCCC TGCGCGTCGG CCAGGTAGCA
5341 ATTGAGCATG AGTTCGTAGT TGAGCGCCTC GGCCGCGTGG CCCTTGGCGC GGAGCTTACC
5401 TTTGGAAGTG TGTCCGCAGA CGGGACAGAG GAGGGACTTG AGGGCGTAGA GCTTGGGGGC
5461 GAGGAAGACG GACTCGGGGG CGTAGGCGTC CGCGCCGCAG CTGGCGCAGA CGGTCTCGCA
5521 CTCCACGAGC CAGGTGAGGT CGGGGCGGTT GGGGTCAAAA ACGAGGTTTC CTCCGTGCTT
5581 TTTGATGCGT TTCTTACCTC TGGTCTCCAT GAGCTCGTGT CCCCGCTGGG TGACAAAGAG
5641 GCTGTCCGTG TCCCCGTAGA CCGACTTTAT GGGCCGGTCC TCGAGCGGGG TGCCGCGGTC
5701 CTCGTCGTAG AGGAACCCCG CCCACTCCGA GACGAAGGCC CGGGTCCAGG CCAGCACGAA
5761 GGAGGCCACG TGGGAGGGGT AGCGGTCGTT GTCCACCAGC GGGTCCACCT TCTCCAGGGT
5821 ATGCAAGCAC ATGTCCCCCT CGTCCACATC CAGGAAGGTG ATTGGCTTGT AAGTGTAGGC
5881 CACGTGACCG GGGGTCCCGG CCGGGGGGGT ATAAAAGGGG GCGGGCCCCT GCTCGTCCTC
5941 ACTGTCTTCC GGATCGCTGT CCAGGAGCGC CAGCTGTTGG GGTAGGTATT CCCTCTCGAA
6001 GGCGGGCATG ACCTCGGCAC TCAGGTTGTC AGTTTCTAGA AACGAGGAGG ATTTGATATT
6061 GACGGTGCCG TTGGAGACGC CTTTCATGAG CCCCTCGTCC ATTTGGTCAG AAAAGACGAT
6121 CTTTTTGTTG TCGAGCTTGG TGGCGAAGGA GCCGTAGAGG GCGTTGGAGA GCAGCTTGGC
6181 GATGGAGCGC ATGGTCTGGT TCTTTTCCTT GTCGGCGCGC TCCTTGGCGG CGATGTTGAG
6241 CTGCACGTAC TCGCGCGCCA CGCACTTCCA TTCGGGGAAG ACGGTGGTGA GCTCGTCGGG
6301 CACGATTCTG ACCCGCCAGC CGCGGTTGTG CAGGGTGATG AGGTCCACGC TGGTGGCCAC
6361 CTCGCCGCGC AGGGGCTCGT TGGTCCAGCA GAGGCGCCCG CCCTTGCGCG AGCAGAAGGG
6421 GGGCAGCGGG TCCAGCATGA GCTCGTCGGG GGGGTCGGCG TCCACGGTGA AGATGCCGGG
6481 CAGGAGCTCG GGGTCGAAGT AGCTGATGCA GGTGCCCAGA TTGTCCAGCG CCGCTTGCCA
6541 GTCGCGCACG GCCAGCGCGC GCTCGTAGGG GCTGAGGGGC GTGCCCCAGG GCATGGGGTG
6601 CGTGAGCGCG GAGGCGTACA TGCCGCAGAT GTCGTAGACG TAGAGGGGCT CCTCGAGGAC
6661 GCCGATGTAG GTGGGGTAGC AGCGCCCCCC GCGGATGCTG GCGCGCACGT AGTCGTACAG
6721 CTCGTGCGAG GGCGCGAGGA GCCCCGTGCC GAGGTTGGAG CGTTGCGGCT TTTCGGCGCG
6781 GTAGACGATC TGGCGGAAGA TGGCGTGGGA GTTGGAGGAG ATGGTGGGCC TTTGGAAGAT
6841 GTTGAAGTGG GCGTGGGGCA GGCCGACCGA GTCCCTGATG AAGTGGGCGT AGGAGTCCTG
6901 CAGCTTGGCG ACGAGCTCGG CGGTGACGAG GACGTCCAGG GCGCAGTAGT CGAGGGTCTC
```

FIG. 7B

```
 6961 TTGGATGATG TCATACTTGA GCTGGCCCTT CTGCTTCCAC AGCTCGCGGT TGAGAAGGAA
 7021 CTCTTCGCGG TCCTTCCAGT ACTCTTCGAG GGGGAACCCG TCCTGATCGG CACGGTAAGA
 7081 GCCCACCATG TAGAACTGGT TGACGGCCTT GTAGGCGCAG CAGCCCTTCT CCACGGGGAG
 7141 GGCGTAAGCT TGCGCGGCCT TGCGCAGGGA GGTGTGGGTG AGGGCGAAGG TGTCGCGCAC
 7201 CATGACCTTG AGGAACTGGT GCTTGAAGTC GAGGTCGTCG CAGCCGCCCT GCTCCCAGAG
 7261 TTGGAAGTCC GTGCGCTTCT TGTAGGCGGG GTTAGGCAAA GCGAAAGTAA CATCGTTGAA
 7321 GAGGATCTTG CCCGCGCGGG GCATGAAGTT GCGAGTGATG CGGAAAGGCT GGGGCACCTC
 7381 GGCCCGGTTG TTGATGACCT GGGCGGCGAG GACGATCTCG TCGAAGCCGT TGATGTTGTG
 7441 CCCGACGATG TAGAGTTCCA CGAATCGCGG GCGGCCCTTG ACGTGGGGCA GCTTCTTGAG
 7501 CTCGTCGTAG GTGAGCTCGG CGGGGTCGCT GAGCCCGTGC TGCTCGAGGG CCCAGTCGGC
 7561 GACGTGGGGG TTGGCGCTGA GGAAGGAAGT CCAGAGATCC ACGGCCAGGG CGGTCTGCAA
 7621 GCGGTCCCGG TACTGACGGA ACTGTTGGCC CACGGCCATT TTTTCGGGGG TGACGCAGTA
 7681 GAAGGTGCGG GGGTCGCCGT GCCANCGGTC CCACTTGAGC TGGAGGGCGA GGTCGTGGGC
 7741 GAGCTCGACG AGCGGCGGGT CCCCGGAGAG TTTCATGACC AGCATGAAGG GGACGAGCTG
 7801 CTTGCCGAAG GACCCCATCC AGGTGTAGGT TTCCACATCG TAGGTGAGGA AGAGCCTTTC
 7861 GGTGCGAGGA TGCGAGCCGA TGGGGAAGAA CTGGATCTCC TGCCACCAGT TGGAGGAATG
 7921 GCTGTTGATG TGATGGAAGT AGAAATGCCG ACGGCGCGCC GAGCACTCGT GCTTGTGTTT
 7981 ATACAAGCGT CCGCAGTGCT CGCAACGCTG CACGGGATGC ACGTGCTGCA CGAGCTGTAC
 8041 CTGGGTTCCT TTGGCGAGGA ATTTCAGTGG GCAGTGGAGC GCTGGCGGCT GCATCTCGTG
 8101 CTGTACTACG TCTTGGCCAT CGGCGTGGCC ATCGTCTGCC TCGATGGTGG TCATGCTGAC
 8161 GAGCCCGCGC GGGAGGCAGG TCCAGACCTC GGCTCGGACG GGTCGGAGAG CGAGGACGAG
 8221 GGCGCGCAGG CCGGAGCTGT CCAGGGTCCT GAGACGCTGC GGAGTCAGGT CAGTGGGCAG
 8281 CGGCGGCGCG CGGTTGACTT GCAGGAGCTT TTCCAGGGCG CGCGGGAGGT CCAGATGGTA
 8341 CTTGATCTCC ACGGCGCCGT TGGTGGCTAC GTCCACGGCT TGCAGGGTGC CGTGCCCCTG
 8401 GGGCGCCACC ACCGTGCCCC GTTTCTTCTT GGGCGCTGCT TCCATGTCGG TCAGAAGCGG
 8461 CGGCGAGGAC GCGCGCCGGG CGGCAGGGGC GGCTCGGGGC CCGGAGGCAG GGGCGGCAGG
 8521 GGCACGTCGG CGCCGCGCGC GGGCAGGTTG TGGTACTGCG CCCGGAGAAG ACTGGCGTGA
 8581 GCGACGACGC GACGGTTGAC GTCCTGGATC TGACGCCTCT GGGTGAAGGC CACGGGACCC
 8641 GTGAGTTTGA ACCTGAAAGA GAGTTCGACA GAATCAATCT CGGTATCGTT GACGGCGGCC
 8701 TGCCGCAGGA TCTCTTGCAC GTCGCCCGAG TTGTCCTGGT AGGCGATCTC GGTCATGAAC
 8761 TGCTCGATCT CCTCCTCCTG AAGGTCTCCG CGGCCGGCGC GCTCGACGGT GGCCGCGAGG
 8821 TCGTTGGAGA TGCGGCCCAT GAGCTGCGAG AAGGCGTTCA TGCCGGCCTC GTTCCAGACG
 8881 CGGCTGTAGA CCACGGCTCC GTCGGGGTCG CGCGCGCGCA TGACCACCTG GGCGAGGTTG
 8941 AGCTCGACGT GGCGCGTGAA GACCGCGTAG TTGCAGAGGC GCTGGTAGAG GTAGTTGAGC
 9001 GTGGTGGCGA TGTGCTCGGT GACGAAGAAG TACATGATCC AGCGGCGGAG CGGCATCTCG
 9061 CTGACGTCGC CCAGGGCTTC CAAGCGTTCC ATGGCCTCGT AGAAGTCCAC GGCGAAGTTG
 9121 AAAAACTGGG AGTTGCGCGC CGAGACGGTC AACTCCTCCT CCAGAAGACG GATGAGCTCG
 9181 GCGATGGTGG CGCGCACCTC GCGCTCGAAG GCCCCGGGGG GCTCCTCTTC CATCTCCTCC
 9241 TCTTCCTCCT CCACTAACAT CTCTTCTACT TCCTCCTCAG GAGGCGGTGG CGGGGGAGGG
 9301 GCCCTGCGTC GCCGGCGGCG CACGGGCAGA CGGTCGATGA AGCGCTCGAT GGTCTCCCCG
 9361 CGCCGGCGAC GCATGGTCTC GGTGACGGCG CGCCCGTCCT CGCGGGGCCG CAGCATGAAG
 9421 ACGCCGCCGC GCATCTCCAG GTGGCCGCCG GGGGGGTCTC CGTTGGGCAG GGAGAGGGCG
 9481 CTGACGATGC ATCTTATCAA TTGACCCGTA GGGACTCCGC GCAAGGACCT GAGCGTCTCG
 9541 AGATCCACGG GATCCGAAAA CCGCTGAACG AAGGCTTCGA GCCAGTCGCA GTCGCAAGGT
 9601 AGGCTGAGCC CGGTTTCTTG TTCTTCGGGT ATTTGGTCGG GAGGCGGCGG GCGATGCTGC
 9661 TGGTGATGAA GTTGAAGTAG GCGGTCCTGA GACGGCGGAT GGTGGCGAGG AGCACCAGGT
 9721 CCTTGGGCCC GGCTTGCTGG ATGCGCAGAC GGTCGGCCAT GCCCCAGGCG TGGTCCTGAC
 9781 ACCTGGCGAG GTCCTTGTAG TAGTCCTGCA TGAGCCGCTC CACGGGCACC TCCTCCTCGC
 9841 CCGCGCGGCC GTGCATGCGC GTGAGCCCGA ACCCGCGCTG CGGCTGGACG AGCGCCAGGT
 9901 CGGCGACGAC GCGCTCGGTG AGGATGGCCT GCTGGATCTG GGTGAGGGTG GTCTGGAAGT
 9961 CGTCGAAGTC GACGAAGCGG TGGTAGGCTC CGGTGTTGAT GGTGTAGGAG CAGTTGGCCA
10021 TGACGGACCA GTTGACGGTC TGGTGGCCGG GTCGCACGAG CTCGTGGTAC TTGAGGCGCG
10081 AGTAGGCGCG CGTGTCGAAG ATGTAGTCGT TGCAGGCGCG CACGAGGTAC TGGTATCCGA
10141 CGAGGAAGTG CGGCGGCGGC TGGCGGTAGA GCGGCCATCG CTCGGTGGCG GGGGCGCCGG
10201 GCGCGAGGTC CTCGAGCATG AGGCGGTGGT AGCCGTAGAT GTACCTGGAC ATCCAGGTGA
10261 TGCCGGCGGC GGTGGTGGAG GCGCGCGGGA ACTCGCGGAC GCGGTTCCAG ATGTTGCGCA
10321 GCGGCAGGAA GTAGTTCATG GTGGCCGCGG TCTGGCCCGT GAGGCGCGCG CAGTCGTGGA
10381 TGCTCTAGAC ATACGGGCAA AAACGAAAGC GGTCAGCGGC TCGACTCCGT GGCCTGGAGG
```

FIG. 7C

```
10441 CTAAGCGAAC GGGTTGGGCT GCGCGTGTAC CCCGGTTCGA ATCTCGAATC AGGCTGGAGC
10501 CGCAGCTAAC GTGGTACTGG CACTCCCGTC TCGACCCAAG CCTGCTAACG AAACCTCCAG
10561 GATACGGAGG CGGGTCGTTT TTTGGCCTTG GTCGCTGGTC ATGAAAAACT AGTAAGCGCG
10621 GAAAGCGGCC GCCCGCGATG GCTCGCTGCC GTAGTCTGGA GAAAGAATCG CCAGGGTTGC
10681 GTTGCGGTGT GCCCCGGTTC GAGCCTCAGC GCTCGGCGCC GGCCGGATTC CGCGGCTAAC
10741 GTGGGCGTGG CTGCCCCGTC GTTTCCAAGA CCCCTTAGCC AGCCGACTTC TCCAGTTACG
10801 GAGCGAGCCC CTCTTTTTTT TTCTTGTGTT TTTGCCAGAT GCATCCCGTA CTGCGGCAGA
10861 TGCGCCCCCA CCCTCCACCA CAACCGCCCC TACCGCAGCA GCAGCAACAG CCGGCGCTTC
10921 TGCCCCCGCC CCAGCAGCAG CCAGCCACTA CCGCGGCGGC CGCCGTGAGC GGAGCCGGCG
10981 TTCAGTATGA CCTGGCCTTG GAAGAGGGCG AGGGGCTGGC GCGGCTGGGG GCGTCGTCGC
11041 CGGAGCGGCA CCCGCGCGTG CAGATGAAAA GGGACGCTCG CGAGGCCTAC GTGCCCAAGC
11101 AGAACCTGTT CAGAGACAGG AGCGGCGAGG AGCCCGAGGA GATGCGCGCC TCCCGCTTCC
11161 ACGCGGGGCG GGAGCTGCGG CGCGGCCTGA ACCGAAAGCG GGTGCTGAGG GACGAGGATT
11221 TCGAGGCGGA CGAGCTGACG GGGATCAGCC CCGTGCGCGC GCACGTGGTC GNGGNCAACC
11281 TGGTCACGGC GTACGAGCAG ACCGTGAAGG AGGAGAGCAA CTTCCAAAAA TCCTTCAACA
11341 ACCACGTGCG CACCTTGATC GCGCGCGAGG AGGTGACCCT GGGCCTGATG CACCTGTGGG
11401 ACCTGCTGGA GGCCATCGTG CAGAACCCCA CGAGCAAGCC GCTGACGGCG CAGCTGTTTC
11461 TGGTGGTGCA GCACAGTCGG GACAACGAGA CGTTCAGGGA GGCGCTGCTG AATATCACCG
11521 AGCCCGAGGG CCGCTGGCTC CTGGACCTGG TGAACATTTT GCAGAGCATC GTGGTGCAGG
11581 AGCGCGGGCT GCCGCTGTCC GAGAAGCTGG CGGCCATCAA CTTCTCGGTG CTGAGTCTGG
11641 GCAAGTACTA CGCTAGGAAG ATCTACAAGA CCCCGTACGT GCCCATAGAC AAGGAGGTGA
11701 AGATCGACGG GTTTTACATG CGCATGACCC TGAAAGTGCT GACCCTGAGC GACGATCTGG
11761 GGGTGTACCG CAACGACAGG ATGCACCGCG CGGTGAGCGC CAGCCGCCGG CGCGAGCTGA
11821 GCGACCAGGA GCTGATGCAC AGCCTGCAGC GGGCCCTGAC CGGGGCCGGG ACCGAGGGGG
11881 AGAGCTACTT TGACATGGGC GCGGACCTGC GCTGGCAGCC CAGCCGCCGG GCCTTGGAAG
11941 CTGCCGGCGG TTCCCCCTAC GTGGAGGAGG TGGACGATGA GGAGGAGGAG GGCGAGTACC
12001 TGGAAGACTG ATGGCGCGAC CGTATTTTTG CTAGATGCAG CAACAGCCAC CGCCGCCGCC
12061 TCCTGATCCC GCGATGCGGG CGGCGCTGCA GAGCCAGCCG TCCGGCATTA ACTCCTCGGA
12121 CGATTGGACC CAGGCCATGC AACGCATCAT GGCGCTGACG ACCCGCAATC CCGAAGCCTT
12181 TAGACAGCAG CCTCAGGCCA ACCGGCTCTC GGCCATCCTG GAGGCCGTGG TGCCCTCGCG
12241 CTCGAACCCC ACGCACGAGA AGGTGCTGGC CATCGTGAAC GCGCTGGTGG AGAACAAGGC
12301 CATCCGCGGT GACGAGGCCG GGCTGGTGTA CAACGCGCTG CTGGAGCGCG TGGCCCGCTA
12361 CAACAGCACC AACGTGCAGA CGAACCTGGA CCGCATGGTG ACCGACGTGC GCGAGGCGGT
12421 GTCGCAGCGC GAGCGGTTCC ACCGCGAGTC GAACCTGGGC TCCATGGTGG CGCTGAACGC
12481 CTTCCTGAGC ACGCAGCCCG CCAACGTGCC CCGGGGCCAG GAGGACTACA CCAACTTCAT
12541 CAGCGCGCTG CGGCTGATGG TGGCCGAGGT GCCCCAGAGC GAGGTGTACC AGTCGGGGCC
12601 GGACTACTTC TTCCAGACCA GTCGCCAGGG CTTGCAGACC GTGAACCTGA GCCAGGCTTT
12661 CAAGAACTTG CAGGGACTGT GGGGCGTGCA GGCCCCGGTC GGGGACCGCG CGACGGTGTC
12721 GAGCCTGCTG ACGCCGAACT CGCGCCTGCT GCTGCTGCTG GTGGCGCCCT TCACGGACAG
12781 CGGCAGCGTG AGCCGCGACT CGTACCTGGG CTACCTGCTT AACCTGTACC GCGAGGCCAT
12841 CGGACAGGCG CACGTGGACG AGCAGACCTA CCAGGAGATC ACCCACGTGA GCCGCGCGCT
12901 GGGCCAGGAG GACCCGGGCA ACCTGGAGGC CACCCTGAAC TTCCTGCTGA CCAACCGGTC
12961 GCAGAAGATC CCGCCCCAGT ACGCGCTGAG CACCGAGGAG GAGCGCATCC TGCGCTACGT
13021 GCAGCAGAGC GTGGGGCTGT TCCTGATGCA GGAGGGGGCC ACGCCCAGCG CGGCGCTCGA
13081 CATGACCGCG CGCAACATGG AGCCCAGCAT GTACGCCCGC AACCGCCCGT TCATCAATAA
13141 GCTGATGGAC TACTTGCATC GGGCGGCCGC CATGAACTCG GACTACTTTA CCAACGCCAT
13201 CTTGAACCCG CACTGGCTCC CGCCGCCCGG GTTCTACACG GGCGAGTACG ACATGCCCGA
13261 CCCCAACGAC GGGTTCCTGT GGGACGACGT GGACAGCAGC GTGTTCTCGC CGCGTCCAGG
13321 AACCAATGCC GTGTGGAAGA AAGAGGGCGG GGACCGGCGG CCGTCCTCGG CGCTGTCCGG
13381 TCGCGCGGGT GCTGCCGCGG CGGTGCCCGA GGCCGCCAGC CCCTTCCCGA GCCTGCCCTT
13441 TTCGCTGAAC AGCGTGCGCA GCAGCGAGCT GGGTCGGCTG ACGCGACCGC GCCTGCTGGG
13501 CGAGGAGGAG TACCTGAACG ACTCCTTGTT GAGGCCCGAG CGCGAGAAGA ACTTCCCCAA
13561 TAACGGGATA GAGAGCCTGG TGGACAAGAT GAGCCGCTGG AAGACGTACG CGCACGAGCA
13621 CAGGGACGAG CCCCGAGCTA GCAGCGCAGG CACCCGTAGA CGCCAGCGGC ACGACAGGCA
13681 GCGGGGACTG GTGTGGGACG ATGAGGATTC CGCCGACGAC AGCAGCGTGT TGGACTTGGG
13741 TGGGAGTGGT GGTAACCCGT TCGCTCACCT GCGCCCCCGT ATCGGGCGCC TGATGTAAGA
13801 ATCTGAAAAA ATAAAAGACG GTACTCACCA AGGCCATGGC GACCAGCGTG CGTTCTTCTC
13861 TGTTGTTTGT AGTAGTATGA TGAGGCGCGT GTACCCGGAG GGTCCTCCTC CCTCGTACGA
```

FIG. 7D

```
13921 GAGCGTGATG CAGCAGGCGG TGGCGGCGGC GATGCAGCCC CCGCTGGAGG CGCCTTACGT
13981 GCCCCCGCGG TACCTGGCGC CTACGGAGGG GCGGAACAGC ATTCGTTACT CGGAGCTGGC
14041 ACCCTTGTAC GATACCACCC GGTTGTACCT GGTGGACAAC AAGTCGGCAG ACATCGCCTC
14101 GCTGAACTAC CAGAACGACC ACAGCAACTT CCTGACCACC GTGGTGCAGA ACAACGATTT
14161 CACCCCCACG GAGGCCAGCA CCCAGACCAT CAACTTTGAC GAGCGCTCGC GGTGGGGCGG
14221 CCAGCTGAAA ACCATCATGC ACACCAACAT GCCCAACGTG AACGAGTTCA TGTACAGCAA
14281 CAAGTTCAAG GCGCGGGTGA TGGTCTCGCG CAAGACCCCC AACGGGGTGG ATGATGATTA
14341 TGATGGTAGT CAGGACGAGC TGACCTACGA GTGGGTGGAG TTTGAGCTGC CCGAGGGCAA
14401 CTTCTCGGTG ACCATGACCA TCGATCTGAT GAACAACGCC ATCATCGACA ACTACTTGGC
14461 GGTGGGGCGG CAGAACGGGG TGCTGGAGAG CGACATCGGC GTGAAGTTCG ACACGCGCAA
14521 CTTCCGGCTG GGCTGGGACC CCGTGACCGA GCTGGTGATG CCGGGCGTGT ACACCAACGA
14581 GGCCTTCCAC CCCGACATCG TCCTGCTGCC CGGCTGCGGC GTGGACTTCA CCGAGAGCCG
14641 CCTCAGCAAC CTGCTGGGCA TCCGCAAGCG GCAGCCCTTC CAGGAGGGCT TCCAGATCCT
14701 GTACGAGGAC CTGGAGGGGG GCAACATCCC CGCGCTCTTG GATGTCGAAG CCTACGAGAA
14761 AAGCAAGGAG GATAGCACCG CCGCGGCGAC CGCAGCCGTG GCCACCGCCT CTACCGAGGT
14821 GCGGGGCGAT AATTTTGCTA GCGCTGCGGC AGCGGCCGAG GCGGCTGAAA CCGAAAGTAA
14881 GATAGTCATC CAGCCGGTGG AGAAGGACAG CAAGGACAGG AGCTACAACG TGCTCGCGGA
14941 CAAGAAAAAC ACCGCCTACC GCAGCTGGTA CCTGGCCTAC AACTACGGCG ACCCCGAGAA
15001 GGGCGTGCGC TCCTGGACGC TGCTCACCAC CTCGGACGTC ACCTGCGGCG TGGAGCAAGT
15061 CTACTGGTCG CTGCCCGACA TGATGCAAGA CCCGGTCACC TTCCGCTCCA CGCGTCAAGT
15121 TAGCAACTAC CCGGTGGTGG GCGCCGAGCT CCTGCCCGTC TACTCCAAGA GCTTCTTCAA
15181 CGAGCAGGCC GTCTACTCGC AGNAGCTGCG CGCCTTCACC TCGCTCACGC ACGTCTTCAA
15241 CCGCTTCCCC GAGAACCAGA TCCTCGTCCG CCGCCGCGCC CACCATTACC ACCGTCAGTG
15301 AAAACGTTCC TGCTCTCACA GATCACGGGA CCCTGCCGCT GCGCAGCAGT ATCCGGGGAG
15361 TCCAGCGCGT GACCGTCACT GACGCCAGAC GCCGCACCTG CCCCTACGTC TACAAGGCCC
15421 TGGGCGTAGT CGCGCCGCGC GTCCTCTCGA GCCGCACCTT CTAAAAAATG TCCATTCTCA
15481 TCTCGCCCAG TAATAACACC GGTTGGGGCC TGCGCGCGCC CAGCAAGATG TACGGAGGCG
15541 CTCGCCAACG CTCCACGCAA CACCCCGTGC GCGTGCGCGG GCACTTCCGC GCTCCCTGGG
15601 GCGCCCTCAA GGGCCGCGTG CGCTCGCGCA CCACCGTCGA CGACGTGATC GACCAGGTGG
15661 TGGCCGACGC GCGCAACTAC ACGCCCGCCG CCGCGCCCGT CTCCACCGTG GACGCCGTCA
15721 TCGACAGCGT GGTGGCCGAC GCGCGCCGGT ACGCCCGCAC CAAGAGCCGG CGGCGGCGCA
15781 TCGCCCGGCG GCACCGGAGC ACCCCCGCCA TGCGCGCGGC GCGAGCCTTG CTGCGCAGGG
15841 CCAGGCGCAC GGGACGCAGG GCCATGCTCA GGGCGGCCAG ACGCGCGGCC TCCGGCAGCA
15901 GCAGCGCCGG CAGGACCCGC AGACGCGCGG CCACGGCGGC GGCGGCGGCC ATCGCCAGCA
15961 TGTCCCGCCC GCGGCGCGGC AACGTGTACT GGGTGCGCGA CGCCGCCACC GGTGTGCGCG
16021 TGCCCGTGCG CACCCGCCCC CCTCGCACTT GAAGATGCTG ACTTCGCGAT GTTGATGTGT
16081 CCCAGCGGCG AGGAGGATGT CCAAGCGCAA ATACAAGGAA GAGATGCTCC AGGTCATCGC
16141 GCCTGAGATC TACGGCCCCG CGGCGGCGGT GAAGGAGGAA AGAAAGCCCC GCAAACTGAA
16201 GCGGGTCAAA AAGGACAAAA AGGAGGAGGA AGATGACGGA CTGGTGGAGT TTGTGCGCGA
16261 GTTCGCCCCC CGGCGGCGCG TGCAGTGGCG CGGGCGGAAA GTGAAACCGG TGCTGCGGCC
16321 CGGCACCACG GTGGTCTTCA CGCCCGGCGA GCGTTCCGGC TCCGCCTCCA AGCGCTCCTA
16381 CGACGAGGTG TACGGGGACG AGGACATCCT CGAGCAGGCG GTCGAGCGTC TGGGCGAGTT
16441 TGCGTACGGC AAGCGCAGCC GCCCCGCGCC CTTGAAAGAG GAGGCGGTGT CCATCCCGCT
16501 GGACCACGGC AACCCCACGC CGAGCCTGAA GCCGGTGACC CTGCAGCAGG TGCTACCGAG
16561 CGCGGCGCCG CGCCGGGGCT TCAAGCGCGA GGGCGGCGAG GATCTGTACC CGACCATGCA
16621 GCTGATGGTG CCCAAGCGCC AGAAGCTGGA GGACGTGCTG GAGCACATGA AGGTGGACCC
16681 CGAGGTGCAG CCCGAGGTCA AGGTGCGGCC CATCAAGCAG GTGGCCCCGG GCCTGGGCGT
16741 GCAGACCGTG GACATCAAGA TCCCCACGGA GCCCATGGAA ACGCAGACCG AGCCCGTGAA
16801 GCCCAGCACC AGCACCATGG AGGTGCAGAC GGATCCCTGG ATGCCAGCAC CAGCTTCCAC
16861 CAGCACTCGC CGAAGACGCA AGTACGGCGC GGCCAGCCTG CTGATGCCCA ACTACGCGGC
16921 TGCATCCTTC CATCATCCCC ACGCCGGGCT ACCGCGGCAC GCGCTTCTAC CGCGGCTACA
16981 CCAGCAGCCG CCGCCGCAAG ACCACCACCC GCCGCCGTCG TCGCAGCCGC CGCAGCAGCA
17041 CCGCGACTTC CGCCTTGGTG CGGAGAGTGT ATCGCAGCGG GCGCGAGCCT CTGACCCTGC
17101 CGCGCGCGCG CTACCACCCG AGCATCGCCA TTTAACTACC GCCTCCTACT TGCAGATATG
17161 GCCCTCACAT GCCGCCTCCG CGTCCCCATT ACGGGCTACC GAGGAAGAAA GCCGCGCCGT
17221 AGAAGGCTGA CGGGGAACGG GCTGCGTCGC CATCACCACC GGCGGCGGCG CGCCATCAGC
17281 AAGCGGTTGG GGGGAGGCTT CCTGCCCGCG CTGATCCCCA TCATCGCCGC GGCGATCGGG
17341 GCGATCCCCG GCATAGCTTC CGTGGCGGTG CAGGCCTCTC AGCGCCACTG AGACACAAAA
```

FIG. 7E

17401 AAGCATGGAT TTGTAATAAA AAAAAAAATG GACTGACGCT CCTGGTCCTG TGATGTGTGT
17461 TTTTAGATGG AAGACATCAA TTTTTCGTCC CTGGCACCGC GACACGGCAC GCGGCCGTTT
17521 ATGGGCACCT GGAGCGACAT CGGCAACAGC CAACTGAACG GGGGCGCCTT CAATTGGAGC
17581 AGTCTCTGGA GCGGGCTTAA GAATTTCGGG TCCACGCTCA AAACCTATGG CAACAAGGCG
17641 TGGAACAGCA GCACAGGGCA GGCGCTGAGG GAAAAGCTGA AAGAACAGAA CTTCCAGCAG
17701 AAGGTGGTTG ATGGCCTGGC CTCAGGCATC AACGGGGTGG TTGACCTGGC CAACCAGGCC
17761 GTGCAGAAAC AGATCAACAG CCGCCTGGAC GCGGTCCCGC CCGCGGGGTC CGTGGAGATG
17821 CCCCAGGTGG AGGAGGAGCT GCCTCCCCTG GACAAGCGCG GCGACAAGCG ACCGCGTCCC
17881 GACGCGGAGG AGACGCTGCT GACGCACACG GACGAGCCGC CCCCGTACGA GGAGGCGGTG
17941 AAACTGGGCC TGCCCACCAC GCGGCCCGTG GCGCCTCTGG CCACCGGAGT GCTGAAACCC
18001 AGCAGCAGCC AGCCCGCGAC CCTGGACTTG CCTCCGCCTC GCCCCTCCAC AGTGGCTAAG
18061 CCCCTGCCGC CGGTGGCCGT CGCGTCGCGC GCCCCCCGAG GCCGCCCCCA GGCGAACTGG
18121 CAGAGCACTC TGAACAGCAT CGTGGGTCTG GGAGTGCAGA GTGTGAAGCG CCGCCGCTGC
18181 TATTAAAAGA CACTGTAGCG CTTAACTTGC TTGTCTGTGT GTATATGTAT GTCCGCCGAC
18241 CAGAAGGAGG AGTGTGAAGA GGCGCGTCGC CGAGTTGCAA GATGGCCACC CCATCGATGC
18301 TGCCCCAGTG GGCGTACATG CACATCGCCG GACAGGACGC TTCGGAGTAC CTGAGTCCGG
18361 GTCTGGTGCA GTTCGCCCGC GCCACAGACA CCTACTTCAG TCTGGGGAAC AAGTTTAGGA
18421 ACCCCACGGT GGCGCCCACG CACAATGTGA CCACCGACCG CAGCCAGCGG CTGACGGTGC
18481 GCTTCGTGCC CGTGGACCGC GAGGACAACA CCTACTCGTA CAAAGTGCGC TACACGCTGG
18541 CCGTGGGCGA CAACCGCGTG CTGGACATGG CCAGCACCTA CTTTGACATC CGCGGCGTGC
18601 TGGACCGGGG CCCTAGCTTC AAACCCTACT CTGGCACCGC CTACAACAGC CTAGCTCCCA
18661 AGGGAGCTCC CAATTCCAGC CAGTGGGAGC AAGCAAAAAC AGGCAATGGG GGAACTATGG
18721 AAACACACAC ATATGGTGTG GCCCCAATGG GCGGAGAGAA TATTACAAAA GATGGTCTTC
18781 AAATTGGAAC TGACGTTACA GCGAATCAGA ATAAACCAAT TTATGCCGAC AAAACATTTC
18841 AACCAGAACC GCAAGTAGGA GAAGAAAATT GGCAAGAAAC TGAAAACTTT TATGGCGGTA
18901 GAGCTCTTAA AAAAGACACA AACATGAAAC CTTGCTATGG CTCCTATGCT AGACCCACCA
18961 ATGAAAAAGG AGGTCAAGCT AAACTTAAAG TTGGAGATGA TGGAGTTCCA ACCAAAGAAT
19021 TCGACATAGA CCTGGCTTTC TTTGATACTC CCGGTGGCAC CGTGAACGGT CAAGACGAGT
19081 ATAAAGCAGA CATTGTCATG TATACCGAAA ACACGTATTT GGAAACTCCA GACACGCATG
19141 TGGTATACAA ACCAGGCAAG GATGATGCAA GTTCTGAAAT TAACCTGGTT CAGCAGTCTA
19201 TGCCCAACAG ACCCAACTAC ATTGGGTTCA GGGACAACTT TATCGGTCTT ATGTACTACA
19261 ACAGCACTGG CAATATGGGT GTGCTTGCTG GTCAGGCCTC CCAGCTGAAT GCTGTGGTTG
19321 ATTTGCAAGA CAGAAACACC GAGCTGTCCT ACCAGCTCTT GCTTGACTCT TTGGGTGACA
19381 GAACCCGGTA TTTCAGTATG TGGAACCAGG CGGTGGACAG TTATGACCCC GATGTGCGCA
19441 TCATCGAAAA CCATGGTGTG GAGGATGAAT TGCCAAACTA TTGCTTCCCC TTGGACGGCT
19501 CTGGCACTAA CGCCGCATAC CAAGGTGTGA AAGTAAAAGA TGGTCAAGAT GGTGATGTTG
19561 AGAGTGAATG GGAAAATGAC GATACTGTTG CAGCTCGAAA TCAATTATGT AAAGGTAACA
19621 TTTTCGCCAT GGAGATTAAT CTCCAGGCTA ACCTGTGGAG AAGTTTCCTC TACTCGAACG
19681 TGGCCCTGTA CCTGCCCGAC TCCTACAAGT ACACGCCGAC CAACGTCACG CTGCCGACCA
19741 ACACCAACAC CTACGATTAC ATGAATGGCA GAGTGACACC TCCCTCGCTG GTAGACGCCT
19801 ACCTCAACAT CGGGGCGCGC TGGTCGCTGG ACCCCATGGA CAACGTCAAC CCCTTCAACC
19861 ACCACCGCAA CGCGGGCCTG CGCTACCGCT CCATGCTCCT GGGCAACGGG CGCTACGTGC
19921 CCTTCCACAT CCAGGTGCCC CAAAAGTTTT TCGCCATCAA GAGCCTCCTG CTCCTGCCCG
19981 GGTCCTACAC CTACGAGTGG AACTTCCGCA AGGACGTCAA CATGATCCTG CAGAGCTCCC
20041 TAGGCAACGA CCTGCGCACG GACGGGGCCT CCATCGCCTT CACCAGCATC AACCTCTACG
20101 CCACCTTCTT CCCCATGGCG CACAACACCG CCTCCACGCT CGAGGCCATG CTGCGCAACG
20161 ACACCAACGA CCAGTCCTTC AACGACTACC TCTCGGCGGC CAACATGCTC TACCCCATCC
20221 CGGCCAACGC CACCAACGTG CCCATCTCCA TCCCCTCGCG CAACTGGGCC GCCTTCCGCG
20281 GATGGTCCTT CACGCGCCTG AAGACCCGCG AGACGCCCTC GCTCGGCTCC GGGTTCGACC
20341 CCTACTTCGT CTACTCGGGC TCCATCCCCT ACCTAGACGG CACCTTCTAC CTCAACCACA
20401 CCTTCAAGAA GGTCTCCATC ACCTTCGACT CCTCCGTCAG CTGGCCCGGC AACGACCGCC
20461 TCCTGACGCC CAACGAGTTC GAAATCAAGC GCACCGTCGA CGGAGAGGGA TACAACGTGG
20521 CCCAGTGCAA CATGACCAAG GACTGGTTCC TGGTCCAGAT GCTGGCCCAC TACAACATCG
20581 GCTACCAGGG CTTCTACGTG CCCGAGGGCT ACAAGGACCG CATGTACTCC TTCTTCCGCA
20641 ACTTCCAGCC CATGAGCCGC CAGGTCGTGG ACGAGGTCAA CTACAAGGAC TACCAGGCCG
20701 TCACCCTGGC CTACCAGCAC AACAACTCGG GCTTCGTCGG CTACCTCGCG CCCACCATGC
20761 GCCAGGGCCA GCCCTACCCC GCCAACTACC CCTACCCGCT CATCGGCAAG AGCGCCGTCG
20821 CCAGCGTCAC CCAGAAAAAG TTCCTCTGCG ACCGGGTCAT GTGGCGCATC CCCTTCTCCA

FIG. 7F

```
21661 GCTTCCTCAA TGCCCACTCC GCCTACTTTC GCTCCCACCG CGCGCGCATC GAGAAGGCCA
20881 GCAACTTCAT GTCCATGGGC GCGCTCACCG ACCTCGGCCA GAACATGCTC TACGCCAACT
20941 CCGCCCACGC GCTAGACATG AATTTCGAAG TCGACCCCAT GGATGAGTCC ACCCTTCTCT
21001 ATGTTGTCTT CGAAGTCTTC GACGTCGTCC GAGTGCACCA GCCCCACCGC GGCGTCATCG
21061 AAGCCGTCTA CCTGCGCACG CCCTTCTCGG CCGGCAACGC CACCACCTAA GCCGCTCTTG
21121 CTTCTTGCAA GATGACGGCG GGCTCCGGCG AGCAGGAGCT CAGGGCCATC CTCCGCGACC
21181 TGGGCTGCGG GCCCTGCTTC CTGGGCACCT TCGACAAGCG CTTCCCTGGA TTCATGGCCC
21241 CGCACAAGCT GGCCTGCGCC ATCGTGAACA CGGCCGGCCG CGAGACCGGG GGCGAGCACT
21301 GGCTGGCCTT CGCCTGGAAC CCGCGCTCCC ACACATGCTA CCTCTTCGAC CCCTTCGGGT
21361 TCTCGGACGA GCGCCTCAAG CAGATCTACC AGTTCGAGTA CGAGGGCCTG CTGCGTCGCA
21421 GCGCCCTGGC CACCGAGGAC CGCTGCGTCA CCCTGGAAAA GTCCACCCAG ACCGTGCAGG
21481 GTCCGCGCTC GGCCGCCTGC GGGCTCTTCT GCTGCATGTT CCTGCACGCC TTCGTGCACT
21541 GGCCCGACCG CCCCATGGAC AAGAACCCCA CCATGAACTT ACTGACGGGG GTGCCCAACG
21601 GCATGCTCCA GTCGCCCCAG GTGGAACCCA CCCTGCGCCG CAACCAGGAA GCGCTCTACC
21721 CCGCCTTCGA CCGCATGAAT CAAGACATGT AAAAAACCGG TGTGTGTATG TGAATGCTTT
21781 ATTCATAATA AACAGCACAT GTTTATGCCA CCTTCTCTGA GGCTCTGACT TTATTTAGAA
21841 ATCGAAGGGG TTCTGCCGGC TCTCGGCATG GCCCGCGGGC AGGGATACGT TGCGGAACTG
21901 GTACTTGGGC AGCCACTTGA ACTCGGGGAT CAGCAGCTTG GGCACGGGGA GGTCGGGGAA
21961 CGAGTCGCTC CACAGCTTGC GCGTGAGTTG CAGGGCGCCC AGCAGGTCGG GCGCGGAGAT
22021 CTTGAAATCG CAGTTGGGAC CCGCGTTCTG CGCGCGAGAG TTGCGGTACA CGGGGTTGCA
22081 GCACTGGAAC ACCATCAGGG CCGGGTGCTT CACGCTTGCC AGCACCGTCG CGTCGGTGAT
22141 GCCCTCCACG TCCAGATCCT CGGCGTTGGC CATCCCGAAG GGGGTCATCT TGCAGGTCTG
22201 CCGCCCCATG CTGGGCACGC AGCCGGGCTT GTGGTTGCAA TCGCAGTGCA GGGGGATCAG
22261 CATCATCTGG GCCTGCTCGG AGCTCATGCC CGGGTACATG GCCTTCATGA AAGCCTTCAG
22321 CTGGCGGAAG GCCTGCTGCG CCTTGCCGCC CTCGGTGAAG AAGACCCCGC AGGACTTGCT
22381 AGAGAACTGG TTGGTGGCGC AGCCGGCGTC GTGCACGCAG CAGCGCGCGT CGTTGTTGGC
22441 CAGCTGCACC ACGCTGCGCC CCCAGCGGTT CTGGGTGATC TTGGCCCGGT TGGGGTTCTC
22501 CTTCAGCGCG CGCTGCCCGT TCTCGCTCGC CACATCCATC TCGATAGTGT GCTCCTTCTG
22561 GATCATCACG GTCCCGTGCA GGCACCGCAG CTTGCCCTCG GCTTCGGTGC AGCCGTGCAG
22621 CCACAGCGCG CAGCCGGTGC ACTCCCAGTT CTTGTGGGCG ATCTGGGAGT GCGAGTGCAC
22681 GAAGCCCTGC AGGAAGCGGC CCATCATCGC GGTCAGGGTC TTGTTGCTGG TGAAGGTCAG
22741 CGGGATGCCG CGGTGCTCCT CGTTCACATA CAGGTGGCAG ATGCGGCGGT ACACCTCGCC
22801 CTGCTCGGGC ATCAGCTGGA AGGCGGACTT CAGGTCGCTC TCCACGCGGT ACCGGTCCAT
22861 CAGCAGCGTC ATCACTTCCA TGCCCTTCTC CCAGGCCGAA ACGATCGGCA GGCTCAGGGG
22921 GTTCTTCACC GCCATTGTCA TCTTAGTCGC CGCCGCCGAG GTCAGGGGGT CGTTCTCGTC
22981 CAGGGTCTCA AACACTCGCT TGCCGTCCTT CTCGATGATG CGCACGGGGG GAAAGCTGAA
23041 GCCCACGGCC GCCAGCTCCT CCTCGGCCTG CCTTTCGTCC TCGCTGTCCT GGCTGATGTC
23101 TTGCAAAGGC ACATGCTTGG TCTTGCGGGG TTTCTTTTTG GGCGGCAGAG GCGGCGGCGA
23161 TGTGCTGGGA GAGCGCGAGT TCTCGTTCAC CACGACTATT TCTTCTTCTT GGCCGTCGTC
23221 CGAGACCACG CGGCGGTAGG CATGCCTCTT CTGGGGCAGA GGCGGAGGCG ACGGGCTCTC
23281 GCGGTTCGGC GGGCGGCTGG CAGAGCCCCT TCCGCGTTCG GGGGTGCGCT CCTGGCGGCG
23341 CTGCTCTGAC TGACTTCCTC CGCGGCCGGC CATTGTGTTC TCCTAGGGAG CAACAACAAG
23401 CATGGAGACT CAGCCATCGT CGCCAACATC GCCATCTGCC CCCGCCGCCA CCGCCGACGA
23461 GAACCAGCAG CAGAATGAAA GCTTAACCGC CCCGCCGCCC AGCCCCACCT CCGACGCCGC
23521 GGCCCCAGAC ATGCAAGAGA TGGAGGAATC CATCGAGATT GACCTGGGCT ACGTGACGCC
23581 CGCGGAGCAC GAGGAGGAGC TGGCAGCGCG CTTTTCAGCC CCGGAAGAGA ACCACCAAGA
23641 GCAGCCAGAG CAGGAAGCAG AGAACGAGCA GAACCAGGCT GGGCACGAGC ATGGCGACTA
23701 CCTGAGCGGG GCAGAGGACG TGCTCATCAA GCATCTGGCC CGCCAATGCA TCATCGTCAA
23761 GGACGCGCTG CTCGACCGCG CCGAGGTGCC CCTCAGCGTG GCGGAGCTCA GCCGCGCCTA
23821 CGAGCGCAAC CTCTTCTCGC CGCGCGTGCC CCCCAAGCGC CAGCCCAACG GCACCTGTGA
23881 GCCCAACCCG CGCCTCAACT TCTACCCGGT CTTCGCGGTG CCCGAGGCCC TGGCCACCTA
23941 CCACCTCTTT TTCAAGAACC AAAGGATCCC CGTCTCCTGC CGCGCCAACC GCACCCGCGC
24001 CGACGCCCTG CTCAACCTGG GCCCCGGCGC CCGCCTACCT GATATCACCT CCTTGGAAGA
24061 GGTTCCCAAG ATCTTCGAGG GTCTGGGCAG CGACGAGACT CGGGCCGCGA ACGCTCTGCA
24121 AGGAAGCGGA GAGGAGCATG AGCACCACAG CGCCCTGGTG GAGTTGGAAG GCGACAACGC
24181 GCGCCTGGCG GTCCTCAAGC GCACGGTCGA GCTGACCCAC TTCGCCTACC CGGCGCTCAA
24241 CCTGCCCCCC AAGGTCATGA GCGCCGTCAT GGACCAGGTG CTCATCAAGC GCGCCTCGCC
24301 CCTCTCGGAG GAGGAGATGC AGGACCCCGA GAGTTCGGAC GAGGGCAAGC CCGTGGTCAG
```

FIG. 7G

```
24361 CGACGAGCAG CTGGCGCGCT GGCTGGGAGC GAGTANCACC CCCCAGAGCC TGGAAGAGCG
24421 GCGCAAGCTC ATGATGGCCG TGGTCCTGGT GACCGTGGAG CTGGAGTGTC TGCGCCGCTT
24481 CTTTGCCGAC GCGGAGACCC TGCGCAAGGT CGAGGAGAAC CTGCACTACC TCTTCAGGCA
24541 CGGGTTCGTG CGCCAGGCCT GCAAGATCTC CAACGTGGAG CTGACCAACC TGGTCTCCTA
24601 CATGGGCATC CTGCACGAGA ACCGCCTGGG GCAAAACGTG CTGCACACCA CCCTGCGCGG
24661 GGAGGCCCGC CGCGACTACA TCCGCGACTG CGTCTACCTG TACCTCTGCC ACACCTGGCA
24721 GACGGGCATG GGCGTGTGGC AGCAGTGCCT GGAGGAGCAG AACCTGAAAG AGCTCTGCAA
24781 GCTCCTGCAG AAGAACCTCA AGGCCCTGTG GACCGGGTTC GACGAGCGTA CCACCGCCTC
24841 GGACCTGGCC GACCTCATCT TCCCCGAGCG CCTGCGGCTG ACGCTGCGCA ACGGGCTGCC
24901 CGACTTTATG AGCCAAAGCA TGTTGCAAAA CTTTCGCTCT TTCATCCTCG AACGCTCCGG
24961 GATCCTGCCC GCCACCTGCT CCGCGCTGCC CTCGGACTTC GTGCCGCTGA CCTTCCGCGA
25021 GTGCCCCCCG CCGCTCTGGA GCCACTGCTA CTTGCTGCGC CTGGCCAACT ACCTGGCCTA
25081 CCACTCGGAC GTGATCGAGG ACGTCAGCGG CGAGGGTCTG CTGGAGTGCC ACTGCCGCTG
25141 CAACCTCTGC ACGCCGCACC GCTCCCTGGC CTGCAACCCC CAGCTGCTGA GCGAGACCCA
25201 GATCATCGGC ACCTTCGAGT TGCAAGGCCC CGGCGACGGC GAGGGCAAGG GGGGTCTGAA
25261 ACTCACCCCG GGGCTGTGGA CCTCGGCCTA CTTGCGCAAG TTCGTGCCCG AGGACTACCA
25321 TCCCTTCGAG ATCAGGTTCT ACGAGGACCA ATCCCAGCCG CCCAAGGCCG AGCTGTCGGC
25381 CTGCGTCATC ACCCAGGGGG CCATCCTGGC CCAATTGCAA GCCATCCAGA AATCCCGCCA
25441 AGAATTTCTG CTGAAAAAGG GCCACGGGGT CTACTTGGAC CCCCAGACCG GAGAGGAGCT
25501 CAACCCCAGC TTCCCCCAGG ATGCCCCGAG GAAGCAGCAA GAAGCTGAAA GTGGAGCTGC
25561 CGCCGCCGGA GGATTTGGAG GAAGACTGGG AGAGCAGTCA GGCAGAGGAG GAGGAGATGG
25621 AAGACTGGGA CAGCACTCAG GCAGAGGAGG ACAGCCTGCA AGACAGTCTG GAGGAGGAAG
25681 ACGAGGTGGA GGAGGCAGAG GAAGAAGCAG CCGCCGCCAG ACCGTCGTCC TCGGCGGAGA
25741 AAGCAAGCAG CACGGATACC ATCTCCGCTC CGGGTCGGGG TCGCGGCGGC CGGGCCCACA
25801 GTAGGTGGGA CGAGACCGGG CGCTTCCGAA CCCCACCACC CAGACCGGTA AGAAGGAGCG
25861 GCAGGGATAC AAGTCCTGGC GGGGGCACAA AAACGCCATC GTCTCCTGCT TGCAAGCCTG
25921 CGGGGGCAAC ATCTCCTTCA-CCCGGCGCTA-CCTGCTCTTT CACCGCGGGG-TGAACTTCCC
25981 CCGCAACATC TTGCATTACT ACCGTCACCT CCACAGCCCC TACTACTGTT TCCAAGAAGA
26041 GGCAGAAACC CAGCAGCAGC AGAAAACCAG CGGCAGCAGC AGCTAGAAAA TCCACAGCGG
26101 CGGCAGGTGG ACTGAGGATC-GCGGCGAACG AGCCGGCGCA GACCCGGGAG CTGAGGAACC
26161 GGATCTTTCC CACCCTCTAT GCCATCTTCC AGCAGAGTCG GGGGCAGGAG CAGGAACTGA
26221 AAGTCAAGAA CCGTTCTCTG CGCTCGCTCA CCCGCAGTTG TCTGTATCAC AAGAGCGAAG
26281 ACCAACTTCA GCGCACTCTC GAGGACGCCG AGGCTCTCTT CAACAAGTAC TGCGCGCTCA
26341 CTCTTAAAGA GTAGCCCGCG CCCGCCCACA CACGGAAAAA GGCGGGAATT ACGTCACCAC
26401 CTGCGCCCTT CGCCCGACCA TCATGAGCAA AGAGATTCCC ACGCCTTACA TGTGGAGCTA
26461 CCAGCCCCAG ATGGGCCTGG CCGCCGGCGC CGCCCAGGAC TACTCCACCC GCATGAACTG
26521 GCTCAGTGCC GGGCCCGCGA TGATCTCACG GGTGAATGAC ATCCGCGCCC ACCGAAACCA
26581 GATACTCCTA GAACAGTCAG CGATCACCGC CACGCCCCGC CATCACCTTA ATCCGCGTAA
26641 TTGGCCCGCC GCCCTGGTGT ACCAGGAAAT TCCCCAGCCC ACGACCGTAC TACTTCCGCG
26701 AGACGCCCAG GCCGAAGTCC AGCTGACTAA CTCAGGTGTC CAGCTGGCCG GCGGCGCCGC
26761 CCTGTGTCGT CACCGCCCCG CTCAGGGTAT AAAGCGGCTG GTGATCCGAG GCAGAGGCAC
26821 ACAGCTCAAC GACGAGGTGG TGAGCTCTTC GCTGGGTCTG CGACCTGACG GAGTCTTCCA
26881 ACTCGCCGGA TCGGGGAGAT CTTCCTTCAC GCCTCGTCAG GCCGTCCTGA CTTTGGAGAG
26941 TTCGTCCTCG CAGCCCCGCT CGGGCGGCAT CGGCACTCTC CAGTTCGTGG AGGAGTTCAC
27001 TCCCTCGGTN TACTTCAACC CCTTCTCCGG CTCCCCCGGC CACTACCCGG ACGAGTTCAT
27061 CCCGAACTTC GACGCCATCA GCGAGTCGGT GGACGGCTAC GATTGAATGT CCCATGGTGG
27121 CGCAGCTGAC CTAGCTCGGC TTCGACACCT GGACCACTGC CGCCGCTTCC GCTGCTTCGC
27181 TCGGGATCTC GCCGAGTTTG CCTACTTTGA GCTGCCCGAG GAGCACCCTC AGGGCCCAGC
27241 CCACGGAGTG CGGATCATCG TCGAAGGGGG CCTCGACTCC CACCTGCTTC GGATCTTCAG
27301 CCAGCGACCG ATCCTGGTCG AGCGCGAACA AGGACAGACC CTTCTTACTT TGTACTGCAT
27361 CTGCAACCAC CCCGGCCTGC ATGAAAGTCT TTGTTGTCTG CTGTGTACTG AGTATAATAA
27421 AAGCTGAGAT CAGCGACTAC TCCGGACTCG ATTGTGGTGT TCCTGCTATC AACCGGTCCC
27481 TGTTCTTCAC CGGGAACGAG ACCGAGCTCC AGCTCCAGTG TAAGCCCCAC AAGAAGTACC
27541 TCACCTGGCT GTTCCAGGGC TCCCCGATCG CCGTTGTCAA CCACTGCGAC AACGACGGAG
27601 TCCTGCTGAG CGGCCCTGCC AACCTTACTT TTTCCACCCG CAGAAGCAAG CTCCAGCTCT
27661 TCCAACCCTT CCTCCCCGGG ACCTATCAGT GCGTCTCAGG ACCCTGCCAT CACACCTTCC
27721 ACCTGATCCC GAATACCACA GCGCCGCTCC CCGCTACTAA CAACCAAACT ACCCACCAAC
27781 GCCACCGTCG CGACCTTTCC TCTGAATCTA ATACCACTAC CGGAGGTGAG CTCCGAGGTC
```

FIG. 7H

```
27841 GACCAACCTC TGGGATTTAC TACGGCCCCT GGGAGGTGGT GGGGTTAATA GCGCTAGGCC
27901 TAGTTGCGGG TGGGCTTTTG GTTCTCTGCT ACCTATACCT CCCTTGCTGT TCGTACTTAG
27961 TGGTGCTGTG TTGCTGGTTT AAGAAATGGG GAAGATCACC CTAGTGAGCT GCGGTGCGCT
28021 GGTGGCGGTG TTGCTTTCGA TTGTGGGACT GGGCGGCGCG GCTGTAGTGA AGGAGAAGGC
28081 CGATCCCTGC TTGCATTTCA ATCCCAACAA ATGCCAGCTG AGTTTTCAGC CCGATGGCAA
28141 TCGGTGCGCG GTACTGATCA AGTGCGGATG GGAATGCGAG AACGTGAGAA TCGAGTACAA
28201 TAACAAGACT CGGAACAATA CTCTCGCGTC CGTGTGGCAG CCCGGGGACC CCGAGTGGTA
28261 CACCGTCTCT GTCCCCGGTG CTGACGGCTC CCCGCGCACC GTGAATAATA CTTTCATTTT
28321 TGCGCACATG TGCAACACGG TCATGTGGAT GAGCAAGCAG TACGATATGT GGCCCCCCAC
28381 GAAGGAGAAC ATCGTGGTCT TCTCCATCGC TTACAGCCTG TGCACGGCGC TAATCACCGC
28441 TATCGTGTGC CTGAGCATTC ACATGCTCAT CGCTATTCGC CCCAGAAATA ATGCCGAGAA
28501 AGAGAAACAG CCATAACACG TTTTTTCACA CACCTTGTTT TTACAGACAA TGCGTCTGTT
28561 AAATTTTTTA AACATTGTGC TCAGTATTGC TTATGCCTCT GGTTATGCAA ACATACAGAA
28621 AACCCTTTAT GTAGGATCTG ATGGTACACT AGAGNGTACC CAATCACAAG CCAAGGTTGC
28681 ATGGTATTTT TATAGAACCA ACACTGATCC AGTTAAACTT TGTAAGGGTG AATTGCCGCG
28741 TACACATAAA ACTCCACTTA CATTTAGTTG CAGCAATAAT AATCTTACAC TTTTTTCAAT
28801 TACAAAACAA TATACTGGTA CTTATTACAG TACAAACTTT CATACAGGAC AAGATAAATA
28861 TTATACTGTT AAGGTAGAAA ATCCTACCAC TCCTAGAACT ACCACCACCA CCACTACTGC
28921 AAAGCCCACT GTGAAAACTA CAACTAGGAC CACCACAACT ACAGAAACCA CCACCAGCAC
28981 AACACTTGCT GCAACTACAC ACACACACAC TAAGCTAACC TTACAGACCA CTAATGATTT
29041 GATCGCCCTG CTGCAAAAGG GGGATAACAG CACCACTTCC AATGAGGAGA TACCCAAATC
29101 CATGATTGGC ATTATTGTTG CTGTAGTGGT GTGCATGTTG ATCATCGCCT TGTGCATGGT
29161 GTACTATGCC TTCTGCTACA GAAAGCACAG ACTGAACGAC AAGCTGGAAC ACTTACTAAG
29221 TGTTGAATTT TAATTTTTTA GAACCATGAA GATCCTAGGC CTTTTTAGTT TTTCTATCAT
29281 TACCTCTGCT CTTTGTGAAT CAGTGGATAG AGATGTTACT ATTACCACTG GTTCTAATTA
29341 TACACTGAAA GGGCCACCCT CAGGTATGCT TTCGTGGTAT TGCTATTTTG GAACTGACAC
29401 TGATCAAACT GAATTATGCA ATTTTCAAAA AGGCAAAACC TCAAACTCTA AAATCTCTAA
29461 TTATCAATGC AATGGCACTG ATCTGATACT ACTCAATGTC ACGAAAGCAT ATGGTGGCAG
29521 TTATTATTGC CCTGGACAAA ACACTGAAGA AATGATTTTT TACAAAGTGG AAGTGGTTGA
29581 TCCCACTACA CCACCCACCA CCACAACTAT TCATACCACA CACACAGAAC AAACACCAGA
29641 GGCAACAGAA GCAGAGTTGG CCTTCCAGGT TCACGGAGAT TCCTTTGCTG TCAATACCCC
29701 TACACCCGAT CAGCGGTGTC CGGGGCCGCT AGTCAGCGGC ATTGTCGGTG TGCTTTCGGG
29761 ATTAGCAGTC ATAATCATCT GCATGTTCAT TTTTGCTTGC TGCTATAGAA GGCTTTACCG
29821 ACAAAAATCA GACCCACTGC TGAACCTCTA TGTTTAATTT TTTCCAGAGC CATGAAGGCA
29881 GTTAGCGCTC TAGTTTTTTG TTCTTTGATT GGCATTGTTT TTAATAGTAA AATTACCAGA
29941 GTTAGCTTTA TTAAACATGT TAATGTAACT GAAGGAGATA ACATCACACT AGCAGGTGTA
30001 GAAGGTGCTC AAAACACCAC CTGGACAAAA TACCATCTAG GATGGAGAGA TATTTGCACC
30061 TGGAATGTAA CTTATTATTG CATAGGAGTT AATCTTACCA TTGTTAACGC TAACCAATCT
30121 CAGAATGGGT TAATTAAAGG ACAGAGTGTT AGTGTGACCA GTGATGGGTA CTATACCCAG
30181 CATAGTTTTA ACTACAACAT TACTGTCATA CCACTGCCTA CGCCTAGCCC ACCTAGCACT
30241 ACCACACAGA CAACCACATA CAGTACATCA AATCAGCCTA CCACCACTAC AGCAGCAGAG
30301 GTTGCCAGCT CGTCTGGGGT CCGAGTGGCA TTTTTGATGT TGGCCCCATC TAGCAGTCCC
30361 ACTGCTAGTA CCAATGAGCA GACTACTGAA TTTTTGTCCA CTGTCGAGAG CCACACCACA
30421 GCTACCTCCA GTGCCTTCTC TAGCACCGCC AATCTCTCCT CGCTTTCCTC TACACCAATC
30481 AGCCCCGCTA CTACTCCTAG CCCCGCTCCT CTTCCCACTC CCCTGAAGCA AACAGACGGC
30541 GGCATGCAAT GGCAGATCAC CCTGCTCATT GTGATCGGGT TGGTCATCCT GGCCGTGTTG
30601 CTCTACTACA TCTTCTGCCG CCGCATTCCC AACGCGCACC GCAAGCCGGC CTACAAGCCC
30661 ATCGTTATCG GGCAGCCGGA GCCGCTTCAG GTGGAAGGGG GTCTAAGGAA TCTTCTCTTC
30721 TCTTTTACAG TATGGTGATT GAANTATGAT TCCTAGACAA TTCTTGATCA CTATTCTTAT
30781 CTGCCTCCTC CAAGTCTGTG CCACCCTCGC TCTGGTGGCC AACGCCAGTC CAGACTGTAT
30841 TGGGCCCTTC GCCTCCTACG TGCTCTTTGC CTTCGTCACC TGCATCTGCT GCTGTAGCAT
30901 AGTCTGCCTG CTTATCACCT TCTTCCAGTT CATTGACTGG ATCTTTGTGC GCATCGCCTA
30961 CCTGCGCCAC CACCCCCAGT ACCGCGACCA GCGAGTGGCG CAGCTGCTCA GGCTCCTCTG
31021 ATAAGCATGC GGGCTCTGCT ACTTNTCGCG CTTCTGCTGT TAGTGCTCCC CCGTCCCGTC
31081 GACCCCCGGT CCCCCACTCA GTCCCCCGAG GAGGTTCGCA AATGCAAATT CCAAGAACCC
31141 TGGAAATTCC TCAAATGCTA CCGCCAAAAA TCAGACATGC ATCCCAGCTG GATCATGATC
31201 ATTGGGATCG TGAACATTCT GGCCTGCACC CTCATCTCCT TTGTGATTTA CCCCTGCTTT
31261 GACTTTGGTT GGAACTCGCC AGAGGCGCTC TATCTCCCGC CTGAACCTGA CACACCACCA
```

FIG. 7I

```
31321 CAGCATCAAC CTCAGGCACA CGCACTACCA CCACCACAGC CTAGGCCACA ATACATGCCC
31381 ATATTAGACT ATGAGGCCGA GCCACAGCGA CCCATGCTCC CCGCTATTAG TTACTTCAAT
31441 CTAACCGGCG GAGATGACTG ACCCACTGGC CAATAACAAC GTCAACGACC TTCTCCTGGA
31501 CATGGACGGC CGCGCCTCGG AGCAGCGACT CGCCCAACTT CGCATTCGTC AGCAGCAGGA
31561 GAGAGCCGTC AAGGAGCTGC AGGACGGCAT AGCCATCCAC CAGTGCAAGA GAGGCATCTT
31621 CTGCCTGGTG AAACAGGCCA AGATCTCCTA CGAGGTCACC CAGACCGACC ATCGCCTCTC
31681 CTACGAGCTC CTGCAGCAGC GCCAGAAGTT CACCTGCCTG GTCGGAGTCA ACCCCATCGT
31741 CATCACCCAG CCAGCAGTCG GGCGATACCA AGGGGTGCAT CCACTGCTCC TGCGACTCCC
31801 CCGACTGCGT CCACACTCTG ATCAAGACCC TCTGCGGCCT CCGCGACCTC CTCCCCATGA
31861 ACTAATCACC CCCTTATCCA GTGAAATAAA GATCATATTG ATGATGATTT AAATAAAAAA
31921 AATAATCATT TGATTTGAAA TAAAGATACA ATCATATTGA TGATTTGAGT TTAACAAAAA
31981 TAAAGAATCA CTTACTTGAA ATCTGATACC AGGTCTCTGT CCATGTTTTC TGCCAACACC
32041 ACCTCACTCC CCTCTTCCCA GCTCTGGTAC TGCAGGCCCC GGCGGGCTGC AAACTTCCTC
32101 CACACGCTGA AGGGGATGTC AAATTCCTCC TGTCCCTCAA TCTTCATTTT ATCTTCTATC
32161 AGATGTCCAA AAAGCGCGTC CGGGTGGATG ATGACTTCGA CCCCGTCTAC CCCTACGATG
32221 CAGACAACGC ACCGACCGTG CCCTTCATCA ACCCCCCCTT CGTCTCTTCA GATGGATTCC
32281 AAGAGAAGCC CCTGGGGGTG TTGTCCCTGC GACTGGCTGA CCCCGTCACC ACCAAGAACG
32341 GGGAAATCAC CCTCAAGCTG GGAGAGGGGG TGGACCTCGA CTCGTCGGGA AAACTCATCT
32401 CCAACACGGC CACCAAGGCC GCCGCCCCTC TCAGTATTTC AAACAACACC ATTTCCCTTA
32461 AAACTGCTGC CCCTTTCTAC AACAACAATG GAACTTTAAG CCTCAATGTC TCCACACCAT
32521 TAGCAGTATT TCCCACATTT AACACTTTAG GCATAAGTCT TGGAAACGGT CTTCAGACTT
32581 CAAATAAGTT GTTGACTGTA CAACTAACTC ATCCTCTTAC ATTCAGCTCA AATAGCATCA
32641 CAGTAAAAAC AGACAAAGGG CTATATATTA ACTCCAGTGG AAACAGAGGA CTTGAGGCTA
32701 ATATAAGCCT AAAAAGAGGA CTAGTTTTTG ACGGTAATGC TATTGCAACA TATATTGGAA
32761 ATGGCTTAGA CTATGGATCT TATGATAGTG ATGGAAAAAC AAGACCCGTA ATTACCAAAA
32821 TTGGAGCAGG ATTAAATTTT GATGCTAACA AAGCAATAGC TGTCAAACTA GGCACAGGTT
32881 TAAGTTTTGA CTCCGCTGGT GCCTTGACAG CTGGAAACAA ACAGGATGAC AAGCTAACAC
32941 TTTGGACTAC CCCTGACCCA AGCCCTAATT GTCAATTACT TTCAGACAGA GATGCCAAAT
33001 TTACTCTCTG TCTTACAAAA TGCGGTAGTC AAATACTAGG CACTGTGGCA GTGGCGGCTG
33061 TTACTGTAGG ATCAGCACTA AATCCAATTA ATGACACAGT CAAAAGCGCC ATAGTTTTCC
33121 TTAGATTTGA TTCCGATGGT GTACTCATGT CAAACTCATC AATGGTAGGT GATTACTGGA
33181 ACTTTAGGGA GGGACAGACC ACTCAAAGTG TAGCCTATAC AAATGCTGTG GGATTCATGC
33241 CAAATATAGG TGCATATCCA AAAACCCAAA GTAAAACACC TAAAAATAGC ATAGTCAGTC
33301 AGGTATATTT AACTGGAGAA ACTACTATGC CAATGACACT AACCATAACT TTCAATGGCA
33361 CTGATGAAAA AGACACAACC CCAGTTAGCA CCTACTCTAT GACTTTTACA TGGCAGTGGA
33421 CTGGAGACTA TAAGGACAAA AATATTACCT TTGCTACCAA CTCATTCTCT TTTTCCTACA
33481 TCGCCCAGGA ATAATCCCAC CCAGCAAGCC AACCCCTTTT CCCACCACCT TTGTCTATAT
33541 GGAAACTCTG AAACAGAAAA ATAAAGTTCA AGTGTTTTAT TGAATCAACA GTTTTACAGG
33601 ACTCGAGCAG TTATTTTTCC TCCACCCTCC CAGGACATGG AATACACCAC CCTCTCCCCC
33661 CGCACAGCCT TGAACATCTG AATGCCATTG GTGATGGACA TGCTTTTGGT CTCCACGTTC
33721 CACACAGTTT CAGAGCGAGC CAGTCTCGGA TCGGTCAGGG AGATGAAACC CTCCGGGCAC
33781 TCCCGCATCT GCACCTCACA GCTCAACAGC TGAGGATTGT CCTCGGTGGT CGGGATCACG
33841 GTTATCTGGA AGAAGCAGAA GAGCGGCGGT GGGAATCATA GTCCGCGAAC GGGATCGGCC
33901 GGTGGTGTCG CATCAGGCCC CGCAGCAGTC GCTGCCGCCG CCGCTCCGTC AAGCTGCTGC
33961 TCAGGGGGTT CGGGTCCAGG GACTCCCTCA GCATGATGCC CACGGCCCTC AGCATCAGTC
34021 GTCTGGTGCG GCGGGCGCAG CAGCGCATGC GAATCTCGCT CAGGTCACTG CAGTACGTGC
34081 AACACAGGAC CACCAGGTTG TTCAACAGTC CATAGTTCAA CACGCTCCAG CCGAAACTCA
34141 TCGCGGGAAG GATGCTACCC ACGTGGCCGT CGTACCAGAT CCTCAGGTAA ATCAAGTGGC
34201 GCTCCCTCCA GAAGACGCTG CCCATGTACA TGATCTCCTT GGGCATGTGG CGGTTCACCA
34261 CCTCCCGGTA CCACATCACC CTCTGGTTGA ACATGCAGCC CCGGATGATC CTGCGGAACC
34321 ACAGGGCCAG CACCGCCCCG CCCGCCATGC AGCGAAGAGA CCCCGGATCC CGGCAATGAC
34381 AATGGAGGAC CCACCGCTCG TACCCGTGGA TCATCTGGGA GCTGAACAAG TCTATGTTGG
34441 CACAGCACAG GCATATGCTC ATGCATCTCT TCAGCACTCT CAGCTCCTCG GGGGTCAAAA
34501 CCATATCCCA GGGCACGGGG AACTCTTGCA GGACAGCGAA CCCCGCAGAA CAGGGCAATC
34561 CTCGCACATA ACTTACATTG TGCATGGACA GGGTATCGCA ATCAGGCAGC ACCGGGTGAT
34621 CCTCCACCAG AGAAGCGCGG GTCTCGGTCT CCTCACAGCG TGGTAAGGGG GCCGGCCGAT
34681 ACGGGTGATG GCGGGACGCG GCTGATCGTG TTCTCGACCG TGTCATGATG CAGTTGCTTT
34741 CGGACATTTT CGTACTTGCT GTAGCAGAAC CTGGTCCGGG CGCTGCACAC CGATCGCCGG
```

FIG. 7J

```
34801 CGGCGGTCTC GGCGCTTGGA ACGCTCGGTG TTAAAGTTGT AAAACAGCCA CTCTCTCAGA
34861 CCGTGCAGCA GATCTAGGGC CTCAGGAGTG ATGAAGATCC CATCATGCCT GATAGCTCTG
34921 ATCACATCGA CCACCGTGGA ATGGGCCAGG CCCAGCCAGA TGATGCAATT TTGTTGGGTT
34981 TCGGTGACGG CGGGGGAGGG AAGAACAGGA AGAACCATGA TTAACTTTTA ATCCAAACGG
35041 TCTCGGAGCA CTTCAAAATG AAGGTCACGG AGATGGCACC TCTCGCCCCC GCTGTGTTGG
35101 TGGAAAATAA CAGCCAGGTC AAAGGTGATA CGGTTCTCGA GATGTTCCAC GGTGGCTTCC
35161 AGCAAAGCCT CCACGCGCAC ATCAGAAACA AGACAATAGC GAAAGCGGGA GGGTTCTCTA
35221 ATTCCTCAAC CATCATGTTA CACTCCTGCA CCATCCCCAG ATAATTTTCA TTTTTCCAGC
35281 CTTGAATGAT TCGAACTAGT TCCTGAGGTA AATCCAAGCC AGCCATGATA AAAAGCTCGC
35341 GCAGAGCACC CTCCACCGGC ATTCTTAAGC ACACCCTCAT AATTCCAAGA TATTCTGCTC
35401 CTGGTTCACC TGCAGCAGAT TGACAAGCGG AATATCAAAA TCTCTGCCGC GATCCCTGAG
35461 CTCCTCCCTC AGCAATAACT GTAAGTACTC TTTCATATCG TCTCCGAAAT TTTTAGCCAT
35521 AGGACCCCCA GGAATAAGAG AAGGGCAAGC CACATTACAG ATAAACCGAA GTCCCCCCCA
35581 GTGAGCATTG CCAAATGTAA GATTGAAATA AGCATGCTGG CTAGACCCGG TGATATCTTC
35641 CAGATAACTG GACAGAAAAT CGGGTAAGCA ATTTTTAAGA AAATCAACAA AAGAAAAATC
35701 TTCCAGGTGC ACGTTTAGGG CCTCGGGAAC AACGATGGAG TAAGTGCAAG GGGTGCGTTC
35761 CAGCATGGTT AGTTAGCTGA TCTGTAAAAA AACAAAAAAT AAAACATTAA ACCATGCTAG
35821 CCTGGCGAAC AGGTGGGTAA ATCGTTCTCT CCAGCACCAG GCAGGCCACG GGGTCTCCGG
35881 CGCGACCCTC GTAAAAATTG TCGCTATGAT TGAAAACCAT CACAGAGAGA CGTTCCCGGT
35941 GGCCGGCGTG AATGATTCGA GAAGAAGCAT ACACCCCCCG GAACATTGGA GTCCGTGAGT
36001 GAAAAAAAGC GGCCGAGGAA GCAATGAGGC ACTACAACGC TCACTCTCAA GTCCAGCAAA
36061 GCGATGCCAT GCGGATGAAG CACAAAATTT TCAGGTGCGT AAAAAATGTA ATTACTCCCC
36121 TCCTGCACAG GCAGCGAAGC TCCCGATCCC TCCAGATACA CATACAAAGC CTCAGCGTCC
36181 ATAGCTTACC GAGCGGCAGC AGCAGCGGCA CACAACAGGC GCAAGAGTCA GAGAAAAGAC
36241 TGAGCTCTAA CCTGTCCGCC CGCTCTCTGC TCAATATATA GCCCCAGATC TACACTGACG
36301 TAAAGGCCAA AGTCTAAAAA TACCCGCCAA ATAATCACAC ACGCCCAGCA CACGCCCAGA
36361 AACCGGTGAC ACACTCAGAA AAATACGCGC ACTTCCTCAA ACGGCCAAAC TGCCGTCATT
36421 TCCGGGTTCC CACGCTACGT CATCAAAACA CGACTTTCAA ATTCCGTCGA CCGTTAAAAA
36481 CATCACCCGC CCCGCCCCTA ACGGTCGCCG CTCCCGCAGC CAATCACCTT CCTCCCTCCC
36541 CAAATTCAAA CAGCTCATTT GCATATTAAC GCGCACCAAA AGTTTGAGGT ATATTATTGA
36601 TGATGG  (SEQ ID NO: 3)
```

FIG. 7K

```
   1 CATCATCAAT AATATACCTC AAACTTTTGG TGCGCGTTAA TATGCAAATG AGCTGTTTGA
  61 ATTTGGGGAG GGAGGAAGGT GATTGGCCGA GAGACGGGCG ACCGTTAGGG GCGGGGCGGG
 121 TGACGTTTTG ATGACGTGGC CGTGAGGCGG AGCCGGTTTG CAAGTTCTCG TGGGAAAAGT
 181 GACGTCAAAC GAGGTGTGGT TTGAACACGG AAATACTCAA TTTTCCCGCG CTCTCTGACA
 241 GGAAATGAGG TGTTTCTGGG CGGATGCAAG TGAAAACGGG CCATTTTCGC GCGAAAACTG
 301 AATGAGGAAG TGAAAATCTG AGTAATTTCG CGTTTATGGC AGGGAGGAGT ATTTGCCGAG
 361 GGCCGAGTAG ACTTTGACCG ATTACGTGGG GGTTTCGATT ACCGTATTTT TCACCTAAAT
 421 TTCCGCGTAC GGTGTCAAAG TCCGGTGTTT TTACGTAGGC GTCAGCTGAT CGCCAGGGTA
 481 TTTAAACCTG CGCTCTCTAG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGTTTTCT
 541 CCTCCGCGCC GCGAGTCAGA TCTACACTTT GAAAGATGAG GCACCTGAGA GACCTGCCCG
 601 GTAATGTTTT CCTGGCTACT GGGAACGAGA TTCTGGAATT GGTGGTGGAC GCCATGATGG
 661 GTGGCGACCC TCCTGAGCCC CCTACCCCAT TTGAGGCGCC TTCGCTGTAC GATTTGTATG
 721 ATCTGGAGGT GGATGTGCCC GAGAACGACC CCAACGAGGA GGCGGTGAAT GATTTGTTTA
 781 GCGATGCCGC GCTGCTGGCT GCCGAGCAGG CTAATACGGA CTCTGGCTCA GACAGCGATT
 841 CCTCTCTCCA TACCCCGAGA CCCGGCAGAG GTGAGAAAAA GATCCCCGAG CTTAAAGGGG
 901 AAGAGCTCGA CCTGCGCTGC TATGAGGAAT GCTTGCCTCC GAGCGATGAT GAGGAGGACG
 961 AGGAGGCGAT TCGAGCTGCA TCGAACCAGG GAGTGAAAGC TGCGGGCGAA AGCTTTAGCC
1021 TGGACTGTCC TACTCTGCCC GGACACGGCT GTAAGTCTTG TGAATTTCAT CGCATGAATA
1081 CTGGAGATAA GAATGTGATG TGTGCCCTGT GCTATATGAG AGCTTACAAC CATTGTGTTT
1141 ACAGTAAGTG TGATTAACTT TAGTTGGGAA GGCAGAGGGT GACTGGGTGC TGACTGGTTT
1201 ATTTATGTAT ATGTTTTTTT ATGTGTAGGT CCCGTCTCTG ACGTAGATGA GACCCCCACT
1261 TCAGAGTGCA TTTCATCACC CCCAGAAATT GGCGAGGAAC CGCCCGAAGA TATTATTCAT
1321 AGACCAGTTG CAGTGAGAGT CACCGGGCGG AGAGCAGCTG TGGAGAGTTT GGATGACTTG
1381 CTACAGGGTG GGGATGAACC TTTGGACTTG TGTACCCGGA AACGCCCCAG GCACTAAGTG
1441 CCACACATGT GTGTTTACTT AAGGTGATGT CAGTATTTAT AGGGTGTGGA GTGCAATAAA
1501 ATCCGTGTTG ACTTTAAGTG CGTGGTTTAT GACTCAGGGG TGGGGACTGT GGGTATATAA
1561 GCAGGTGCAG ACCTGTGTGG TCAGTTCAGA GCAGGACTCA TGGAGATCTG GACGGTCTTG
1621 GAAGACTTTC ACCAGACTAG ACAGCTGCTA GAGAACTCAT CGGAGGGGGT CTCTTACCTG
1681 TGGAGATTCT GCTTCGGTGG GCCTCTAGCT AAGCTAGTCT ATAGGGCCAA ACAGGATTAT
1741 AAGGATCAAT TTGAGGATAT TTTGAGAGAG TGTCCTGGTA TTTTTGACTC TCTCAACTTG
1801 GGCCATCAGT CTCACTTTAA CCAGAGTATT CTGAGAGCCC TTGACTTTTC TACTCCTGGC
1861 AGAACTACCG CCGCGGTAGC CTTTTTTGCC TTTATCCTTG ACAAATGGAG TCAAGAAACC
1921 CATTTCAGCA GGGATTACCG TCTGGACTGC TTAGCAGTAG CTTTGTGGAG AACATGGAGG
1981 TGCCAGCGCC TGAATGCAAT CTCCGGCTAC TTGCCAGTAC AGCCGGTAGA CACGCTGAGG
2041 ATCCTGAGTC TCCAGTCACC CCAGGAACAC CAACGCCGCC AGCAGCCGCA GCAGGAGCAG
2101 CAGCAAGAGG AGGAGGAGGA TCGAGAAGAG AACCCGAGAG CCGGTCTGGA CCCTCCGGTG
2161 GCGGAGGAGG AGGAGTAGCT GACTTGTTTC CCGAGCTGCG CCGGGTGCTG ACTAGGTCTT
2221 CCAGTGGACG GGAGAGGGGG ATTAAGCGGG AGAGGCATGA GGAGACTAGC CACAGAACTG
2281 AACTGACTGT CAGTCTGATG AGCCGCAGGC GCCCAGAATC GGTGTGGTGG CATGAGGTTC
2341 AGTCGCAGGG GATAGATGAG GTCTCGGTGA TGCATGAGAA ATATTCCCTG GAACAAGTCA
2401 AGACTTGTTG GTTGGAGCCT GAGGATGATT GGGAGGTAGC CATCAGGAAT TATGCCAAGC
2461 TGGCTCTGAA GCCAGACAAG AAGTACAAGA TTACCAAACT GATTAATATC AGAAATTCCT
2521 GCTACATTTC AGGGAATGGG GCCGAGGTGG AGATCAGTAC CCAGGAGAGG GTGGCCTTCA
2581 GATGTTGTAT GATGAATATG TACCCGGGGG TGGTGGGCAT GGAGGGAGTC ACCTTTATGA
2641 ACGCGAGGTT CAGGGGTGAT GGGTATAATG GGGTGGTCTT TATGGCCAAC ACCAAGCTGA
2701 CAGTGCACGG ATGCTCCTTC TTTGGGTTCA ATAACATGTG CATCGAGGCC TGGGGCAGTG
2761 TTTCAGTGAG GGGATGCAGC TTTTCAGCCA ACTGGATGGG GGTCGTGGGC AGAACCAAGA
2821 GCAAGGTGTC AGTGAAGAAA TGCCTGTTCG AGAGGTGCCA CCTGGGGGTG ATGAGCGAGG
2881 GCGAAGCCAA AGTCAAACAC TGCGCCTCTA CTGAGACGGG CTGCTTTGTG CTGATCAAGG
2941 GCAATGCCCA AGTCAAGCAT AACATGATCT GTGGGGCCTC GGATGAGCGC GGCTACCAGA
3001 TGCTGACCTG CGCCGGTGGG AACAGCCATA TGCTGGCCAC CGTGCATGTG ACCTCGCACC
3061 CCCGCAAGAC ATGGCCCGAG TTCGAGCACA ACGTCATGAC CCGATGCAAT GTGCACCTGG
3121 GGTCCCGCCG AGGCATGTTC ATGCCCTACC AGTGCAACAT GCAATTTGTG AAGGTGCTGC
3181 TGGAGCCCGA TGCCATGTCC AGAGTGAGCC TGACGGGGGT GTTTGACATG AATGTGGAGC
3241 TGTGGAAAAT TCTGAGATAT GATGAATCCA AGACCAGGTG CCGGGCCTGC GAATGCGGAG
3301 GCAAGCACGC CAGGCTTCAG CCCGTGTGTG TGGAGGTGAC GGAGGACCTG CGACCCGATC
3361 ATTTGGTGTT GTCCTGCAAC GGGACGGAGT TCGGCTCCAG CGGGGAAGAA TCTGACTAGA
3421 GTGAGTAGTG TTTGGGGGAG GTGGAGGGCT TGTATGAGGG GCAGAATGAC TAAAATCTGT
```

FIG. 8A

3481 GTTTTTCTGT GTGTTGCAGC AGCATGAGCG GAAGCGCCTC CTTTGAGGGA GGGGTATTCA
3541 GCCCTTATCT GACGGGGCGT CTCCCCTCCT GGGCGGGAGT GCGTCAGAAT GTGATGGGAT
3601 CCACGGTGGA CGGCCGGCCC GTGCAGCCCG CGAACTCTTC AACCCTGACC TACGCGACCC
3661 TGAGCTCCTC GTCCGTGGAC GCAGCTGCCG CCGCAGCTGC TGCTTCCGCC GCCAGCGCCG
3721 TGCGCGGAAT GGCCCTGGGC GCCGGCTACT ACAGCTCTCT GGTGGCCAAC TCGACTTCCA
3781 CCAATAATCC CGCCAGCCTG AACGAGGAGA AGCTGCTGCT GCTGATGGCC CAGCTCGAGG
3841 CCCTGACCCA GCGCCTGGGC GAGCTGACCC AGCAGGTGGC TCAGCTGCAG GCGGAGACGC
3901 GGGCCGCGGT TGCCACGGTG AAAACCAAAT AAAAAATGAA TCAATAAATA AACGGAGACG
3961 GTTGTTGATT TTAACACAGA GTCTTGAATC TTTATTTGAT TTTTCGCGCG CGGTAGGCCC
4021 TGGACCACCG GTCTCGATCA TTGAGCACCC GGTGGATTTT TTCCAGGACC CGGTAGAGGT
4081 GGGCTTGGAT GTTGAGGTAC ATGGGCATGA GCCCGTCCCG GGGGTGGAGG TAGCTCCATT
4141 GCAGGGCCTC GTGCTCGGGG GTGGTGTTGT AAATCACCCA GTCATAGCAG GGGCGCAGGG
4201 CGTGGTGCTG CACGATGTCC TTGAGGAGGA GACTGATGGC CACGGGCAGC CCCTTGGTGT
4261 AGGTGTTGAC GAACCTGTTG AGCTGGGAGG GATGCATGCG GGGGGAGATG AGATGCATCT
4321 TGGCCTGGAT CTTGAGATTG GCGATGTTCC CGCCCAGATC CCGCCGGGGG TTCATGTTGT
4381 GCAGGACCAC CAGCACGGTG TATCCGGTGC ACTTGGGGAA TTTGTCATGC AACTTGGAAG
4441 GGAAGGCGTG AAAGAATTTG GAGACGCCCT TGTGACCGCC CAGGTTTTCC ATGCACTCAT
4501 CCATGATGAT GGCGATGGGC CCGTGGGCGG CGGCCTGGGC AAAGACGTTT CGGGGGTCGG
4561 ACACATCGTA GTTGTGGTCC TGGGTGAGCT CGTCATAGGC CATTTTAATG AATTTGGGGC
4621 GGAGGGTGCC CGACTGGGGG ACGAAGGTGC CCTCGATCCC GGGGGCGTAG TTGCCCTCGC
4681 AGATCTGCAT CTCCCAGGCC TTGAGCTCGG AGGGGGGGAT CATGTCCACC TGCGGGGCGA
4741 TGAAAAAAAC GGTTTCCGGG GCGGGGGAGA TGAGCTGGGC CGAAAGCAGG TTCCGGAGCA
4801 GCTGGGACTT GCCGCAGCCG GTGGGGCCGT AGATGACCCC GATGACCGGC TGCAGGTGGT
4861 AGTTGAGGGA GAGACAGCTG CCGTCCTCGC GGAGGAGGGG GGCCACCTCG TTCATCATCT
4921 CGCGCACATG CATGTTCTCG CGCACGAGTT CCGCCAGGAG GCGCTCGCCC CCCAGCGAGA
4981 GGAGCTCTTG CAGCGAGGCG AAGTTTTTCA GCGGCTTGAG CCCGTCGGCC ATGGGCATTT
5041 TGGAGAGGGT CTGTTGCAAG AGTTCCAGAC GGTCCCAGAG CTCGGTGATG TGCTCTAGGG
5101 CATCTCGATC CAGCAGACCT CCTCGTTTCG CGGGTTGGGG CGACTGCGGG AGTAGGGCAC
5161 CAGGCGATGG GCGTCCAGCG AGGCCAGGGT CCGGTCCTTC CAGGGTCGCA GGGTCCGCGT
5221 CAGCGTGGTC TCCGTCACGG TGAAGGGGTG CGCGCCGGGC TGGGCGCTTG CGAGGGTGCG
5281 CTTCAGGCTC ATCCGGCTGG TCGAGAACCG CTCCCGGTCG GCGCCCTGCG CGTCGGCCAG
5341 GTAGCAATTG AGCATGAGTT CGTAGTTGAG CGCCTCGGCC GCGTGGCCCT TGGCGCGGAG
5401 CTTACCTTTG GAAGTGTGTC CGCAGACGGG ACAGAGGAGG GACTTGAGGG CGTAGAGCTT
5461 GGGGGCGAGG AAGACGGACT CGGGGGCGTA GGCGTCCGCG CCGCAGCTGG CGCAGACGGT
5521 CTCGCACTCC ACGAGCCAGG TGAGGTCGGG CCGGTTGGGG TCAAAAACGA GGTTTCCTCC
5581 GTGCTTTTTG ATGCGTTTCT TACCTCTGGT CTCCATGAGC TCGTGTCCCC GCTGGGTGAC
5641 AAAGAGGCTG TCCGTGTCCC CGTAGACCGA CTTTATGGGC CGGTCCTCGA GCGGGGTGCC
5701 GCGGTCCTCG TCGTAGAGGA ACCCCGCCCA CTCCGAGACG AAGGCCCGGG TCCAGGCCAG
5761 CACGAAGGAG GCCACGTGGG AGGGGTAGCG GTCGTTGTCC ACCAGCGGGT CCACCTTCTC
5821 CAGGGTATGC AAGCACATGT CCCCCTCGTC CACATCCAGG AAGGTGATTG GCTTGTAAGT
5881 GTAGGCCACG TGACCGGGGG TCCCGGCCGG GGGGGTATAA AAGGGGGCGG GCCCCTGCTC
5941 GTCCTCACTG TCTTCCGGAT CGCTGTCCAG GAGCGCCAGC TGTTGGGGTA GGTATTCCCT
6001 CTCGAAGGCT GGCATAACCT CGGCACTCAG GTTGTCAGTT TCTAGAAACG AGGAGGATTT
6061 GATATTGACG GTGCCGTTGG AGACGCCTTT CATGAGCCCC TCGTCCATCT GGTCAGAAAA
6121 GACGATCTTT TTGTTGTCGA GCTTGGTGGC GAAGGAGCCG TAGAGGGCGT TGGAGAGGAG
6181 CTTGGCGATG GAGCGCATGG TCTGGTTCTT TTCCTTGTCG GCGCGCTCCT TGGCGGCGAT
6241 GTTGAGCTGC ACGTACTCGC GCGCCACGCA CTTCCATTCG GGGAAGACGG TGGTGAGCTC
6301 GTCGGGCACG ATTCTGACCC GCCAGCCGCG GTTGTGCAGG GTGATGAGGT CCACGCTGGT
6361 GGCCACCTCG CCGCGCAGGG GCTCGTTGGT CCAGCAGAGG CGCCCGCCCT TGCGCGAGCA
6421 GAAGGGGGGC AGCGGGTCCA GCATGAGCTC GTCGGGGGGG TCGGCGTCCA CGGTGAAGAT
6481 GCCGGGCAGA AGCTCGGGGT CGAAGTAGCT GATGCAGGTG TCCAGATCGT CCAGCGCCGC
6541 TTGCCAGTCG CGCACGGCCA GCGCGCGCTC GTAGGGGCTG AGGGGCGTGC CCCAGGGCAT
6601 GGGGTGCGTG AGCGCGGAGG CGTACATGCC GCAGATGTCG TAGACGTAGA GGGGCTCCTC
6661 GAGGACGCCG ATGTAGGTGG GGTAGCAGCG CCCCCCGCGG ATGCTGGCGC GCACGTAGTC
6721 GTACAGCTCG TGCGAGGGCG CGAGGAGCCC CGTGCCGAGG TTGGAGCGTT GCGGCTTTTC
6781 GGCGCGGTAG ACGATCTGGC GGAAGATGGC GTGGGAGTTG GAGGAGATGG TGGGCCTCTG
6841 GAAGATGTTG AAGTGGGCGT GGGGCAGGCC GACCGAGTCC CTGATGAAGT GGGCGTAGGA
6901 GTCCTGCAGC TTGGCGACGA GCTCGGCGGT GACGAGGACG TCCAGGGCGC AGTAGTCGAG

FIG. 8B

```
 6961 GGTCTCTTGG ATGATGTCGT ACTTGAGCTG GCCCTTCTGC TTCCACAGCT CGCGGTTGAG
 7021 AAGGAACTCT TCGCGGTCCT TCCAGTACTC TTCGAGGGGG AACCCGTCCT GATCGGCACG
 7081 GTAAGAGCCC ACCATGTAGA ACTGGTTGAC GGCCTTGTAG GCGCAGCAGC CCTTCTCCAC
 7141 GGGGAGGGCG TAAGCTTGTG CGGCCTTGCG CAGGGAGGTG TGGGTGAGGG CGAAGGTGTC
 7201 GCGCACCATG ACCTTGAGGA ACTGGTGCTT GAAGTCGAGG TCGTCGCAGC CGCCCTGCTC
 7261 CCAGAGCTGG AAGTCCGTGC GCTTCTTGTA GGCGGGGTTG GGCAAAGCGA AAGTAACATC
 7321 GTTGAAGAGG ATCTTGCCCG CGCGGGGCAT GAAGTTGCGA GTGATGCGGA AAGGCTGGGG
 7381 CACCTCGGCC CGGTTGTTGA TGACCTGGGC GGCGAGGACG ATCTCGTCGA AGCCGTTGAT
 7441 GTTGTGCCCG ACGATGTAGA GTTCCACGAA TCGCGGGCGG CCCTTAACGT GGGGCAGCTT
 7501 CTTGAGCTCG TCGTAGGTGA GCTCGGCGGG GTCGCTGAGC CCGTGCTGCT CGAGGGCCCA
 7561 GTCGGCGACG TGGGGGTTGG CGCTGAGGAA GGAAGTCCAG AGATCCACGG CCAGGGCGGT
 7621 CTGCAAGCGG TCCCGGTACT GACGGAACTG CTGGCCCACG GCCATTTTTT CGGGGGTGAC
 7681 GCAGTAGAAG GTGCGGGGGT CGCCGTGCCA GCGGTCCCAC TTGAGCTGGA GGGCGAGGTC
 7741 GTGGGCGAGC TCGACGAGCG GCGGGTCCCC GGAGAGTTTC ATGACCAGCA TGAAGGGGAC
 7801 GAGCTGCTTG CCGAAGGACC CCATCCAGGT GTAGGTTTCC ACATCGTAGG TGAGGAAGAG
 7861 CCTTTCGGTG CGAGGATGCG AGCCGATGGG GAAGAACTGG ATCTCCTGCC ACCAGTTGGA
 7921 GGAATGGCTG TTGATGTGAT GGAAGTAGAA ATGCCGACGG CGCGCCGAGC ACTCGTGCTT
 7981 GTGTTTATAC AAGCGTCCGC AGTGCTCGCA ACGCTGCACG GGATGCACGT GCTGCACGAG
 8041 CTGTACCTGG GTTCCTTTGA CGAGGAATTT CAGTGGGCAG TGGAGCGCTG GCGGCTGCAT
 8101 CTGGTGCTGT ACTACGTCCT GGCCATCGGC GTGGCCATCG TCTGCCTCGA TGGTGGTCAT
 8161 GCTGACGAGC CCGCGCGGGA GGCAGGTCCA GACTTCGGCT CGGACGGGTC GGAGAGCGAG
 8221 GACGAGGGCG CGCAGGCCGG AGCTGTCCAG GGTCCTGAGA CGCTGCGGAG TCAGGTCAGT
 8281 GGGCAGCGGC GGCGCGCGGT TGACTTGCAG GAGCTTTTCC AGGGCGCGCG GGAGGTCCAG
 8341 ATGGTACTTG ATCTCCACGG CGCCGTTGGT GGCGACGTCC ACGGCTTGCA GGGTCCCGTG
 8401 CCCCTGGGGC GCCACCACCG TGCCCCGTTT CTTCTTGGGC GCTGCTTCCA TGCCGGTCAG
 8461 AAGCGGCGGC GAGGACGCGC GCCGGGCGGC AGGGGCGGCT CGGGACCCGG AGGCAGGGGC
 8521 GGCAGGGGCA CGTCGGCGCC GCGCGCGGGC AGGTTCTGGT ACTGCGCCCG GAGAAGACTG
 8581 GCGTGAGCGA CGACGCGACG GTTGACGTCC TGGATCTGAC GCCTCTGGGT GAAGGCCACG
 8641 GGACCCGTGA GTTTGAACCT GAAAGAGAGT TCGACAGAAT CAATCTCGGT ATCGTTGACG
 8701 GCGGCCTGCC GCAGGATCTC TTGCACGTCG CCCGAGTTGT CCTGGTAGGC GATCTCGGTC
 8761 ATGAACTGCT CGATCTCCTC CTCCTGAAGG TCTCCGCGGC CGGCGCGCTC GACGGTGGCC
 8821 GCGAGGTCGT TGGAGATGCG GCCCATGAGC TGCGAGAAGG CGTTCATGCC GGCCTCGTTC
 8881 CAGACGCGGC TGTAGACCAC GGCTCCGTCG GGGTCGCGCG CGCGCATGAC CACCTGGGCG
 8941 AGGTTGAGCT CGACGTGGCG CGTGAAGACC GCGTAGTTGC AGAGGCGCTG GTAGAGGTAG
 9001 TTGAGCGTGG TGGCGATGTG CTCGGTGACG AAGAAGTACA TGATCCAGCG GCGGAGCGGC
 9061 ATCTCGCTGA CGTCGCCCAG GGCTTCCAAG CGCTCCATGG CCTCGTAGAA GTCCACGGCG
 9121 AAGTTGAAAA ACTGGGAGTT GCGCGCCGAG ACGGTCAACT CCTCCTCCAG AAGACGGATG
 9181 AGCTCAGCGA TGGTGGCGCG CACCTCGCGC TCGAAGGCCC CGGGGGGCTC CTCTTCTTCC
 9241 ATCTCTTCCT CCTCCACTAA CATCTCTTCT ACTTCCTCCT CAGGAGGCGG CGGCGGGGGA
 9301 GGGGCCCTGC GTCGCCGGCG GCGCACGGGC AGACGGTCGA TGAAGCGCTC GATGGTCTCC
 9361 CCGCGCCGGC GACGCATGGT CTCGGTGACG GCGCGCCCGT CCTCGCGGGG CCGCAGCGTG
 9421 AAGACGCCGC CGCGCATCTC CAGGTGGCCG CCGGGGGGGT CTCCGTTGGG CAGGGAGAGG
 9481 GCGCTGACGA TGCATCTTAT CAATTGGCCC GTAGGGACTC CGCGCAAGGA CCTGAGCGTC
 9541 TCGAGATCCA CGGGATCCGA AAACCGCTGA ACGAAGGCTT CGAGCCAGTC GCAGTCGCAA
 9601 GGTAGGCTGA GCCCGGTTTC TTGTTCTTCG GGGATTTCGG GAGGCGGGCG GGCGATGCTG
 9661 CTGGTGATGA AGTTGAAGTA GGCGGTCCTG AGACGGCGGA TGGTGGCGAG GAGCACCAGG
 9721 TCCTTGGGCC CGGCTTGCTG GATGCGCAGA CGGTCGGCCA TGCCCCAGGC GTGGTCCTGA
 9781 CACCTGGCGA GGTCCTTGTA GTAGTCCTGC ATGAGCCGCT CCACGGGCAC CTCCTCCTCG
 9841 CCCGCGCGGC CGTGCATGCG CGTGAGCCCG AACCCGCGCT GGGGCTGGAC GAGCGCCAGG
 9901 TCGGCGACGA CGCGCTCGGC GAGGATGGCC TGCTGTATCT GGGTGAGGGT GGTCTGGAAG
 9961 TCGTCGAAGT CGACGAAGCG GTGGTAGGCT CCGGTGTTGA TGGTATAGGA GCAGTTGGCC
10021 ATGACGGACC AGTTGACGGT CTGGTGGCCG GGTCGCACGA GCTCGTGGTA CTTGAGGCGC
10081 GAGTAGGCGC GCGTGTCGAA GATGTAGTCG TTGCAGGTGC GCACGAGGTA CTGGTATCCG
10141 ACGAGGAAGT GCGGCGGCGG CTGGCGGTAG AGCGGCCATC GCTCGGTGGC GGGGGCGCCG
10201 GGCGCGAGGT CCTCGAGCAT GAGGCGGTGG TAGCCGTAGA TGTACCTGGA CATCCAGGTG
10261 ATGCCGGCGG CGGTGGTGGA GGCGCGCGGG AACTCGCGGA CGCGGTTCCA GATGTTGCGC
10321 AGCGGCAGGA AGTAGTTCAT GGTGGCCGCG GTCTGGCCCG TGAGGCGCGC GCAGTCGTGG
10381 ATGCTCTAGA CATACGGGCA AAAACGAAAG CGGTCAGCGG CTCGACTCCG TGGCCTGGAG
```

FIG. 8C

10441 GCTAAGCGAA CGGGTTGGGC TGCGCGTGTA CCCCGGTTCG AATCTCGAAT CAGGCTGGAG
10501 CCGCAGCTAA CGTGGTACTG GCACTCCCGT CTCGACCCAA GCCTGCTAAC GAAACCTCCA
10561 GGATACGGAG GCGGGTCGTT TTTTGGCCTT GGTCGCTGGT CATGAAAAAC TAGTAAGCGC
10621 GGAAAGCGAC CGCCCGCGAT GGCTCGCTGC CGTAGTCTGG AGAAAGAATC GCCAGGGTTG
10681 CGTTGCGGTG TGCCCCGGTT CGAGCCTCAG CGCTCGGCGC CGGCCGGATT CCGCGGCTAA
10741 CGTGGGCGTG GCTGCCCCGT CGTTTCCAAG ACCCCTTAGC CAGCCGACTT CTCCAGTTAC
10801 GGAGCGAGCC CCTCTTTTTC TTGTGTTTTT GCCAGATGCA TCCCGTACTG CGGCAGATGC
10861 GCCCCCACCC TCCACCTCAA CCGCCCCTAC CGCCGCAGCA GCAGCAACAG CCGGCGCTTC
10921 TGCCCCCGCC CCAGCAGCAG CCAGCCACTA CCGCGGCGGC CGCCGTGAGC GGAGCCGGCG
10981 TTCAGTATGA CCTGGCCTTG GAAGAGGGCG AGGGGCTGGC GCGGCTGGGG GCGTCGTCGC
11041 CGGAGCGGCA CCCGCGCGTG CAGATGAAAA GGGACGCTCG CGAGGCCTAC GTGCCCAAGC
11101 AGAACCTGTT CAGAGACAGG AGCGGCGAGG AGCCCGAGGA GATGCGCGCC TCCCGCTTCC
11161 ACGCGGGGCG GGAGCTGCGG CGCGGCCTGG ACCGAAAGCG GGTGCTGAGG GACGAGGATT
11221 TCGAGGCGGA CGAGCTGACG GGGATCAGCC CCGCGCGCGC GCACGTGGCC GCGGCCAACC
11281 TGGTCACGGC GTACGAGCAG ACCGTGAAGG AGGAGAGCAA CTTCCAAAAA TCCTTCAACA
11341 ACCACGTGCG CACGCTGATC GCGCGCGAGG AGGTGACCCT GGGCCTGATG CACCTGTGGG
11401 ACCTGCTGGA GGCCATCGTG CAGAACCCCA CGAGCAAGCC GCTGACGGCG CAGCTGTTTC
11461 TGGTGGTGCA GCACAGTCGG GACAACGAGA CGTTCAGGGA GGCGCTGCTG AATATCACCG
11521 AGCCCGAGGG CCGCTGGCTC CTGGACCTGG TGAACATTCT GCAGAGCATC GTGGTGCAGG
11581 AGCGCGGGCT GCCGCTGTCC GAGAAGCTGG CGGCTATCAA CTTCTCGGTG CTGAGCCTGG
11641 GCAAGTACTA CGCTAGGAAG ATCTACAAGA CCCCGTACGT GCCCATAGAC AAGGAGGTGA
11701 AGATCGACGG GTTTTACATG CGCATGACCC TGAAAGTGCT GACCCTGAGC GACGATCTGG
11761 GGGTGTACCG CAACGACAGG ATGCACCGCG CGGTGAGCGC CAGCCGCCGG CGCGAGCTGA
11821 GCGACCAGGA GCTGATGCAC AGCCTGCAGC GGGCCCTGAC CGGGGCCGGG ACCGAGGGGG
11881 AGAGCTACTT TGACATGGGC GCGGACCTGC GCTGGCAGCC CAGCCGCCGG GCCTTGGAAG
11941 CTGCCGGCGG TTCCCCCTAC GTGGAGGAGG TGGACGATGA GGAGGAGGAG GGCGAGTACC
12001 TGGAAGACTG ATGGCGCGAC CGTATTTTTG CTAGATGCAG CAACAGCCAC CGCCTCCTGA
12061 TCCCGCGATG CGGGCGGCGC TGCAGAGCCA GCCGTCCGGC ATTAACTCCT CGGACGATTG
12121 GACCCAGGCC ATGCAACGCA TCATGGCGCT GACGACCCGC AATCCCGAAG CCTTTAGACA
12181 GCAGCCTCAG GCCAACCGGC TCTCGGCCAT CCTGGAGGCC GTGGTGCCCT CGCGCTCGAA
12241 CCCCACGCAC GAGAAGGTGC TGGCCATCGT GAACGCGCTG GTGGAGAACA AGGCCATCCG
12301 CGGCGACGAG GCCGGGCTGG TGTACAACGC GCTGCTGGAG CGCGTGGCCC GCTACAACAG
12361 CACCAACGTG CAGACGAACC TGGACCGCAT GGTGACCGAC GTGCGCGAGG CGGTGTCGCA
12421 GCGCGAGCGG TTCCACCGCG AGTCGAACCT GGGCTCCATG GTGGCGCTGA ACGCCTTCCT
12481 GAGCACGCAG CCCGCCAACG TGCCCCGGGG CCAGGAGGAC TACACCAACT TCATCAGCGC
12541 GCTGCGGCTG ATGGTGGCCG AGGTGCCCCA GAGCGAGGTG TACCAGTCGG GGCCGGACTA
12601 CTTCTTCCAG ACCAGTCGCC AGGGCTTGCA GACCGTGAAC CTGAGCCAGG CTTTCAAGAA
12661 CTTGCAGGGA CTGTGGGGCG TGCAGGCCCC GGTCGGGGAC CGCGCGACGG TGTCGAGCCT
12721 GCTGACGCCG AACTCGCGCC TGCTGCTGCT GCTGGTGGCG CCCTTCACGG ACAGCGGCAG
12781 CGTGAGCCGC GACTCGTACC TGGGCTACCT GCTTAACCTG TACCGCGAGG CCATCGGGCA
12841 GGCGCACGTG GACGAGCAGA CCTACCAGGA GATCACCCAC GTGAGCCGCG CGCTGGGCCA
12901 GGAGGACCCG GGCAACCTGG AGGCCACCCT GAACTTCCTG CTGACCAACC GGTCGCAGAA
12961 GATCCCGCCC CAGTACGCGC TGAGCACCGA GGAGGAGCGC ATCCTGCGCT ACGTGCAGCA
13021 GAGCGTGGGG CTGTTCCTGA TGCAGGAGGG GGCCACGCCC AGCGCCGCGC TCGACATGAC
13081 CGCGCGCAAC ATGGAGCCCA GCATGTACGC TCGCAACCGC CCGTTCATCA ATAAGCTGAT
13141 GGACTACTTG CATCGGGCGG CCGCCATGAA CTCGGACTAC TTTACCAACG CCATCTTGAA
13201 CCCGCACTGG CTCCCGCCGC CCGGGTTCTA CACGGGCGAG TACGACATGC CCGACCCCAA
13261 CGACGGGTTC CTGTGGGACG ACGTGGACAG CAGCGTGTTC TCGCCGCGCC CCGCCACCAC
13321 CGTGTGGAAG AAAGAGGGCG GGGACCGGCG GCCGTCCTCG GCGCTGTCCG GTCGCGCGGG
13381 TGCTGCCGCG GCGGTGCCTG AGGCCGCCAG CCCCTTCCCG AGCCTGCCCT TTTCGCTGAA
13441 CAGCGTGCGC AGCAGCGAGC TGGGTCGGCT GACGCGGCCG CGCCTGCTGG GCGAGGAGGA
13501 GTACCTGAAC GACTCCTTGT TGAGGCCCGA GCGCGAGAAG AACTTCCCCA ATAACGGGAT
13561 AGAGAGCCTG GTGGACAAGA TGAGCCGCTG GAAGACGTAC GCGCACGAGC ACAGGGACGA
13621 GCCCCGAGCT AGCAGCAGCG CAGGCACCCG TAGACGCCAG CGACACGACA GGCAGCGGGG
13681 TCTGGTGTGG GACGATGAGG ATTCCGCCGA CGACAGCAGC GTGTTGGACT TGGGTGGGAG
13741 TGGTGGTGGT AACCCGTTCG CTCACTTGCG CCCCCGTATC GGGCGCCTGA TGTAAGAATC
13801 TGAAAAAATA AAAAACGGTA CTCACCAAGG CCATGGCGAC CAGCGTGCGT TCTTCTCTGT
13861 TGTTTGTAGT AGTATGATGA GGCGCGTGTA CCCGGAGGGT CCTCCTCCCT CGTACGAGAG

FIG. 8D

```
13921 CGTGATGCAG CAGGCGGTGG CGGCGGCGAT GCAGCCCCCG CTGGAGGCGC CTTACGTGCC
13981 CCCGCGGTAC CTGGCGCCTA CGGAGGGGCG GAACAGCATT CGTTACTCGG AGCTGGCACC
14041 CTTGTACGAT ACCACCCGGT TGTACCTGGT GGACAACAAG TCGGCGGACA TCGCCTCGCT
14101 GAACTACCAG AACGACCACA GCAACTTCCT GACCACCGTG GTGCAGAACA ACGATTTCAC
14161 CCCCACGGAG GCCAGCACCC AGACCATCAA CTTTGACGAG CGCTCGCGGT GGGGCGGCCA
14221 GCTGAAAACC ATCATGCACA CCAACATGCC CAACGTGAAC GAGTTCATGT ACAGCAACAA
14281 GTTCAAGGCG CGGGTGATGG TCTCGCGCAA GACCCCCAAT GGGGTCGCGG TGGATGAGAA
14341 TTATGATGGT AGTCAGGACG AGCTGACTTA CGAGTGGGTG GAGTTTGAGC TGCCCGAGGG
14401 CAACTTCTCG GTGACCATGA CCATCGATCT GATGAACAAC GCCATCATCG ACAACTACTT
14461 GGCGGTGGGG CGTCAGAACG GGGTGCTGGA GAGCGACATC GGCGTGAAGT TCGACACGCG
14521 CAACTTCCGG CTGGGCTGGG ACCCCGTGAC CGAGCTGGTG ATGCCGGGCG TGTACACCAA
14581 CGAGGCCTTC CACCCCGACA TCGTCCTGCT GCCCGGCTGC GGCGTGGACT TCACCGAGAG
14641 CCGCCTCAGC AACCTGCTGG GCATCCGCAA GCGGCAGCCC TTCCAGGAGG GCTTCCAGAT
14701 CCTGTACGAG GACCTGGAGG GGGGCAACAT CCCCGCGCTC TTGGATGTCG AAGCCTATGA
14761 GAAAAGCAAG GAGGAGGCCG CCGCAGCGGC GACCGCAGCC GTGGCCACCG CCTCTACCGA
14821 GGTGCGGGGC GATAATTTTG CTAGCGCCGC GGCAGTGGCC GAGGCGGCTG AAACCGAAAG
14881 TAAGATAGTC ATCCAGCCGG TGGAGAAGGA CAGCAAGGAC AGGAGCTACA ACGTGCTCGC
14941 GGACAAGAAA AACACCGCCT ACCGCAGCTG GTACCTGGCC TACAACTACG GCGACCCCGA
15001 GAAGGGCGTG CGCTCCTGGA CGCTGCTCAC CACCTCGGAC GTCACCTGCG GCGTGGAGCA
15061 AGTCTACTGG TCGCTGCCCG ACATGATGCA AGACCCGGTC ACCTTCCGCT CCACGCGTCA
15121 AGTTAGCAAC TACCCGGTGG TGGGCGCCGA GCTCCTGCCC GTCTACTCCA AGAGCTTCTT
15181 CAACGAGCAG GCCGTCTACT CGCAGCAGCT GCGCGCCTTC ACCTCGCTCA CGCACGTCTT
15241 CAACCGCTTC CCCGAGAACC AGATCCTCGT CCGCCCGCCC GCGCCCACCA TTACCACCGT
15301 CAGTGAAAAC GTTCCTGCTC TCACAGATCA CGGGACCCTG CCGCTGCGCA GCAGTATCCG
15361 GGGAGTCCAG CGCGTGACCG TCACTGACGC CAGACGCCGC ACCTGCCCCT ACGTCTACAA
15421 GGCCCTGGGC GTAGTCGCGC CGCGCGTCCT CTCGAGCCGC ACCTTCTAAA AAATGTCCAT
15481 TCTCATCTCG CCCAGTAATA ACACCGGTTG GGGCCTGCGC GCGCCCAGCA AGATGTACGG
15541 AGGCGCTCGC CAACGCTCCA CGCAACACCC CGTGCGCGTG CGCGGGCACT TCCGCGCTCC
15601 CTGGGGCGCC CTCAAGGGCC GCGTGCGCTC GCGCACCACC GTCGACGACG TGATCGACCA
15661 GGTGGTGGCC GACGCGCGCA ACTACACGCC CGCCGCCGCG CCCGCCTCCA CCGTGGACGC
15721 CGTCATCGAC AGCGTGGTGG CCGATGCGCG CCGGTACGCC CGCGCCAAGA GCCGGCGGCG
15781 GCGCATCGCC CGGCGGCACC GGAGCACCCC CGCCATGCGC GCGGCGCGAG CCTTGCTGCG
15841 CAGGGCCAGG CGCACGGGAC GCAGGGCCAT GCTCAGGGCG GCCAGACGCG CGGCCTCCGG
15901 CAGCAGCAGC GCCGGCAGGA CCCGCAGACG CGCGGCCACG GCGGCGGCGG CGGCCATCGC
15961 CAGCATGTCC CGCCCGCGGC GCGGCAACGT GTACTGGGTG CGCGACGCCG CCACCGGTGT
16021 GCGCGTGCCC GTGCGCACCC GCCCCCCTCG CACTTGAAGA TGCTGACTTC GCGATGTTGA
16081 TGTGTCCCAG CGGCGAGGAG GATGTCCAAG CGCAAATACA AGGAAGAGAT GCTCCAGGTC
16141 ATCGCGCCTG AGATCTACGG CCCCGCGGTG AAGGAGGAAA GAAAGCCCCG CAAACTGAAG
16201 CGGGTCAAAA AGGACAAAAA GGAGGAGGAA GATGTGGACG GACTGGTGGA GTTTGTGCGC
16261 GAGTTCGCCC CCCGGCGGCG CGTGCAGTGG CGCGGGCGGA AAGTGAAACC GGTGCTGCGG
16321 CCCGGCACCA CGGTGGTCTT CACGCCCGGC GAGCGTTCCG GCTCCGCCTC CAAGCGCTCC
16381 TACGACGAGG TGTACGGGGA CGAGGACATC CTCGAGCAGG CGGTCGAGCG TCTGGGCGAG
16441 TTTGCTTACG GCAAGCGCAG CCGCCCCGCG CCCTTGAAAG AGGAGGCGGT GTCCATCCCG
16501 CTGGACCACG GCAACCCCAC GCCGAGCCTG AAGCCGGTGA CCCTGCAGCA GGTGCTGCCG
16561 AGCGCGGCGC CGCGCCGGGG CTTCAAGCGC GAGGGCGGCG AGGATCTGTA CCCGACCATG
16621 CAGCTGATGG TGCCCAAGCG CCAGAAGCTG GAGGACGTGC TGGAGCACAT GAAGGTGGAC
16681 CCCGAGGTGC AGCCCGAGGT CAAGGTGCGG CCCATCAAGC AGGTGGCCCC GGGCCTGGGC
16741 GTGCAGACCG TGGACATCAA GATCCCCACG GAGCCCATGG AAACGCAGAC CGAGCCCGTG
16801 AAGCCCAGCA CCAGCACCAT GGAGGTGCAG ACGGATCCCT GGATGCCGGC GCCGGCTTCC
16861 ACCACTCGCC GAAGACGCAA GTACGGCGCG GCCAGCCTGC TGATGCCCAA CTACGCGCTG
16921 CATCCTTCCA TCATCCCCAC GCCGGGCTAC CGCGGCACGC GCTTCTACCG CGGCTACACC
16981 AGCAGCCGCC GCAAGACCAC CACCCGCCGC CGCCGTCGTC GCACCCGCCG CAGCAGCACC
17041 GCGACTTCCG CCGCCGCCCT GGTGCGGAGA GTGTACCGCA GCGGGCGCGA GCCTCTGACC
17101 CTGCCGCGCG CGCGCTACCA CCCGAGCATC GCCATTTAAC TCTGCCGTCG CCTCCTACTT
17161 GCAGATATGG CCCTCACATG CCGCCTCCGC GTCCCCATTA CGGGCTACCG AGGAAGAAAG
17221 CCGCGCCGTA GAAGGCTGAC GGGGAACGGG CTGCGTCGCC ATCACCACCG GCGGCGGCGC
17281 GCCATCAGCA AGCGGTTGGG GGGAGGCTTC CTGCCCGCGC TGATCCCCAT CATCGCCGCG
17341 GCGATCGGGG CGATCCCCGG CATAGCTTCC GTGGCGGTGC AGGCCTCTCA GCGCCACTGA
```

FIG. 8E

17401 GACACAGCTT GGAAAATTTG TAATAAAAAA ATGGACTGAC GCTCCTGGTC CTGTGATGTG
17461 TGTTTTTAGA TGGAAGACAT CAATTTTTCG TCCCTGGCAC CGCGACACGG CACGCGGCCG
17521 TTTATGGGCA CCTGGAGCGA CATCGGCAAC AGCCAACTGA ACGGGGGCGC CTTCAATTGG
17581 AGCAGTCTCT GGAGCGGGCT TAAGAATTTC GGGTCCACGC TCAAAACCTA TGGCAACAAG
17641 GCGTGGAACA GCAGCACAGG GCAGGCGCTG AGGGAAAAGC TGAAAGAGCA GAACTTCCAG
17701 CAGAAGGTGG TCGATGGCCT GGCCTCGGGC ATCAACGGGG TGGTGGACCT GGCCAACCAG
17761 GCCGTGCAGA AACAGATCAA CAGCCGCCTG GACGCGGTCC CGCCCGCGGG GTCCGTGGAG
17821 ATGCCCCAGG TGGAGGAGGA GCTGCCTCCC CTGGACAAGC GCGGCGACAA GCGACCGCGT
17881 CCCGACGCGG AGGAGACGCT GCTGACGCAC ACGGACGAGC CGCCCCCGTA CGAGGAGGCG
17941 GTGAAACTGG GTCTGCCCAC CACGCGGCCC GTGGCGCCTC TGGCCACCGG GGTGCTGAAA
18001 CCCAGCAGCA GCAGCCAGCC CGCGACCCTG GACTTGCCTC CGCCTGCTTC CCGCCCCTCC
18061 ACAGTGGCTA AGCCCCTGCC GCCGGTGGCC GTCGCGTCGC GCGCCCCCCG AGGCCGCCCC
18121 CAGGCGAACT GGCAGAGCAC TCTGAACAGC ATCGTGGGTC TGGGAGTGCA GAGTGTGAAG
18181 CGCCGCCGCT GCTATTAAAA GACACTGTAG CGCTTAACTT GCTTGTCTGT GTGTATATGT
18241 ATGTCCGCCG ACCAGAAGGA GGAAGAGGCG CGTCGCCGAG TTGCAAGATG GCCACCCCAT
18301 CGATGCTGCC CCAGTGGGCG TACATGCACA TCGCCGGACA GGACGCTTCG GAGTACCTGA
18361 GTCCGGGTCT GGTGCAGTTC GCCCGCGCCA CAGACACCTA CTTCAGTCTG GGGAACAAGT
18421 TTAGGAACCC CACGGTGGCG CCCACGCACG ATGTGACCAC CGACCGCAGC CAGCGGCTGA
18481 CGCTGCGCTT CGTGCCCGTG GACCGCGAGG ACAACACCTA CTCGTACAAA GTGCGCTACA
18541 CGCTGGCCGT GGGCGACAAC CGCGTGCTGG ACATGGCCAG CACCTACTTT GACATCCGCG
18601 GCGTGCTGGA TCGGGGGCCC AGCTTCAAAC CCTACTCCGG CACCGCCTAC AACAGCCTGG
18661 CTCCCAAGGG AGCGCCCAAC ACTTGCCAGT GGACATATAA AGCTGGTGAT ACTGATACAG
18721 AAAAAACCTA TACATATGGA AATGCACCTG TGCAAGGCAT TAGCATTACA AAGGATGGTA
18781 TTCAACTTGG AACTGACAGC GATGGTCAGG CAATCTATGC AGACGAAACT TATCAACCAG
18841 AGCCTCAAGT GGGTGATGCT GAATGGCATG ACATCACTGG TACTGATGAA AAATATGGAG
18901 GCAGAGCTCT TAAGCCTGAC ACCAAAATGA AGCCTTGCTA TGGTTCTTTT GCCAAGCCTA
18961 CCAATAAAGA AGGAGGCCAG GCAAATGTGA AAACCGAAAC AGGCGGTACC AAAGAATATG
19021 ACATTGACAT GGCATTCTTC GATAATCGAA GTGCAGCTGC CGCCGGCCTA GCCCCAGAAA
19081 TTGTTTTGTA TACTGAGAAT GTGGATCTGG AAACTCCAGA TACCCATATT GTATACAAGG
19141 CAGGTACAGA TGACAGTAGC TCTTCTATCA ATTTGGGTCA GCAGTCCATG CCCAACAGAC
19201 CCAACTACAT TGGCTTCAGA GACAACTTTA TCGGTCTGAT GTACTACAAC AGCACTGGCA
19261 ATATGGGTGT ACTGGCTGGA CAGGCCTCCC AGCTGAATGC TGTGGTGGAC TTGCAGGACA
19321 GAAACACCGA ACTGTCCTAC CAGCTCTTGC TTGACTCTCT GGGTGACAGA ACCAGGTATT
19381 TCAGTATGTG GAATCAGGCG GTGGACAGTT ATGACCCCGA TGTGCGCATT ATTGAAAATC
19441 ACGGTGTGGA GGATGAACTT CCTAACTATT GCTTCCCCCT GGATGCTGTG GGTAGAACTG
19501 ATACTTACCA GGGAATTAAG GCCAATGGTG ATAATCAAAC CACCTGGACC AAAGATGATA
19561 CTGTTAATGA TGCTAATGAA TTGGGCAAGG GCAATCCTTT CGCCATGGAG ATCAACATCC
19621 AGGCCAACCT GTGGCGGAAC TTCCTCTACG CGAACGTGGC GCTGTACCTG CCCGACTCCT
19681 ACAAGTACAC GCCGGCCAAC ATCACGCTGC CCACCAACAC CAACACCTAC GATTACATGA
19741 ACGGCCGCGT GGTGGCGCCC TCGCTGGTGG ACGCCTACAT CAACATCGGG GCGCGCTGGT
19801 CGCTGGACCC CATGGACAAC GTCAACCCCT TCAACCACCA CCGCAACGCG GGCCTGCGAT
19861 ACCGCTCCAT GCTCCTGGGC AACGGGCGCT ACGTGCCCTT CCACATCCAG GTGCCCCAAA
19921 AGTTTTTCGC CATCAAGAGC CTCCTGCTCC TGCCCGGGTC CTACACCTAC GAGTGGAACT
19981 TCCGCAAGGA CGTCAACATG ATCCTGCAGA GCTCCCTCGG CAACGACCTG CGCACGGACG
20041 GGGCCTCCAT CGCCTTCACC AGCATCAACC TCTACGCCAC CTTCTTCCCC ATGGCGCACA
20101 ACACCGCCTC CACGCTCGAG GCCATGCTGC GCAACGACAC CAACGACCAG TCCTTCAACG
20161 ACTACCTCTC GGCGGCCAAC ATGCTCTACC CCATCCCGGC CAACGCCACC AACGTGCCCA
20221 TCTCCATCCC CTCGCGCAAC TGGGCCGCCT TCCGCGGCTG GTCCTTCACG CGCCTCAAGA
20281 CCCGCGAGAC GCCCTCGCTC GGCTCCGGGT TCGACCCCTA CTTCGTCTAC TCGGGCTCCA
20341 TCCCCTACCT CGACGGCACC TTCTACCTCA ACCACACCTT CAAGAAGGTC TCCATCACCT
20401 TCGACTCCTC CGTCAGCTGG CCCGGCAACG ACCGCCTCCT GACGCCCAAC GAGTTCGAAA
20461 TCAAGCGCAC CGTCGACGGA GAGGGGTACA ACGTGGCCCA GTGCAACATG ACCAAGGACT
20521 GGTTCCTGGT CCAGATGCTG GCCCACTACA ACATCGGCTA CCAGGGCTTC TACGTGCCCG
20581 AGGGCTACAA GGACCGCATG TACTCCTTCT TCCGCAACTT CCAGCCCATG AGCCGCCAGG
20641 TCGTGGACGA GGTCAACTAC AAGGACTACC AGGCCGTCAC CCTGGCCTAC CAGCACAACA
20701 ACTCGGGCTT CGTCGGCTAC CTCGCGCCCA CCATGCGCCA GGGCCAGCCC TACCCCGCCA
20761 ACTACCCCTA CCCGCTCATC GGCAAGAGCG CCGTCGCCAG CGTCACCCAG AAAAAGTTCC
20821 TCTGCGACCG GGTCATGTGG CGCATCCCCT TCTCCAGCAA CTTCATGTCC ATGGGCGCGC

FIG. 8F

```
20881 TCACCGACCT CGGCCAGAAC ATGCTCTACG CCAACTCCGC CCACGCGCTA GACATGAATT
20941 TCGAAGTCGA CCCCATGGAT GAGTCCACCC TTCTCTATGT TGTCTTCGAA GTCTTCGACG
21001 TCGTCCGAGT GCACCAGCCC CACCGCGGCG TCATCGAGGC CGTCTACCTG CGCACGCCCT
21061 TCTCGGCCGG CAACGCCACC ACCTAAGCCT CTTGCTTCTT GCAAGATGAC GGCCTGCGCG
21121 GGCTCCGGCG AGCAGGAGCT CAGGGCCATC CTCCGCGACC TGGGCTGCGG GCCCTGCTTC
21181 CTGGGCACCT TCGACAAGCG CTTCCCGGGA TTCATGGCCC CGCACAAGCT GGCCTGCGCC
21241 ATCGTCAACA CGGCCGGCCG CGAGACCGGG GGCGAGCACT GGCTGGCCTT CGCCTGGAAC
21301 CCGCGCTCCC ACACCTGCTA CCTCTTCGAC CCCTTCGGGT TCTCGGACGA GCGCCTCAAG
21361 CAGATCTACC AGTTCGAGTA CGAGGGCCTG CTGCGTCGCA GCGCCCTGGC CACCGAGGAC
21421 CGCTGCGTCA CCCTGGAAAA GTCCACCCAG ACCGTGCAGG GTCCGCGCTC GGCCGCCTGC
21481 GGGCTCTTCT GCTGCATGTT CCTGCACGCC TTCGTGCACT GGCCCGACCG CCCCATGGAC
21541 AAGAACCCCA CCATGAACTT GCTGACGGGG GTGCCCAACG GCATGCTCCA GTCGCCCCAG
21601 GTGGAACCCA CCCTGCGCCG CAACCAGGAG GCGCTCTACC GCTTCCTCAA CGCCCACTCC
21661 GCCTACTTTC GCTCCCACCG CGCGCGCATC GAGAAGGCCA CCGCCTTCGA CCGCATGAAT
21721 CAAGACATGT AATCCGGTGT GTGTATGTGA ATGCTTTATT CATCATAATA AACAGCACAT
21781 GTTTATGCCA CCTTCTCTGA GGCTCTGACT TTATTTAGAA ATCGAAGGGG TTCTGCCGGC
21841 TCTCGGCATG GCCCGCGGGC AGGGATACGT TGCGGAACTG GTACTTGGGC AGCCACTTGA
21901 ACTCGGGGAT CAGCAGCTTC GGCACGGGGA GGTCGGGGAA CGAGTCGCTC CACAGCTTGC
21961 GCGTGAGTTG CAGGGCGCCC AGCAGGTCGG GCGCGGAGAT CTTGAAATCG CAGTTGGGAC
22021 CCGCGTTCTG CGCGCGAGAG TTACGGTACA CGGGGTTGCA GCACTGGAAC ACCATCAGGG
22081 CCGGGTGCTT CACGCTCGCC AGCACCGTCG CGTCGGTGAT GCCCTCCACG TCCAGATCCT
22141 CGGCGTTGGC CATCCCGAAG GGGGTCATCT TGCAGGTCTG CCGCCCCATG CTGGGCACGC
22201 AGCCGGGCTT GTGGTTGCAA TCGCAGTGCA GGGGGATCAG CATCATCTGG GCCTGCTCGG
22261 AGCTCATGCC CGGGTACATG GCCTTCATGA AAGCCTCCAG CTGGCGGAAG GCCTGCTGCG
22321 CCTTGCCGCC CTCGGTGAAG AAGACCCCGC AGGACTTGCT AGAGAACTGG TTGGTGGCGC
22381 AGCCAGCGTC GTGCACGCAG CAGCGCGCGT CGTTGTTGGC CAGCTGCACC ACGCTGCGCC
22441 CCCAGCGGTT CTGGGTGATC TTGGCCCGGT CGGGGTTCTC CTTCAGCGCG CGCTGCCCGT
22501 TCTCGCTCGC CACATCCATC TCGATCGTGT GCTCCTTCTG GATCATCACG GTCCCGTGCA
22561 GGCACCGCAG CTTGCCCTCG GCCTCGGTGC ACCCGTGCAG CCACAGCGCG CAGCCGGTGC
22621 TCTCCCAGTT CTTGTGGGCG ATCTGGGAGT GCGAGTGCAC GAAGCCCTGC AGGAAGCGGC
22681 CCATCATCGT GGTCAGGGTC TTGTTGCTGG TGAAGGTCAG CGGAATGCCG CGGTGCTCCT
22741 CGTTCACATA CAGGTGGCAG ATACGGCGGT ACACCTCGCC CTGCTCGGGC ATCAGCTGGA
22801 AGGCGGACTT CAGGTCGCTC TCCACGCGGT ACCGGTCCAT CAGCAGCGTC ATCACTTCCA
22861 TGCCCTTCTC CCAGGCCGAA ACGATCGGCA GGCTCAGGGG GTTCTTCACC GTTGTCATCT
22921 TAGTCGCCGC CGCCGAAGTC AGGGGGTCGT TCTCGTCCAG GGTCTCAAAC ACTCGCTTGC
22981 CGTCCTTCTC GGTGATGCGC ACGGGGGGAA AGCTGAAGCC CACGGCCGCC AGCTCCTCCT
23041 CGGCCTGCCT TTCGTCCTCG CTGTCCTGGC TGATGTCTTG CAAAGGCACA TGCTTGGTCT
23101 TGCGGGGTTT CTTTTTGGGC GGCAGAGGCG GCGGCGGAGA CGTGCTGGGC GAGCGCGAGT
23161 TCTCGCTCAC CACGACTATT TCTTCTCCTT GGCCGTCGTC CGAGACCACG CGGCGGTAGG
23221 CATGCCTCTT CTGGGGCAGA GGCGGAGGCG ACGGGCTCTC GCGGTTCGGC GGGCGGCTGG
23281 CAGAGCCCCT TCCGCGTTCG GGGGTGCGCT CCTGGCGGCG CTGCTCTGAC TGACTTCCTC
23341 CGCGGCCGGC CATTGTGTTC TCCTAGGGAG CAAGCATGGA GACTCAGCCA TCGTCGCCAA
23401 CATCGCCATC TGCCCCCGCC GCCGCCGACG AGAACCAGCA GCAGCAGAAT GAAAGCTTAA
23461 CCGCCCCGCC GCCCAGCCCC ACCTCCGACG CCGCAGCCCC AGACATGCAA GAGATGGAGG
23521 AATCCATCGA GATTGACCTG GGCTACGTGA CGCCCGCGGA GCACGAGGAG GAGCTGGCAG
23581 CGCGCTTTTC AGCCCCGGAA GAGAACCACC AAGAGCAGCC AGAGCAGGAA GCAGAGAGCG
23641 AGCAGAACCA GGCTGGGCTC GAGCATGGCG ACTACCTGAG CGGGGCAGAG GACGTGCTCA
23701 TCAAGCATCT GGCCCGCCAA TGCATCATCG TCAAGGACGC GCTGCTCGAC CGCGCCGAGG
23761 TGCCCCTCAG CGTGGCGGAG CTCAGCCGCG CCTACGAGCG CAACCTCTTC TCGCCGCGCG
23821 TGCCCCCCAA GCGCCAGCCC AACGGCACCT GCGAGCCCAA CCCGCGCCTC AACTTCTACC
23881 CGGTCTTCGC GGTGCCCGAG GCCCTGGCCA CCTACCACCT CTTTTTCAAG AACCAAAGGA
23941 TCCCCGTCTC CTGCCGCGCC AACCGCACCC GCGCCGACGC CCTGCTCAAC CTGGGCCCCG
24001 GCGCCCGCCT ACCTGATATC GCCTCCTTGG AAGAGGTTCC CAAGATCTTC GAGGGTCTGG
24061 GCAGCGACGA GACTCGGGCC GCGAACGCTC TGCAAGGAAG CGGAGAGGAG CATGAGCACC
24121 ACAGCGCCCT GGTGGAGTTG GAAGGCGACA ACGCGCGCCT GGCGGTCCTC AAGCGCACGG
24181 TCGAGCTGAC CCACTTCGCC TACCCGGCGC TCAACCTGCC CCCCAAGGTC ATGAGCGCCG
24241 TCATGGACCA GGTGCTCATC AAGCGCGCCT CGCCCCTCTC GGAGGAGGAG ATGCAGGACC
24301 CCGAGAGCTC GGACGAGGGC AAGCCCGTGG TCAGCGACGA GCAGCTGGCG CGCTGGCTGG
```

FIG. 8G

25561 GAGAGCAGTC AGGCAGAGGA GGAGGAGATG GAAGACTGGG ACAGCACTCA GGCAGAGGAG
24361 GAGCGAGTAG CACCCCCCAG AGCCTGGAAG AGCGGCGCAA GCTCATGATG GCCGTGGTCC
24421 TGGTGACCGT GGAGCTGGAG TGTCTGCGCC GCTTCTTCGC CGACGCGGAG ACCCTGCGCA
24481 AGGTCGAGGA GAACCTGCAC TACCTCTTCA GACACGGGTT CGTGCGCCAG GCCTGCAAGA
24541 TCTCCAACGT GGAGCTGACC AACCTGGTCT CCTACATGGG CATCCTGCAC GAGAACCGCC
24601 TGGGGCAGAA CGTGCTGCAC ACCACCCTGC GCGGGGAGGC CCGCCGCGAC TACATCCGCG
24661 ACTGCGTCTA CCTGTACCTC TGCCACACCT GGCAGACGGG CATGGGCGTG TGGCAGCAGT
24721 GCCTGGAGGA GCAGAACCTG AAAGAGCTCT GCAAGCTCCT GCAGAAGAAC CTCAAGGCCC
24781 TGTGGACCGG GTTCGACGAG CGCACCACCG CCGCGGACCT GGCCGACCTC ATCTTCCCCG
24841 AGCGCCTGCG GCTGACGCTG CGCAACGGGC TGCCCGACTT TATGAGCCAA AGCATGTTGC
24901 AAAACTTTCG CTCTTTCATC CTCGAACGCT CCGGGATCCT GCCCGCCACC TGCTCCGCGC
24961 TGCCCTCGGA CTTCGTGCCG CTGACCTTCC GCGAGTGCCC CCCGCCGCTC TGGAGCCACT
25021 GCTACCTGCT GCGCCTGGCC AACTACCTGG CCTACCACTC GGACGTGATC GAGGACGTCA
25081 GCGGCGAGGG CCTGCTCGAG TGCCACTGCC GCTGCAACCT CTGCACGCCG CACCGCTCCC
25141 TGGCCTGCAA CCCCCAGCTG CTGAGCGAGA CCCAGATCAT CGGCACCTTC GAGTTGCAAG
25201 GCCCCGGCGA GGGCAAGGGG GGTCTGAAAC TCACCCCGGG GCTGTGGACC TCGGCCTACT
25261 TGCGCAAGTT CGTGCCCGAG GACTACCATC CCTTCGAGAT CAGGTTCTAC GAGGACCAAT
25321 CCCAGCCGCC CAAGGCCGAG CTGTCGGCCT GCGTCATCAC CCAGGGGGCC ATCCTGGCCC
25381 AATTGCAAGC CATCCAGAAA TCCCGCCAAG AATTTCTGCT GAAAAAGGGC CACGGGGTCT
25441 ACTTGGACCC CCAGACCGGA GAGGAGCTCA ACCCCAGCTT CCCCCAGGAT GCCCCGAGGA
25501 AGCAGCAAGA AGCTGAAAGT GGAGCTGCCG CCGCCGCCGG AGGATTTGGA GGAAGACTGG
25621 GACAGCCTGC AAGACAGTCT GGAGGAGGAA GACGAGGTGG AGGAGGCAGA GGAAGAAGCA
25681 GCCGCCGCCA GACCGTCGTC CTCGGCGGAG GAGGAGAAAG CAAGCAGCAC GGATACCATC
25741 TCCGCTCCGG GTCGGGGTCG CGGCGGCCGG GCCCACAGTA GATGGGACGA GACCGGGCGC
25801 TTCCCGAACC CCACCACCCA GACCGGTAAG AAGGAGCGGC AGGGATACAA GTCCTGGCGG
25861 GGGCACAAAA ACGCCATCGT CTCCTGCTTG CAAGCCTGCG GGGGCAACAT CTCCTTCACC
25921 CGGCGCTACC TGCTCTTCCA CCGCGGGGTG AACTTCCCCC GCAACATCTT GCATTACTAC
25981 CGTCACCTCC ACAGCCCCTA CTACTGTTTC CAAGAAGAGG CAGAAACCCA GCAGCAGCAG
26041 CAGCAGCAGA AAACCAGCGG CAGCAGCTAG AAAATCCACA GCGGCGGCAG GTGGACTGAG
26101 GATCGCGGCG AACGAGCCGG CGCAGACCCG GGAGCTGAGG AACCGGATCT TTCCCACCCT
26161 CTATGCCATC TTCCAGCAGA GTCGGGGGCA AGAGCAGGAA CTGAAAGTCA AGAACCGTTC
26221 TCTGCGCTCG CTCACCCGCA GTTGTCTGTA TCACAAGAGC GAAGACCAAC TTCAGCGCAC
26281 TCTCGAGGAC GCCGAGGCTC TCTTCAACAA GTACTGCGCG CTCACTCTTA AAGAGTAGCC
26341 CGCGCCCGCC CACACACGGA AAAAGGCGGG AATTACGTCA CCACCTGCGC CCTTCGCCCG
26401 ACCATCATCA TGAGCAAAGA GATTCCCACG CCTTACATGT GGAGCTACCA GCCCCAGATG
26461 GGCCTGGCCG CCGGCGCCGC CCAGGACTAC TCCACCCGCA TGAACTGGCT CAGTGCCGGG
26521 CCCGCGATGA TCTCACGGGT GAATGACATC CGCGCCCACC GAAACCAGAT ACTCCTAGAA
26581 CAGTCAGCGA TCACCGCCAC GCCCCGCCAT CACCTTAATC CGCGTAATTG GCCCGCCGCC
26641 CTGGTGTACC AGGAAATTCC CCAGCCCACG ACCGTACTAC TTCCGCGAGA CGCCCAGGCC
26701 GAAGTCCAGC TGACTAACTC AGGTGTCCAG CTGGCCGGCG GCGCCGCCCT GTGTCGTCAC
26761 CGCCCCGCTC AGGGTATAAA GCGGCTGGTG ATCCGAGGCA GAGGCACACA GCTCAACGAC
26821 GAGGTGGTGA GCTCTTCGCT GGGTCTGCGA CCTGACGGAG TCTTCCAACT CGCCGGATCG
26881 GGGAGATCTT CCTTCACGCC TCGTCAGGCC GTCCTGACTT TGGAGAGTTC GTCCTCGCAG
26941 CCCCGCTCGG GTGGCATCGG CACTCTCCAG TTCGTGGAGG AGTTCACTCC CTCGGTCTAC
27001 TTCAACCCCT TCTCCGGCTC CCCCGGCCAC TACCCGGACG AGTTCATCCC GAACTTCGAC
27061 GCCATCAGCG AGTCGGTGGA CGGCTACGAT TGAATGTCCC ATGGTGGCGC GGCTGACCTA
27121 GCTCGGCTTC GACACCTGGA CCACTGCCGC CGCTTCCGCT GCTTCGCTCG GGATCTCGCC
27181 GAGTTTGCCT ACTTTGAGCT GCCCGAGGAG CACCCTCAGG GCCCGGCCCA CGGAGTGCGG
27241 ATCGTCGTCG AAGGGGGTCT CGACTCCCAC CTGCTTCGGA TCTTCAGCCA GCGTCCGATC
27301 CTGGCCGAGC GCGAGCAAGG ACAGACCCTT CTGACCCTGT ACTGCATCTG CAACCACCCC
27361 GGCCTGCATG AAAGTCTTTG TTGTCTGCTG TGTACTGAGT ATAATAAAAG CTGAGATCAG
27421 CGACTACTCC GGACTTCCGT GTGTTCCTGC TATCAACCAG TCCCTGTTCT TCACCGGGAA
27481 CGAGACCGAG CTCCAGCTCC AGTGTAAGCC CCACAAGAAG TACCTCACCT GGCTGTTCCA
27541 GGGCTCTCCG ATCGCCGTTG TCAACCACTG CGACAACGAC GGAGTCCTGC TGAGCGGCCC
27601 TGCCAACCTT ACTTTTTCCA CCCGCAGAAG CAAGCTCCAG CTCTTCCAAC CCTTCCTCCC
27661 CGGGACCTAT CAGTGCGTCT CGGGACCCTG CCATCACACC TTCCACCTGA TCCCGAATAC
27721 CACAGCGTCG CTCCCCGCTA CTAACAACCA AACTACCCAC CAACGCCACC GTCGCGACCT
27781 TTCCTCTGGG TCTAATACCA CTACCGGAGG TGAGCTCCGA GGTCGACCAA CCTCTGGGAT

FIG. 8H

```
27841 TTACTACGGC CCCTGGGAGG TGGTAGGGTT AATAGCGCTA GGCCTAGTTG CGGGTGGGCT
27901 TTTGGCTCTC TGCTACCTAT ACCTCCCTTG CTGTTCGTAC TTAGTGGTGC TGTGTTGCTG
27961 GTTTAAGAAA TGGGGAAGAT CACCCTAGTG AGCTGCGGTG TGCTGGTGGC GGTGGTGCTT
28021 TCGATTGTGG GACTGGGCGG CGCGGCTGTA GTGAAGGAGA AGGCCGATCC CTGCTTGCAT
28081 TTCAATCCCG ACAAATGCCA GCTGAGTTTT CAGCCCGATG GCAATCGGTG CGCGGTGCTG
28141 ATCAAGTGCG GATGGGAATG CGAGAACGTG AGAATCGAGT ACAATAACAA GACTCGGAAC
28201 AATACTCTCG CGTCCGTGTG GCAGCCCGGG GACCCCGAGT GGTACACCGT CTCTGTCCCC
28261 GGTGCTGACG GCTCCCCGCG CACCGTGAAT AATACTTTCA TTTTTGCGCA CATGTGCGAC
28321 ACGGTCATGT GGATGAGCAA GCAGTACGAT ATGTGGCCCC CCACGAAGGA GAACATCGTG
28381 GTCTTCTCCA TCGCTTACAG CGTGTGCACG GCGCTAATCA CCGCTATCGT GTGCCTGAGC
28441 ATTCACATGC TCATCGCTAT TCGCCCCAGA AATAATGCCG AAAAAGAAAA ACAGCCATAA
28501 CACGTTTTTT CACACACCTT TTTCAGACCA TGGCCTCTGT TAAATTTTTG CTTTTATTTG
28561 CCAGTCTCAT TGCCGTCATT CATGGAATGA GTAATGAGAA AATTACTATT TACACTGGCA
28621 CTAATCACAC ATTGAAAGGT CCAGAAAAAG CCACAGAAGT TTCATGGTAT TGTTATTTTA
28681 ATGAATCAGA TGTATCTACT GAACTCTGTG GAAACAATAA CAAAAAAAAT GAGAGCATTA
28741 CTCTCATCAA GTTTCAATGT GGATCTGACT TAACCCTAAT TAACATCACT AGAGACTATG
28801 TAGGTATGTA TTATGGAACT ACAGCAGGCA TTTCGGACAT GGAATTTTAT CAAGTTTCTG
28861 TGTCTGAACC CACCACGCCT AGAATGACCA CAACCACAAA AACTACACCT GTTACCACTA
28921 TACAGCTCAC TACCAATGGC TTTCTTGCCA TGCTTCAAGT GGCTGAAAAT AGCACCAGCA
28981 TTCAACCCAC CCCACCCAGT GAGGAAATTC CCAGATCCAT GATTGGCATT ATTGTTGCTG
29041 TAGTGGTGTG CATGTTGATC ATCGCCTTGT GCATGGTGTA CTATGCCTTC TGCTACAGAA
29101 AGCACAGACT GAACGACAAG CTGGAACACT TACTAAGTGT TGAATTTTAA TTTTTTAGAA
29161 CCATGAAGAT CCTAGGCCTT TTAGTTTTTT CTATCATTAC CTCTGCTCTA TGCAATTCTG
29221 ACAATGAGGA CGTTACTGTC GTTGTCGGAT CAAATTATAC ACTAAAAGGT CCAGCAAAAG
29281 GTATGCTTTC GTGGTATTGT TGGTTCGGAA CTGACGAGCA ACAGACAGAA CTTTGCAATG
29341 CTCAAAAAGG CAAAACCTCA AATTCTAAAA TCTCTAATTA TCAATGCAAT GGCACTGACT
29401 TAGTATTGCT CAATGTCACG AAAGCATATG CTGGCAGTTA CACCTGCCCT GGAGATGATG
29461 CCGACAATAT GATTTTTTAC AAAGTGGAAG TGGTTGATCC CACTACTCCA CCGCCCACCA
29521 CCACAACTAC TCATACCACA CACACAGAAC AAACACCAGA GGCAGCAGAA GCAGAGTTGG
29581 CCTTCCAGGT TCACGGAGAT TCCTTTGCTG TCAATACCCC TACACCCGAT CAGCGGTGTC
29641 CGGGGCTGCT CGTCAGCGGC ATTGTCGGTG TGCTTTCGGG ATTAGCAGTC ATAATCATCT
29701 GCATGTTCAT TTTTGCTTGC TGCTATAGAA GGCTTTACCG ACAAAAATCA GACCCACTGC
29761 TGAACCTCTA TGTTTAATTT TTTCCAGAGC CATGAAGGCA GTTAGCGCTC TAGTTTTTTG
29821 TTCTTTGATT GGCATTGTTT TTAGTGCTGG GTTTTTGAAA AATCTTACCA TTTATGAAGG
29881 TGAGAATGCC ACTCTAGTGG GCATCAGTGG TCAAAATGTC AGCTGGCTAA AATACCATCT
29941 AGATGGGTGG AAAGACATTT GCGATTGGAA TGTCACTGTG TATACATGTA ATGGAGTTAA
30001 CCTCACCATT ACTAATGCCA CCCAAGATCA GAATGGTAGG TTTAAGGGCC AGAGTTTCAC
30061 TAGAAATAAT GGGTATGAAT CCCATAACAT GTTTATCTAT GACGTCACTG TCATCAGAAA
30121 TGAGACTGCC ACCACCACAC AGATGCCCAC TACACACAGT TCTACCACTA CTACCATGCA
30181 AACCACACAG ACAACCACTA CATCAACTCA GCATATGACC ACCACTACAG CAGCAAAGCC
30241 AAGTAGTGCA GCGCCTCAGC CCCAGGCTTT GGCTTTGAAA GCTGCACAAC CTAGTACAAC
30301 TACTAGGACC AATGAGCAGA CTACTGAATT TTTGTCCACT GTCGAGAGCC ACACCACAGC
30361 TACCTCCAGT GCCTTCTCTA GCACCGCCAA TCTCTCCTCG CTTTCCTCTA CACCAATCAG
30421 TCCCGCTACT ACTCCCACCC CAGCTCTTCT CCCCACTCCC CTGAAGCAAA CTGAGGACAG
30481 CGGCATGCAA TGGCAGATCA CCCTGCTCAT TGTGATCGGG TTGGTCATCC TGGCCGTGTT
30541 GCTCTACTAC ATCTTCTGCC GCCGCATTCC CAACGCGCAC CGCAAACCGG CCTACAAGCC
30601 CATCGTTATC GGGCAGCCGG AGCCGCTTCA GGTGGAAGGG GGTCTAAGGA ATCTTCTCTT
30661 CTCTTTTACA GTATGGTGAT TGAACTATGA TTCCTAGACA ATTCTTGATC ACTATTCTTA
30721 TCTGCCTCCT CCAAGTCTGT GCCACCCTCG CTCTGGTGGC CAACGCCAGT CCAGACTGTA
30781 TTGGGCCCTT CGCCTCCTAC GTGCTCTTTG CCTTCATCAC CTGCATCTGC TGCTGTAGCA
30841 TAGTCTGCCT GCTTATCACC TTCTTCCAGT TCATTGACTG GATCTTTGTG CGCATCGCCT
30901 ACCTGCGCCA CCACCCCCAG TACCGCGACC AGCGAGTGGC GCGGCTGCTC AGGCTCCTCT
30961 GATAAGCATG CGGGCTCTGC TACTTCTCGC GCTTCTGCTG TTAGTGCTCC CCCGCCCCGT
31021 CGACCCCCGG TCCCCCACTC AGTCCCCCGA AGAGGTCCGC AAATGCAAAT TCCAAGAACC
31081 CTGGAAATTC CTCAAATGCT ACCGCCAAAA ATCAGACATG CTTCCCAGCT GGATCATGAT
31141 CATTGGGATC GTGAACATTC TGGCCTGCAC CCTCATCTCC TTTGTGATTT ACCCCTGCTT
31201 TGACTTTGGT TGGAACTCGC CAGAGGCGCT CTATCTCCCG CCTGAACCTG ACACACCACC
31261 ACAGCAACCT CAGGCACACG CACTACCACC ACCACAGCCT AGGCCACAAT ACATGCCCAT
```

FIG. 8I

```
31321 ATTAGACTAT GAGGCCGAGC CACAGCGACC CATGCTCCCC GCTATTAGTT ACTTCAATCT
31381 AACCGGCGGA GATGACTGAC CCACTGGCCA ACAACAACGT CAACGACCTT CTCCTGGACA
31441 TGGACGGCCG CGCCTCGGAG CAGCGACTCG CCCAACTTCG CATTCGCCAG CAGCAGGAGA
31501 GAGCCGTCAA GGAGCTGCAG GACGGCATAG CCATCCACCA GTGCAAGAAA GGCATCTTCT
31561 GCCTGGTGAA ACAGGCCAAG ATCTCCTACG AGGTCACCCC GACCGACCAT CGCCTCTCCT
31621 ACGAGCTCCT GCAGCAGCGC CAGAAGTTCA CCTGCCTGGT CGGAGTCAAC CCCATCGTCA
31681 TCACCCAGCA GTCGGGCGAT ACCAAGGGGT GCATCCACTG CTCCTGCGAC TCCCCCGACT
31741 GCGTCCACAC TCTGATCAAG ACCCTCTGCG GCCTCCGCGA CCTCCTCCCC ATGAACTAAT
31801 CACCCCCTTA TCCAGTGAAA TAAATATCAT ATTGATGATG ATTTAAATAA AAAATAATCA
31861 TTTGATTTGA AATAAAGATA CAATCATATT GATGATTTGA GTTTTAAAAA ATAAAGAATC
31921 ACTTACTTGA AATCTGATAC CAGGTCTCTG TCCATGTTTT CTGCCAACAC CACCTCACTC
31981 CCCTCTTCCC AGCTCTGGTA CTGCAGACCC CGGCGGGCTG CAAACTTCCT CCACACGCTG
32041 AAGGGGATGT CAAATTCCTC CTGTCCCTCA ATCTTCATTT TATCTTCTAT CAGATGTCCA
32101 AAAAGCGCGT CCGGGTGGAT GATGACTTCG ACCCCGTCTA CCCCTACGAT GCAGACAACG
32161 CACCGACCGT GCCCTTCATC AACCCCCCCT TCGTCTCTTC AGATGGATTC CAAGAGAAGC
32221 CCCTGGGGGT GCTGTCCCTG CGACTGGCTG ACCCCGTCAC CACCAAGAAC GGGGAAATCA
32281 CCCTCAAGCT GGGAGAGGGG GTGGACCTCG ACTCCTCGGG AAAACTCATC TCCAACACGG
32341 CCACCAAGGC CGCCGCCCCT CTCAGTTTTT CCAACAACAC CATTTCCCTT AACATGGATA
32401 CCCCTCTTTA TACCAAAGAT GGAAAATTAT CCTTACAAGT TTCTCCACCG TTAAACATAT
32461 TAAAATCAAC CATTCTGAAC ACATTAGCTG TAGCTTATGG ATCAGGTTTA GGACTGAGTG
32521 GTGGCACTGC TCTTGCAGTA CAGTTGGCCT CTCCACTCAC TTTTGATGAA AAAGGAAATA
32581 TTAAAATTAA CCTAGCCAGT GGTCCATTAA CAGTTGATGC AAGTCGACTT AGTATCAACT
32641 GCAAAAGAGG GGTCACTGTC ACTACCTCAG GAGATGCAAT TGAAAGCAAC ATAAGCTGGC
32701 CTAAAGGTAT AAGATTTGAA GGTAATGGCA TAGCTGCAAA CATTGGCAGA GGATTGGAAT
32761 TTGGAACCAC TAGTACAGAG ACTGATGTCA CAGATGCATA CCCAATTCAA GTTAAATTGG
32821 GTACTGGCCT TACCTTTGAC AGTACAGGCG CCATTGTTGC TTGGAACAAA GAGGATGATA
32881 AACTTACATT ATGGACCACA GCCGACCCCT CGCCAAATTG CAAAATATAC TCTGAAAAAG
32941 ATGCCAAACT CACACTTTGC TTGACAAAGT GTGGAAGTCA AATTCTGGGT ACTGTGACTG
33001 TATTGGCAGT GAATAATGGA AGTCTCAACC CAATCACAAA CACAGTAAGC ACTGCACTCG
33061 TCTCCCTCAA GTTTGATGCA AGTGGAGTTT TGCTAAGCAG CTCCACATTA GACAAAGAAT
33121 ATTGGAACTT CAGAAAGGGA GATGTTACAC CTGCTGAGCC CTATACTAAT GCTATAGGTT
33181 TTATGCCTAA CATAAAGGCC TATCCTAAAA ACACATCTGC AGCTTCAAAA AGCCATATTG
33241 TCAGTCAAGT TTATCTCAAT GGGGATGAGG CCAAACCACT GATGCTGATT ATTACTTTTA
33301 ATGAAACTGA GGATGCAACT TGCACCTACA GTATCACTTT TCAATGGAAA TGGGATAGTA
33361 CTAAGTACAC AGGTGAAACA CTTGCTACCA GCTCCTTCAC CTTCTCCTAC ATCGCCCAAG
33421 AATGAACACT GTATCCCACC CTGCATGCCA ACCCTTCCCA CCCCACTCTG TCTATGGAAA
33481 AAACTCTGAA GCACAAAATA AAATAAAGTT CAAGTGTTTT ATTGATTCAA CAGTTTTACA
33541 GGATTCGAGC AGTTATTTTT CCTCCACCCT CCCAGGACAT GGAATACACC ACCCTCTCCC
33601 CCCGCACAGC CTTGAACATC TGAATGCCAT TGGTGATGGA CATGCTTTTG GTCTCCACGT
33661 TCCACACAGT TTCAGAGCGA GCCAGTCTCG GGTCGGTCAG GGAGATGAAA CCCTCCGGGC
33721 ACTCCCGCAT CTGCACCTCA CAGCTCAACA GCTGAGGATT GTCCTCGGTG GTCGGGATCA
33781 CGGTTATCTG GAAGAAGCAG AAGAGCGGCG GTGGGAATCA TAGTCCGCGA ACGGGATCGG
33841 CCGGTGGTGT CGCATCAGGC CCCGCAGCAG TCGCTGCCGC CGCCGCTCCG TCAAGCTGCT
33901 GCTCAGGGGG TCCGGGTCCA GGGACTCCCT CAGCATGATG CCCACGGCCC TCAGCATCAG
33961 TCGTCTGGTG CGGCGGGCGC AGCAGCGCAT GCGGATCTCG CTCAGGTCGC TGCAGTACGT
34021 GCAACACAGG ACCACCAGGT TGTTCAACAG TCCATAGTTC AACACGCTCC AGCCGAAACT
34081 CATCGCGGGA AGGATGCTAC CCACGTGGCC GTCGTACCAG ATCCTCAGGT AAATCAAGTG
34141 GCGCTCCCTC CAGAACACGC TGCCCACGTA CATGATCTCC TTGGGCATGT GGCGGTTCAC
34201 CACCTCCCGG TACCACATCA CCCTCTGGTT GAACATGCAG CCCCGGATGA TCCTGCGGAA
34261 CCACAGGGCC AGCACCGCCC CGCCCGCCAT GCAGCGAAGA GACCCCGGGT CCCGGCAATG
34321 GCAATGGAGG ACCCACCGCT CGTACCCGTG GATCATCTGG GAGCTGAACA AGTCTATGTT
34381 GGCACAGCAC AGGCATATGC TCATGCATCT CTTCAGCACT CTCAGCTCCT CGGGGGTCAA
34441 AACCATATCC CAGGGCACGG GGAACTCTTG CAGGACAGCG AACCCCGCAG AACAGGGCAA
34501 TCCTCGCACA TAACTTACAT TGTGCATGGA CAGGGTATCG CAATCAGGCA GCACCGGGTG
34561 ATCCTCCACC AGAGAAGCGC GGGTCTCGGT CTCCTCACAG CGTGGTAAGG GGGCCGGCCG
34621 ATACGGGTGA TGGCGGGACG CGGCTGATCG TGTTCGCGAC CGTGTCATGA TGCAGTTGCT
34681 TTCGGACATT TTCGTACTTG CTGTAGCAGA ACCTGGTCCG GGCGCTGCAC ACCGATCGCC
34741 GGCGGCGGTC CCGGCGCTTG GAACGCTCGG TGTTGAAATT GTAAAACAGC CACTCTCTCA
```

FIG. 8J

```
34801 GACCGTGCAG CAGATCTAGG GCCTCAGGAG TGATGAAGAT CCCATCATGC CTGATAGCTC
34861 TGATCACATC GACCACCGTG GAATGGGCCA GACCCAGCCA GATGATGCAA TTTTGTTGGG
34921 TTTCGGTGAC GGCGGGGGAG GGAAGAACAG GAAGAACCAT GATTAACTTT TAATCCAAAC
34981 GGTCTCGGAG CACTTCAAAA TGAAGGTCGC GGAGATGGCA CCTCTCGCCC CCGCTGTGTT
35041 GGTGGAAAAT AACAGCCAGG TCAAAGGTGA TACGGTTCTC GAGATGTTCC ACGGTGGCTT
35101 CCAGCAAAGC CTCCACGCGC ACATCCAGAA ACAAGACAAT AGCGAAAGCG GGAGGGTTCT
35161 CTAATTCCTC AATCATCATG TTACACTCCT GCACCATCCC CAGATAATTT TCATTTTTCC
35221 AGCCTTGAAT GATTCGAACT AGTTCCTGAG GTAAATCCAA GCCAGCCATG ATAAAGAGCT
35281 CGCGCAGAGC GCCCTCCACC GGCATTCTTA AGCACACCCT CATAATTCCA AGATATTCTG
35341 CTCCTGGTTC ACCTGCAGCA GATTGACAAG CGGAATATCA AAATCTCTGC CGCGATCCCT
35401 AAGCTCCTCC CTCAGCAATA ACTGTAAGTA CTCTTTCATA TCCTCTCCGA AATTTTTAGC
35461 CATAGGACCA CCAGGAATAA GATTAGGGCA AGCCACAGTA CAGATAAACC GAAGTCCTCC
35521 CCAGTGAGCA TTGCCAAATG CAAGACTGCT ATAAGCATGC TGGCTAGACC CGGTGATATC
35581 TTCCAGATAA CTGGACAGAA AATCACCCAG GCAATTTTTA AGAAAATCAA CAAAAGAAAA
35641 ATCCTCCAGG TGCACGTTTA GAGCCTCGGG AACAACGATG AAGTAAATGC AAGCGGTGCG
35701 TTCCAGCATG GTTAGTTAGC TGATCTGTAA AAAACAAAAA ATAAAACATT AAACCATGCT
35761 AGCCTGGCGA ACAGGTGGGT AAATCGTTCT CTCCAGCACC AGGCAGGCCA CGGGGTCTCC
35821 GGCGCGACCC TCGTAAAAAT TGTCGCTATG ATTGAAAACC ATCACAGAGA GACGTTCCCG
35881 GTGGCCGGCG TGAATGATTC GACAAGATGA ATACACCCCC GGAACATTGG CGTCCGCGAG
35941 TGAAAAAAAG CGCCCGAGGA AGCAATAAGG CACTACAATG CTCAGTCTCA AGTCCAGCAA
36001 AGCGATGCCA TGCGGATGAA GCACAAAATC CTCAGGTGCG TACAAAATGT AATTACTCCC
36061 CTCCTGCACA GGCAGCGAAG CCCCCGATCC CTCCAGATAC ACATACAAAG CCTCAGCGTC
36121 CATAGCTTAC CGAGCAGCAG CACACAACAG GCGCAAGAGT CAGAGAAAGG CTGAGCTCTA
36181 ACCTGTCCAC CCGCTCTCTG CTCAATATAT AGCCCAGATC TACACTGACG TAAAGGCCAA
36241 AGTCTAAAAA TACCCGCCAA ATAATCACAC ACGCCCAGCA CACGCCCAGA AACCGGTGAC
36301 ACACTCAAAA AAATACGCGC ACTTCCTCAA ACGCCCAAAC TGCCGTCATT TCCGGGTTCC
36361 CACGCTACGT CATCGGAATT CGACTTTCAA ATTCCGTCGA CCGTTAAAAA CGTCACCCGC
36421 CCCGCCCCTA ACGGTCGCCC GTCTCTCGGC CAATCACCTT CCTCCCTCCC CAAATTCAAA
36481 CAGCTCATTT GCATATTAAC GCGCACCAAA AGTTTGAGGT ATATTATTGA TGATG
```

(SEQ ID NO: 4)

FIG. 8K

```
   1 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg aggtatttga
  61 atttggggat gcggggcggt gattggctgc gggagcggcg accgttaggg gcggggcggg
 121 tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt
 181 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca
 241 ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg
 301 aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag
 361 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat
 421 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggt gtcagctgat cgccagggta
 481 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct
 541 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg
 601 gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg
 661 gtgacgaccc tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgtatg
 721 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta
 781 gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt
 841 cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg
 901 aagagctcga cctgcgctgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg
 961 aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggctttagcc
1021 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata
1081 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt
1141 acagtaagtg tgattaactt tagctgggga ggcagagggt gactgggtgc tgactggttt
1201 atttatgtat atgtttttta tgtgtaggtc ccgtctctga cgtagatgag acccccacta
1261 cagagtgcat ttcatcaccc ccagaaattg gcgaggaacc gcccgaagat attattcata
1321 gaccagttgc agtgagagtc accgggcgta gagcagctgt ggagagtttg gatgacttgc
1381 tacagggtgg ggatgaacct ttggacttgt gtacccggaa acgccccagg cactaagtgc
1441 cacacatgtg tgtttactta aggtgatgtc agtatttata gggtgtggag tgcaataaaa
1501 tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag
1561 caggtgcaga cctgtgtggt cagttcagag caggactcat ggagatctgg acagtcttgg
1621 aagactttca ccagactaga cagctgctag agaactcatc ggagggagtc tcttacctgt
1681 ggagattctg cttcggtggg cctctagcta agctagtcta tagggccaag caggattata
1741 aggatcaatt tgaggatatt ttgagagagt gtcctggtat ttttgactct ctcaacttgg
1801 gccatcagtc tcactttaac cagagtattc tgagagccct tgacttttct actcctggca
1861 gaactaccgc cgcggtagcc ttttttgcct ttatccttga caaatggagt caagaaaccc
1921 atttcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt
1981 gccagcgcct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga
2041 tcctgagtct ccagtcaccc caggaacacc aacgccgcca gcagccgcag caggagcagc
2101 agcaagagga ggaccgagaa gagaacctga gagccggtct ggaccctccg gtggcggagg
2161 aggaggagta gctgacttgt ttcccgagct gcgccgggtg ctgactaggt cttccagtgg
2221 acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac
2281 tgtcagtctg atgagtcgca ggcgcccaga atcggtgtgg tggcatgagg tgcagtcgca
2341 ggggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg
2401 ttggttggag cccgaggatg attgggaggt agccatcagg aattatgcca agctggctct
2461 gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat
2521 ttcagggaat ggggccgagg tggagatcag tacccaggag agggtggcct tcagatgctg
2581 catgatgaat atgtacccgg gggtggtggg catggaggga gtcaccttta tgaacgcgag
2641 gttcaggggt gatgggtata atggggtggt ctttatggcc aacaccaagc tgacagtgca
2701 cggatgctcc ttctttggct tcaataacat gtgcattgag gcctggggca gtgtttcagt
2761 gaggggatgc agtttttcag ccaactggat gggggtcgtg ggcagaacca agagcatggt
2821 gtcagtgaag aaatgcctgt tcgagaggtg ccacctgggg gtgatgagcg agggcgaagc
2881 caaagtcaaa cactgcgcct ctaccgagac gggctgcttt gtactgatca agggcaatgc
2941 caaagtcaag cataatatga tctgtggggc ctcggatgag cgcggctacc agatgctgac
3001 ctgcgccggt gggaacagcc atatgctagc caccgtgcat gtggcctcgc acccccgcaa
3061 gacatggccc gagttcgagc acaacgtcat gacccgctgc aatgtgcacc tggggtcccg
3121 ccgaggcatg ttcatgccct accagtgcaa catgcaattt gtgaaggtgc tgctggagcc
3181 cgatgccatg tccagagtga gcctgacggg ggtgtttgac atgaatgtgg agctgtggaa
3241 aattctgaga tatgatgaat ccaagaccag gtgccgggcc tgcgaatgcg gaggcaagca
3301 cgccaggctt cagcccgtgt gtgtggaggt gacggaggac ctgcgacccg atcatttggt
3361 gttgtcctgc aacgggacgg agttcggctc cagcggggaa gaatctgact agagtgagta
3421 gtgtttggga ctgggtggga gcctgcatga tgggcagaat gactaaaatc tgtgtttttc
```

FIG. 9A

3481 tgcgcagcag catgagcgga agcgcctcct ttgagggagg ggtattcagc ccttatctga
3541 cggggcgtct cccctcctgg gcgggagtgc gtcagaatgt gatgggatcc acggtggacg
3601 gccggcccgt gcagcccgcg aactcttcaa ccctgaccta cgcgaccctg agctcctcgt
3661 ccgtggacgc agctgccgcc gcagctgctg cttccgccgc cagcgccgtg cgcggaatgg
3721 ccctgggcgc cggctactac agctctctgg tggccaactc gagttccacc aataatcccg
3781 ccagcctgaa cgaggagaag ctgctgctgc tgatggccca gctcgaggcc ctgacccagc
3841 gcctgggcga gctgacccag caggtggctc agctgcaggc ggagacgcgg gccgcggttg
3901 ccacggtgaa aaccaaataa aaaatgaatc aataaataaa cggagacggt tgttgatttt
3961 aacacagagt cttgaatctt tatttgattt ttcgcgcgcg gtaggccctg gaccaccggt
4021 ctcgatcatt gagcacccgg tggatctttt ccaggacccg gtagaggtgg gcttggatgt
4081 tgaggtacat gggcatgagc ccgtcccggg ggtggaggta gctccattgc agggcctcgt
4141 gctcgggggt ggtgttgtaa atcacccagt catagcaggg gcgcagggcg tggtgctgca
4201 cgatgtcctt gaggaggaga ctgatggcca cgggcagccc cttggtgtag gtgttgacga
4261 acctgttgag ctgggaggga tgcatgcggg gggagatgag atgcatcttg gcctggatct
4321 tgagattggc gatgttcccg cccagatccc gccgggggtt catgttgtgc aggaccacca
4381 gcacggtgta tccggtgcac ttggggaatt tgtcatgcaa cttggaaggg aaggcgtgaa
4441 agaatttgga gacgcccttg tgaccgccca ggttttccat gcactcatcc atgatgatgg
4501 cgatgggccc gtgggcggcg gcttgggcaa agacgtttcg ggggtcggac acatcgtagt
4561 tgtggtcctg ggtgagctcg tcataggcca ttttaatgaa tttggggcgg agggtgcccg
4621 actgggggac gaaggtgccc tcgatcccgg gggcgtagtt gccctcgcag atctgcatct
4681 cccaggcctt gagctcggag gggggatca tgtccacctg cggggcgatg aaaaaaacgg
4741 tttccggggc gggggagatg agctgggccg aaagcaggtt ccggagcagc tgggacttgc
4801 cgcagccggt ggggccgtag atgaccccga tgaccggctg caggtggtag ttgagggaga
4861 gacagctgcc gtcctcgcgg aggaggggg ccacctcgtt catcatctcg cgcacatgca
4921 tgttctcgcg cacgagttcc gccaggaggc gctcgccccc aagcgagagg agctcttgca
4981 gcgaggcgaa gtttttcagc ggcttgagcc cgtcggccat gggcattttg gagagggtct
5041 gttgcaagag ttccagacgg tcccagagct cggtgatgtg ctctagggca tctcgatcca
5101 gcagacctcc tcgtttcgcg ggttggggcg actgcgggag tagggcacca ggcgatgggc
5161 gtccagcgag gccagggtcc ggtccttcca ggggcgcagg gtccgcgtca gcgtggtctc
5221 cgtcacggtg aaggggtgcg cgccgggctg ggcgcttgcg agggtgcgct tcaggctcat
5281 ccggctggtc gagaaccgct cccggtcggc gccctgcgcg tcggccaggt agcaattgag
5341 catgagttcg tagttgagcg cctcggccgc gtggcccttg gcgcggagct taccttttgga
5401 agtgtgtccg cagacgggac agaggaggga cttgagggcg tagagcttgg gggcgaggaa
5461 gacggactcg ggggcgtagg cgtccgcgcc gcagctggcg cagacggtct cgcactccac
5521 gagccaggtg aggtctggcc ggtcggggtc aaaaacgagg tttcctccgt gctttttgat
5581 gcgtttctta cctctggtct ccatgagctc gtgtccccgc tgggtgacaa agaggctgtc
5641 cgtgtccccg tagaccgact ttatgggccg gtcctcgagc ggggtgccgc ggtcctcgtc
5701 gtagaggaac cccgcccact ccgagacgaa ggcccgggtc caggccagca cgaaggaggc
5761 cacgtgggag gggtagcggt cgttgtccac cagcgggtcc accttctcca gggtatgcaa
5821 gcacatgtcc ccctcgtcca catccaggaa ggtgattggc ttgtaagtgt aggccacgtg
5881 accgggggtc ccggccgggg gggtataaaa gggggcgggc ccctgctcgt cctcactgtc
5941 ttccggatcg ctgtccagga gcgccagctg ttggggtagg tattccctct cgaaggcggg
6001 catgacctcg gcactcaggt tgtcagtttc tagaaacgag gaggatttga tattgacggt
6061 gccgttggag acgcctttca tgagcccctc gtccatctgg tcagaaaaga cgatcttttt
6121 gttgtcgagc ttggtggcga aggagccgta gagggcgttg gagagcagct tggcgatgga
6181 gcgcatggtc tggttctttt ccttgtcggc gcgctccttg gcggcgatgt tgagctgcac
6241 gtactcgcgc gccacgcact tccattcggg gaagacggtg gtgagcttgt cgggcacgat
6301 tctgacccgc cagccgcggt tgtgcagggt gatgaggtcc acgctggtgg ccacctcgcc
6361 gcgcaggggc tcgttggtcc agcagaggcg cccgcccttg cgcgagcaga agggggggcag
6421 cgggtccagc atgagctcgt cgggggggtc ggcgtccacg gtgaagatgc cgggcaggag
6481 ctcggggtcg aagtagctga tgcaggtgcc cagatcgtcc agcgccgctt gccagtcgcg
6541 cacggccagc gcgcgctcgt aggggctgag gggcgtgccc cagggcatgg ggtgcgtgag
6601 cgcggaggcg tacatgccgc agatgtcgta gacgtagagg ggctcctcga ggacgccgat
6661 gtaggtgggg tagcagcgcc ccccgcggat gctggcgcgc acgtagtcgt acagctcgtg
6721 cgagggcgcg aggagcccgg tgccgaggtt ggagcgctgc ggcttttcgg cgcggtagac
6781 gatctggcgg aagatggcgt gggagttgga ggagatggtg ggcctctgga agatgttgaa
6841 gtgggcgtgg ggcagtccga ccgagtccct gatgaagtgg gcgtaggagt cctgcagctt
6901 ggcgacgagc tcggcggtga cgaggacgtc cagggcgcag tagtcgaggg tctcttggat

FIG. 9B

```
6961 gatgtcgtac ttgagctggc ccttctgctt ccacagctcg cggttgagaa ggaactcttc
7021 gcggtccttc cagtactctt cgaggggaa cccgtcctga tcggcacggt aagagcccac
7081 catgtagaac tggttgacgg ccttgtaggc gcagcagccc ttctccacgg ggagggcgta
7141 agcttgcgcg gccttgcgca gggaggtgtg ggtgagggcg aaggtgtcgc gcaccatgac
7201 cttgaggaac tggtgcttga agtcgaggtc gtcgcagccg ccctgctccc agagctggaa
7261 gtccgtgcgc ttcttgtagg cggggttggg caaagcgaaa gtaacatcgt tgaagaggat
7321 cttgcccgcg cggggcatga agttgcgagt gatgcggaaa ggctggggca cctcggcccg
7381 gttgttgatg acctgggcgg cgaggacgat ctcgtcgaag ccgttgatgt tgtgcccgac
7441 gatgtagagt tccacgaatc gcgggcggcc cttgacgtgg ggcagcttct tgagctcgtc
7501 gtaggtgagc tcggcggggt cgctgaggcc gtgctgctcg agggcccagt cggcgaggtg
7561 ggggttggcg ccgaggaagg aagtccagag atccacggcc agggcggtct gcaagcggtc
7621 ccggtactga cggaactgct ggcccacggc catttttttcg ggggtgacgc agtagaaggt
7681 gcgggggtcg ccgtgccagc ggtcccactt gagctggagg gcgaggtcgt gggcgagctc
7741 gacgagcggc gggtccccgg agagtttcat gaccagcatg aaggggacga gctgcttgcc
7801 gaaggacccc atccaggtgt aggtttccac gtcgtaggtg aggaagagcc tttcggtgcg
7861 aggatgcgag ccgatgggga agaactggat ctcctgccac cagttggagg aatggctgtt
7921 gatgtgatgg aagtagaaat gccgacggcg cgccgagcac tcgtgcttgt gtttatacaa
7981 gcgtccgcag tgctcgcaac gctgcacggg atgcacgtgc tgcacgagct gtacctgggt
8041 tcctttgacg aggaatttca gtgggcagtg gagcgctggc ggctgcatct ggtgctgtac
8101 tacgtcctgg ccatcggcgt ggccatcgtc tgcctcgatg gtggtcatgc tgacgaggcc
8161 gcgcgggagg caggtccaga cctcggctcg gacgggtcgg agagcgagga cgagggcgcg
8221 caggccggag ctgtccaggg tcctgagacg ctgcggagtc aggtcagtgg gcagcggcgg
8281 cgcgcggttg acttgcagga gcttttccag ggcgcgcggg aggtccagat ggtacttgat
8341 ctccacggcg ccgttggtgg cgacgtccac ggcttgcagg gtcccgtgcc cctggggcgc
8401 caccaccgtg ccccgtttct tcttgggtgc tggcggcggc ggctccatgc ttagaagcgg
8461 cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg
8521 ggcacgtcgg cgccgcgcgc gggcaggttc tggtactgcg cccggagaag actggcgtga
8581 gcgacgacgc gacggttgac gtcctggatc tgacgcctct gggtgaaggc cacgggaccc
8641 gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc
8701 tgccgcagga tctcttgcac gtcgcccgag ttgtcctggt aggcgatctc ggtcatgaac
8761 tgctcgatct cctcctcctg aaggtctccg cgaccggcgc gctcgacggt ggccgcgagg
8821 tcgttggaga tgcggcccat gagctgcgag aaggcgttca tgccggcctc gttccagacg
8881 cggctgtaga ccacggctcc gtcggggtcg cgcgcgcgca tgaccacctg ggcgaggttg
8941 agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc
9001 gtggtggcga tgtgctcggt gacgaagaag tacatgatcc agcggcggag cggcatctcg
9061 ctgacgtcgc ccagggcttc caagcgctcc atggcctcgt agaagtccac ggcgaagttg
9121 aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg
9181 gcgatggtgg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc ttccatctcc
9241 tcctcctctt ccatctcctc cactaacatc tcttctactt cctcctcagg aggcggcggc
9301 gggggagggg ccctgcgtcg ccggcggcgc acgggcagac ggtcgatgaa gcgctcgatg
9361 gtctccccgc gccggcgacg catggtctcg gtgacggcgc gcccgtcctc gcggggccgc
9421 agcgtgaaga cgccgccgcg catctccagg tggccgccgg gggggtctcc gttgggcagg
9481 gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg
9541 agcgtctcga gatccacggg atccgaaaac cgctgaacga aggcttcgag ccagtcgcag
9601 tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg
9661 gcgatgctgc tggtgatgaa gttgaagtag gcggtcctga gacggcggat ggtggcgagg
9721 agcaccaggt ccttgggccc ggcttgctgg atgcgcagac ggtcggccat gccccaggcg
9781 tggtcctgac acctggcgag gtccttgtag tagtcctgca tgagccgctc cacgggcacc
9841 tcctcctcgc ccgcgcggcc gtgcatgcgc gtgagcccga acccgcgctg cggctggacg
9901 agcgccaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgagggtg
9961 gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cggtgttgat ggtgtaggag
10021 cagttggcca tgacggacca gttgacggtc tggtggccgg ggcgcacgag ctcgtggtac
10081 ttgaggcgcg agtaggcgcg cgtgtcgaag atgtagtcgt tgcaggtgcg cacgaggtac
10141 tggtatccga cgaggaagtg cggcggcggc tggcggtaga gcggccatcg ctcggtggcg
10201 ggggcgccgg gcgcgaggtc ctcgagcatg aggcggtggt agccgtagat gtacctggac
10261 atccaggtga tgccggcggc ggtggtggag gcgcgcggga actcgcggac gcggttccag
10321 atgttgcgca gcggcaggaa gtagttcatg gtggccgcgg tctggcccgt gaggcgcgcg
10381 cagtcgtgga tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactccgt
```

FIG. 9C

```
10441 ggcctggagg ctaagcgaac gggttgggct gcgcgtgtac cccggttcga gtccctgctc
10501 gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc
10561 taacgaaacc tccaggatac ggaggcgggt cgttttggcc attttcgtca ggccggaaat
10621 gaaactagta agcgcggaaa gcggccgtcc gcgatggctc gctgccgtag tctggagaaa
10681 gaatcgccag ggttgcgttg cggtgtgccc cggttcgagc ctcagcgctc ggcgccggcc
10741 ggattccgcg gctaacgtgg gcgtggctgc cccgtcgttt ccaagacccc ttagccagcc
10801 gacttctcca gttacggagc gagcccctct ttttcttgtg tttttgccag atgcatcccg
10861 tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac
10921 agccggcgct tctgcccccg ccccagcagc agcagccagc cactaccgcg gcggccgccg
10981 tgagcggagc cggcgttcag tatgacctgg ccttggaaga gggcgagggg ctggcgcggc
11041 tgggggcgtc gtcgccggag cggcacccgc gcgtgcagat gaaaagggac gctcgcgagg
11101 cctacgtgcc caagcagaac ctgttcagag acaggagcgg cgaggagccc gaggagatgc
11161 gcgcctcccg cttccacgcg gggcgggagc tgcggcgcgg cctggaccga aagcgggtgc
11221 tgagggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgcacg
11281 tggccgcggc caacctggtc acggcgtacg agcagaccgt gaaggaggag agcaacttcc
11341 aaaaatcctt caacaaccac gtgcgcacgc tgatcgcgcg cgaggaggtg accctgggcc
11401 tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa ccccacgagc aagccgctga
11461 cggcgcagct gtttctggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc
11521 tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga
11581 gcatcgtggt gcaggagcgc gggctgccgc tgtccgagaa gctggcggcc atcaacttct
11641 cggtgctgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgccca
11701 tagacaagga ggtgaagatc gacgggtttt acatgcgcat gaccctgaaa gtgctgaccc
11761 tgagcgacga tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgccagcc
11821 gccggcgcga gctgagcgac caggagctga tgcacagcct gcagcgggcc ctgaccgggg
11881 ccgggaccga gggggagagc tactttgaca tgggcgcgga cctgcgctgg cagcctagcc
11941 gccgggcctt ggaagctgcc ggcggttccc cctacgtgga ggaggtggac gatgaggagg
12001 aggagggcga gtacctggaa gactgatggc gcgaccgtat ttttgctaga tgcagcaaca
12061 gccaccgccg cctcctgatc ccgcgatgcg ggcggcgctg cagagccagc cgtccggcat
12121 taactcctcg gacgattgga cccaggccat gcaacgcatc atggcgctga cgacccgcaa
12181 tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt
12241 ggtgccctcg cgctcgaacc ccacgcacga gaaggtgctg gccatcgtga acgcgctggt
12301 ggagaacaag gccatccgcg gcgacgaggc cgggctggtg tacaacgcgc tgctggagcg
12361 cgtggcccgc tacaacagca ccaacgtgca gacgaacctg gaccgcatgg tgaccgacgt
12421 gcgcgaggcg gtgtcgcagc gcgagcggtt ccaccgcgag tcgaacctgg gctccatggt
12481 ggcgctgaac gccttcctga gcacgcagcc cgccaacgtg ccccggggcc aggaggacta
12541 caccaacttc atcagcgcgc tgcggctgat ggtggccgag gtgccccaga gcgaggtgta
12601 ccagtcgggg ccggactact tcttccagac cagtcgccag ggcttgcaga ccgtgaacct
12661 gagccaggct ttcaagaact tgcagggact gtggggcgtg caggccccgg tcggggaccg
12721 cgcgacggtg tcgagcctgc tgacgccgaa ctcgcgcctg ctgctgctgc tggtggcgcc
12781 cttcacggac agcggcagcg tgagccgcga ctcgtacctg ggctacctgc ttaacctgta
12841 ccgcgaggcc atcgggcagg cgcacgtgga cgagcagacc taccaggaga tcacccacgt
12901 gagccgcgcg ctgggccagg aggacccggg caacctggag gccaccctga acttcctgct
12961 gaccaaccgg tcgcagaaga tcccgcccca gtacgcgctg agcaccgagg aggagcgcat
13021 cctgcgctac gtgcagcaga gcgtggggct gttcctgatg caggaggggg ccacgcccag
13081 cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc
13141 gttcatcaat aagctgatgg actacttgca tcgggcggcc gccatgaact cggactactt
13201 taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta
13261 cgacatgccc gaccccaacg acgggttcct gtgggacgac gtggacagca gcgtgttctc
13321 gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggccgtcctc
13381 ggcgctgtcc ggtcgcgcgg gtgctgccgc ggcggtgccc gaggccgcca gcccttccc
13441 gagcctgccc ttttcgctga acagcgtgcg cagcagcgag ctgggtcggc tgacgcggcc
13501 gcgcctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccg agcgcgagaa
13561 gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta
13621 cgcgcacgag cacagggacg agccccgagc tagcagcagc accggcgcca cccgtagacg
13681 ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacag
13741 cagcgtgttg gacttgggtg ggagtggtgg tggtaacccg ttcgctcacc tgcgcccccg
13801 tatcgggcgc ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg
13861 cgaccagcgt gcgttcttct ctgttgtttg tagtagtatg atgaggcgcg tgtacccgga
```

FIG. 9D

```
13921 gggtcctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcggcgg cgatgcagcc
13981 cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag
14041 cattcgttac tcggagctgg cacccttgta cgataccacc cggttgtacc tggtggacaa
14101 caagtcggcg gacatcgcct cgctgaacta ccagaacgac cacagcaact tcctgaccac
14161 cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga
14221 cgagcgctcg cggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt
14281 gaacgagttc atgtacagca acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc
14341 caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt gggtggagtt
14401 tgagctgccc gagggcaact tctcggtgac catgaccatc gatctgatga acaacgccat
14461 catcgacaac tacttggcgg tggggcggca gaacggggtg ctggagagcg acatcggcgt
14521 gaagttcgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tggtgatgcc
14581 gggcgtgtac accaacgagg ccttccaccc cgacatcgtc ctgctgcccg gctgcggcgt
14641 ggacttcacc gagagccgcc tcagcaacct gctgggcatc cgcaagcggc agcccttcca
14701 ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga
14761 cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcggcgaccg cagccgtggc
14821 caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcagcggc
14881 cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga
14941 gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc
15001 ctacaactac ggcgaccccg agaagggcgt gcgctcctgg acgctgctca ccacctcgga
15061 cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagacccggt
15121 caccttccgc tccacgcgtc aagttagcaa ctacccggtg gtgggcgccg agctcctgcc
15181 cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt
15241 cacctcgctc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgcccgcc
15301 cgcgcccacc attaccaccg tcagtgaaaa cgttcctgct ctcacagatc acgggaccct
15361 gccgctgcgc agcagtatcc ggggagtcca gcgcgtgacc gtcactgacg ccagacgccg
15421 cacctgcccc tacgtctaca aggccctggg cgtagtcgcg ccgcgcgtcc tctcgagccg
15481 caccttctaa aaaatgtcca ttctcatctc gcccagtaat aacaccggtt ggggcctgcg
15541 cgcgcccagc aagatgtacg gaggcgctcg ccaacgctcc acgcaacacc ccgtgcgcgt
15601 gcgcgggcac ttccgcgctc cctggggcgc cctcaagggc cgcgtgcgct cgcgcaccac
15661 cgtcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgccgccgc
15721 gcccgtctcc accgtggacg ccgtcatcga cagcgtggtg gccgacgcgc gccggtacgc
15781 ccgcgccaag agccggcggc ggcgcatcgc ccggcggcac cggagcaccc ccgccatgcg
15841 cgcggcgcga gccttgctgc gcagggccag gcgcacggga cgcagggcca tgctcagggc
15901 ggccagacgc gcggcctccg gcagcagcag cgccggcagg acccgcagac gcgcggccac
15961 ggcggcggcg gcggccatcg ccagcatgtc ccgcccgcgg cgcggcaacg tgtactgggt
16021 gcgcgacgcc gccaccggtg tgcgcgtgcc cgtgcgcacc cgccccctc gcacttgaag
16081 atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaattc
16141 aaggaagaga tgctccaggt catcgcgcct gagatctacg gcccggcggc ggtgaaggag
16201 gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg
16261 gacggactgg tggagtttgt gcgcgagttc gccccccggc ggcgcgtgca gtggcgcggg
16321 cggaaagtga aaccggtgct gcgacccggc accacggtgg tcttcacgcc cggcgagcgt
16381 tccggctccg cctccaagcg ctcctacgac gaggtgtacg gggacgagga catcctcgag
16441 caggcggccg aacgtctggg cgagtttgct tacggcaagc gcagccgccc cgcgcccttg
16501 aaagaggagg cggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg
16561 gtgaccctgc agcaggtgct gcctggtgcg gcgccgcgcc ggggcttcaa gcgcgagggc
16621 ggcgaggatc tgtacccgac catgcagctg atggtgccca agcgccagaa gctggaggac
16681 gtgctggagc acatgaaggt ggaccccgag gtgcagcccg aggtcaaggt gcggcccatc
16741 aagcaggtgg ccccgggcct gggcgtgcag accgtggaca tcaagatccc cacggagccc
16801 atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat
16861 ccctggatgc cggcaccggc ttccaccacc cgccgaagac gcaagtacgg cgcggccagc
16921 ctgctgatgc ccaactacgc gctgcatcct tccatcatcc ccacgccggg ctaccgcggc
16981 acgcgcttct accgcggcta caccagcagc cgccgccgca agaccaccac ccgccgccgc
17041 cgtcgtcgca cccgccgcag cagcaccgcg acttccgccg ccgccctggt gcggagagtg
17101 taccgcagcg ggcgcgagcc tctgaccctg ccgcgcgcgc gctaccaccc gagcatcgcc
17161 atttaactac cgcctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtccccat
17221 tacgggctac cgaggaagaa agccgcgccg tagaaggctg acggggaacg ggctgcgtcg
17281 ccatcaccac cggcggcggc gcgccatcag caagcggttg gggggaggct tcctgcccgc
17341 gctgatgccc atcatcgccg cggcgatcgg ggcgatcccc ggcatagctt ccgtggcggt
```

FIG. 9E

17401 gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga
17461 cgctcctggt cctgtgatgt gtgtttttag atggaagaca tcaatttttc gtccctggca
17521 ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggcaa cagccaactg
17581 aacgggggcg ccttcaattg gagcagtctc tggagcgggc ttaagaattt cgggtccacg
17641 ctcaaaacct atggcaacaa ggcgtggaac agcagcacag ggcaggcgct gagggaaaag
17701 ctgaaagagc agaacttcca gcagaaggtg gtcgatggcc tggcctcggg catcaacggg
17761 gtggtggacc tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc
17821 ccgcccgcgg ggtccgtgga gatgccccag gtggaggagg agctgcctcc cctggacaag
17881 cgcggcgaca agcgaccgcg tcccgacgcg gaggagacgc tgctgacgca cacggacgag
17941 ccgcccccgt acgaggaggc ggtgaaactg ggtctgccca ccacgcggcc cgtggcgcct
18001 ctggccaccg gggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg
18061 cctccgcctg cttcccgccc ctccacagtg gctaagcccc tgccgccggt ggccgtcgcg
18121 tcgcgcgccc cccgaggccg cccccaggcg aactggcaga gcactctgaa cagcatcgtg
18181 ggtctgggag tgcagagtgt gaagcgccgc cgctgctatt aaaagacact gtagcgctta
18241 acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgcgt
18301 cgccgagttg caagatggcc accccatcga tgctgcccca gtgggcgtac atgcacatcg
18361 ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag
18421 acacctactt cagtctgggg aacaagttta ggaaccccac ggtggcgccc acgcacgatg
18481 tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccgtggac cgcgaggaca
18541 acacctactc gtacaaagtg cgctacacgc tggccgtggg cgacaaccgc gtgctggaca
18601 tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaaccct
18661 actccggcac cgcttacaac agcctggctc ccaagggagc gcccaacact tgccagtgga
18721 catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgcctg
18781 tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc
18841 ccatttatgc agataaaact tatcaaccag agcctcaagt gggtgatgct gaatggcatg
18901 acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga
18961 agccctgcta tggttctttt gccaagccta ccaataaaga aggaggtcag gcaaatgtga
19021 aaaccgaaac aggcggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa
19081 gtgcagctgc ggctggcctg gccccagaaa ttgttttgta tactgagaat gtggatctgg
19141 aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca
19201 atttgggtca gcagtccatg cccaacagac ccaactacat tggctttaga gacaacttta
19261 tcgggctcat gtactacaac agcactggca acatgggcgt gctggctggt caggcctccc
19321 agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc
19381 ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct
19441 atgaccccga tgtgcgcatt attgaaaatc acggtgtgga ggatgaactc cctaactatt
19501 gcttcccccc ggatgctgtg ggtagaactg atacttacca gggaattaag gccaatggtg
19561 ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaatgaa ttgggcaagg
19621 gcaatccttt cgccatggag atcaacatcc aggccaacct gtggcggaac ttcctctacg
19681 cgaacgtggc gctgtacctg cccgactcct acaagtacac gccggccaac atcacgctgc
19741 cgaccaacac caacacctac gattacatga acggccgcgt ggtggcgccc tcgctggtgg
19801 acgcctacat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaacccct
19861 tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gctcctgggc aacgggcgct
19921 acgtgccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctcctgctcc
19981 tgcccgggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga
20041 gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc
20101 tctacgccac cttcttcccc atggcgcaca acaccgcctc cacgctcgag gccatgctgc
20161 gcaacgacac caacgaccag tccttcaacg actacctctc ggcggccaac atgctctacc
20221 ccatcccggc caacgccacc aacgtgccca tctccatccc ctcgcgcaac tgggccgcct
20281 tccgcggatg gtccttcacg cgcctcaaga cccgcgagac gccctcgctc ggctccgggt
20341 tcgaccccta cttcgtctac tcgggctcca tcccctacct cgacggcacc ttctacctca
20401 accacacctt caagaaggtc tccatcacct tcgactcctc cgtcagctgg cccggcaacg
20461 accgcctcct gacgcccaac gagttcgaaa tcaagcgcac cgtcgacgga gaggggtaca
20521 acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gcccactaca
20581 acatcggcta ccagggcttc tacgtgcccg agggctacaa ggaccgcatg tactccttct
20641 tccgcaactt ccagcccatg agccgccagg tcgtggacga ggtcaactac aaggactacc
20701 aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctcgcgccca
20761 ccatgcgcca gggacagccc taccccgcca actaccccta cccgctcatc ggcaagagcg
20821 ccgtcgccag cgtcacccag aaaaagttcc tctgcgaccg ggtcatgtgg cgcatccccct

FIG. 9F

20881 tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg
20941 ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccaccc
21001 ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg
21061 tcatcgaggc cgtctacctg cgcacgccct tctcggccgg caacgccacc acctaagccc
21121 cgctcttgct tcttgcaaga tgacggcctg tgcgggctcc ggcgagcagg agctcagggc
21181 catcctccgc gacctgggct gcgggccctg cttcctgggc accttcgaca agcgcttccc
21241 gggattcatg gccccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac
21301 cgggggcgag cactggctgg ccttcgcctg gaacccgcgc tcccacacct gctacctctt
21361 cgaccccttc gggttctcgg acgagcgcct caagcagatc taccagttcg agtacgaggg
21421 cctgctgcgc cgcagcgccc tggccaccga ggaccgctgc gtcaccctgg aaaagtccac
21481 ccagaccgtg cagggtccgc gctcggccgc ctgcgggctc ttctgctgca tgttcctgca
21541 cgccttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgctgac
21601 gggggtgccc aacggcatgc tccagtcgcc ccaggtggaa cccaccctgc gccgcaacca
21661 ggaggcgctc taccgcttcc tcaacgccca ctccgcctac tttcgctccc accgcgcgcg
21721 catcgagaag gccaccgcct tcgaccgcat gaatcaagac atgtaaaccg tgtgtgtatg
21781 tgaatgcttt attcataata aacagcacat gtttatgcca ccttttctga ggctctgact
21841 ttatttagaa atcgaagggg ttctgccggc tctcggcgtg ccccgcgggc agggatacgt
21901 tgcggaactg gtacttgggc agccacttga actcggggat cagcagcttc ggcacgggga
21961 ggtcggggaa cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg
22021 gcgcggagat cttgaaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca
22081 cggggttgca gcactggaac accatcaggg ccgggtgctt cacgctcgcc agcaccgtcg
22141 cgtcggtgat gccctccacg tccagatcct cggcgttggc catcccgaag ggggtcatct
22201 tgcaggtctg ccgccccatg ctgggcacgc agccgggctt gtggttgcaa tcgcagtgca
22261 gggggatcag catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga
22321 aagcctccag ctggcggaag gcctgctgcg ccttgccgcc ctcggtgaag aagaccccgc
22381 aggacttgct agagaactgg ttggtggcgc agccggcgtc gtgcacgcag cagcgcgcgt
22441 cgttgttggc cagctgcacc acgctgcgcc cccagcggtt ctgggtgatc ttggcccggt
22501 cggggttctc cttcagcgcg cgctgcccgt tctcgctcgc cacatccatc tcgatcgtgt
22561 gctccttctg gatcatcacg gtcccgtgca ggcatcgcag cttgccctcg gcctcggtgc
22621 acccgtgcag ccacagcgcg cagccggtgc actcccagtt cttgtgggcg atctgggagt
22681 gcgagtgcac gaagccctgc aggaagcggc ccatcatcgt ggtcagggtc ttgttgctgg
22741 tgaaggtcag cgggatgccg cggtgctcct cgttcacata caggtggcag atgcggcggt
22801 acacctcgcc ctgctcgggc atcagctgga aggcggactt caggtcgctc tccacgcggt
22861 accggtccat cagcagcgtc atgacttcca tgcccttctc ccaggccgag acgatcggca
22921 ggctcagggg gttcttcacc gccgttgtca tcttagtcgc cgccgctgag gtcaggggggt
22981 cgttctcgtc cagggtctca aacactcgct tgccgtcctt ctcggtgatg cgcacggggg
23041 gaaagctgaa gcccacggcc gccagctcct cctcggcctg cctttcgtcc tcgctgtcct
23101 ggctgatgtc ttgcaaaggc acatgcttgg tcttgcgggg tttcttttg ggcggcagag
23161 gcggcggcgg agacgtgctg ggcgagcgcg agttctcgct caccacgact atttcttctt
23221 cttggccgtc gtccgagacc acgcggcggt aggcatgcct cttctggggc agaggcggag
23281 gcgacgggct ctcgcggttc ggcgggcggc tggcagagcc ccttccgcgt tcgggggtgc
23341 gctcctggcg gcgctgctct gactgacttc ctccgcggcc ggccattgtg ttctcctagg
23401 gagcaacaag catggagact cagccatcgt cgccaacatc gccatctgcc cccgccgccg
23461 ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gccgcccagc cccacctccg
23521 acgccgccgc ggccccagac atgcaagaga tggaggaatc catcgagatt gacctgggct
23581 acgtgacgcc cgcggagcac gaggaggagc tggcagcgcg cttttcagcc ccggaagaga
23641 accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcgagc
23701 atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcatctggcc cgccaatgca
23761 tcatcgtcaa ggacgcgctg ctcgaccgcg ccgaggtgcc cctcagcgtg gcggagctca
23821 gccgcgccta cgagcgcaac ctcttctcgc cgcgcgtgcc ccccaagcgc cagcccaacg
23881 gcacctgcga gcccaacccg cgcctcaact tctacccggt cttcgcggtg cccgaggccc
23941 tggccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc
24001 gcacccgcgc cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatcgcct
24061 ccttggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggccgcga
24121 acgctctgca aggaagcgga gaggagcatg agcaccacag cgccctggtg gagttggaag
24181 gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgacccac ttcgcctacc
24241 cggcgctcaa cctgcccccc aaggtcatga gcgccgtcat ggaccaggtg ctcatcaagc
24301 gcgcctcgcc cctctcggat gaggacatgc aggaccccga gagctcggac gagggcaagc

FIG. 9G

```
24361 ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gagtagcacc ccccagagct
24421 tggaagagcg gcgcaagctc atgatggccg tggtcctggt gaccgtggag ctggagtgtc
24481 tgcgccgctt cttcgccgac gcagagaccc tgcgcaaggt cgaggagaac ctgcactacc
24541 tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc
24601 tggtctccta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca
24661 ccctgcgcgg ggaggcccgc cgcgactaca tccgcgactg cgtctacctg tacctctgcc
24721 acacctggca gacgggcatg ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag
24781 agctctgcaa gctcctgcag aagaacctga aggccctgtg gaccgggttc gacgagcgca
24841 ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca
24901 acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct ttcatcctcg
24961 aacgctccgg gatcctgccc gccacctgct ccgcgctgcc ctcggacttc gtgccgctga
25021 ccttccgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact
25081 acctggccta ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc
25141 actgccgctg caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga
25201 gcgagaccca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aaggggggtc
25261 tgaaactcac cccggggctg tggacctcgg cctacttgcg caagttcgtg cccgaggact
25321 accatccctt cgagatcagg ttctacgagg accaatccca gccgcccaag gccgaactgt
25381 cggcctgcgt catcacccag ggggccatcc tggcccaatt gcaagccatc cagaaatccc
25441 gccaagaatt tctgctgaaa aagggccacg gggtctacct ggacccccag accggagagg
25501 agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag
25561 ctgccgccgc cggaggattt ggaggaagac tgggagagca gtcaggcaga ggaggaggag
25621 atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac
25681 gaggtggagg aggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctcggcggag
25741 aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcggcga ccgggcccac
25801 agtaggtggg acgagaccgg gcgcttcccg aaccccacca cccagaccgg taagaaggag
25861 cggcagggat acaagtcctg gcgggggcac aaaaacgcca tcgtctcctg cttgcaagcc
25921 tgcgggggca acatctcctt cacccgccgc tacctgctct tccaccgcgg ggtgaacttc
25981 ccccgcaaca tcttgcatta ctaccgtcac ctccacagcc cctactactg tttccaagaa
26041 gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc
26101 ggcggcaggt ggactgagga tcgcagcgaa cgagccggcg cagacccggg agctgaggaa
26161 ccggatcttt cccaccctct atgccatctt ccagcagagt cgggggcagg agcaggaact
26221 gaaagtcaag aaccgttctc tgcgctcgct cacccgcagt tgtctgtatc acaagagcga
26281 agaccaactt cagcgcactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgct
26341 cactcttaaa gagtagcccg cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc
26401 acctgcgccc ttcgcccgac catcatcatg agcaaagaga ttcccacgcc ttacatgtgg
26461 agctaccagc cccagatggg cctggccgcc ggcgccgccc aggactactc cacccgcatg
26521 aactggctca gcgccgggcc cgcgatgatc tcacgggtga atgacatccg cgcccgccga
26581 aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg
26641 cgtaattggc ccgccgccct ggtgtaccag gaaattcccc agcccacgac cgtactactt
26701 ccgcgagacg cccaggccga agtccagctg actaactcag gtgtccagct ggccggcggc
26761 gccgccctgt gtcgtcaccg ccccgctcag ggtataaagc ggctggtgat ccgaggcaga
26821 ggcacacagc tcaacgacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc
26881 ttccaactcg ccggatcggg gagatcttcc ttcacgcctc gtcaggccgt cctgactttg
26941 gagagttcgt cctcgcagcc ccgctcgggt ggcatcggca ctctccagtt cgtggaggag
27001 ttcactccct cggtctactt caacccccttc tccggctccc ccggccacta cccggacgag
27061 ttcatcccga acttcgacgc catcagcgag tcggtggacg gctacgattg aatgtcccat
27121 ggtggcgcag ctgacctagc tcggcttcga cacctggacc actgtcgcct ctcctacgag
27181 ctcctgcagc agcgccagaa gttcacctgc ctggtcggag tcaaccccat cgtcatcacc
27241 cagcagtcgg gcgataccaa ggggtgcatc cactgctcct gcgactcccc cgactgcgtc
27301 cacactctga tcaagaccct ctgcggcctc cgcgacctcc tccccatgaa ctaatcaccc
27361 ccttatccag tgaaataaag atcatattga tgatttgagt ttaataaaaa taaagaatca
27421 cttacttgaa atctgatacc aggtctctgt ccatgttttc tgccaacacc acttcactcc
27481 cctcttccca gctctggtac tgcaggcccc ggcgggctgc aaacttcctc cacaccctga
27541 aggggatgtc aaattcctcc tgtccctcaa tcttcatttt atcttctatc agatgtccaa
27601 aaagcgcgtc cgggtggatg atgacttcga ccccgtctac ccctacgatg cagacaacgc
27661 accgaccgtg cccttcatca acccccccctt cgtctcttca gatggattcc aagagaagcc
27721 cctgggggtg ctgtccctgc gtctggccga tcccgtcacc accaagaacg gggaaatcac
27781 cctcaagctg ggagatgggg tggacctcga ctcctcggga aaactcatct ccaacacggc
```

FIG. 9H

27841 caccaaggcc gccgcccctc tcagtttttc caacaacacc atttccctta acatggatac
27901 ccctttttac aacaacaatg gaaagttagg catgaaagtc actgctccac tgaagatact
27961 agacacagac ttgctaaaaa cacttgttgt agcttatgga caaggtttag gaacaaacac
28021 cactggtgcc cttgttgccc aactagcatc cccacttgct tttgatagca atagcaaaat
28081 tgcccttaat ttaggcaatg gaccattgaa agtggatgca aatagactga acatcaattg
28141 caatagagga ctctatgtta ctaccacaaa agatgcactg gaagccaata taagttgggc
28201 taatgctatg acatttatag gaaatgccat gggtgtcaat attgatacac aaaaaggctt
28261 gcaatttggc accactagta ccgtcgcaga tgttaaaaac gcttacccca tacaaatcaa
28321 acttggagct ggtctcacat ttgacagcac aggtgcaatt gttgcatgga acaaagatga
28381 tgacaagctt acactatgga ccacagccga cccctctcca aattgtcaca tatattctga
28441 aaaggatgct aagcttacac tttgcttgac aaagtgtggc agtcagattc tgggcactgt
28501 ttccctcata gctgttgata ctggcagttt aaatcccata acaggaacag taaccactgc
28561 tcttgtctca cttaaattcg atgcaaatgg agttttgcaa agcagctcaa cactagactc
28621 agactattgg aatttcagac agggagatgt tacacctgct gaagcctata ctaatgctat
28681 aggtttcatg cccaatctaa aagcataccc taaaaacaca agtggagctg caaaaagtca
28741 cattgttggg aaagtgtacc tacatgggga tacaggcaaa ccactggacc tcattattac
28801 tttcaatgaa acaagtgatg aatcttgcac ttactgtatt aactttcaat ggcagtgggg
28861 ggctgatcaa tataaaaatg aaacacttgc cgtcagttca ttcacctttt cctatattgc
28921 taaagaataa accccactct gtaccccatc tctgtctatg gaaaaaactc tgaaacacaa
28981 aataaaataa agttcaagtg ttttattgat tcaacagttt tacaggattc gagcagttat
29041 ttttcctcca ccctcccagg acatggaata caccaccctc tccccccgca cagccttgaa
29101 catctgaatg ccattggtga tggacatgct tttggtctcc acgttccaca cagtttcaga
29161 gcgagccagt ctcgggtcgg tcagggagat gaaaccctcc gggcactccc gcatctgcac
29221 ctcacagctc aacagctgag gattgtcctc ggtggtcggg atcacggtta tctggaagaa
29281 gcagaagagc ggcggtggga atcatagtcc gcgaacggga tcggccggtg gtgtcgcatc
29341 aggccccgca gcagtcgctg tcgccgccgc tccgtcaagc tgctgctcag ggggtccggg
29401 tccagggact ccctcagcat gatgcccacg gccctcagca tcagtcgtct ggtgcggcgg
29461 gcgcagcagc gcatgcggat ctcgctcagg tcgctgcagt acgtgcaaca caggaccacc
29521 aggttgttca acagtccata gttcaacacg ctccagccga aactcatcgc gggaaggatg
29581 ctacccacgt ggccgtcgta ccagatcctc aggtaaatca agtggcgccc cctccagaac
29641 acgctgccca tgtacatgat ctccttgggc atgtggcggt tcaccacctc ccggtaccac
29701 atcaccctct ggttgaacat gcagccccgg atgatcctgc ggaaccacag ggccagcacc
29761 gccccgcccg ccatgcagcg aagagacccc gggtcccgac aatggcaatg gaggacccac
29821 cgctcgtacc cgtggatcat ctgggagctg aacaagtcta tgttggcaca gcacaggcat
29881 atgctcatgc atctcttcag cactctcagc tcctcggggg tcaaaaccat atcccagggc
29941 acggggaact cttgcaggac agcgaacccc gcagaacagg gcaatcctcg cacataactt
30001 acattgtgca tggacagggt atcgcaatca ggcagcaccg ggtgatcctc caccagagaa
30061 gcgcgggtct cggtctcctc acagcgtggt aaggggccg gccgatacgg gtgatggcgg
30121 gacgcggctg atcgtgttcg cgaccgtgtt atgatgcagt tgctttcgga cattttcgta
30181 cttgctgtag cagaacctgg tccgggcgct gcacaccgat cgccggcggc ggtcccggcg
30241 cttggaacgc tcggtgttga agttgtaaaa cagccactct ctcagaccgt gcagcagatc
30301 tagggcctca ggagtgatga agatcccatc atgcctgatg gctctaatca catcgaccac
30361 cgtggaatgg gccagaccca gccagatgat gcaattttgt tgggtttcgg tgacggcggg
30421 ggagggaaga acaggaagaa ccatgattaa cttttaatcc aaacggtctc ggagcacttc
30481 aaaatgaaga tcgcggagat ggcacctctc gccccgctg tgttggtgga aaataacagc
30541 caggtcaaag gtgatacggt tctcgagatg ttccacggtg gcttccagca aagcctccac
30601 gcgcacatcc agaaacaaga caatagcgaa agcgggaggg ttctctaatt cctcaatcat
30661 catgttacac tcctgcacca tccccagata attttcattt ttccagcctt gaatgattcg
30721 aactagttcc tgaggtaaat ccaagccagc catgataaag agctcgcgca gagcgccctc
30781 caccggcatt cttaagcaca ccctcataat tccaagatat tctgctcctg gttcacctgc
30841 agcagattga caagcggaat atcaaaatct ctgccgcgat ccctaagctc ctccctcagc
30901 aataactgta agtactcttt catatcctct ccgaaatttt tagccatagg accaccagga
30961 ataagattag ggcaagccac agtacagata aaccgaagtc ctccccagtg agcattgcca
31021 aatgcaagac tgctataagc atgctggcta gacccggtga tatcttccag ataactggac
31081 agaaaatcgc ccaggcaatt tttaagaaaa tcaacaaaag aaaaatcctc caggtgcacg
31141 tttagagcct cgggaacaac gatggagtaa atgcaagcgg tgcgttccag catggttagt
31201 tagctgatct gtagaaaaaa acaaaaatga acattaaacc atgctagcct ggcgaacagg
31261 tgggtaaatc gttctctcca gcaccaggca ggccacgggg tctccggcac gaccctcgta

FIG. 9I

```
31321 aaaattgtcg ctatgattga aaaccatcac agagagacgt tcccggtggc cggcgtgaat
31381 gattcgacaa gatgaataca cccccggaac attggcgtcc gcgagtgaaa aaaagcgccc
31441 aaggaagcaa taaggcacta caatgctcag tctcaagtcc agcaaagcga tgccatgcgg
31501 atgaagcaca aaattctcag gtgcgtacaa aatgtaatta ctcccctcct gcacaggcag
31561 caaagccccc gatccctcca ggtacacata caaagcctca gcgtccatag cttaccgagc
31621 agcagcacac aacaggcgca agagtcagag aaaggctgag ctctaacctg tccacccgct
31681 ctctgctcaa tatatagccc agatctacac tgacgtaaag gccaaagtct aaaaataccc
31741 gccaaataat cacacacgcc cagcacacgc ccagaaaccg gtgacacact caaaaaaata
31801 cgcgcacttc ctcaaacgcc caaactgccg tcatttccgg gttcccacgc tacgtcatca
31861 aaattcgact ttcaaattcc gtcgaccgtt aaaaacgtcg cccgccccgc ccctaacggt
31921 cgccgctccc gcagccaatc accgccccgc atccccaaat tcaaatacct catttgcata
31981 ttaacgcgca ccaaaagttt gaggtatatt attgatgatg  (SEQ ID NO: 5)
```

FIG. 9J

```
   1  ATGAAGCGCA CCAAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA TGACACGGAA
  61  AGCGGCCCTC CCTCCGTCCC TTTCCTCACC CCTCCCTTCG TGTCTCCCGA TGGATTCCAA
 121  GAAAGTCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC TTCCCACGGC
 181  ATGCTCGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCTGGCAA CCTCACCTCT
 241  CAAGATATCA CCACCGCTAG CCCTCCCCTC AAAAAAACCA AGACCAACCT CAGCCTAGAA
 301  ACCTCATCCC CCCTAACTGT GAGCACCTCA GGCGCCCTCA CCGTAGCAGC CGCCGCTCCC
 361  CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC AGTACAGGAT
 421  GCAAAACTCA CCCTGGCCAC CAAAGGCCCC CTGACCGTGT CTGAAGGCAA ACTGGCCTTG
 481  CAAACATCGG CCCCGCTGAC GGCCGCTGAC AGCAGCACCC TCACAGTCAG TGCCACACCA
 541  CCCCTTAGCA CAAGCAATGG CAGCTTGGGT ATTGACATGC AAGCCCCCAT TTACACCACC
 601  AATGGAAAAC TAGGACTTAA CTTTGGCGCT CCCCTGCATG TGGTAGACAG CCTAAATGCA
 661  CTGACTGTAG TTACTGGCCA AGGTCTTACG ATAAACGGAA CAGCCCTACA AACTAGAGTC
 721  TCAGGTGCCC TCAACTATGA CACATCAGGA AACCTAGAAT TGAGAGCTGC AGGGGGTATG
 781  CGAGTTGATG CAAATGGTCA ACTTATCCTT GATGTAGCTT ACCCATTTGA TGCACAAAAC
 841  AATCTCAGCC TTAGGCTTGG ACAGGGACCC CTGTTTGTTA ACTCTGCCCA CAACTTGGAT
 901  GTTAACTACA ACAGAGGCCT CTACCTGTTC ACATCTGGAA ATACCAAAAA GCTAGAAGTT
 961  AATATCAAAA CAGCCAAGGG TCTCATTTAT GATGACACTG CTATAGCAAT CAATGCGGGT
1021  GATGGGCTAC AGTTTGACTC AGGCTCAGAT ACAAATCCAT TAAAAACTAA ACTTGGATTA
1081  GGACTGGATT ATGACTCCAG CAGAGCCATA ATTGCTAAAC TGGGAACTGG CCTAAGCTTT
1141  GACAACACAG GTGCCATCAC AGTAGGCAAC AAAAATGATG ACAAGCTCAC CTTGTGGACC
1201  ACACCAGACC CATCTCCTAA CTGTAGAATC TATTCAGAGA AAGATGCTAA ATTCACACTT
1261  GTTTTGACTA AATGCGGCAG TCAGGTGTTG GCCAGCGTTT CTGTTTTATC TGTAAAAGGT
1321  AGCCTTGCGC CCATCAGTGG CACAGTAACT AGTGCTCAGA TTGTCCTCAG ATTTGATGAA
1381  AATGGAGTTC TACTAAGCAA TTCTTCCCTT GACCCTCAAT ACTGGAACTA CAGAAAAGGT
1441  GACCTTACAG AGGGCACTGC ATATACCAAC GCAGTGGGAT TTATGCCCAA CCTCACAGCA
1501  TACCCAAAAA CACAGAGCCA AACTGCTAAA AGCAACATTG TAAGTCAGGT TTACTTGAAT
1561  GGGGACAAAT CCAAACCCAT GACCCTCACC ATTACCCTCA ATGGAACTAA TGAAACAGGA
1621  GATGCCACAG TAAGCACTTA CTCCATGTCA TTCTCATGGA ACTGGAATGG AAGTAATTAC
1681  ATTAATGAAA CGTTCCAAAC CAACTCCTTC ACCTTCTCCT ACATCGCCCA AGAATAA
 (SEQ ID NO: 6)
```

FIG. 10

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGTT GTCCCTGCGA CTGGCCGACC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGATT CCTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTATCCT TACAAGTTTC TCCACCATTA
 361  AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC AGGTTTAGGA
 421  CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTTGA TACTGATGGA
 481  AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC AATTGAAAGC
 541  AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC CAACATTGGA
 601  AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TTGATGATGC TTACCCAATC
 661  CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT GGCTGGTAAC
 721  AAAGAAGACG ATAAACTCAC TTTGTGGACA ACACCTGATC CATCACCAAA CTGTCAAATA
 781  CTCGCAGAAA ATGATGCAAA ACTAACACTT TGCTTGACTA AATGTGGTAG TCAAATACTG
 841  GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAAACCTAA ACCCCATTAC TGGCACCGTA
 901  AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAC AGAACATTCT
 961  ACACTAAAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC TCCATATACC
1021  AATGCTGTAG GATTCATGCC CAATTTAAAA GCTTATCCAA AGTCACAAAG TTCTACTACT
1081  AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAAACC TATGCTTCTC
1141  ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC ATTTTCATAC
1201  ACCTGGACTA ATGGAAGCTA TGTTGGAGCA ACATTTGGGG CTAACTCTTA TACCTTCTCA
1261  TACATCGCCC AAGAATGA  (SEQ ID NO: 7)
```

FIG. 11

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGCT GTCCCTGCGT CTGGCCGATC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGATGGGGTG GACCTCGACG ACTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATACCC CTCTTTACAA CAACAATGGA AAGCTAGGTA TGAAGGTAAC CGCACCATTA
 361  AAGATATTAG ACACAGATCT ACTAAAAACA CTTGTTGTTG CTTATGGGCA GGGATTAGGA
 421  ACAAACACCA ATGGTGCTCT TGTTGCCCAA CTAGCATACC CACTTGTTTT TAATACCGCT
 481  AGCAAAATTG CCCTTAATTT AGGCAATGGA CCATTAAAAG TGGATGCAAA TAGACTGAAC
 541  ATTAATTGCA AAAGAGGTAT CTATGTCACT ACCACAAAAG ATGCACTGGA GATTAATATC
 601  AGTTGGGCAA ATGCTATGAC ATTTATAGGA AATGCCATTG GTGTCAATAT TGACACAAAA
 661  AAAGGCCTAC AGTTCGGCAC TTCAAGCACT GAAACAGATG TTAAAAATGC TTTTTCACTC
 721  CAAGTAAAAC TTGGAGCTGG TCTTACATTT GACAGCACAG GTGCCATTGT TGCTTGGAAC
 781  AAAGAAGATG ACAAACTTAC ACTGTGGACC ACAGCCGATC CATCTCCAAA CTGTCACATA
 841  TATTCTGCAA AGGATGCTAA GCTTACACTC TGCTTGACAA AGTGTGGTAG TCAAATCCTA
 901  GGCACTGTCT CCCTATTAGC AGTCAGTGGC AGCTTGGCTC CTATCACAGG GGCTGTTAGA
 961  ACTGCACTTG TATCACTCAA ATTCAATGCT AATGGAGCCC TTTTGGACAA ATCAACTCTG
1021  AACAAAGAAT ACTGGAACTA CAGACAAGGA GATCTAATTC CAGGTACACC ATATACACAT
1081  GCTGTGGGTT TCATGCCTAA CAAAAAAGCC TACCCTAAAA ACACAACTGC AGCTTCCAAG
1141  AGCCACATTG TGGGTGATGT GTATTTAGAT GGAGATGCAG ATAAACCTTT ATCTCTTATC
1201  ATCACTTTCA ATGAAACTGA TGATGAAACC TGTGATTACT GCATCAACTT TCAATGGAAA
1261  TGGGGAGCTG ATCAATATAA GGATAAGACA CTCGCAACCA GTTCATTCAC CTTCTCATAC
1321 ATCGCCCAAG AATAA  (SEQ ID NO: 8)
```

FIG. 12

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCCGACC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATACCC CTTTTTACAA CAATAATGGA AAGTTAGGCA TGAAAGTCAC TGCTCCACTG
 361  AAGATACTCG ACACAGACTT GCTAAAAACA CTTGTTGTAG CTTATGGACA AGGTTTAGGA
 421  ACAAACACCA CTGGTGCCCT TGTTGCCCAA CTAGCAGCCC CACTTGCTTT TGATAGCAAT
 481  AGCAAAATTG CCCTTAATTT AGGCAATGGA CCATTGAAAG TGGATGCAAA TAGACTGAAC
 541  ATCAATTGCA ATAGAGGACT CTATGTTACT ACCACAAAAG ATGCACTGGA AACCAACATA
 601  AGTTGGGCTA ATGCTATGAC ATTTATAGGA AATGCCATGG GTGTCAATAT TGATACACAA
 661  AAAGGCTTGC AATTTGGCAC CACTAGTACC GTCGCAGATG TTAAAAACGC TTACCCCATA
 721  CAAGTCAAAC TGGGAGCTGG TCTCACATTT GACAGCACAG GTGCAATTGT CGCTTGGAAC
 781  AAAGAAGATG ACAAACTTAC ACTGTGGACC ACAGCCGATC CATCTCCAAA CTGTCACATA
 841  TATTCTGACA AGGATGCTAA GCTTACACTC TGCTTGACAA AGTGTGGCAG TCAGATACTG
 901  GGCACTGTTT CTCTCATAGC TGTTGATACT GGTAGCTTAA ATCCAATAAC AGGACAAGTA
 961  ACCACTGCTC TTGTTTCACT TAAATTCGAT GCCAATGGAG TTTTGCAAAC CAGTTCAACA
1021  TTGGACAAAG AATATTGGAA TTTTAGAAAA GGAGATGTGA CACCTGCTGA GCCATATACT
1081  AATGCTATAG GTTTCATGCC CAATCTAAAG GCATACCCTA AAAACACAAG TGGAGCTGCA
1141  AAAAGTCACA TTGTTGGGAA AGTGTACCTA CATGGGGATA CAGACAAACC ACTGGACCTG
1201  ATTATTACTT TCAATGAAAC AAGTGATGAA TCTTGCACTT ACTGTATTAA CTTTCAATGG
1261  AAATGGGATA GTACTAAGTA CACAGGTGAA ACACTTGCTA CAAGCTCCTT CACCTTCTCC
1321  TACATTGCCC AAGAATGA  (SEQ ID NO: 9)
```

FIG. 13

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGTT GTCCCTGCGA CTGGCCGACC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC TGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTAGCCT TACAAGTTTC TCCACCATTA
 361  AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC AGGTTTAGGA
 421  CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTTGA TACTGATGGA
 481  AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC AATTGAAAGC
 541  AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC CAACATTGGA
 601  AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TCGATGATGC TTACCCAATC
 661  CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT GGCTGGTAAC
 721  AAAGAAGACG ATAAACTCAC TTTGTGGACA ACACCTGATC CATCACCAAA CTGTCAAATA
 781  CTCGCAGAAA ATGATGCAAA ACTAACACTT TGCTTGACTA AATGTGGTAG TCAAATACTG
 841  GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAGACCTAA ACCCCATTAC TGGCACCGTA
 901  AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAC AGAACATTCT
 961  ACACTAAAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC TCCATATGCC
1021  AATGCTGTAG GATTCATGCC CAATTTAAAA GCTTATCCAA AGTCACAAAG TTCTACTACT
1081  AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAAACC TATGCTTCTC
1141  ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC ATTTTCATAC
1201  ACCTGGACTA ATGGAAGCTA TGTTGGAGCA ACATTTGGAG CTAACTCTTA TACCTTCTCC
1261  TACATCGCCC AAGAATGA  (SEQ ID NO: 10)
```

FIG. 14

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCCGACC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATCACC CCTTTTACAC TAAAGATGGA AAATTATCCT TACAAGTTTC TCCACCATTA
 361  AATATACTGA GAACAAGCAT TCTAAACACA CTAGCTTTAG GTTTTGGATC AGGTTTAGGA
 421  CTCCGTGGCT CTGCCTTGGC AGTACAGTTA GTCTCTCCAC TTACATTTGA TACTGATGGA
 481  AACATAAAGC TTACCTTAGA CAGAGGTTTG CATGTTACAA CAGGAGATGC AATTGAAAGC
 541  AACATAAGCT GGGCTAAAGG TTTAAAATTT GAAGATGGAG CCATAGCAAC CAACATTGGA
 601  AATGGGTTAG AGTTTGGAAG CAGTAGTACA GAAACAGGTG TTGATGATGC TTACCCAATC
 661  CAAGTTAAAC TTGGATCTGG CCTTAGCTTT GACAGTACAG GAGCCATAAT GGCTGGTAAC
 721  AAAGAAGACG ATAAACTCAC TTTGTGGACA ACACCTGATC CATCGCCAAA CTGTCAAATA
 781  CTCGCAGAAA ATGATGCAAA ACTAACACTT TGCTTGACTA AATGTGGTAG TCAAATACTG
 841  GCCACTGTGT CAGTCTTAGT TGTAGGAAGT GGAAACCTAA ACCCCATTAC TGGCACCGTA
 901  AGCAGTGCTC AGGTGTTTCT ACGTTTTGAT GCAAACGGTG TTCTTTTAAC AGAACATTCT
 961  ACACTAAAAA AATACTGGGG GTATAGGCAG GGAGATAGCA TAGATGGCAC TCCATATACC
1021  AATGCTGTAG GATTCATGCC CAATTTAAAA GCTTATCCAA AGTCACAAAG TTCTACTACT
1081  AAAAATAATA TAGTAGGGCA AGTATACATG AATGGAGATG TTTCAAAACC TATGCTTCTC
1141  ACTATAACCC TCAATGGTAC TGATGACAGC AACAGTACAT ATTCAATGTC ATTTTCATAC
1201  ACCTGGACTA ATGGAAGCTA TGTTGGAGCA ACATTTGGGG CTAACTCTTA TACCTTCTCA
1261  TACATCGCCC AAGAATGA  (SEQ ID NO: 11)
```

FIG. 15

```
   1  ATGAAGCGCA CCAAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA TGACACGGAA
  61  AACGGTCCTC CCTCCGTCCC TTTCCTCACC CCTCCCTTCG TGTCTCCCGA TGGATTCCAA
 121  GAGAGCCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC TTCCCACGGC
 181  ATGCTCGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCCGGCAA CCTCACCTCT
 241  CAAGATGTCA CCACCACTAC CCCTCCCCTG AAAAAAACCA AGACCAACCT CAGCCTAGAA
 301  ACCTCAGCCC CCCTGACTGT GAGCACCTCA GGCGCCCTCA CCCTAGCAGC CGCCGTTCCC
 361  CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC AGTCCAAGAT
 421  GCAAAACTCA CCCTGGCCAC CAAGGGCCCC CTGACCGTGT CTGAAGGCAA ACTAGCCTTG
 481  CAGACCTCGG CCCCGCTGAC GGCCGCTGAC AGCAGCACCC TCACAATCAG CGCCACACCG
 541  CCCCTTAGCA CAAGCAATGG CAGCTTGGGT ATTGACATGC AAGCCCCCAT TTACACTACT
 601  AACGGAAAAC TGGGACTTAA CTTTGGTGCT CCCCTGCATG TGGTAGACAG CCTAAATGCA
 661  CTGACTGTAG TGACTGGCCA AGGTCTTACG ATAAACGGTA CAGCCCTACA AACTAGAGTC
 721  TCAGGTGCCC TCAACTATGA CTCATCAGGA AACCTAGAAT TGAGAGCTGC AGGGGGTATG
 781  CGAGTTGATG CAAATGGCAA ACTTATCCTT GACGTAGCTT ACCCATTTGA TGCTCAAAAC
 841  AACCTCAGCC TTAGACTTGG ACAGGGACCC CTGTTTGTTA ACTCTGCCCA CAACTTGGAT
 901  GTTAACTACA ACAGAGGCCT CTACCTGTTC ACATCTGGAA ATACCAAAAA GCTAGAAGTT
 961  AATATCAAAA CAGCCAAAGG CCTCATTTAT GATGACACTG CTATAGCAAT CAATCCAGGC
1021  GATGGGCTAG AGTTTGGCTC AGGCTCAGAT ACAAATCCAT TAAAAACTAA ACTTGGATTG
1081  GGACTAGAGT ATGACTCCAG CAGAGCCATA ATTGCTAAGC TGGGAACCGG CCTAAGCTTT
1141  GACAACACAG GTGCCATCAC AGTGGGCAAC AAAAATGATG ACAAGCTTAC CTTGTGGACC
1201  ACACCAGACC CCTCTCCCAA CTGTAGAATT TATTCAGAAA AAGATGCTAA ATTTACACTA
1261  GTTTTAACTA AATGCGGCAG TCAGGTGTTG GCCAGCGTTT CTGTTTTATC TGTAAAAGGC
1321  AGCCTTGCGC CCATCAGTGG CACAGTAACT AGCGCTCAGA TTATTCTCAG ATTTGATGAA
1381  AATGGAGTTC TACTAAGCAA TTCTTCTCTT GACCCCCAAT ACTGGAACTA CAGAAAAGGT
1441  GACCTTACAG AGGGCACTGC ATATACCAAC GCAGTGGGAT TTATGCCCAA CCTCACAGCA
1501  TACCCAAAAA CACAGAGTCA AACTGCTAAA AGCAACATTG TAAGCCAGGT TTACTTGAAT
1561  GGGGACAAAT CCAAACCCAT GATCCTCACC ATTACCCTCA ATGGAACTAA TGAAACAGGG
1621  GATGCTACAG TTAGCACTTA CTCCATGTCA TTCTCATGGA ATTGGAATGG AAGTAATTAC
1681  ATTAATGAAA CGTTCCAAAC CAACTCTTTC ACCTTCTCCT ACATCGCCCA AGAATAA
(SEQ ID NO: 12)
```

FIG. 16

```
   1  ATGTCCAAAA AGCGCGTCCG GGTGGATGAT GACTTCGACC CCGTCTACCC CTACGATGCA
  61  GACAACGCAC CGACCGTGCC CTTCATCAAC CCCCCCTTCG TCTCTTCAGA TGGATTCCAA
 121  GAGAAGCCCC TGGGGGTGCT GTCCCTGCGA CTGGCTGACC CCGTCACCAC CAAGAACGGG
 181  GAAATCACCC TCAAGCTGGG AGAGGGGGTG GACCTCGACT CCTCGGGAAA ACTCATCTCC
 241  AACACGGCCA CCAAGGCCGC CGCCCCTCTC AGTTTTTCCA ACAACACCAT TTCCCTTAAC
 301  ATGGATACCC CTTTTTACAC CAAAGATGGA AAATTAACCA TGCAGGTCAC TGCACCACTA
 361  AAGTTAGCAA ACACAGCCAT ATTGAACACA CTAGCTATGG CATATGGAAA TGGATTAGGT
 421  CTAAGCAACA ACGCTCTTAC CGTTCAGTTA CAATCTCCAC TCACCTTTAA CAACAGCAAG
 481  GTTGCAATCA ACCTGGGAAA TGGACCACTA AATGTAACAT CAAACAGACT TAGCATTAAT
 541  TGCAAGAGGG GTGTCTATGT CACCACCACA GGAGATGCAA TTGAAACCAA CATAAGTTGG
 601  TCAAATGCTA TTAAATTTAT AGGAAATGCC ATGGGTGTCA ACATTGATAC AAACAAAGGC
 661  TTGCAATTTG GCACCACCAG CACTGTCACA GATGTGACCA ATGCTTTCCC CATACAAGTC
 721  AAACTTGGGG CTGGTCTTGC ATTTGATAGC ACTGGAGCTA TTGTTGCATG GAACAAAGAG
 781  GATGACAGTC TCACTTTGTG GACTACACCA GATCCATCTC CAAATTGCAA GATAGCATCT
 841  GACAAAGATG CTAAACTCAC ACTTTGCTTG ACAAAATGTG GTAGTCAGAT ACTGGGCACT
 901  GTCTCCTTGT TAGCTGTGAG TGGCAGTTTA GCTCCTATCA CTGGAGCTGT GAGCACTGCA
 961  CTTGTATCAC TTAAATTCGA TGCCAATGGA GCACTCTTGG AAAAATCAAC CCTAAACAGA
1021  GAATATTGGA ACTATAGACA AGGAGATCTT ATTCCAGGTA CGCCATATAC TCACGCAGTA
1081  GGTTTCATGC CCAACAAGAA AGCCTACCCT AAAAACACAA CTGCAGCTTC CAAAAGCCAC
1141  ATTGTGGGAG AAGTCTATCT AGACGGAGAT GCAGATAAGC CCCTATCTCT CATAATCACT
1201  TTTAATGAAA CTGATGATGA ATCATGTGAC TATTGCATGA ACTTTCAATG GAAATGGGGT
1261  GCTGATCAAT ACAAGGACAA AACACTCGCT ACCAGCTCCT TCACCTTCTC CTACATTGCC
1321  CAAGAATGA  (SEQ ID NO: 13)
```

FIG. 17

```
   1  ATGAAGCGCA CCAAAACGTC TGACGAGAGC TTCAACCCCG TGTACCCCTA TGACACGGAA
  61  AGCGGCCCTC CCTCCGTCCC TTTCCTCACC CCTCCCTTCG TGTCTCCCGA TGGATTCCAA
 121  GAAAGCCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC TTCCCACGGC
 181  ATGCTTGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCTGGCAA CCTTACCTCT
 241  CAAGATATTA CCTCCACTAC CCCTCCCCTC AAAAAAACCA AGACCAACCT CAGCCTAGAA
 301  ACCTCATCCC CCCTAACTGT AAGCACCTCA GGCGCCCTCA CCGTAGCAGC CGCCGCTCCC
 361  CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGGC AGTACAGGAT
 421  GCAAAACTCA CCCTGGCCAC CAAAGGCCCC CTGACCGTGT CTGAAGGCAA ACTGGCCTTG
 481  CAAACATCGG CCCCGCTGAC GGCCGCTGAC AGCAGCACCC TCACCGTTAG CTCCACTCCA
 541  CCAATTAGTG TAAGCAGTGG AAGTTTGGGC TTGGACATGG AAGACCCCAT GTATACTCAC
 601  GATGGAAAAC TGGGAATAAG AATTGGGGGT CCACTAAGAG TAGTAGACAG CTTGCACACA
 661  CTCACTGTAG TTACCGGAAA TGGACTAACT GTAGATAACA ATGCCCTCCA AACTAGAGTT
 721  ACGGGCGCCC TAGGTTATGA CACATCAGGA AATCTACAAC TGAGAGCCGC AGGGGGTATG
 781  CGAATTGATG CAAATGGCCA ACTTATCCTT GATGTGGCAT ACCCATTTGA TGCTCAAAAC
 841  AATCTCAGCC TTAGACTTGG TCAGGGACCC CTGTATGTAA ATACAGACCA CAACCTGGAT
 901  TTAAATTGCA ACAGAGGTCT AACCACAACT ACCACCAACA ACACAAAAAA ACTTGAGACT
 961  AAAATTAGCT CAGGCTTAGA CTATGACACC AATGGTGCTG TCATTATTAA ACTTGGCACT
1021  GGTCTAAGCT TCGACAACAC AGGCGCCCTA ACTGTGGGAA ACACTGGTGA TGATAAACTG
1081  ACTCTGTGGA CGACCCCAGA CCCATCTCCA AATTGCAGAA TTCACTCAGA CAAAGACTGC
1141  AAGTTTACTC TCGTCCTAAC TAAGTGTGGA AGCCAAATCC TGGCCTCTGT CGCCGCCCTA
1201  GCGGTATCAG GAAATCTGGC TTCGATAACA GGCACCGTTG CCAGCGTTAC CATCTTTCTT
1261  AGATTTGATC AGAATGGAGT GCTTATGGAA AACTCCTCAC TAGACAAGCA GTACTGGAAC
1321  TTCAGAAATG GCAATTCAAC TAATGCTGCC CCCTACACCA ACGCAGTTGG GTTCATGCCA
1381  AACCTCGCAG CGTACCCCAA AACGCAAAGC CAGACTGCTA AAAACAACAT TGTAAGTCAG
1441  GTTTACTTGA ATGGAGACAA ATCCAAACCC ATGACCCTTA CCATCACCCT CAATGGAACT
1501  AATGAATCCA GTGAAACTAG TCAGGTGAGT CACTACTCCA TGTCATTTAC ATGGGCTTGG
1561  GAAAGCGGGC AATATGCCAC TGAAACCTTT GCCACCAACT CCTTCACCTT TTCTTACATT
1621  GCTGAACAAT AA  (SEQ ID NO: 14)
```

FIG. 18

```
   1  ATGAAGCGCA CCAAAACGTC TGACAAGAGC TTCAACCCCG TGTACCCCTA TGACACGGAA
  61  AACGGTCCTC CCTCCGTCCC TTTCCTCACC CCTCCCTTCG TGTCTCCCGA TGGATTCCAA
 121  GAGAGCCCCC CCGGGGTCCT GTCTCTGAAC CTGGCCGAGC CCCTGGTCAC TTCCCACGGC
 181  ATGCTCGCCC TGAAAATGGG AAGTGGCCTC TCCCTGGACG ACGCCGGCAA CCTCACCTCT
 241  CAAGATGTCA CCACCACTAC CCCTCCCCTG AAAAAAACCA AGACCAACCT CAGCCTAGAA
 301  ACCTCAGCCC CCCTGACTGT GAGCACCTCA GGCGCCCTCA CCCTAGCAGC CGCCGCCCCC
 361  CTGGCGGTGG CCGGCACCTC CCTCACCATG CAATCAGAGG CCCCCCTGAC AGTCCAAGAT
 421  GCAAAACTCA CCCTGGCCAC CAAGGGCCCC CTGACCGTGT CTGAAGGCAA ACTGGCCTTG
 481  CAGACCTCGG CCCCGCTGAC GGCCGCTGAC AGCAGCACCC TCACCGTTAG CGCCACACCA
 541  CCCATCAGTG TAAGCAGTGG AAGTTTGGGC TTAGACATGG AAGACCCAAT GTATACTCAT
 601  GATGGAAAAC TGGGAATAAG AATTGGGGGC CCACTGAGAG TAGTAGACAG CCTGCACACA
 661  CTGACTGTAG TTACCGGAAA TGGAATAGCT GTAGATAACA ATGCCCTCCA AACTAGAGTT
 721  ACGGGCGCCC TGGGTTATGA CACATCAGGA AACCTACAAC TGAGAGCCGC GGGGGGTATG
 781  CGAATTGATG CAAATGGCCA ACTTATCCTT GATGTGGCAT ACCCATTTGA TGCTCAAAAC
 841  AATCTCAGCC TTAGACTTGG TCAGGGACCC CTGTATGTAA ACACAGACCA CAACCTAGAT
 901  TTGAATTGCA ACAGAGGTCT GACCACAACT ACCACCAACA ACACAAAAAA ACTTGAAACT
 961  AAAATTGGCT CAGGCTTAGA CTATGATACC AATGGTGCTG TTATTATTAA ACTTGGCACT
1021  GGTGTCAGCT TTGACAGCAC AGGTGCCCTA AGTGTGGGAA ACACTGGCGA TGATAAACTG
1081  ACTCTGTGGA CAACCCCAGA CCCATCTCCA AATTGCAGAA TTCACTCAGA CAAAGACTGC
1141  AAGTTTACTC TAGTCCTAAC TAAGTGTGGA AGTCAAATCC TGGCTTCTGT CGCCGCCCTA
1201  GCGGTGTCAG GAAATCTGGC TTCAATAACA GGCACCGTTT CCAGCGTTAC CATCTTTCTC
1261  AGATTTGATC AGAATGGAGT GCTTATGGAA AACTCCTCGC TAGACAAGCA GTACTGGAAC
1321  TTCAGAAATG GTAATTCAAC CAATGCCACC CCCTACACCA ATGCAGTTGG GTTTATGCCA
1381  AACCTCGCAG CATACCCCAA GACACAGAGC CAGACTGCAA AAAACAACAT TGTAAGTCAG
1441  GTTTACTTGA ATGGGGACAA ATCCAAACCC ATGACCCTTA CCATTACCCT CAATGGAACT
1501  AATGAATCCA GTGAAACTAG CCAGGTGAGT CACTACTCCA TGTCATTTAC GTGGGCTTGG
1561  GAGAGTGGGC AATATGCCAC CGAAACCTTT GCCACCAATT CCTTTACCTT CTCTTACATT
1621  GCTGAACAAT AA  (SEQ ID NO: 15)
```

FIG. 19

```
1                                                                                                   100
     C1 MAKRTRLSSS .FNPVYPYED EN..SSHPFI NPGFISPNGF TQSPDGVLTL NCVAPLTTAN GALDIKVGGG LKVNSTDGFL EENIN..... ..........
   CV68 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
   PAN5 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGDG VDLDSSGKLI SNTAT..... ..........
   PAN6 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
   PAN7 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
  CHAD3 MKRTKTSDES .FNPVYPYDT ESGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDITTASPP LKKTKTNLSL
  CHAD4 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
  CHAD5 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGDG VDLDDSGKLI SNTAT..... ..........
  CHAD6 MSKKRAARVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GAVTLKLGEG VDLDDSGKLI SKNAT..... ..........
  CHAD7 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
  CHAD8 MT.KRVRLSS SFNPVYPYED ES.SSQHPFI NPGFISSNGF TQSPDGVLTL KCLSPLTTTG GSLQLKVGGG LSVDDTDGSL EENIS..... ..........
  CHAD9 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD10 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD11 MKRTKTSDES .FNPVYPYDT ENGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDVTTTTPP LKKTKTNLSL
 CHAD16 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD17 MKRTKTSDES .FNPVYPYDT ESGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDITSTTPP LKKTKTNLSL
 CHAD19 MKRTKTSDKS .FNPVYPYDT ENGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDVTTTTPP LKKTKTNLSL
 CHAD20 MKRTKTSDES .FNPVYPYDT ESGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDITTASPP LKKTKTNLSL
 CHAD22 MA.KRARLSS SFNPVYPYED ES.SSQHPFI NPGFISSNGF AQSPDGVLTL KCVNPLTTAS GPLQLKVGSS LTVDNIDGSL EENIT..... ..........
 CHAD24 MKRTKTSDES .FNPVYPYDT ENGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDVTTTTPP LKKTKTNLSL
 CHAD26 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGDG VDLDDSGKLI SNTAT..... ..........
 CHAD30 MA.KRARLSS SFNPVYPYED ES.SSQHPFI NPGFISSNGF AQSPDGVLTL KCVNPLTTAS GPLQLKVGSS LTVDTIDGSL EENIT..... ..........
 CHAD31 MKRTKTSDES .FNPVYPYDT ESGPPSVPFL TPPFVSPDGF QESPPGVLSL NLAEPLVTSH GMLALKMGSG LSLDDAGNLT SQDITTASPP LKKTKTNLSL
 CHAD37 MA.KRARLSS SFNPVYPYED ES.SSQHPFI NPGFISPDGF TQSPDGVLTL KCVSPLTTTS GALDIKVGRG LKVDSTDGSL EENID..... ..........
 CHAD38 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD44 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD63 MSKKRVRVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GEITLKLGEG VDLDSSGKLI SNTAT..... ..........
 CHAD82 MSKKRAARVDD DFDPVYPYDA DN.APTVPFI NPPFVSSDGF QEKPLGVLSL RLADPVTTKN GAVTLKLGEG VDLDDSGKLI SKNAT..... ..........
```

FIG. 20A

```
                                                
```

```
                201                                                
```

```
                                                
```

```
401                                                                                                 500
C1      ....NTLWTG AKPSANCVIK EGEDSPDCKL TLVLVKNGGL VNGYITLMGD SE.YTNTLFK NKQVTIDVNL AFDNTGQIIT YLSSLKS.NL NFKDNQNMAT
CV68    EDDKLTLWTT PDPSPNCQIL AE...NDAKL TLCLTKCGSQ ILATVSVLVV GS.GNLNPIT GTVSSAQVFL RFDANGVLLT E.HSTLKKYW GYRQ.GDSID
PAN5    DDDKLTLWTT ADPSPNCHIY SE...KDAKL TLCLTKCGSQ ILGTVSLIAV DT.GSLNPIT GTVTTALVSL KFDANGVLQS S.STLDSDYW NFRQ.GDVTP
PAN6    QDDKLTLWTT PDPSPNCQLL SD...RDAKF TLCLTKCGSQ ILGTVAVAAV TVGSALNPIN DTVKSAIVFL RFDSDGVLMS N.SSMVGDYW NFRE.GQTTQ
PAN7    EDDKLTLWTT ADPSPNCKIY SE...KDAKL TLCLTKCGSQ ILGTVTVLAV NN.GSLNPIT NTVSTALVSL KFDASGVLLS S.STLDKEYW NFRK.GDVTP
CHAD3   GDDKLTLWTT PDPSPNCRIH SD...KDCKF TLVLTKCGSQ ILASVAALAV S..GNLASIT GTVASVTIFL RFDQNGVLME N.SSLDRQYW NFRN.GNSTN
CHAD4   EDDKLTLWTT PDPSPNCQIL AE...NDAKL TLCLTKCGSQ ILATVSVLVV GS.GNLNPIT GTVSSAQVFL RFDANGVLLT E.HSTLKKYW GYRQ.GDSID
CHAD5   EDDKLTLWTT ADPSPNCHIY SA...KDAKL TLCLTKCGSQ ILGTVSLLAV S..GSLAPIT GAVRTALVSL KFNANGALLD K.STLNKEYW NYRQ.GDLIP
CHAD6   EDDKLTLWTT ADPSPNCHIY SA...KDAKL TLCLTKCGSQ ILGTVSLIAV DT.GSLNPIT GKVTTALVSL KFDANGVLQA S.STLDKEYW NFRK.GDVTP
CHAD7   EDDKLTLWTT ADPSPNCHIY SD...KDAKL TLCLTKCGSQ ILGTVSLIAV DT.GSLNPIT GQVTTALVSL KFDANGVLQT S.STLDKEYW NFRK.GDVTP
CHAD8   N.....TLWTG ATPDANCLVL GTES.NDCKL TLALVKSGAL VNAYVALVGA SD.AVNDLTT ETSAQIIADI YFDAQGKLLP DLSALKTELK HKSG..QGTS
CHAD9   EDDKLTLWTT PDPSPNCQIL AE...NDAKL TLCLTKCGSQ ILATVSVLVV GS.GDLNPIT GTVSSAQVFL RFDANGVLLT E.HSTLKKYW GYRQ.GDSID
CHAD10  EDDKLTLWTT PDPSPNCQIL AE...NDAKL TLCLTKCGSQ ILATVSVLVV GS.GNLNPIT GTVSSAQVFL RFDANGVLLT E.HSTLKKYW GYRQ.GDSID
CHAD11  NDDKLTLWTT PDPSPNCRIY SE...KDAKF TLVLTKCGSQ VLASVSVLSV K..GSLAPIS GTVTSAQIIL RFDENGVLLS N.SSLDPQYW NYRK.GDLTE
CHAD16  EDDSLTLWTT PDPSPNCKIA SD...KDAKL TLCLTKCGSQ ILGTVSLLAV S..GSLAPIT GAVSTALVSL KFDANGALLE K.STLNREYW NYRQ.GDLIP
CHAD17  GDDKLTLWTT PDPSPNCRIH SD...KDCKF TLVLTKCGSQ ILASVAALAV S..GNLASIT GTVASVTIFL RFDQNGVLME N.SSLDKQYW NFRN.GNSTN
CHAD19  GDDKLTLWTT PDPSPNCRIH SD...KDCKF TLVLTKCGSQ ILASVAALAV S..GNLASIT GTVSSVTIFL RFDQNGVLME N.SSLDKQYW NFRN.GNSTN
CHAD20  NDDKLTLWTT PDPSPNCRIY SE...KDAKF TLVLTKCGSQ VLASVSVLSV K..GSLAPIS GTVTSAQIVL RFDENGVLLS N.SSLDPQYW NYRK.GDLTE
CHAD22  ....NTLWTG AKPSANCVIK EGEDSPDCKL TLVLVKNGGL INGYITLMGA S.EYTNTLFK NKQVTIDVNL AFDNTGQIIT YLSSLKS.NL NFKDNQNMAT
CHAD24  GDDKLTLWTT PDPSPNCRIH SD...KDCKF TLVLTKCGSQ ILASVAALAV S..GNLASIT GTVSSVTIFL RFDQNGVLME N.SSLDKQYW NFRN.GNSTN
CHAD26  EDDKLTLWTT ADPSPNCHIY SA...KDAKL TLCLTKCGSQ ILGTVSLLAV S..GSLAPIT GAVRTALVSL KFNANGALLD K.STLNKEYW NYRQ.GDLIP
CHAD30  ....NTLWTG AKPSANCVIK EGEDSPDCKL TLVLVKNGGL INGYITLMGA S.EYTNTLFK NKQVTIDVNL AFDNTGQIIT YLSSLKS.NL NFKDNQNMAT
CHAD31  NDDKLTLWTT PDPSPNCRIY SE...KDAKF TLVLTKCGSQ VLASVSVLSV K..GSLAPIS GTVTSAQIVL RFDENGVLLS N.SSLDPQYW NYRK.GDLTE
CHAD37  N.....TLWTG VNPSANCIIT DNGETNDSKL TLILVKSGGL INAYVSLMGD SD.TVNKLTT EKSAQITVDI YFDNQGKVLT ELSALKTDLK HKFG..QNMA
CHAD38  EDDKLTLWTT ADPSPNCKIY SE...KDAKL TLCLTKCGSQ ILGTVTVLAV NN.GSLNPIT NTVSTALVSL KFDASGVLLS S.STLDKEYW NFRK.GDVTP
CHAD44  QDDKLTLWTT PDPSPNCQLL SD...RDAKF TLCLTKCGSQ IIGTVAVAAV TVGSALNPIN DTVKSAIVFL RFDSDGVLMS N.SSMVGDYW NFRE.GQTTQ
CHAD63  EDDKLTLWTT PDPSPNCQIL AEN...DAKL TLCLTKCGSQ ILATVSVLVV GS.GNLNPIT GTVSSAQVFL RFDANGVLLT EHSTLKK.YW GYRQ.GDSID
CHAD82  EDDKLTLWTT ADPSPNCHIY SA...KDAKL TLCLTKCGSQ ILGTVSLIAV DT.GSLNPIT GKVTTALVSL KFDANGVLQA S.STLDKEYW NFRK.GDVTP
```

FIG. 20E

```
                                                
```

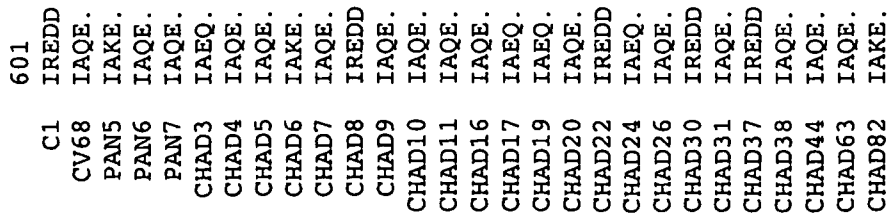
601
C1 IREDD
CV68 IAQE.
PAN5 IAKE.
PAN6 IAQE.
PAN7 IAQE.
CHAD3 IAEQ.
CHAD4 IAQE.
CHAD5 IAQE.
CHAD6 IAKE.
CHAD7 IAQE.
CHAD8 IREDD
CHAD9 IAQE.
CHAD10 IAQE.
CHAD11 IAQE.
CHAD16 IAQE.
CHAD17 IAEQ.
CHAD19 IAEQ.
CHAD20 IAQE.
CHAD22 IREDD
CHAD24 IAEQ.
CHAD26 IAQE.
CHAD30 IREDD
CHAD31 IAQE.
CHAD37 IREDD
CHAD38 IAQE.
CHAD44 IAQE.
CHAD63 IAQE.
CHAD82 IAKE.

FIG. 20G

```
   1  ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG CCAGGACGCC
  61  TCGGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCGCG CCACCGAGAG CTACTTCAGC
 121  CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGG
 181  TCTCAGCGCC TGACGCTGCG GTTCATTCCC GTGGACCGCG AGGACACCGC GTACTCGTAC
 241  AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CTCCACCTAC
 301  TTTGACATCC GCGGGGTGCT GGACCGGGGT CCCACTTTCA AGCCCTACTC TGGCACCGCC
 361  TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACCCATGCG AGTGGGATGA GGCTGCTACT
 421  GCCCTTGACA TTGATTTGAA CGCAGAAGAC GATGAAGAAA GCGACGAAGC TCAAGGGGAA
 481  GCAGATCAGC AGAAAACTCA TGTATTTGGC CAGGCGCCCT ACTCCGGACA GAACATTACA
 541  AAAGAAGGCA TACAGATAGG CATAGATGCT GCCAGTCAAG CCCAGACACC TGTATATGCC
 601  GATAAAACAT TCCAACCAGA ACCTCAAGTT GGAGAATCAC AGTGGAATGA GACAGAGATT
 661  AGTTATGGAG CGGGACGGGT GCTTAAAAAA ACCACTCTCA TGAAACCTTG CTATGGGTCG
 721  TATGCAAGGC CTACTAATGA GAACGGAGGT CAGGGCATCC TCTTGGAACA AGATGGAAAG
 781  AAAGAAAGTC AAGTGGAAAT GCAATTTTTC TCTACTACTC AGGCAGCCGC GGGTAATTCA
 841  GATAATCCTA CCCCAAAGGT TGTTTTGTAC AGCGAGGATG TTAACCTGGA AACACCAGAT
 901  ACACACATTT CATACATGCC CACCAACAAC GAGACAAATT CAAGAGAGCT TTTGGGACAA
 961  CAGGCCATGC CCAACAGGCC TAATTACATT GGCTTCAGAG ACAACTTTAT CGGTCTCATG
1021  TATTACAACA GCACTGGCAA CATGGGAGTG CTTGCAGGTC AGGCCTCTCA GTTGAACGCA
1081  GTGGTGGACT TGCAAGACAG AAACACAGAA CTGTCATACC AGCTCTTGCT TGATTCCATG
1141  GGTGACAGAA CCAGATACTT TTCCATGTGG AATCAGGCAG TGGACAGTTA TGACCCAGAT
1201  GTCAGAATTA TTGAAAATCA TGGAACTGAA GACGAGCTCC CCAACTATTG TTTCCCTCTG
1261  GGCGGCGTAA TCAATACGGA AACTTTCACA AAAGTAAAAC CTAAAGCTGC ACAGGACGCT
1321  CAGTGGGAAA AAGATTCAGA ATTTTCAGAT AAAAATGAAA TAAGGGTGGG AAACAACTTC
1381  GCCATGGAAA TTAACCTCAA TGCCAATCTG TGGAGGAACT TTTTGTACTC CAACGTAGCC
1441  CTCTACTTGC CTGACAAGCT TAAGTATACT CCATCCAATG TGCAAATTTC CAACAATCCC
1501  AACTCCTACG ATTACATGAA CAAGCGAGTG GTGGCCCCGG GGCTGGTGGA CTGCTACATC
1561  AACCTGGGCG CGCGCTGGTC GCTGGACTAC ATGGACAACG TCAACCCCTT CAACCACCAC
1621  CGCAATGCGG GCCTGCGCTA CCGCTCCATG CTCCTGGGCA ACGGGCGCTA CGTGCCCTTC
1681  CACATCCAGG TGCCCCAGAA GTTCTTTGCC ATCAAGAACC TCCTCCTCCT GCCGGGCTCC
1741  TACACCTACG AGTGGAACTT CAGGAAGGAT GTCAACATGG TCCTCCAGAG CTCTCTGGGT
1801  AACGATCTCA GGGTGGACGG GGCCAGCATC AAGTTCGAGA GCATCTGCCT CTACGCCACC
1861  TTCTTCCCCA TGGCCCACAA CACGGCCTCC ACGCTCGAGG CCATGCTCAG GAACGACACC
1921  AACGACCAGT CCTTCAATGA CTACCTTTCC GCCGCCAACA TGCTCTACCC CATACCCGCC
1981  AACGCCACCA ACGTCCCCAT CTCCATCCCC TCGCGCAACT GGGCGGCCTT CCGCGGCTGG
2041  GCCTTCACCC GCCTCAAGAC CAAGGAGACC CCCTCCCTGG GCTCGGGATT CGACCCCTAC
2101  TACACCTACT CGGGCTCCAT TCCCTACCTG GACGGCACCT TCTACCTCAA CCACACTTTC
2161  AAGAAGGTCT CGGTCACCTT CGACTCCTCG GTCAGCTGGC CGGGCAACGA CCGTCTGCTC
2221  ACCCCCAACG AGTTCGAGAT CAAGCGCTCG GTCGACGGGG AGGGCTACAA CGTGGCCCAG
2281  TGCAACATGA CCAAGGACTG GTTCCTGGTC CAGATGCTGG CCAACTACAA CATCGGCTAC
2341  CAGGGCTTCT ACATCCCAGA GAGCTACAAG GACAGGATGT ACTCCTTCTT CAGGAACTTC
2401  CAGCCCATGA GCCGGCAGGT GGTGGACCAG ACCAAGTACA AGGACTACCA GGAGGTGGGC
2461  ATCATCCACC AGCACAACAA CTCGGGCTTC GTGGGCTACC TCGCCCCCAC CATGCGCGAG
2521  GGACAGGCCT ACCCCGCCAA CTTCCCCTAC CCGCTCATAG GCAAGACCGC GGTCGACAGC
2581  ATCACCCAGA AAAAGTTCCT CTGCGACCGC ACCCTCTGGC GCATCCCCTT CTCCAGCAAC
2641  TTCATGTCCA TGGGTGCGCT CTCGGACCTG GGCCAGAACT TGCTCTACGC CAACTCCGCC
2701  CACGCCCTCG ACATGACCTT CGAGGTCGAC CCCATGGACG AGCCCACCCT TCTCTATGTT
2761  CTGTTCGAAG TCTTTGACGT GGTCCGGGTC CACCAGCCGC ACCGCGGCGT CATCGAGACC
2821  GTGTACCTGC GTACGCCCTT CTCGGCCGGC AACGCCACCA CCTAA (SEQ ID NO: 16)
```

FIG. 21

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCCTACTC CGGCACCGCC
 361  TACAACAGCC TGGCTCCCAA GGGAGCGCCC AATTCCAGCC AGTGGGAGCA AAAAAAGACT
 421  GGCAATAATG CCAATGGAGA TACGGAGAAT GTCACTTATG GTGTAGCTGC CATGGGAGGA
 481  ATTGACATCG ATAAAAATGG CCTTCAAATT GGAACCGATG ACACCAAAGA TGACGATAAT
 541  GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAAA TAGGAGAGGA AAACTGGCAA
 601  GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAAG ATACCAAAAT GAAGCCATGC
 661  TATGGCTCAT TTGCCAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAAT AAAAACAGAT
 721  GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTTG ATATTCCCAA TTCTGGCGCG
 781  GGAAATGGCA CAAATGTTAA CGATGATCCA GATATGGTTA TGTATACAGA AAATGTAAAT
 841  CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTCAGATGA CAGCTCAAAG
 901  GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT CAGAGACAAT
 961  TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC TGGTCAGGCC
1021  TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC CTACCAGCTC
1081  TTGCTTGACT CTCTGGGTGA CAGAACCAGG TATTTCAGTA TGTGGAATCA GGCGGTGGAC
1141  AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA ATTGCCAAAC
1201  TATTGCTTCC CCTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT TAAACCAAAA
1261  ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA CAATCAGATA
1321  TGCAAAGGCA ATATCTATGC CATGGAGATC AACCTCCAGG CCAACCTGTG GAGAAGTTTT
1381  CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC GGCCAACATC
1441  ACCCTGCCCA CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT GCCTCCCTCG
1501  CTGGTGGACG CCTACATCAA CATCGGGGCG CGCTGGTCGC TGGACCCCAT GGACAACGTG
1561  AATCCCTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT CCTGGGCAAC
1621  GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTCGCCAT CAAGAGCCTC
1681  CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAACTTCC GCAAGGACGT CAACATGATC
1741  CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC CTTCACCAGC
1801  ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGCACAACA CGGCCTCCAC GCTCGAGGCC
1861  ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC GGCCAACATG
1921  CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC GCGCAACTGG
1981  GCCGCCTTCC GCGGCTGGTC CTTCACGCGC CTCAAGACCC GCGAGACGCC CTCGCTGGGC
2041  TCCGGGTTCG ACCCCTACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA CGGCACCTTC
2101  TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ACTCCTCCGT CAGCTGGCCC
2161  GGCAACGACC GCCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT CGACGGAGAG
2221  GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA GATGCTGGCC
2281  CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCGAGG GCTACAAGGA CCGCATGTAC
2341  TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTGG TGGACGAGGT CAACTACAAG
2401  GACTACCAGG CCGTCACCCT GGCCTACCAG CACAACAACT CGGGCTTCGT CGGCTACCTC
2461  GCGCCCACCA TGCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCGTACCC GCTCATCGGA
2521  AAGAGCGCCG TCACCAGCGT CACCCAGAAA AAGTTCCTCT GCGACAGGGT CATGTGGCGC
2581  ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG CCAGAACATG
2641  CTCTATGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTCG AAGTCGACCC CATGGATGAG
2701  TCCACCCTTC TCTATGTTGT CTTCGAAGTC TTCGACGTCG TCCGAGTGCA CCAGCCCCAC
2761  CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA CGCCACCACC
2821  TAA  (SEQ ID NO: 17)
```

FIG. 22

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCCTACTC CGGCACCGCC
 361  TACAACAGCC TGGCTCCCAA GGGAGCGCCC AATTCCAGCC AGTGGGAGCA AAAAAAGACT
 421  GGCAATAATG CCAATGGAGA TACGGAGAAT GTCACTTATG GTGTAGCTGC CATGGGAGGA
 481  ATTGACATCG ATAAAAATGG CCTTCAAATT GGAACCGATG ACACCAAAGA TGACGATAAT
 541  GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAAA TAGGAGAGGA AAACTGGCAA
 601  GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAAG ATACCAAAAT GAAGCCATGC
 661  TATGGCTCAT TTGCCAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAAT AAAAACAGAT
 721  GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTTG ATATTCCCAA TTCTGGCGCG
 781  GGAAATGGCA CAAATGTTAA CGATGATCCA GATATGGTTA TGTATACAGA AAATGTAAAT
 841  CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTCAGATGA CAGCTCAAAG
 901  GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT CAGAGACAAT
 961  TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC TGGTCAGGCC
1021  TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC CTACCAGCTC
1081  TTGCTTGACT CTCTGGGTGA CAGAACCAGG TATTTCAGTA TGTGGAATCA GGCGGTGGAC
1141  AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA ATTGCCAAAC
1201  TATTGCTTCC CCTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT TAAACCAAAA
1261  ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA CAATCAGATA
1321  TGCAAAGGCA ATATCTATGC CATGGAGATC AACCTCCAGG CCAACCTGTG GAGAAGTTTT
1381  CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC GGCCAACATC
1441  ACCCTGCCCA CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT GCCTCCCTCG
1501  CTGGTGGACG CCTACATCAA CATCGGGGCG CGCTGGTCGC TGGACCCCAT GGACAACGTG
1561  AATCCCTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT CCTGGGCAAC
1621  GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTTGCCAT CAAGAGCCTC
1681  CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAACTTCC GCAAGGACGT CAACATGATC
1741  CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC CTTCACCAGC
1801  ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGCACAACA CGGCCTCCAC GCTCGAGGCC
1861  ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC GGCCAACATG
1921  CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC GCGCAACTGG
1981  GCCGCCTTCC GCGGCTGGTC CTTCACGCGC CTCAAGACCC GCGAGACGCC CTCGCTGGGC
2041  TCCGGGTTCG ACCCCTACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA CGGCACCTTC
2101  TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ACTCCTCCGT CAGCTGGCCC
2161  GGCAACGACC GCCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT CGACGGAGAG
2221  GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA GATGCTGGCC
2281  CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCGAGG GCTACAAGGA CCGCATGTAC
2341  TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTCG TGGACGAGGT CAACTACAAG
2401  GACTACCAGG CCGTCACCCT GGCCTACCAG CACAACAACT CGGGCTTCGT CGGCTACCTC
2461  GCGCCCACCA TGCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCCTACCC GCTCATCGGC
2521  AAGAGCGCCG TCGCCAGCGT CACCCAGAAA AAGTTCCTCT GCGACCGGGT CATGTGGCGC
2581  ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG CCAGAACATG
2641  CTCTACGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTCG AAGTCGACCC CATGGATGAG
2701  TCCACCCTTC TCTATGTTGT CTTCGAAGTC TTCGACGTCG TCCGAGTGCA CCAGCCCCAC
2761  CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA AGCCACCACC
2821  TAA  (SEQ ID NO: 18)
```

FIG. 23

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGACCGGGGC CCTAGCTTCA AACCTTACTC CGGCACCGCT
 361  TACAACAGCC TGGCCCCCAA GGGAGCACCC AATTCCAGCC AGTGGGAGCA AAAAAAGACT
 421  GGCAAAAATG CCAATGGAGA TACGGAGAAT GTCACTTATG GTGTAGCTGC CATGGGAGGA
 481  ATTGACATCG ATAAAAATGG CCTTCAAATT GGAACCGATG ACACCAAAGA TGGCGATAAT
 541  GAAATTTATG CAGACAAAAC ATATCAGCCT GAGCCGCAAA TAGGAGAGGA AAACTGGCAA
 601  GAAACATATT CCTACTATGG AGGTAGAGCT CTTAAAAAAG ATACCAAAAT GAAGCCATGC
 661  TATGGCTCAT TTGCTAGACC TACCAATGTG AAAGGAGGAC AGGCAAAAAT AAAAACAGAT
 721  GGAGATGTTA AGTCATTTGA CATAGACCTA GCCTTCTTTG ATATTCCAAA TTCTGGCGCG
 781  GGAAATGGCA CAAATGTTAA CGATGATCCA GATATGGTTA TGTATACAGA AAATGTAAAT
 841  CTGGAAACCC CAGATACTCA TATTGTGTAC AAACCAGGAA CTTCAGATGA CAGCTCCGAG
 901  GTCAACTTGT GTCAGCAATC CATGCCTAAC AGACCCAATT ATATTGGCTT CAGAGACAAT
 961  TTTATTGGGC TTATGTACTA CAACAGCACT GGCAATATGG GTGTGCTGGC TGGTCAGGCC
1021  TCTCAACTGA ATGCCGTGGT GGACTTGCAA GACAGAAACA CAGAGCTGTC CTACCAGCTC
1081  TTGCTTGACT CTCTGGGTGA CAGAACCAGG TATTTCAGTA TGTGGAATCA GGCGGTGGAC
1141  AGTTATGATC CTGATGTGCG CATTATTGAA AACCATGGTG TGGAGGATGA ATTGCCAAAC
1201  TATTGCTTCC CCTTGGATGG AGCAGGCACC AATTCGGTTT ACCAAGGTGT TAAACCAAAA
1261  ACTGACAATG GCAACGATCA GTGGGAAACA GATTCCACAG TTTCAAGTCA CAATCAGATA
1321  TGCAAAGGCA ATATCTATGC CATGGAGATC AATCTCCAGG CCAACCTGTG GAGAAGTTTC
1381  CTCTACTCGA ACGTGGCCCT GTACCTGCCC GATTCTTACA AGTACACGCC GGCCAACATC
1441  ACCCTGCCCA CCAACACCAA CACCTACGAT TACATGAACG GGAGAGTGGT GCCTCCCTCG
1501  CTGGTGGATG CCTACATCAA CATCGGAGCG CGCTGGTCGC TGGACCCCAT GGACAACGTC
1561  AATCCCTTCA ACCACCACCG CAATGCGGGG CTGCGCTACC GCTCCATGCT CCTGGGCAAC
1621  GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAAT TTTTCGCCAT CAAGAGCCTT
1681  CTGCTCCTGC CCGGGTCCTA CACCTACGAG TGGAACTTCC GCAAGGACGT CAACATGATC
1741  CTGCAGAGCT CCCTCGGCAA CGACCTGCGC ACGGACGGGG CCTCCATCTC CTTCACCAGC
1801  ATCAACCTCT ACGCCACCTT CTTCCCCATG GCGCACAACA CGGCCTCCAC GCTCGAGGCC
1861  ATGCTGCGCA ACGACACCAA CGACCAGTCC TTCAACGACT ACCTCTCGGC GGCCAACATG
1921  CTCTACCCCA TCCCGGCCAA CGCCACCAAC GTGCCCATCT CCATCCCCTC GCGCAACTGG
1981  GCCGCCTTCC GCGGCTGGTC CTTCACGCGC CTCAAGACCA AGGAGACGCC CTCGCTGGGC
2041  TCCGGGTTCG ACCCATACTT CGTCTACTCG GGCTCCATCC CCTACCTCGA CGGCACCTTC
2101  TACCTCAACC ACACCTTCAA GAAGGTCTCC ATCACCTTCG ATTCCTCCGT CAGCTGGCCC
2161  GGCAACGACC GGCTCCTGAC GCCCAACGAG TTCGAAATCA AGCGCACCGT CGACGGCGAG
2221  GGATACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA GATGCTGGCC
2281  CACTACAACA TCGGCTACCA GGGCTTCTAC GTGCCCGAGG GCTACAAGGA CCGCATGTAC
2341  TCCTTCTTCC GCAACTTCCA GCCCATGAGC CGCCAGGTGG TGGACGAGGT CAACTACAAG
2401  GACTACCAGG CCGTCACCCT GGCCTACCAG CACAACAACT CGGGCTTCGT CGGCTACCTC
2461  GCGCCCACCA TGCGCCAGGG CCAGCCCTAC CCCGCCAACT ACCCGTACCC GCTCATCGGC
2521  AAGAGCGCCG TCACCAGCGT CACCCAGAAA AAGTTCCTCT GCGACAGGGT CATGTGGCGC
2581  ATCCCCTTCT CCAGCAACTT CATGTCCATG GGCGCGCTCA CCGACCTCGG GCAGAACATG
2641  CTCTATGCCA ACTCCGCCCA CGCGCTAGAC ATGAATTTCG AAGTCGACCC CATGGATGAG
2701  TCCACCCTTC TCTATGTTGT CTTCGAAGTC TTCGACGTCG TCCGAGTGCA CCAGCCCCAC
2761  CGCGGCGTCA TCGAGGCCGT CTACCTGCGC ACCCCCTTCT CGGCCGGTAA CGCCACCACC
2821  TAA (SEQ ID NO: 19)
```

FIG. 24

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCACCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCCTACTC CGGCACCGCT
 361  TACAACAGCC TGGCTCCCAA GGGAGCGCCC AACACTTGCC AGTGGACATA TACTGATAAC
 421  CAAACTGAGA AAACAGCCAC ATATGGAAAT GCACCCGTAG AGGGCATTAA CATTACAAAA
 481  GATGGCATTC AACTTGGAAC TGACAGCGAT GGTCAGGCAA TCTATGCAGA CGAAACTTAT
 541  CAGCCCGAAC CTCAGGTGGG AGATCCTGAA TGGCATGATA CCACAGGTAC AGAAGAAAAA
 601  TATGGAGGCA GAGCGCTTAA ACCTGCCACC GACATGAAAC CTTGCTATGG CTCTTTTGCC
 661  AAGCCAACTA ATGTTAAGGG AGGTCAGGCC AAAAGCAGAA CAAAAACTGA TGGAACAACT
 721  GAGCCTGATA TTGACATGGC CTTTTTTGAT GGCAGAAATG CAACAACAGC TGGTTTGACT
 781  CCAGAAATTG TTTTGTATAC TGAAAATGTG GATCTGGAAA CTCCAGATAC CCATATTGTA
 841  TACAAGGCAG GCACAGATGA CAGCAGCTCT TCTATCAATT TGGGTCAGCA GTCCATGCCC
 901  AACAGACCCA ACTACATTGG CTTCAGAGAC AACTTTATCG GGCTCATGTA CTACAACAGC
 961  ACTGGCAATA TGGGTGTACT GGCTGGACAG GCCTCCCAGC TGAATGCTGT GGTGGACTTG
1021  CAGGACAGAA ACACTGAACT GTCCTACCAG CTCTTGCTTG ACTCTCTGGG TGACAGAACC
1081  AGGTATTTCA GTATGTGGAA TCAGGCGGTG GACAGTTATG ACCCCGATGT GCGCATTATT
1141  GAAAATCACG GTGTGGAGGA TGAACTCCCC AACTATTGCT TCCCCCTGAA TGCTGTGGGT
1201  AGAACAAATA GTTATCAGGG AATTAAACCC AATGGAGGCG ATCCAGCTAC ATGGGCCAAA
1261  GATGAAAGCG TCAATGATTC TAATGAATTG GGCAAGGGCA ATCCTTTCGC CATGGAGATC
1321  AACATCCAGG CCAACCTGTG GCGGAACTTC CTCTACGCGA ACGTGGCGCT GTACCTGCCC
1381  GACTCCTACA AGTACACGCC GGCCAACATC ACGCTGCCCG CCAACACCAA CACCTACGAT
1441  TACATGAACG GCCGCGTGGT GGCGCCCTCG CTGGTGGACG CCTACATCAA CATCGGGGCG
1501  CGCTGGTCGC TGGACCCCAT GGACAACGTC AACCCCTTCA ACCACCACCG CAACGCGGGC
1561  CTGCGCTACC GCTCCATGCT CCTGGGCAAC GGGCGCTACG TGCCCTTCCA CATCCAGGTG
1621  CCCCAAAAGT TTTTCGCCAT CAAGAGCCTC CTGCTCCTGC CCGGGTCCTA CACCTACGAG
1681  TGGAACTTCC GCAAGGACGT CAACATGATC CTGCAGAGCT CCCTCGGCAA CGACCTGCGC
1741  ACGGACGGGG CCTCCATCGC CTTCACCAGC ATCAACCTCT ACGCCACCTT CTTCCCCATG
1801  GCGCACAACA CCGCCTCCAC GCTCGAGGCC ATGCTGCGCA ACGACACCAA CGACCAGTCC
1861  TTCAACGACT ACCTCTCGGC GGCCAACATG CTCTACCCCA TCCCGGCCAA CGCCACCAAC
1921  GTGCCCATCT CCATCCCCTC GCGCAACTGG GCCGCCTTCC GCGGATGGTC CTTCACGCGC
1981  CTCAAGACCC GCGAGACGCC CTCGCTAGGC TCCGGGTTCG ACCCCTACTT CGTCTACTCG
2041  GGCTCCATCC CCTACCTCGA CGGCACCTTC TACCTCAACC ACACCTTCAA GAAGGTCTCC
2101  ATCACCTTCG ACTCCTCCGT CAGCTGGCCC GGCAACGACC GCCTCCTGAC GCCCAACGAG
2161  TTCGAAATCA AGCGCACCGT CGACGGAGAG GGATACAACG TGGCCCAGTG CAACATGACC
2221  AAGGACTGGT TCCTGGTCCA GATGCTGGCC CACTACAACA TCGGCTACCA GGGCTTCTAC
2281  GTGCCCGAGG GCTACAAGGA CCGCATGTAC TCCTTCTTCC GCAACTTCCA GCCCATGAGC
2341  CGCCAGGTCG TGGACGAGGT CAACTACAAG GACTACCAGG CCGTCACCCT GGCCTACCAG
2401  CACAACAACT CGGGCTTCGT CGGCTACCTC GCGCCCACCA TGCGCCAGGG CCAGCCCTAC
2461  CCCGCCAACT ACCCCTACCC GCTCATCGGC AAGAGCGCCG TCGCCAGCGT CACCCAGAAA
2521  AAGTTCCTCT GCGACCGGGT CATGTGGCGC ATCCCCTTCT CCAGCAACTT CATGTCCATG
2581  GGCGCGCTCA CCGACCTCGG CCAGAACATG CTCTACGCCA ACTCCGCCCA CGCGCTAGAC
2641  ATGAATTTCG AAGTCGACCC CATGGATGAG TCCACCCTTC TCTATGTTGT CTTCGAAGTC
2701  TTCGACGTCG TCCGAGTGCA CCAGCCCCAC CGCGGCGTCA TCGAGGCCGT CTACCTGCGC
2761  ACGCCCTTCT CGGCCGGCAA CGCCACCACC TAA  (SEQ ID NO: 20)
```

FIG. 25

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGATCGGGGC CCTAGCTTCA AACCCTACTC CGGCACCGCC
 361  TACAACAGCC TGGCTCCCAA GGGAGCGCCC AACACTTGCC AGTGGACATA TACTGATAAC
 421  CAAACTGAGA AAACAGCCAC ATATGGAAAT GCGCCTGTGC AAGGCATTAG TATTACAAAA
 481  GATGGTATTC AACTTGGAAC TGACACTGAT GATCAGCCCA TTTATGCAGA TAAAACTTAT
 541  CAACCAGAGC CTCAAGTGGG TGATGCTGAA TGGCATGACA TCACTGGTAC TGATGAAAAA
 601  TATGGAGGCA GAGCTCTCAA GCCTGACACC AAAATGAAGC CCTGCTATGG TTCTTTTGCC
 661  AAGCCTACCA ATAAAGAAGG AGGTCAGGCA AATGTGAAAA CCGAAACAGG CGGTACCAAA
 721  GAATATGACA TTGACATGGC ATTCTTCGAT AATCGAAGTG CAGCTGCGGC TGGCCTGGCC
 781  CCAGAAATTG TTTTGTATAC TGAGAATGTG GATCTGGAAA CTCCAGATAC TCATATTGTA
 841  TACAAGGCAG GCACAGATGA CAGCAGCTCT TCTATCAATT TGGGTCAGCA GTCCATGCCC
 901  AACAGACCCA ACTACATTGG CTTCAGAGAC AACTTTATCG GTCTCATGTA CTACAACAGC
 961  ACTGGCAATA TGGGTGTACT GGCTGGTCAG GCCTCCCAGC TGAATGCTGT GGTGGACTTG
1021  CAGGACAGAA ACACTGAACT GTCCTACCAG CTCTTGCTTG ACTCTCTGGG TGACAGAACC
1081  AGGTATTTTA GTATGTGGAA TCAGGCGGTG GACAGTTATG ACCCCGATGT GCGCATTATT
1141  GAAAATCACG GTGTGGAGGA TGAACTCCCT AATTATTGCT TCCCCCTTAA TGCTGTGGGT
1201  AGAACTGATA CTTACCAGGG AATTAAGGCC AATGGTGCTG ATCAAACCAC ATGGACCAAA
1261  GATGATACTG TTAATGATGC TAATGAATTG GGCAAGGGCA ATCCTTTCGC CATGGAGATC
1321  AACATCCAGG CCAACCTGTG GCGGAACTTC CTCTACGCGA ACGTGGCCCT GTACCTGCCC
1381  GACTCCTACA AGTACACGCC GGCCAACATC ACGCTGCCCA CCAACACCAA CACCTACGAT
1441  TACATGAACG GCCGCGTGGT GGCGCCCTCG CTGGTGGACG CCTACATCAA CATCGGGGCG
1501  CGCTGGTCGC TGGACCCCAT GGACAACGTC AACCCCTTCA ACCACCACCG CAACGCGGGC
1561  CTGCGCTACC GCTCCATGCT CCTGGGCAAC GGGCGCTACG TGCCCTTCCA CATCCAGGTG
1621  CCCCAAAAGT TCTTCGCCAT CAAGAGCCTC CTGCTCCTGC CCGGGTCCTA CACCTACGAG
1681  TGGAACTTCC GCAAGGACGT CAACATGATC CTGCAGAGCT CCCTCGGCAA CGACCTGCGC
1741  ACGGACGGGG CCTCCATCGC CTTCACCAGC ATCAACCTCT ACGCCACCTT CTTCCCCATG
1801  GCGCACAACA CCGCCTCCAC GCTCGAGGCC ATGCTGCGCA ACGACACCAA CGACCAGTCC
1861  TTCAACGACT ACCTCTCGGC GGCCAACATG CTCTACCCCA TCCCGGCCAA TGCCACCAAC
1921  GTGCCCATCT CCATCCCCTC GCGCAACTGG GCCGCCTTCC GCGGATGGTC CTTCACGCGC
1981  CTCAAGACCC GCGAGACGCC CTCGCTAGGC TCCGGGTTCG ACCCCTACTT CGTCTACTCG
2041  GGCTCCATCC CCTACCTCGA CGGCACCTTC TACCTCAACC ACACCTTCAA GAAGGTCTCC
2101  ATCACCTTCG ACTCCTCCGT CAGCTGGCCC GGCAACGACC GCCTCCTGAC GCCCAACGAG
2161  TTCGAAATCA AGCGCACCGT CGACGGAGAG GGGTACAACG TGGCCCAGTG CAACATGACC
2221  AAGGACTGGT TCCTGGTCCA GATGCTGGCC CACTACAACA TCGGCTACCA GGGCTTCTAC
2281  GTGCCCGAGG GCTACAAGGA CCGCATGTAC TCCTTCTTCC GCAACTTCCA GCCCATGAGC
2341  CGCCAGGTCG TGGACGAGGT CAACTACAAG GACTACCAGG CCGTCACCCT GGCCTACCAG
2401  CACAACAACT CGGGCTTCGT CGGCTACCTC GCGCCCACCA TGCGCCAGGG CCAGCCCTAC
2461  CCCGCCAACT ACCCCTACCC GCTCATCGGC AAGAGCGCCG TCGCCAGCGT CACCCAGAAA
2521  AAGTTCCTCT GCGACCGGGT CATGTGGCGC ATCCCCTTCT CCAGCAACTT CATGTCCATG
2581  GGCGCGCTCA CCGACCTCGG CCAGAACATG CTCTACGCCA ACTCCGCCCA CGCGCTAGAC
2641  ATGAATTTCG AAGTCGACCC CATGGATGAG TCCACCCTTC TCTATGTTGT CTTCGAAGTC
2701  TTCGACGTCG TCCGAGTGCA CCAGCCCCAC CGCGGCGTCA TCGAGGCCGT CTACCTGCGC
2761  ACGCCCTTCT CGGCCGGCAA CGCCACCACC TAA  (SEQ ID NO: 21)
```

FIG. 26

```
   1  ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG CCAGGACGCC
  61  TCGGAGTACC TGAGTCCCGG GCTGGTGCAG TTCGCTCGCG CCACCGAGAG CTACTTCAGT
 121  CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGG
 181  TCCCAGCGCC TGACGCTGCG GTTCATCCCC GTGGACCGCG AGGACACCGC GTACTCGTAC
 241  AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CTCCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGACCGCGGC CCCACCTTCA AGCCCTACTC CGGCACCGCY
 361  TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACTCCTGCG AGTGGGAGCA AGAGGAAACT
 421  CAGGCAGTTG AAGAAGCAGC AGAAGAGGAG GAAGAAGATG CTGACGGTCA AGCTGAGGAA
 481  GAGCAAGCAG CTACCAAAAA GACTCATGTA TATGCTCAGG CTCCCCTTTC CGGCGAAAAA
 541  ATTAGCAAAG ACGGTCTGCA GATAGGAACG GACGCTACAG CAACCGAACA AAAACCTATT
 601  TATGCAGACC CTACATTCCA GCCCGAACCC CAAATCGGGG AGTCCCAGTG GAATGAGGCA
 661  GATGCTACAG TCGCTGGTGG TAGAGTGCTC AAGAAAACCA CTCCCATGAA ACCATGCTAT
 721  GGTTCCTATG CAAGACCCAC GAATGCTAAT GGAGGTCAGG GTGTACTAGC GGCAAATGCC
 781  CAAGGACAGC TAGAATCTCA GGTTGAAATG CAATTCTTTT CAACTTCTGA AAACGCCCGT
 841  AACGAGGCTA ACAACATTCA GCCCAAATTG GTGCTGTATA GCGAGGATGT GCACATGGAG
 901  ACCCCGGATA CACACCTCTC TTACAAGCCC ACAAAAAGCG ATGACAATTC TAAAGTTATG
 961  CTGGGCCAAC AGGCCATGCC CAACAGGCCT AATTACATTG GCTTCAGAGA CAACTTTATC
1021  GGTCTCATGT ACTACAACAG CACTGGCAAC ATGGGAGTGC TTGCAGGTCA GGCCTCTCAG
1081  TTGAATGCAG TGGTGGACTT GCAAGACAGA AACACAGAAC TGTCCTACCA GCTCTTGCTT
1141  GATTCCATGG GTGACAGAAC CAGATATTTC TCCATGTGGA ATCAGGCAGT GGACAGTTAT
1201  GACCCAGATG TCAGAATTAT TGAAAATCAT GGAACTGAAG ACGAGCTCCC CAACTATTGT
1261  TTCCCTCTGG GCGGCATAGG GGTAACTGAC ACTTACCAGG CTGTTAAGAC CAACAATGGC
1321  AATAATGGGG GTCAGGTGAC TTGGACAAAA GATGAAACTT TTGCAGAGCG CAATGAGATA
1381  GGGGTGGGAA ACAATTTCGC CATGGAGATC AACCTCAATG CCAACCTGTG GAGGAACTTC
1441  CTGTACTCCA ACGTGGCCCT GTACCTGCCA GACAAGCTTA AGTACAACCC CTCCAACGTG
1501  GACATCTCTG ACAACCCCAA CACCTACGAT TACATGAACA AGCGAGTGGT GGCCCCGGGG
1561  CTGGTGGACT GCTACATCAA CCTGGGCGCG CGCTGGTCGC..TGGACTACAT GGACAACGTC
1621  AACCCTTTCA ACCACCACCG CAACGCGGGC CTGCGCTACC GCTCCATGCT CCTGGGCAAC
1681  GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAGT TCTTTGCCAT CAAGAACCTC
1741  CTCCTCCTGC CGGGCTCCTA CACCTACGAG TGGAACTTCA GGAAGGATGT CAACATGGTC
1801  CTCCAGAGCT CTCTGGGCAA CGATCTCAGG GTGGACGGGG CCAGCATCAA GTTCGAGAGC
1861  ATCTGCCTCT ACGCCACCTT CTTCCCCATG GCCCACAACA CCGCCTCCAC GCTCGAGGCC
1921  ATGCTCAGGA ACGACACCAA CGACCAGTCC TTCAATGACT ACCTCTCCGC CGCCAACATG
1981  CTCTACCCCA TCCCCGCCAA CGCCACCAAC GTCCCCATCT CCATCCCCTC GCGCAACTGG
2041  GCGGCCTTCC GCGGCTGGGC CTTCACCCGC CTCAAGACCA AGGAGACCCC CTCCCTGGGC
2101  TCGGGATTCG ACCCCTACTA CACCTACTCG GGATCCATTC CCTACCTGGA CGGCACCTTC
2161  TACCTCAACC ACACTTTCAA GAAGGTCTCG GTCACCTTCG ACTCCTCGGT CAGCTGGCCG
2221  GGCAACGACC GCCTGCTCAC CCCCAACGAG TTCGAGATCA AGCGCTCGGT CGACGGGGAG
2281  GGCTACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA GATGCTGGCC
2341  AACTACAACA TCGGCTACCA GGGCTTCTAC ATCCCAGAGA GCTACAAGGA CAGGATGTAC
2401  TCCTTCTTCA GGAACTTCCA GCCCATGAGC CGGCAGGTGG TGGACCAGAC CAAGTACAAG
2461  GACTACCAGG AGGTGGGCAT CATCCACCAG CACAACAACT CGGGCTTCGT GGGCTACCTC
2521  GCCCCCACCA TGCGCGAGGG ACAGGCCTAC CCCGCCAACT TCCCCTACCC GCTCATAGGC
2581  AAGACCGCGG TCGACAGCAT CACCCAGAAA AAGTTCCTCT GCGACCGCAC CCTCTGGCGC
2641  ATCCCCTTCT CCAGCAACTT CATGTCCATG GGTGCGCTCA·CGGACCTGGG CCAGAACCTG
2701  CTCTATGCCA ACTCCGCCCA CGCGCTCGAC ATGACCTTCG AGGTCGACCC CATGGACGAG
2761  CCCACCCTTC TCTATGTTCT GTTCGAAGTC TTTGACGTGG TCCGGGTCCA CCAGCCGCAC
2821  CGCGGCGTCA TCGAGACCGT GTACCTGCGC ACGCCCTTCT CGGCCGGCAA CGCCACCACC
2881  TAA  (SEQ ID NO: 22)
```

FIG. 27

```
   1  ATGGCCACCC CATCGATGCT GCCCCAGTGG GCGTACATGC ACATCGCCGG ACAGGACGCT
  61  TCGGAGTACC TGAGTCCGGG TCTGGTGCAG TTCGCCCGCG CCACAGACAC CTACTTCAGT
 121  CTGGGGAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGC
 181  AGCCAGCGGC TGACGCTGCG CTTCGTGCCC GTGGACCGCG AGGACAACAC CTACTCGTAC
 241  AAAGTGCGCT ACACGCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CAGCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGACCGGGGC CCTAGCTTCA AACCCTACTC CGGCACCGCC
 361  TACAACAGCC TGGCCCCCAA GGGAGCTCCC AATTCCAGTC AGTGGGAGCA GACGGAGAAC
 421  GGGGGCGGAC AGGCTACGAC TAAAACACAC ACCTATGGAG TTGCCCCAAT GGGTGGAACT
 481  AATATTACAG TCGACGGACT ACAAATTGGA ACTGACGCTA CAGCTGATAC GGAAAAACCA
 541  ATTTATGCTG ATAAAACATT CCAACCTGAG CCTCAGATAG GAGAGGAAAA CTGGCAAGAA
 601  ACTGAAAGCT TTTATGGCGG TAGGGCTCTT AAGAAAGACA CAAACATGAA GCCTTGTTAT
 661  GGCTCATTTG CCAGACCTAC CAATGAAAAG GGAGGTCAAG CTAAACTTAA AGTTGGAGCT
 721  GATGGGCTGC CGACCAAAGA ATTTGACATA GACCTAGCAT TCTTTGATAC TCCTGGTGGC
 781  ACTGTGACCG GAGGTACAGA GGAGTATAAA GCAGATATTG TTATGTATAC CGAAAACACG
 841  TATCTGGAAA CTCCAGACAC ACATGTGGTG TATAAACCAG GCAAGGATAA CACAAGTTCT
 901  AAAATTAACC TGGTCCAGCA GTCTATGCCC AACAGGCCCA ACTACATTGG GTTTAGGGAC
 961  AACTTTATTG GGCTCATGTA TTACAACAGC ACTGGCAATA TGGGTGTGCT GGCCGGTCAG
1021  GCTTCTCAGT TGAATGCTGT GGTTGACTTG CAAGACAGAA ACACTGAACT GTCTTACCAG
1081  CTCTTGCTTG ACTCTTTGGG TGACAGAACC AGGTATTTCA GTATGTGGAA TCAGGCGGTG
1141  GACAGTTATG ATCCTGATGT GCGCATTATT GAAAACCATG GTGTGGAAGA TGAACTTCCC
1201  AACTATTGCT TCCCCCTGGA TGGGTCTGGC ACTAACGCCG CTTACCAAGG TGTGAAAGTA
1261  AAAAATGGTC AAGATGGTGA TGTTGAGAGC GAATGGGAAA AAGATGATAC TGTCGCAGCT
1321  CGAAATCAAT TATGCAAGGG CAACATTTTT GCCATGGAGA TCAATCTCCA GGCCAACCTG
1381  TGGAGAAGTT TTCTCTACTC GAACGTGGCC CTGTACCTGC CCGATTCTTA CAAGTACACG
1441  CCGGCCAACA TCACCCTGCC CACCAACACC AACACCTACG ATTACATGAA CGGGAGAGTG
1501  GTGCCTCCCT CGCTGGTGGA CGCCTACATC AACATCGGGG CGCGCTGGTC GCTGGACCCC
1561  ATGGACAACG TCAATCCCTT CAACCACCAT CGCAACGCGG GGCTGCGCTA CCGCTCCATG
1621  CTCCTGGGCA ACGGGCGCTA CGTGCCCTTC CACATCCAGG TGCCCCAGAA ATTTTTCGCC
1681  ATTAAGAGCC TCCTGCTCCT GCCCGGGTCC TACACCTACG AGTGGAACTT CCGCAAGGAC
1741  GTCAACATGA TCCTGCAGAG CTCCCTCGGC AACGACCTGC GCACGGACGG GGCCTCCATC
1801  TCCTTCACCA GCATCAACCT CTACGCCACC TTCTTCCCCA TGGCGCACAA CACCGCCTCC
1861  ACGCTCGAGG CCATGCTGCG CAACGACACC AACGACCAGT CCTTCAACGA CTACCTCTCG
1921  GCGGCCAACA TGCTCTACCC CATCCCGGCC AACGCCACCA ACGTGCCCAT CTCCATCCCC
1981  TCGCGCAACT GGGCCGCCTT CCGCGGCTGG TCCTTCACGC GCCTCAAGAC CAAGGAGACG
2041  CCCTCGCTGG GCTCCGGGTT CGACCCCTAC TTCGTCTACT CGGGCTCCAT CCCCTACCTC
2101  GACGGCACCT TCTACCTCAA CCACACCTTC AAGAAGGTCT CCATCACCTT CGACTCCTCC
2161  GTCAGCTGGC CCGGCAACGA CCGGCTCCTG ACGCCCAACG AGTTCGAAAT CAAGCGCACC
2221  GTCGACGGCG AGGGCTACAA CGTGGCCCAG TGCAACATGA CCAAGGACTG GTTCCTGGTC
2281  CAGATGCTGG CCCACTACAA CATCGGCTAC CAGGGCTTCT ACGTGCCCGA GGGCTACAAG
2341  GACCGCATGT ACTCCTTCTT CCGCAACTTC CAGCCCATGA GCCGCCAGGT CGTGGACGAG
2401  GTCAACTACA AGGACTACCA GGCCGTCACC CTGGCCTACC AGCACAACAA CTCGGGCTTC
2461  GTCGGCTACC TCGCGCCCAC CATGCGCCAG GGCCAGCCCT ACCCCGCCAA CTACCCCTAC
2521  CCGCTCATCG GCAAGAGCGC CGTCGCCAGC GTCACCCAGA AAAAGTTCCT CTGCGACCGG
2581  GTCATGTGGC GCATCCCCTT CTCCAGCAAC TTCATGTCCA TGGGCGCGCT CACCGACCTC
2641  GGCCAGAACA TGCTCTACGC CAACTCCGCC CACGCGCTAG ACATGAATTT CGAAGTCGAC
2701  CCCATGGATG AGTCCACCCT TCTCTATGTT GTCTTCGAAG TCTTCGACGT CGTCCGAGTG
2761  CACCAGCCCC ACCGCGGCGT CATCGAGGCC GTCTACCTGC GCACCCCCTT CTCGGCCGGT
2821  AACGCCACCA CCTAA (SEQ ID NO: 23)
```

FIG. 28

```
   1  ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG CCAGGACGCC
  61  TCNGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCGCG CCACCGAGAG CTACTTCAGC
 121  CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGG
 181  TCTCAGCGCC TGACGCTGCG GTTCATTCCC GTGGACCGCG AGGACACCGC GTACTCGTAC
 241  AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CTCCACCTAC
 301  TTTGACATCC GCGGGGTGCT GGACCGGGGT CCCACTTTCA AGCCCTACTC TGGCACCGCC
 361  TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACTCCTGCG AGTGGGAGCA AGAGGAAACT
 421  CAGGCAGTTG AAGAAGCAGC AGAAGAGGAA GAAGAAGATG CTGACGGTCA AGCTGAGGAA
 481  GAGCAAGCAG CTACCAAAAA GACTCATGTA TATGCTCAGG CTCCCCTTTC TGGCGAAAAA
 541  ATTAGTAAAG ATGGTCTGCA AATAGGAACG GACGCTACAG CTACAGAACA AAAACCTATT
 601  TATGCAGACC CTACATTCCA GCCCGAACCC CAAATCGGGG AGTCACAGTG GAATGAGGCA
 661  GATGCTACAG TCGCCGGCGG TAGAGTGCTA AAGAAATCTA CTCCCATGAA ACCATGCTAT
 721  GGTTCCTATG CAAGACCCAC AAATGCTAAT GGAGGTCAGG GTGTACTAAC GGCAAATGCC
 781  CAGGGACAGC TAGAATCTCA GGTTGAAATG CAATTCTTTT CAACTTCTGA AAACGCCCGT
 841  AACGAGACTA ACAACATTCA GCCCAAATTG GTGCTGTATA GTGAGGATGT GCACATGGAG
 901  ACCCCGGATA CGCACCTTTC TTACAAGCCC GCAAAAAGCG ATGACAATTC AAAAATCATG
 961  CTGGGTCAGC AGTCCATGCC CAACAGACCT AATTACATCG GCTTCAGAGA TAACTTTATC
1021  GGCCTCATGT ATTACAATAG CACTGGCAAC ATGGGAGTGC TTGCAGGTCA GGCCTCTCAG
1081  TTGAATGCAG TGGTGGACTT GCAAGACAGA AACACAGAAC TGTCCTACCA GCTCTTGCTT
1141  GATTCCATGG GTGACAGAAC CAGATACTTT TCCATGTGGA ATCAGGCAGT GGACAGTTAT
1201  GACCCAGATG TTAGAATTAT TGAAAATCAT GGAACTGAAG ACGAGCTCCC CAACTATTGT
1261  TTCCCTCTGG GTGGCATAGG GGTAACTGAC ACTTACCAGG CTGTTAAAAC CAACAATGGC
1321  AATAACGGGG GCCAGGTGAC TTGGACAAAA GATGAAACTT TTGCAGATCG CAATGAAATA
1381  GGGGTGGGAA ACAATTTCGC TATGGAGATA AACCTCAGTG CCAACCTGTG GAGAAACTTC
1441  CTGTACTCCA ACGTGGCGCT GTACCTACCA GACAAGCTTA AGTACAACCC CTCCAATGTG
1501  GACATCTCTG ACAACCCCAA CACCTACGAT TACATGAACA AGCGAGTGGT GGCCCCGGGG
1561  CTGGTGGACT GCTACATCAA CCTGGGCGCG CGCTGGTCGC TGGACTACAT GGACAACGTC
1621  AACCCCTTCA ACCACCACCG CAATGCGGGC CTGCGCTACC GCTCCATGCT CCTGGGCAAC
1681  GGGCGCTACG TGCCCTTCCA CATCCAGGTG CCCCAGAAGT TCTTTGCCAT CAAGAACCTC
1741  CTCCTCCTGC CGGGCTCCTA CACCTACGAG TGGAACTTCA GGAAGGATGT CAACATGGTC
1801  CTCCAGAGCT CTCTGGGTAA CGATCTCAGG GTGGACGGGG CCAGCATCAA GTTCGAGAGC
1861  ATCTGCCTCT ACGCCACCTT CTTCCCCATG GCCCACAACA CGGCCTCCAC GCTCGAGGCC
1921  ATGCTCAGGA ACGACACCAA CGACCAGTCC TTCAATGACT ACCTCTCCGC CGCCAACATG
1981  CTCTACCCCA TACCCGCCAA CGCCACCAAC GTCCCCATCT CCATCCCCTC GCGCAACTGG
2041  GCGGCCTTCC GCGGCTGGGC CTTCACCCGC CTCAAGACCA AGGAGACCCC CTCCCTGGGC
2101  TCGGGATTCG ACCCCTACTA CACCTACTCG GGCTCCATTC CCTACCTGGA CGGCACCTTC
2161  TACCTCAACC ACACTTTCAA GAAGGTCTCG GTCACCTTCG ACTCCTCGGT CAGCTGGCCG
2221  GGCAACGACC GTCTGCTCAC CCCCAACGAG TTCGAGATCA AGCGCTCGGT CGACGGGGAG
2281  GGCTACAACG TGGCCCAGTG CAACATGACC AAGGACTGGT TCCTGGTCCA GATGCTGGCC
2341  AACTACAACA TCGGCTACCA GGGCTTCTAC ATCCCAGAGA GCTACAAGGA CAGGATGTAC
2401  TCCTTCTTCA GGAACTTCCA GCCCATGAGC CGGCAGGTGG TGGACCAGAC CAAGTACAAG
2461  GACTACCAGG AGGTGGGCAT CATCCACCAG CACAACAACT CGGGCTTCGT GGGCTACCTC
2521  GCCCCCACCA TGCGCGAGGG ACAGGCCTAC CCCGCCAACT TCCCCTATCC GCTCATAGGC
2581  AAGACCGCGG TCGACAGCAT CACCCAGAAA AAGTTCCTCT GCGACCGCAC CCTCTGGCGC
2641  ATCCCCTTCT CCAGCAACTT CATGTCCATG GGTGCGCTCT CGGACCTGGG CCAGAACTTG
2701  CTCTACGCCA ACTCCGCCCA CGCCCTCGAC ATGACCTTCG AGGTCGACCC CATGGACGAG
2761  CCCACCCTTC TCTATGTTCT GTTCGAAGTC TTTGACGTGG TCCGGGTCCA CCAGCCGCAC
2821  CGCGGCGTCA TCGAGACCGT GTACCTGCGT ACGCCCTTCT CGGCCGGCAA CGCCACCACC
2881  TAA (SEQ ID NO: 24)
```

FIG. 29

```
   1  ATGGCGACCC CATCGATGAT GCCGCAGTGG TCGTACATGC ACATCTCGGG CCAGGACGCC
  61  TCGGAGTACC TGAGCCCCGG GCTGGTGCAG TTCGCCCGCG CCACCGAGAG CTACTTCAGT
 121  CTGAGTAACA AGTTTAGGAA CCCCACGGTG GCGCCCACGC ACGATGTGAC CACCGACCGG
 181  TCCCAGCGCC TGACGCTGCG GTTCATCCCC GTGGACCGCG AGGACACCGC GTACTCGTAC
 241  AAGGCGCGGT TCACCCTGGC CGTGGGCGAC AACCGCGTGC TGGACATGGC CTCCACCTAC
 301  TTTGACATCC GCGGCGTGCT GGACCGCGGC CCCACCTTCA AGCCCTACTC CGGCACCGCC
 361  TACAACTCCC TGGCCCCCAA GGGCGCTCCC AACTCTTGTG AGTGGGAGCA ATTAGAAGAA
 421  GCCCAGGCCG CTTTGGAAGA CGAAGAATTA GAAGATGAAG ACGAGGAACC ACAGGATGAG
 481  GCGCCTGTGA AAAAGACCCA TGTATACGCT CAGGCTCCCC TTTCTGGAGA AGAAATTACT
 541  AAAGACGGTT TGCAAATAGG GTCAGATAAC ACAGAAGCTC AGTCTAAGCC TATATATGCA
 601  GACCCTACAT TCCAGCCCGA ACCCCAAATC GGGGAGTCCC AGTGGAACGA GGCAGATGCT
 661  ACAGTCGCTG GTGGTAGAGT GCTCAAGAAA ACCACTCCCA TGAAACCATG CTATGGTTCC
 721  TATGCAAGAC CCACGAATGC TAATGGAGGT CAGGGTGTGC TGGTGGCTGA TGATAAGGGG
 781  GTCCTTCAAT CTAAAGTTGA ATTGCAATTT TTTTCAAATA CTACTACTCT TAATCAGCGG
 841  GAGGGTAATG ATACAAAACC AAAAGTAGTG CTGTATAGCG AGGATGTGCA CATGGAAACA
 901  CCAGACACCC ACATTTCTTA CAAGCCCACA AAAAGCGATG ACAATTCTAA AGTTATGCTG
 961  GGCCAACAGT CCATGCCCAA CAGGCCTAAT TACATCGGCT TCAGAGACAA CTTTATCGGT
1021  CTCATGTACT ACAACAGCAC TGGCAACATG GGAGTGCTTG CAGGTCAGGC CTCTCAGTTG
1081  AATGCAGTGG TGGACTTGCA AGACAGAAAC ACAGAACTGT CCTACCAGCT CTTGCTTGAT
1141  TCCATGGGTG ACAGAACCAG ATATTTCTCC ATGTGGAATC AGGCAGTGGA CAGTTATGAC
1201  CCGGATGTCA GAATTATTGA AAATCATGGA ACCGAAGACG AGCTCCCCAA CTATTGTTTT
1261  CCTCTGGGTG GCATAGGGGT AACTGACACT TACCAGGTCA TTAAAACTAA TGGCAATGGT
1321  CAAGCAGACC CAACCTGGGA AAAAGATACA GAGTTTGCAG ACCGCAATGA AATAGGGGTG
1381  GGAAACAATT TCGCCATGGA GATCAACCTC AATGCCAACC TGTGGAGGAA CTTCCTGTAC
1441  TCCAACGTGG CCCTGTACCT GCCAGACAAG CTTAAGTACA ACCCCTCCAA CGTGGACATC
1501  TCTGACAACC CCAACACCTA CGATTACATG AACAAGCGAG TGGTGGCCCC GGGGCTGGTG
1561  GACTGCTACA TCAACCTGGG CGCGCGCTGG TCGCTGGACT ACATGGACAA CGTCAACCCC
1621  TTCAACCACC ACCGCAACGC GGGCCTGCGC TACCGCTCCA TGCTCCTGGG CAACGGGCGC
1681  TACGTGCCCT TCCACATCCA GGTGCCCCAG AAGTTCTTTG CCATCAAGAA CCTCCTCCTC
1741  CTGCCGGGCT CCTACACCTA CGAGTGGAAC TTCAGGAAGG ATGTCAACAT GGTCCTCCAG
1801  AGCTCTTTGG GCAACGATCT CAGGGTGGAC GGGGCCAGCA TCAAGTTCGA GAGCATCTGC
1861  CTCTACGCCA CCTTCTTCCC CATGGCCCAC AACACCGCCT CCACGCTCGA GGCCATGCTC
1921  AGGAACGACA CCAACGACCA GTCCTTCAAT GACTACCTCT CCGCCGCCAA CATGCTCTAC
1981  CCCATCCCCG CCAACGCCAC CAACGTCCCT ATCTCCATCC CCTCGCGCAA CTGGGCGGCC
2041  TTCCGCGGCT GGGCCTTCAC CCGCCTCAAG ACCAAGGAGA CACCCTCCCT GGGCTCGGGA
2101  TTCGACCCCT ACTACACCTA CTCGGGATCC ATTCCCTACC TGGACGGCAC CTTCTACCTC
2161  AACCACACTT TCAAGAAGGT CTCGGTCACC TTCGACTCCT CGGTCAGCTG GCCGGGCAAC
2221  GACCGCCTGC TCACCCCCAA CGAGTTCGAG ATCAAGCGCT CGGTCGACGG GGAGGGCTAC
2281  AACGTGGCCC AGTGCAACAT GACCAAGGAC TGGTTCCTGG TCCAGATGCT GGCCAACTAC
2341  AACATCGGCT ACCAGGGCTT CTACATCCCA GAGAGCTACA AGGACAGGAT GTACTCCTTC
2401  TTCAGGAACT TCCAGCCCAT GAGCCGGCAG GTGGTGGACC AAACCAAGTA CAAGGACTAC
2461  CAGGAGGTGG GCATCATCCA CCAGCACAAC AACTCGGGCT TCGTGGGCTA CCTCGCCCCC
2521  ACCATGCGCG AGGGACAGGC CTACCCCGCC AACTTCCCCT ACCCGCTCAT AGGCAAGACC
2581  GCGGTCGACA GCATCACCCA GAAAAAGTTC CTCTGCGACC GCACCCTCTG GCGCATCCCC
2641  TTCTCCAGCA ACTTCATGTC CATGGGTGCG CTCACGGACC TGGGCCAGAA CCTGCTCTAT
2701  GCCAACTCCG CCCACGCGCT CGACATGACC TTCGAGGTCG ACCCCATGGA CGAGCCCACC
2761  CTTCTCTATG TTCTGTTCGA AGTCTTTGAC GTGGTCCGGG TCCACCAGCC GCACCGCGGC
2821  GTCATCGAGA CCGTGTACCT GCGCACGCCC TTCTCGGCCG GCAACGCCAC CACCTAA
(SEQ ID NO: 25)
```

FIG. 30

```
                                                                                                  100
1
     C1 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFN LGNKFRNPTV APTHDVTTDR SQRLMLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTF
   CV68 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
  CHAD3 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
  CHAD4 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
  CHAD5 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
  CHAD6 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
  CHAD7 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
  CHAD8 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFN LGNKFRNPTV APTHDVTTDR SQRLMLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTF
  CHAD9 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD10 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD11 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD16 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD17 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD19 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD20 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD22 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFN LGNKFRNPTV APTHDVTTDR SQRLMLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTF
 CHAD24 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD26 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD30 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFN LGNKFRNPTV APTHDVTTDR SQRLMLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTF
 CHAD31 MATPSMMPQW SYMHISGQDA SEYLSPGLVQ FARATESYFS LSNKFRNPTV APTHDVTTDR SQRLTLRFIP VDREDTAYSY KARFTLAVGD NRVLDMASTY
 CHAD37 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFN LGNKFRNPTV APTHDVTTDR SQRLMLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTF
 CHAD38 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD44 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD82 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
 CHAD63 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
   PAN5 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
   PAN6 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
   PAN7 MATPSMLPQW AYMHIAGQDA SEYLSPGLVQ FARATDTYFS LGNKFRNPTV APTHDVTTDR SQRLTLRFVP VDREDNTYSY KVRYTTLAVGD NRVLDMASTY
```

FIG. 31A

```
                                                
```

```
                                                
```

```
                                                
```

```
401                                                
```

```
                                                
```

```
                                                
```

```
                                                
```

```
                                                
```

```
                                                
```

AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTG (SEQ ID NO: 26)

ATGGAATTCGTTTAAACCATCATCAATAATATACCTC (SEQ ID NO: 27)

CGCTGGCACTCAAGAGTGGCCTC (SEQ ID NO: 28)

ATGAAGCTTGTTTAAACCCATCATCAATAATATACCT (SEQ ID NO: 29)

ATCTAGACAGCGTCCATAGCTTACCG (SEQ ID NO: 30)

ATGCTACGTAGCGATCGCGTGAGTAGTGTTTGGGGGTGGGTGGG (SEQ ID NO: 31)

TAGGCGCGCCGCTTCTCCTCGTTCAGGCTGGCG (SEQ ID NO: 32)

GATCTAGTTAGTTTAAACGAATTCGGATCTGCGACGCG (SEQ ID NO: 33)

TTCGATCATGTTTAAACGAAATTAAGAATTCGGATCC (SEQ ID NO: 34)

TATTCTGCGATCGCTGAGGTGGGTGAGTGGGCG (SEQ ID NO: 35)

TAGGCGCGCCCTTAAACGGCATTTGTGGGAG (SEQ ID NO: 36)

CGTCTAGAAGACCCGAGTCTTACCAGT (SEQ ID NO: 37)

CGGGATCCGTTTAAACCATCATCAATAATATACCTTATT (SEQ ID NO: 38)

ATGGAATTCGTTTAAACCATCATCAATAATATACCTT (SEQ ID NO: 39)

ATGACGCGATCGCTGATATCCTATAATAATAAAACGCAGACTTTG (SEQ ID NO: 40)

TGTCCTACCARCTCTTGCTTGA (SEQ ID NO: 45)

GTGGAARGGCACGTAGCG (SEQ ID NO: 46)

FIG. 32

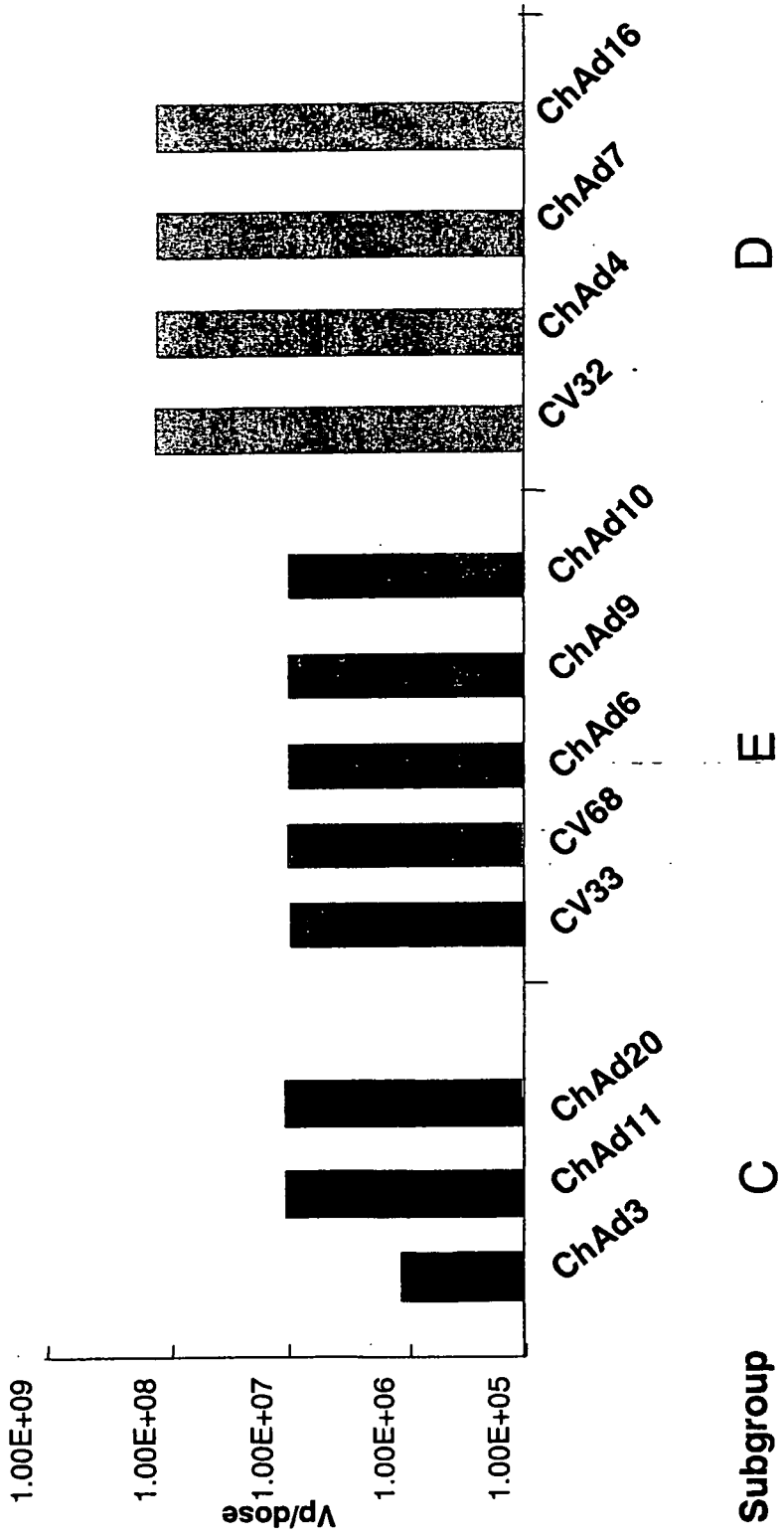
ChAd16
ChAd7
ChAd4
CV32
ChAd10
ChAd9
ChAd6
CV68
CV33
ChAd20
ChAd11
ChAd3
D
E
C
Subgroup
1.00E+09
1.00E+08
1.00E+07
1.00E+06
1.00E+05
Vp/dose

FIG. 33

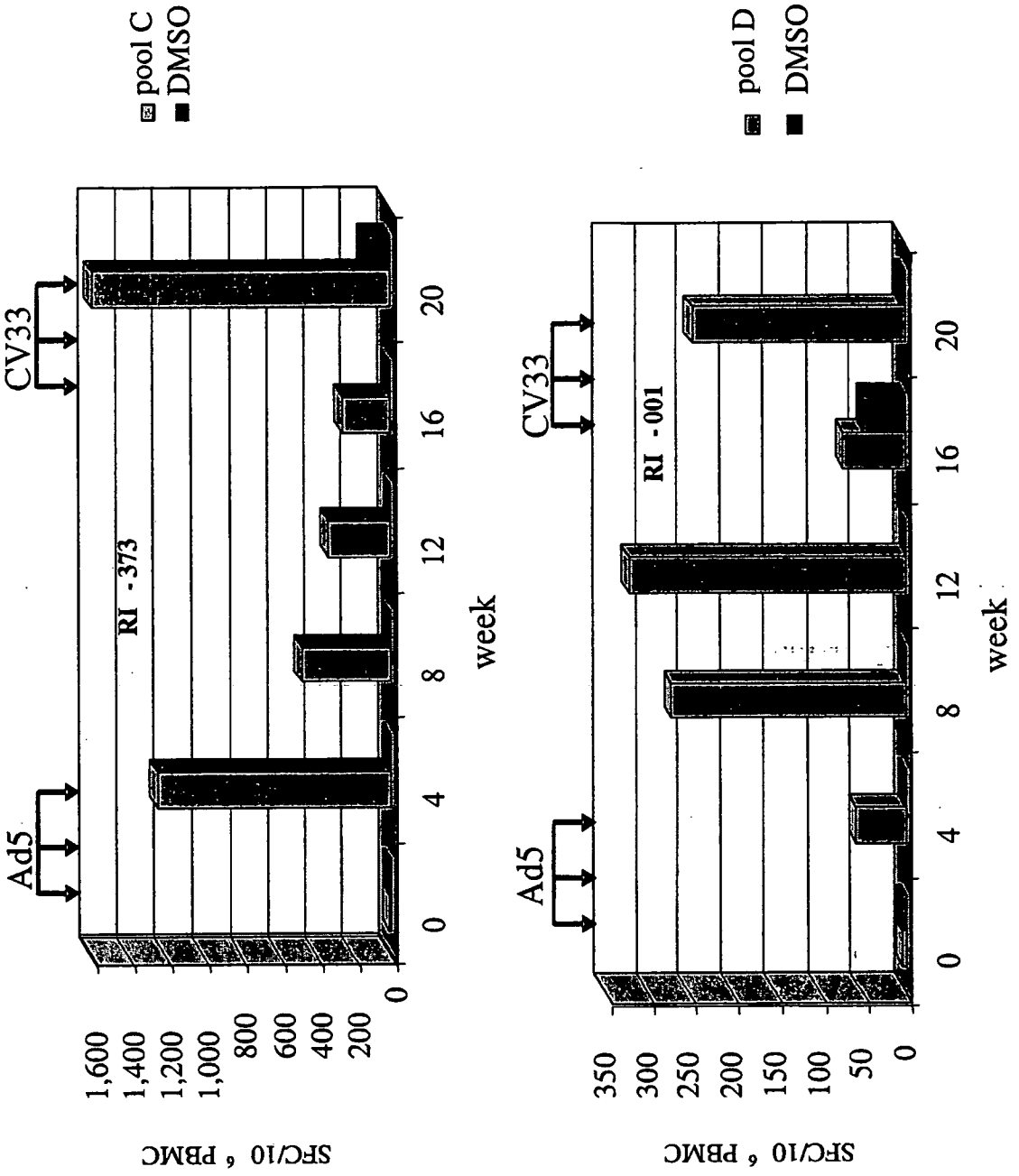
pool C
DMSO
RI -373
Ad5
CV33
SFC/10 ⁶ PBMC
week
pool D
DMSO
RI -001
Ad5
CV33
SFC/10 ⁶ PBMC
week

FIG. 34

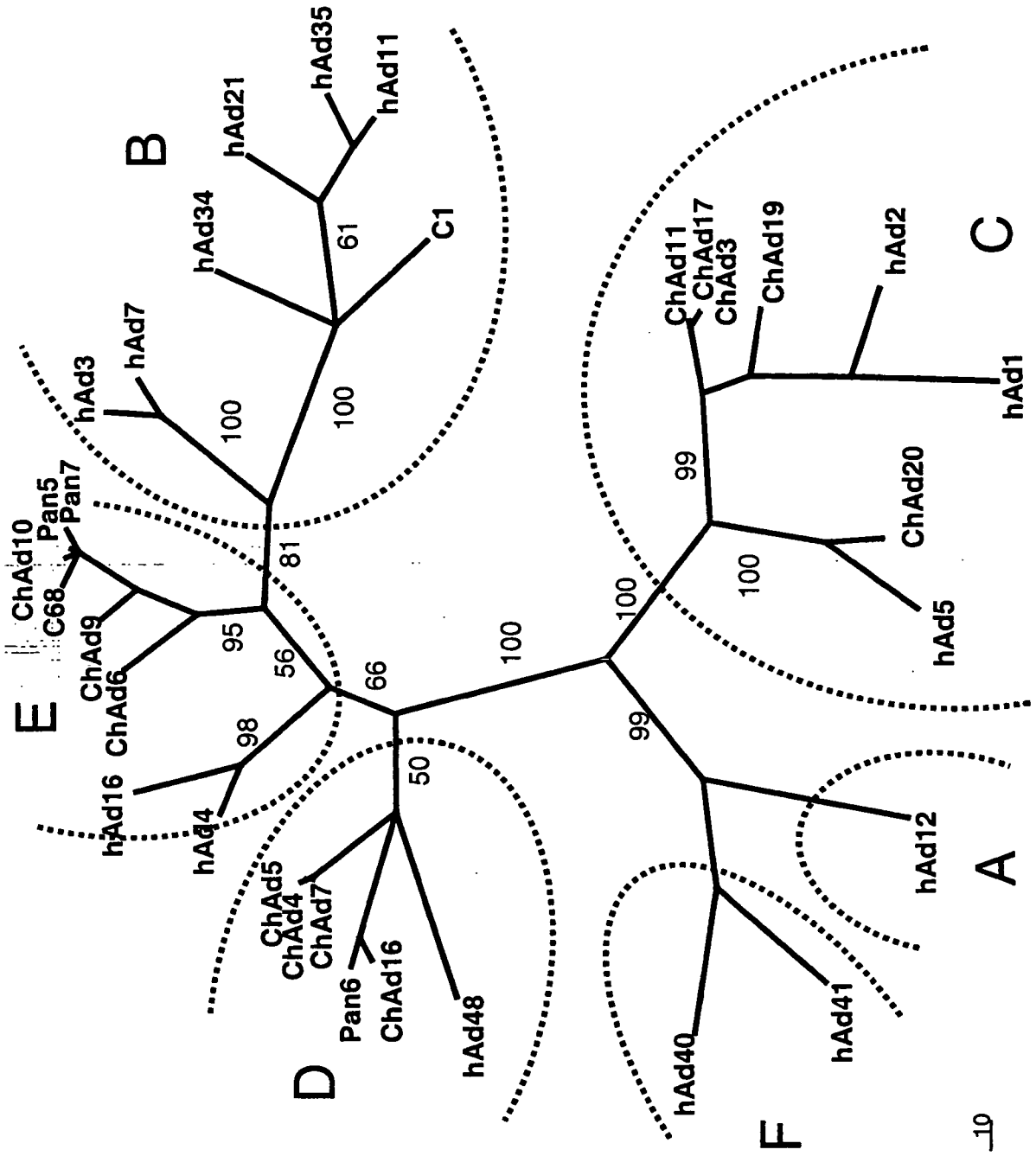
B
hAd35
hAd11
hAd21
hAd34
C1
61
100
100
hAd7
hAd3
ChAd10
Pan5
Pan7
C68
ChAd9
ChAd6
E
95
81
56
66
100
98
hAd16
hAd44
50
ChAd5
ChAd4
ChAd7
Pan6
ChAd16
hAd48
D
99
hAd40
hAd41
F
hAd12
A
100
99
ChAd11
ChAd17
ChAd3
ChAd19
hAd2
hAd1
C
100
ChAd20
hAd5
10

FIG. 35

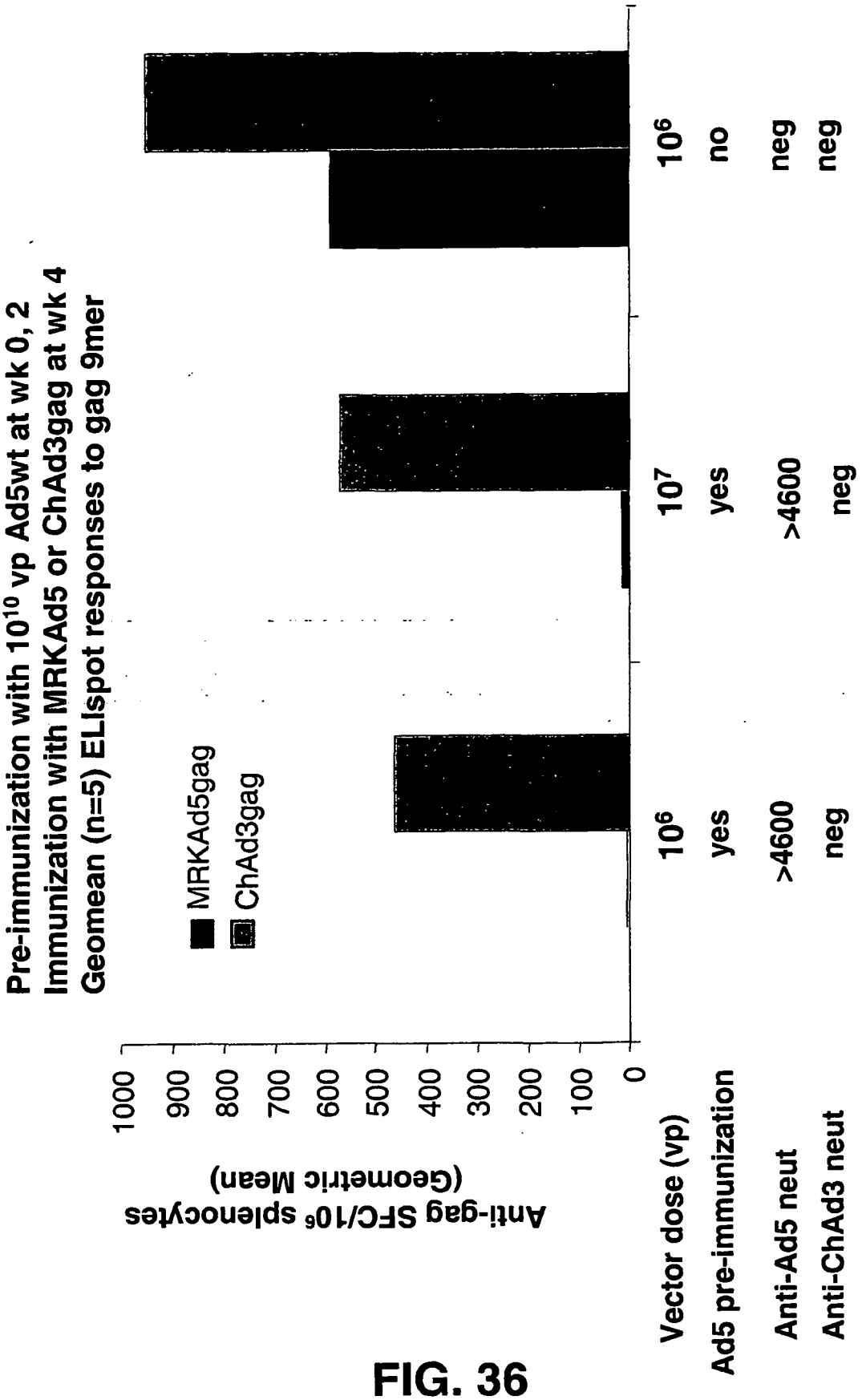
Pre-immunization with 10^10 vp Ad5wt at wk 0, 2
Immunization with MRKAd5 or ChAd3gag at wk 4
Geomean (n=5) ELIspot responses to gag 9mer
MRKAd5gag
ChAd3gag
Anti-gag SFC/10^6 splenocytes (Geometric Mean)
1000
900
800
700
600
500
400
300
200
100
0
Vector dose (vp)
10^6
10^7
10^6
Ad5 pre-immunization
yes
yes
no
Anti-Ad5 neut
>4600
>4600
neg
Anti-ChAd3 neut
neg
neg
neg

FIG. 36

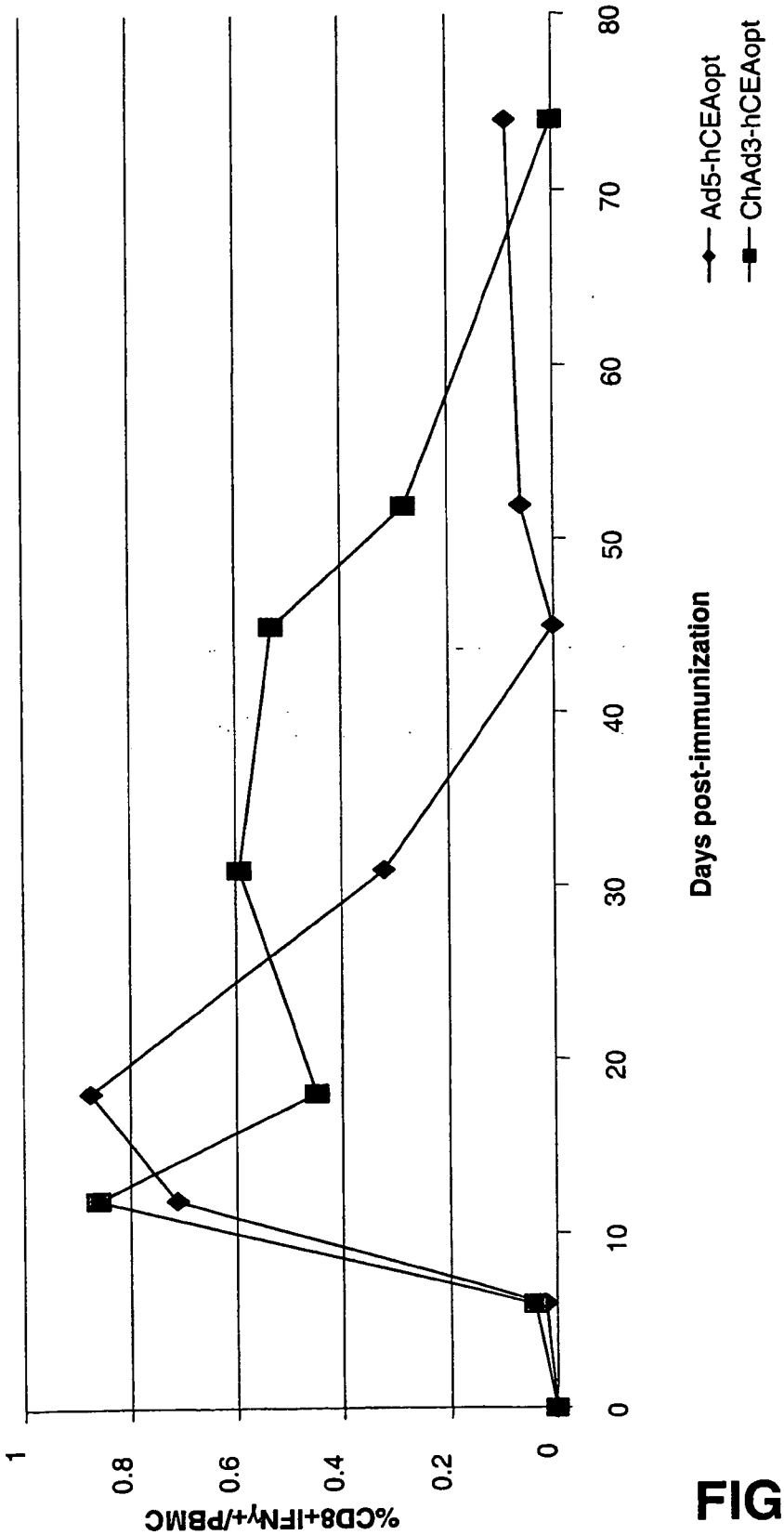
%CD8+IFNγ+/PBMC
0
0.2
0.4
0.6
0.8
1
0
10
20
30
40
50
60
70
80
Days post-immunization
Ad5-hCEAopt
ChAd3-hCEAopt

FIG. 37

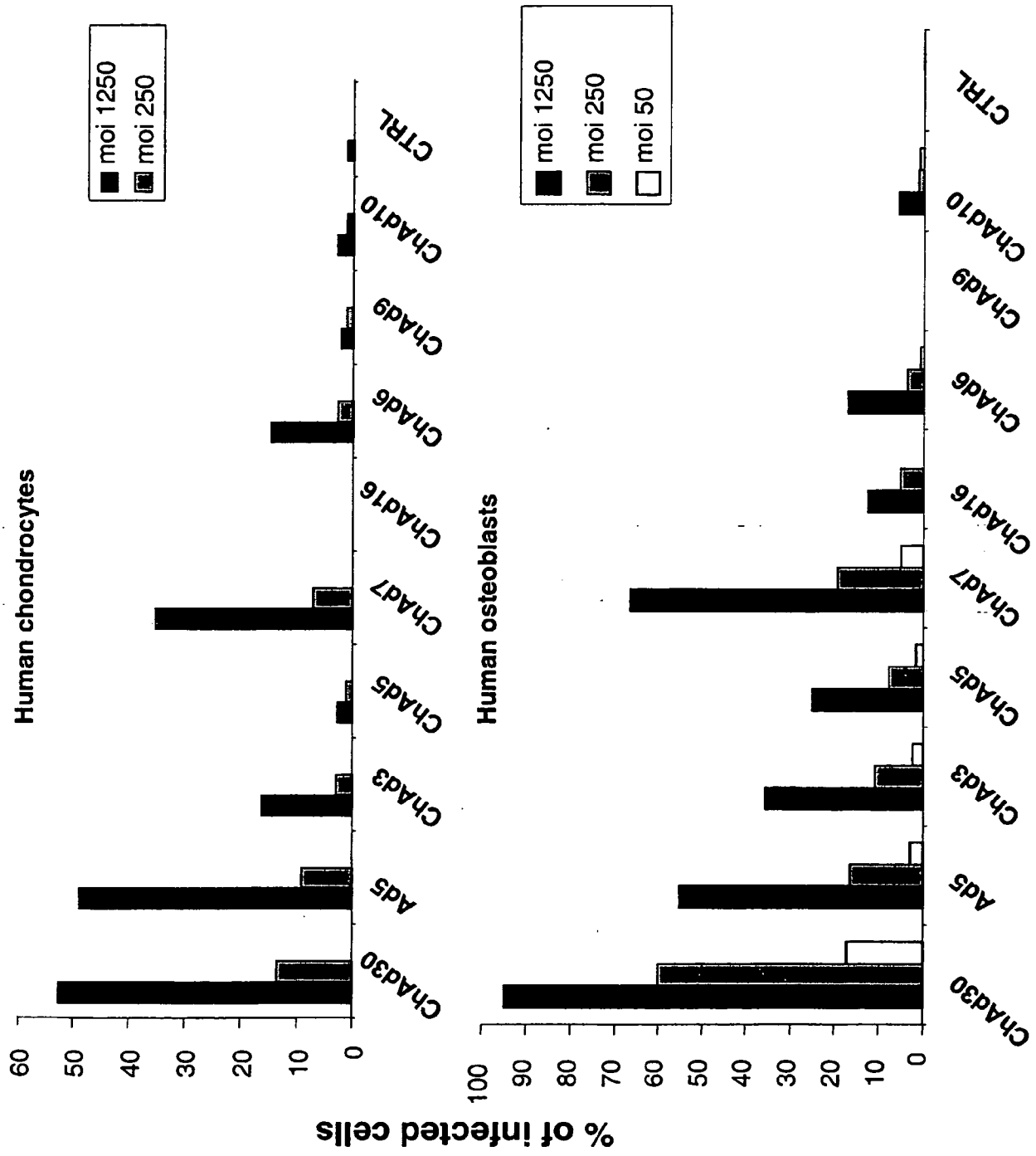
Human chondrocytes
moi 1250
moi 250
CTRL
ChAd10
ChAd9
ChAd6
ChAd16
ChAd7
ChAd5
ChAd3
Ad5
ChAd30
60
50
40
30
20
10
0
Human osteoblasts
moi 1250
moi 250
moi 50
CTRL
ChAd10
ChAd9
ChAd6
ChAd16
ChAd7
ChAd5
ChAd3
Ad5
ChAd30
100
90
80
70
60
50
40
30
20
10
0
% of infected cells

FIG. 38A

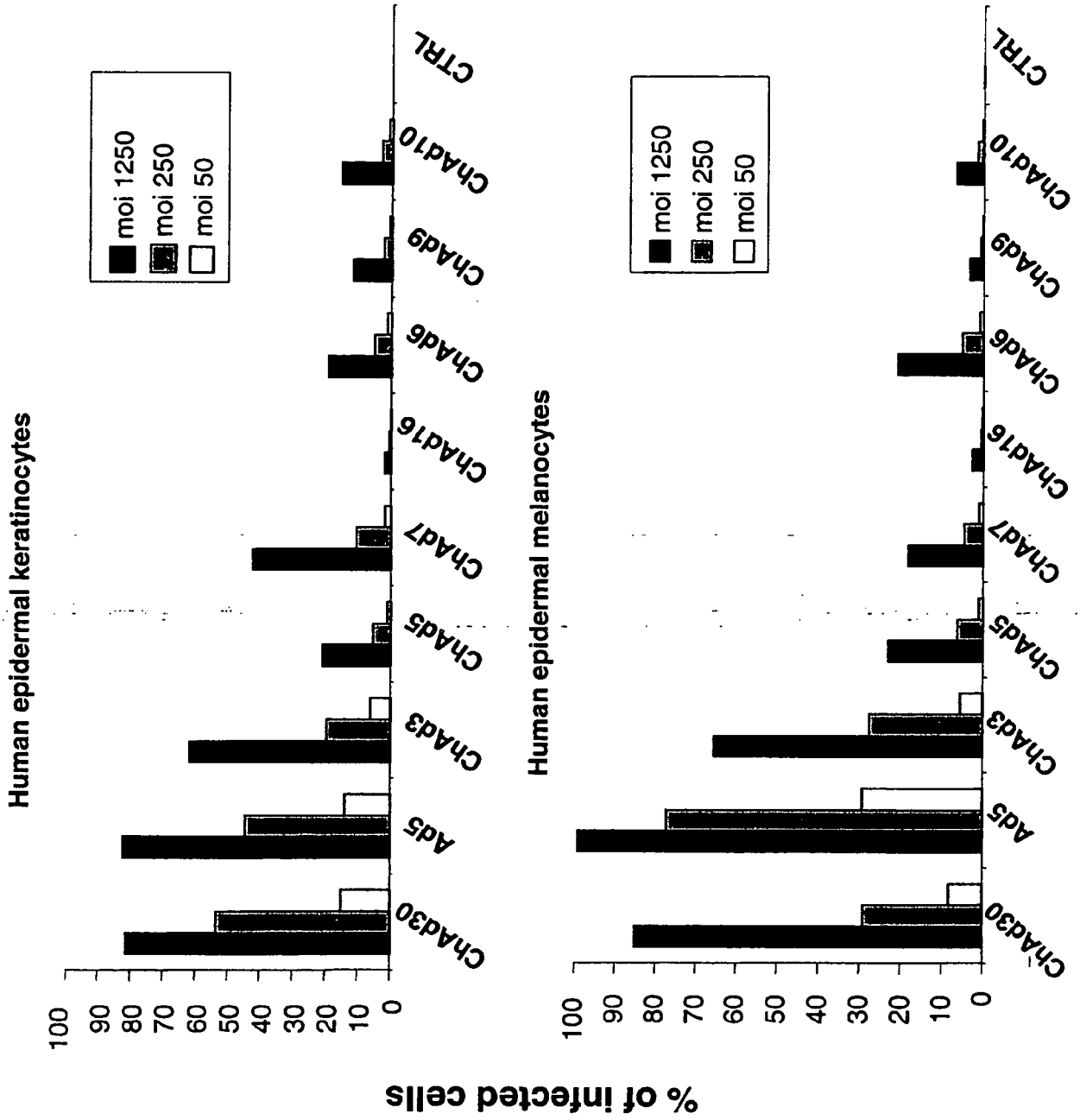
Human epidermal keratinocytes
Human epidermal melanocytes
moi 1250
moi 250
moi 50
% of infected cells
0
10
20
30
40
50
60
70
80
90
100
CTRL
ChAd10
ChAd9
ChAd6
ChAd16
ChAd7
ChAd5
ChAd3
Ad5
ChAd30

FIG. 38B

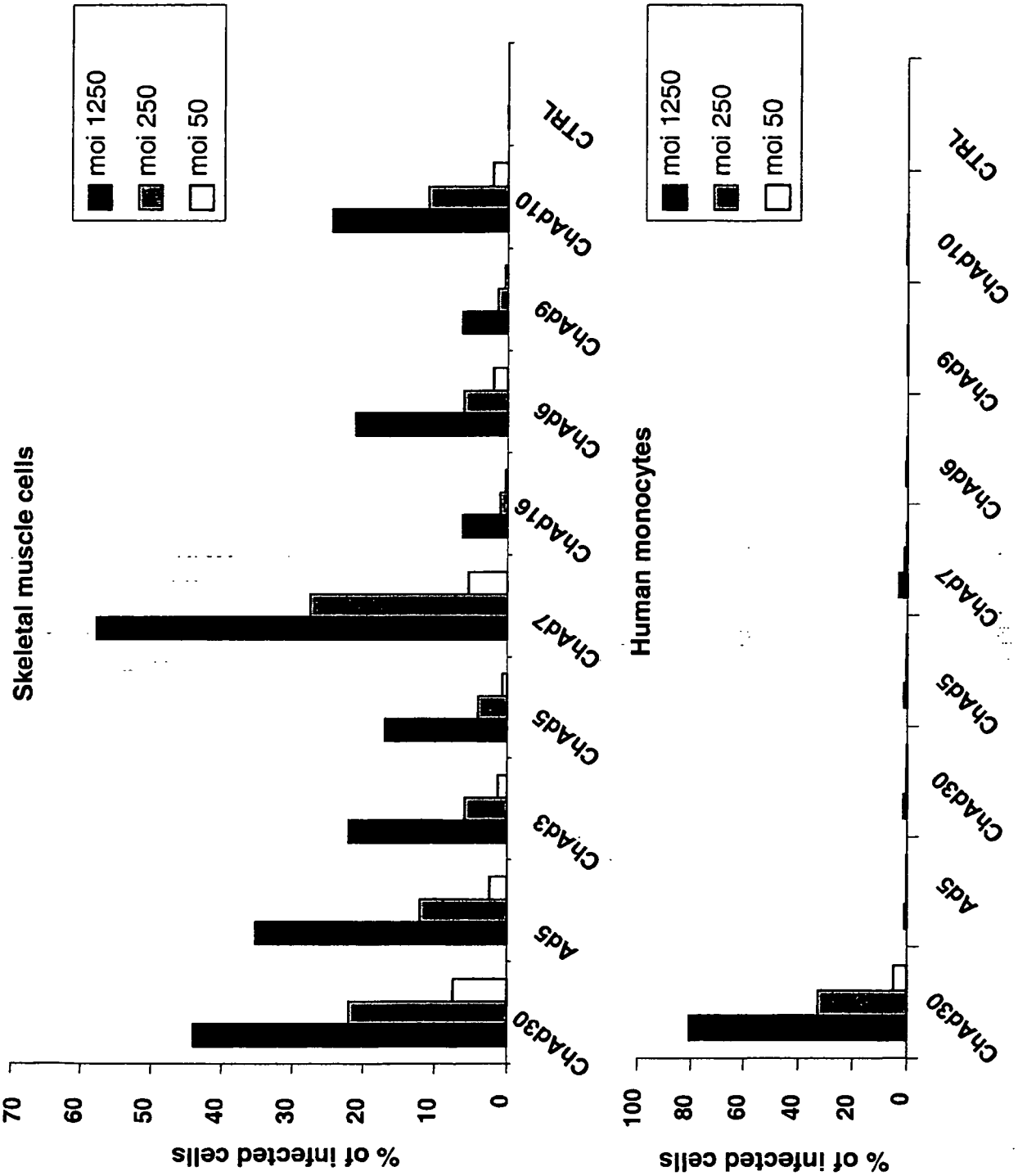
Skeletal muscle cells
% of infected cells
0
10
20
30
40
50
60
70
moi 1250
moi 250
moi 50
CTRL
ChAd10
ChAd9
ChAd6
ChAd16
ChAd7
ChAd5
ChAd3
Ad5
ChAd30
Human monocytes
% of infected cells
0
20
40
60
80
100
moi 1250
moi 250
moi 50
CTRL
ChAd10
ChAd9
ChAd6
ChAd7
ChAd5
ChAd30
Ad5
ChAd30

FIG. 38C

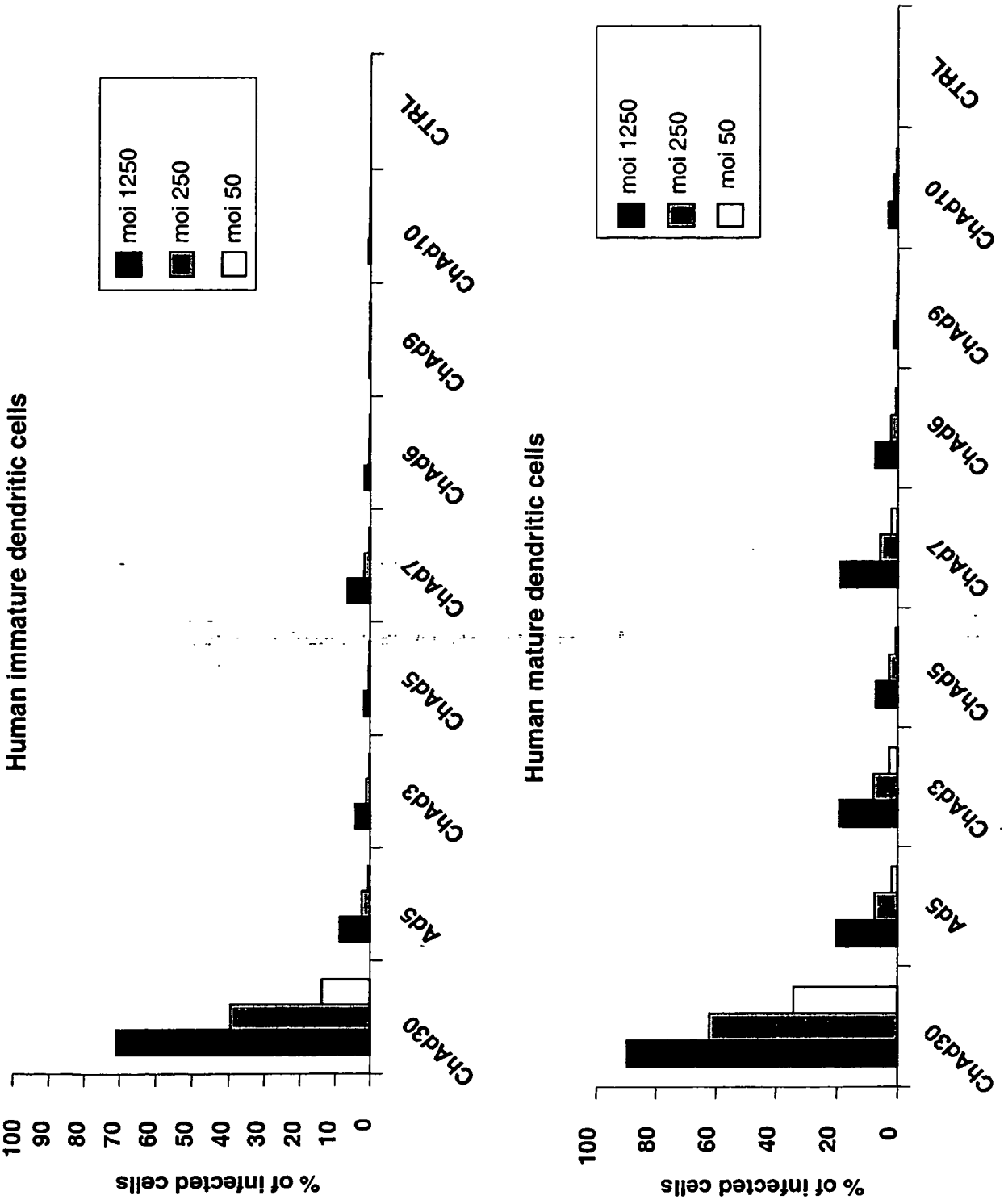
Human immature dendritic cells
moi 1250
moi 250
moi 50
% of infected cells
100
90
80
70
60
50
40
30
20
10
0
Chad30
Ad5
Chad3
Chad5
Chad7
Chad6
Chad9
Chad10
CTRL
Human mature dendritic cells
moi 1250
moi 250
moi 50
% of infected cells
100
80
60
40
20
0
Chad30
Ad5
Chad3
Chad5
Chad7
Chad6
Chad9
Chad10
CTRL

FIG. 38D

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 53879904 P **[0001]**
- US 6083716 A **[0057]**
- WO 03000851 PCT **[0057]**
- WO 9700326 A **[0100]**
- US 5849561 A **[0104]**
- US 6127175 A, Vigne **[0105]**
- WO 60538799 A **[0179]**

Non-patent literature cited in the description

- *In The Adenoviruses,* 1984, 451-498 **[0003]**
- **HIERHOLZER et al.** *J. Infect. Dis.,* 1988, vol. 158, 804-813 **[0003]**
- **SCHNURR ; DONDERO.** *Intervirology,* 1993, vol. 36, 79-83 **[0003]**
- **JONG et al.** *J Clin Microbiol.,* 1999, vol. 37, 3940-3945 **[0003]**
- **HE, Z et al.** *Virology,* 2001, vol. 270, 3583-3590 **[0005]**
- **XIANG, Z. et al.** *Virolgy,* 1996, vol. 219, 220-227 **[0005]**
- **RODRIGUEZ, E. et al.** *J. Immunol.,* 1997, vol. 158, 1268-1274 **[0005]**
- **MASTRANGELI et al.** *Human Gene Therapy,* 1996, vol. 7, 79-87 **[0005]**
- **GRAHAM, F.L. et al.** *J. Gen. Virol.,* 1977, vol. 36, 59-72 **[0007]**
- **FALLUX, F. J. et al.** *Human Gene Therapy,* 1998, vol. 9, 1909-1917 **[0007]**
- **WIGAND, R et al.** *Intervirology,* 1989, vol. 30, 1-9 **[0012]**
- **RUX, J.J. et al.** *J. Virol.,* 2003, vol. 77, 9553-9566 **[0020]**
- **W.C. RUSSEL.** *J. Gen. Virol.,* 2000, vol. 81, 2573-2604 **[0047]**
- **DAMBROSIO, E.** *J. Hyg.,* 1982, vol. 89, 209-219 **[0048]**
- **TOOGOOD, C.I.A. ; CROMPTON, J. ; HAY R.T.** *J. Gen. Virol.,* 1992, vol. 73, 1429-1435 **[0048]**
- **CRAWFORD-MIKSZA, L ; D.P.SCHNURR.** *J. Virol.,* 1996, vol. 70, 1836-1844 **[0048]**
- **SOUDOIS, C. et al.** *J. Virology,* 2000, vol. 74, 10639-10649 **[0052]**
- **FARINA, S.F. et al.** *J. Virology,* 2001, vol. 75, 11603-11613 **[0052]**
- **COHEN, C.J. et al.** *J. Gen. Virology,* 2002, vol. 83, 151-155 **[0052]**
- **BASNIGHT, M.** *Am. J. Epidemiol.,* 1971, vol. 94, 166-171 **[0053]**
- **FARINA, S. F. et al.** *J. Virol.,* 2001, vol. 75 (23), 11603-11613 **[0053] [0054]**
- **FARINA, S. F. et al.** *J. Virol.,* 2001, vol. 75 (23), 11603-13 **[0053]**
- **COHEN, C. et al.** *J. Gen. Virol.,* 2002, vol. 83, 151-155 **[0054]**
- **XIANG, Z et al.** *J. Virol.,* 2002, vol. 76 (5), 2667-2675 **[0055]**
- **FITZGERALD, J. C. et al.** *J. Immunol.,* 2003, vol. 170, 1416-1422 **[0055]**
- **BRODY et al.** *Ann N YAcad Sci.,* 1994, vol. 716, 90-101 **[0058]**
- **AUSUBEL et al.** Current Protocols in Molecular Biology. John Wiley & Sons, 1989 **[0070]**
- **CHAPMAN et al.** *Nucl. Acids Res,* 1991, vol. 19, 3979-3986 **[0071]**
- **GREENBERG, P. D.** *Adv. Immunol.,* 1991, vol. 49, 281-355 **[0086]**
- **PARDOLL, D. M. et al.** *Curr. Opin. Immunol.,* 1998, vol. 10, 588-94 **[0086]**
- *J. Clin. Oncol.,* 2002, vol. 20 (11), 2624-2632 **[0088]**
- **BERINSTEIN, N. L.** *J. Clin. Oncol.,* 2002, vol. 29 (8), 2197 **[0089] [0090]**
- **TSANG, K. Y. et al.** *J. Natl. Cancer Inst.,* 1995, vol. 87, 982-990 **[0090]**
- **SAMANCI, A. et al.** *Cancer Immunol. Immunother,* 1998, vol. 47, 131-142 **[0090]**
- **FOON, L, A. et al.** *J. Clin. Invest..,* 1995, vol. 96, 334-342 **[0090]**
- **PRESS et al.** *Oncogene,* 1990, vol. 5, 953-962 **[0092]**
- **DISIS ; CHEEVER.** *Adv. Cancer Research,* 1997, vol. 71, 343-371 **[0092]**
- **SLAMON et al.** *Science,* 1989, vol. 244, 707-712 **[0092]**
- **AKIYAMA, T et al.** *Science,* 1986, vol. 232, 1644-1646 **[0093]**
- **GOLDENBERG MM.** *Clin. Ther.,* 1999, vol. 21, 309-318 **[0093]**
- **SAMBROOK et al.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, 1989 **[0094]**
- **GRAHAM et al.** *J. Gen. Virol,* 1977, vol. 36, 59-72 **[0100]**

- **FALLAUX, F.J. et al.** *Human Gene Therapy,* 1998, vol. 9, 1909-1917 **[0101]**
- **VOGELS R.** Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. *J Virol.,* 2003, vol. 77 (15), 8263-71 **[0103]**
- **ABRAHAMSEN et al.** *J. Virol.,* 1997, 8946-8951 **[0104]**
- **WILLIAM R. PENSON.** CABIOS. University of Virginia, 1989, vol. 4, 11-17 **[0128]**
- Adenoviridae and their replication. **HOROWITZ, MS.** Virology. Raven Press, 1990, 1679-1740 **[0131]**
- Adenoviridae and their replication. **HOROWITZ, MS.** Virology. raven Press, 1990, 1679-1740 **[0132]**
- **ROMANI, N. et al.** *J. Immunol. Methods,* 1996, vol. 196, 137 **[0148]**
- **CLARKE, P et al.** *Cancer Res.,* 1998, vol. 58 (7), 1469-77 **[0165]**
- Guide for Care and Use of Laboratory Animals. Institute of Laboratory Animal Resources, National Research Council **[0169]**
- **ALLEN et al.** *J. Virol.,* 2001, vol. 75 (2), 738-749 **[0170]**

Csimpánz adenovírus vakcina hordozók

Szabadalmi igénypontok

1. Replikáció-defektív csimpánz adenovirális vektor, amely tartalmazza a következőket:

a SEQ ID NO: 116. azonosítószámú szekvencia szerinti peptidet kódoló nukleotid-szekvenciát tartalmazó rekombináns adenovirális genom; és

legalább egy immunogént kódoló transzgén, amely működőképesen a transzgén emlőssejtben való expressziójának irányítására szolgáló szabályozó szekvenciákhoz van kapcsolva, ahol a vektor úgy módosított, hogy az E1-gén funkcionálisan deletálva lett.

2. Az 1. igénypont szerinti replikáció-defektív csimpánz adenovírus vektor, ahol a rekombináns adenovirális genom SEQ ID NO: 115. azonosítószámú szekvencia szerinti nukleotid-szekvenciát tartalmaz.

3. Az 1. vagy 2. igénypont szerinti, replikáció-defektív csimpánz adenovirális vektor, amelyben a rekombináns adenovirális genom SEQ ID NO: 76. azonosítószámú szekvencia szerinti nukleotid-szekvenciát tartalmaz.

4. Gazdasejt, amely 1-3. igénypontok bármelyike szerinti replikáció-defektív csimpánz adenovirális vektort tartalmaz, ahol a gazdasejt egy vagy több adenovirális régiót expresszál a következők által alkotott csoportból kiválasztva: E1a, E1b, E2a, E2b, E4 1., 2., 3., 4., 5., 6., 6/7. orf-ok, pIX, IVa2, L1, L2, L3, L4, L5 régiók.

5. Izolált rekombináns adenovírusgenom, amely tartalmaz:

a) SEQ ID NO: 115. azonosítószámú szekvencia szerinti nukleotid-szekvenciát tartalmazó hexongén-szekvenciát, és

b) SEQ ID NO: 76. azonosítószámú szekvencia szerinti nukleotid-szekvenciát tartalmazó fibergén-szekvenciát.

6. Az 1-3. igénypontok bármelyike szerinti, replikáció-defektív, csimpánz adenovirális vektor, antigén-specifikus immunválasz emlősben történő megerősítésére szolgáló eljárásban történő alkalmazásra, amely eljárás magában foglalja a rekombináns ChAd vektor elegendő mennyiségének beadását az emlősnek, ahol a ChAd vektor beadása megerősített választ vált ki.

7. A 6. igénypont szerinti vektor, ahol a csimpánz adenovirális vektor az E1 génjeinek a teljes delécióját tartalmazza, és továbbá ahol a vektor adott esetben az E3 génjeinek a delécióját is tartalmazza.

8. A 6. igénypont szerinti vektor, ahol az immunválasz TAA-ra specifikus megerősített immunválasz.

9. A 6. vagy 8. igénypont szerinti vektor, ahol a megerősített immunválasz magában foglalja antigén-specifikus CD8+ T-sejtek termelődését.

10. Az 1-3. igénypontok bármelyike szerinti, replikáció-defektív csimpánz adenovirális vektor, antigén-specifikus immunválasz naiv emlősben történő kiváltására szolgáló eljárásban történő alkalmazásra, amely eljárás magában foglalja az adenovirális vektor elegendő mennyiségének beadását az emlősnek, ahol az adenovirális vektor beadása primer immunválaszt vált ki.

11. A 6. vagy 10. igénypont szerinti vektor, ahol rendre a megerősített vagy primer immunválasz fertőző ágensekből származó antigénre specifikus, amelyek például, de nem kizárólagosan, a következők: HIV, HCV, HPV, HSV1, HSV2, SARS CoV, *Plasmodium malariae*, Ebola vírus, West Nile vírus, Dengue vírus, Influenza A, Influenza B, *Mycobacterium tuberculosis* CMV és RSV.

SZTNH-100052586

12. A 6. igénypont szerinti vektor, ahol az immunválasz az emlős által tolerált TAA-ra specifikus primer immunválasz.

13. Az 1-3. igénypontok bármelyike szerinti, replikáció-defektív, csimpánz adenovirális vektor fertőző ágensből származó antigénre specifikus immunválasz indukálására szolgáló eljárásban történő alkalmazásra, amely eljárás tartalmazza a következő lépéseket:

(a) gazda fertőzőágens-antigénre adott válaszának beindítása, olyan fertőzőágens-antigént kódoló nukleotid-szekvenciát tartalmazó első vakcina-készítmény beadásával, amely ellení antigén-specifikus immunválaszt kívánunk; és

(b) az (a) lépésbeli immunválasz megerősítése egy második vakcina-készítmény beadásával, amely 1. igénypont szerinti olyan rekombináns ChAd-vektort tartalmaz, amely az E1 génjében legalább egy funkcionális deléciót, és az E1 gén deléciója helyében olyan szekvenciát tartalmaz, amely a beindítási lépésben bejuttatottal azonos fertőzőágens-antigént kódoló DNS expresszióját irányítani képes promóter tartalmaz;

ahol a megerősítő készítmény beadása olyan immunválasz vált ki, amelynek védő immunitást biztosító hatása van, ahol a fertőző ágensből származó antigén a következők alkotta csoportól választott: HIV, HBV, HCV, HPV, HSV1, HSV2, SARS, CoV, *Leishmania major, Plasmodium malaria,* Ebolavírus, West Nile vírus, Dengue vírus, Influenza A, Influenza B, RSV, CMV és *Mycobacterium tuberculosis.*

14. A 13. igénypont szerinti vektor, ahol az immunválasz tartalmazza antigén-specifikus CD8+ T-sejtek termelését.

15. Az 1-3. igénypontok bármelyike szerinti, replikáció-defektív, csimpánz adenovirális vektor, gazda saját-antigén iránti toleranciájának megtörésére szolgáló eljárásban történő alkalmazásra, amely eljárás magában foglalja a következőket:

(a) gazda saját-antigénre adott válaszának beindítása, olyan saját-antigént kódoló nukleotid-szekvenciát tartalmazó első vakcina-készítmény beadásával, amely ellení antigén-specifikus immunválaszt kívánunk, ezáltal beindított választ váltva ki; és

(b) az (a) lépésbeli beindított immunválasz megerősítése egy második vakcina-készítmény beadásával, amely 1. igénypont szerinti olyan rekombináns vektort tartalmaz, amely az E1 génjében legalább egy funkcionális deléciót, és az E1 gén deléciója helyében olyan szekvenciát tartalmaz, amely a beindítási lépésben bejuttatottal azonos saját-antigént kódoló DNS expresszióját irányítani képes promóter tartalmaz;

ahol a megerősítő készítmény beadása olyan immunválasz vált ki, amelynek a gazda saját-antigén iránti toleranciáját letörő hatása van.

16. A 15. igénypont szerinti replikáció-defektív csimpáz adenovirális vektor az ott meghatározott eljárásban történő alkalmazásra, ahol a saját antigén a következők által alkotott csoportból kiválasztott: HER2 NEU, CEA, HEPCAM, PSA, PSMA, TELOMERÁZ, GP100, MELAN-A/MART-1, MUC-1, NY-ESO-1, SURVIVIN, STROMELIZIN 3, TIROZINÁZ, MAGE3, CML68, CML66, OY-TES-1, SSX-2, SART-1, SART-2, SART-3, NY-CO-58, NY-BR-62, HKLP2, 5T4 és VEGFR2.

17. A 15. igénypont szerinti replikáció-defektív csimpánz adenovirális vektor, ahol az immunválasz magában foglalja antigén-specifikus CD8+ T-sejtek termelődését.

18. A 15. vagy 16. igénypont szerinti, replikáció-defektív, csimpánz adenovirális vektor, amelyben az első és a második vakcinakészítmény egyaránt adenovirális vektor, amelyek különböző szerotípusok által jellemzettek.