



(86) Date de dépôt PCT/PCT Filing Date: 2001/09/14
(87) Date publication PCT/PCT Publication Date: 2002/03/21
(85) Entrée phase nationale/National Entry: 2003/03/14
(86) N° demande PCT/PCT Application No.: US 2001/028861
(87) N° publication PCT/PCT Publication No.: 2002/022080
(30) Priorité/Priority: 2000/09/15 (60/233,180) US

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 15/86

(71) Demandeur/Applicant:
MERCK & CO., INC., US

(72) Inventeurs/Inventors:
EMINI, EMILIO A., US;
YOUIL, RIMA, US;
BETT, ANDREW J., US;
CHEN, LING, US;
KASLOW, DAVID C., US;
SHIVER, JOHN W., US;
TONER, TIMOTHY J., US;
CASIMIRO, DANILO R., US

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : VACCINS ADENOVIRAUX DE PREMIERE GENERATION EVOLUES, EXPRIMANT LES PROTEINES GAG, POL ET NEF DU VIH-1 A OPTIMISATION DES CODONS ET LEURS MODIFICATIONS

(54) Title: ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

(57) **Abrégé/Abstract:**

First generation adenoviral vectors and associated recombinant adenovirus-based HIV vaccines which show enhanced stability and growth properties and greater cellular-mediated immunity are described within this specification. These adenoviral vectors are utilized to generate and produce through cell culture various adenoviral-based HIV-1 vaccines which contain HIV-1 gag, HIV-1 pol and/or HIV-1 nef polynucleotide pharmaceutical products, and biologically relevant modifications thereof. These adenovirus vaccines, when directly introduced into living vertebrate tissue, preferably a mammalian host such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV1- Gag, Pol and/or Nef protein or biologically modification thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding HIV-1 Gag, encoding codon optimized HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNase H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The adenoviral vaccines of the present invention, when administered alone or in a combined modality regime, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

TITLE OF THE INVENTION

ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING
CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit, under 35 U.S.C. §119(e), of U.S.
provisional applications 60/233,180, 60/279,056, and Attorney Docket 20867PV2
(serial number unassigned), filed September 15, 2000, March 27, 2001, and
September 7, 2001, respectively.

10

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not Applicable

REFERENCE TO MICROFICHE APPENDIX

15

Not Applicable

FIELD OF THE INVENTION

The present invention relates to recombinant, replication-deficient first
generation adenovirus vaccines found to exhibit enhanced growth properties and
20 greater cellular-mediated immunity as compared to other replication-deficient vectors.
The invention also relates to the associated first generation adenoviral vectors
described herein, which, through the incorporation of additional 5' adenovirus
sequence, enhance large scale production efficiency of the recombinant, replication-
defective adenovirus described herein. Another aspect of the instant invention is the
25 surprising discovery that the intron A portion of the human cytomegalovirus (hCMV)
promoter constitutes a region of instability in adenoviral vector constructs. Removal
of this region from adenoviral expression constructs results in greatly improved vector
stability. Therefore, improved vectors expressing a transgene under the control of an
intron A-deleted CMV promoter constitute a further aspect of this invention. These
30 adenoviral vectors are useful for generating recombinant adenovirus vaccines against
human immunodeficiency virus (HIV). In particular, the first generation adenovirus
vectors disclosed herein are utilized to construct and generate adenovirus-based HIV-
1 vaccines which contain HIV-1 Gag, HIV-1 Pol and/or HIV-1 Nef polynucleotide
pharmaceutical products, and biologically active modifications thereof. Host
35 administration of the recombinant, replication-deficient adenovirus vaccines described
herein results in expression of HIV-1 Gag, HIV-1- Pol and/or Nef protein or

immunologically relevant modifications thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding codon optimized HIV-1 Gag, HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNase H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef, and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The HIV adenovirus vaccines of the present invention, when administered alone or in a combined modality and/or prime/boost regimen, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

BACKGROUND OF THE INVENTION

Human Immunodeficiency Virus-1 (HIV-1) is the etiological agent of acquired human immune deficiency syndrome (AIDS) and related disorders. HIV-1 is an RNA virus of the Retroviridae family and exhibits the 5' LTR-*gag-pol-env*-LTR 3' organization of all retroviruses. The integrated form of HIV-1, known as the provirus, is approximately 9.8 Kb in length. Each end of the viral genome contains flanking sequences known as long terminal repeats (LTRs). The HIV genes encode at least nine proteins and are divided into three classes; the major structural proteins (Gag, Pol, and Env), the regulatory proteins (Tat and Rev); and the accessory proteins (Vpu, Vpr, Vif and Nef).

The *gag* gene encodes a 55-kilodalton (kDa) precursor protein (p55) which is expressed from the unspliced viral mRNA and is proteolytically processed by the HIV protease, a product of the *pol* gene. The mature p55 protein products are p17 (matrix), p24 (capsid), p9 (nucleocapsid) and p6.

The *pol* gene encodes proteins necessary for virus replication; a reverse transcriptase, a protease, integrase and RNase H. These viral proteins are expressed as a Gag-Pol fusion protein, a 160 kDa precursor protein which is generated via a ribosomal frame shifting. The viral encoded protease proteolytically cleaves the Pol polypeptide away from the Gag-Pol fusion and further cleaves the Pol polypeptide to the mature proteins which provide protease (Pro, P10), reverse transcriptase (RT, P50), integrase (IN, p31) and RNase H (RNase, p15) activities.

The *nef* gene encodes an early accessory HIV protein (Nef) which has been shown to possess several activities such as down regulating CD4 expression, disturbing T-cell activation and stimulating HIV infectivity.

5 The *env* gene encodes the viral envelope glycoprotein that is translated as a 160-kilodalton (kDa) precursor (gp160) and then cleaved by a cellular protease to yield the external 120-kDa envelope glycoprotein (gp120) and the transmembrane 41-kDa envelope glycoprotein (gp41). Gp120 and gp41 remain associated and are displayed on the viral particles and the surface of HIV-infected cells.

10 The *tat* gene encodes a long form and a short form of the Tat protein, a RNA binding protein which is a transcriptional transactivator essential for HIV-1 replication.

The *rev* gene encodes the 13 kDa Rev protein, a RNA binding protein. The Rev protein binds to a region of the viral RNA termed the Rev response element (RRE). The Rev protein promotes transfer of unspliced viral RNA from the nucleus
15 to the cytoplasm. The Rev protein is required for HIV late gene expression and in turn, HIV replication.

Gp120 binds to the CD4/chemokine receptor present on the surface of helper T-lymphocytes, macrophages and other target cells in addition to other co-receptor molecules. X4 (macrophage tropic) virus show tropism for CD4/CXCR4 complexes
20 while a R5 (T-cell line tropic) virus interacts with a CD4/CCR5 receptor complex. After gp120 binds to CD4, gp41 mediates the fusion event responsible for virus entry. The virus fuses with and enters the target cell, followed by reverse transcription of its single stranded RNA genome into the double-stranded DNA via a RNA dependent DNA polymerase. The viral DNA, known as provirus, enters the cell nucleus, where
25 the viral DNA directs the production of new viral RNA within the nucleus, expression of early and late HIV viral proteins, and subsequently the production and cellular release of new virus particles. Recent advances in the ability to detect viral load within the host shows that the primary infection results in an extremely high generation and tissue distribution of the virus, followed by a steady state level of virus
30 (albeit through a continual viral production and turnover during this phase), leading ultimately to another burst of virus load which leads to the onset of clinical AIDS. Productively infected cells have a half life of several days, whereas chronically or latently infected cells have a 3-week half life, followed by non-productively infected cells which have a long half life (over 100 days) but do not significantly contribute to
35 day to day viral loads seen throughout the course of disease.

Destruction of CD4 helper T lymphocytes, which are critical to immune defense, is a major cause of the progressive immune dysfunction that is the hallmark of HIV infection. The loss of CD4 T-cells seriously impairs the body's ability to fight most invaders, but it has a particularly severe impact on the defenses against viruses, fungi, parasites and certain bacteria, including mycobacteria.

Effective treatment regimens for HIV-1 infected individuals have become available recently. However, these drugs will not have a significant impact on the disease in many parts of the world and they will have a minimal impact in halting the spread of infection within the human population. As is true of many other infectious diseases, a significant epidemiologic impact on the spread of HIV-1 infection will only occur subsequent to the development and introduction of an effective vaccine. There are a number of factors that have contributed to the lack of successful vaccine development to date. As noted above, it is now apparent that in a chronically infected person there exists constant virus production in spite of the presence of anti-HIV-1 humoral and cellular immune responses and destruction of virally infected cells. As in the case of other infectious diseases, the outcome of disease is the result of a balance between the kinetics and the magnitude of the immune response and the pathogen replicative rate and accessibility to the immune response. Pre-existing immunity may be more successful with an acute infection than an evolving immune response can be with an established infection. A second factor is the considerable genetic variability of the virus. Although anti-HIV-1 antibodies exist that can neutralize HIV-1 infectivity in cell culture, these antibodies are generally virus isolate-specific in their activity. It has proven impossible to define serological groupings of HIV-1 using traditional methods. Rather, the virus seems to define a serological "continuum" so that individual neutralizing antibody responses, at best, are effective against only a handful of viral variants. Given this latter observation, it would be useful to identify immunogens and related delivery technologies that are likely to elicit anti-HIV-1 cellular immune responses. It is known that in order to generate CTL responses antigen must be synthesized within or introduced into cells, subsequently processed into small peptides by the proteasome complex, and translocated into the endoplasmic reticulum/Golgi complex secretory pathway for eventual association with major histocompatibility complex (MHC) class I proteins. CD8⁺ T lymphocytes recognize antigen in association with class I MHC via the T cell receptor (TCR) and the CD8 cell surface protein. Activation of naive CD8⁺ T cells into activated effector or memory cells generally requires both TCR engagement of antigen as described above as well as engagement of costimulatory proteins. Optimal

induction of CTL responses usually requires "help" in the form of cytokines from CD4⁺ T lymphocytes which recognize antigen associated with MHC class II molecules via TCR and CD4 engagement.

European Patent Applications 0 638 316 (Published February 15, 1995) and 0
5 586 076 (Published March 9, 1994), (both assigned to American Home Products Corporation) describe replicating adenovirus vectors carrying an HIV gene, including *env* or *gag*. Various treatment regimens were used with chimpanzees and dogs, some of which included booster adenovirus or protein plus alum treatments.

Replication-defective adenoviral vectors harboring deletions in the E1 region
10 are known, and recent adenoviral vectors have incorporated the known packaging repeats into these vectors; e.g., see EP 0 707 071, disclosing, *inter alia*, an adenoviral vector deleted of E1 sequences from base pairs 459 to 3328; and U.S. Patent No. 6,033,908, disclosing, *inter alia*, an adenoviral vector deleted of base pairs 459-3510. The packaging efficiency of adenovirus has been taught to depend on the number of
15 incorporated individual A (packaging) repeats; *see, e.g.*, Gräble and Hearing, 1990 *J. Virol.* 64(5):2047-2056; Gräble and Hearing, 1992 *J. Virol.* 66(2):723-731.

Larder, et al., (1987, *Nature* 327: 716-717) and Larder, et al., (1989, *Proc. Natl. Acad. Sci.* 86: 4803-4807) disclose site specific mutagenesis of HIV-1 RT and the effect such changes have on *in vitro* activity and infectivity related to interaction
20 with known inhibitors of RT.

Davies, et al. (1991, *Science* 252:, 88-95) disclose the crystal structure of the RNase H domain of HIV-1 Pol.

Schatz, et al. (1989, *FEBS Lett.* 257: 311-314) disclose that mutations Glu478Gln and His539Phe in a complete HIV-1 RT/RNase H DNA fragment results
25 in defective RNase activity without effecting RT activity.

Mizrahi, et al. (1990, *Nucl. Acids. Res.* 18: pp. 5359-5353) disclose additional mutations Asp443Asn and Asp498Asn in the RNase region of the *pol* gene which also results in defective RNase activity. The authors note that the Asp498Asn mutant was difficult to characterize due to instability of this mutant protein.

Leavitt, et al. (1993, *J. Biol. Chem.* 268: 2113-2119) disclose several
30 mutations, including a Asp64Val mutation, which show differing effect on HIV-1 integrase (IN) activity.

Wiskerchen, et al. (1995, *J. Virol.* 69: 376-386) disclose single and double mutants, including mutation of aspartic acid residues which effect HIV-1 IN and viral
35 replication functions.

It would be of great import in the battle against AIDS to produce a prophylactic- and/or therapeutic-based HIV vaccine which generates a strong cellular immune response against an HIV infection. The present invention addresses and meets these needs by disclosing a class of adenovirus vaccines which, upon host administration, express codon optimized and modified versions of the HIV-1 genes, *gag*, *pol* and *nef*. These recombinant, replication-defective adenovirus vaccines may be administered to a host, such as a human, alone or as part of a combined modality regimen and/or prime-boost vaccination regimen with components of the present invention and/or a distinct viral HIV DNA vaccine, non-viral HIV DNA vaccine, HIV subunit vaccine, an HIV whole killed vaccine and/or a live attenuated HIV vaccine.

SUMMARY OF THE INVENTION

The present invention relates to enhanced replication-defective recombinant adenovirus vaccine vectors and associated recombinant, replication-deficient adenovirus vaccines which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef. The adenovirus vaccines of the present invention express HIV antigens and provide for improved cellular-mediated immune responses upon host administration. Potential vaccinees include but are not limited to primates and especially humans and non-human primates, and also include any non-human mammal of commercial or domestic veterinary importance. An effect of the improved recombinant adenovirus-based vaccines of the present invention should be a lower transmission rate to previously uninfected individuals (i.e., prophylactic applications) and/or reduction in the levels of the viral loads within an infected individual (i.e., therapeutic applications), so as to prolong the asymptomatic phase of HIV-1 infection. In particular, the present invention relates to adenoviral-based vaccines which encode various forms of codon optimized HIV-1 Gag (including but in no way limited to p55 versions of codon optimized full length (FL) Gag and tPA-Gag fusion proteins), HIV-1 Pol, HIV-1 Nef, and selected modifications of immunological relevance. The administration, intracellular delivery and expression of these adenovirus vaccines elicit a host CTL and Th response. The preferred replication-defective recombinant adenoviral vaccine vectors include but are not limited to synthetic DNA molecules which (1) encode codon optimized versions of wild type HIV-1 Gag; (2) encode codon optimized versions of HIV-1 Pol; (3) encode codon optimized versions of HIV-1 Pol fusion proteins; (4) encode codon optimized versions of modified HIV-1 Pol proteins and fusion proteins, including but not limited

to *pol* modifications involving residues within the catalytic regions responsible for RT, RNase and IN activity within the host cell; (5) encode codon optimized versions of wild type HIV-1 Nef; (6) codon optimized versions of HIV-1 Nef fusion proteins; and/or (7) codon optimized versions of HIV-1 Nef derivatives, including but not limited to *nef* modifications involving introduction of an amino-terminal leader sequence, removal of an amino-terminal myristylation site and/or introduction of dileucine motif mutations. The Nef-based fusion and modified proteins, disclosed within this specification and expressed from an adenoviral-based vector vaccine this specification, may possess altered trafficking and/or host cell function while retaining the ability to be properly presented to the host MHC I complex and in turn elicit a host CTL and Th response. Examples of HIV-1 Gag, Pol and/or Nef fusion proteins include but are not limited to fusion of a leader or signal peptide at the NH₂-terminal portion of the viral antigen coding region. Such a leader peptide includes but is not limited to a tPA leader peptide.

The adenoviral vector utilized in construction of the HIV-1 Gag-, HIV-1 Pol- and/or HIV-1 Nef- based vaccines of the present invention may comprise any replication-defective adenoviral vector which provides for enhanced genetic stability of the recombinant adenoviral genome through large scale production and purification of the recombinant virus. In other words, an HIV-1 Gag-, Pol- or Nef-based adenovirus vaccine of the present invention is a purified recombinant, replication-defective adenovirus which is shown to be genetically stable through multiple passages in cell culture and remains so during large scale production and purification procedures. Such a recombinant adenovirus vector and harvested adenovirus vaccine lends itself to large scale dose filling and subsequent worldwide distribution procedures which will be demanded of an efficacious monovalent or multivalent HIV vaccine. The present invention meets this basic requirement with description of a replication-defective adenoviral vector and vectors derived therefrom, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome. A preferred embodiment of the instant invention comprises base pairs 1-450 of a wildtype adenovirus. In other preferred embodiments, the replication -defective adenoviral vector has, in addition thereto, a region 3' to the E1-deleted region comprising base pairs 3511-3523. Basepairs 342-450 (more particularly, 400-450) constitute an extension of the 5' region of previously disclosed vectors carrying viral antigens, particularly HIV antigens (see, e.g., PCT International Application PCT/US00/18332, published

January 11, 2001 (WO 01/02067), which claims priority to U.S. Provisional Application Serial Nos. 60/142,631 and 60/148,981, filed 7/6/1999 and 8/13/1999, respectively; these documents herein incorporated by reference. Applicants have found that extending the 5' region further into the E1 gene into the disclosed vaccine
5 vectors incorporated elements found to be important in optimizing the packaging of the virus.

As compared to previous vectors not comprising basepairs from about 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome, vectors comprising the above region exhibited enhanced
10 growth characteristics, with approximately 5-10 fold greater amplification rates, a more potent virus effect, allowing lower doses of virus to be used to generate equivalent immunity; and a greater cellular-mediated immune response than replication-deficient vectors not comprising this region (basepairs 1-450). Even more important, adenoviral constructs derived therefrom are very stable genetically in
15 large-scale production, particularly those comprising an expression cassette under the control of a hCMV promoter devoid of intron A. This is because Applicants have surprisingly found that the intron A portion of the hCMV promoter constituted a region of instability when employed in adenoviral vectors. Applicants have, therefore, identified an enhanced adenoviral vector which is particularly suited for use
20 in gene therapy and nucleotide-based vaccine vectors which, favorably, lends itself to large scale propagation.

A preferred embodiment of this invention is a replication-defective adenoviral vector in accordance with the above description wherein the gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or
25 biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

In preferred embodiments, the E1 gene, other than that contained within basepairs 1-450 or, alternatively, that contained within base pairs 1-450 and 3511-
30 3523 has been deleted from the adenoviral vector, and the gene expression cassette has replaced the deleted E1 gene. In other preferred embodiments, the replication defective adenovirus genome does not have a functional E3 gene, or the E3 gene has been deleted. Most preferably, the E3 region is present within the adenoviral genome. Further preferred embodiments are wherein the gene expression cassette is in an E1
35 anti-parallel (transcribed in a 3' to 5' direction relative to the vector backbone)

orientation or, more preferably, an E1 parallel (transcribed in a 5' to 3' direction relative to the vector backbone) orientation.

Further embodiments relate to a shuttle plasmid vector comprising: an adenoviral portion and a plasmid portion, wherein said adenovirus portion comprises:

5 a) a replication defective adenovirus genome, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) of the wildtype adenovirus genome and, preferably, in addition thereto, basepairs 3511-3523 of a wildtype adenovirus sequence; and b) a gene

10 expression cassette comprising: (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and (c) a transcription terminator and/or a polyadenylation site.

Other aspects of this invention include a host cell comprising said adenoviral

15 vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors.

To this end, the present invention particularly relates to harvested

20 recombinant, replication defective virus derived from a host cell, such as but not limited to 293 cells or PER.C6[®] cells, including but not limited to harvested virus related to any of the MRKAd5 vector backbones, with or without an accompanying transgene, including but not limited to the HIV-1 antigens described herein. An HIV-1 vaccine is represented by any harvested, recombinant adenovirus material

25 which expresses any one or more of the HIV-1 antigens disclosed herein. This harvested material may then be purified, formulated and stored prior to host administration.

Another aspect of this invention is a method of generating a cellular immune response against a protein in an individual comprising administering to the individual

30 an adenovirus vaccine vector comprising:

a) a recombinant, replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting adenovirus packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) and, preferably in addition thereto,

35 base pairs 3511-3523 of a wildtype adenovirus sequence, and,

b) a gene expression cassette comprising: (i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a polyadenylation site.

5 In view of the efficacious nature of the adenoviral and/or DNA plasmid vaccines described herein, the present invention relates to all methodology regarding administration of one or more of these adenoviral and/or DNA plasmid vaccines to provide effective immunoprophylaxis, to prevent establishment of an HIV-1 infection following exposure to this virus, or as a post-HIV infection therapeutic vaccine to
10 mitigate the acute HIV-1 infection so as to result in the establishment of a lower virus load with beneficial long term consequences. As discussed herein, such a treatment regimen may include a monovalent or multivalent composition, various combined modality applications, and/or a prime/boost regimen to as to optimize antigen expression and a concomitant cellular-mediated and/or humoral immune response
15 upon inoculation into a living vertebrate tissue. Therefore, the present invention provides for methods of using the adenoviral and/or DNA plasmid vaccines disclosed herein within the various parameters disclosed herein as well as any additional parameters known in the art, which, upon introduction into mammalian tissue induces intracellular expression of the gag, pol and/or nef-based vaccines.

20 To this end, the present invention relates in part to methods of generating a cellular immune response in a vaccinee, preferably a human vaccinee, wherein the individual is given more than one administration of adenovirus vaccine vector, and it may be given in a regimen accompanied by the administration of a plasmid vaccine. The plasmid vaccine (also referred to herein as a "DNA plasmid vaccine" or "vaccine
25 plasmid" comprises a nucleic acid encoding a protein or an immunologically relevant portion thereof, a heterologous promoter operably linked to the nucleic acid sequence, and a transcription terminator or a polyadenylation signal (such as bGH or SPA, respectively). There may be a predetermined minimum amount of time separating the administrations. The individual can be given a first dose of plasmid vaccine, and then
30 a second dose of plasmid vaccine. Alternatively, the individual may be given a first dose of adenovirus vaccine, and then a second dose of adenovirus vaccine. In other embodiments, the plasmid vaccine is administered first, followed after a time by administration of the adenovirus vaccine. Conversely, the adenovirus vaccine may be administered first, followed by administration of plasmid vaccine after a time. In
35 these embodiments, an individual may be given multiple doses of the same adenovirus serotype in either viral vector or plasmid form, or the virus may be of

differing serotypes. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephalitis virus.

5 The present invention also relates to multivalent adenovirus vaccine compositions which comprise Gag, Pol and Nef components described herein; see, e.g., Example 29 and Table 25. Such compositions will provide for an enhanced cellular immune response subsequent to host administration, particularly given the genetic diversity of human MHCs and of circulating virus. Examples, but not
10 limitations, include MRKAd5-vector based multivalent vaccine compositions which provide for a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components) composition. Such a multivalent vaccine may be filled for a single dose or may consist of multiple inoculations of each individually filled component; and may in addition be part of a prime/boost regimen
15 with viral or non-viral vector vaccines as introduced in the previous paragraph. To this end, preferred compositions are MRKAd5 adenovirus used in combination with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of
20 such combined modalities vaccine formulation and administration increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

 The concept of a "combined modality" as disclosed herein also covers the alternative mode of administration whereby multiple HIV-1 viral antigens may be
25 ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example, a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possibly a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol; Example 29 and Table 25)
30 within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES). Therefore, a multivalent vaccine delivered as a single, or possibly a second
35 harvested recombinant, replication-deficient adenovirus is contemplated as part of the present invention.

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly if pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by administering the adenoviral vaccine. Multiple primings 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, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine. Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

It is an object of the present invention to provide for enhanced replication-defective recombinant adenoviral vaccine vector backbones. These recombinant adenoviral backbones may accept one or more transgenes, which may be passaged through cell culture for growth, amplification and harvest.

It is a further object to provide for enhanced replication-defective recombinant adenoviral vaccine vectors which encode various transgenes.

It is also an object of the present invention to provide for a harvested recombinant, replication-deficient adenovirus which shows enhanced growth and amplification rates while in combination with increased virus stability after continuous passage in cell culture. Such a recombinant adenovirus is particularly suited for use in gene therapy and nucleotide-based vaccine vectors which, favorably, lends itself to large scale propagation.

To this end, it is an object of the present invention to provide for (1) enhanced replication-defective recombinant adenoviral vaccine vectors as described herein which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef, and (2) harvested, purified recombinant replication-deficient adenovirus generated by passage of the adenoviral vectors of (1) through one or multiple passages through cell culture, including but not limited to passage through 293 cells or PER.C6[®] cells.

It is also an object of the present invention to provide for recombinant adenovirus harvested by one or multiple passages through cell culture. As relating to recombinant adenoviral vaccine vector, this recombinant virus is harvested and formulated for subsequent host administration.

5 It is also an object of the present invention to provide for replication-defective adenoviral vectors wherein at least one gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

10 It is also an object of the present invention to provide for a host cell comprising said adenoviral vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors.

15 It is a further object of the present invention to provide for methods of generating a cellular immune response against a protein in an individual comprising administering to the individual an adenovirus vaccine vector comprising a) a replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting packaging region from about base pair 1 to between from about base pair
20 342 (more preferably, 400) to about 450 (preferably, 1-450) and, preferably, 3511-3523 of a wildtype adenovirus sequence, and, b) a gene expression cassette comprising: (i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a
25 polyadenylation site.

It is also an object of the present invention to provide various alternatives for vaccine administration regimes, namely administration of one or more adenoviral and/or DNA plasmid vaccines described herein to provide effective immunoprophylaxis for uninfected individuals or a therapeutic treatment for HIV
30 infected patients. Such processes include but are not limited to multivalent HIV-1 vaccine compositions, various combined modality regimes as well as various prime/boost alternatives. These methods of administration, relating to vaccine composition and/or scheduled administration, will increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a
35 single modality regimen.

As used throughout the specification and claims, the following definitions and abbreviations are used:

"HAART" refers to -- highly active antiretroviral therapy --.

"first generation" vectors are characterized as being replication-defective.

5 They typically have a deleted or inactivated E1 gene region, and preferably have a deleted or inactivated E3 gene region as well.

"AEX" refers to Anion Exchange chromatography.

"QPA" refers to Quick PCR-based Potency Assay.

"bps" refers to basepairs.

10 "s" or "str" denotes that the transgene is in the E1 parallel or "straight" orientation.

"PBMCs" refers to peripheral blood monocyte cells.

"FL" refers to full length.

"FLgag" refers to a full-length optimized gag gene, as shown in Figure 2.

15 "Ad5-Flgag" refers to an adenovirus serotype 5 replication deficient virus which carries an expression cassette which comprises a full length optimized gag gene under the control of a CMV promoter.

"Promoter" means 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.

20 "Leader" means a DNA sequence at the 5' end of a structural gene which is transcribed along with the gene. This usually results a protein having an N-terminal peptide extension, often referred to as a pro-sequences.

25 "Intron" means a section of DNA occurring in the middle of a gene which does not code for an amino acid in the gene product. The precursor RNA of the intron is excised and is therefore not transcribed into mRNA not translated into protein.

"Immunologically relevant" or "biologically active" means (1) with regards to a viral protein, that the protein is capable, upon administration, of eliciting a measurable immune response within an individual sufficient to retard the propagation and/or spread of the virus and/or to reduce the viral load present within the individual; or (2) with regards to a nucleotide sequence, that the sequence is capable of encoding for a protein capable of the above.

30 "Cassette" refers to a nucleic acid sequence which is to be expressed, along with its transcription and translational control sequences. By changing the cassette, a vector can express a different sequence.

"bGHpA" refers to the bovine growth hormone transcription terminator/polyadenylation sequence.

"tPAgag" refers to a fusion between the leader sequence of the tissue plasminogen activator leader sequence and an optimized HIV gag gene, as exemplified in Figure 30A-B, whether in a DNA or adenovirus-based vaccine vector.

Where utilized, "IA" or "inact" refers to an inactivated version of a gene (e.g. IApol).

"MCS" is "multiple cloning site".

In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example:

"Ad5 HIV-1 gag", also referred to as the original HIV-1 gag adenoviral vector, is a vector containing a transgene cassette composed of a hCMV intron A promoter, the full length version of the human codon-optimized HIV-1 gag gene, and the bovine growth hormone polyadenylation signal. The transgene was inserted in the E1 antiparallel orientation in an E1 and E3 deleted adenovector.

"MRK Ad5 HIV-1 gag" also referred to as "MRKAd5gag" or "Ad5gag2" is an adenoviral vector taught herein which is deleted of E1, comprises basepairs 1-450 and 3511-3523, and has a human codon-optimized HIV-1 gene in an E1 parallel orientation under the control of a CMV promoter without intron A. The construct also comprises a bovine growth hormone polyadenylation signal.

"pV1JnsHIVgag", also referred to as "HIVFLgagPR9901", is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone.

"pV1JnsCMV(no intron)-FLgag-bGHpA" is a plasmid derived from pV1JnsHIVgag which is deleted of the intron A portion of CMV and which comprises the full length HIV gag gene. This plasmid is also referred to as "pV1JnsHIVgag-bGHpA", pV1Jns-hCMV-FL-gag-bGHpA" and "pV1JnsCMV(no intron) + FLgag + bGHpA".

"pV1JnsCMV(no intron)-FLgag-SPA" is a plasmid of the same composition as pV1JnsCMV(no intron)-FLgag-bGHpA except that the SPA termination sequence replaces that of bGHpA. This plasmid is also referred to as "pV1Jns-HIVgag-SPA" and pV1Jns-hCMV-FLgag-SPA".

"pdelE1sp1A" is a universal shuttle vector with no expression cassette (i.e., no promoter or polyA). The vector comprises wildtype adenovirus serotype 5 (Ad5) sequences from bp 1 to bp 341 and bp 3524 to bp 5798, and has a multiple cloning

site between the Ad5 sequences ending 341 bp and beginning 3524 bp. This plasmid is also referred to as the original Ad 5 shuttle vector.

"MRKpdelE1sp1A" or "MRKpdelE1(Pac/pIX/pack450)" or "MRKpdelE1(Pac/pIX/pack450)Cla1" is a universal shuttle vector with no expression cassette (i.e. no promoter or polyA) comprising wildtype adenovirus serotype 5 (Ad5) sequences from bp1 to bp450 and bp 3511 to bp 5798. The vector has a multiple cloning site between the Ad5 sequence ending 450 bp and beginning 3511 bp. This shuttle vector may be used to insert the CMV promoter and the bGHpA fragments in both the straight ("str". or E1 parallel) orientation or in the opposite (opp. or E1 antiparallel) orientation)

"MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" is still another shuttle vector which is the modified vector that contains the CMV promoter (no intronA) and the bGHpA fragments. The expression unit containing the hCMV promoter (no intron A) and the bovine growth hormone polyadenylation signal has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *Bgl*III site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1/E3+)Cla1 pre-plasmid. This shuttle vector, as shown in Figures 22 and 23, was used to insert the respective IAPol and G2A,LLAA nef genes directly into.

"MRKpdelE1-CMV(no intron)-FLgag-bGHpA" is a shuttle comprising Ad5 sequences from basepairs 1-450 and 3511-5798, with an expression cassette containing human CMV without intron A, the full-length human codon-optimized HIV gag gene and bovine growth hormone polyadenylation signal. This plasmid is also referred to as "MRKpdelE1 shuttle +hCMV-FL-gag-BGHPA"

"MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA" is an adenoviral vector comprising all Ad5 sequences except those nucleotides encompassing the E1 region (from 451-3510), a human CMV promoter without intron A, a full-length human codon-optimized HIV gag gene, and a bovine growth hormone polyadenylation signal. This vector is also referred to as "MRKpAdHVE3 + hCMV-FL-gag-BGHPA", "MRKpAd5HIV-1gag", "MRKpAd5gag", "pMRKAd5gag" or "pAd5gag2".

"pV1Jns-HIV-pol inact(opt)" or "pV1Jns-HIV IA pol (opt)" is the inactivated Pol gene (contained within SEQ ID NO:3) cloned into the *Bgl*III site of V1Jns (Figure 17A-C). As noted herein, various derivatives of HIV-1 pol may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdel+hCMVmin+FL-pol+bGHpA(s)" is the "MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" shuttle mentioned above which contains the LA pol gene in the proper orientation. This shuttle vector is used in a bacterial recombination with MRKpAd(E1-/E3+)Cla1.

5 "MRKpAd+hCMVmin+FL-pol+bGHpA(S)E3+", also referred to herein as "pMRKAd5pol", is the pre-adenovirus plasmid which comprises a CMV-pol inact(opt)-pGHpA construct. The construction of this pre-adenovirus plasmid is shown in Figure 22.

10 "pV1Jns/nef (G2A,LLAA)" or "V1Jns/opt nef (G2A,LLAA)" comprises codon optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175 (SEQ ID NO:13; which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662). This fragment is subcloned into the Bgl II site of V1Jns
15 and/or V1Jns-tPA (Figures 16A-B). As noted above for HIV-1 pol, HIV-1 nef constructs may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdelE1hCMVminFL-nefBGHpA(s)", also referred to herein as
20 "pMRKAd5nef", is the pre-adenovirus plasmid which comprises a CMV-nef (G2A,LLAA) codon optimized sequence. The construction of this pre-adenovirus plasmid is shown in Figure 23.

BRIEF DESCRIPTION OF THE FIGURES

25 Figure 1 shows the original HIV-1 gag adenovector (Ad5HIV-1gag). This vector is disclosed in PCT International Application No. PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S. Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S. Application Serial No. 60/148,981, filed August 13, 1999, all three applications which are hereby incorporated by reference.

30 Figure 2 shows the nucleic acid sequence (SEQ ID NO: 29) of the optimized human HIV-1 gag open reading frame.

Figure 3 shows diagrammatically the new transgene constructs in comparison with the original gag transgene.

35 Figure 4 shows the modifications made to the original adenovector backbone in the generation of the novel vectors of the instant invention.

Figure 5 shows the virus mixing experiments that were carried out to determine the effects of the addition made to the packaging signal region (Expt. #1) and the E3 gene on viral growth (Expt. #2). The bars denote the region of modifications made to the E1 deletion.

5 Figure 6 shows an autoradiograph of viral DNA analysis following the viral mixing experiments described in Examples 6 and 7.

Figures 7A, 7B and 7C are as follows: Figure 7A shows the hCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7B shows the hCMV-Flgag-SPA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Again, both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7C shows the mCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Once again, both E1 parallel and E1 antiparallel transgene orientation are represented.

Figure 8A shows the experiment designed to test the effect of transgene orientation.

Figure 8B shows the experiments designed to test the effect of polyadenylation signal.

20 Figure 9 shows viral DNA from the four adenoviral vectors tested (Example 12) at P5, following *Bst*E11 digestion.

Figure 10 shows viral DNA analysis of passages 11 and 12 of MRKpAdHVE3, MRKAd5HIV-1gag, and MRKAd5HIV-1gagE3-.

Figure 11 shows viral DNA analysis (*Hind*III digestion) of passage 6 MRKpAdHVE3 and MRKAd5HIV-1gag used to initiate the viral competition study. The last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI of 280 viral particles).

Figure 12 shows viral DNA analysis by *Hind* III digestion on high passage numbers for MRKAd5HIV-1gag in serum-containing media with collections made at specified times. The first lane shows the 1kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and *Hind*III), MRKAd5HIV-1gag at P16, P19, and P21.

Figure 13 shows serum anti-p24 levels at 3 wks post i.m. immunization of balb/c mice (n=10) with varying doses of several Adgag constructs: (A) MRK Ad5 HIV-1 gag (through passage 5); (B) MRKAd5 hCMV-FLgag-bGHpA (E3-); (C) MRKAd5 hCMV-FLgag-SPA (E3+); (D) MRKAd5 mCMV-FLgag-bGHpA (E3+);

(E) research lot (293 cell-derived) of Ad5HIV-1 gag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-1 gag. Reported are the geometric mean titers (GMT) for each cohort along with the standard error bars.

Figure 14 shows a restriction map of the pMRKAd5HIV-1gag vector.

5 Figures 15A-X illustrates the nucleotide sequence of the pMRKAd5HIV-1gag vector (SEQ ID NO:27.[coding] and SEQ ID NO:28 [non-coding]).

Figures 16A-B shows a schematic representation of DNA vaccine expression vectors V1Jns (A) and V1Jns-tPA (B), which are utilized for HIV-1 gag, pol and nef constructs in various DNA/viral vector combined modality regimens as disclosed
10 herein.

Figures 17A-C shows the nucleotide (SEQ ID NO:3) and amino acid sequence (SEQ ID NO:4) of IA-Pol. Underlined codons and amino acids denote mutations, as listed in Table 1.

Figure 18 shows codon optimized nucleotide and amino acid sequences
15 through the fusion junction of tPA-pol inact(opt) (contained within SEQ ID NOs: 7 and 8, respectively). The underlined portion represents the NH₂-terminal region of IA-Pol.

Figures 19A-B show a nucleotide sequence comparison between wild type nef(jrfl) and codon optimized nef. The wild type nef gene from the jrfl isolate
20 consists of 648 nucleotides capable of encoding a 216 amino acid polypeptide. WT, wild type sequence (SEQ ID NO:19); opt, codon-optimized sequence (contained within SEQ ID NO:1). The Nef amino acid sequence is shown in one-letter code (SEQ ID NO:2).

Figures 20A-C show nucleotide sequences at junctions between nef coding
25 sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as
30 underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine174
35 and 175 are the sites involved in myristylation and dileucine motif, respectively. For both versions of the tpanef fusion genes, the putative leader peptide cleavage sites are

indicated with “*”, and a exogenous serine residue introduced during the construction of the mutants is underlined.

Figure 22 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Pol.

5 Figure 23 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Nef.

Figure 24 shows a comparison of clade B vs. clade C anti-gag T cell responses in clade B HIV-infected subjects.

10 Figure 25 shows a comparison of clade B vs. clade C anti-nef T cell responses in clade B HIV-infected subjects.

Figures 26A-AO illustrates the nucleotide sequence of the pMRKAd5HIV-1pol adenoviral vector (SEQ ID NO:32 [coding] and SEQ ID NO:33 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO3).

15 Figures 27A-AM illustrates the nucleotide sequence of the pMRKAd5HIV-1 nef adenoviral vector (SEQ ID NO:34 [coding] and SEQ ID NO:35 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO13).

Figure 28 shows the stability of MRKAd5 vectors comprising various promoter fragments (hCMV or mCMV) and terminations signals (bGH or SPA) in E3(+) or E3(-) backbones.

20 Figures 29A and B shows the anion-exchange HPLC viral particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36, 48, and 60 hpi time points (Figure 29A) and the timcourse QPA supernatant titers (Figure 29B) for MRKAd5gag, MRKAd5pol and MRKAd5nef.

25 Figure 30 shows the nucleotide sequence (SEQ ID NO:36) and amino acid sequence (SEQ ID NO:37) comprising the open reading frame of a representative tPA-gag fusion for use in the DNA and/or adenoviral vaccine disclosed herein.

30 Figure 31 shows the intracellular γ IFN staining of PBMCs collected at week 10 (post DNA prime) and week 30 (post Ad boost). The cells were stimulated overnight in the presence or absence of the gag peptide pool. They were subsequently stained using fluorescence-tagged anti-CD3, anti-CD8, anti-CD4, and anti- γ IFN monoclonal antibodies. Each plot shows all CD3+ T cells which were segregated in terms of positive staining for surface CD8 and γ IFN production. The numbers in the upper right and lower right quadrants of each plot are the percentages of CD3+ cells that were CD8+ γ IFN+ and CD4+ γ IFN+, respectively.

35 Figure 32 shows a comparison of single-modality adenovirus immunization with DNA + adjuvant prime/adenovirus boost immunization.

Figures 33A-B show the nucleotide sequence (SEQ ID NO: 38) of the open reading frame for the gag-IAPol fusion of Example 29.

Figures 34A-B show the protein sequence (SEQ ID NO:39) of the gag-IAPol fusion frame.

5

DETAILED DESCRIPTION OF THE INVENTION

A novel replication-defective, or "first generation," adenoviral vector suitable for use in gene therapy or nucleotide-based vaccine vectors is described. This vector is at least partially deleted in E1 and comprises a wildtype adenovirus *cis*-acting packaging region from about base pair 1 to between about base pair 342 (more preferably, 400) to about 458 (preferably, 1-450) and, preferably, 3511-3523 of a wild-type adenovirus sequence. It has been found that a vector of this description possesses enhanced growth characteristics, with approximately 5-10 fold greater amplification rates, and is more potent allowing lower doses of virus to be used to generate equivalent immunity. The vector, furthermore, generates a harvested recombinant adenovirus which shows greater cellular-mediated immune responses than replication-deficient vectors not comprising this region (basepairs 342-450). Adenoviral constructs derived from these vectors are, further, very stable genetically, particularly those comprising a transgene under the control of a hCMV promoter devoid of intron A. Viruses in accordance with this description were passaged continually and analyzed; see Example 12. Each virus analyzed maintained its correct genetic structure. Analysis was also carried out under propagation conditions similar to that performed in large scale production. Again, the vectors were found to possess enhanced genetic stability; see Figure 12. Following 21 passages, the viral DNA showed no evidence of rearrangement, and was highly reproducible from one production lot to the next. The outcome of all relevant tests indicate that the adenoviral vector is extremely well suited for large-scale production of recombinant, replication-deficient adenovirus, as shown herein with the data associated with Figure 28.

A preferred adenoviral vector in accordance with this description is a vector comprising basepairs 1-450, which is deleted in E3. This vector can accommodate up to approximately 7,500 base pairs of foreign DNA inserts (or exogenous genetic material). Another preferred vector is one retaining E3 which comprises basepairs 1-450. A preferred vector of this description is an E3+ vector comprising basepairs 1-450 and 3511-3523. This vector, when deleted of the region spanning basepairs 451-3510, can accommodate up to approximately, 4,850 base pairs of foreign DNA inserts

(or exogenous genetic material). The cloning capacities of the above vectors have been determined using 105% of the wildtype Ad5 sequence as the upper genome size limit.

Wildtype adenovirus serotype 5 is used as the basis for the specific basepair numbers provided throughout the specification. The wildtype adenovirus serotype 5 sequence is known and described in the art; *see*, Chroboczek *et al.*, 1992 *J. Virology* 186:280, which is hereby incorporated by reference. Accordingly, a particular embodiment of the instant invention is a vector based on the adenovirus serotype 5 sequence. One of skill in the art can readily identify the above regions in other adenovirus serotypes (e.g., serotypes 2, 4, 6, 12, 16, 17, 24, 31, 33, and 42), regions defined by basepairs corresponding to the above basepair positions given for adenovirus serotype 5. Accordingly, the instant invention encompasses all adenoviral vectors partially deleted in E1 comprising basepairs corresponding to 1-450 (particularly, 342-450) and, preferably, 3511-3523 of a wild-type adenovirus serotype 5 (Ad5) nucleic acid sequence. Particularly preferred embodiments of the instant invention are those derived from adenoviruses like Ad5 which are classified in subgroup C (e.g., Ad2).

Vectors in accordance with the instant invention are at least partially deleted in E1. Preferably the E1 region is completely deleted or inactivated. Most preferably, the region deleted of E1 is within basepairs 451-3510. It is to be noted that the extended 5' and 3' regions of the disclosed vectors are believed to effectively reduce the size of the E1 deletion of previous constructs without overlapping any part of the E1A/E1B gene present in the cell line used, i.e., the PER.C6[®] cell line transfected with base pairs 459-3510. Overlap of adenoviral sequences is avoided because of the possibility of recombination. One of ordinary skill in the art can certainly appreciate that the instant invention can, therefore, be modified if a different cell line transfected with a different segment of adenovirus DNA is utilized. For purposes of exemplification, a 5' region of base pairs 1 to up to 449 is more appropriate if a cell line is transfected with adenoviral sequence from base pairs 450-3510. This holds true as well in the consideration of segments 3' to the E1 deletion.

Preferred embodiments of the instant invention possess an intact E3 region (i.e., an E3 gene capable of encoding a functional E3). Alternate embodiments have a partially deleted E3, an inactivated E3 region, or a sequence completely deleted of E3. Applicants have found, in accordance with the instant invention, that virus comprising the E3 gene were able to amplify more rapidly compared with virus not comprising an E3 gene; see Figure 6 wherein a diagnostic CsCl band corresponding to the E3+ virus

tested (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. These results were obtained following a virus competition study involving mixing equal MOI ratio (1:1) of adenovectors both comprising the E3 gene and not comprising the E3 gene. This increased amplification capacity of the E3+ adenovectors was subsequently confirmed with growth studies; see Table 4A, wherein the E3+ virus exhibit amplification ratios of 470, 420 and 320 as compared with the 115 and 40-50 of the E3- constructs.

As stated above, vectors in accordance with the instant invention can accommodate up to approximately 4,850 base pairs of exogenous genetic material for an E3+ vector and approximately 7,500 base pairs for an E3- vector. Preferably, the insert brings the adenoviral vector as close as possible to a wild-type genomic size (e.g., for Ad5, 35,935 basepairs). It is well known that adenovirus amplifies best when they are close to their wild-type genomic size.

The genetic material can be inserted in an E1-parallel or an E1 anti-parallel orientation, as such is illustrated in Figure 7A, 7B, 7C and Figure 8A. Particularly preferred embodiments of the instant invention, have the insert in an E1-parallel orientation. Applicants have found, via competition experiments with plasmids containing transgenes in differing orientation (Figure 8A), that vector constructs with the foreign DNA insert in an E1-parallel orientation amplify better and actually out-compete E1-antiparallel-oriented transgenes. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation as compared with the E1 anti-parallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested.

Adenoviral vectors in accordance with the instant invention are particularly well suited to effectuate expression of desired proteins, one example of which is an HIV protein, particularly an HIV full length gag protein. Exogenous genetic material encoding a protein of interest can exist in the form of an expression cassette. A gene expression cassette preferably comprises (a) a nucleic acid encoding a protein of interest; (b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and (c) a transcription terminator.

The transcriptional promoter is preferably 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), preferably without intronic sequences. Most preferred

for use within the instant adenoviral vector is a human CMV promoter without intronic sequences, like intron A. Applicants have found that intron A, a portion of the human cytomegalovirus promoter (hCMV), constitutes a region of instability for adenoviral vectors. CMV without intron A has been found to effectuate (Examples 1-3) comparable expression capabilities *in vitro* when driving HIV gag expression and, furthermore, behaved equivalently to intron A-containing constructs in Balb/c mice *in vivo* with respect to their antibody and T-cell responses at both dosages of plasmid DNA tested (20 µg and 200 µg). Those skilled in the art will appreciate that any of a number of other known promoters, such as the strong immunoglobulin, or other eukaryotic gene promoters may also 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.

In preferred embodiments, the promoter may also comprise a regulatable sequence such as the Tet operator sequence. This would be extremely useful, for example, in cases where the gene products are effecting a result other than that desired and repression is sought.

Preferred transcription termination sequences present within the gene expression cassette are the bovine growth hormone terminator/polyadenylation signal (bGHpA) and the short synthetic polyA signal (SPA) of 50 nucleotides in length, defined as follows: AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGT-TTTTGTGTG (SEQ ID NO:26).

The combination of the CMV promoter (devoid of the intron A region) with the BGH terminator is particularly preferred although other promoter/terminator combinations in the context of FG adenovirus may also be used.

Other embodiments incorporate a leader or signal peptide into the transgene. A preferred leader is that from the tissue-specific plasminogen activator protein, tPA. Examples include but are not limited to the various tPA-gag, tPA-pol and tPA-nef adenovirus-based vaccines disclosed throughout this specification.

In view of the improved adenovirus vectors described herein, an essential portion of the present invention are adenoviral-based HIV vaccines comprising said adenovirus backbones which may be administered to a mammalian host, preferably a human host, in either a prophylactic or therapeutic setting. The HIV vaccines of the present invention, whether administered alone or in combination regimens with other viral- or non-viral-based DNA vaccines, should elicit potent and broad cellular immune responses against HIV that will either lessen the likelihood of persistent virus infection and/or lead to the establishment of a clinically significant lowered virus load

subject to HIV infection or in combination with HAART therapy, mitigate the effects of previously established HIV infection (antiviral immunotherapy(ARI)). While any HIV antigen (e.g., gag, pol, nef, gp160, gp41, gp120, tat, rev, etc.) may be utilized in the herein described recombinant adenoviral vectors, preferred embodiments include
5 the codon optimized p55 gag antigen (herein exemplified as MRKAd5gag), pol and nef. Sequences based on different Clades of HIV-1 are suitable for use in the instant invention, most preferred of which are Clade B and Clade C. Particularly preferred embodiments are those sequences (especially, codon-optimized sequences) based on consensus Clade B sequences. Preferred versions of the MRKAd5pol and
10 MRKAd5nef series of adenoviral vaccines will encode modified versions of pol or nef, as discussed herein. Preferred embodiments of the MRKAd5HIV-1 vectors carrying HIV envelope genes and modifications thereof comprise the HIV codon-optimized *env* sequences of PCT International Applications PCT/US97/02294 and PCT/US97/10517, published August 28, 1997 (WO 97/31115) and December 24,
15 1997, respectively; both documents of which are hereby incorporated by reference.

A most preferred aspect of the instant invention is the disclosed use of the adenoviral vector described above to effectuate expression of HIV gag. Sequences for many genes of many HIV strains are publicly available in GENBANK and primary, field isolates of HIV are available from the National Institute of Allergy and
20 Infectious Diseases (NIAID) which has contracted with Quality Biological (Gaithersburg, MD) to make these strains available. Strains are also available from the World Health Organization (WHO), Geneva Switzerland. It is preferred that the gag gene be from an HIV-1 strain (CAM-1; Myers et al, eds. "Human Retroviruses and AIDS: 1995, IIA3-IIA19, which is hereby incorporated by reference). This gene
25 closely resembles the consensus amino acid sequence for the clade B (North American/European) sequence. Therefore, it is within the purview of the skilled artisan to choose an appropriate nucleotide sequence which encodes a specific HIV gag antigen, or immunologically relevant portion thereof. As shown in Example 25, a clade B or clade C based p55 gag antigen will potentially be useful on a global scale.
30 As noted herein, the transgene of choice for insertion in to a DNA or MRKAd-based adenoviral vector of the present invention is a codon optimized version of p55 gag. Such a MRKAd5gag adenoviral vector is documented in Example 11 and is at least referred to herein as MRKAd5HIV-1gag. Of course, additional versions are contemplated, including but not limited to modifications such as promoter (e.g.,
35 mCMV for hCMV) and/or pA-terminations signal (SPA for bGH) switching, as well as generating MRK Ad5 backbones with or without deletion of the Ad5 E3 gene.

The present invention also relates a series of MRKAd5pol-based adenoviral vaccines which are shown herein to generate cellular immune responses subsequent to administration in mice and non-human primate studies. Several of the MRKAd5pol series are exemplified herein. One such adenoviral vector is referred to as

5 MRKAd5hCMV-inact opt pol(E3+), which comprises the MRKAd5 backbone, the hCMV promoter (no intron A), an inactivated pol transgene, and contains the Ad5 E3 gene in the adenoviral backbone. A second exemplified pre-adenovirus plasmid and concomitant virus is referred to as MRKAd5hCMV-inact opt pol(E3-), which is identical to the former adenoviral vector except that the E3 is deleted. Both

10 constructions contain a codon optimized, inactivated version of HIV-1 Pol, wherein at least the entire coding region is disclosed herein as SEQ ID NO:3 and the expressed protein is shown as SEQ ID NO:4 (see also Figure 17A-C and Table 1, which show targeted deletion for inactivated pol. This and other preferred codon optimized versions of HIV Pol as disclosed herein are essentially as described in U.S.

15 Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference. As disclosed in the above-mentioned documents, the open reading frame for these codon-optimized HIV-1 Pol-based DNA vaccines are represented by codon optimized DNA molecules encoding codon

20 optimized HIV-1 Pol (e.g. SEQ ID NO:2), codon optimized HIV-1 Pol fused to an amino terminal localized leader sequence (e.g. SEQ ID NO:6), and especially preferable, and exemplified by the MRKAd5-Pol construct in e.g., Example 19, biologically inactivated pol ("inact opt Pol"; e.g., SEQ ID NO:4) which is devoid of significant PR, RT, RNase or IN activity associated with wild type Pol. In addition, a

25 construct related to SEQ ID NO:4 is contemplated which contains a leader peptide at the amino terminal region of the IA Pol protein. A specific construct is ligated within an appropriate DNA plasmid vector containing regulatory regions operatively linked to the respective HIV-1 Pol coding region, with or without a nucleotide sequence encoding a functional leader peptide. To this end, various HIV-1 Pol constructs

30 disclosed herein relate to open reading frames for cloning to the enhanced first generation Ad vectors of the present invention (such a series of MRKAd5pol adenoviral vaccine vectors), including but not limited to wild type Pol (comprising the DNA molecule encoding WT opt Pol, as set forth in SEQ ID NO:2), tPA-opt WTPol, (comprising the DNA molecule encoding tPA Pol, as set forth in SEQ ID NO:6), inact

35 opt Pol (comprising the DNA molecule encoding IA Pol, as set forth in SEQ ID NO:4), and tPA-inact opt Pol, (comprising the DNA molecule encoding tPA-inact opt

Pol, as set forth in SEQ ID NO:8). The pol-based versions of enhanced first generation adenovirus vaccines elicit CTL and Th cellular immune responses upon administration to the host, including primates and especially humans. As noted in the above, an effect of the cellular immune-directed vaccines of the present invention should be a lower transmission rate to previously uninfected individuals and/or reduction in the levels of the viral loads within an infected individual, so as to prolong the asymptomatic phase of HIV-1 infection.

The present invention further relates to a series of MRKAd5nef-based adenoviral vaccines which, similar to HIV gag and pol antigens, generate cellular immune responses subsequent to administration in mice and non-human primate studies. The MRKAd5nef series are exemplified herein by utilizing the improved MRK adenoviral backbone in combination with modified versions of HIV nef. These exemplified MRKAd5nef vectors are as follows: (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+), which comprises the improved MRKAd5 backbone, a human CMV promoter an intact Ad5 E3 gene and a modified nef gene: (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+), which is the same as (1) above but substituting a murine CMV promoter for a human CMV promoter; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+), which is the same as (2) except that the nef transgene is tpanef(LLAA). Codon optimized versions of HIV-1 Nef and HIV-1 Nef modifications are essentially as described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference. Particular embodiments of codon optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding regions for use in the adenoviral vectors of the present invention comprise a codon optimized DNA molecule encoding a protein containing the human plasminogen activator (tpa) leader peptide fused with the NH₂-terminus of the HIV-1 Nef polypeptide. The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:11, while the expressed open reading frame is disclosed herein as SEQ ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule encoding optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein

described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:13, while the expressed open reading frame is disclosed herein as SEQ ID NO:14. MRKAd5nef vectors (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) and (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) contain this transgene. An additional embodiment relates to a DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEQ ID NO:16. The MRKAd5nef vector "MRKAd5mCMV-tpanef(LLAA) (E3+)" contains this transgene.

Along with the improved MRKAd5gag adenovirus vaccine vector described herein, generation of a MRKAd5pol and MRKAd5nef adenovirus vector provide for enhanced HIV vaccine capabilities. Namely, the generation of this trio of adenoviral vaccine vectors, all shown to generate effective cellular immune responses subsequent to host administration, provide for the ability to administer these vaccine candidates not only alone, but preferably as part of a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components). Therefore, a preferred aspect of the present invention are vaccine formulations and associated methods of administration and concomitant generation of host cellular immune responses associated with formulating three separate series of MRKAd5-based adenoviral vector vaccines. Of course, this MRKAd5 vaccine series based on distinct HIV antigens promotes expanded opportunities for formulation of a divalent or trivalent vaccine, or possibly administration of separate formulations of one or more monovalent or divalent formulations within a reasonable window of time. It is also within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen. An example, but certainly not a limitation, would be separate MRKAd5pol vectors, with one vaccine vector expressing wild type Pol (SEQ ID NO:2) and another MRKAd5pol vector expressing inactivated Pol (SEQ ID NO:6). Another example might be separate MRKAd5nef vectors, with one vaccine vector expressing the tPA/LLAA version of Nef (SEQ ID NO:16) and another MRKAd5nef vector expressing the G2A,LLAA modified version of Nef (SEQ ID NO:14). Therefore, the MRKAd5 adenoviral vectors of the present invention may be used in combination

with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities vaccine formulation and administration
5 increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

The present invention also relates to application of a mono-, dual-, or tri-modality administration regime of the MRKAd5gag, pol and nef adenoviral vaccine series in a prime/boost vaccination schedule. This prime/boost schedule may include
10 any reasonable combination of the MRKAd5gag, pol and nef adenoviral vaccine series disclosed herein. In addition, a prime/boost regime may also involve other viral and/or non-viral DNA vaccines. A preferable addition to an adenoviral vaccine vector regime includes but is not limited to plasmid DNA vaccines, especially DNA plasmid vaccines that contain at least one of the codon optimized gag, pol and nef
15 constructions, as disclosed herein.

Therefore, one aspect of this invention is the administration of the adenoviral vector containing the optimized gag gene in a prime/boost regiment in conjunction with a plasmid DNA encoding gag. To distinguish this plasmid from the adenoviral-containing shuttle plasmids used in the construction of an adenovirus vector, this
20 plasmid will be referred to as a "vaccine plasmid" or "DNA plasmid vaccine". Preferred vaccine plasmids for use in this administration protocol are disclosed in pending U.S. patent application 09/017,981, filed February 3, 1998 and WO98/34640, published August 13, 1998, both of which are hereby incorporated by reference. Briefly, the preferred vaccine plasmid is designated V1Jns-FLgag, which expresses
25 the same codon-optimized gag gene as the adenoviral vectors of this invention (see Figure 2 for the nucleotide sequence of the exemplified optimized codon version of full length p55 gag). The vaccine plasmid backbone, designated V1Jns contains the CMV immediate-early (IE) promoter and intron A, a bovine growth hormone-derived polyadenylation and transcription termination sequence as the gene expression
30 regulatory elements, and a minimal pUC backbone; see Montgomery *et al.*, 1993, *DNA Cell Biol.* 12:777-783. The pUC sequence permits high levels of plasmid production in *E. coli* and has a neomycin resistance gene in place of an ampicillin resistance gene to provide selected growth in the presence of kanamycin.

Alternatively, a vaccine plasmid which has the CMV promoter deleted of intron A can
35 be used. Those of skill in the art will recognize that alternative vaccine plasmid

vectors may be easily substituted for these specific constructs, and this invention specifically envisions use of such alternative plasmid DNA vaccine vectors.

Another aspect of the present invention is a prime/boost regimen which includes a vaccine plasmid which encodes an HIV pol antigen, preferably a codon
5 optimized form of pol and also preferably a vaccine plasmid which comprises a nucleotide sequence which encodes a Pol antigen selected from the group of Pol antigens as shown in SEQ ID NOs: 2, 4, 6 and 8. The variety of potential DNA plasmid vaccines which encode various biologically active forms of HIV-1 Pol, wherein administration, intracellular delivery and expression of the HIV-1 Pol gene of
10 interest elicits a host CTL and Th response. The preferred synthetic DNA molecules of the present invention encode codon optimized wild type Pol (without Pro activity) and various codon optimized inactivated HIV-1 Pol proteins. The HIV-1 *pol* open reading disclosed herein are especially preferred for pharmaceutical uses, especially for human administration as delivered via a recombinant adenoviral vaccine,
15 especially an enhanced first generation recombinant adenoviral vaccine as described herein. Several embodiments of this portion of the invention are provided in detail below, namely DNA molecules which comprise a HIV-1 pol open reading frame, whether encoding full length pol or a modification or fusion as described herein, wherein the codon usage has been optimized for expression in a mammal, especially a
20 human. Again, these DNA sequences are positioned appropriately within a recombinant adenoviral vector, such as the exemplified recombinant adenoviral vector described herein, so as to promote expression of the respective HIV-1 Pol gene of interest, and subsequent to administration, elicit a host CTL and Th response. Again, these preferred, but in no way limiting, pol genes are as disclosed herein and
25 essentially as described in U.S. Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference.

A third series of vaccine plasmids which are useful in a combined modality and/or prime/boost regimen are vaccine plasmids which encode an HIV nef antigen or
30 biologically and/or immunologically relevant modification thereof. As noted elsewhere, preferred vaccine plasmids contain a codon optimized form of nef and also preferably comprise a nucleotide sequence which encodes a Nef antigen selected from the group of Nef antigens as shown in SEQ ID NOs: 10, 12, 14 and 16. These preferred nef coding regions are disclosed herein, as well as being described in U.S.
35 Application Serial No. 09/738,782, filed December 15, 2000 and PCT International

Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference.

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost
5 administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly if pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by
10 administering the adenoviral vaccine. Multiple primings 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, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine.
15 Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

Furthermore and in the alternative, multiple HIV-1 viral antigens, such as the
20 MRKAd5 adenoviral vaccines disclosed herein, may be ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possibly a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and
25 inactivated optimized pol; Example 29 and Table 25) within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES), as disclosed in
30 International Publication No. WO 95/24485, which is hereby incorporated by reference. Figure 9 shows that the use of multiple promoters and termination sequences provide for similar growth properties, while Figure 28 shows that these MRKAd5gag-based vectors are also stable at least through passage 21. In the absence of the use of IRES-based technology, it is preferred that a distinct promoter be used to
35 support each respective open reading frame, so as to best preserve vector stability. As examples, and certainly not as limitations, potential multiple transgene vaccines may

include a three transgene vector such as hCMV-gagpol-bGHpA + mCMV-nef-SPA in an E3 deleted backbone or hCMV-gagpol-bGHpA + mCMV-nef-SPA(E3+). Potential "2+1" divalent vaccines of the present invention might be a hCMV-gag-bGHpA + mCMV-nef-SPA in an E3+ backbone (vector #1) in combination with
5 hCMV-pol-bGHpA in an E3+ backbone (vector #2), with all transgenes in the E1 parallel orientation. Fusion constructs other than the gag-pol fusion described above are also suitable for use in various divalent vaccine strategies and can be composed of any two HIV antigens fused to one another (*e.g.*, nef-pol and gag-nef). These adenoviral compositions are, as above, preferably delivered along with an adenoviral
10 composition comprising an additional HIV antigen in order to diversify the immune response generated upon administration. Therefore, a multivalent vaccine delivered in a single, or possible second, adenoviral vector is certainly contemplated as part of the present invention. Again, this mode of administration is another example of whereby an efficacious adenovirus-based HIV-1 vaccine may be administered via a
15 combined modality regime. It is important to note, however, that in terms of deciding on an insert for the disclosed adenoviral vectors, due consideration must be dedicated to the effective packaging limitations of the adenovirus vehicle. Adenovirus has been shown to exhibit an upper cloning capacity limit of approximately 105% of the wildtype Ad5 sequence.

20 Regardless of the gene chosen for expression, it is preferred that the sequence be "optimized" for expression in a human cellular environment. A "triplet" codon of four possible nucleotide bases can exist in 64 variant forms. That these forms provide the message for only 20 different amino acids (as well as transcription initiation and termination) means that some amino acids can be coded for by more than one codon.
25 Indeed, some amino acids have as many as six "redundant", alternative codons while some others have a single, required codon. For reasons not completely understood, alternative codons are not at all uniformly present in the endogenous DNA of differing types of cells and there appears to exist variable natural hierarchy or "preference" for certain codons in certain types of cells. As one example, the amino
30 acid leucine is specified by any of six DNA codons including CTA, CTC, CTG, CTT, TTA, and TTG (which correspond, respectively, to the mRNA codons, CUA, CUC, CUG, CUU, UUA and UUG). Exhaustive analysis of genome codon frequencies for microorganisms has revealed endogenous DNA of *E. coli* most commonly contains the CTG leucine-specifying codon, while the DNA of yeasts and slime molds most
35 commonly includes a TTA leucine-specifying codon. In view of this hierarchy, it is generally held that the likelihood of obtaining high levels of expression of a leucine-

rich polypeptide by an *E. coli* host will depend to some extent on the frequency of codon use. For example, a gene rich in TTA codons will in all probability be poorly expressed in *E. coli*, whereas a CTG rich gene will probably highly express the polypeptide. Similarly, when yeast cells are the projected transformation host cells for expression of a leucine-rich polypeptide, a preferred codon for use in an inserted DNA would be TTA.

The implications of codon preference phenomena on recombinant DNA techniques are manifest, and the phenomenon may serve to explain many prior failures to achieve high expression levels of exogenous genes in successfully transformed host organisms--a less "preferred" codon may be repeatedly present in the inserted gene and the host cell machinery for expression may not operate as efficiently. This phenomenon suggests that synthetic genes which have been designed to include a projected host cell's preferred codons provide a preferred form of foreign genetic material for practice of recombinant DNA techniques. Thus, one aspect of this invention is an adenovirus vector or adenovirus vector in some combination with a vaccine plasmid where both specifically include a gene which is codon optimized for expression in a human cellular environment. As noted herein, a preferred gene for use in the instant invention is a codon-optimized HIV gene and, particularly, HIV gag, pol or nef.

Adenoviral vectors in accordance with the instant invention can be constructed using known techniques, such as those reviewed in Hitt et al, 1997 "Human Adenovirus Vectors for Gene Transfer into Mammalian Cells" *Advances in Pharmacology* 40:137-206, which is hereby incorporated by reference.

In constructing the adenoviral vectors of this invention, it is often convenient to insert them into a plasmid or shuttle vector. These techniques are known and described in Hitt et al., *supra*. This invention specifically includes both the adenovirus and the adenovirus when inserted into a shuttle plasmid.

Preferred shuttle vectors contain an adenoviral portion and a plasmid portion. The adenoviral portion is essentially the same as the adenovirus vector discussed *supra*, containing adenoviral sequences (with non-functional or deleted E1 and E3 regions) and the gene expression cassette, flanked by convenient restriction sites. The plasmid portion of the shuttle vector 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 shuttle vector 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
5 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.

In one embodiment of this invention, the pre-plasmids (e.g., pMRKAd5pol,
10 pMRKAd5nef and pMRKAd5gag were generated by homologous recombination using the MRKHVE3 (and MRKHVO for the E3- version) backbones and the appropriate shuttle vector, as shown for pMRKAd5pol in Figure 22 and for pMRKAd5nef in Figure 23. The plasmid in linear form is capable of replication after entering the PER.C6[®] cells and virus is produced. The infected cells and media were
15 harvested after viral replication was complete.

Viral vectors can be propagated in various 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) and issued U.S. Patent No. 6,033,908, both of which are hereby
20 incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient (FG) adenovirus, but is designed to prevent generation of replication competent adenovirus by homologous recombination. Cells of particular interest have been stably transformed with a transgene that encodes the AD5E1A and E1B gene, like PER.C6[®],
25 from 459 bp to 3510 bp inclusive. 293 cells are described in Graham et al., 1977 *J. Gen. Virol* 36:59-72, which is hereby incorporated by reference. As stated above, consideration must be given to the adenoviral sequences present in the complementing cell line used. It is important that the sequences not overlap with that present in the vector if the possibility of recombination is to be minimized.

30 It has been found that vectors generated in accordance with the above description are more effective in inducing an immune response and, thus, constitute very promising vaccine candidates. More particularly, it has been found that first generation adenoviral vectors in accordance with the above description carrying a codon-optimized HIV gag gene, regulated with a strong heterologous promoter can be
35 used as human anti-HIV vaccines, and are capable of inducing immune responses.

Standard techniques of molecular biology for preparing and purifying DNA constructs enable the preparation of the DNA immunogens of this invention.

A vaccine composition comprising an adenoviral vector in accordance with the instant invention may contain physiologically acceptable components, such as
5 buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. One preferred formulation has: 2.5-10 mM TRIS buffer, preferably about 5 mM TRIS buffer; 25-100 mM NaCl, preferably about 75 mM NaCl; 2.5-10% sucrose, preferably about 5% sucrose; 0.01 -2 mM MgCl₂; and 0.001%-0.01% polysorbate 80 (plant derived). The pH should range from about 7.0-9.0, preferably
10 about 8.0. One skilled in the art will appreciate that other conventional vaccine excipients may also be used it make the formulation. The preferred formulation contains 5mM TRIS, 75 mM NaCl, 5% sucrose, 1mM MgCl₂, 0.005% polysorbate 80 at pH 8.0 This has a pH and divalent cation composition which is near the optimum for Ad5 stability and minimizes the potential for adsorption of virus to a glass surface.
15 It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use.

The amount of adenoviral 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
20 product. In general, an immunologically or prophylactically effective dose of 1×10^7 to 1×10^{12} particles and preferably about 1×10^{10} to 1×10^{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. It is also
25 contemplated that booster vaccinations are to be provided. Following vaccination with HIV adenoviral vector, boosting with a subsequent HIV adenoviral vector and/or plasmid may be desirable. Parenteral administration, such as intravenous, intramuscular, subcutaneous or other means of administration of interleukin-12 protein, concurrently with or subsequent to parenteral introduction of the vaccine
30 compositions of this invention is also advantageous.

The adenoviral vector and/or vaccine plasmids of this invention polynucleotide may be unassociated with any proteins, adjuvants or other agents which impact on the recipients' immune system. In this case, it is desirable for the vector to be in a physiologically acceptable solution, such as, but not limited to, sterile
35 saline or sterile buffered saline. Alternatively, the vector may be associated with an adjuvant known in the art to boost immune responses (i.e., a "biologically effective"

adjuvant), such as a protein or other carrier. Vaccine plasmids of this invention may, for instance, be delivered in saline (e.g., PBS) with or without an adjuvant. Preferred adjuvants are Alum or CRL1005 Block Copolymer. Agents which assist in the cellular uptake of DNA, such as, but not limited to, calcium ions, may also be used to advantage. These agents are generally referred to herein as transfection facilitating reagents and pharmaceutically acceptable carriers. Techniques for coating microprojectiles coated with polynucleotide are known in the art and are also useful in connection with this invention.

This invention also includes a prime and boost regimen wherein a first adenoviral vector is administered, then a booster dose is given. The booster dose may be repeated at selected time intervals. Alternatively, a preferred inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype. More preferably, the inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype, wherein the first and second adenovirus serotypes are classified within separate subgroups of adenoviruses. The above prime/boost schemes are particularly preferred in those situations where a preexisting immunity is identified to the adenoviral vector of choice. In this type of scheme, the individual or population of individuals is primed with an adenovirus of a serotype other than that to which the preexisting immunity is identified. This enables the first adenovirus to effectuate sufficient expression of the transgene while evading existing immunity to the second adenovirus (the boosting adenovirus) and, further, allows for the subsequent delivery of the transgene via the boosting adenovirus to be more effective. Adenovirus serotype 5 is one example of a virus to which such a scheme might be desirable. In accordance with this invention, therefore, one might decide to prime with a non-group C adenovirus (e.g., Ad12, a group A adenovirus, Ad24, a group D adenovirus, or Ad35, a group B adenovirus) to evade anti-Ad5 immunity and then boost with Ad5, a group C adenovirus. Another preferred embodiment involves administration of a different adenovirus (including non-human adenovirus) vaccine followed by administration of the adenoviral vaccines disclosed. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephalitis virus.

A large body of human and animal data supports the importance of cellular immune responses, especially CTL in controlling (or eliminating) HIV infection. In humans, very high levels of CTL develop following primary infection and correlate

with the control of viremia. Several small groups of individuals have been described who are repeatedly exposed to HIV by remain uninfected; CTL has been noted in several of these cohorts. In the SIV model of HIV infection, CTL similarly develops following primary infection, and it has been demonstrated that addition of anti-CD8 monoclonal antibody abrogated this control of infection and leads to disease progression. This invention uses adenoviral vaccines alone or in combination with plasmid vaccines to induce CTL.

The following non-limiting Examples are presented to better illustrate the invention.

EXAMPLE 1

Removal of the Intron A Portion of the hCMV Promoter

GMP grade pVII_{ns}HIVgag was used as the starting material to amplify the hCMV promoter. PVI_{ns}HIVgag is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone; see Montgomery *et al.*, *supra* for a description of the plasmid backbone. The amplification was performed with primers suitably positioned to flank the hCMV promoter. A 5' primer was placed upstream of the *MscI* site of the hCMV promoter and a 3' primer (designed to contain the *Bgl*III recognition sequence) was placed 3' of the hCMV promoter. The resulting PCR product (using high fidelity *Taq* polymerase) which encompassed the entire hCMV promoter (minus intron A) was cloned into TOPO PCR blunt vector and then removed by double digestion with *MscI* and *Bgl*III. This fragment was then cloned back into the original GMP grade pV1_{ns}HIVgag plasmid from which the original promoter, intron A, and the gag gene were removed following *MscI* and *Bgl*III digestion. This ligation reaction resulted in the construction of a hCMV promoter (minus intron A) + bGHpA expression cassette within the original pV1_{ns}HIVgag vector backbone. This vector is designated pVII_{ns}CMV(no intron).

The FLgag gene was excised from pV1_{ns}HIVgag using *Bgl*III digestion and the 1,526 bp gene was gel purified and cloned into pV1_{ns}CMV(no intron) at the *Bgl*III site. Colonies were screened using *Sma*I restriction enzymes to identify clones that carried the FLgag gene in the correct orientation. This plasmid, designated pV1_{ns}CMV(no intron)-FLgag-bGHpA, was fully sequenced to confirm sequence integrity.

Two additional transgenes were also constructed. The plasmid, pV1JnsCMV(no intron)-FLgag-SPA, is identical to pV1JnsCMV(no intron)-FLgag-bGHpA except that the bovine growth hormone polyadenylation signal has been replaced with a short synthetic polyA signal (SPA) of 50 nucleotides in length. The sequence of the SPA is as shown, with the essential components (poly(A) site, (GT)_n, and (T)_n; respectively) underlined:

AATAAAAGATCTTTATTTTCATTAGATCTGTGTG TTGGTTTTTTGTGTG
(SEQ ID NO:18).

The plasmid, pV1Jns-mCMV-FLgag-bGHpA, is identical to the pV1JnsCMV(no intron)-FLgag-bGHpA except that the hCMV promoter has been removed and replaced with the murine CMV (mCMV) promoter.

Figure 3 diagrammatically shows the new transgene constructs in comparison with the original transgene.

15 EXAMPLE 2

Gag Expression Assay for Modified Gag Transgenes

Gag Elisa was performed on culture supernatants obtained from transient tissue culture transfection experiments in which the two new hCMV-containing plasmid constructs, pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no intron)-FLgag-SPA, both devoid of intron A, were compared to pV1JnsHIVgag which, as noted above possesses the intron A as part of the hCMV promoter. Table 2 below shows the *in vitro* gag expression data of the new gag plasmids compared with the GMP grade original plasmid. The results displayed in Table 2 show that both of the new hCMV gag plasmid constructs have expression capacities comparable to the original plasmid construct which contains the intron A portion of the hCMV promoter.

Table 2: *In vitro* DNA transfection of original and new plasmid HIV-1 gag constructs.

Plasmid	μg gag/10e6 COS cells/5 μg DNA/48 hr
HIVFL-gagPR9901 ^a	10.8
PV1Jns-hCMV-FLgag-bGHpA ^b	16.6
pV1Jns-hCMV-FLgag-SPA ^{b,c}	12.0

^a GMP grade pV1Jns-hCMVintronA-FLgag-bGHpA.

5 ^b New plasmid constructions that have the intron A portion removed from the hCMV promoter.

^c In this construct the bGH terminator has been replaced with the short synthetic polyadenylation signal (SPA)

10

EXAMPLE 3

Rodent (Balb/c) Study for Modified gag Transgenes

A rodent study was performed on the two new plasmid constructs described above – pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no intron)-FLgag-SPA - in order to compare them with the construct described above possessing the intron A portion of the CMV promoter, pV1JnsHIVgag. Gag antibody and Elispot responses (described in PCT International Application No. 15 PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S. Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S. Application Serial No. 60/148,981, filed August 13, 1999, all three applications which 20 are hereby incorporated by reference) were measured. The results displayed in Table 3 below, show that the new plasmid constructs behaved equivalently to the original construct in Balb/c mice with respect to their antibody and T-cell responses at both dosages of plasmid DNA tested, 20 μg and 200 μg .

EXAMPLE 4

Table 3: HIV191: Immunogenicity of V1Jns-gag under different promoter and termination control elements.

DNA ^a Promoter/terminator	Dose, ug ^b	Anti-p24 Titers (3 Wk PD1) ^c			SFC/10 ⁶ Cells (4 Wk PD1) ^d		
		GMT	+SE	-SE	Media	gag197-205	p24
HIVFL-gagPR9901 (GMP grade)	200	12800	4652	3412	2(2)	129(19)	30(11)
	20	5572	1574	1227	0	56(9)	25(6)
pV1Jns-hCMV- FL-gag-bGHpA	200	11143	2831	2257	0	98(5)	12(6)
	20	7352	2808	2032	0	73(9)	11(6)
pV1Jns-hCMV- FL-gag-SPA	200	16890	5815	4326	1(1)	94(4)	26(7)
	20	5971	5361	2825	0	85(17)	38(10)
Naïve	0	123	50	36	0	0	0

^ain PBS^bi.m. Injections into both quads, 50 µL per quad^cn=10; GMT, geometric mean titer; SE, standard. error^dn=5, pooled spleens; mean of triplicate wells and standard. deviation. in parentheses;

Construction of the Modified Shuttle Vector -“MRKpdelE1 Shuttle”

- The modifications to the original Ad5 shuttle vector (pdelE1sp1A; a vector comprising Ad5 sequences from basepairs 1-341 and 3524-5798, with a multiple cloning region between nucleotides 341 and 3524 of Ad5, included the following three manipulations carried out in sequential cloning steps as follows:
- (1) The left ITR region was extended to include the *Pac1* site at the junction between the vector backbone and the adenovirus left ITR sequences. This allow for easier manipulations using the bacterial homologous recombination system.
 - (2) The packaging region was extended to include sequences of the wild-type (WT) adenovirus from 342 bp to 450 bp inclusive.
 - (3) The area downstream of pIX was extended 13 nucleotides (i.e., nucleotides 3511-3523 inclusive).

These modifications (Figure 4) effectively reduced the size of the E1 deletion without overlapping with any part of the E1A/E1B gene present in the transformed PER.C6[®] cell line. All manipulations were performed by modifying the Ad shuttle vector pdelE1sp1A.

Once the modifications were made to the shuttle vector, the changes were incorporated into the original Ad5 adenovector backbones (pAdHVO and pAdHVE3) by bacterial homologous recombination using *E. coli* BJ5183 chemically competent cells.

EXAMPLE 5

Construction of Modified Adenovector Backbones (E3+ and E3-)

The original adenovectors pAdHVO (comprising all Ad5 sequences except those nucleotides encompassing the E1 and E3 regions) and pAdHVE3 (comprising all Ad5 sequences except those nucleotides encompassing the E1 region), were each reconstructed so that they contained the modifications to the E1 region. This was accomplished by digesting the newly modified shuttle vector (MRKpdeIE1 shuttle) with *Pac*I and *Bst*Z1101 and isolating the 2,734 bp fragment which corresponds to the adenovirus sequence. This fragment was co-transformed with DNA from either *Cla*I linearized pAdHVO (E3- adenovector) or *Cla*I linearized pAdHVE3 (E3+adenovector) into *E. coli* BJ5183 competent cells. At least two colonies from each transformation were selected and grown in Terrific™ broth for 6-8 hours until turbidity was reached. DNA was extracted from each cell pellet and then transformed into *E. coli* XL1 competent cells. One colony from each transformation was selected and grown for plasmid DNA purification. The plasmid was analyzed by restriction digestions to identify correct clones. The modified adenovectors were designated MRKpAdHVO (E3- plasmid) and MRKpAdHVE3 (E3+ plasmid). Virus from these new adenovectors (MRKHVO and MRKHVE3, respectively) as well as the old version of the adenovectors were generated in the PER.C6® cell lines to accommodate the following series of viral competition experiments. In addition, the multiple cloning site of the original shuttle vector contained *Cla*I, *Bam*HI, *Xho* I, *Eco*RV, *Hind*III, *Sal* I, and *Bgl* II sites. This MCS was replaced with a new MCS containing *Not* I, *Cla* I, *Eco*RV and *Asc* I sites. This new MCS has been transferred to the MRKpAdHVO and MRKpAdHVE3 pre-plasmids along with the modification made to the packaging region and pIX gene.

EXAMPLE 6

Analysis of the Effect of the Packaging Signal Extension

To study the effects of the modifications made to the E1 deletion region, the viruses obtained from the original backbone (pAdHVE3) and the new backbone (MRKpAdHVE3) were mixed together in equal MOI ratios (1:1 and 5:5) and passaged through several rounds; see Figure 5, Expt.#1. Both of the viruses in the experiment contained the E3 gene intact and did not contain a transgene. The only difference between the two viruses was within the region of the E1 deletion. Following the coinfection of the viruses at P1 (passage 1), the mixtures were propagated through an additional 4 passages at which time the cells were harvested

and the virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with *HindIII* and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids (pAdHVE3 ("OLD E3+"); MRKpAdHVE3 ("NEW E3+")) were also digested with *HindIII* (and *PacI* to remove the vector backbone) and subsequently labeled with [³³P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 clearly shows that the new adenovirus which has the addition made to the packaging signal region has a growth advantage compared with the original adenovirus. In the experiments performed (at either ratio tested), only the digestion bands pertaining to the newly modified virus were present. The diagnostic band of size 3,206 (from the new virus) was clearly present. However, there was no evidence of the diagnostic band of size 2,737 bp expected from the original virus.

EXAMPLE 7

Analysis of the Effect of the E3 Gene

The second set of the virus competition study involved mixing equal MOI ratio (1:1) of the newly modified viruses, that obtained from MRKpAdHVO and MRKpAdHVE3 (Figure 5, Expt. #2). In this set, both viruses had the new modifications made to the E1 deletion. The first virus (that from MRKpAdHVO) does not contain an E3 gene. The second virus (that from MRKpAdHVE3) does contain the E3 gene. Neither of the viruses contain a transgene. Following co-infection of the viruses, the mixtures were propagated through an additional 4 passages at which time the cells were harvested and the total virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with *HindIII* and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids MRKpAdHVO ("NEW E3-"); MRKpAdHVE3 ("NEW E3+") were also digested with *HindIII* (and *PacI* to remove the vector backbone) and then labeled with [³³P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 shows the results of the viral DNA analysis of the E3+ virus and E3- virus mixing experiment. The diagnostic band corresponding to the E3+ virus (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. This indicates that the virus that contains the E3 gene is able to amplify more rapidly

EXAMPLE 8

The modified plasmid pV1JnsCMV(no intron)-FLgag-bGHpA was digested with *MscI* overnight and then digested with *SfiI* for 2 hours at 50°C. The DNA was then treated with Mungbean nuclease for 30 mins at 30°C. The DNA mixture was desalted using the Qiaex II kit and then Klenow treated for 30 mins at 37°C to fully blunt the ends of the transgene fragment. The 2,559 bp transgene fragment was then gel purified. The modified shuttle vector (MRKpdeIE1 shuttle) was linearized by digestion with *EcoRV*, treated with calf intestinal phosphatase and the resulting 6,479 bp fragment was then gel purified. The two purified fragments were then ligated together and several dozen clones were screened to check for insertion of the transgene within the shuttle vector. Diagnostic restriction digestion was performed to identify those clones carrying the transgene in the E1 parallel and E1 anti-parallel orientation. This strategy was followed to clone in the other gag transgenes in the MRKpdeIE1 shuttle vector.

Construction of the MRK FG Adenovectors

43

within the gag gene as well as the plasmid backbone. The pre-plasmid clone is designated MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA and is 37,498 bp in size. This strategy was followed to generate E3- and E3+ versions of each of the other gag transgene constructions in both E1 parallel and E1 anti-parallel versions. Figures 7A, 7B and 7C show the various combinations of adenovectors constructed.

EXAMPLE 10

Plasmid Competition Studies

A series of plasmid competition studies was carried out. Briefly, the screening of the various combinations of new constructs was performed by mixing equal amounts of each of two competing plasmids. In the experiment shown in Figure 8A, plasmids containing the same transgene but in different orientations were mixed together to create a "competition" between the two plasmids. The aim was to look at the effects of transgene orientation. In the experiment shown in Figure 8B, plasmids containing different polyadenylation signals (but in the same orientation) were mixed together in equal amounts. The aim was to assess effects of polyA signals. Following the initial transfection, the virus was passaged through ten rounds and the viral DNA analyzed by radioactive restriction analysis.

Analysis of the viral species from the plasmid mixing experiment (Figure 8A) showed that adenovectors which had the transgene inserted in the E1 parallel orientation amplified better and were able to out-compete the adenovirus which had the transgene inserted in the E1 anti-parallel orientation. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation compared with the E1 antiparallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested (hCMV(no intron)-FLgag-bGHpA and hCMV(no intron)-FLgag-SPA).

Analysis of the viral species from the plasmid mixing experiment #2 (Figure 8B) at passages 3 and 6 showed that the polyadenylation signals tested (bGHpA and SPA) did not have an effect on the growth of the virus. Even at passage 10 the two viral species in the mixture were still present in equal amounts.

EXAMPLE 11

Virus generation of an enhanced adenoviral construct – “MRK Ad5 HIV-1 gag”

The results obtained from the competition study allowed us to make the following conclusions: (1) The packaging signal extension is beneficial; (2) Presence of E3 does enhance viral growth; (3) E1 parallel orientation is recommended; and (4) PolyA signals have no effect on the growth of the adenovirus.

MRK Ad5 HIV-1 gag exhibited the most desirable results. This construct contains the hCMV(no intron)-FLgag-bGHpA transgene inserted into the new E3+ adenovector backbone, MRKpAdHVE3, in the E1 parallel orientation. We have designated this adenovector MRK Ad5 HIV-1 gag. This construct was prepared as outlined below:

The pre-plasmid MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA was digested with *Pac1* to release the vector backbone and 3.3 µg was transfected by calcium phosphate method (Amersham Pharmacia Biotech.) in a 6 cm dish containing PER.C6[®] cells at ~60% confluence. Once CPE was reached (7-10 days), the culture was freeze/thawed three times and the cell debris pelleted. 1 ml of this cell lysate was used to infect into a 6 cm dish containing PER.C6[®] cells at 80-90% confluence. Once CPE was reached, the culture was freeze/thawed three times and the cell debris pelleted. The cell lysate was then used to infect a 15 cm dish containing PER.C6[®] cells at 80-90% confluence. This infection procedure was continued and expanded at passage 6. The virus was then extracted from the cell pellet by CsCl method. Two bandings were performed (3-gradient CsCl followed by a continuous CsCl gradient). Following the second banding, the virus was dialyzed in A105 buffer. Viral DNA was extracted using pronase treatment followed by phenol chloroform. The viral DNA was then digested with *HindIII* and radioactively labeled with [³³P]dATP. Following gel electrophoresis to separate the digestion products the gel was dried down on Whatman paper and then subjected to autoradiography. The digestion products were compared with the digestion products from the pre-plasmid (that had been digested with *Pac1/HindIII* prior to labeling). The expected sizes were observed, indicating that the virus had been successfully rescued. This strategy was used to rescue virus from each of the various adenovector plasmid constructs prepared.

EXAMPLE 12

Stability Analyses

To determine whether the various adenovector constructs (e.g., MRK Ad5 HIV-1 gag) show genetic stability, the viruses were each passaged continually. The viral DNA was analyzed at passages 3, 6 and 10. Each virus maintained its correct genetic structure. In addition, the stability of the MRK Ad5 HIV-1 gag was analyzed under propagation conditions similar to that performed in large scale production. For this analysis, the transfections of MRK Ad5 HIV-1 gag as well as three other adenoviral vectors were repeated and the virus was purified at P3. The three other adenovectors were as follows: (1) that comprising hCMV(no intron)-Flgag with a bGHpA terminator in an E3- adenovector backbone; (2) that comprising hCMV(no intron)-Flgag with a SPA termination signal in an E3+ adenovector backbone, and that comprising a mCMV-Flgag with a bGHpA terminator in an E3+ adenovector backbone. All of the vectors have the transgene inserted in the E1 parallel orientation. Viral DNA was analyzed by radioactive restriction analysis to confirm that it was correct before being delivered to fermentation cell culture for continued passaging in serum-free media. At P5 each of the four viruses were purified and the viral DNA extracted for analysis by the restriction digestion and radiolabeling procedure. This virus has subsequently been used in a series of studies (*in vitro* gag expression in COS cells, rodent study and rhesus monkey study) as will be described below. The viruses from P5 are shown in Figure 9.

The passaging under serum-free conditions was continued for the MRKHVE3 (transgene-less, obtained from MRKpAdHVE3 pre-plasmid) and the MRKAd5HIV-1gag (obtained from MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA pre-plasmid) viruses. Figure 10 shows viral DNA analysis by radioactive restriction digestion at passage 11 for MRKHVE3, MRKAd5HIV-1gagE3-, and passage 11 and 12 for MRKAd5HIV-1gag. Aside from the first lane which is the DNA marker lane, the next three lanes are virus from the pre-plasmid controls (controls based on the original virus) - MRKpAdHVE3 (also referred to as "pMRKHVE3"), MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA, and pMRKAd5gag(E3-), respectively. As seen in Figure 10, each of the viral DNA samples show the expected bands with no extraneous bands showing. This signifies that there are no major variant adenovirus species present that can be detected by autoradiography.

Figure 11 shows the results of viral competition study between MRKHVE3 and MRKAd5HIV-1gag. These viruses were mixed together at equal MOI (140 viral

particles each; 280 vp total) at passage 6 and continued to be passaged until P11.

Aside from the first lane which is the DNA marker lane, the next two lanes are the pre-plasmid controls obtained from MRKpAdHVE3 and MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA. The next two lanes are the viral DNA from the starting viral material at passage six. The last two lanes are the competition studies performed in duplicate. The data in Figure 11 shows the effect the gag transgene in culture.

Growth of a MRKAd5gag virus was compared with growth of a "transgene-less" MRKHVE3. These two viruses were infected at the same MOI (i.e. 140 vp each) at passage 6 and then passaged through to passage 11 and the viral pool was analyzed by radioactive restriction analysis. The data shows that one virus did not out compete the other. Therefore, the gag transgene did not show obvious signs of toxicity to the adenovirus.

Analysis by *Hind*III digestion shows that each virus specie is present in approximately equal amounts. As above, there does not appear to be signs of any extraneous bands. Figure 12 shows higher passage numbers for MRKAd5HIV-1gag grown under serum-containing conditions. The genome integrity again has been maintained and there is no evidence of rearrangements, even at the highest passage level (P21).

Each of the four vectors shown in Figure 9 were analyzed for amplification capacity. Table 4 below shows the QPA analysis used in the estimation of viral amplification ratios at P4. The determination of the amplification ratio for the original HIV-1 gag construct is based on the clinical lot at P12. It has been shown that amplification rates increases with higher passage number for the original virus. The reason for this observation is due to the emergence of variants which exhibit increased growth rates compared to the intact adenovector. With continued passaging of the original Ad gag vector, the level of variants increases and hence amplification rates increase also.

The MRK Ad5 HIV-1 gag virus has also been continually passaged under process conditions (i.e., serum-free media). Viral DNA extracted from passages 11 and 12 show no evidence of rearrangement.

Table 4:
Amplification Ratios Based on AEX and QPA Analysis of
Virus Amplification from Passage 3 to Passage 4.

Ad gag construct	Amplification Ratio
MRKAd5gag	470
HCMV-Flgag-bGHpA [E3-]	115
HCMV-Flgag-SPA [E3+]	320
mCMV-FLgag-bGHpA [E3+]	420
Original construct *	40 - 50

5

* This estimation is based on the clinical lot growth characteristics at Passage 12.

EXAMPLE 13

10

Analytical Evaluation of the enhanced Ad5 Constructs

15

To study the effects of the transgene and the E3 gene on virus amplification, the enhanced adenoviral vector, MRK Ad5 HIV-1 gag, along with its transgene-less version (MRKpAdHVE3) and its E3- version (MRK Ad5 HIV-1 gag E3-), was studied for several passages under serum-free conditions. Table 5A shows the amplification ratios determined for passages P3 to P8 for MRK Ad5 HIV-1 gag. Within a certain MOI range, it has been determined that the virus output is directly proportional to the virus input. Therefore, the greater the number of virus particles per cell at infection, the greater the virus amount produced. Viral amplification ratios, on the other hand, are inversely proportional to the virus input. The lower the virus input, the greater the amplification ratio.

20

25

Table 5B shows the amplification rates of the new E3+ vector backbone MRKpAdHVE3. It has a significantly lower rate of amplification compared with the gag transgene containing version. This may be contributed to the larger size MRK Ad5 HIV-1 gag since it contains the transgene. This inclusion of the transgene brings the size of the adenovirus closer to the size of a wild type Ad5 virus. It is well known that adenoviruses amplify best when they are at close to their wild type genomic size.

Wild type Ad5 is 35,935 bp. The MRKpAdHVE3 is 32,905 bp in length. The enhanced adenovector MRK Ad5 HIV-1 gag is 35,453bp (See Figure 14 for vector map; see also Figure 15A-X show the complete pre-adenoviral vector sequence, which includes an additional 2,021 bp of the vector backbone).

- 5 Table 5C shows the amplification rates of the new E3- gag containing virus MRK Ad5 HIV-1 gag E3-. Once again, this virus shows lower growth rate than the enhanced adenoviral vector. This may be attributed to the decreased sized of this virus (due to the E3 gene deletion) compared with wild type Ad5. The MRK Ad5 HIV-1 gag E3- virus is 32,810 bp in length. This can be compared with the wild type
- 10 Ad5 which is 35,935 bp and MRK Ad5 HIV-1 gag which is 35,453 bp in length.

Table 5A: Amplification ratios determined by AEX and QPA for MRKAd5gag over several continuous passaging in serum free media. Following P5, two replicate samples were taken (rep-1 and rep-2) and analyzed.

MRKAd5gag rep1

	Xv (10 ⁶ cells/ml), Infection	Viability (%), Harvest	Harvest Time h.p.i.	Cell Passage Number	Titer 10 ¹⁰ vp/ml culture	Titer 10 ⁴ vp/cell	QPA 10 ⁶ TCID ₅₀ /ml	Ratio AEX:QPA	Amplification Ratio	AEX Internal Control
P4	1.49, 81%	0.58, 50%	44	46	8.7	5.9	1.72	50	470 (MOI = 125)	
P5	1.38, 93%	0.66, 47%	48	49	6.7	4.9	1.38	49	170	
P6	1.04, 94%	0.68, 77%	47	48	5.8	5.6	1.42	41	200	
P7	1.50, 84%	0.96, 61%	49.5	50	3.9	1.4	0.97	40	50	
P7	1.09, 97%	0.76, 69%	50	52	5.2	4.7	1.70	31	170	
P8	1.03, 94%	0.86, 64%	47.5	54	9.0	8.7	1.10	82	310	
P9	0.89, 95%	0.99, 73%	47.5	56	4.4	4.9	1.03	43	175	3.12 2.84
P10	1.09, 91%	1.06, 66%	47.5	58	3.0	2.8	1.16	26	100	2.70 2.60
P11	1.19, 88%	0.98, 65%	47	60	3.6	3.0	1.15	31	110	2.70 2.70
P12	0.98, 91%	0.85, 63%	47.5	47	5.4	5.5	1.20	45	200	2.86 2.60
P13	1.00, 88%	0.70, 67%	49	49	5.8	5.8	1.11	52	210	3.18 3.18
P14	1.94, 92%	0.88, 67%	46	53	8.6	4.4			160	3.28 3.27
P15	0.97, 96%	0.64, 66%	47	47	6.9	7.1			250	3.12 2.91

Table 5B: Amplification ratios determined by AEX and QPA for MRKHVE3 over several continuous passaging in serum free media. MRKHVE3 is the new vector backbone which does NOT carry a transgene.

MRKHVE3

	Xv (10 ⁶ cells/ml), Infection	Viability (%), Harvest	Harvest Time h.p.i.	Cell Passage Number	Titer 10 ¹⁰ vp/ml culture	Titer 10 ⁴ vp/cell	QPA 10 ⁶ TCID ₅₀ /ml	Ratio AEX:QPA	Amplification Ratio	AEX Internal Control
P4	1.10, 97%	1.28, 79%	49	54	4.1	3.8	1.70	25	300 (MOI = 125)	
P5	0.92, 89%	1.18, 77%	47	48	4.3	4.7	1.24	35	170	
P6	1.55, 86%	1.26, 76%	49.5	50	1.2	0.8	0.56	21	30	
P6	1.09, 87%	1.11, 81%	49	52	4.0	3.6	1.16	34	130	
P7	1.17, 91%	1.22, 91%	47.5	54	3.7	3.2	0.50	74	110	
P8	0.98, 88%	1.41, 83%	48	56	2.1	2.1	0.47	45	75	3.12 2.84
P9	1.20, 89%	1.26, 81%	47.5	58	0.8	0.7	0.29	28	25	2.70 2.60
P10	0.99, 82%	1.55, 86%	47	60	2.3	2.3	0.43	53	80	2.70 2.70
P11	1.07, 96%	1.25, 83%	48	47	2.7	2.5	0.41	66	90	2.86 2.60
P12	0.80, 91%	1.14, 80%	49.5	49	5.9	7.4	0.48	123	260	3.18 3.18
P13	1.96, 95%	1.14, 85%	45.5	53	5.8	3.0			110	3.28 3.27
P14	0.97, 96%	1.03, 98%	48.5	47	9.4	9.7			350	3.12 2.91
P15	0.87, 99%	0.97, 69%	49.5	49	5.3	6.1			218	2.78 2.52

Table 5C. Amplification ratios determined by AEX and QPA for MRKAd5gag(E3-) over several continuous passaging in serum free media. This construct is identical to the MRKAd5gag construct except that this version is DELETED of the E3 gene.

5

MRKAd5gag(E3-)

	Xv (10 ⁶ cells/ml), Infection	Viability (%), Harvest	Harvest Time h.p.i.	Cell Passage Number	Titer 10 ¹⁰ vp/ml culture	Titer 10 ⁴ vp/cell	QPA 10 ⁶ TCID ₅₀ /ml	Ratio AEX:QPA	Amplification Ratio	AEX Internal Control
P4	1.62, 77%	1.12, 62%	47.5	46	2.0	1.2	0.92	20	100 (MOI=125)	
P5	1.16, 92%	0.62, 43%	49	49	3.3	2.9	0.99	34	100	
P6	1.71, 86%	0.20, 10%	49	50	4.7	2.7	1.70	28	100	
P6	1.09, 97%	0.63, 54%	49.5	52	5.4	5.0	1.76	31	180	
P7	1.17, 91%	0.98, 72%	47.50	54	7.1	6.1	0.67	106	220	
P8	0.98, 88%	0.77, 48%	48	56	3.1	3.2	0.66	47	115	3.12 2.84
P9	1.20, 89%	1.03, 72%	48	58	1.8	1.5	0.57	32	55	2.70 2.60
P10	0.99, 82%	0.80, 62%	46.5	60	3.2	3.2	0.68	47	115	2.70 2.70
P11	1.07, 96%	0.88, 70%	48.5	47	5.9	5.5	0.68	87	200	2.86 2.60
P12	0.80, 91%	0.67, 59%	50	49	5.1	6.4	0.72	71	230	3.18 3.18
P13	1.96, 95%	0.91, 59%	45.5	53	7.4	3.8			135	3.28 3.27
P14	0.97, 96%	0.81, 74%	48	47	6.8	7.0			250	3.12 2.91
P15	0.87, 99%	0.84, 56%	49	49	4.8	5.5			196	2.78 2.52

EXAMPLE 14

Gag Expression Analysis of the Novel Constructs

In vitro gag analysis of the MRK Ad5 HIV-1 gag and the original HIV-gag vectors (research and clinical lot) show comparable gag expression. The clinical lot shows only a slightly reduced gag expression level. The most noticeable difference is with the mCMV vector. This vector shows roughly 3 fold lower expression levels compared with the other vectors tested (which all contain hCMV promoters). The mCMV-FLgag with bGHpA assay was performed three times using different propagation and purification lots and it consistently exhibited weaker gag expression.

EXAMPLE 15

Evaluation of MRK Ad5 HIV-1 gag and Other gag-Containing Adenovectors in Balb/c Mice

Cohorts of 10 balb/c mice were vaccinated intramuscularly with escalating doses of MRK Ad5 HIV-1 gag, and the research and clinical lots of original Ad5HIV-1gag. Serum samples were collected 3 weeks post dose 1 and analyzed by anti-p24 sandwich ELISA.

Anti-p24 titers in mice that received MRK Ad5 HIV-1 gag (10^7 and 10^9 vp(viral particle) doses) were comparable (Figure 13) to those of the research lot of Ad5HIV-1 gag, for which much of the early rhesus data were generated on. These titers were also comparable when E3 is deleted (MRKAd5hCMVgagbGHpA(E3-)) or SPA is substituted for bGHpA terminator (MRKAd5 hCMV-gag-SPA (E3+)) or murine CMV promoter is used in place of hCMV (MRKAd5 mCMV-gag-bGHpA (E3+)) in the MRKAd5 backbone.

The results shown in Table 7 indicate that the three other vectors (in addition to the preferred vector, MRK Ad5 HIV-1 gag, are also capable of inducing strong anti-gag antibody responses in mice. Interestingly enough, while the mCMV-FLgag construct containing bGHpA and E3+ in an E1 parallel orientation showed lowest gag expression in the COS cell *in vitro* infection (Table 6) in comparison with the other vectors tested, it generated the greatest anti-gag antibody response this *in vivo* Balb/c study. Table 7 also shows a dose response in anti-gag antibody production in both the research and the clinical lot. As expected, the clinical lot shows reduced anti-gag antibody induction at each dosage level compared to the same dosage used for the research lot.

Table 6: *In vitro* analysis for gag expression in COS cells by Elisa assay.

Viral Vectors ^a	$\mu\text{g gag}/4.8 \times 10^5 \text{ COS}/10^8 \text{ parts}/48\text{hr}$
MRKAd5gag ^b	1.40
Clinical lot Ad5gag ^c	1.28
Research lot Ad5gag ^d	1.32
MCMVFL-gagbGHpA ^e	0.42

^a $A_{260\text{nm}}$ absorbance readings taken for viral particle determinations.

^b MRKAd5gag was produced in serum free conditions and purified at P5.

^c Clinical lot# Ad5gagFN0001

^d Research Ad5FLgag lot# 6399

^e mCMVFL-gagbGHpA was produced in serum free conditions and purified at P5.

Table 7: mHIV020 Anti-p24 Ab Titers in Balb/c mice (n=10) vaccinated with various Adgag constructs and lots (3 week post dose1).

Group ID	Vaccine	Dose (vp)	GMT	SE upper	SE lower
1	^a MRKAd5gag	10 ⁷	25600	5877	4780
2	"	10 ⁹	409600	94028	76473
3	hCMV FL-gag bGHpA [E3-] →	10 ⁷	7352	2077	1620
4	"	10 ⁹	235253	59767	47659
5	hCMV FL-gag SPA [E3+] →	10 ⁷	12800	9905	236
6	"	10 ⁹	310419	99181	75165
7	^b mCMV FL-gag bGHpA [E3+] →	10 ⁷	44572	23504	15389
8	"	10 ⁹	941014	239068	190636
9	^c hCMV FL-gag bGHpA [E3-] ←	10 ⁷	3676	934	745
10	"	10 ⁹	117627	17491	15227
11	research lot hCMV intronA FL-gag bGHpA [E3-] <-	10 ⁶	528	262	175
12	"	10 ⁷	14703	5274	3882
13	"	10 ⁸	58813	14942	11915
14	"	10 ⁹	204800	53232	42250
15	clinical lot hCMVintronA FL-gag bGHpA [E3-] <-	10 ⁶	230	82	61
16	"	10 ⁷	4222	3405	1138
17	"	10 ⁸	19401	3939	3274
18	"	10 ⁹	89144	25187	19639
19	Naïve	none	93	7	6

*2x50 µL i.m. (quad) injections/animal

P.I.s: Youil, Chen, Casimiro

Vaccination: T. Toner, Q. Su

Assay: M. Chen

^aThe structure of MRKAd5gag is: hCMVFL-gagbGHpA [E3+] → The same lot of MRKAd5gag used in this rodent study was used in the Rhesus monkey study (Tables 7 and 8).

^bThe same lot of mCMVFL-gagbGHpA[E3+] used in the *in vitro* study (Table 6) was used here.

^cThis construct was designed by Volker Sandig. It contains a shorter version of the hCMV promoter than that used in the MRK constructs. The adenovector backbone is identical to the original backbone used in the original Adgag vector. Expression at 10e7 dose from this vector is 7 fold lower than the same dose of the MRKAd5gag and 4 fold lower than the research lot.

EXAMPLE 16

Comparison of Humoral and Cellular Responses Towards the Original Ad-gag Construct with the New MRK Ad5 HIV-1 gag in Rhesus Monkeys

- 5 Cohorts of 3 rhesus monkeys were vaccinated intramuscularly with MRK Ad5 HIV-1 gag or the clinical Ad5gag bulk at two doses, 10¹¹ vp and 10⁹ vp. Immunizations were conducted at week 0, 4, and 25. Serum and PBMC samples were collected at selected time points. The serum sample were assayed for anti-p24 Ab titers (using competitive based assay) and the PBMCs for antigen-specific IFN-
10 gamma secretion following overnight stimulation with gag 20-mer peptide pool (via ELISpot assay).

The results shown in Table 8 indicate comparable responses with respect to the generation of anti-gag antibodies. The frequencies of gag-specific T cells in

peripheral blood as summarized in Table 9 demonstrate a strong cellular immune response generated after a single dose with the new construct MRK Ad5 HIV-1 gag. The responses are also boostable with second dose of the same vector. The vector is also able to induce CD8+ T cell responses (as evident by remaining spot counts after CD4+ depletion of PBMCs) which are responsible for cytotoxic activity.

Table 8 Anti-p24 antibody titers (in mMU/mL) in rhesus macaques immunized with gag-expressing adenovectors (Protocol HIV203).

Vaccine	Pre	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 25	Wk 28
MRKAd5gag^a, 10¹¹ vp								
97N010	<10	118	5528	11523	7062	21997	ND	51593
97N116	<10	62	772	1447	1562	2174	ND	20029
98X007	<10	66	3353	6156	6845	3719	ND	24031
MRKAd5gag, 10⁹ vp								
97N120	<10	51	204	318	366	482	ND	6550
97N144	<10	18	118	274	706	888	ND	7136
98X008	<10	15	444	386	996	1072	ND	12851
Ad5gag^b, Clinical Lot, 10¹¹ vp								
97X001	<10	87	2579	4718	7174	7250	ND	69226
97N146	<10	72	3604	7380	7526	18906	ND	60283
98X009	<10	78	4183	3946	3124	6956	ND	26226
Ad5gag, Clinical Lot, 10⁹ vp								
97N020	<10	<10	143	371	390	1821	ND	17177
97X003	<10	<10	39	93	156	596	ND	2053
98X012	<10	81	342	717	956	1558	ND	11861
^a MRKAd5gag (hCMV, bGHPA, E3+)								
^b original Ad5gag vector (hCMV/Intron A, bGHPA, E3-), lot#FN0001								
ND, not determined								

Table 9. Number of gag-specific T cells per million peripheral blood mononuclear cells (PBMCs) in rhesus monkeys immunized with gag-expressing adenovectors. Also included are those frequencies in PBMCs depleted of CD4⁺ T cells.

Grp #	Vaccination T=0,4,25 wks	Monkey ID	T=4 Wk		T=6 Wk		T=11 Wk		T=16 Wk		T=25 Wk		T=28 Wk	
			Media ^a	Gag H ^b	Media	Gag H	Media	Gag H	Media	Gag H	Media	Gag H	Media	Gag H
1	MRKAd5gag 10 ¹¹ vp	97N010	6	89	0	395	0	1058	0	1174	3	775	4	1074
		97N010(CD4-)	4	38			3	993			0	76	0	594
		97N116	1	396	1	609	0	534	4	395	1	261	0	408
		97N116(CD4-)	11	676			0	593			0	184	0	666
		98X007	10	579	0	1304	3	2193	1	2118	3	1588	0	2113
		98X007(CD4-)	20	965			0	2675			0	1656	0	1278
2	MRKAd5gag 10 ⁹ vp	97N120	5	275	1	249	4	141	4	119	9	206	4	219
		97N120(CD4-)	11	170			0	85			0	75	1	219
		97N144	3	236	6	438	1	318	3	256	1	98	5	373
		97N144(CD4-)	6	148			0	285			ND	ND	0	625
		98X008	4	368	1	1090	3	891	4	673	3	473	5	735
		98X008(CD4-)	14	696			0	1175			0	391	4	848
3	Ad5gag clinical lot 10 ¹¹ vp	97X001	0	261	1	485	0	817	0	1220b	1	894	0	1858
		97X001(CD4-)	10	283			3	996			0	1010	0	1123
		97N146	3	150	1	465	0	339	1	1272	3	1238	3	1785
		97N146(CD4-)	6	133			0	370			0	654	0	971
		98X009	0	93	3	339	3	559	0	896	1	384	0	1748
		98X009(CD4-)	0	73			0	333			0	225	0	644
4	Ad5gag clinical lot 10 ⁹ vp	97N020	3	30	1	101	0	66	0	36	0	26	0	41
		97N020(CD4-)	10	29			0	15			0	1	0	16
		97X003	4	68	5	134	0	18	1	38	4	38	6	81
		97X003(CD4-)	9	40			0	6			0	4	0	19
		98X012	5	95	3	54	1	34	0	18	0	20	1	121
		98X012(CD4-)	11	70			0	11			0	8	0	41
5	Naïve	96R041	6	8	1	1	0	0	0	0	0	0	1	0
		053F	14	18	5	16	20	14	19	15	10	15	24	9

Based on either 4x10⁶ or 2x10⁶ cells per well (depending on spot density)

ND, not determined

^aMock or no peptide control

^bPool of 20-aa peptides overlapping by 10 aa and encompassing the gag sequence

5

The adenovectors described herein and, particularly, MRK Ad5 HIV-1 gag, represent very promising HIV-gag adenovectors with respect to their enhanced growth characteristics in both serum and, more importantly, in serum-free media conditions. In comparison with the current HIV-1 gag adenovector construct, MRK Ad5 HIV-1 gag shows a 5-10 fold increased amplification rate. We have shown that it is genetically stable at passage 21. This construct is able to generate significant cellular immune responses *in vivo* even at a relatively low dose of 10⁹ vp. The potency of the MRKAd5gag construct is comparable to, if not better than the original HIV-1gag vector as shown in this rhesus monkey study.

15

EXAMPLE 17

CODON OPTIMIZED HIV-1 POL AND CODON OPTIMIZED HIV-1 POL MODIFICATIONS

20

The open reading frames for the various synthetic *pol* genes disclosed herein comprise coding sequences for the reverse transcriptase (or RT which consists of a polymerase and RNase H activity) and integrase (IN). The protein sequence is based

on that of Hxb2r, a clonal isolate of IIIB; this sequence has been shown to be closest to the consensus clade B sequence with only 16 nonidentical residues out of 848 (Korber, et al., 1998, Human retroviruses and AIDS, Los Alamos National Laboratory, Los Alamos, New Mexico). The skilled artisan will understand after review of this specification that any available HIV-1 or HIV-2 strain provides a potential template for the generation of HIV pol DNA vaccine constructs disclosed herein. It is further noted that the protease gene is excluded from the DNA vaccine constructs of the present invention to insure safety from any residual protease activity in spite of mutational inactivation. The design of the gene sequences for both wild-type (wt-pol) and inactivated pol (IA-pol) incorporates the use of human preferred ("humanized") codons for each amino acid residue in the sequence in order to maximize *in vivo* mammalian expression (Lathe, 1985, *J. Mol. Biol.* 183:1-12). As can be discerned by inspecting the codon usage in SEQ ID NOs: 1, 3, 5 and 7, the following codon usage for mammalian optimization is preferred: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG); Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which, as noted elsewhere in this specification, is hereby incorporated by reference. It is intended that the skilled artisan may use alternative versions of codon optimization or may omit this step when generating HIV pol vaccine constructs within the scope of the present invention. Therefore, the present invention also relates to non-codon optimized versions of DNA molecules and associated recombinant adenoviral HIV vaccines which encode the various wild type and modified forms of the HIV Pol protein disclosed herein. However, codon optimization of these constructs is a preferred embodiment of this invention.

A particular embodiment of this portion of the invention comprises codon optimized nucleotide sequences which encode wt-pol DNA constructs (herein, "wt-pol" or "wt-pol (codon optimized)") wherein DNA sequences encoding the protease (PR) activity are deleted, leaving codon optimized "wild type" sequences which encode RT (reverse transcriptase and RNase H activity) and IN integrase activity. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:1, the open reading frame being contained from an initiating Met residue at nucleotides 10-12 to a termination codon from nucleotides 2560-2562. SEQ ID NO:1 is as follows:

AGATCTACCA TGGCCCCCAT CTCCCCCATT GAGACTGTGC CTGTGAAGCT GAAGCCTGGC
ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG

GAAATCTGCA CTGAGATGGA GAAGGAGGGC AAAATCTCCA AGATTGGCCC CGAGAACCCC
 TACAACACCC CTGTGTTTGC CATCAAGAAG AAGGACTCCA CCAAGTGGAG GAAGCTGGTG
 GACTTCAGGG AGCTGAACAA GAGGACCCAG GACTTCTGGG AGGTGCAGCT GGGCATCCCC
 CACCCCGCTG GCCTGAAGAA GAAGAAGTCT GTGACTGTGC TGGATGTGGG GGATGCCTAC
 5 TTCTCTGTGC CCCTGGATGA GGACTTCAGG AAGTACACTG CCTTCACCAT CCCCTCCATC
 AACAAATGAGA CCCCTGGCAT CAGGTACCAG TACAATGTGC TGCCCCAGGG CTGGAAGGGC
 TCCCCTGCCA TCTTCCAGTC CTCCATGACC AAGATCCTGG AGCCCTTCAG GAAGCAGAAC
 CCTGACATTG TGATCTACCA GTACATGGAT GACCTGTATG TGGGCTCTGA CCTGGAGATT
 GGGCAGCACA GGACCAAGAT TGAGGAGCTG AGGCAGCACC TGCTGAGGTG GGGCCTGACC
 10 ACCCCTGACA AGAAGCACCA GAAGGAGCCC CCCTTCCTGT GGATGGGCTA TGAGCTGCAC
 CCCGACAAGT GGACTGTGCA GCCCATTTGTG CTGCCTGAGA AGGACTCCTG GACTGTGAAT
 GACATCCAGA AGCTGGTGGG CAAGCTGAAC TGGGCCTCCC AAATCTACCC TGGCATCAAG
 GTGAGGCAGC TGTGCAAGCT GCTGAGGGGC ACCAAGGCCC TGACTGAGGT GATCCCCCTG
 ACTGAGGAGG CTGAGCTGGA GCTGGCTGAG AACAGGGAGA TCCTGAAGGA GCCTGTGCAT
 15 GGGGTGTACT ATGACCCCTC CAAGGACCTG ATTGCTGAGA TCCAGAAGCA GGGCCAGGGC
 CAGTGGACCT ACCAAATCTA CCAGGAGCCC TTCAAGAACC TGAAGACTGG CAAGTATGCC
 AGGATGAGGG GGGCCACAC CAATGATGTG AAGCAGCTGA CTGAGGCTGT GCAGAAGATC
 ACCACTGAGT CCATTGTGAT CTGGGGCAAG ACCCCCAAGT TCAAGCTGCC CATCCAGAAG
 GAGACCTGGG AGACCTGGTG GACTGAGTAC TGGCAGGCCA CCTGGATCCC TGAGTGGGAG
 20 TTTGTGAACA CCCCCCCCCT GGTGAAGCTG TGGTACCAGC TGGAGAAGGA GCCCATTTGTG
 GGGGCTGAGA CCTTCTATGT GGATGGGGCT GCCAACAGGG AGACCAAGCT GGGCAAGGCT
 GGCTATGTGA CCAACAGGGG CAGGCAGAAG GTGGTGACCC TGACTGACAC CACCAACCAG
 AAGACTGAGC TCCAGGCCAT CTACCTGGCC CTCCAGGACT CTGGCCTGGA GGTGAACATT
 GTGACTGACT CCCAGTATGC CCTGGGCATC ATCCAGGCCC AGCCTGATCA GTCTGAGTCT
 25 GAGCTGGTGA ACCAGATCAT TGAGCAGCTG ATCAAGAAGG AGAAGGTGTA CCTGGCCTGG
 GTGCCTGCCC ACAAGGGCAT TGGGGGCAAT GAGCAGGTGG ACAAGCTGGT GTCTGCTGGC
 ATCAGGAAGG TGCTGTTTCT GGATGGCATT GACAAGGCCC AGGATGAGCA TGAGAAGTAC
 CACTCCAACCT GGAGGGCTAT GGCCTCTGAC TTCAACCTGC CCCCTGTGGT GGCTAAGGAG
 ATTGTGGCCT CCTGTGACAA GTGCCAGCTG AAGGGGGAGG CCATGCATGG GCAGGTGGAC
 30 TGCTCCCCTG GCATCTGGCA GCTGGACTGC ACCACCTGG AGGGCAAGGT GATCCTGGTG
 GCTGTGCATG TGGCCTCCGG CTACATTGAG GCTGAGGTGA TCCCTGCTGA GACAGGCCAG
 GAGACTGCCT ACTTCCTGCT GAAGCTGGCT GGCAGGTGGC CTGTGAAGAC CATCCACACT
 GACAATGGCT CCAACTTCAC TGGGGCCACA GTGAGGGCTG CCTGCTGGTG GGCTGGCATC
 AAGCAGGAGT TTGGCATCCC CTACAACCCC CAGTCCCAGG GGGTGGTGGG GTCCATGAAC
 35 AAGGAGCTGA AGAAGATCAT TGGGCAGGTG AGGGACCAGG CTGAGCACCT GAAGACAGCT
 GTGCAGATGG CTGTGTTTCT CCACAACCTC AAGAGGAAGG GGGGCATCGG GGGCTACTCC

GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG
 CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCCTGTGG
 AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGGCTG TGGTGATCCA GGACAACTCT
 GACATCAAGG TGGTGCCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAGATG
 5 GCTGGGGATG ACTGTGTGGC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ
 ID NO:1) .

The open reading frame of the wild type pol construct disclosed as SEQ ID
 NO:1 contains 850 amino acids, disclosed herein as SEQ ID NO:2, as follows:

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro
 10 Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
 Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys
 Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala
 Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg
 Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile
 15 Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp
 Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys
 Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile
 Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala
 Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln
 20 Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly
 Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg
 Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln
 Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys
 Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val
 25 Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile
 Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
 Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu
 Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr
 Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln
 30 Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys
 Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys
 Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile
 Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp
 Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp
 35 Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu
 Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala

Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 5 Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile
 Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
 Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys
 Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys
 Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro
 10 Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys
 Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln
 Leu Asp Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His
 Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly
 Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val
 15 Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe Thr Gly Ala Thr Val
 Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro
 Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys Glu Leu
 Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
 Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
 20 Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr
 Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn
 Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro
 Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
 Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
 25 Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp
 Glu Asp (SEQ ID NO:2) .

The present invention especially relates to an adenoviral vector vaccine which
 comprises a codon optimized HIV-1 DNA pol construct wherein, in addition to
 deletion of the portion of the wild type sequence encoding the protease activity, a
 30 combination of active site residue mutations are introduced which are deleterious to
 HIV-1 pol (RT-RH-IN) activity of the expressed protein. Therefore, the present
 invention preferably relates to an adenoviral HIV-1 DNA pol-based vaccine wherein
 the construct is devoid of DNA sequences encoding any PR activity, as well as
 containing a mutation(s) which at least partially, and preferably substantially,
 35 abolishes RT, RNase and/or IN activity. One type of HIV-1 pol mutant which is part
 and parcel of an adenoviral vector vaccine may include but is not limited to a mutated

DNA molecule comprising at least one nucleotide substitution which results in a point mutation which effectively alters an active site within the RT, RNase and/or IN regions of the expressed protein, resulting in at least substantially decreased enzymatic activity for the RT, RNase H and/or IN functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a HIV-1 DNA pol construct contains a mutation or mutations within the Pol coding region which effectively abolishes RT, RNase H and IN activity. An especially preferable HIV-1 DNA pol construct in a DNA molecule which contains at least one point mutation which alters the active site of the RT, RNase H and IN domains of Pol, such that each activity is at least substantially abolished. Such a HIV-1 Pol mutant will most likely comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. To this end, an especially preferred HIV-1 DNA pol construct is exemplified herein and contains nine codon substitution mutations which results in an inactivated Pol protein (IA Pol: SEQ ID NO:4, Figure 17A-C) which has no PR, RT, RNase or IN activity, wherein three such point mutations reside within each of the RT, RNase and IN catalytic domains. Therefore, an especially preferred exemplification is an adenoviral vaccine which comprises, in an appropriate fashion, a DNA molecule which encodes IA-pol, which contains all nine mutations as shown below in Table 1. An additional preferred amino acid residue for substitution is Asp551, localized within the RNase domain of Pol. Any combination of the mutations disclosed herein may suitable and therefore may be utilized as an IA-Pol-based vaccine of the present invention. While addition and deletion mutations are contemplated and within the scope of the invention, the preferred mutation is a point mutation resulting in a substitution of the wild type amino acid with an alternative amino acid residue.

Table 1

	<u>wt aa</u>	<u>aa residue</u>	<u>mutant aa</u>	<u>enzyme function</u>
	Asp	112	Ala	RT
	Asp	187	Ala	RT
30	Asp	188	Ala	RT
	Asp	445	Ala	RNase H
	Glu	480	Ala	RNase H
	Asp	500	Ala	RNase H
	Asp	626	Ala	IN
35	Asp	678	Ala	IN
	Glu	714	Ala	IN

It is preferred that point mutations be incorporated into the LApol mutant adenoviral vaccines of the present invention so as to lessen the possibility of altering epitopes in and around the active site(s) of HIV-1 Pol.

To this end, SEQ ID NO:3 discloses the nucleotide sequence which codes for a codon optimized pol in addition to the nine mutations shown in Table 1, disclosed as follows, and referred to herein as "LApol":

```

AGATCTACCA TGGCCCCCAT CTCCCCCATT GAGACTGTGC CTGTGAAGCT GAAGCCTGGC
ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG
GAAATCTGCA CTGAGATGGA GAAGGAGGGC AAAATCTCCA AGATTGGCCC CGAGAACCCC
10 TACAACACCC CTGTGTTTGC CATCAAGAAG AAGGACTCCA CCAAGTGGAG GAAGCTGGTG
GACTTCAGGG AGCTGAACAA GAGGACCCAG GACTTCTGGG AGGTGCAGCT GGGCATCCCC
CACCCCGCTG GCCTGAAGAA GAAGAAGTCT GTGACTGTGC TGGCTGTGGG GGATGCCTAC
TTCTCTGTGC CCCTGGATGA GGAATTCAGG AAGTACACTG CCTTCACCAT CCCCTCCATC
AACAATGAGA CCCCTGGCAT CAGGTACCAG TACAATGTGC TGCCCCAGGG CTGGAAGGGC
15 TCCCCTGCCA TCTTCCAGTC CTCCATGACC AAGATCCTGG AGCCCTTCAG GAAGCAGAAC
CCTGACATTG TGATCTACCA GTACATGGCT GCCCTGTATG TGGGCTCTGA CCTGGAGATT
GGGCAGCACA GGACCAAGAT TGAGGAGCTG AGGCAGCACC TGCTGAGGTG GGGCCTGACC
ACCCCTGACA AGAAGCACCA GAAGGAGCCC CCCTTCCTGT GGATGGGCTA TGAGCTGCAC
CCCGACAAGT GGAATGTGCA GCCCATTGTG CTGCCTGAGA AGGACTCCTG GACTGTGAAT
20 GACATCCAGA AGCTGGTGGG CAAGCTGAAC TGGGCCTCCC AAATCTACCC TGGCATCAAG
GTGAGGCAGC TGTGCAAGCT GCTGAGGGGC ACCAAGGCCC TGAATGAGGT GATCCCCCTG
ACTGAGGAGG CTGAGCTGGA GCTGGCTGAG AACAGGGAGA TCCTGAAGGA GCCTGTGCAT
GGGGTGTACT ATGACCCCTC CAAGGACCTG ATTGCTGAGA TCCAGAAGCA GGGCCAGGGC
CAGTGGACCT ACCAAATCTA CCAGGAGCCC TTCAAGAACC TGAAGACTGG CAAGTATGCC
25 AGGATGAGGG GGGCCACAC CAATGATGTG AAGCAGCTGA CTGAGGCTGT GCAGAAGATC
ACCACTGAGT CCATTGTGAT CTGGGGCAAG ACCCCCAAGT TCAAGCTGCC CATCCAGAAG
GAGACCTGGG AGACCTGGTG GACTGAGTAC TGGCAGGCCA CCTGGATCCC TGAGTGGGAG
TTTGTGAACA CCCCCCCCCT GGTGAAGCTG TGGTACCAGC TGGAGAAGGA GCCCATTGTG
GGGGCTGAGA CCTTCTATGT GGCTGGGGCT GCCAACAGGG AGACCAAGCT GGGCAAGGCT
30 GGCTATGTGA CCAACAGGGG CAGGCAGAAG GTGGTGACCC TGAATGACAC CACCAACCAG
AAGACTGCCC TCCAGGCCAT CTACCTGGCC CTCCAGGACT CTGGCCTGGA GGTGAACATT
GTGACTGCCT CCCAGTATGC CCTGGGCATC ATCCAGGCCC AGCCTGATCA GTCTGAGTCT
GAGCTGGTGA ACCAGATCAT TGAGCAGCTG ATCAAGAAGG AGAAGGTGTA CCTGGCCTGG
GTGCCTGCCC ACAAGGGCAT TGGGGGCAAT GAGCAGGTGG ACAAGCTGGT GTCTGCTGGC
35 ATCAGGAAGG TGCTGTTTCT GGATGGCATT GACAAGGCCC AGGATGAGCA TGAGAAGTAC
CACTCCAACCT GGAGGGCTAT GGCTCTGAC TTCAACCTGC CCCCTGTGGT GGCTAAGGAG

```

ATGTGTGGCCT CCTGTGACAA GTGCCAGCTG AAGGGGGAGG CCATGCATGG GCAGGTGGAC
 TGCTCCCCTG GCATCTGGCA GCTGGCCTGC ACCCACCTGG AGGGCAAGGT GATCCTGGTG
 GCTGTGCATG TGGCCTCCGG CTACATTGAG GCTGAGGTGA TCCCTGCTGA GACAGGCCAG
 GAGACTGCCT ACTTCCTGCT GAAGCTGGCT GGCAGGTGGC CTGTGAAGAC CATCCACACT
 5 GCCAATGGCT CCAACTTCAC TGGGGCCACA GTGAGGGCTG CCTGCTGGTG GGCTGGCATC
 AAGCAGGAGT TTGGCATCCC CTACAACCCC CAGTCCCAGG GGGTGGTGGC CTCCATGAAC
 AAGGAGCTGA AGAAGATCAT TGGGCAGGTG AGGGACCAGG CTGAGCACCT GAAGACAGCT
 GTGCAGATGG CTGTGTTTCA CCACAACCTT AAGAGGAAGG GGGGCATCGG GGGCTACTCC
 GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG
 10 CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCTGTGG
 AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGGCTG TGGTGATCCA GGACAACTCT
 GACATCAAGG TGGTGCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAGATG
 GCTGGGGATG ACTGTGTGGC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ ID
 NO:3) .

15 In order to produce the IA-pol-based adenoviral vaccines of the present
 invention, inactivation of the enzymatic functions was achieved by replacing a total of
 nine active site residues from the enzyme subunits with alanine side-chains. As
 shown in Table 1, all residues that comprise the catalytic triad of the polymerase,
 namely Asp112, Asp187, and Asp188, were substituted with alanine (Ala) residues
 20 (Larder, et al., *Nature* 1987, 327: 716-717; Larder, et al., 1989, *Proc. Natl. Acad. Sci.*
 1989, 86: 4803-4807). Three additional mutations were introduced at Asp445,
 Glu480 and Asp500 to abolish RNase H activity (Asp551 was left unchanged in this
 IA Pol construct), with each residue being substituted for an Ala residue, respectively
 (Davies, et al., 1991, *Science* 252:, 88-95; Schatz, et al., 1989, *FEBS Lett.* 257: 311-
 25 314; Mizrahi, et al., 1990, *Nucl. Acids. Res.* 18: pp. 5359-5353). HIV pol integrase
 function was abolished through three mutations at Asp626, Asp678 and Glu714.
 Again, each of these residues has been substituted with an Ala residue (Wiskerchen,
 et al., 1995, *J. Virol.* 69: 376-386; Leavitt, et al., 1993, *J. Biol. Chem.* 268: 2113-
 2119). Amino acid residue Pro3 of SEQ ID NO:4 marks the start of the RT gene.
 30 The complete amino acid sequence of IA-Pol is disclosed herein as SEQ ID NO:4 and
 Figure 17A-C, as follows:

35 Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro
 Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys
 Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys
 Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala
 Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg

Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile
 Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala
 Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys
 Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile
 5 Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala
 Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln
 Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly
 Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg
 Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln
 10 Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys
 Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val
 Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile
 Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr
 Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu
 15 Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr
 Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln
 Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys
 Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys
 Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile
 20 Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp
 Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp
 Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu
 Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala
 Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly
 25 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile
 Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
 30 Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys
 Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys
 Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro
 Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys
 Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln
 35 Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His
 Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly

Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val
 Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val
 Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro
 Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu
 5 Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
 Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
 Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr
 Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn
 Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro
 10 Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
 Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
 Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp
 Glu Asp (SEQ ID NO:4) .

As noted above, it will be understood that any combination of the mutations
 15 disclosed above may be suitable and therefore be utilized as an IA-pol-based
 adenoviral HIV vaccine of the present invention, either when administered alone or in
 a combined modality regime and/or a prime-boost regimen. For example, it may be
 possible to mutate only 2 of the 3 residues within the respective reverse transcriptase,
 RNase-H, and integrase coding regions while still abolishing these enzymatic
 20 activities. However, the IA-pol construct described above and disclosed as SEQ ID
 NO:3, as well as the expressed protein (SEQ ID NO:4;) is preferred. It is also
 preferred that at least one mutation be present in each of the three catalytic domains.

Another aspect of this portion of the invention are codon optimized HIV-1
 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal
 25 peptide such as from tPA (tissue-type plasminogen activator) or by a leader peptide
 such as is found in highly expressed mammalian proteins such as immunoglobulin
 leader peptides. Any functional leader peptide may be tested for efficacy. However,
 a preferred embodiment of the present invention, as with HIV-1 Nef constructs shown
 herein, is to provide for a HIV-1 Pol mutant adenoviral vaccine construction wherein
 30 the pol coding region or a portion thereof is operatively linked to a leader peptide,
 preferably a leader peptide from human tPA. In other words, a codon optimized
 HIV-1 Pol mutant such as IA-Pol (SEQ ID NO:4) may also comprise a leader peptide
 at the amino terminal portion of the protein, which may effect cellular trafficking and
 hence, immunogenicity of the expressed protein within the host cell. As noted in
 35 Figure 16A-B, a DNA vector which may be utilized to practice the present invention
 may be modified by known recombinant DNA methodology to contain a leader signal

peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Pol protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading
 5 frame for the Pol protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1 Pol protein of interest, including but not limited to a HIV-1 Pol protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized
 10 herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17). Therefore, another aspect of the present invention is to generate HIV-1 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal peptide such as from tPA. To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity
 15 is deleted and a human tPA leader sequence is fused to the 5' end of the coding region. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame disclosed herein as SEQ ID NO:6.

To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity is deleted
 20 and a human tPA leader sequence is fused to the 5' end of the coding region (herein, "tPA-wt-pol"). A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame being contained from an initiating Met residue at nucleotides 8-10 to a termination codon from nucleotides 2633-2635. SEQ ID NO:5 is as follows:

25 GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGTG CTGCTGCTGT GTGGAGCAGT
 CTTCGTTTCG CCCAGCGAGA TCTCCGCCCC CATCTCCCCC ATTGAGACTG TGCCTGTGAA
 GCTGAAGCCT GGCATGGATG GCCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT
 CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAATCT CCAAGATTGG
 CCCCAGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG
 30 GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA
 GCTGGGCATC CCCCACCCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGATGT
 GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC
 CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA
 GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT
 35 CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GATGACCTGT ATGTGGGCTC
 TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG

GTGGGGCCTG ACCACCCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG
 CTATGAGCTG CACCCCGACA AGTGGACTGT GCAGCCCATT GTGCTGCCTG AGAAGGACTC
 CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA
 CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA
 5 GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA
 GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA
 GCAGGGCCAG GGCCAGTGGG CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC
 TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGA CTGAGGCT
 TGTGCAGAAG ATCACC ACTG AGTCCATTGT GATCTGGGGC AAGACCCCCA AGTTCAAGCT
 10 GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT
 CCCTGAGTGG GAGTTTGTGA ACACCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA
 GGAGCCCATT GTGGGGGCTG AGACCTTCTA TGTGGATGGG GCTGCCAACA GGGAGACCAA
 GCTGGGCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA
 CACCACCAAC CAGAAGACTG AGCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT
 15 GGAGGTGAAC ATTGTGACTG ACTCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA
 TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT
 GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT
 GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA
 GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT
 20 GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA
 TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGAC TGCACCCACC TGGAGGGCAA
 GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC
 TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA
 GACCATCCAC ACTGACAATG GCTCCAACCT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG
 25 GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCAGTCCC AGGGGGTGGT
 GGAGTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA
 CCTGAAGACA GCTGTGCAGA TGGCTGTGTT CATCCACAAC TTCAAGAGGA AGGGGGGCAT
 CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA
 GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG
 30 GAACCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT
 CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA
 TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC
 GGGCAGATCT (SEQ ID NO:5).

The open reading frame of the wild type tPA-pol construct disclosed as SEQ
 35 ID NO:5 contains 875 amino acids, disclosed herein as SEQ ID NO:6, as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile
 Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 5 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 10 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
 Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
 15 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
 20 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 25 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 30 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 35 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile

Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 5 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn
 Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala
 Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 10 Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 15 Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:6).

The present invention also relates to a codon optimized HIV-1 Pol mutant
 20 contained within a recombinant adenoviral vector such as IA-Pol (SEQ ID NO:4)
 which comprises a leader peptide at the amino terminal portion of the protein, which
 may effect cellular trafficking and hence, immunogenicity of the expressed protein
 within the host cell. Any such adenoviral-based HIV-1 DNA pol mutant disclosed in
 the above paragraphs is suitable for fusion downstream of a leader peptide, such as a
 25 leader peptide including but not limited to the human tPA leader sequence. Therefore,
 any such leader peptide-based HIV-1 pol mutant construct may include but is not
 limited to a mutated DNA molecule which effectively alters the catalytic activity of
 the RT, RNase and/or IN region of the expressed protein, resulting in at least
 substantially decreased enzymatic activity one or more of the RT, RNase H and/or IN
 30 functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a
 leader peptide/HIV-1 DNA pol construct contains a mutation or mutations within the
 Pol coding region which effectively abolishes RT, RNase H and IN activity. An
 especially preferable HIV-1 DNA pol construct is a DNA molecule which contains at
 least one point mutation which alters the active site and catalytic activity within the
 35 RT, RNase H and IN domains of Pol, such that each activity is at least substantially
 abolished, and preferably totally abolished. Such a HIV-1 Pol mutant will most likely

comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. An especially preferred embodiment of this portion of the invention relates to a human tPA leader fused to the IA-Pol protein comprising the nine mutations shown in Table 1. The DNA molecule is disclosed
 5 herein as SEQ ID NO:7 and the expressed tPA-IA Pol protein comprises a fusion junction as shown in Figure 18. The complete amino acid sequence of the expressed protein is set forth in SEQ ID NO:8. To this end, SEQ ID NO:7 discloses the nucleotide sequence which codes for a human tPA leader fused to the IA Pol protein comprising the nine mutations shown in Table 1 (herein, "tPA-opt-IApol"). The open
 10 reading frame begins with the initiating Met (nucleotides 8-10) and terminates with a "TAA" codon at nucleotides 2633-2635. The nucleotide sequence encoding tPA-IAPol is also disclosed as follows:

GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGTG CTGCTGCTGT GTGGAGCAGT
 CTTCTGTTTCG CCCAGCGAGA TCTCCGCCCC CATCTCCCC ATTGAGACTG TGCCTGTGAA
 15 GCTGAAGCCT GGCATGGATG GCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT
 CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAAATCT CCAAGATTGG
 CCCCAGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG
 GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA
 GCTGGGCATC CCCCACCCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGCTGT
 20 GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC
 CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA
 GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT
 CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GCTGCCCTGT ATGTGGGCTC
 TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG
 25 GTGGGGCCTG ACCACCCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG
 CTATGAGCTG CACCCCGACA AGTGGACTGT GCAGCCCAT TGTGCTGCCTG AGAAGGACTC
 CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA
 CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA
 GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA
 30 GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA
 GCAGGGCCAG GGCCAGTGGA CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC
 TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGAAGGAGGC
 TGTGCAGAAG ATCACCCTG AGTCCATTGT GATCTGGGGC AAGACCCCA AGTTCAAGCT
 GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT
 35 CCCTGAGTGG GAGTTTGTGA ACACCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA
 GGAGCCCAT TGTGGGGCTG AGACCTTCTA TGTGGCTGGG GCTGCCAACA GGGAGACCAA

GCTGGGCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA
 CACCACCAAC CAGAAGACTG CCCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT
 GGAGGTGAAC ATTGTGACTG CCTCCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA
 TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT
 5 GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT
 GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA
 GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT
 GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA
 TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGCC TGCACCCACC TGGAGGGCAA
 10 GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC
 TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA
 GACCATCCAC ACTGCCAATG GCTCCAACCT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG
 GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCCAGTCCC AGGGGGTGGT
 GGCCTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA
 15 CCTGAAGACA GCTGTGCAGA TGGCTGTGTT CATCCACAAC TTCAAGAGGA AGGGGGGCAT
 CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA
 GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG
 GAACCCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT
 CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA
 20 TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC
 GGGCAGATCT (SEQ ID NO:7).

The open reading frame of the tPA-IA-pol construct disclosed as SEQ ID NO:7 contains 875 amino acids, disclosed herein as tPA-IA-Pol and SEQ ID NO:8, as follows:

25 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile
 Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 30 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser
 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 35 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
 Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 5 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 10 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 15 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
 20 Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 25 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu
 30 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn
 Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala
 Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 35 Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 5 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:8) .

EXAMPLE 18

10 CODON OPTIMIZED HIV-1 NEF AND CODON OPTIMIZED HIV-1 NEF MODIFICATIONS

Codon optimized version of HIV-1 Nef and HIV-1 Nef modifications are
 essentially as described in U.S. Application Serial No. 09/738,782, filed
 December 15, 2000 and PCT International Application PCT/US00/34162, also filed
 15 December 15, 2000, both documents which are hereby incorporated by reference. As
 disclosed within the above-mentioned documents, particular embodiments of codon
 optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef
 from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a
 mammalian system such as a human. The DNA molecule which encodes this protein
 20 is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is
 disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding
 regions for use in the adenoviral vectors of the present invention comprise a codon
 optimized DNA molecule encoding a protein containing the human plasminogen
 activator (tpa) leader peptide fused with the NH₂-terminus of the HIV-1 Nef
 25 polypeptide. The DNA molecule which encodes this protein is disclosed herein as
 SEQ ID NO:11, while the expressed open reading frame is disclosed herein as SEQ
 ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule
 encoding optimized HIV-1 Nef wherein the open reading frame codes for
 modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and
 30 substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein
 described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is
 disclosed herein as SEQ ID NO:13, while the expressed open reading frame is
 disclosed herein as SEQ ID NO:14. An additional embodiment relates to a DNA
 molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation
 35 site and dileucine motif have been deleted, as well as comprising a tPA leader peptide.
 This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which

encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEQ ID NO:16.

5 As disclosed in the above-identified documents (U.S. Application Serial No. 09/738,782 and PCT International Application PCT/US00/34162) and reiterated herein, the following nef-based nucleotide and amino acid sequences which comprise the respective open reading frame are as follows:

1. The nucleotide sequence of the codon optimized version of HIV-1 jfrl nef gene is disclosed herein as SEQ ID NO:9, as shown herein:

10 GATCTGCCAC CATGGGCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA
GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG CCGACAGGGT GAGGAGGACC GAGCCCGCCG
CCGTGGGCGT GGGCGCCGTG TCCAGGGACC TGGAGAAGCA CGGCGCCATC ACCTCCTCCA
ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG
15 GCTTCCCCGT GAGGCCCCAG GTGCCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC
TGTCCCCTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC
AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGGCTA CTTCCCCGAC TGGCAGAACT
ACACCCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCGG CTGGTGCTTC AAGCTGGTGC
CCGTGGAGCC CGAGAAGGTG GAGGAGGCCA ACGAGGGCGA GAACAACTGC CTGCTGCACC
20 CCATGTCCCA GCACGGCATC GAGGACCCCG AGAAGGAGGT GCTGGAGTGG AGGTTGACT
CCAAGCTGGC CTTCCACCAC GTGGCCAGGG AGCTGCACCC CGAGTACTAC AAGGACTGCT
AAAGCCCGGG C (SEQ ID NO:9).

Preferred codon usage is as follows: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG);
25 Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which is hereby incorporated by reference. See also Figure 19A-B for a comparison of wild type vs. codon optimized nucleotides comprising the open reading frame of HIV-Nef.

30 The open reading frame for SEQ ID NO:9 above comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The open reading frame of SEQ ID NO:9 provides for a 216 amino acid HIV-1 Nef protein expressed through utilization of a codon optimized DNA vaccine vector. The 216 amino acid HIV-1 Nef (jfrl) protein is disclosed herein as SEQ ID
35 NO:10, and as follows:

Met Gly Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val

Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg
 Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu
 Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp
 Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val
 5 Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp
 Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His
 Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln
 Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg
 Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro
 10 Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Leu Leu His
 Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu
 Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu
 His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:10).

HIV-1 Nef is a 216 amino acid cytosolic protein which associates with the
 15 inner surface of the host cell plasma membrane through myristylation of Gly-2
 (Franchini et al., 1986, *Virology* 155: 593-599). While not all possible Nef functions
 have been elucidated, it has become clear that correct trafficking of Nef to the inner
 plasma membrane promotes viral replication by altering the host intracellular
 environment to facilitate the early phase of the HIV-1 life cycle and by increasing the
 20 infectivity of progeny viral particles. In one aspect of the invention regarding
 codon-optimized, protein-modified polypeptides, the nef-encoding region of the
 adenovirus vector of the present invention is modified to contain a nucleotide
 sequence which encodes a heterologous leader peptide such that the amino terminal
 region of the expressed protein will contain the leader peptide. The diversity of
 25 function that typifies eukaryotic cells depends upon the structural differentiation of
 their membrane boundaries. To generate and maintain these structures, proteins must
 be transported from their site of synthesis in the endoplasmic reticulum to
 predetermined destinations throughout the cell. This requires that the trafficking
 proteins display sorting signals that are recognized by the molecular machinery
 30 responsible for route selection located at the access points to the main trafficking
 pathways. Sorting decisions for most proteins need to be made only once as they
 traverse their biosynthetic pathways since their final destination, the cellular location
 at which they perform their function, becomes their permanent residence.
 Maintenance of intracellular integrity depends in part on the selective sorting and
 35 accurate transport of proteins to their correct destinations. Defined sequence motifs
 exist in proteins which can act as 'address labels'. A number of sorting signals have

been found associated with the cytoplasmic domains of membrane proteins. An effective induction of CTL responses often required sustained, high level endogenous expression of an antigen. As membrane-association via myristylation is an essential requirement for most of Nef's function, mutants lacking myristylation, by glycine-to-
5 alanine change, change of the dileucine motif and/or by substitution with a tpa leader sequence as described herein, will be functionally defective, and therefore will have improved safety profile compared to wild-type Nef for use as an HIV-1 vaccine component.

In another embodiment of this portion of the invention, either the DNA vector
10 or the HIV-1 nef nucleotide sequence is modified to include the human tissue-specific plasminogen activator (tPA) leader. As shown in Figure 16A-B, a DNA vector may be modified by known recombinant DNA methodology to contain a leader signal peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Nef
15 protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading frame for the Nef protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1
20 Nef protein of interest, including but not limited to a HIV-1 Nef protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17).

It has been shown that myristylation of Gly-2 in conjunction with a dileucine motif in the carboxy region of the protein is essential for Nef-induced down
25 regulation of CD4 (Aiken et al., 1994, *Cell* 76: 853-864) via endocytosis. It has also been shown that Nef expression promotes down regulation of MHCI (Schwartz et al., 1996, *Nature Medicine* 2(3): 338-342) via endocytosis. The present invention relates in part to DNA vaccines which encode modified Nef proteins altered in trafficking and/or functional properties. The modifications introduced into the adenoviral vector
30 HIV vaccines of the present invention include but are not limited to additions, deletions or substitutions to the nef open reading frame which results in the expression of a modified Nef protein which includes an amino terminal leader peptide, modification or deletion of the amino terminal myristylation site, and modification or deletion of the dileucine motif within the Nef protein and which alter
35 function within the infected host cell. Therefore, a central theme of the DNA molecules and recombinant adenoviral HIV vaccines of the present invention is (1)

host administration and intracellular delivery of a codon optimized nef-based
 adenoviral HIV vaccine; (2) expression of a modified Nef protein which is
 immunogenic in terms of eliciting both CTL and Th responses; and, (3) inhibiting or
 at least altering known early viral functions of Nef which have been shown to
 5 promote HIV-1 replication and load within an infected host. Therefore, the nef
 coding region may be altered, resulting in a DNA vaccine which expresses a modified
 Nef protein wherein the amino terminal Gly-2 myristylation residue is either deleted
 or modified to express alternate amino acid residues. Also, the nef coding region may
 be altered so as to result in a DNA vaccine which expresses a modified Nef protein
 10 wherein the dileucine motif is either deleted or modified to express alternate amino
 acid residues. In addition, the adenoviral vector HIV vaccines of the present
 invention also relate to an isolated DNA molecule, regardless of codon usage, which
 expresses a wild type or modified Nef protein as described herein, including but not
 limited to modified Nef proteins which comprise a deletion or substitution of Gly 2, a
 15 deletion or substitution of Leu 174 and Leu 175 and/or inclusion of a leader sequence.

Therefore, specific Nef-based constructs further include the following, as
 exemplification's and not limitations. For example, the present invention relates to an
 adenoviral vector vaccine which encodes modified forms of HIV-1, an open reading
 frame which encodes a Nef protein which comprises a tPA leader sequence fused to
 20 amino acid residue 6-216 of HIV-1 Nef (jfr1) is referred to herein as opt tpanef. The
 nucleotide sequence comprising the open reading frame of opt tpanef is disclosed
 herein as SEQ ID NO:11, as shown below:

CATGGATGCA ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
 TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG
 25 GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG
 CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC
 CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCAGGAG GACGAGGAGG TGGGCTTCCC
 CGTGAGGCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GGCGCCGTGG ACCTGTCCCA
 CTTCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT
 30 CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC
 CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA
 GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCCTGCTGC ACCCATGTC
 CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTGCTGGAG TGGAGGTTTCG ACTCCAAGCT
 GGCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCAGGTAC TACAAGGACT GCTAAAGCC
 35 (SEQ ID NO:11).

The open reading frame for SEQ ID NO:11 comprises an initiating methionine

residue at nucleotides 2-4 and a "TAA" stop codon from nucleotides 713-715. The open reading frame of SEQ ID NO:3 provides for a 237 amino acid HIV-1 Nef protein which comprises a tPA leader sequence fused to amino acids 6-216 of HIV-1 Nef, including the dileucine motif at amino acid residues 174 and 175. This 237 amino acid tPA/Nef (jfrl) fusion protein is disclosed herein as SEQ ID NO:12, and is shown as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro
Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala
10 Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val
Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala
Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu
Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr
Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu
15 Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp
Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro
Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu
Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn
Asn Cys Leu Leu His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu
20 Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His
Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:12).

Therefore, this exemplified Nef protein, Opt tPA-Nef, contains both a tPA leader sequence as well as deleting the myristylation site of Gly-2A DNA molecule encoding HIV-1 Nef from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a mammalian system such as a human.

In another specific embodiment of the present invention, a DNA molecule is disclosed which encodes optimized HIV-1 Nef wherein the open reading frame of a recombinant adenoviral HIV vaccine encodes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175. This open reading frame is herein described as opt nef (G2A,LLAA) and is disclosed as SEQ ID NO:13, which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The nucleotide sequence of this codon optimized version of HIV-1 jfrl nef gene with the above mentioned modifications is disclosed herein as SEQ ID NO:13, as follows:

GATCTGCCAC CATGGCCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA
 GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG CCGACAGGGT GAGGAGGACC GAGCCCGCCG
 CCGTGGGCGT GGGCGCCGTG TCCAGGGACC TGGAGAAGCA CGGCGCCATC ACCTCCTCCA
 ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG
 5 GCTTCCCCGT GAGGCCCCAG GTGCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC
 TGTCCCACTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC
 AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGCTA CTTCCCCGAC TGGCAGAACT
 ACACCCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCGG CTGGTGCTTC AAGCTGGTGC
 CCGTGGAGCC CGAGAAGGTG GAGGAGGCCA ACGAGGGCGA GAACAACTGC GCCGCCACC
 10 CCATGTCCCA GCACGGCATC GAGGACCCCG AGAAGGAGGT GCTGGAGTGG AGGTTGACT
 CCAAGCTGGC CTTCCACCAC GTGGCCAGGG AGCTGCACCC CGAGTACTAC AAGGACTGCT
 AAAGCCCGGG C (SEQ ID NO:13) .

The open reading frame of SEQ ID NO:13 encodes Nef (G2A,LLAA), disclosed herein as SEQ ID NO:14, as follows:

15 Met Ala Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val
 Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg
 Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu
 Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp
 Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val
 20 Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp
 Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His
 Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln
 Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg
 Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro
 25 Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Ala Ala His
 Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu
 Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu
 His Pro Glu Tyr Tyr Lys Asp Cys Ser (SEQ ID NO:14) .

30 An additional embodiment of the present invention relates to another DNA
 molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation
 site and dileucine motif have been deleted, as well as comprising a tPA leader peptide.
 This DNA molecule, opt tpanef (LLAA) comprises an open reading frame which
 encodes a Nef protein containing a tPA leader sequence fused to amino acid residue
 6-216 of HIV-1 Nef (jfr1), wherein Leu-174 and Leu-175 are substituted with Ala-174
 35 and Ala-175 (Ala-195 and Ala-196 in this tPA-based fusion protein). The nucleotide

sequence comprising the open reading frame of opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, as shown below:

CATGGATGCA ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
 TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG
 5 GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG
 CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC
 CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCCAGGAG GACGAGGAGG TGGGCTTCCC
 CGTGAGGCCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GCGCCGTGG ACCTGTCCCA
 CTCCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT
 10 CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC
 CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA
 GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCGCCGCC ACCCCATGTC
 CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTGCTGGAG TGGAGGTTCG ACTCCAAGCT
 GGCCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCAGTAC TACAAGGACT GCTAAAGCCC
 15 (SEQ ID NO:15).

The open reading frame of SEQ ID NO:7 encoding tPA-Nef (LLAA), disclosed herein as SEQ ID NO:16, is as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro
 20 Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala
 Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val
 Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala
 Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu
 Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr
 25 Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu
 Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp
 Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro
 Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu
 Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn
 30 Asn Cys Ala Ala His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu
 Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His
 Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:16).

An adenoviral vector of the present invention may comprise a DNA sequence,
 regardless of codon usage, which expresses a wild type or modified Nef protein as
 35 described herein, including but not limited to modified Nef proteins which comprise a
 deletion or substitution of Gly 2, a deletion or substitution of Leu 174 and Leu 175

and/or inclusion of a leader sequence. Therefore, partial or fully codon optimized DNA vaccine expression vector constructs are preferred since such constructs should result in increased host expression. However, it is within the scope of the present invention to utilize "non-codon optimized" versions of the constructs disclosed herein, especially modified versions of HIV Nef which are shown to promote a substantial cellular immune response subsequent to host administration.

Figure 20A-C show nucleotide sequences at junctions between nef coding sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine 174 and 175 are the sites involved in myristylation and dileucine motif, respectively.

EXAMPLE 19

MRKAd5Pol Construction and Virus Rescue

Construction of vector: shuttle plasmid and pre-adenovirus plasmid - Key steps performed in the construction of the vectors, including the pre-adenovirus plasmid denoted MRKAd5pol, is depicted in Figure 22. Briefly, the adenoviral shuttle vector for the full-length inactivated HIV-1 pol gene is as follows. The vector MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is a derivative of the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. The vector contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *Bgl*II site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 (or MRKpAdHVE3) pre-plasmid. The vector, similar to the original shuttle vector contains the *Pac*I site, extension to the packaging signal region, and extension to the pIX gene. The synthetic full-length codon-optimized HIV-1 pol gene was isolated directly from the plasmid pV1Jns-HIV-pol-inact(opt). Digestion of this plasmid with *Bgl* II releases the pol

gene intact (comprising a codon optimized IA pol sequence as disclosed in SEQ ID NO:3). The pol fragment was gel purified and ligated into the MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) shuttle vector at the *Bgl*III site. The clones were checked for the correct orientation of the gene by using
 5 restriction enzymes *Dra*III/*Not*I. A positive clone was isolated and named MRKpdel+hCMVmin+FL-pol+bGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdel+hCMVmin+FL-pol+bGHpA(S) was digested with restriction enzymes *Pac*I and *Bst*1107 I (or its
 10 isoschizomer, *Bst*Z107 I) and then co-transformed into *E. coli* strain BJ5183 with linearized (*Cla*I digested) adenoviral backbone plasmid, MRKpAd(E1-/E3+)ClaI. The resulting pre-plasmid originally named MRKpAd+hCMVmin+FL-pol+bGHpA(S)E3+ is now referred to as "pMRKAd5pol". The genetic structure of the resulting pMRKAd5pol was verified by PCR, restriction enzyme and DNA
 15 sequence analysis. The vectors were transformed into competent *E. coli* XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the pol transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1pol adenoviral vector is shown in Figure 26 A-AO.

20 *Generation of research-grade recombinant adenovirus* - The pre-adenovirus plasmid, pMRKAd5pol, was rescued as infectious virions in PER.C6[®] adherent monolayer cell culture. To rescue infectious virus, 12 µg of pMRKAd5pol was digested with restriction enzyme *Pac*I (New England Biolabs) and 3.3 µg was transfected per 6 cm dish of PER.C6[®] cells using the calcium phosphate co-
 25 precipitation technique (Cell Pfect Transfection Kit, Amersham Pharmacia Biotech Inc.). *Pac*I digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6[®] cells. Infected cells and media were harvested 6 -10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at ≤ -60°C. This pol containing
 30 recombinant adenovirus is referred to herein as "MRKAd5pol". This recombinant adenovirus expresses an inactivated HIV-1 Pol protein as shown in SEQ ID NO:6.

EXAMPLE 20

MRKAd5Nef Construction and Virus Rescue

35 *Construction of vector: shuttle plasmid and pre-adenovirus plasmid* - Key steps performed in the construction of the vectors, including the pre-adenovirus

plasmid denoted MRKAd5nef, is depicted in Figure 23. Briefly, as shown in Example 19 above, the vector MRKpdeIE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. It has been modified to contain the *Pac*1 site, extension to the packaging signal region, and extension to the pIX gene. It contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *Bgl*11 site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 pre-plasmid. The synthetic full-length codon-optimized HIV-1 nef gene was isolated directly from the plasmid pV1Jns/nef (G2A,LLAA). Digestion of this plasmid with *Bgl*11 releases the pol gene intact, which comprises the nucleotide sequence as disclosed in SEQ ID NO:13. The nef fragment was gel purified and ligated into the MRKpdeIE1+CMVmin+BGHpA(str.) shuttle vector at the *Bgl*11 site. The clones were checked for correction orientation of the gene by using restriction enzyme *Sca*1. A positive clone was isolated and named MRKpdeIE1hCMVminFL-nefBGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdeIE1hCMVminFL-nefBGHpA(s) was digested with restriction enzymes *Pac*1 and *Bst*1107 I (or its isoschizomer, *Bst*Z107 I) and then co-transformed into *E. coli* strain BJ5183 with linearized (*Cla*1 digested) adenoviral backbone plasmid, MRKpAd(E1/E3+)Cla1. The resulting pre-plasmid originally named MRKpdeIE1hCMVminFL-nefBGHpA(s) is now referred to as "pMRKAd5nef". The genetic structure of the resulting pMRKAd5nef was verified by PCR, restriction enzyme and DNA sequence analysis. The vectors were transformed into competent *E. coli* XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the nef transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1nef adenoviral vector is shown in Figure 27A-AM.

Generation of research-grade recombinant adenovirus - The pre-adenovirus plasmid, pMRKAd5nef, was rescued as infectious virions in PER.C6[®] adherent monolayer cell culture. To rescue infectious virus, 12 µg of pMRKAdnef was digested with restriction enzyme *Pac*1 (New England Biolabs) and 3.3 µg was transfected per 6 cm dish of PER.C6[®] cells using the calcium phosphate co-precipitation technique (Cell Pfect Transfection Kit, Amersham Pharmacia Biotech

Inc.). *Pac1* digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6[®] cells. Infected cells and media were harvested 6 -10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at $\leq -60^{\circ}\text{C}$. This nef containing recombinant adenovirus is now referred to as "MRKAd5nef".

EXAMPLE 21

Construction of Murine CMV Promoter Containing Shuttle Vectors for Inactivated Pol and Nef/G2A,LLAA

10 The murine CMV (mCMV) was amplified from the plasmid pMH4 (supplied by Frank Graham, McMaster University) using the primer set: mCMV (*Not* I) Forward: 5'-ATA AGA ATG CGG CCG CCA TAT ACT GAG TCA TTA GG-3' (SEQ ID NO: 20); mCMV (*Bgl* II) Reverse: 5'-AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C-3' (SEQ ID NO:21). The underlined nucleotides represent the *Not* I and the *Bgl* II sites respectively for each primer. This PCR amplicon was used for the construction of the mCMV shuttle vector containing the transgene in the E1 parallel orientation. The hCMV promoter was removed from the original shuttle vector (containing the hCMV-gag-bGHpA transgene in the E1 parallel orientation) by digestion with *Not* I and *Bgl* II. The mCMV promoter (*Not* I/*Bgl* II digested PCR product) was inserted into the shuttle vector in a directional manner. The shuttle vector was then digested with *Bgl* II and the gag reporter gene (*Bgl* II fragment) was re-inserted back into the shuttle vector. Several clones were screened for correct orientation of the reporter gene. For the construction of the mCMV-gag in the E1 antiparallel orientation, the mCMV promoter was amplified from the plasmid pMH4 using the following primer set: mCMV (*Asc* I) Forward: 5'- ATA AGA ATG GCG CGC CAT ATA CTG AGT CAT TAG G (SEQ ID NO:22); mCMV (*Bgl* II) Reverse: 5' AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C (SEQ ID NO:23). The underlined nucleotides represent the *Asc* I and *Bgl* II sites, respectively for each primer. The shuttle vector containing the hCMV-gag transgene in the E1 antiparallel orientation was digested with *Asc* I and *Bgl* II to remove the hCMV-gag portion of the transgene. The mCMV promoter (*Asc* I/*Bgl* II digested PCR product) was inserted into the shuttle vector in a directional manner. The vector was then digested with *Bgl* II and the gag reporter gene (*Bgl* II fragment) was re-inserted. Several clones were screened for correct orientation of the reporter gene. For each of the full length IA pol and full length nef/G2A,LLAA genes, cloning was performed using the unique

Bgl II site within the mCMV-bGHpA shuttle vector. The pol and nef genes were excised from their respective pV1Jns plasmids by *Bgl* II digestion.

EXAMPLE 22

5 Construction of mCMV Full Length Inactivated Pol and Full Length nef/G2A.LLAA Adenovectors

Each of these transgenes of Example 21 were inserted into the modified shuttle vector in both the E1 parallel and E1 anti-parallel orientations. *Pac*I and *Bst*Z110I digestion of each shuttle vector was performed and each specific transgene
10 fragment containing the flanking Ad5 sequences was isolated and co-transformed with *Cla* I digested MRKpAd5(E3+) or MRKpAd5(E3-) adenovector plasmids via bacterial homologous recombination in BJ5183 *E. coli* cells. Recombinant pre-plasmid adenovectors containing the various transgenes in both the E3- and E3+ versions (and in the E1 parallel and E1 antiparallel orientations) were subsequently
15 prepared in large scale following transformation into XL-1 Blue *E. coli* cells and analyzed by restriction analysis and sequencing.

EXAMPLE 23

Construction of hCMV-tpa-nef (LLAA) Adenovector

20 The tpa-nef gene was amplified out from GMP grade pV1Jns-tpanef (LLAA) vector using the primer sets: Tpanef (*Bam*HI) F 5'-ATT GGA TCC ATG GAT GCA ATG AAG AGA GGG (SEQ ID 24); Tpanef (*Bam*HI) R 5'-ATA GGA TCC TTA GCA GTC CTT GTA GTA CTC G (SEQ ID NO:25). The resulting PCR product was digested with *Bam*HI, gel purified and cloned into the *Bgl* II site of MRKAd5CMV-bGHpA shuttle vector (*Bgl* II digested and calf intestinal phosphatase treated).
25 Clones containing the tpanef (LLAA) gene (see SEQ ID NO:15 for complet coding region) in the correct orientation with respect to the hCMV promoter were selected following *Sca* I digestion. The resulting MRKAd5tpanef shuttle vector was digested with *Pac* I and *Bst* Z1101 and cloned into the E3+ MRKAd5 adenovector via bacterial
30 homologous recombination techniques.

EXAMPLE 24

Immunogenicity of MRKAd5pol and MRKAd5nef Vaccine

Materials and Methods - Rodent Immunization - Groups of N=10 BALB/c
35 mice were immunized i.m. with the following vectors: (1) MRKAd5hCMV-IApol (E3+) at either 10^7 vp and 10^9 vp; and (2) MRKAd5hCMV-IApol (E3-) at either

10⁷ vp and 10⁹ vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second dose, sera and spleens were collected from all the animals for RT ELISA and IFN γ ELIspot analyses, respectively. For all rodent immunizations, the Ad5 vectors were
 5 diluted in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl₂, 0.005% polysorbate 80, pH 8.0. The total dose was injected to both quadricep muscles in 50 μ L aliquots using a 0.3-mL insulin syringe with 28-1/2G needles (Becton-Dickinson, Franklin Lakes, NJ).

Groups of N=10 C57/BL6 mice were immunized i.m. with the following
 10 vectors: (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10⁷ vp and 10⁹ vp; (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10⁷ vp and 10⁹ vp; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10⁷ vp and 10⁹ vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second dose, sera and spleens were
 15 collected from all the animals for RT ELISA and IFN γ ELIspot analyses, respectively.

Non-human Primate immunization - Cohorts of 3 rhesus macaques (2-3 kg) were vaccinated with the following Ad vectors: (1) MRKAd5hCMV-IApol (E3+) at either 10⁹ vp and 10¹¹ vp dose; and (2) MRKAd5hCMV-IApol (E3-) at either
 20 10⁹ vp and 10¹¹ vp; (3) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10⁹ vp and 10¹¹ vp; and (4) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10⁹ vp and 10¹¹ vp. The vaccine was administered to chemically restrained monkeys (10 mg/kg ketamine) by needle injection of two 0.5 mL aliquots of the Ad vectors (in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl₂, 0.005% polysorbate 80, pH 8.0)
 25 into both deltoid muscles. The animals were immunized twice at a 4 week interval (T=0, 4 weeks).

Murine anti-RT and anti-nef ELISA - Anti-RT titers were obtained following standard secondary antibody-based ELISA. Maxisorp plates (NUNC, Rochester, NY) were coated by overnight incubation with 100 μ L of 1 μ g /mL HIV-1 RT protein
 30 (Advanced Biotechnologies, Columbia, MD) in PBS. For anti-nef ELISA, 100 μ L of 1 μ g/mL HIV-1 nef (Advanced Biotechnologies, Columbia, MD) was used to coat the plates. The plates were washed with PBS/0.05% Tween 20 using Titertek MAP instrument (Hunstville, AL) and incubated for 2 h with 200 μ L/well of blocking solution (PBS/0.05% tween/1% BSA). An initial serum dilution of 100-fold was
 35 performed followed by 4-fold serial dilution. 100- μ L aliquots of serially diluted samples were added per well and incubated for 2 h at room temperature. The plates

were washed and 100 μ L of 1/1000-diluted HRP-rabbit anti-mouse IgG (ZYMED, San Francisco, CA) were added with 1 h incubation. The plates were washed thoroughly and soaked with 100 μ L 1,2-phenylenediamine dihydrochloride/hydrogen peroxide (DAKO, Norway) solution for 15 min. The reaction was quenched by adding 100 μ L of 0.5M H₂SO₄ per well. OD₄₉₂ readings were recorded using Titertek Multiskan MCC/340 with S20 stacker. Endpoint titers were defined as the highest serum dilution that resulted in an absorbance value of greater than or equal to 0.1 OD₄₉₂ (2.5 times the background value).

Non-human primate and murine ELISpot assays - The enzyme-linked immuno-spot (ELISpot) assay was utilized to enumerate antigen-specific INF γ -secreting cells from mouse spleens (Miyahira, et al.1995, *J. Immunol. Methods* 181:45-54) or macaque PBMCs. Mouse spleens were pooled from 5 mice/cohort and single cell suspensions were prepared at 5x10⁶/mL in complete RPMI media (RPMI1640, 10% FBS, 2mM L-glutamine, 100U/mL Penicillin, 100 u/mL streptomycin, 10 mM Hepes, 50 uM β -ME). Rhesus PBMCs were prepared from 8-15 mL of heparinized blood following standard Ficoll gradient separation (Coligan, et al, 1998, *Current Protocols in Immunology*. John Wiley & Sons, Inc.). Multiscreen opaque plates (Millipore, France) were coated with 100 μ L/well of either 5 μ g/mL purified rat anti-mouse IFN- γ IgG1, clone R4-6A2 (Pharmingen, San Diego, CA), or 15 μ g/mL mouse anti-human IFN- γ IgG_{2a} (Cat. No. 1598-00, R&D Systems, Minneapolis, MN) in PBS at 4°C overnight for murine or monkey assays, respectively. The plates were washed with PBS/penicillin/streptomycin and blocked with 200 μ L/well of complete RPMI media for 37 °C for at least 2 h.

To each well, 50 μ L of cell samples (4-5x10⁵ cells per well) and 50 μ L of the antigen solution were added. To the control well, 50 μ L of the media containing DMSO were added; for specific responses, either selected peptides or peptide pools (4 μ g/mL per peptide final concentration) were added. For BALB/c mice immunized with the pol constructs, stimulation was conducted using a pool of CD4⁺-epitope containing 20-mer peptides (aa21-40, aa411-430, aa641-660, aa731-750, aa771-790) or a pool of CD8⁺-epitope containing peptides (aa201-220, aa311-330, aa781-800). For C57/BL6 mice immunized with the nef construct, either aa51-70 (CD8⁺ T cell epitope) or aa81-100 (CD4⁺) peptide derived from the nef sequence was added for specific stimulation. In monkeys, the responses against pol were evaluated using two pools (L and R) of 20-aa peptides that encompass the entire pol sequence and overlap by 10 amino acids. In monkeys vaccinated with the nef constructs, a single pool containing 20-mer peptides covering the entire HIV-1 nef sequence and overlapping

by 10 aa was used. Each sample/antigen mixture was performed in triplicate wells for murine samples or in duplicate wells for rhesus PBMCs. Plates were incubated at 37°C, 5% CO₂, 90% humidity for 20-24 h. The plates were washed with PBS/0.05% Tween 20 and incubated with 100 µL/well of either 1.25 µg/mL biotin-conjugated rat anti-mouse IFN-γ mAb, clone XMG1.2 (Pharmingen) or of 0.1 ug/mL biotinylated anti-human IFN-gamma goat polyclonal antibody (R&D Systems) at 4°C overnight. The plates were washed and incubated with 100 µL/well 1/2500 dilution of streptavidin-alkaline phosphatase conjugate (Pharmingen) in PBS/0.005% Tween/5% FBS for 30 min at 37 °C. Spots were developed by incubating with 100 µL/well 1-step NBT/BCIP (Pierce Chemicals) for 6-10 min. The plates were washed with water and allowed to air dry. The number of spots in each well was determined using a dissecting microscope and the data normalized to 10⁶ cell input.

Non-human Primate anti-RT ELISA - The pol-specific antibodies in the monkeys were measured in a competitive RT EIA assay, wherein sample activity is determined by the ability to block RT antigen from binding to coating antibody on the plate well. Briefly, Maxisorp plates were coated with saturating amounts of pol positive human serum (#97111234). 250 uL of each sample is incubated with 15 uL of 266 ng/mL RT recombinant protein (in RCM 563, 1% BSA, 0.1% tween, 0.1% NaN₃) and 20 uL of lysis buffer (Coulter p24 antigen assay kit) for 15 min at room temperature. Similar mixtures are prepared using serially diluted samples of a standard and a negative control which defines maximum RT binding. 200 uL/well of each sample and standard were added to the washed plate and the plate incubated 16-24 h at room temperature. Bound RT is quantified following the procedures described in Coulter p24 assay kit and reported in milliMerck units per mL arbitrarily defined by the chosen standard.

Results - Rodent Studies - BALB/c mice (n=5 mice/cohort) were immunized once or twice with varying doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). At 3 weeks after the second dose, Anti-pol IgG levels were determined by an ELISA assay using RT as a surrogate antigen. Cellular response were quantified via IFNγ ELISpot assay against pools of pol-epitope containing peptides. The results of these assays are summarized in Table 10. The results indicate that the mouse vaccinees exhibited detectable anti-RT IgGs with an adenovector dose as low as 10⁷ vp. The humoral responses are highly dose-dependent and are boostable with a second immunization. One or two doses of either pol vectors elicit high frequencies of antigen-specific CD4⁺ and CD8⁺ T cells; the responses are weakly dose-dependent but are boostable with a second immunization.

Table 10. Immunogenicity of MRKAd5pol Vectors in BALB/c mice.

Group	Vaccine	Dose	No. of Doses	Anti-RT IgG Titers ^a			SFC/10 ⁶ cells ^c		
				GMT	+SE	-SE	Medium	CD4+ peptide pool	CD8+ peptide pool
1	MRKAd5hCMVFLpol (E3+)	10 ⁷ vp	2 1	310419 919	301785 372	153020 265	1(1) 1(1)	75(4) 72(9)	2313(67) 533(41)
2	MRKAd5hCMVFLpol (E3+)	10 ⁹ vp	2 1	1638400 ^b 713155	0 528520	0 303555	2(2) 1(1)	114(9) 48(7)	2063(182) 733(89)
3	MRKAd5hCMVFLpol (E3-)	10 ⁷ vp	2 1	310419 6400	386218 14013	172097 4393	0(0) 10(8)	223(7) 141(21)	2607(27) 409(28)
4	MRKAd5hCMVFLpol (E3-)	10 ⁹ vp	2 1	1638400 ^b 1241675 ^b	0 396725	0 300661	1(1) 0(0)	160(13) 39(13)	2385(11) 833(83)
5	Naïve	none	none	57	9	7	9(2)	11(4)	10(1)

^aGMT, geometric mean titer of the cohort of 5 mice; SE, standard error of the geometric mean^bNear or at the upper limit of the serial dilution; hence, could be greater than this value^cNo. of Spot-forming Cells per million splenocytes; mean values of triplicates are reported along with standard errors in parenthesis.

- 5 C57/BL6 mice were immunized once or twice with varying doses of MRKAd5hCMV-nef(G2A,LLAA) (E3+), MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10⁷ vp and (3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10⁷ vp and 10⁹ vp. The immune response were analyzed using similar protocols and the results are listed in Table 11. While anti-nef IgG responses could not be detected in this
- 10 model system with any of the constructs, there are strong indications of a cellular immunity generated against nef using the ELIspot assay.

Table 11. Immunogenicity of MRKAd5nef Vectors in C57/BL6 mice.

Group	Vaccine	Dose	No. of Doses	Anti-nef IgG Titers ^a			SFC/10 ⁶ cells ^b		
				GMT	+SE	-SE	Medium	aa51-70 CD8+	aa81-100 CD4+
1	MRKAd5hCMVFLnef (E3+)	10 ⁷ vp	2 1	174 132	70 42	50 32	1(1) 0(0)	23(1) 0(0)	1(1) 0(0)
2	MRKAd5hCMVFLnef (E3+)	10 ⁹ vp	2 1	174 132	70 42	50 32	0(0) 1(1)	61(7) 62(7)	4(2) 3(1)
3	MRKAd5mCMVFLnef (E3+)	10 ⁷ vp	2 1	132 115	42 46	32 33	3(1) 3(2)	15(5) 3(2)	5(2) 4(2)
4	MRKAd5mCMVFLnef (E3+)	10 ⁹ vp	2 1	132 132	42 42	32 32	4(2) 2(1)	83(13) 29(2)	5(1) 4(0)
5	MRKAd5mCMVtpanef(E3+)	10 ⁷ vp	2 1	132 100	42 0	32 0	3(2) 3(1)	14(2) 13(4)	5(1) 10(3)
6	MRKAd5mCMVtpanef(E3+)	10 ⁹ vp	2 1	230 115	170 46	98 33	3(2) 7(1)	145(29) 151(14)	4(0) 10(0)
7	Naïve	none	none	152	78	52	21(2)	18(6)	26(3)

^aGMT, geometric mean titer of the cohort of 5 mice; SE, standard error of the geometric mean^bNo. of spot-forming cells per million splenocytes; mean values of triplicates are reported along with standard errors in parenthesis.

15

Monkey Studies - Cohorts of 3 rhesus macaques were immunized with 2 doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). The number of antigen-specific T cells (per million PBMCs) were enumerated using one of two

peptide pools (L and R) that cover the entire pol sequence; the results are listed in Table 12. Moderate-to-strong T cell responses were detected in the vaccinees using either constructs even at a low dose of 10^9 vp. Longitudinal analyses of the anti-RT antibody titers in the animals suggest that the pol transgene product is expressed efficiently to elicit a humoral response (Table 13). It would appear that generally higher immune responses were observed in animals that received the E3- construct compared to the E3+ virus.

Table 12. Pol-specific T Cell Responses in MRKAd5pol Immunized Rhesus Macaques.

Vaccine (T=0,4 wks)	Monkey #	Prebleed			T=4			T=7			T=16		
		Mock	Pol L	Pol R	Mock	Pol L	Pol R	Mock	Pol L	Pol R	Mock	Pol L	Pol R
MRKAd5hCMV-IAPol(E3+) 10^{11} vp	99C100	1	0	0	1	38	31	0	52	146	0	49	715
	99C215	1	2	2	10	98	249	1	109	305	22	88	250
	99D201	5	5	4	6	149	85	0	40	35	0	35	18
MRKAd5hCMV-IAPol(E3+) 10^9 vp	99D212	0	2	0	4	331	114	0	58	14	0	6	6
	99D180	0	4	2	0	19	192	4	36	156	5	38	108
	99C201	8	5	21	6	62	62	0	18	32	1	14	65
MRKAd5hCMV-IAPol(E3-) 10^{11} vp	99D239	5	2	2	20	82	172	1	68	114	9	21	40
	99C186	4	12	6	5	120	421	2	271	489	16	875	530
	99C084	1	8	9	8	84	464	0	14	236	1	24	264
MRKAd5hCMV-IAPol(E3-) 10^9 vp	CC7C	10	10	8	12	724	745	4	322	376	4	188	176
	CD1G	2	0	1	5	474	468	0	232	212	0	101	121
	CD11	6	6	12	10	98	110	5	60	80	8	25	34
Naïve	083Q	nd	nd	nd	nd	nd	nd	4	2	2	2	1	2

nd, not determined
Reported are SFC per million PBMCs; mean of duplicate wells.

Table 13. Anti-RT Ig Levels in MRKAd5pol Immunized macaques.

RT ANTIBODY ASSAY TITERS IN mMU/mL				
Vaccine/Monkey Tag	T=4	T=7	T=12	T=16
MRKAd5hCMV-IAPol(E3+), 10^{11} vp				
99C100	61	1999	5928	4768
99C215	81	1541	2356	2767
99D201	53	336	539	387
MRKAd5hCMV-IAPol(E3+), 10^9 vp				
99D212	10	40	49	68
99D180	<10	36	79	93
99C201	<10	37	71	76
MRKAd5hCMV-IAPol(E3-), 10^{11} vp				
99D239	44	460	1234	1015
99C186	21	233	480	345
99C084	235	2637	2858	1626
MRKAd5hCMV-IAPol(E3-), 10^9 vp				
CC7C	32	175	306	235
CD1G	20	140	273	419
CD11	15	112	149	237

When rhesus macaques were immunized i.m. with two doses of MRKAd5nef constructs, vigorous T cell responses ranging from 100 to as high as 1100 per million were observed in 8 of 12 vaccinees (Table 14). The efficacies of the mCMV- and hCMV- driven nef constructs are comparable on the basis of the data generated thus far.

Table 14. Nef-specific T cell Responses in MRKAd5nef Immunized Rhesus Macaques.

Vaccine (T=0,4 wks)	Monk #	Pre		T=4		T=7		T=16	
		Mock	Nef	Mock	Nef	Mock	Nef	Mock	Nef
MRKAd5hCMV-nef(G2A,LLAA) (E3+) 10 ¹¹ vp	CD2D	0	4	31	440	4	368	1	251
	CC7B	0	0	2	521	0	178	1	1522
	CC61	2	9	31	112	0	108	11	100
MRKAd5hCMV-nef(G2A,LLAA) (E3+) 10 ⁹ vp	CC2K	9	9	6	52	0	35	0	15
	CD15	5	4	30	998	2	586	0	434
	CD16	6	1	6	1146	0	369	1	212
MRKAd5mCMV-nef(G2A,LLAA) (E3+) 10 ¹¹ vp	99D191	1	5	4	614	0	298	2	419
	99D144	4	6	5	434	0	1100	2	932
	99C193	1	2	1	58	1	22	0	64
MRKAd5mCMV-nef(G2A,LLAA) (E3+) 10 ⁹ vp	99D224	1	11	14	231	1	125	0	70
	99D250	8	9	4	108	0	54	0	5
	99C120	1	6	20	299	0	92	0	79
Naive	083Q	nd	nd	18	22	4	5	2	1

EXAMPLE 25

Comparison of Clade B vs. Clade C T Cell Responses in HIV-Infected Subjects

PBMC samples collected from two dozens of patients infected with HIV-1 in US were tested in ELISPOT assays with peptide pools of 20-mer peptides overlapping by 10 amino acids. Four different peptide pools were tested for cross-clade recognition, and they were either derived from a clade B-based isolate (gag H-b; nef-b) or a clade C-based isolate (gag H-c, nef-c). Data in Table 15 shows that T cells from these patients presumably infected with clade B HIV-1 could recognize clade C gag and nef antigens in ELISPOT assay. Correlation analysis further demonstrated that these T cell responses against clade C gag peptide pool were about 60% of the clade B counterpart (Figure 24), while the T cell responses against clade C nef were about 85% of the clade B counterpart (Figure 25). These results suggest that cellular immune responses generated in patients infected with clade B HIV-1 can recognize gag and nef antigens derived from clade C HIV-1. These data show that a HIV vaccine, such as a DNA or MRKAd5-based adenoviral vaccine expressing a clade B

gag and/or nef antigen will potentially have the ability to provide a prophylactic and/or therapeutic advantage on a global scale.

5

Table 15

Responses Shown as the Number of gIFN-Secreting T Cells per Million PBMCs

subject	bleed date	gag epitope # (from mapping)	mock	gag H-b	gagH-c	nef-b	nef-c
#100	19-Jul-99	12	10	3950	1385	1295	1300
#101	25-Jul-99	3	15	3885	1280	na	1020
#102	25-Jul-99	4	15	1740	850	1255	1785
#104	7-Jun-99	2	5	1355	1185	na	1060
#107	11-Oct-99	2	25	3305	2795	670	870
#405	11-Jul-99	2	15	4575	3180	1700	1500
#501	19-Jul-99	2	15	1100	570	3365	3460
#505	18-Jul-99	5	10	2145	1725	1235	na
#506	28-Feb-99	2	25	150	45	400	610
#701	28-Mar-99	5	30	7620	4775	3320	2780
#709	17-May-99	3	15	2785	1945	1090	1630
#710	24-May-99	4	5	1055	1080	2210	2140

10

EXAMPLE 26

Characterization and Production of MRKAd5pol and MRKAd5nef
Vectors in Roller Bottles

Expansion of nef and pol Adenovectors - Nef and pol CsCl purified MRKAd5 seeds were used to infect roller bottles to produce P4 virus to be used as a seed for further experiments. P4 MRKAd5 pol and nef vectors were used to infect roller bottles at an MOI 280 vp/cell, except for hCMV-tpa-nef [E3+] which was infected at an MOI of 125 due to low titers of seed obtained at P4.

20

Table 16 Viral particle concentrations for P5 nef and pol adenovectors

Adenovector	AEX Titer (10 ¹⁰ vp/ml culture)	AEX Titer (10 ⁴ vp/cell)	Amplification Ratio
hCMV-FL-nef [E3+]	1.1	0.9	30
mCMV-FL-nef [E3+]	2.2	2.1	75
hCMV-tpa-nef [E3+]	0.07	0.1	5
mCMV-tpa-nef [E3+]	1.3	0.9	35
hCMV-FL-pol [E3+]	2.7	2.1	75
hCMV-FL-pol [E3-]	1.9	1.3	45

5 *Roller Bottle Passaging* - Passaging of the *pol* and *nef* constructs continued through passage seven. Cell-associated (freeze/thaw lysis) and whole broth (triton-lysis) titers obtained in all passages were very consistent. In general, MRKAd5pol is ca. 70% as productive as MRKAd5gag while MRKAd5nef is ca. 25% as productive as MRKAd5gag. Samples of P7 virus for both constructs were analyzed by V&CB by
10 restriction digest analysis and did not show any rearrangements.

Table 17. Passage Six Viral Productivity for MRKAd5pol and MRKAd5nef

		Xviable (10^6 cells/ml), Viability (%)		Cell Passage Number	AEX Titer (Cell Associated) 10^{10} vp/ml culture	Titer 10^4 vp/cell	Amplification Ratio	Triton Lysis Titer 10^{10} vp/ml culture
		Infection	Harvest					
hCMV-FL-nef [E3+]	pool	1.22, 85%		62	0.8	0.7	25	1.6
	1		0.99, 62%					
	2		1.10, 72%					
hCMV-FL-pol [E3+]	pool	1.42, 89%		62	4.5	3.2	115	7.0
	1		1.22, 70%					
	2		1.42, 74%					

15 Table 18. Passage Seven Viral Productivity for MRKAd5pol and MRKAd5nef

		Xviable (10^6 cells/ml), Viability (%)		Cell Passage Number	AEX Titer (Cell Associated) 10^{10} vp/ml culture	Titer 10^4 vp/cell	Amplification Ratio	Triton Lysis Titer 10^{10} vp/ml culture
		Infection	Harvest					
hCMV-FL-nef [E3+]	Pool	1.33, 90%		66	1.0	0.8	29	2.1
	1		0.96, 70%					
	2		1.18, 73%					
hCMV-FL-pol [E3+]	Pool	0.90*, 90%		56	4.2	4.7	168	6.5
	1		1.18, 88%					
	2		1.04, 80%					

MRKAd5nef and MRKAd5pol Viral Production Kinetics - A timecourse experiment was carried out in roller bottles to determine if the viral production kinetics of the MRKAd5pol and MRKAd5nef vectors were similar to those of MRKAd5gag. PER.C6[®] cells in roller bottle cultures were infected at an MOI of 280
20 vp/cells with P5 MRKAd5pol, P5 MRKAd5nef and P7 MRKAd5gag; for each adenovector, two infected bottles were sampled at 24, 36, 48, and 60 hours post infection. In addition, two bottles were left unsampled until 48 hpi when they were harvested under the Phase I process conditions. The anion-exchange HPLC viral
25 particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36,

48, and 60 hpi timepoints are shown in Figure 29A-B. The QPA titers show a similar trend (data not shown).

Comparison of hCMV- and mCMV-FL-nef - As the titers obtained with the MRKAd5nef construct (hCMV-FL-nef) were lower than those obtained with MRKAd5gag or MRKAd5pol, a viral productivity comparison experiment was performed with mCMV-FL-nef. For each of the two adenovectors (hCMV- and mCMV-FL-nef), two roller bottles were infected at an MOI of 280 vp/cell with passage five clarified lysate. The macroscopic and microscopic observations of the four roller bottles were identical at the time of harvest. Analysis of the clarified lysate produced indicated a higher viral particle concentration in the bottles infected with mCMV-FL-nef, as shown in Table 19. It is stipulated that the higher productivity with mCMV promoter driven nef vector is due to lower nef expression levels in PER.C6[®] cells- experiments are underway at V&CB to measure nef expression levels.

Table 19. Passage Six Viral Productivity Comparison of hCMV- and mCMV-FL-nef

		Xv (10 ⁶ cells/ml), Viability (%)		Cell Passage	AEX Titer	Titer	Amplification	Triton Lysis Titer
		Infection	Harvest	Number	10 ¹⁰ vp/ml culture	10 ⁴ vp/cell	Ratio	10 ¹⁰ vp/ml culture
hCMV-FL-nef (MRKAd5nef)	Pool	1.11, 91%		60	1.5	1.4	50	2.8
	1		1.23, 75%					
	2		1.34, 74%					
mCMV-FL-nef	Pool	1.11, 91%		60	2.3	2.1	75	4.6
	1		1.49, 84%					
	2		1.18, 77%					

20

EXAMPLE 27

Characterization and Large Scale Production of MRKAd5nef Virus in Bioreactors

Materials and Methods - The experiment of the present example was run twice under the following conditions: 36.5°C, DO 30%, pH 7.30, 150rpm agitation rate, no sparging, Life Technologies (Gibco, Invitrogen) 293 SFM II (with 6mM L-glutamine), 0.5M NaOH as base for pH control. During the first run (B20010115), two 10L stirred vessel bioreactors were inoculated with PER.C6[®] cells at a concentration of 0.2x10⁶ cells/ml. Cells were grown until they reached a cell concentration of approximately 1x10⁶ cells/ml. The cells were infected with uncloned MRKAd5nef (G2A,LLAA) at a MOI of 280 virus particles (vp)/cell. For the second batch (B20010202), the same procedure as the first run was used, except the cells

were infected with cloned MRAd5nef. During both runs, the bioreactors were harvested 48 hours post-infection. Samples were taken and virus concentrations were determined from whole broth (with triton lysis), supernatant, and cell pellets (3 X freeze/thaw) with the AEX and QPA assays. Metabolites were measured with BioProfile 250 throughout the process.

Table 20: Experimental Conditions

Temperature	36.5 °C
DO	30%
PH	7.30
Agitation	150 rpm
Sparging	None

Table 21: Virus source used for experiments.

Run	Batch ID	Cloned/Uncloned MRKAd5nef	MOI (vp/cells)
#1	B20010115-1	Uncloned	280
	B20010115-2	Uncloned	280
#2	B20010202-1	Cloned	280
	B20010202-2	Cloned	280

Results - Table 22 and 23 show an the ability to scale up production of MRKAd5nef by growth in a bioreactor.

Table 22: Virus Concentration as measured by the AEX assay

Run	Batch ID	Cloned/Uncloned MRKAd5nef	Virus Concentration @ 48hpi (1×10^{13} vp/L)			
			Supernatant	Clarified Lysate	Total	Triton Lysate
#1	B20010115-1	Uncloned	0.72	3.26	3.98	5.76
	B20010115-2	Uncloned	0.38	1.67	2.05	2.46
#2	B20010202-1	Cloned	0.80	6.00	6.80	8.88
	B20010202-2	Cloned	0.50	6.00	6.50	8.47

Table 23: Virus Titers as measured by the QPA assay

Run	Batch ID	Cloned/Uncloned MRKAd5nef	Virus Concentration @ 48hpi (1×10^{11} IU/L)				
			Whole Broth	Supernatant	Clarified Lysate	Total	Triton Lysate
#1	B20010115-1	Uncloned	0.13	1.12	1.76	2.88	11.28
	B20010115-2	Uncloned	0.14	0.73	1.54	2.27	5.86
#2	B20010202-1	Cloned	0.14	0.97	1.62	2.69	11.89
	B20010202-2	Cloned	0.14	1.17	1.70	2.97	12.47

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art

from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

EXAMPLE 28

5 MRKAd5HIV-1gag Boosting of DNA-Primed Animals

Groups of 3-5 rhesus macaques were immunized with (a) 5 mgs of V1Jns-Flgag (pVIJnsCMV(no intron)-FL-gag-bGHpA), (b) 5 mgs of V1Jns-Flgag formulated with 45 mgs of a non-ionic block copolymer CRL1005, or (c) 5 mgs of
10 V1Jns-Flgag formulated with 7.5 mgs of CRL1005 and 0.6 mM benzalkonium chloride at weeks 0, 4, and 8. All animals received a single dose of 10^7 viral particles (vp) of the MRKAd5HIV-1gag at week 26. Note: 10^7 is too low to prime or boost effectively when used as a single modality (dose is selected to mimic preexposure to adenovirus); see Figure 32.

15 Blood samples were collected from all animals at several time points and peripheral blood mononuclear cells (PBMCs) were prepared using standard Ficoll method. The PBMCs were counted and analyzed for gamma-interferon secretion using the ELISpot assay (Table 24). For each monkey, the PBMCs were incubated overnight either in the absence (medium) or presence of a pool (called "gag H") of 50
20 20-aa long peptides that encompass the entire HIV-1 gag sequence.

The results indicate that MRKAd5HIV-1gag was very effective in boosting the T cell immune responses in these monkeys. At week 28 or 2 weeks after the viral boost, the number of gag-specific T cells per million PBMCs increased 2-48 fold compared to the levels observed at week 24 or 2 weeks prior to the boost.

25 The PBMCs were also analyzed by intracellular gamma-interferon staining prior to (at week 10) and after the MRKAd5gag boost (at week 30). The results for select animals are shown on Figure 31. The results indicate that (a) immunization with DNA/adjuvant formulation elicited T cell responses which can either be balanced, $CD4^+$ -biased or $CD8^+$ -biased, and (b) boosting with the MRKAd5gag
30 construct produced in all cases a strongly $CD8^+$ -biased response. These results suggest that boosting with MRKAd5HIV-1gag construct is able to improve the levels of antigen-specific $CD8^+$ T cells.

Table 24. Boosting of DNA/Adjuvant-Primed Rhesus Monkeys with MRKAd5gag

Number of SFC/million PBMCs

Grp#	Priming T=0, 4, 8 wks	Boost T=26 wks	Monkey	T=0		T=4		T=6		T=10		T=17		T=24		T=28		T=30	
				Medium	gag H	Medium	gag H	Medium	gag H	Medium	gag H	Medium	gag H	Medium	gag H	Medium	gag H	Medium	gag H
1	DNA/5 mgs PBS (D101)	MRKAd5gag(E3+) 10 ⁷ vp	CB5H CC6X AW3G	NA 0 5	NA 0 11	3 0 0	35 15 36	15 0 3	71 46 51	4 0 3	224 68 46	8 0 2	115 75 89	6 0 8	85 35 65	19 3 10	956 1705 989	0 1 0	316 755 395
2	DNA/5mgs + CRL1005/45mgs	MRKAd5gag(E3+) 10 ⁷ vp	CC1C CC1K AW3P CB5F AK8B	0 4 9 NA 9	4 0 8 NA 12	1 1 1 0 4	60 101 10 31 36	0 0 4 0 1	111 254 71 288 119	5 0 4 0 0	270 791 154 530 439	4 5 8 19 0	280 452 104 374 425	8 0 5 9 0	232 321 85 251 316	3 0 11 8 4	959 1915 836 1549 1229	19 1 6 20 5	1345 1099 241 1734 1354
3	DNA/5 mgs + CRL1005/7.5 mgs + 0.6 mM BAK	MRKAd5gag(E3+) 10 ⁷ vp	AW20 CA4R CB58 CB5W CB7D	10 1 8 4 1	4 0 6 3 0	1 3 0 0 0	59 121 6 26 136	5 1 3 1 0	264 135 119 91 316	19 1 0 0 1	425 270 274 139 609	6 5 6 0 5	105 130 282 164 626	9 1 1 1 1	205 105 208 62 759	18 14 0 5 0	565 1384 636 543 2278	8 10 1 1 4	404 978 828 349 1831
4	none	None	98D201	3	0	0	0	1	0	0	0	0	1	1	2	3	0	0	0

NA, not available

EXAMPLE 29

Construction of gagpol fusion for MRKAd5gagpol fusion constructs

The open reading frames for the codon-optimized HIV-1 gag gene was fused directly to the open reading frame of the IA pol gene (consisting of RT, RNaseH and integrase domains) by stepwise PCR. Because the gene (SEQ ID NO: 38) does not include the protease gene and the frameshift sequence, it encodes a single polypeptide of the combined size of p55, RT, RNase H and integrase (1350 amino acids; SEQ ID NO: 39).

The fragment that extends from the BstEII site within the gag gene to the last non-stop codon was ligated via PCR to a fragment that extends from the start codon of the IApol to a unique BamHI site. This fragment was digested with BstEII and BamHI. Construction of gag-IApol fusion was achieved via three-fragment ligation involving the PstI-BstEII gag digestion fragment, the BstEII/BamHI digested PCR product and long PstI/BamHI V1R-FLpol backbone fragment.

The MRKAd5-gagpol adenovirus vector was constructed using the BglII fragment of the V1R-gagpol containing the entire ORF of gag-IApol fusion gene.

EXAMPLE 30

Immunogenicity Studies in Non-Human Primates

Cohorts of three (3) macaques were immunized with 10e8 or 10e10 viral particles (vp) of one of the following MRKAd5 HIV-1 vaccines: (1) MRKAd5gag; (2) MRKAd5pol; (3) MRKAd5nef; (4) a mixture containing equal amounts of MRKAd5gag, MRKAd5pol, and MRKAd5nef, or (5) a mixture of equal amounts of MRKAd5gagpol and MRKAd5nef. The vaccines were administered at weeks 0 and 4.

The T cell responses against each of the HIV-1 antigens were assayed by IFN-gamma ELISpot assay using pools of 20-aa peptides that encompass the entire protein sequence of each antigen. The results (Table 25) are expressed as the number of spot-forming cells (sfc) per million peripheral blood mononuclear cells (PBMC) that respond to each of the peptide pools.

Results indicate the following observations: (1) each of the single gene constructs (MRKAd5gag, MRKAd5pol, or MRKAd5nef) is able to elicit high levels of antigen-specific T cells in monkeys; (2) the single-gene MRKAd5 constructs can be mixed as a multi-cocktail formulation capable of eliciting very broad T cell responses against gag, pol, and nef; (3) the MRKAd5 vector expressing the fusion

protein of gag plus IA pol is capable of inducing strong T cell responses to both gag and pol.

Table 25. Evaluation of Mixtures of MRKAd5 vectors expressing humanized HIV-1 gag, pol, gagpol, nef in rhesus macaques

5

Grp #	Vaccine T=0, 4 wks	Monk #	T=6 wks				
			Mock	Gag H	Pol - 1	Pol - 2	Nef
1	MRKAd5 gag 10 ¹⁰ vp	CB9V	0	15	-	-	-
		CD19	0	374	-	-	-
		109H	1	843	-	-	-
2	MRKAd5 gag 10 ⁸ vp	99D130	1	948	-	-	-
		W277	16	324	-	-	-
		143H	4	595	-	-	-
3	MRKAd5 pol 10 ¹⁰ vp	CC1X	4	-	46	256	-
		AW3W	3	-	463	550	-
		AV43	6	-	95	1333	-
4	MRKAd5 pol 10 ⁸ vp	AW38	1	-	19	30	-
		CC8K	0	-	50	995	-
		CC21	1	-	33	436	-
5	MRKAd5 nef 10 ¹⁰ vp	076Q	9	-	-	-	1204
		091Q	4	-	-	-	85
		083Q	0	-	-	-	176
6	MRKAd5 nef 10 ⁸ vp	00C029	1	-	-	-	114
		98D022	6	-	-	-	170
		98D160	3	-	-	-	198
7	MRKAd5gag+MRKAd5pol+MRKAd5nef 10 ¹⁰ vp each	99D251	3	206	15	193	120
		05H	3	135	21	9	638
		00C016	3	26	4	51	23
8	MRKAd5gag+MRKAd5pol+MRKAd5nef 10 ⁸ vp each	99D215	1	171	18	193	240
		81H	5	73	6	14	243
		12H	8	1140	115	811	719
9	MRKAd5gagpol +MRKAd5 nef 10 ¹⁰ vp each	99D211	0	83	56	838	725
		22H	4	385	119	1194	1915
		61H	4	343	11	765	853
10	MRKAd5gagpol +MRKAd5 nef 10 ⁸ vp each	34H	3	78	19	5	75
		48H	1	65	105	46	43
		70H	5	158	15	220	191

Indicated are numbers of spot-forming cells per million PBMCs against the peptide pools. Mock, no peptides; gag H, fifty 20-aa peptides encompassing p55 sequence; pol-1, 20-aa peptides representing N-terminal half of IA pol; pol-2, 20-aa peptides representing the carboxy-terminal half of IA pol; nef, 20-aa peptides encompassing the entire wild-type nef sequence. Responses to the antigens prior to the first immunization did not exceed 40 sfc/10⁶ PBMC.

10

WHAT IS CLAIMED IS

:

1. A recombinant adenoviral vaccine vector at least partially deleted in
5 E1 and devoid of E1 activity, comprising:
 - a) an adenovirus *cis*-acting packaging region corresponding to from
about base pair 1 to between from about base pair 400 to about
base pair 458 of a wildtype adenovirus genome; and
 - b) a gene encoding an HIV protein or immunologically relevant
10 modification thereof.
2. A vector in accordance with claim 1 comprising a packaging region
corresponding to from about base pair 1 to about base pair 450 of a wildtype
adenovirus genome.
3. A vector in accordance with claim 1 further comprising nucleotides
15 corresponding to between from about base pair 3511 to about 3524 to about base pair
5798 of a wildtype adenovirus genome.
4. A vector in accordance with claim 3 comprising base pairs
corresponding to 1-450 and 3511-5798 of a wildtype adenovirus genome.
5. A vector in accordance with claim 4 which is deleted of base pairs
20 451-3510.
6. A vector in accordance with claim 1 which is at least partially
deleted in E3.
7. A vector in accordance with claim 6 wherein the E3 deleted region
is from base pairs 28,133-30,818.

8. A vector in accordance with claim 1 wherein the gene encoding the HIV protein or modification thereof comprises codons optimized for expression in a human.

9. A vector in accordance with claim 1 wherein the vector comprises a
5 gene expression cassette comprising:

a) a nucleic acid encoding a protein;

b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and

(c) a transcription termination sequence.

10. A vector in accordance with claim 9 wherein the gene expression cassette is inserted into the E1 region.

11. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette is in an E1 parallel orientation

12. An adenoviral vector in accordance with claim 9 wherein the gene
15 expression cassette is in an E1 antiparallel orientation.

13. An adenoviral vector in accordance with claim 9 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

14. An adenoviral vector in accordance with claim 13 wherein the promoter is an immediate early human cytomegalovirus promoter.

20 15. An adenoviral vector in accordance with claim 9 wherein the promoter is a murine cytomegalovirus promoter.

16. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

17. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a synthetic polyadenylation signal (SPA).

18. A cell comprising the adenoviral vector of claim 1.

19. Recombinant, replication-defective adenovirus particles harvested
5 and purified subsequent to transfection of the adenoviral vector of claim 1 into a cell line which expresses adenovirus E1 protein at complementing levels.

20. An HIV vaccine composition comprising purified adenovirus particles of claim 19.

21. An HIV vaccine composition of claim 20 which comprises a
10 physiologically acceptable carrier.

22. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 1 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant,
15 replication-defective adenovirus.

23. A method according to claim 22 wherein the cell is a PER.C6[®] cell.

24. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of
20 claim 21.

25. A method according to claim 24 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

26. A method according to claim 25 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

27. A method according to claim 24 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

28. A method according to claim 24 which comprises administering and readministering the adenovirus vaccine vector to the individual.

29. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV gag or an immunologically relevant modification thereof.

30. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV gag protein or immunologically relevant modification thereof.

31. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:

a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and

b) a gene expression cassette comprising

- i) SEQ ID NO: 29;
- ii) a heterologous promoter operatively linked to i); and
- iii) a transcription termination sequence.

32. An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 parallel orientation.

33. An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 antiparallel orientation.

5 34. An adenoviral vector in accordance with claim 31 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

35. An adenoviral vector in accordance with claim 31 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

10 36. An adenoviral vector in accordance with claim 31 which is at least partially deleted in E3.

37. A cell comprising the adenoviral vector of claim 30.

38. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 30 into a cell
15 line which expresses adenovirus E1 protein at complementing levels.

39. An HIV vaccine composition comprising purified adenovirus particles of claim 38.

40. An HIV vaccine composition of claim 39 which comprises a physiologically acceptable carrier.

20 41. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 30 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

42. A method according to claim 41 wherein the cell is a PER.C6[®] cell.

43. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of
5 claim 21.

44. A method according to claim 43 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

10 45. A method according to claim 44 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

46. A method according to claim 43 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

15 47. A method according to claim 43 which comprises administering and readministering the adenovirus vaccine vector to the individual.

48. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV pol or an immunologically relevant modification thereof.

49. An adenoviral vector in accordance with claim 9 wherein the gene
20 expression cassette comprises an open reading frame encoding an HIV pol protein or immunologically relevant modification thereof.

50. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:

a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and

b) a gene expression cassette comprising

i) a nucleotide sequence selected the group consisting of

SEQ ID NO: 1, SEQ ID NO: 5 and SEQ ID NO: 7;

ii) a heterologous promoter operatively linked to i); and

iii) a transcription termination sequence.

51. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 parallel orientation.

52. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 antiparallel orientation.

53. An adenoviral vector in accordance with claim 50 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

54. An adenoviral vector in accordance with claim 50 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

55. An adenoviral vector in accordance with claim 50 which is at least partially deleted in E3.

56. A cell comprising the adenoviral vector of claim 49.

57. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 49 into a cell line which expresses adenovirus E1 protein at complementing levels.

58. An HIV vaccine composition comprising purified adenovirus
5 particles of claim 57.

59. An HIV vaccine composition of claim 58 which comprises a physiologically acceptable carrier.

60. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of
10 claim 49 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

61. A method according to claim 60 wherein the cell is a PER.C6[®] cell.

62. A method of generating a cellular-mediated immune response
15 against HIV in an individual comprising administering to the individual a vaccine of claim 59.

63. A method according to claim 62 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with
20 a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

64. A method according to claim 63 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

65. A method according to claim 62 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

66. A method according to claim 62 which comprises administering and readministering the adenovirus vaccine vector to the individual.

5 67. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV nef or an immunologically relevant modification thereof.

68. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV nef protein or immunologically relevant modification thereof.

10 69. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:

a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and

b) a gene expression cassette comprising

i) a nucleotide sequence selected the group consisting of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13 and SEQ ID NO: 15;

ii) a heterologous promoter operatively linked to i); and

iii) a transcription termination sequence.

70. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 parallel orientation.

71. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 antiparallel orientation.

72. An adenoviral vector in accordance with claim 69 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

5 73. An adenoviral vector in accordance with claim 69 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

74. An adenoviral vector in accordance with claim 69 which is at least partially deleted in E3.

10 75. A cell comprising the adenoviral vector of claim 68.

76. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 68 into a cell line which expresses adenovirus E1 protein at complementing levels.

15 77. An HIV vaccine composition comprising purified adenovirus particles of claim 76.

78. An HIV vaccine composition of claim 77 which comprises a physiologically acceptable carrier.

79. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 68 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

20

80. A method according to claim 79 wherein the cell is a PER.C6[®] cell.

81. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 78.

82. A method according to claim 81 which further comprises
5 administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

83. A method according to claim 82 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus
10 vaccine.

84. A method according to claim 81 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

85. A method according to claim 81 which comprises administering and readministering the adenovirus vaccine vector to the individual.

86. A multivalent adenovirus vaccine composition comprising
15 recombinant, replication-defective adenovirus particles, wherein the adenovirus particles are harvested and purified from a cell line expressing adenovirus E1 protein, and wherein the particles are harvested subsequent to transfection of the cells with an adenoviral vector or vectors in accordance with claim 9; said vector(s) comprising a
20 gene expression cassette or cassettes comprising nucleotide sequences encoding HIV proteins selected from the group consisting of:

- a) gag, pol, and nef, expressed independently from three individual vectors;

- b) gag, pol, and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- c) gag, pol, and nef, expressed via two vectors, one expressing a pol-nef fusion, and another expressing gag;
- d) gag, pol, and nef, expressed via two vectors, one expressing a gag-pol fusion and another expressing nef;
- e) gag, pol and nef, expressed via two vectors, one expressing a nef-gag fusion and another expressing pol;
- f) gag, pol, and nef, expressed via one vector expressing a gag-pol-nef fusion;
- g) gag and pol, expressed independently from two individual vectors;
- h) gag and pol, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- i) pol and nef, expressed independently from two individual vectors;
- j) pol and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- k) nef and gag, expressed independently from two individual vectors;
- l) nef and gag, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- m) gag and pol, expressed via one vector expressing a gag-pol fusion;

n) pol and nef, expressed via one vector expressing a pol-nef fusion;
and

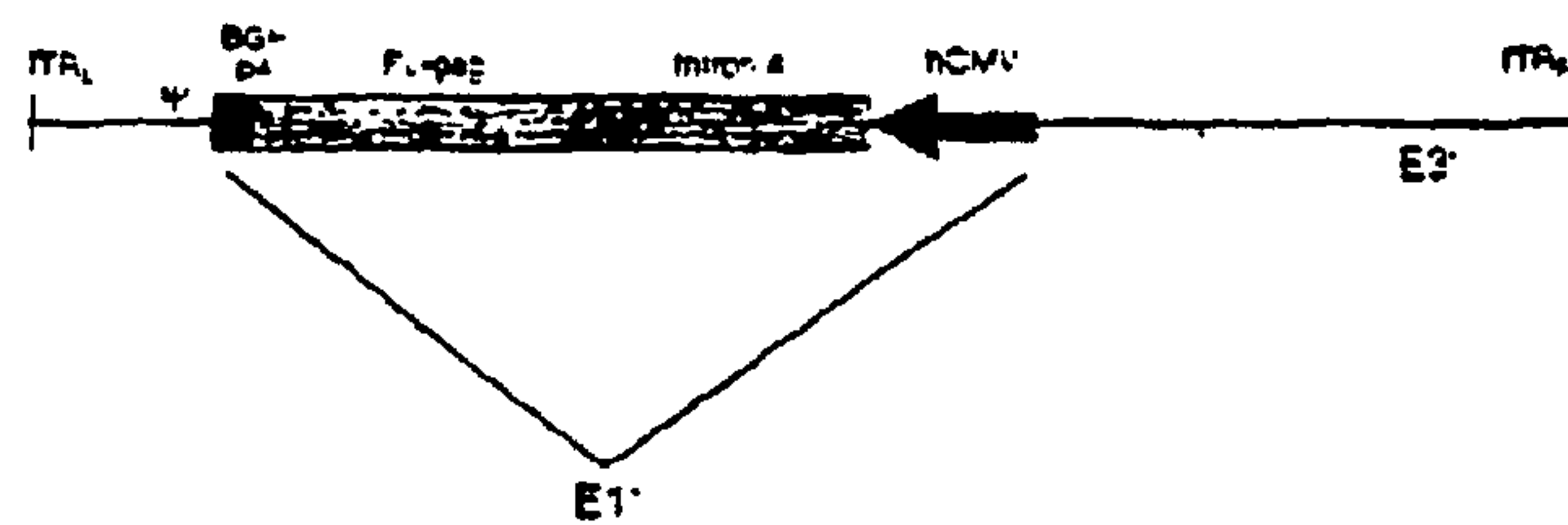
o) nef and gag, expressed via one vector expressing a nef-gag fusion.

87. A multivalent adenovirus vaccine composition in accordance with
5 claim 86 wherein the gag-pol fusion consists of SEQ ID NO: 39.

88. A multivalent adenovirus vaccine composition in accordance with
claim 86 wherein the fused sequences have the encoding nucleic acid sequences
operatively linked to distinct promoters and transcription termination sequences.

89. A multivalent adenovirus vaccine composition in accordance with
10 claim 86 wherein the fused sequences have the encoding nucleic acid sequences
operatively linked to a single promoter; and the encoding nucleic acid sequences
operatively linked by an internal ribosome entry sequence ("IRES").

Original Adenovector Construct:

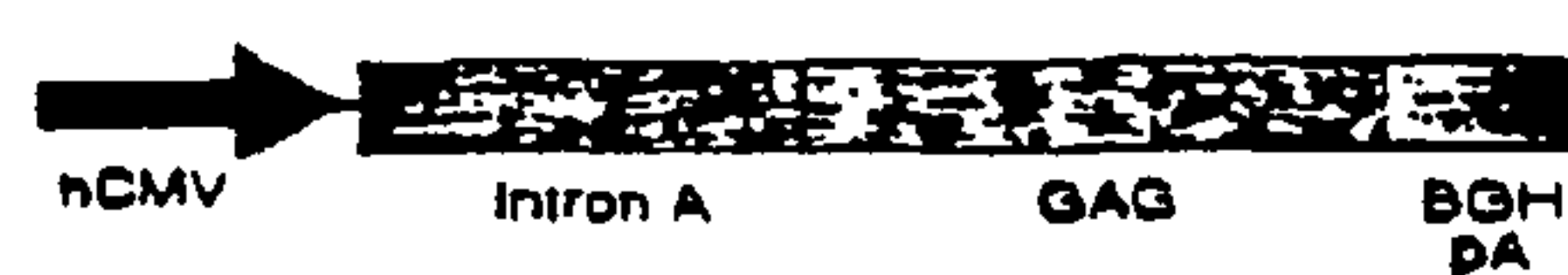
**Figure 1:** Original HIV-1 gag adenovector.

Sequence of the open reading frame for FL-gag (human codon optimized)

atgggtgctagggcttctgtgctgtctgggtggtagctggacaagtgggagaagatcaggctgaggcctggtagg
caagaagaagtaacaagctaaagcacatgtgtgtggccctccagggagctggagaggtttgtgtgaacctggc
ctgtgtggagacctctgaggggtgcaggcagatcctgggccagctccagccctccctgcaaacaggctctgagg
agctgagggtccctgtacaacacagtggctacctgtactgtgtgcaccagaagatgatgtgaaggacaccaag
gaggccctggagaagatgaggaggagcagaacaagtccaagaagaaggcccagcaggctgtgtgtggc
acaggcaactccagccagggtgtcccagaactaccccatgtgtgcagaacctccagggccagatgggtgcaccag
gccatctccccccggacctgaatgcctgggtgaagggtggaggagaaggccttctccctgagggtgatccc
catgttctctgcccgtgtgtgaggggtgccacccccaggacctgaacaccatgtgtgaacacagtggggggccatc
aggctgccatgcagatgtgtgaaggagacctcaatgaggaggctgtgtgagtgaggacaggctgtcatcctgtgc
acgtgtggcccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
ccaggagcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
cctggggcctgaacaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
ccttcagggaactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
ggatgacagagacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
ctgccacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gctgaggccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gaagacagtgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
agggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
ggcaaaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
cccgaggagtccttcagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
agctgtacccccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gat (SEQ ID NO: 29)

Figure 2

Old Transgene:



New Transgenes:

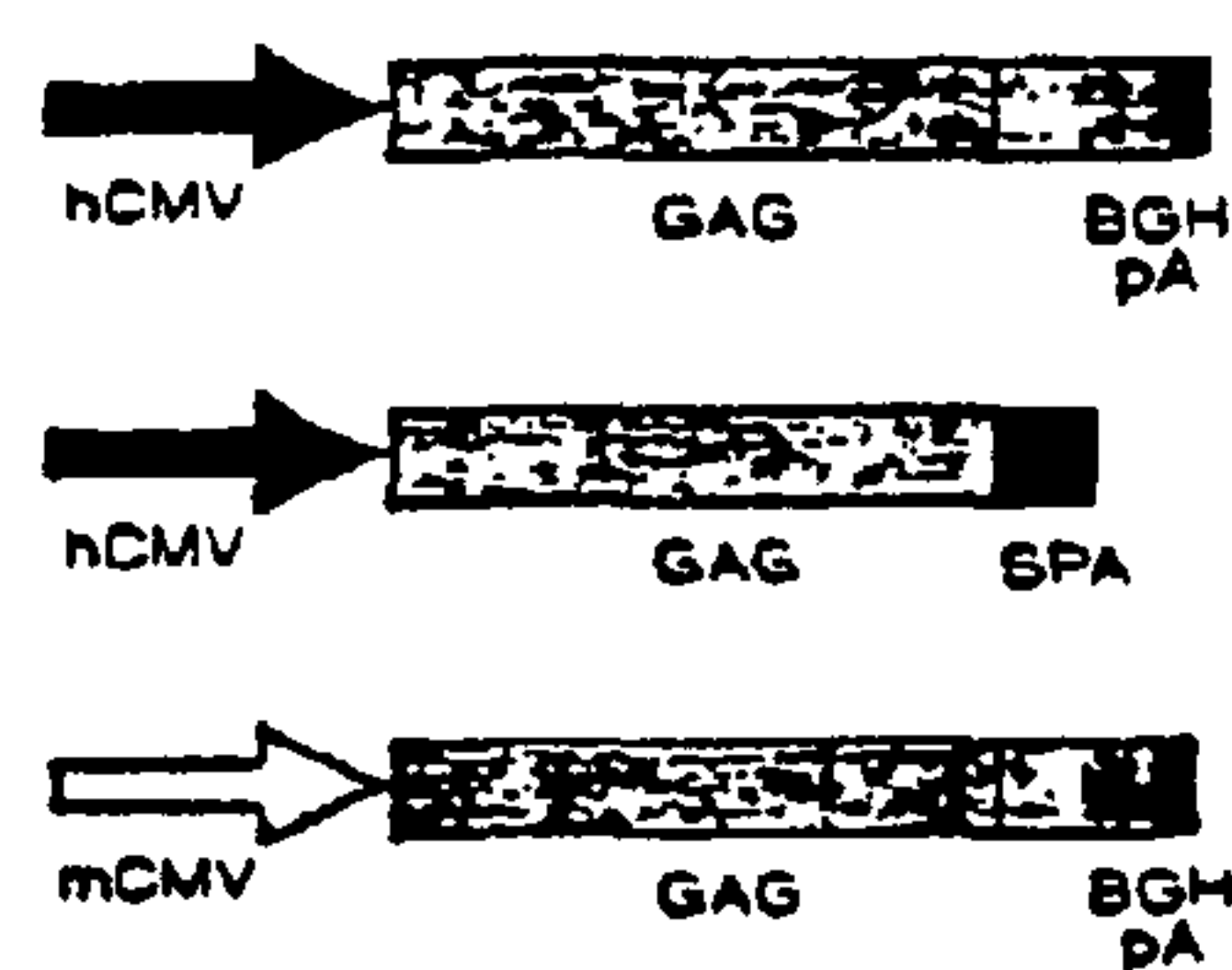


Figure 3: Diagrammatic representation of the original HIV-1 gag transgene and the series of new transgene constructions.

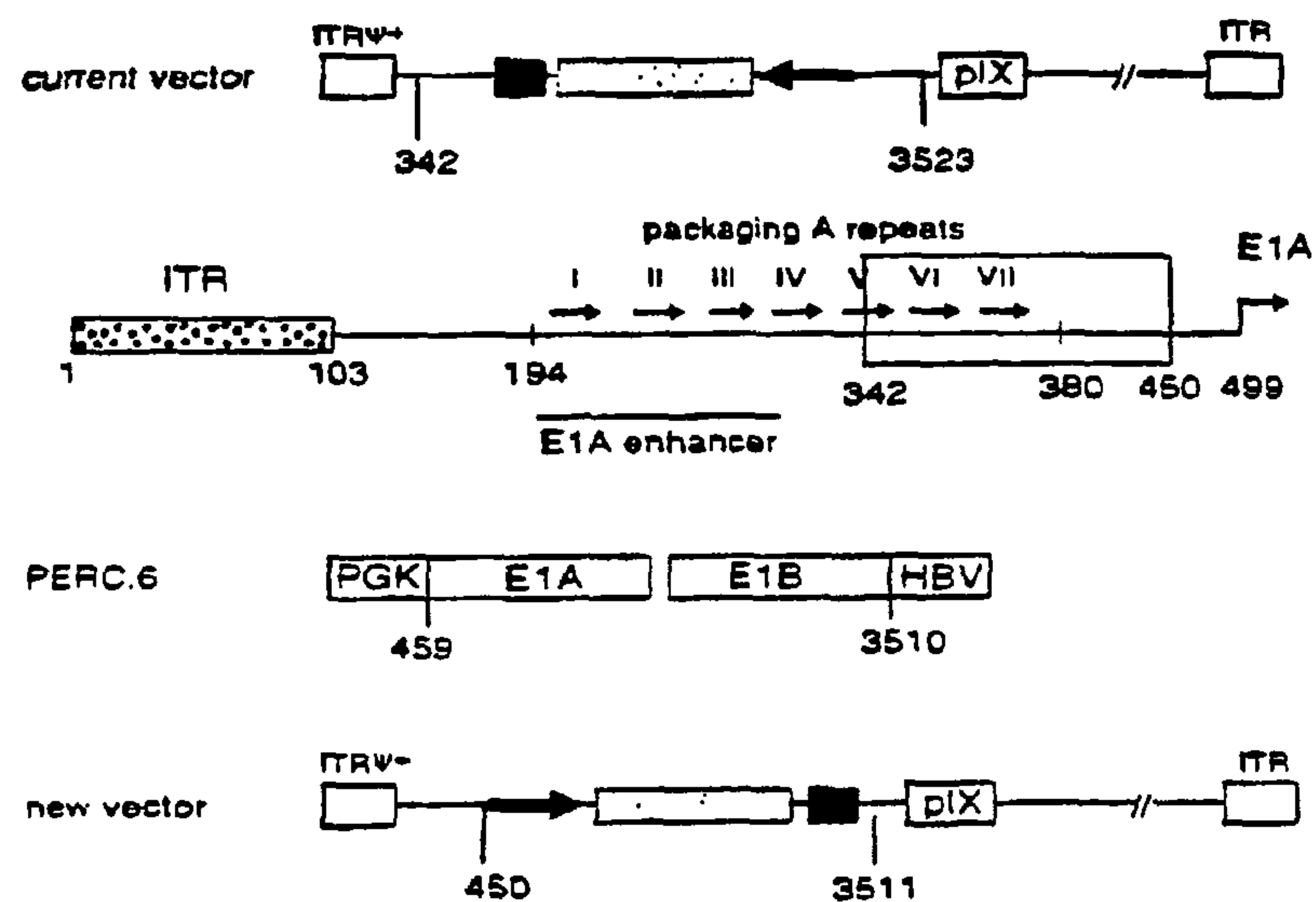


Figure 4: Modifications made to the current adenovector backbone in the generation of the new vector.

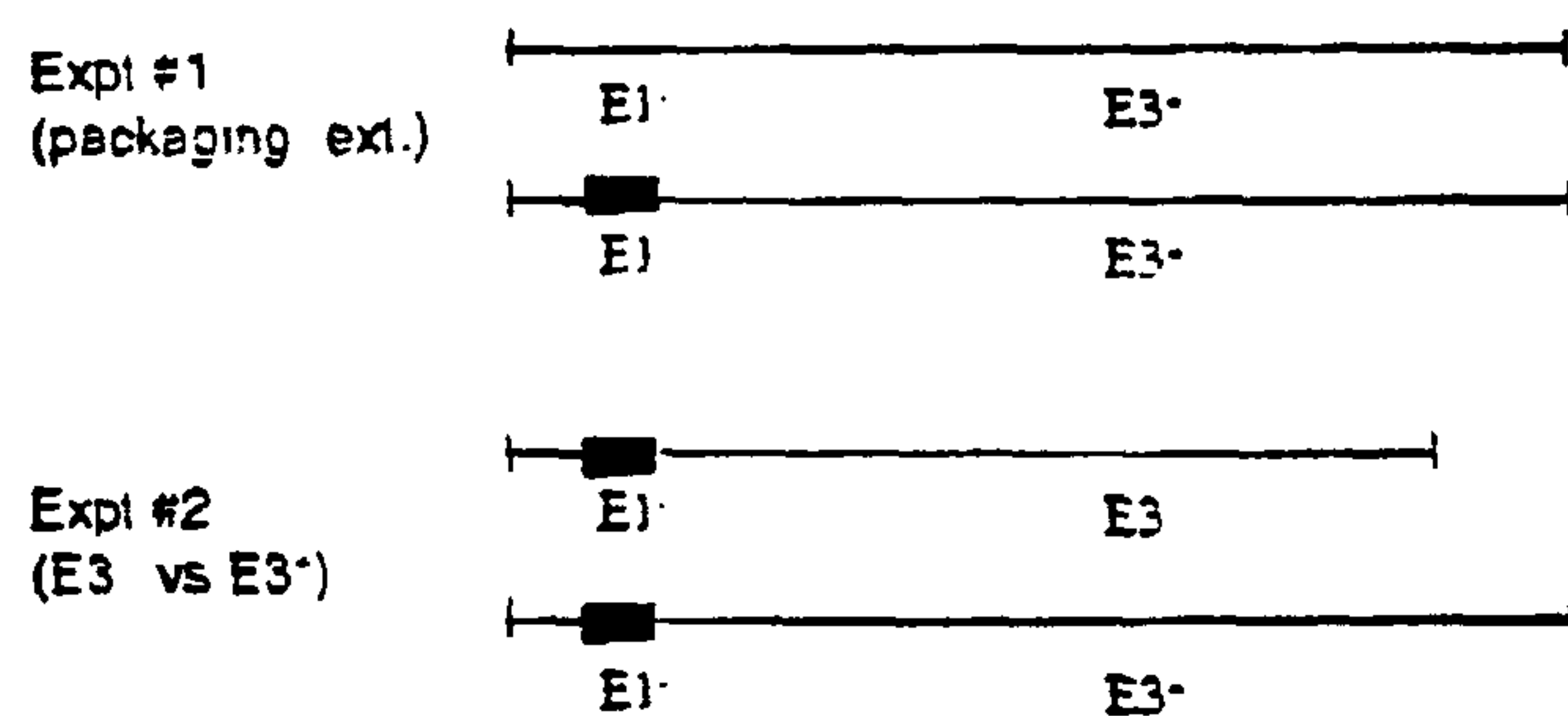


Figure 5: Virus mixing experiments to determine the effects of the addition made to the packaging signal region (Expt #1) and analysis of the effects of the E3 gene on viral growth (Expt. #2). The red bars denote the region of modifications made to the E1 deletion.

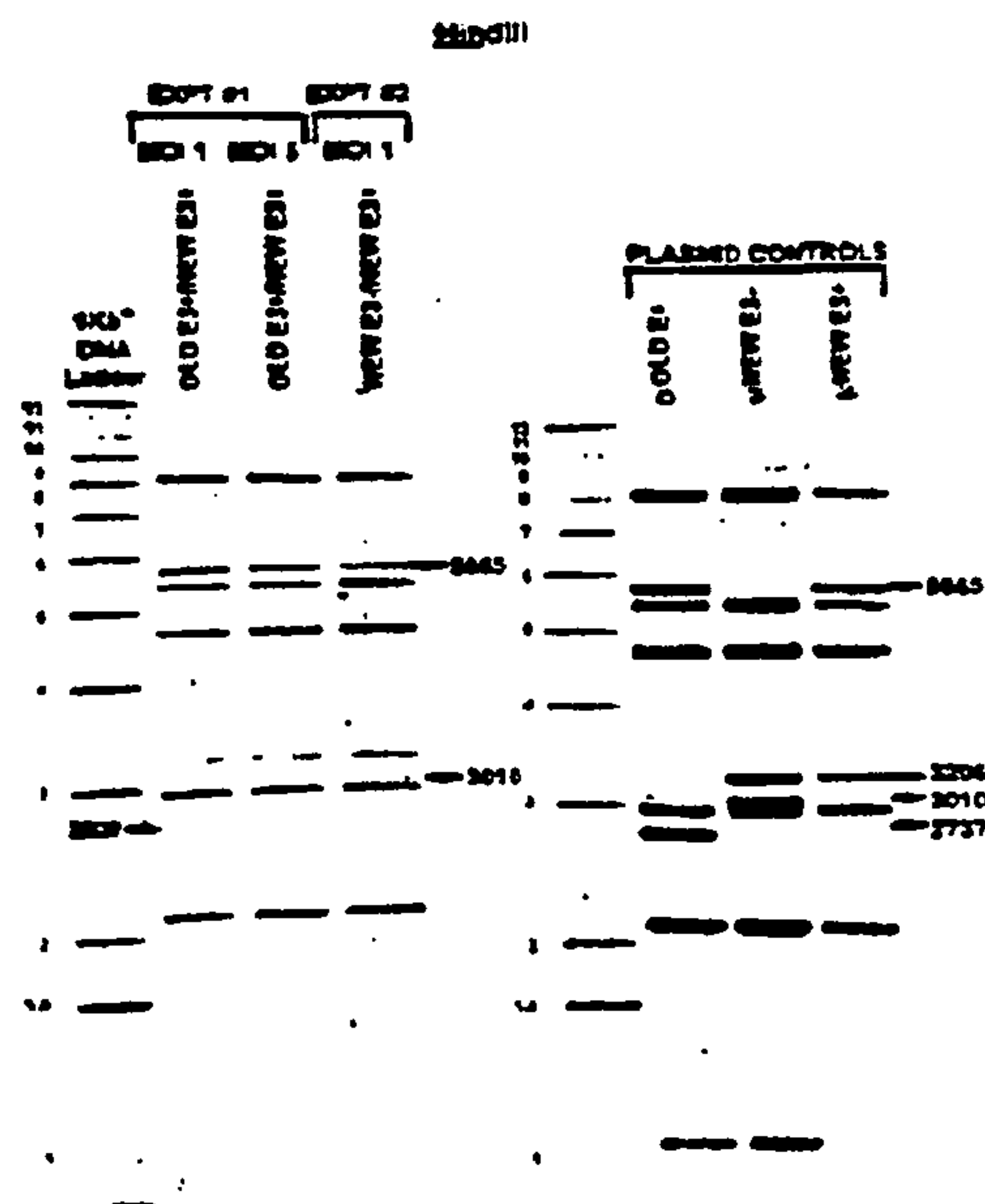


Figure 6: Autoradiograph of viral DNA analysis following viral mixing experiments (expts. #1 and #2) as detailed in the text.

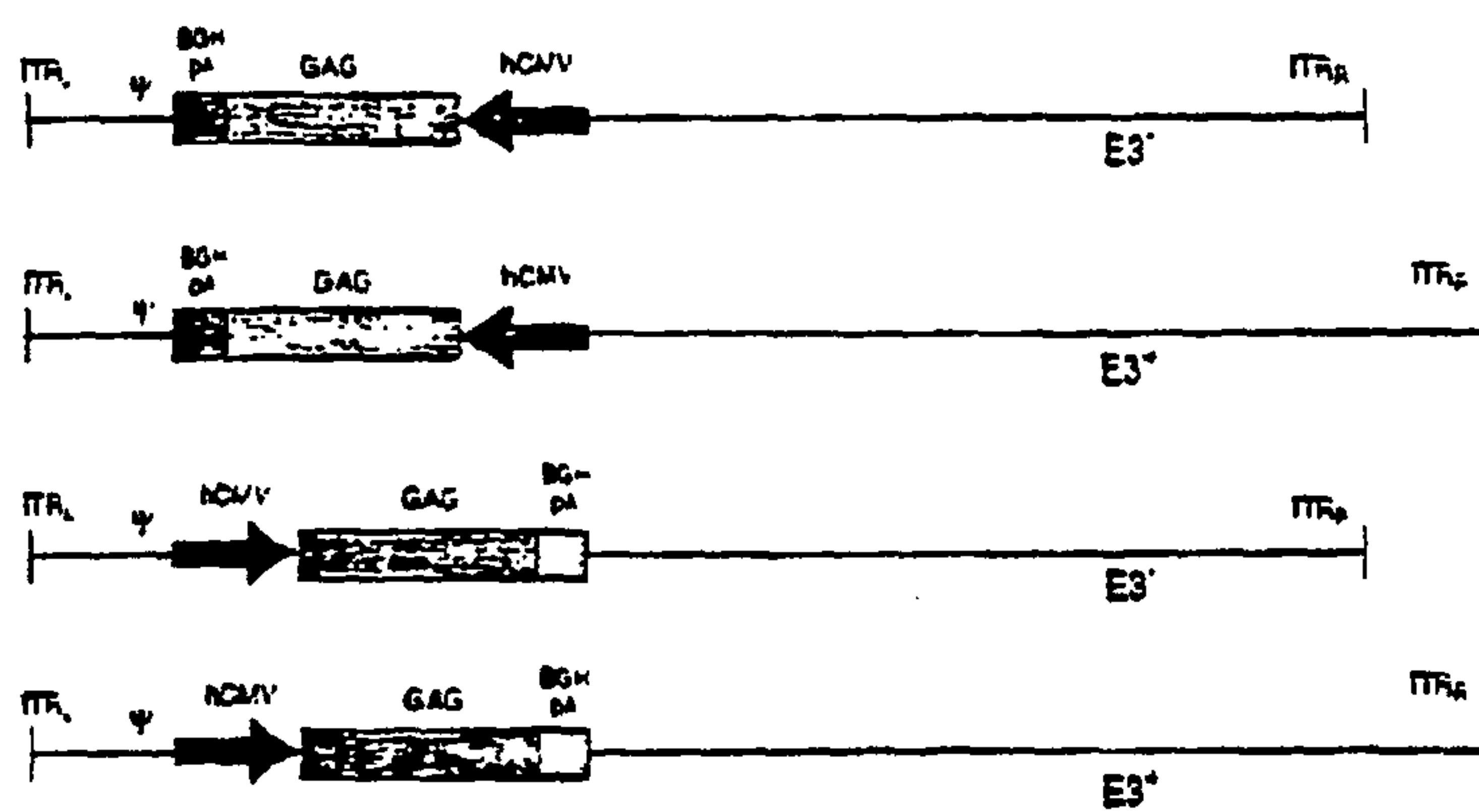


Figure 7A: hCMV-FLgag-bGHpA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

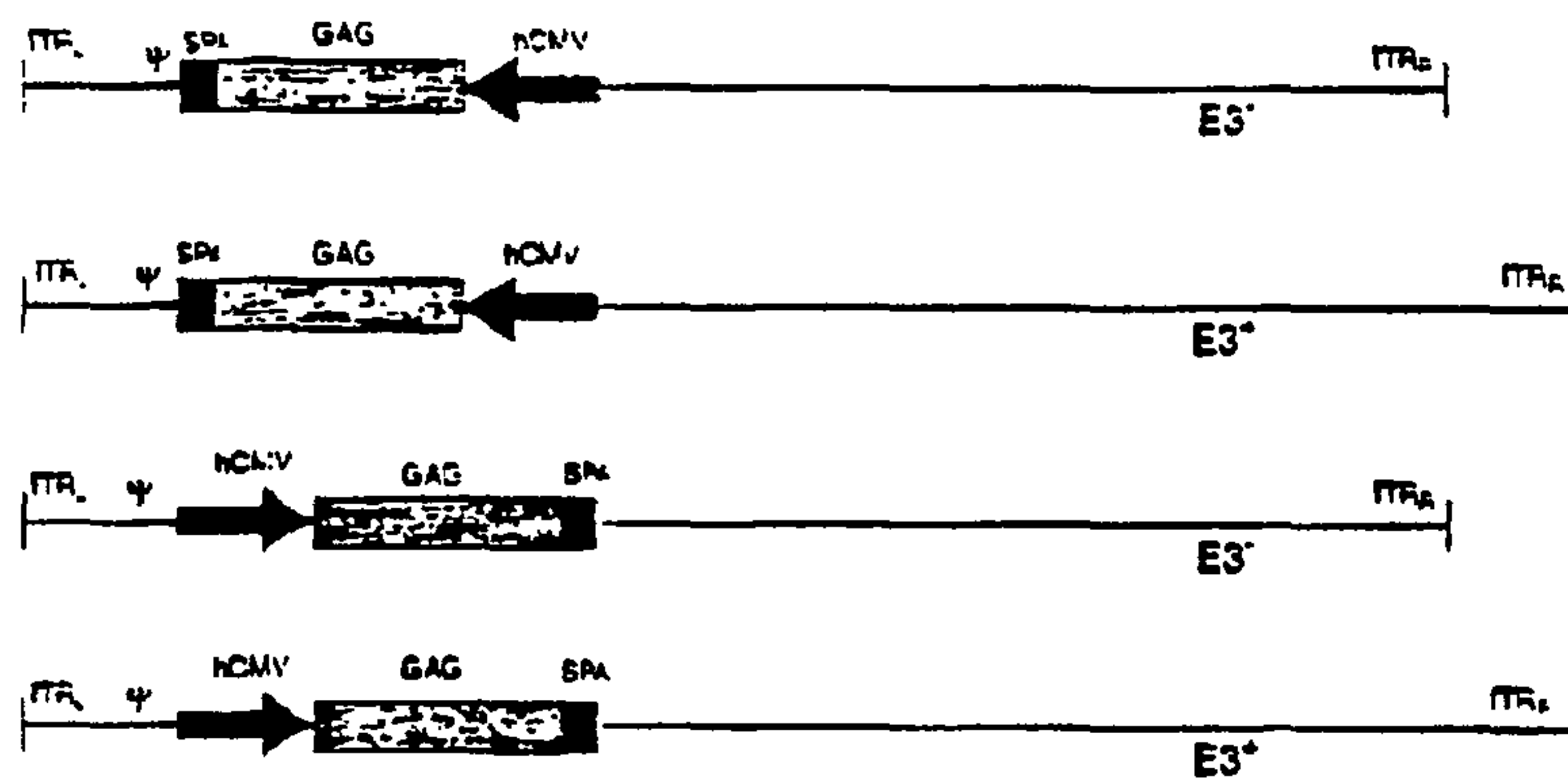


Figure 7B: hCMV-FLgag-SPA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3⁻ and E3⁺ backbones were constructed.

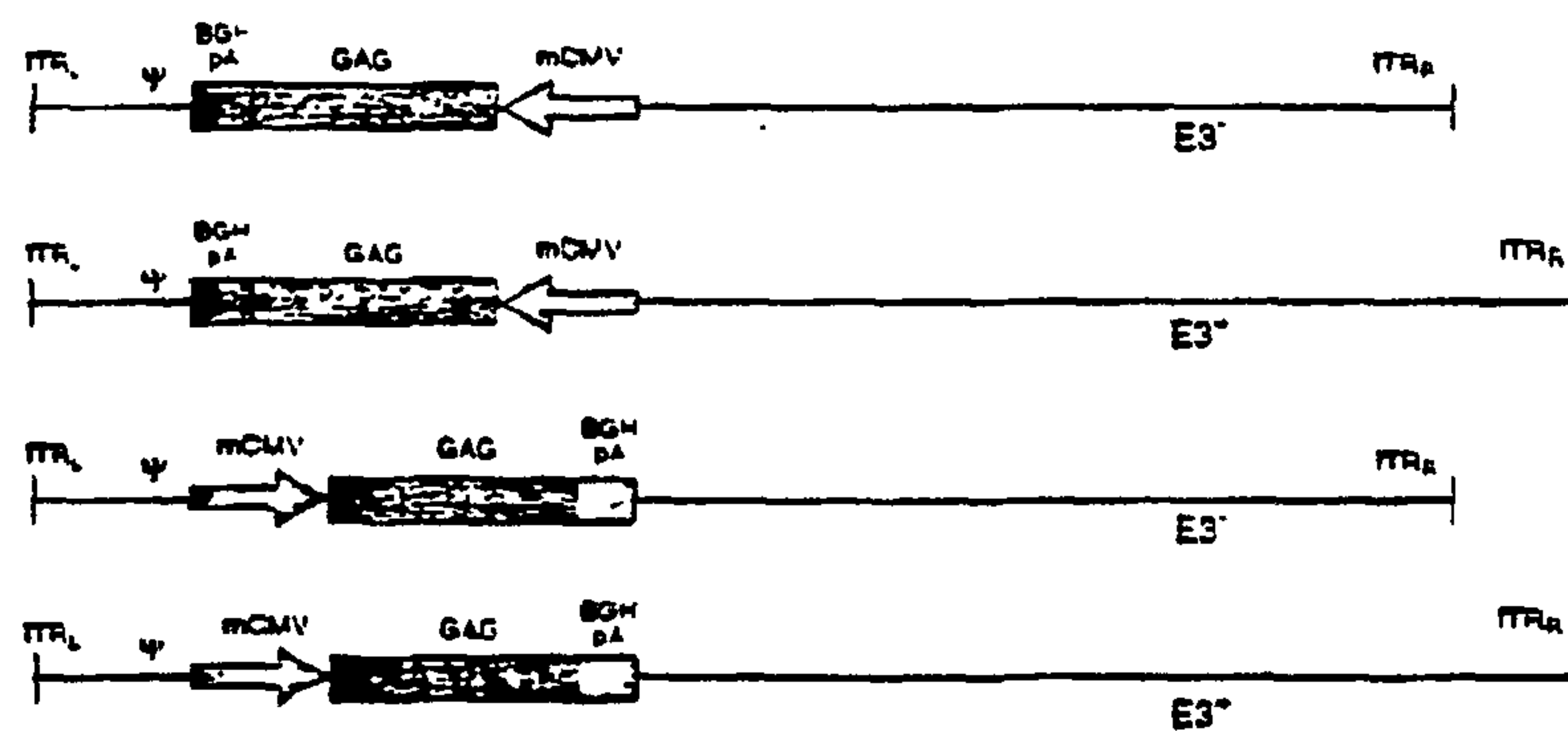


Figure 7C: mCMV-FLgag-bGHpA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3⁻ and E3⁺ backbones were constructed.

Plasmid mixing expt: (orientation)

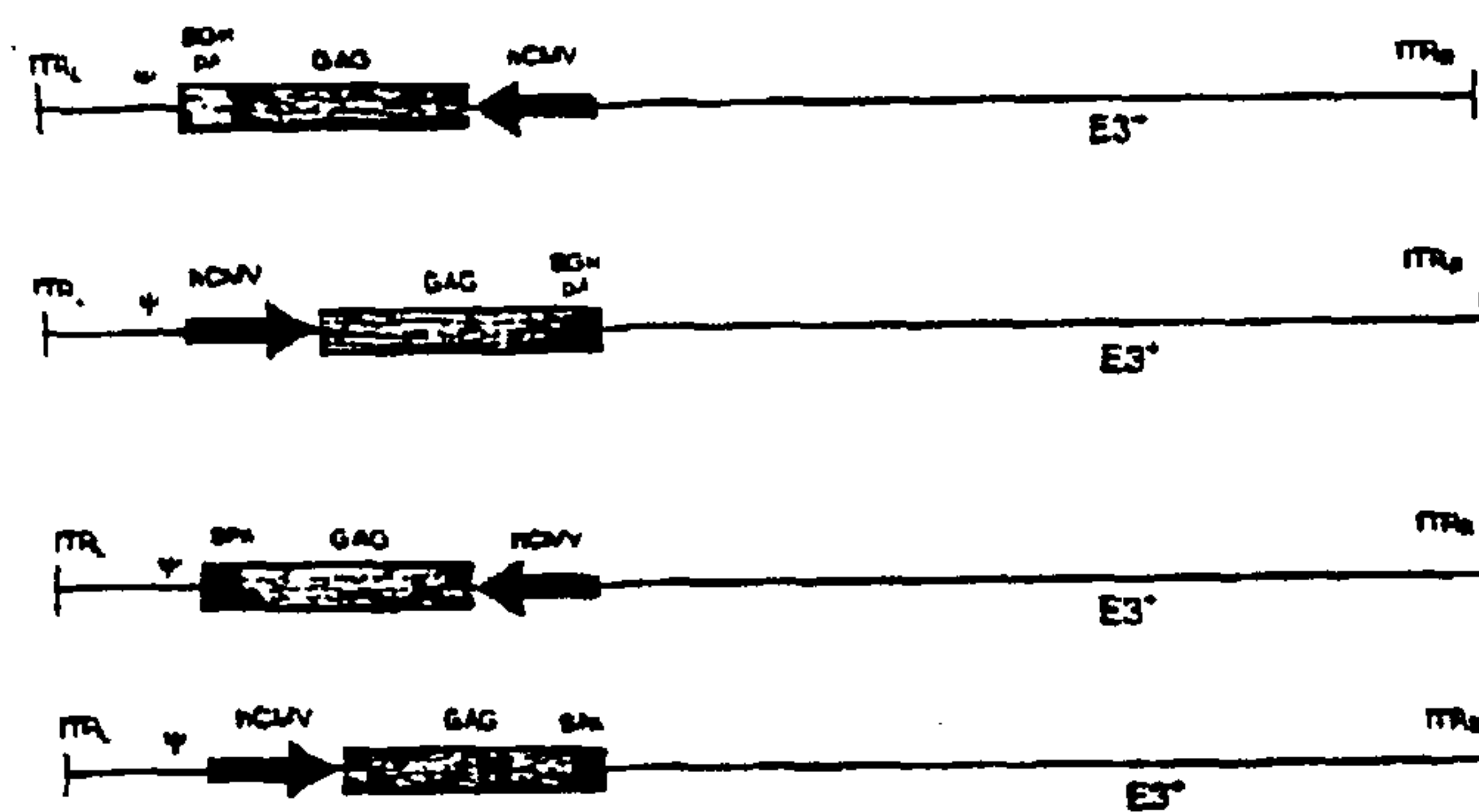


Figure 8A: Effect of transgene orientation

Plasmid Mixing expt: (poly A signal)

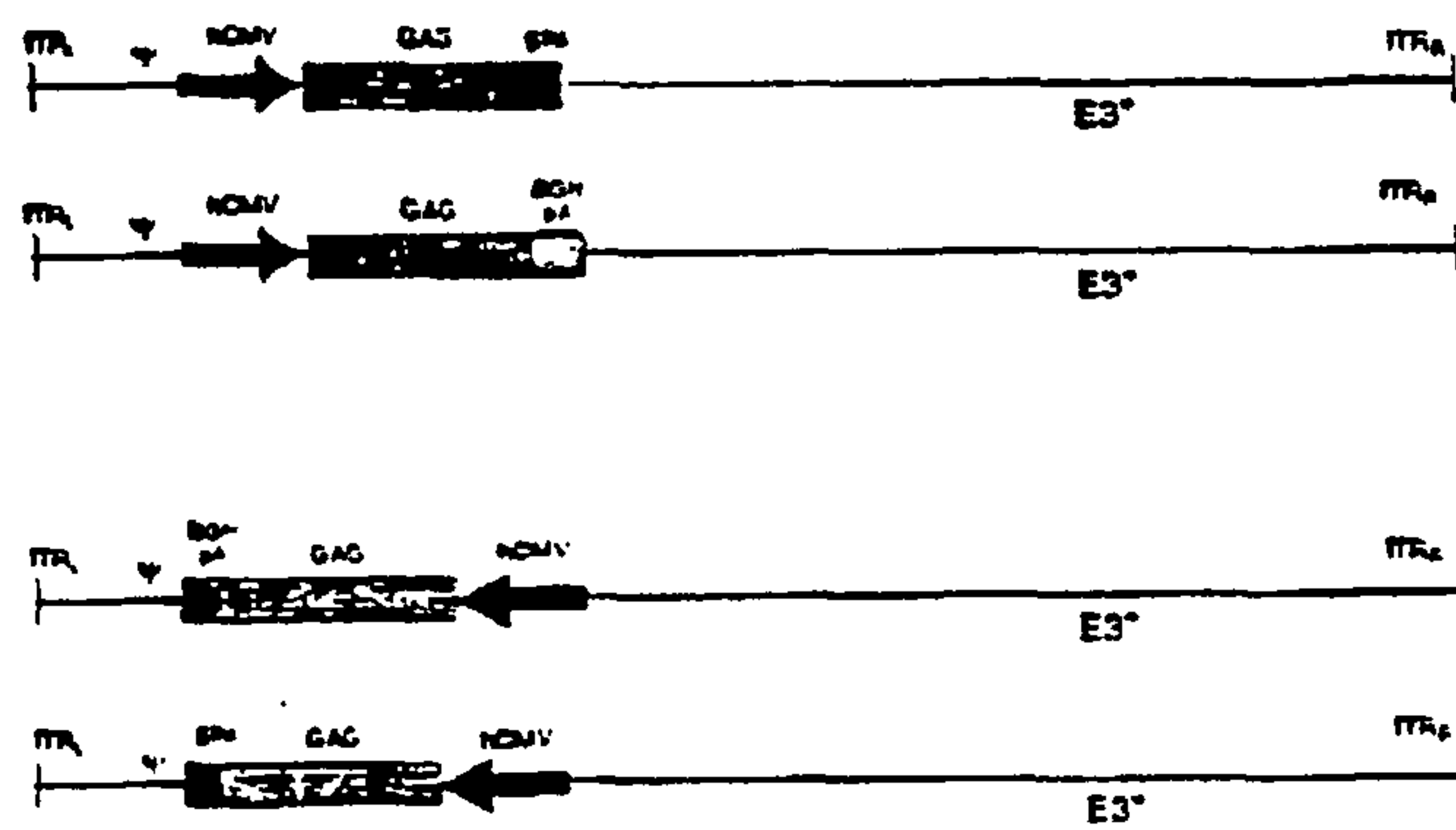


Figure 8B: Effect of polyadenylation signal

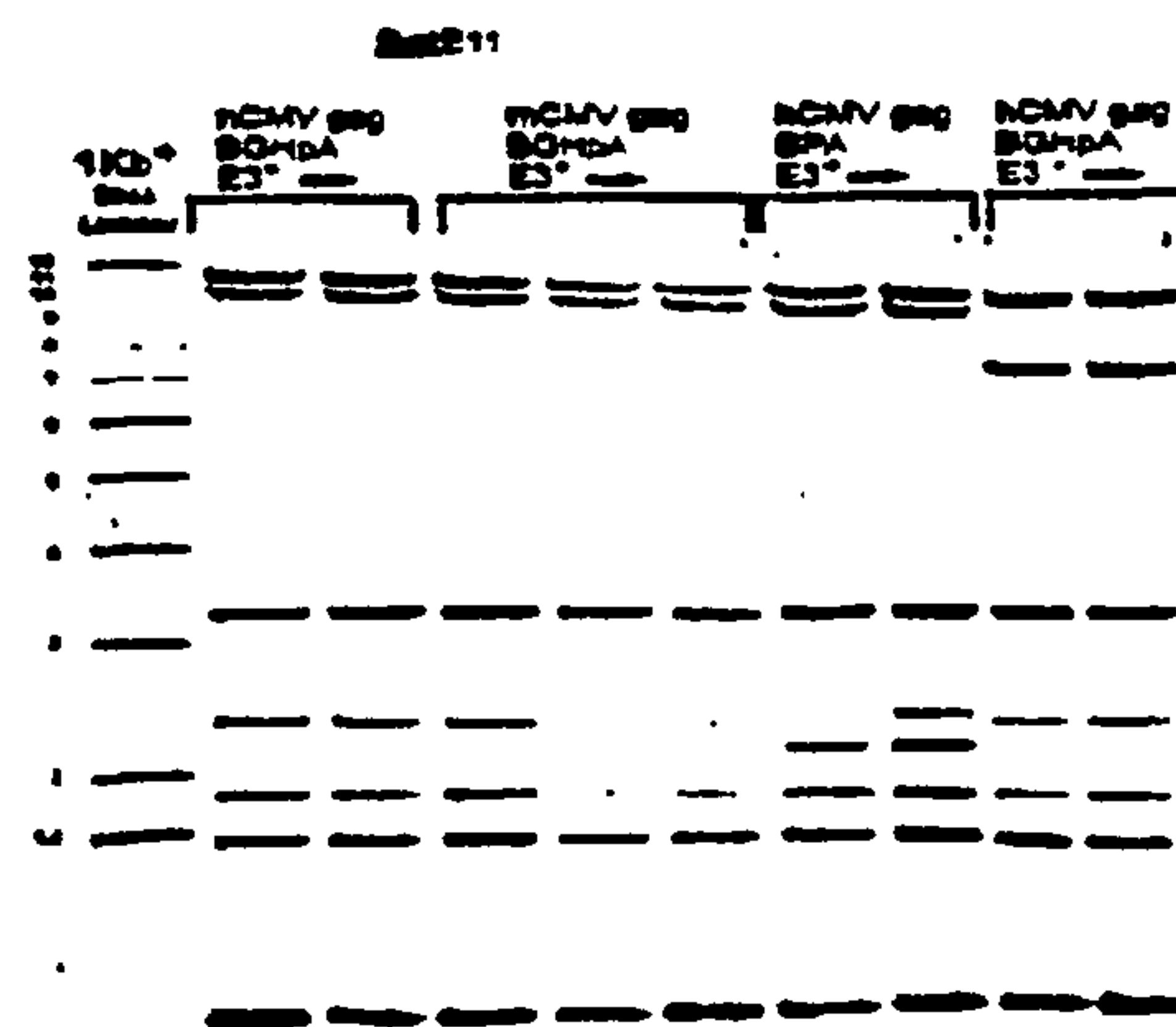


Figure 9: Viral DNA from the four Adgag candidates at P5, following BstE11 digestion.

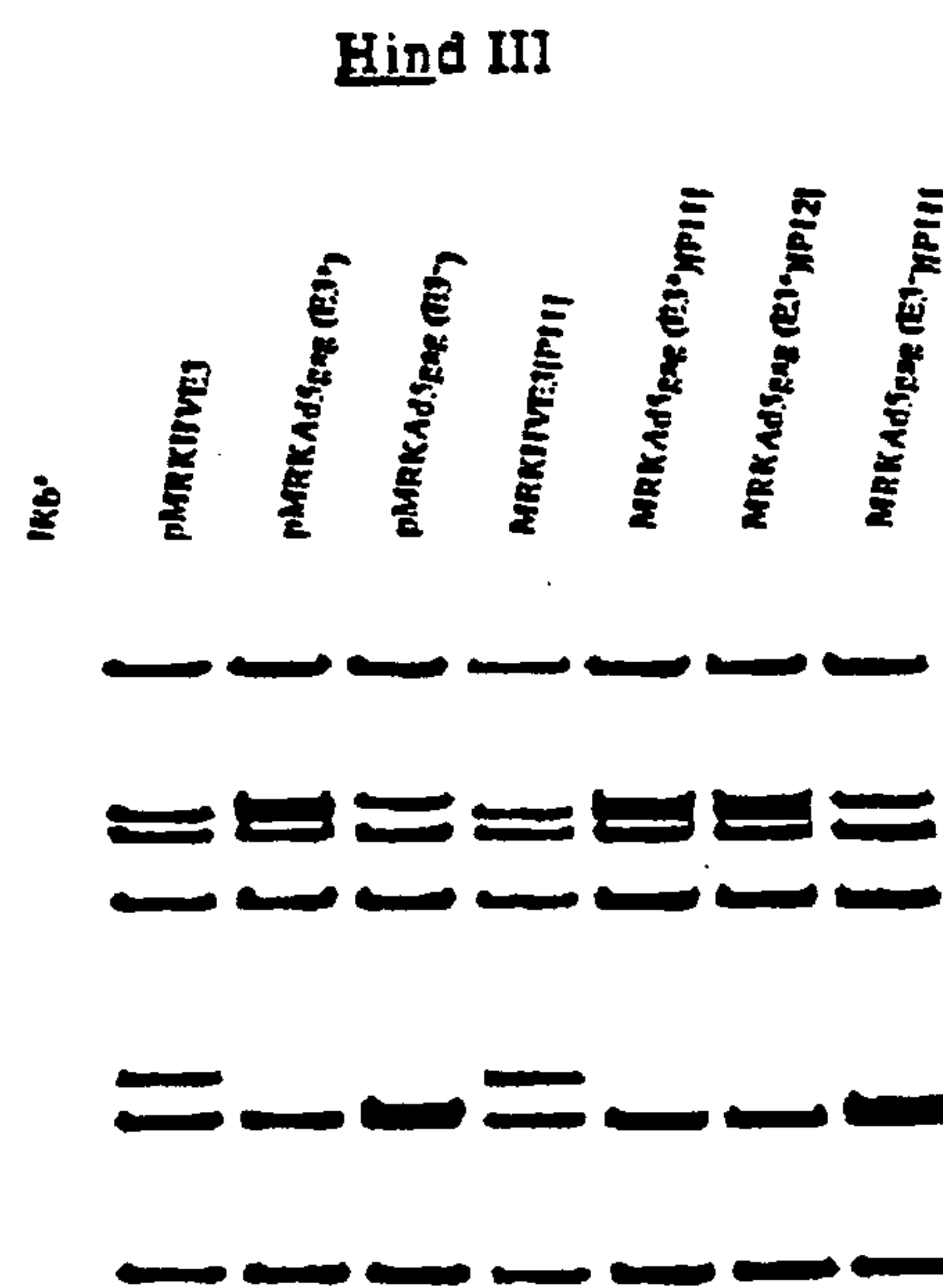


Figure 10: Viral DNA analysis of passage 11 and/or 12 of MRKHVE3, MRKAd5gag and MRKAd5gag(E3-).

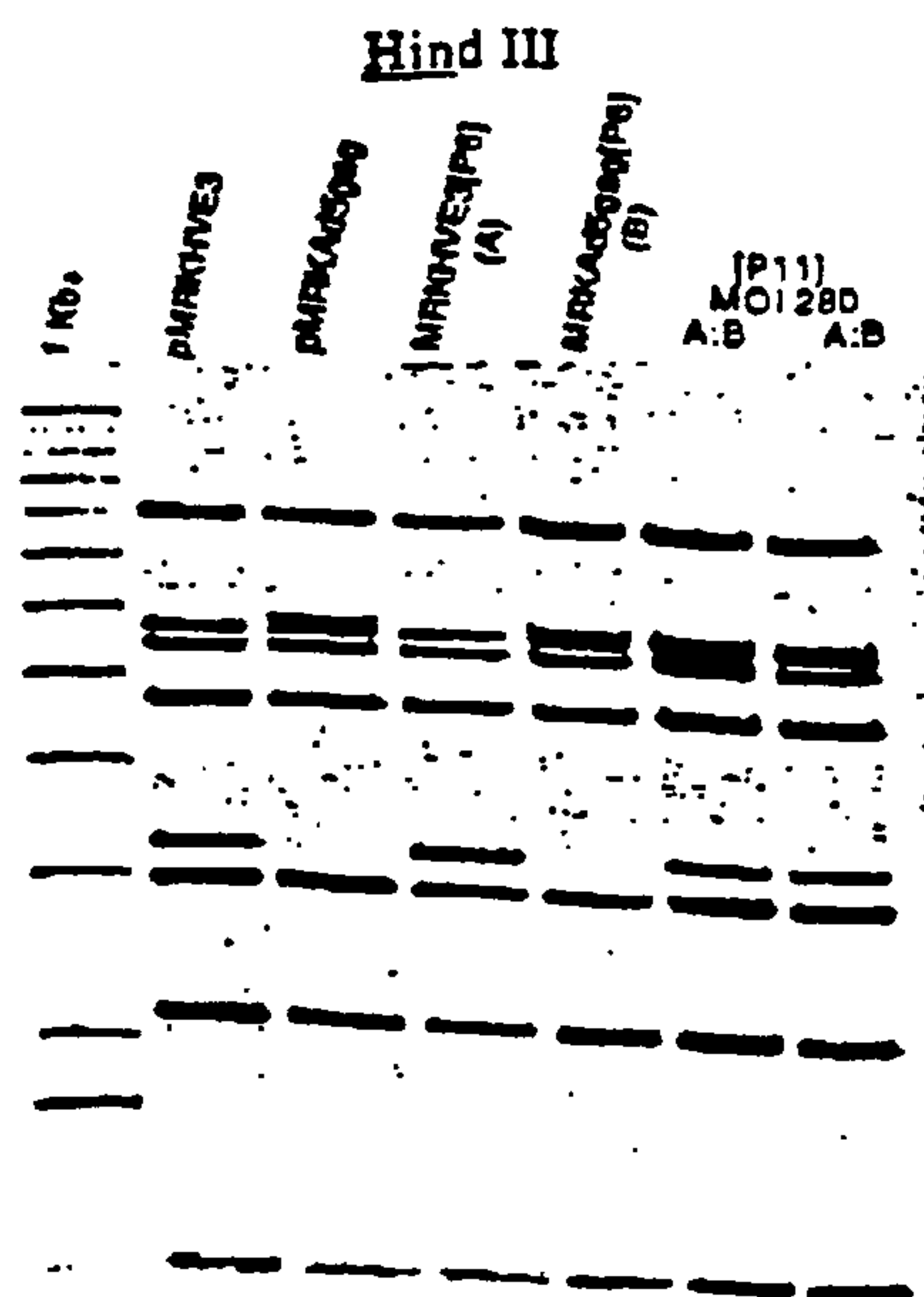


Figure 11: Viral DNA analysis (*Hind*III digestion) of passage 6 MRKHVE3 and MRKAd5gag used to initiate the viral competition study. Last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI 280 vp).

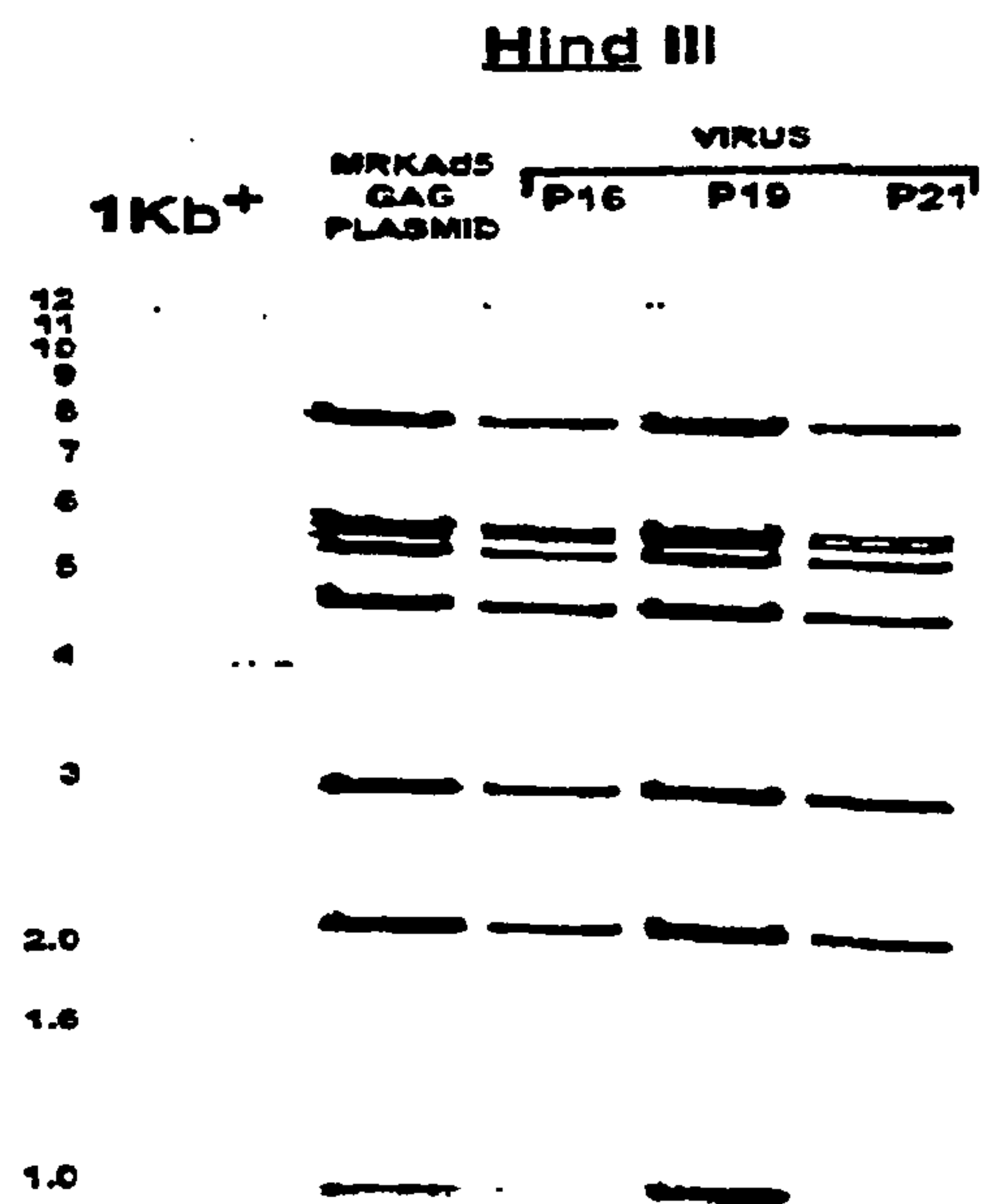
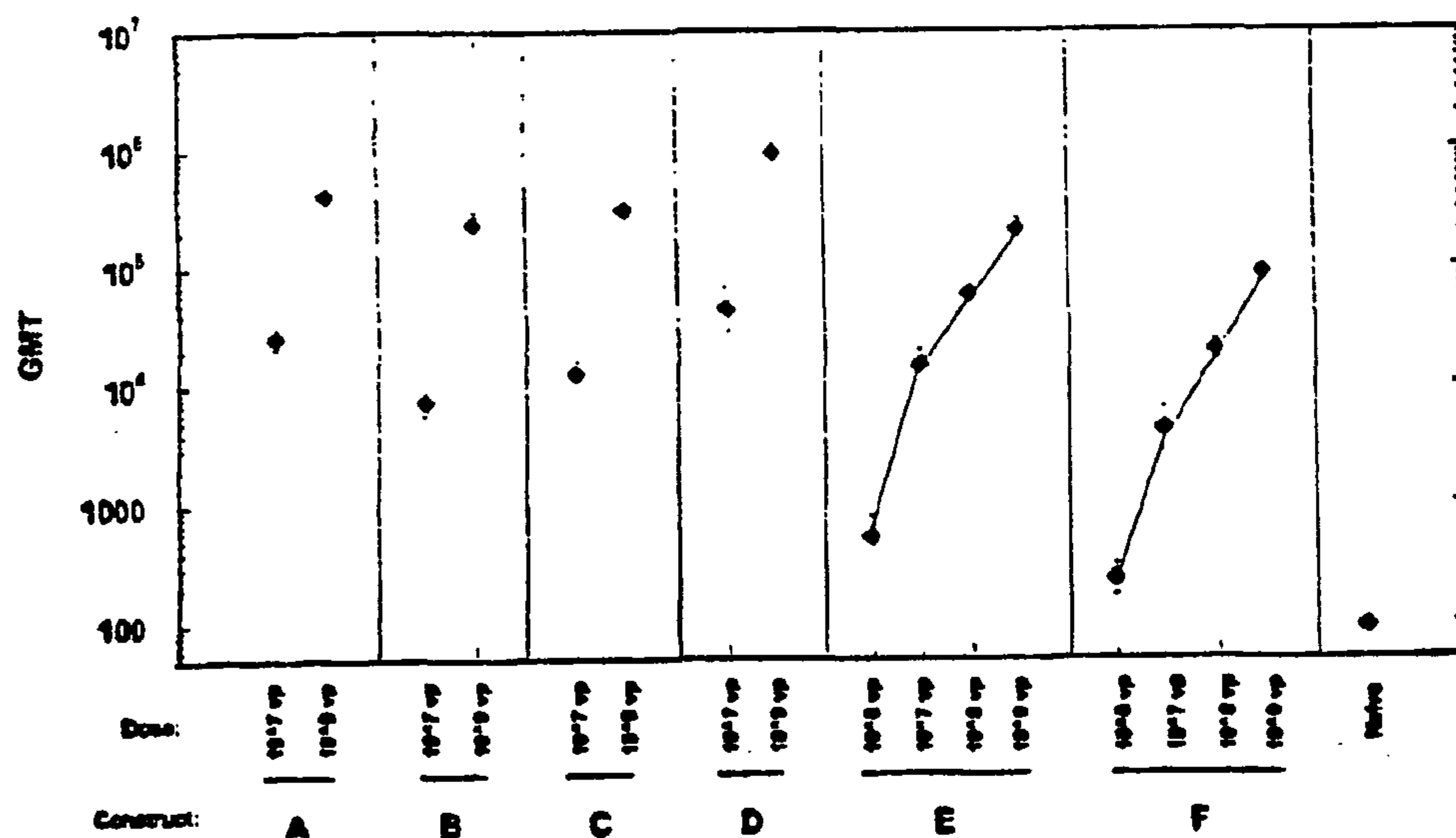


Figure 12: Viral DNA analysis by *Hind*III digestion on high passage numbers for MRKAd5gag in serum containing media with collections made at specified times. The first lane shows the 1 Kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and *Hind*III), and MRKAd5gag virus continually passaged to P16, P19 and P21 (serum containing media).

13
 Figure 13. Serum anti-p24 Levels at 3 Wks post i.m. immunization of balb/c mice (n=10) with Varying Doses of Several Adgag constructs: (A) MRKAd5gag (through passage 5); (B) MRKAd5 E3⁺ hCMV-FLgag-bGHpA; (C) MRKAd5 E3⁺ hCMV-FLgag-SPA; (D) MRKAd5 E3⁺ mCMV-FLgag-bGHpA; (E) research Lot (293 cell-derived) of Ad5HIV-1gag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-1gag. Reported are the geometric mean titers (GMT) for each cohort.



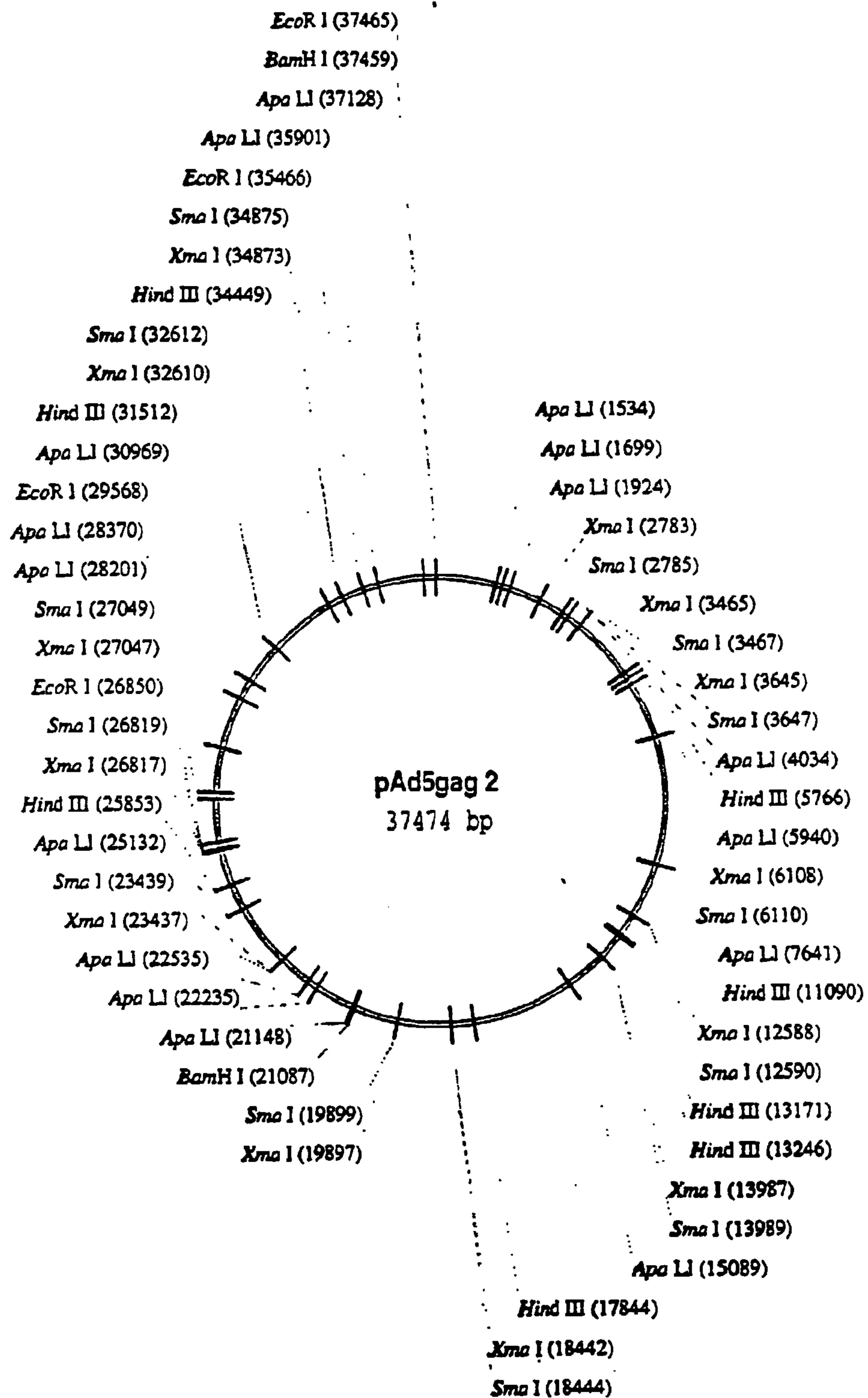


Figure 14

pMRKA159a9 MER682

1 TTTCTTAATTA ACATCATCAA TAAATATACCT TATTTTGAT TAAATCAAT ATGATAATCA GGGGTGGAG TTTGTGACGT GGGCCCGCG GGGGACCGG GTGGGACCG
 101 AAGAAATTAAT TGTAGTAGTT ATTATATGTA ACTTCTGTTA TACTATTACT CCCCACCTC AACACTGCA CCCCACCGG CACCTTCTG GTGTGTGGT GGTGTGTGGT
 201 GGGGGGTGAC GTAGTAGTGT GGGGTGATG GGGGTGATG TGAATGATG GGTAAATGGG GGTAAATGGG GGTAAATGGG GGTAAATGGG GGTAAATGGG GGTAAATGGG
 301 CCGCCCACTG CATCATCACA CCGCCCTCAC ACTACAACTG TCACACCGCC TTGTGTATAT TGTGTATAT TGTGTATAT TGTGTATAT TGTGTATAT TGTGTATAT
 401 GGTGTACACA GGAAGTGACA ATTTTCTGCT TAAATGCTG TAAATGCTG TAAATGCTG TAAATGCTG TAAATGCTG TAAATGCTG TAAATGCTG TAAATGCTG
 501 GAATAGAGG AGTGAAATC TGAATATAT TGAATATAT TGAATATAT TGAATATAT TGAATATAT TGAATATAT TGAATATAT TGAATATAT TGAATATAT
 601 CTTATTTCTC TTCACTTTAG ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA ACTTATTAA
 701 CAGGTGTTT TCTCAGGTGT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT TTTTCTGCT
 801 GTCCACAAA AGAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA AAGTCCACA
 901 ATATGTACAT TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC TTATATTGGC
 1001 TAGCCCATAT ATGAGATTCC GGTATACATA ACTTACGGT GTATATGGG TACAACTGTA ACTAAATNACT GATCAATAT GATCAATAT GATCAATAT GATCAATAT
 1101 ATCGGGTATA TACCTCAAGG CCGAATGTAT TGAATGGCT TGAATGGCT TGAATGGCT TGAATGGCT TGAATGGCT TGAATGGCT TGAATGGCT TGAATGGCT
 1201 GTTCCCATAG TAACGCCAAT AGGACTTTC GTACTGCTG GTACTGCTG GTACTGCTG GTACTGCTG GTACTGCTG GTACTGCTG GTACTGCTG GTACTGCTG
 1301 CAAGTACCCC CCTATTGAC GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG GTCAATGACG
 1401 GTTCAATGCG GGGATAACTG CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC CAGTTACTGC
 1501 CGTATTAGTC ATCGCTATTG CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT CAGTGGTGT
 1601 GCATAATCAG TAGCGATAAT GGTACCATA GGTACCATA GGTACCATA GGTACCATA GGTACCATA GGTACCATA GGTACCATA GGTACCATA GGTACCATA
 1701 ATTGACGTCA ATGGAGTTT GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG GTTTTGGCAG
 1801 TAACTGCACT TACCTTCAAA CAAACCGTG GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT GTTTTAGTGT
 1901 TACCGTGGGA GGTCTATATA AGCAGAGCTC GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA GGTCTAGTGA
 2001 ATGCCACCTT CCAGATATAT TCGTCTCGAG CAATCAGCTT GGCAGTCTAG CCGACTCTG CCGACTCTG CCGACTCTG CCGACTCTG CCGACTCTG CCGACTCTG
 2101 CCGATCCAGC CTCGCGGCG CCTCGGAGG CATTTGGAG GGTATTCCTT CCGATTCCTT CCGATTCCTT CCGATTCCTT CCGATTCCTT CCGATTCCTT CCGATTCCTT
 2201 GGCTAGGTGG GAGGCGGCG CCTTGGCAG GTAACCTTGC GTAACCTTGC GTAACCTTGC GTAACCTTGC GTAACCTTGC GTAACCTTGC GTAACCTTGC GTAACCTTGC
 2301 TGGTGAGCTG GACAAAGTGG AGAAGATCAG GGTGAGGCT GGTGAGGCT GGTGAGGCT GGTGAGGCT GGTGAGGCT GGTGAGGCT GGTGAGGCT GGTGAGGCT
 2401 ACCACTCGAC CTGTTTACCC CTCTCTAGTC CCACTCCGGA CCACTCCGGA CCACTCCGGA CCACTCCGGA CCACTCCGGA CCACTCCGGA CCACTCCGGA CCACTCCGGA
 2501 TTTGCTGTGA ACCCTGGCT GCTGGAGACC TCTGAGGGT TCTGAGGGT TCTGAGGGT TCTGAGGGT TCTGAGGGT TCTGAGGGT TCTGAGGGT TCTGAGGGT
 2601 AAACGACACT TGGGACCGGA GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG GACCTCTGG
 2701 CCTGTACAA CACAGTGGCT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT ACCCTGACT
 2801 GGGACATGTT GTGTACCGGA TGGGACATGA CACAGTGT CACAGTGT CACAGTGT CACAGTGT CACAGTGT CACAGTGT CACAGTGT CACAGTGT
 2901 GTCCAAGNAG AAGGCGCAG AGATGCTTC TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT
 3001 CAGGTCTTTC TTCCGGTCC TCCGACGAGT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT TGGACAGCTT

Figure 1SA

pMRKAI5494 MER682

1701 CACCAGGCCA TCCTCCCCCG GACCTGTAAT GCTTGGTGA AGCTGCTGGA GAGAGAGGCC TTCTCCCTCG AGGTGATCCC CATGTTCTCT GGCCTGCTG
 1801 GTGGTCCGGT AGAGGGGGCG CTGGGACTTA CGGACCCACT TTACAGACTT CTCTTCCCG AGAGGGGAC TCCACTAGCG GTACAAGAGA CCGACAGAC
 1901 AGGTGCCCAC CCCCCAGGAC CTGACACTA TCTGACACT TCTGACACT CATAGGCTG CATTGAGAT GCTGAGGAG ACCATCAATG AGAGGCTG
 2001 TCCCACGGTG GGGGTCTCTG GACTTGTGT ACTGACTTGT TCACTTCTG GCTGAGGAG CCGTACCTTA CGACTTCTC TGGTAGTTAC TCCCTCGAC
 2101 TGAGTGGGAC AGGCTGATC CTGTCGATC TATGCTGAT TATGCTGAT AGCTGAGGA CCGGCTGCTC TCTGACATTG CTGACACCAC CTCACCCCT
 2201 ACTCACCCTG TCCGACGTAG GACACGCTG ACCGGGTAA CCGGCTGCTC TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2301 CAGGAGCAGA TTGGCTGAT GACCAACAC CCCCCATC CTTGAGGGA TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2401 GTCTCTGCTT AACCGACCTA CTGGTCTG GGGGGTGG GAGGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2501 ACTCCCCCAG CTCCTATCTG GACCTGATC AGGCTGCTC CCGGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2601 TGAAGGCTG GAGGTAGGAC CTGTAGTCCG TCCCGGGT TCCCGGGT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2701 CAGGAGGCTG TGAAGGCTG GAGGTAGGAC CTGTAGTCCG TCCCGGGT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2801 GTCTCTGCTT AACCGACCTA CTGGTCTG GGGGGTGG GAGGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 2901 ACTCCCCCAG CTCCTATCTG GACCTGATC AGGCTGCTC CCGGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 3001 TGAAGGCTG GAGGTAGGAC CTGTAGTCCG TCCCGGGT TCCCGGGT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 3101 CAGGAGGCTG TGAAGGCTG GAGGTAGGAC CTGTAGTCCG TCCCGGGT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT
 3201 GTCTCTGCTT AACCGACCTA CTGGTCTG GGGGGTGG GAGGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT TATGCTGAT

Figure 15B

pMRKAd549ag MER682

3301 TTTGAGACTG CAGCTCCGC CCGCCCTTCA GTGCTGAG CAGCTCCGC CAGCTCCGC CAGCTCCGC CAGCTCCGC
 AACCTCTGAC GTCCGAGGCG GCGCCGAGGT CCGCTGAG CAGCTCCGC CAGCTCCGC CAGCTCCGC CAGCTCCGC
 3401 CTTCCCGTTC ATCCGCGCG CAGTCAAGT TCACTGCTT TTTGACAG TTTGACAG TTTGACAG TTTGACAG
 GAGGCGAAG TAGGCGGCG CTACTGTTCA ACCTGAGGAA AACCTGAGT AACCTGAGT AACCTGAGT AACCTGAGT
 3501 TCTGCGCCAG CAGGTTCTG CCGTTAAGG TCTGCGCCAG TCTGCGCCAG TCTGCGCCAG TCTGCGCCAG TCTGCGCCAG
 AGACCGCGTC GTCCNAGAC GCGACTTCG AGGAGAGGGA AGGAGAGGGA AGGAGAGGGA AGGAGAGGGA AGGAGAGGGA
 3601 GTGCTCTGCT GCTCTTATT AGGGTTTTC CAGGCTTTC CAGGCTTTC CAGGCTTTC CAGGCTTTC CAGGCTTTC
 CACGAGCGA CAGGATANA TCCCCANAC GCGCGGCCA TCCCGGCCA TCCCGGCCA TCCCGGCCA TCCCGGCCA
 3701 AAGGTGACT CTGGATGTC AGATACATG GCATACGCC CTCCTGCGG TCGAGGTAG ACCACTGAG ACCACTGAG ACCACTGAG
 TTTCCACTGA GACCTACAG TCTATGTACC CGTATTCGAG CAGAGAGCC CAGAGAGCC CAGAGAGCC CAGAGAGCC
 3801 GATCCAGTCT TAGCAGGAG CCGTGGCGG GTGCTTANA ATGCTCTTCA GTAGCAAGT GATGGCCAG GATGGCCAG GATGGCCAG
 CTAGGTCAGC ATGCTCTTCG CGACCTGAC CAGGATTTT TACAGGAGT CAGCTTCCA CAGCTTCCA CAGCTTCCA CAGCTTCCA
 3901 CCGTTAAGCT GCGATGGTG CATACGCTG GATATGAGT GCATCTTGA CTGTATTTT AGTTGCTA TGTCCAGC CATATCCCTC CCGCTTAT
 GCCAATTGGA CCTACCCAC GTATGCCAC GTATGACTA CAGTACACT CAGTACACT CAGTACACT CAGTACACT
 4001 TGTGTCAGC AACCAACAG ACAGGTATC CCGTCACTT GCGGATTTG TCACTTAGT TCACTTAGT TCACTTAGT TCACTTAGT
 ACACACGCTC TTGCTGCTG TGTACATAG GCGAGGTGA CCGTTTANAC ACTACATCGA ATCTCTCTT ATCTCTCTT ATCTCTCTT
 4101 ACCTCCAGCA TTTTCCATG ATCTGCTCT ATCTGCTCT ATCTGCTCT ATCTGCTCT ATCTGCTCT ATCTGCTCT
 TGGAGGTTCT AAGGTACG TAAAGGCTA TTAAGGCTA TTAAGGCTA TTAAGGCTA TTAAGGCTA TTAAGGCTA
 4201 TGTTCAGCA TGAGATGTC ATAGGCTAT TTAAGGCTA TTAAGGCTA TTAAGGCTA TTAAGGCTA TTAAGGCTA
 ACAAGGTTCT ACTCTAGCAG TATCCGTA AATGTTTTC CCGCGCTCT CCGCGCTCT CCGCGCTCT CCGCGCTCT
 4301 CCTCACAGAT TTGCTTTTC CACGCTTGA GTTCAGATG GCGATCATG TCTACCTGCG GCGGATGGA GCGGATGGA GCGGATGGA
 GGAGTGCTTA AACGTAAAG GTGCGNACT CAGTCTACC CCGCTAGTAC AGATGAGCG CCGCTACTT CCGCTACTT CCGCTACTT
 4401 CTGGGAGAA AGCAGGTTCC TGAGCAGCTG CAGCTGCTG CAGCTGCTG CAGCTGCTG CAGCTGCTG CAGCTGCTG
 GACCTCTCTT TCGTCCAAG ACTCGTCGAC GCTGATGCG GCTGATGCG GCTGATGCG GCTGATGCG GCTGATGCG
 4501 CAGCTGCGGT CATCCCTGAG CAGGCGGCC ACTTCTTAA GCATGCTCT GATCTCTAG TTTTCCCTG CAGAAATCCG CAGAAATCCG
 GTGAGCGCA GTAGGACTC GTGCGCCCG GTGCGCCCG GTGAGCAAT TGTAGCTGAC CAGAAATCCG CAGAAATCCG CAGAAATCCG
 4601 GCGATAGCAG TCTTGCNAG GAGCAAGT TTTCAAGCG TTTGAGAGCG TTTGAGAGCG TTTGAGAGCG TTTGAGAGCG
 CCGTATCTCT AAGACGTTT CTTCTGTTCA AAGATGCGC AAGATGCGC AAGATGCGC AAGATGCGC AAGATGCGC
 4701 CCACAGCTCG GTACCTGCT CTACGCAAT TCGATCCAG ATATCTCTC GTTCTCTCT GTTCTCTCT GTTCTCTCT
 GGTGTCAGC CAGTGGAGA GATGCGCTG AGTATGCTG TATAGAGAG CAGGAGGCG CAGGAGGCG CAGGAGGCG
 4801 CCAGACGCG CAGGCTCAT TCTTTCAGC CAGGCTCAT CAGGCTCAT CAGGCTCAT CAGGCTCAT CAGGCTCAT
 GGTCTGCCCC GTCCAGTAC AGAAGGTC CCGGTTCCA CCGGTTCCA CCGGTTCCA CCGGTTCCA CCGGTTCCA

Figure 15c

pMRKad5gag MER682

4901 GGTGGCGCTTG AGGCTGATCC TCTCTGTGCT GAAAGATTCG CCGTCTTTCG CTTGGTGTTC GGGCAGTTAG GGTGATTA GATTTGACCA TGGTGTGATA GTCCAGTCCC
 CCACGGGAAAC TCCGACCAAG AGACCAAGG CTTGGGAGG GACAGAGAGG CCGGTTCATC GTAAACTGGT ACCACAGTAT CAGGTGCGGG
 5001 TCCCGCGCGT GCGCTTTGDC GCGCAGTTG CCGTTTAAAG AGTGTTCGCA CTAAGTATCAG TTAGGGCGTA GAGCTTGGG GGTAGTAATA
 AGGCGCGCGCA CCGGGAACCG CCGGTGCAAC GCGATCAGC TCCGCGCTT GTTCGCGTC AGTCTGTAAC ACTCCCGCAT CTCGAACCGG CGCTCTTTAT
 5101 CCGATTCCGG GGTGTAGACA TCCCGCGCGC AGGCGCTTCA GATGTCTTTC CATTCGACGA GGTAGTTTAG CTCTGGCCGT TCTGGGTCAA AATCAGGTT
 GGTAAAGGC CTTCAATCGT AGCGCGCGCG TCCGGGCGT CTGCGAGAGC GTAGGTTGCT CGGTCCACTC GAGACCGGCA AGCCCGAGTT TTGTTGTTCA
 5201 TCCCGCAATGC TTTTGTATGC GTTCTTACC TCTGTGTTCC ATGAGCGCT GTCCAGGCTC GGTGACGAAA AGGCTGTCCG TGTCCCGTA TACAGACTT
 AGGCGGTACG AAAAATAAG CAAGATAGG AGACCAAGG TACTGTGCGA CAGGTGCGAG CCACTGCTTT TCCGACAGGC ACAGGGGCAT ATGTTGTGAA
 5301 AGAGGCCCTGT CCTCGAGCGG TGTTCCTGCT ATAGAACTC GGTACACTCT GATGACAAAG GTGCGTTCCA GGTCCAGCAC AGGAGGCTTA
 TCTTCGGACA GAGCTGCGC ACAGCGCGC AGAGGAGCA TATCTTTGAG CTTGTGTTCC GAGCGCAGGT CCGGTGCTGC TTCTTCCGAT
 5401 AGTGGAGGG GTAGCGGTCG TTGTCTACTA GGGGTCCAC TCTCTCCAGG GTGTGAGAGC AGATGTGCGC CTCTTGGGCA TCAAGGAAAG TGAATTTGTT
 TCACCCCTCC CATCGCAGC AACAGGTGAT CCGCCAGGTG AGGAGGTCC CACACTTCTG TGTACAGCGG GAGNAGCGGT AGTTCTCTCC ACTAACCAAA
 5501 GTAGGTGTAG GCCAGGTGAC CCGGTGTTCC TGAAGGAGG CTATATAAGG GGTGTGCGC CCGTTCTCTT CCGCATGCTT GTCTGCGAGG
 CATCCACATC CCGTGCATG GCCACAAAG ACTTCCCGC GATATTTTCC CCGACCGCGC CCGAAGCAG AGTGAGAGAA GGTGTAGGGA CAGACGCTCC
 5601 GCCAGCTGTT GGGGTAGTA CTCTCTTGA AAGCGGGA TCACTTCTGC GTTAAGATTG TCACTTTTCCA AATCAGGGA GATTTTGTATA TTCACCTGTT
 CGGTGAGCAA CCGCACTCAT GAGGTAGACT TTTCGCGCGT ACTGAGAGCG CGATTCTTAC AGTCAAAAGT TTTTCTCTCT CCTAAACTAT AGTGTGAGCT
 5701 CCGCGGTGAT GCTTTGAGG GTGCGCGCAT CCATCTGTTC AGAAGAGACA ATCTTTTGTG TGTCAAGCTT GGTGCGAAC GATCCGTAGA GGTGTTGTA
 GCGCGCACTA CCGAACTCC CACCGCGGTA GGTAGACCG TCTTTTCTGT TAGAANACA ACAGTTGGA CCACCGTTTG CTGGGCATCT CCGGCAACCT
 5801 CAGCAACTTG GGTATGAGC GCAGGTGTTG GTTTTGTGCG CGATGCGCGC GCTCCTTGGC CCGGATGTTT AGCTGCAGCT ATTGCGCGC AACGCACCT
 GTCTGTGAAAC CCGTACCTCG CGTCCCAAC CAAAAACAG CGTAGCGCGG CCGAGAACCG CGGTACAAA TCGACGTGCA TAAGCGCGCG TTGCGTGGG
 5901 CATTCGGGA AGACGGTGT GCGCTCGTCC GGCACAGGT GCACGTGCTA ACCGTGGTTG TGCAGGGTGA CAGGTCAAC GCTGTGCTGT ACCTCTCTCG
 GTAGGCCCTT TCTGCCACA CCGGAGCAGC CCGTGTGCTA CGTGCGCGT TGTGCGCAGC ACGTCCACT GTTCCAGTTG CGACCAACGA TCGAGAGG
 6001 GTAGCGCTC GTTGTCCAG CAGAGCGGC CCGCTTGGG CCGAGCAGAT GGTGTAGCTG GGTCTAGCTG CGTCTGCTCC GGTGGGTCTG CGTCCACCTT
 CATCCCGCAG CAGCAGGTC GTCTCGCGC GCGGAACTC GCTGTCTTA CCGCATCC CCGATCGAC CCGATCGAC GCGAGCAGG CCGCCAGAG CAGGTGCTTA
 6101 AAGACCCCG GCGAGCAGC CCGGTGGA GTAGTCTATC TTGATCTT TCGATCTAG GGTGTAGCTG GGTCTAGCTG CGTCTGCTCC GGTGGGTCTG CGTCCACCTT
 TTTCTGGGC CCGTGTCC CCGGCGCTT CATCAGATAG AACGTAGGA CTTTCAGATC GGTGAGGAG GTACGCGCC GTAGTTGCG CCGGAGCATA
 6201 GGTGTAGTG GCGACCCA TCGCATGCG TGGGTAGCG CCGATGCTTA CATGCGCA ATGTCTTAAA CGTAGAGGG CTCTCTGAGT ATTCCAAGAT
 CCCAACTCAC CCGTGGGT ACCGTACCC ACCCACTGC GCTTCCCAT GTACGGGCTT TACAGCATTT GCATCTCCCG CAGAGACTCA TAAGTTCTTA
 6301 ATGTAGGTA GCATCTTCA CCGGTATTC TGTGTGAC GTATCTTAT GTTGTGCG ATGTGCGG AGGTGCGA CCGAGTTGC TACGGCGG
 TACATCCCAT CGTAGAGGT GCGGCTTCC ACCGCGTG CATTCGATA TCAAGCAGC TCCCTGCTC CTCAGCGCT CTCAGAGCT ATGCGCGC
 6401 CTGCTCTGT CCGAAGACTA TCTGCTGAA GATGATGT GATTTGATG ATATGTTG ACCTGTAG AGCTTCAAG CCGTCTCTGT GAGACTTAC
 GACGAGACGA GCTTCTGAT AGACGACTT CTACCGTACA CTCACCTAC TATACCAAC TCGGACCTTC TCGAACTTC ACCGAGACA CTCTGATG

Figure 150

pMRKad5gag MFR682

6501 GCGTCACGCA CGAAGGAGGC GAGAGAGTGC CCGAGCTTGT TGACCAAGTC GCGCTGACC TCACCTCTA GCGCCAGTA GTCCAGGGTT TCCCTGATGA
 CCGAGTGGT GCTTCTCCG CATCTCAGC GCTTCGACA ACCTCTCAG CCGCACTG ACCTGAGAT CCGCGTCAT CAGGTCCCA AGGACTACT
 6601 TGTCATACTT ATCTGTCC TTTTCTTC ACAGCTGAG GTTGAGTACA AACTTTCG GCTCTTCA GTACTCTGG ATCGGAACC CGTCGCTT
 ACAGTATGAA TAGGACAGG AAAAAAAGG TCTGAGGC CACTCTCT TTGAGAGCG CAGNAAGGT CATGAGAACC TAGCCTTTGG GCAGCCCGAN
 6701 CGAACGTAA GNGCTTAGCA TGTGACTG GTTAGGKX TGTAGGCG AACTCTCT TTCTACGGT AGCGGTATG CCGCGCGGC CTTCGAG
 GCTTCCCATTT CTGGAATCT ACATCTTAC CACTGCGG CACTGCGG TCTTAGGGA AAGATGCCA TGCGCATAC GCGCGCGCG GAGGCCCTT
 6801 GAGGTGTGG TGAGCGCAA GGCTGCTG ACCATGACTT TGAGTACTG GTATTGAG TCACTGCTT CCGATCCGC CTGCTCCAG AGCAAAAGT
 CTCCACACCC ACTCGCTTT CCACAGGAC TGTACTGMA ACTTCATAC CATANCTTC AGTCACAGCA GCGTAGGCG GCGTAGGCTC TCGTTTTC
 6901 CCGTGGCTT TTGGAACGC GGATTGGCA GCGGAGGT GACATCTCTT AAGATATCT TTTCTGCGG AGCATTAAG TTGCGTGTGA TCGGNAAGG
 GGCACCGGA AAGCTTTGG CCTAACCCT CCGCTTCCA CTGTAGCA CTTGTAGCA AAGGCGGC TTTCTGAA TCCGTATTT AACCCACACT ACCCTTCT
 7001 TCCCGGACC TCGGAACGT TGTAAATAC CTGGGCGG AGCAGATCT CCGTAAAGC GTTGATCTTG TCGCCACAA TGTAAAGTTC CAGNAAGCG
 AGGCGGTGG AACTTTCCA ACATTAATG GAGCGCGC TCTGTCTAGA GCACTTTCG CACTACAMC ACCGGGTGT ACATTTCAAG GTTCTTCTCT
 7101 GCGATGCCCT TGATGNAAG CAATTTTTA AGTCTCTGT AGGTAGCTC TCTAGGAG TTTCTGCGG CTTCTGAA GGTCTGAAAG GCGCCAGTCT GCAAGATGAG
 CCTACCGGA ACTACCTTC CTCAACAGT CACGCGCAT TAGTATTTG TCGATTTG ACCTGCTCT GACTCGCA CAGACTTTT CCGGTTCAG COTTTCTACT
 7201 GGTGGAAGC GAGGATGAG CTCAACAGT CACGCGCAT TAGTATTTG TCGATTTG ACCTGCTCT GACTCGCA CAGACTTTT CCGGTTCAG COTTTCTACT
 CCAACCTTC CTGCTTACT GAGGTGTCA GTGCGCGTA ATCTTAACG TCGACAGCG CTTTCAAG CTTCTGCGG TTTTCTG CCAAAAGACC
 7301 GGTGATGAG TGAAGTAA GCGGTCTTG TTCCAGCG GTTCGCGG TAGGTCTGC GNAAGTCTT AACTTGGCA CTTGACCGT GGTATCCGCT ATCTCTGCG
 CCACTAGCT ATCTTCCAT CCGCCAGAC AAGGTCTGC AGGTAGTTT CTAAAGCGG TCGATCTGC ATCCAGAGG ATCCAGAGG CCGCTCAGT TAGAGCGCG
 7401 AACTTCATGA CAGCATGA GCGCAGAT TCTTCCCA AGGCTTCTAT CCAAGTATG GTCTCTACAT CGTAGTGA AAGAGAGCG TCGGTGCGA
 TTGAAGTACT GGTCTACTT CCGGTCTCG ACGAAGGT TCGGCGTA GGTTCATATC CAGAGATGA GCATCCACTG TTTCTCTCG AGCCACGCT
 7501 GATGGAGCC GATCGGAG AACTGGATCT CCGCCACCA ATTGAGTAG TGTGTGAA GTAGAGTCC GTAGAGTCC CTGCGAGCG CCGAACACTC
 CTACGCTCG CTAGCCCTTC TTGACCTAGA GCGCGTGT TAACCTCTC ACCGATTAAT ACACACTTT CATCTTCA GACGCTGCC GCGTTCTGAG GCGTTCTGAG
 7601 GTGCTGGCTT TTGTAAAC GTGCGCAGTA CTGCGAGCG TCGACGCT GTACATCTG CACGAGTTG ACCTGACGAC CCGGCACAG GAGGACAGT
 CAGGACGGA AACATTTTG CAGCGTCTAT GACGTCGCC ACCTCCCA CATGTAGAC GTGCTCCAC TGGACTGCTG GCGGTGTTT CTTCGCTCTA
 7701 GGGATTTGA CCCCCGCC TGGCGGTT TTTCTACTT GGTGCTGT CTTGACGT CTGCTGCTC GAGGAGGT GAGGAGGT ACCGTGAT
 CCTTTAACT CCGGAGCGG ACCGCCCAA GAGATGAG CCGACGACA GAACTTGA CCACTTGA GACCGAGG CTCCCTCA TCGCACCTA
 7801 GAGCCACCAC GCGCGGAG CCAAGTCC AGATGTCGC GCGCGCTT CTGAGCTTA TCAACATC GCGCAGATG GAGCTGTCCA TGGTCTGAG
 CTTGCTGCTG CCGCGGCTC GGGTTTCAG TCTACAGCG CCGCCGCA GCTTCACT ACTGTTGTAG CCGTCTACC CTGACAGGT ACCAGACTC
 7901 CTCGCGCGC GTCAGTCA GCGGAGCT CTCAGTTT ACCTGCTA GAGGCTCAG GCGCGGCT AGATCCAGT GATACCTAAT TTCCAGGCG
 GAGGCGCGC CAGTCCAGT CCGCTTCAG GCGCTTCAG GACGTCCAA TCGACCTAT CTGCGCAGT CCGCGCGCA TCTAGTCCA CTATGATTA AAGGTCCCG
 8001 TGGTTGCTG CCGGTGAT GGTTCAG AGGCTGATC CCGCGGCT GACTTCTTA CTGCGGCT CCGCGGCT GCGGCTGAG TCCCTGAT
 ACCAACCC GCGCAGCTA CCGNACCTT TCGGGGTAG GCGGCTGCT CTGATGCTAT GCGCGGCT CCGCCACCG GCGCGGCTC AGGACTTA

pmrkad59ag MFR6B2

```

8101 ATGCATCTAA AGCGGTGAC GCGGCGGAC CCGCGAGCT AGTATATCT CCGGACCGC CCGGAGAGG GGCAGGGGA CGTGGGGCC GCGGGGGT
TACGTAGATT TTGCGCACTG CCGCGCTCG GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG GCGGCGCGT
8201 AGGAGCTGCT GCTGCGCGG TACTTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG GCGGCGCGT
TCTTCGACCA CGACGCGCG ATCCAGGAC CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
8301 GCTTGAACCT GAAAGAGAT TCGACAGAT CAATTTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
CGAACTTGA CTTTCTCTCA AGCTGTCTTA GTTAAGACA CAGAACTTCA GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT
8401 GATCTCGGC ATGAATGCT GATCTCTCT CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
CTAGAGCGCG TACTTGACGA GCTAGAGAG GATGAGCTCT AGAGGCGCG GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT
8501 TCGAGAGAG GCTTGAAGC TCGTCTCTCT CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
ACGCTCTTCC GCAACTTCC GAGAGCGAG GCTTGGCGCG ACATCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT
8601 CCACGTCGG GCGGAGAGC GCGTCTCTCT CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
GCTGACCGC CCGTCTCTCT CCGTCTCTCT CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT CAGCGCGCG
8701 TCGCAACGT GATTCGTTG TATCCCGCA GCGCTCTCT CCGTCTCTCT GCGGACTTCA TCTCTCTCA GCGCTGGCG GCGCTCTCC CCGTCCCGT
AGGTTGAC CTRAGCACT ATAGGGGTT CCGGAGTTCC GCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT
8801 ACGTTAACT CTTCTTCCG AAGACGGAT AGCTCGCGA CAGTCTCTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT
TGCCAATTGA GAGGAGGTC TTCTGCTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT
8901 CTTCCATAG GCGCTCTCT TCTCTCTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT CCGGACTTCT
GAGGTATT CCGGAGGGA AGAGAGAGA GAGGAGGGA GAGGAGGGA GAGGAGGGA GAGGAGGGA GAGGAGGGA GAGGAGGGA GAGGAGGGA
9001 GATCATCTC CCGGCGGAC GCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
CTAGTAGAG GCGGCGGCT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
9101 GCTGCGCGG GCTGCGCAT GCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
CAACCGCGC CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
9201 CATCGACCG ATCGGAAC CTTCTGAGA AGGCTCTAA CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
GTAGCTGCG TAGCTTTTG GAGAGCTCT TCGGAGGAT TCGGAGGAT TCGGAGGAT TCGGAGGAT TCGGAGGAT TCGGAGGAT TCGGAGGAT
9301 GTTGTCTCT GCGGAGGTC TGCTGATAT GTATTTAAG TATGCTCTCT TATGAGGCG GATGCTCTCT TATGAGGCG GATGCTCTCT TATGAGGCG
CAACAAAGAC CCGCTCCAC CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
9401 TGAATGCGA GCGGAGGTC CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
ACTTACCGT CCGGAGGTC CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
9501 CTTCTCTCT TGTCTCTCA TCTCTCTCT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
GAGGAGGAG AACAGGAGT AGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT
9601 GCGGAGGAG GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA GCGTGAACA
CGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT CCGGAGGAT

```

Fig III

Fig IV

Fig V

Fig VI

Fig VII

Fig VIII

Fig IX

Fig X

Fig XI

Fig XII

Fig XIII

Fig XIV

Fig XV

Fig XVI

Fig XVII

Fig XVIII

Fig XIX

Fig XX

Fig XXI

Fig XXII

Fig XXIII

Fig XXIV

Fig XXV

Fig XXVI

Fig XXVII

Fig XXVIII

Fig XXIX

Fig XXX

Fig XXXI

Fig XXXII

Fig XXXIII

Fig XXXIV

Fig XXXV

Fig XXXVI

Fig XXXVII

Fig XXXVIII

Fig XXXIX

Fig XL

Fig XLI

Fig XLII

Fig XLIII

Fig XLIV

Fig XLV

Fig XLVI

Fig XLVII

Fig XLVIII

Fig XLIX

Fig L

Fig LI

Fig LII

Fig LIII

Fig LIV

Fig LV

Fig LVI

Fig LVII

Fig LVIII

Fig LIX

Fig LX

Fig LXI

Fig LXII

Fig LXIII

Fig LXIV

Fig LXV

Fig LXVI

Fig LXVII

Fig LXVIII

Fig LXIX

Fig LXX

Fig LXXI

Fig LXXII

Fig LXXIII

Fig LXXIV

Fig LXXV

Fig LXXVI

Fig LXXVII

Fig LXXVIII

Fig LXXIX

Fig LXXX

Fig LXXXI

Fig LXXXII

Fig LXXXIII

Fig LXXXIV

Fig LXXXV

Fig LXXXVI

Fig LXXXVII

Fig LXXXVIII

Fig LXXXIX

Fig LXXXX

Fig LXXXXI

Fig LXXXXII

Fig LXXXXIII

Fig LXXXXIV

Fig LXXXXV

Fig LXXXXVI

Fig LXXXXVII

Fig LXXXXVIII

Fig LXXXXIX

Fig LXXXXX

Fig LXXXXXI

Fig LXXXXXII

Fig LXXXXXIII

Fig LXXXXXIV

Fig LXXXXXV

Fig LXXXXXVI

Fig LXXXXXVII

Fig LXXXXXVIII

Fig LXXXXXIX

Fig LXXXXXX

Fig LXXXXXXI

Fig LXXXXXXII

Fig LXXXXXXIII

Fig LXXXXXXIV

Fig LXXXXXXV

Fig LXXXXXXVI

Fig LXXXXXXVII

Fig LXXXXXXVIII

Fig LXXXXXXIX

Fig LXXXXXXX

Fig LXXXXXXXI

Fig LXXXXXXXII

Fig LXXXXXXXIII

Fig LXXXXXXXIV

Fig LXXXXXXXV

Fig LXXXXXXXVI

Fig LXXXXXXXVII

Fig LXXXXXXXVIII

Fig LXXXXXXXIX

Fig LXXXXXXX

Figure 15F

pMRKad59ag MER6B2

9701 ACAAAAGCGGT GGTATGCGCC CGGTGTGATG GTGTAAATGC AGTTGGCCAT AATGGACAG TTAACGGTCT GGTACCCCG CAGCCTCTCG AGCCACATATG
 TGTTCGGCCA CCATACCGCG GCACAACTAC CAAATTCAG TCAACGCTTA TTGCTTGGTC AATGGCCAGA CCACTGGGCC GAGCCTCTCG AGCCACATATG
 9801 TGAGACGGGA GTAAAGCCCTC GAGTCAATAA CTTATGCTTT GAAAGTCTGC AATGATCCAC GGTATGCCAC CAANAATGTC GGCCTGGGCT GGCCTTAGAC
 ACTCTGGCCT CATTGGGAG CTTAGGTTAT GCATCACTAA CATTACAGCG TCTTCACTCA CCAATAGGTC GTTTTTCACG CCGCCGCCGA CCGCCATCTC
 9901 GGGCCAGCGT AGGTGCGCG GGGCTCCGAG GCGCAATCT TCCAACTATA GCGGATGATA TCTGTAGATG TACCTGGACA TCCAGGTGAT GCGCGGCGCG
 CCGGTCGCA TCCACCGGC CCGGAGGCC CCGCTCTAGA AGGTTGTATT CCGCTACTAT AGGATCTAC ATGGACCTGT AGGTCCACTA CCGCCGCCGA
 10001 GTGTGGAGG CCGCGGAMA GTGCGGAGC CCGTTCGAGA TGTTCGCGAG CCGCAAMAG TGTTCATAG TCGGACGCT CCGCGGCGCT AGGCGGCGCT
 CACCACCTCC GCGCGCCTT CAGCGCCTGC GCGAAGGCT ACAAGCGTC GCGGTTTTC ACCAGGTACC AGCCCTGCGA GACCGGCCAG TCCCGCGCG
 10101 AATGTTGAC GCTCTAGACC GTCCAAAGG AGAGCCCTGA AGCGGCACT CTTCCTGCT CTGTGGATA AATTCGCAG GGTATCATCG CCGACGACCG
 TTAGCAACTG CGAGATCTGG CACGTTTTC TCTGGACAT TCGCCCGTGA GAGGACCTAT TTAAGCGTTC CCACTAGTACC GCTTGTCTG
 10201 GGTTCGAGC CCGTATCCG GCGTCCCTC GTGATCCATG CGGTTACCG CGGCTGCTG AATCCAGTG TCGGACGCTA GACACCGCG GAGTGTCTCT
 CCGAAGCTCG GCGCATAGG CCGCAGGCG CACTAGGTAC GCGAATGAG GCGGACAGC TTGCTGCCAC ACCTGCAAT CCGTTCGCTT CCGTACGAGA
 10301 TTGCGCTTC TTCCAGGCG GCGGCTCTCT GCGTAGCTT TTGCTCCAC TTGCTGCTG TCGGCGCGC CAGCTGATC CAGATGCTC CCGTGTCTCG
 AACCGAAGG AAGTCCGCG CCGCGGACGA CCGGATCGAA AATCCGCTG AATCCGCTG AATCCGCTG CAGATGCTC CAGATGCTC CCGTGTCTCG
 10401 GCTGCTTCC TGTAGCGGA GGTATATTT CCAAGGCTG AGTTCGGA CCGCGGCTC GAGTCTGGA CAGTCTGGA CCGCGGCTG CCGCGGCTG
 CAGCGGAGG ACATCGGCT ACATCGGCT CCGGATGGA CCGGATGGA CCGGATGGA CCGGATGGA CCGGATGGA CCGGATGGA CCGGATGGA
 10501 CTCCCGCTC TCCAGGCTC CCGTTCGGA TTCTTCGGA AACAGGAGG AGCCCTCTT TTCTTTTTC CAGATGCTC CAGATGCTC CCGTGTCTCG
 GAGGCGAGT ACGTCTGCG CCGAAGCTT AAGGAGCTT TTCTCCCTG TCGGGAAMA AATCCGCTG CAGATGCTC CAGATGCTC CCGTGTCTCG
 10601 CCGCTCTCT AGCAGCGGA CAGCGGAGA CAGCGGAGA CAGCGGAGA CAGCGGAGA CAGCGGAGA CAGCGGAGA CAGCGGAGA CAGCGGAGA
 GCGGAGGAG TCGTGGCTT TCTCTCTC GTGCGCTCT GTGCGCTCT GTGCGCTCT GTGCGCTCT GTGCGCTCT GTGCGCTCT GTGCGCTCT
 10701 CCGCAGCAG TGTGTATTAC GAACCCCGC GCGCGGCGC CCGGACTAC CTTGACTTGG AGGAGGCGA CCGGAGGCGA CCGGAGGCGA CCGGAGGCGA
 GCGTCTCTT ACCACTAATG CTTGCGCGC CCGCGGCGC CCGGACTAC CTTGACTTGG AGGAGGCGA CCGGAGGCGA CCGGAGGCGA CCGGAGGCGA
 10801 TGAGCGGCAC CCAAGGCTG AGCTTAAGC TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT
 ACTGCGCGTG GGTTCCTACG TCGACTTCT ACTATGCGA CTTGCGGAGT TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT TGTATAGCGT
 10901 ATGCGGATC GAAAGTTCCA CCGAGGCGC GAGCTGCGC ATGCGGCTGA TGTGAGGCG TGTGAGGCG TGTGAGGCG TGTGAGGCG TGTGAGGCG
 TACGCTCTAG CTTTCAGGT GCGTCCCGC CTTGAGGCG TACGCGACTT AGCGCTCGC AATGAGCGC TCTCTCTGAA ACTCGGCTG CCGCTTGGC
 11001 GATATTAGTC CCGCGCGCA CAGGTGCGG CCGCGGACT CCGTACCGCA TACGAGCAGA CCGTGAATCA GCGATTAAC TTTCAAAMA GCTTTAACA
 CCTAATCAGG GCGCGCGGT GTGACCGCG GTGCGCTGA CCAATGGGT ATGCTGCTCT GCGACTTGT CCGCTTAATG AATGTTTTT CCAATTTCT
 11101 CCACCTGCGT ACGTCTGCG CCGCGGAGGA GGTGCTATA GCGCTGATG CCGTCTGGA ATCTGCGA CTTGTAGC GCGCTGAGC AANCCGAA TAGCAGCTG
 GGTGCAACA TCGGACACC CCGCGCTCT CCACTGATAT CCGTCTGGA TAGACACCTT GAACTATCG CCGGAGCTCT TTTTGGTTT ATCTGCTGCG
 11201 CTCATGCGC ACGTCTCTT TATAGTGCAG CAGAGTATG ACAGTATG ATTCAGGAT GCGTCTGA ACATAGTAG GCGGAGGCG CCGTGGCTG
 GATACCGCG TCGACAGGA ATATCAGTC GTGCTGCTC TATGCTCTA CCGGAGGAT TGTATCATCT TGTATCATCT CCGGCTCTCG GCGGCTGCG

Figure 156

pMRKAd5gag MER6B2

11301	TCGATTTGAT AACATCTCTG CAGAGGATAG TGGTTCAGGA GGTGAGTTTG AGCTTGGCTG ACAGGTGCG CGCCATCAAC TATTCCATGC TTAGCTGCG
11401	AGCTAAACTA TTGTAGGAC CTCCTGATG ACCAGTTCT ACCATTGAC TTGTACAGG TTTTCCACCG GCGGTAGTTG ATAGGTACG ATCGGACCT
11501	CAAGTTTATC GCGCGAAGA TATATCATAC CCGTTACGTT CCGATATAC ACGATATAC TCTTCCATTT CTAGCTCCG AAGATGTACG CGTACCAGCA CTTCACCG
11601	GTTCATAATG CCGCGTTCT ATATGATAG GCGAATGCA GGTATCTCT TCTTCCATTT CTAGCTCCG AAGATGTACG CGTACCAGCA CTTCACCG
11701	ACCTTGAGCG AGGACCTGGG CGTTTATGCG AACGAGGCA TCCATATGCG CCGTACGCTG GCACTGCGAC TCGGCTGAGTC CGCTGAGTC TCGGCTGAGTC
11801	TCGAACTCGC TCGTGGACCG GCAATATGCG TTTCTGCGT TTTCTGATAG AGGAGGCGAG TCCCTACTTTG ACCGCGGCG TCGGCTGAGTC CGCTGAGTC
11901	GGCTGCAAG GCGCTTGGCT GGCATGCGCA CCGTACGCT CCGTACGCT TCCCTACTTTG ACCGCGGCG TCGGCTGAGTC CGCTGAGTC
12001	CGGACGTTTC CCGGACCGA CCGTACGCT GGCATGCGCA TCCCTACTTTG ACCGCGGCG TCGGCTGAGTC CGCTGAGTC
12101	CTCTCCGCA TTCTGGAGC GGTGTTCCG GCGCGCGCA ACCGACGCA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
12201	GAGAGCGGTT AAGACCTTCG CCACCGGCG GGTGTTCCG TCGGCTGAGT CTACTAGCTT CTGCTTCCG TCGGCTGAGT TCGGCTGAGT TCGGCTGAGT
12301	GGCCCGGCA GCGCGGCTG GTCATGAGC GCGCGGCGA ACCGACGCA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
12401	CCGCGCTGCT CCGCGCGGAC CAGATGCTGC GCGCGGCGA ACCGACGCA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
12501	TGTGCGGAG GCGCTGCGC AGCGTGAGCG CCGCGGCGA ACCGACGCA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
12601	ACACGCGCTC CCGCACCGCG TCGCATCTGC GCGCGGCGA ACCGACGCA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
12701	GTGCGCGCGG GACAGGAGGA CTACACCAAC TTTGTTAGCG CACTGCGCT ATGTTGAGT GAGACACCGC AAGTGTAGGT GTTACTTCCA CATGCTCTGA
12801	CACGCGCGCC CTGTCTCTCT GATGTGGTTG AACACTCGC GTGACGCGCA TTTCCACTGA CTCTGTGCGG TTTACTTCCA CATGCTCTGA
12901	ATTTTCTCCA GACCAAGTAG CAAGGCTCG AGACCGTAA CCGTACGCGA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13001	TAAATAAGGT CTGCTCATCT GTTCCGAGCG TCTGCTATTT GCACTCTGTC CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13101	CCGCGCGGAC GTGCTTAGCT TCGTACGCG CAACTGCTC CTGTTTCTG CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13201	GGCGCGCTGG CACAGATCGA ACGATGCGG GTTACGCGG GTTACGCGG GAGACGCGA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC
13301	CTAGGTCACT TCGTACACT GTACGCGGAG GCCATAGGTC AAGTGTAGT GAGACGCGA CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC
13401	GATCCAGTGA ACGACTGTGA CATGCGCTC CGGTATCCAG TCGGCTGAGT CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13501	AGGAGGACAC GCGCAGCTG GAGGCACTC TAACTACTT GTTACTTAC CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13601	TCTCTCTGTG CCGCTCGGAC CTCCTGTTCT ATTTCTATGA CCACTTCTG GAGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13701	CATTCTGCGC TACGTCAGC AGACCTTCTG CCTTACCTG ATGTTGAGT GAGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC
13801	GTAAACGCG ATGCACTCG TCTTCTACT CCGATTTGAC TACTCTTTC CCGTACGCT TTTCTGATAG AGGAGGCGAG TCGGCTGAGTC CGCTGAGTC

Figure 15H

pMRKAd5₉ag MER6R2

12901 GGCATGTATG CCTCAGACCG GCGCTTTATC AACGCGCTAA TGTATCTACTT GATATGCTGT GGTGCTCTTA ACCCCGAGTA ACCCCGAGTA TTTCACCAAT TTTCACCAAT GGTATCTTGA
 CCTATACATAC GGAGTTTGGC CCGCAATAG TTTCTCTATT ACCCTATCA ACCCTATCA CCGCTCTAT TGGCTCTAT TGGCTCTAT TGGCTCTAT TGGCTCTAT TGGCTCTAT TGGCTCTAT TGGCTCTAT
 13001 ACCCGACATG GCTACCTGCG CCGCTTTCT ACACGCTGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT ATTCGATGCT
 TGGCGGTAC CGATGGCGG GACCAAGGA TGTGCGCGC TAAGCTCCAC GGTCTCTAT TGTCTCTAT TGTCTCTAT TGTCTCTAT TGTCTCTAT TGTCTCTAT TGTCTCTAT TGTCTCTAT
 13101 TTCCCGCAA CCGCAGACC TGTAGAGTT GACAGAGTC GAGCAGGCG AGCTGCTCTT TCCCAAGGA AGCTTCTCTA GGCCTAGCAG GGCCTAGCAG GGCCTAGCAG GGCCTAGCAG GGCCTAGCAG
 AAGCGCGTT GCGCTCTGG GCGCTCTGA ACATCTCTA CCGCTCTCT CCGCTCTCT TCGCTCTCT CCGCTCTCT CCGCTCTCT CCGCTCTCT CCGCTCTCT CCGCTCTCT CCGCTCTCT
 13201 CTAGCGCTG CCGCGCGCG GTCAGATGCT AGTAGCCAT TTCCCAAGTA TGTCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 GATCGCGAC GCGCGCGCG CAGCTCTGA TCATCGCTA AGCTTCTCTA CTATCTCTA GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 13301 AGGAGGCTA CTTAAGACAC TCGCTCTCT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 TCGCTCTCT GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 13401 GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 CCGCTCTCT GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 13501 GAGGAGGAT ACTCGGAGA CCGCAGGAC GTCCTGAT TCGGAGTAT AGCTTCTCT CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 CCGCTCTCT GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 13601 AAAAAAAG GCATGATGA AATATATGA CCGCAGGAC CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 TTTTCTCTT CCGCTCTCT GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 13701 TTAGGAGGT CCGCTCTCT CCGCTCTCT GATATGCTT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 ACTCTCTCT GAGGAGGGA GATCTCTCT ACCTCTCT CCGCGCGCG GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT
 13801 GCGCTCTCT GGTACCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 CCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 13901 ACAAGTCAAC GATATGCTT TCGCTCTCT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 TGTCTCTCT CCGCTCTCT AGCGGAGTC CCGCAAGGA CCGCTCTCT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT GATATGCTT
 14001 CACACAGAC ATCAGTCTT AGCGGAGTC GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 GGTCTCTCT TAGTATGAC TCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 14101 AATAGTTTA AGCGGAGTC GATCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 TTATCTAAT TCGCGCGCG CCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 14201 ACTACTCTA GACCATGAC ATAGACTTA TGAACAGTC GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT
 TGTATGACT CCGCTCTCT TATCTCTAT ACTCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT GATCTCTCT
 14301 GGTAAAGTT GACACCTCT ACTCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 CCGCTCTCT GATCTCTCT TGAAGTCTA CCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 14401 AATCTCTCT CAGGATGCT GGTCTCTCT ACTCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT
 TAAACGAGC GCGCTCTCT CCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT GCGCTCTCT

Figure 15I

pMIRAd5qag MER682

14501 CCTACGATGA TCTGGAGGGT GGTACATCTC CCGCACTGTT GGAATGGAG GGTATCCAGG CGAGCTTGA AGATGACACC GAAACAGGCG GGGGTGCTGC
 GGATGCTACT AGACCTCCCA CCATTGTAA GGGCTACAA CCTACACTG CCGATCTTC CCGATCTTC TCTACTGTGG CTGTCCCGC CCCCACCGCT
 AGCGGCAGC MCAGCAGTG GCAAGGCGC GGAAGACAC TCACAGCTG CAGTCTGAC ATGCTAGGCG GTGAGGACCA TGAACGATCA TCGCATTTG
 TCCCGCGTG TTGTGCTCNC CTTGCGCGC CTTCTCTTG AGTTTCCGCG CTGCTGCTG TTACTGCTG CACCTCCTGT ACTTCTTAGT ACGGTAAAGC
 GCGGACACCT TTGCCACACG GCGTAGGAG MAGGCGCG MAGGCGCG AGGCTGAGC AGTCTGGA GCTACCTCC CCGCTGCGA ACCGAGGTC GAGNAGCC
 CCGCTGTGGA AACGGTGTG CCGACTCTC TTGCGCGAC TCGGCTTC TCGGCTTC TCGGCTTC GCGGAGCGT TGGCTCCAG CTCTTCCG
 KpnI
 14801 AGAAGAAACC GGTGATCANA CCGCTGACAG AGACAGCAA GAACGCAAT TACACCTAA TTAGCAATGA CAGCACCCTC ACCAGTACC GCAGTGTGTA
 TCTTCTTTG CCACTAGTTT GGGACTGTC TCGTGTCTT TCTGCTCA ATCTTGATTT ATTCTTACT GTCTGGAAG TGGTCTATG CGTCTGACCA
 KpnI
 14901 CCTTGCATAC AACTAGCGG ACCCTCAGAC CCGATCTCG TCATGCTCC TCTTTGAC TCTTACCTA ACCTGCGCT CCGAGCAGGT CTACTGCTC
 GGAACGTATG TTGATGCGC TGGAGTCTG GCGTACGCG AGTACCTG AGTACCTG ACGTACCTG TGGACGCGA GCGTGTCCA GATGACCAAC
 TTGCCAGACA TGATGCAAGA CCGCTGACC TTGCGCTCA CCGCTCAGT CAGCAGCTT CCGTGTGCG GCGGCGAGCT GTTCCCGCTG CACTCCAGA
 AACGCTCTGT ACTAGTTCT GGGGCACTG AAGGCGAGT GCGGCTCTA GTCTTGANA GCGGCAACC CCGGCTCGA CAACGCGCAC GTGAGGTTCT
 Ascl
 15101 GCTTCTACAA GACACAGGC GCTACTCCC AACTCATCG AACTCATCG CCAGTTTACC TCTCTGACC ACTTGTCTA TCGCTTTCC GAGAACGAGA TTTTGGCTC
 CGAAGATGTT GCTGCTCCG CAGATGAGG TTGAGTAGG GTTCANATG AGAGACTGG TGCACAGTT AGCGAAGGG CTCTTGTCTT AAGACCGCT
 Ascl
 15201 CCCCAGCC CCAACCACTA CCACGTCAG TGAACCTT CTTCTCTCA CAGATCAGG GACGCTACG CTGCGCAACA GCATCGAGG AGTCCAGCA
 GCGCGCTCG GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT GGTGCTAGT
 GTGACCATTA CTGACGCGC AGCGGCGC TCGGCGCTG ACCGCTAGT TTTACAGAG TTTACAGAG TTTACAGAG TTTACAGAG TTTACAGAG
 CACTGCTAAT GACTGCTG CACTGCTG CACTGCTG CACTGCTG CACTGCTG CACTGCTG CACTGCTG CACTGCTG CACTGCTG
 CCATGCTCAT CCTATATCG CCACGCAATA ACACAGCTG TTTCCAGCA AGATGTTTGG CCGGCTTGG CCGGCTTGG CCGGCTTGG CCGGCTTGG
 CGTACAGGTA GGAATATAG GGTGCTTAT TGTGCTGAC CCGGCTTGG CCGGCTTGG CCGGCTTGG CCGGCTTGG CCGGCTTGG CCGGCTTGG
 AGTGGCGTG CCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA
 TCACGCGC CCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA
 GAGGCGGCA ACTACAGCC CACGCTTGA CAGTGTGCA CAGTGTGCA CAGTGTGCA CAGTGTGCA CAGTGTGCA CAGTGTGCA CAGTGTGCA
 CTGCGCGCT TGAATGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG GTGCTGCTG
 GAGGCTGAG GCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG CCGCTGAG
 CTGCGCGCT CCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA TCGGCTTGA
 SfiI
 15801 ACGCGCGCC ATGCGCGCG CCGGCTTGA CCGGCTTGA CCGGCTTGA CCGGCTTGA CCGGCTTGA CCGGCTTGA CCGGCTTGA CCGGCTTGA
 TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG TCGCGCGCG
 AGTGTATGA CTCAGGCTG CAGGCTTGA CAGGCTTGA CAGGCTTGA CAGGCTTGA CAGGCTTGA CAGGCTTGA CAGGCTTGA CAGGCTTGA
 TCAGGATAT GAGTCCGCG GTGCGCGCT CAGTATAT CAGTATAT CAGTATAT CAGTATAT CAGTATAT CAGTATAT CAGTATAT
 TTGCAAGAA AACTACTTA GACTGCTAT GACTGCTAT GACTGCTAT GACTGCTAT GACTGCTAT GACTGCTAT GACTGCTAT GACTGCTAT
 AACGTTCTT TTTGATGAT CTGAGCATG CAGCATAT CAGCATAT CAGCATAT CAGCATAT CAGCATAT CAGCATAT CAGCATAT CAGCATAT

Figure 15J

pMRKAd5qag MER682

16101 CCAGGTGATC GGGCCGGAGA TCATATGGCC CCCGAGGAG GAGGAGCAGG ATTACAGGC CTGAAGCTA AAGCGGCTCA AAAGAGAAAA GAAAGATGAT
 GGTCCAGTAG CCGGCTCT AGATACCGGG GGGCTCTTC CTATATGACC TAATGTTGCG GCTTTGAT TTGCCCCAGT TTTTCTTTT CTTCTACT
 BglII
 16201 GATGATGAC TTGACGACCA GGTGAACTG CTTCACGCTA CTGCTCCAG GCGATGCTA CACTGGAAAG GTCGACGCTT AAACGTTGT TTGCGACCC
 CTACTACTTG AACTGCTGCT CCACCTTGAC GACGTGCTGAT GCGCTGGCTC GCGTCCCTAT GTACACCTTTC CAGCTGCGCA TTTTGCACAA AACCTGCTG
 16301 GCACCAACGT AGTCTTTTACG CCGCTGAGC GCTCCACCCG CACTTACAAAG CCGGTGTATG ATGAGGTGTA CCGCGACTAG GACCTGCTTG ACCAGGCCAA
 CGTGGTGGCA TCAGAAATGC TGGCTACTCG CAGGTGCTG GTGATGTTTC GGTACATATC TACTCCACAT GCGCTGCTC CTGGACGAAC TCGTCCGCTT
 16401 CGACCGCTC GGGAGTTTG CCTACGGAAA GCGCATATG GACATGCTTG GGTTCGCTT GACGAGGCG AACCCACAC CTAGCCTAAA GCGCGTAACA
 GTCGCGGAG CCGCTCAAAC GATGCTCTT CCGCTATTC CTGTAGACC GCAATGCGTA CCTGCTCCCG TTGGGTTGTG GATCGGATTT CCGGCAATTG
 PstI
 16501 CTGACGACG TCGTGGCCG GCTTGACCG TCGGAGAA AGCGCGCTT AAATGCGAG TCTGCTACT TGGACCCAG CGTGCAGCTG ATGCTTACCA
 GACGTGCTC ACACCGGCG CGACGTGCG AGGCTTCTT TCGCGCCGA TTTTGGCTC AGACCTCTA ACCGTGGTG GCACGTGAG TACCATTGCT
 16601 AGCGCAGCG ACTGGAGAT GTCTTGAAA AATGACCTG GAACTTGG CTGAGCCCG AGGTCCGCT GGGCCAAATC AAGCAGGTG GCGCGGACT
 TCGGGTGC TGACCTTCTA CAGAACCTT TTTACTGCA CTTTGACTC GACCTCGGC TCCAGCGCA CCGGCTTAG TTGCTCCACC GCGGCCCTGA
 16701 GGGGTGACG ACCGTGACG TTCAGATACC CACTACCTT AGCACATTA TTGCCACCG CACAGAGGGC ATGGAACAC AAACGTCCC GGTGCTCTA
 CCGCACGTC TGGCACCTG AGTCTATCG GTGATGCTA TCGTGTCTC TCGTCTCCG GTGCTCTCC CACGTTTCC TACCTCTGTT TTGACAGGG CCAACGAGT
 16801 GCGTGGCGG ATGCGCGCT GCAGCGGTC GTGCGCGG CGTCCAGAC CTTACTAGG GTGCAACCG ACCGTGGAT GTTTCGCTT TCAGCCCTC
 CGCCACGCG TACGGCGCA CGTCCGCGC CGACGCCCG GCAGTCTCT GAGATGCTC CACGTTTCC TGGCACCTA CAAGCGCA AGTCTGGGG
 16901 GCGGCCGCG CGGTTGAGG AAGTACGGG CCGCAGCGC GCTACTGCC GATATGCTC TACATCTTC CATTGCGCTT ACCCCCGCT ATCGTGGCT
 CCGCGGCGC GGCAGCTCC TTCATGCCG GCGGTGCG CCGTACCGG CTTATACGG ATGTAGGAG GTAACGGGA TGGGGGCCBA TAGCACCGAT
 17001 CACCTACCG CCCAGAGAC GAGCACTAC CCGACCGCA ACCACTACT GAACTGCGG CCGTCTCGC CCGTCTCGC CCGGATTTCC
 GTGATGCGG GGTCTCTCT CTGTTGATG GGTGCGCT TGGTGGTAC CTTGGCGGC GCGCGCTG GGCACGACCG GGCCTAAGG GGGCTAAGG
 17101 GTGCGCAGG TGGCTCGGA AAGAGCGAG ACCCTGCTG TCGCAACAC CCGCATACC CCGAGCATG CCGTCTTGTG TTTAAGGCT GTTCTTGGAG
 CACCGCTCC ACCGAGCGT TCGTCCGTC TGGACCCAG CCGTCTGTC CCGATGGTG GGTCTGTAG AATTTTGG CCAGAAACAC CAAGAACGT
 SphI
 17201 ATATGGCTT CACCTGCGC CTCGTTTCC CCGTCCCGG ATTCCAGCA AGATGCACC GTAGGAGGG CATGGCCCG CACGCGCTGA CCGCGCTCAT
 TATACCGGA GTGACGGG GAGGCAAGG GCCACGGCC TAACTCTCT TCTTACGTG CATCTCCCG GTACCGGCG GTGCGGACT GCGCGCGCTA
 SphI
 17301 GCGTCTGCG CACCAACCG GGGCGCGC GTGCAACCT CTAATGCGG GCGGTATCTT GCGCTCTCTT ATTCCACTGA TCGCGCGCG GATTGGCGC
 CCGACACCG GTGGTGGCG CCGCGCGCG CAGCTGGCA GCGTACGCG CCGCATNGA CCGGAGGAA TAAGTACTT AGCGCGCGG CTAACCGCG
 17401 GTGCGCGAA TTGATCTCT GCGCTGCG GCGCAGAGC ACTGATTAA ACGAGTTTC ATGTGAAA AATCAAAATMA AAGTCTGGA CTCCTACCT
 CACCGGCTT AACGTAGCA CCGGACGTC CCGTCTCTG TGACTAATT TTGTTCAAG TACACCTTT TAGTTTTATT TTTCAGACCT TTTGAGCTGA GAGGTGCGA
 EcoRV
 17501 CCGTCTGCG TGTAACTATT TTGTAGATG GAGCATCA ACTTTGCTC TCTGCTCG TCTGCTGAG GCTGTGCGA GCGCGCGCA GTACCTTTG ACCGTCTAT
 GCGAACCGG ACATTGATA AACTCTTAC CTTCTGTGT TGAACCGAG AGACCTGAC GCTGTGCGA GCGCGCGCA GTACCTTTG ACCGTCTAT

Figure 15K

pMRKAI15gag MER682

17601	TCGBCACCAG	CNATATGAGC	GTTGGGGGCT	TCAGCTGGGG	CTGGCTGGGG	AGCTGCTGGG	AAATTTTGG	TTCCACCGTT	AGGAACTATG	GCAGCAAGG
17701	AGCCGTGGTC	GTTATATCTG	CCACCGGGGA	AGTGGACCCC	GAGGACACAC	TCGGCTGAAT	TTTTAAAGCC	AGGTGGCAAA	TTCTTGTATAC	CGTGGTTCCG
	CTGGAACAGC	AGCACAGGCC	AGATGCTTGG	GGATNAGTTG	ANAGAGCAAA	ATTTTCACAA	AAAGCTGCTA	GATGGCTTGG	CCCTCTGGCAT	TAGCGGGGTM
	GACCTTGTCT	TCGTGTCCGG	TCTAGGACTC	CCATATTCAC	TTTCTCTGTT	TNAGGTTGT	TTTCCACCAT	CTACCGGACC	GGAGACCGTA	ATCGCCCCCN
17801	GTGGACCTTG	CCAAACAGGC	AGTGCNAAAT	AAATTTTGA	GTAAATTTGA	TCCCCGGCT	CCGTAGAGG	AGCTTCCACC	GGCGGTGGAG	ACAGTGTCT
	CACCTGGACC	GGTTGGTCCG	TCACGTTTGA	TTCTNATTTG	CATCTTAAT	AGGCGGGA	GTTCATCTCC	TCGGAGGTGG	CCGGCACCTC	TGTCACAGT
17901	CAGAGGGGGG	TGGCGAAAG	CGTCTGGTCC	CCGACNAGGA	AGNAACTCTG	GTGATGCANA	TAGACGAGCC	TCCCTCGTAC	GAGGAGGCAC	TAAAGCAAT
	GTCTCCCGGC	ACCGCTTTTC	GCAGGCGCGG	GGCTGTCCCT	TCCTTTAGAC	CATCTGGTTT	ATCTGCTGG	AGGAGCATG	CTCCTCGTGG	ATTTCTGTTT
18001	CTTCCCCACC	ACCGCTCCCA	TGGGCGCCAT	GGCTACCGGA	GTGTTGGCC	AGCACACACC	CGTAACGCTG	GACCTGCCCTC	CCCCCGCCGA	CACCCAGCAN
	GGACGGGTGG	TGGGACGGGT	AGCGCGGGTA	CCGATGGCT	CACGACCCGG	TGCTGTGTGG	GCAATGGGAC	CTGGACGGAG	GGGCGCGGGT	GTGGCTCGTT
18101	AAACCTGTGC	TGCCAGGGCC	GACCGCCGTT	GTGTAAACC	GTCTAGGCG	CGCTTCCCTG	CGCGCGCGCG	CCAGGCTGCC	GGGATCGTTG	CGGCCCCGTG
	TTTGGACAGG	ACGTTCCGGG	CTGGCGGCA	CAACATTTGG	CAGGATCGGC	GGCGAGGAC	GGCGCGCGCG	GGTGGCAGG	CGCTAGCAAC	GGCGGCGATC
18201	CCAGTGGGAA	CTGGCANAGC	ACACTGNACA	GCATCGTGGG	TCGGGGGG	CAATCTCTTA	AGCGCGGACG	ATGCTTCTGA	TAGCTAACGT	GTCTGTATGTM
	GGTCACCGTT	GACCGTTTCT	TGTGACTTGT	CGTAGCACCC	AGACCCCGAC	GTTAGGGAAT	TCGGGGCTGC	TACGAGACT	ATCGATTGCA	CACCATACAT
18301	TGTCATGTAT	GGTTCATGT	CGCGCGCAGA	GGAGCTGCTG	AGCGCGCTGG	CGCGCGCTTT	CCAGATGTC	TACCGCTTGG	ATGATGCGCG	AGTGGCTTGA
	ACAGTACATA	CGCAGGTACA	GGCGGGTCT	CTCGAGGAC	TGCGCGCGCG	GGCGGCGGAA	GGTCTTACCG	ATGGGAGGC	TACTACGGCG	TGACCAAGAT
18401	CATGCACATC	TGGGGCCAGG	ACGCTTCGGA	GTACCTGAGC	CGCGGGCTGG	TGCAATTTGG	CGCGGCGCAC	GAGACGTACT	TCAGGCTGMA	ATTGTTTCANA
	GTACGTGTAG	AGCGCGGTCC	TGGGAGGCT	CATGACTCG	GGCGCGGACC	ACGTCNAAAG	GGCGCGGTGG	CTCTGCAATGA	AGTGGGACTT	ACTGCGGTACT
18501	AGAAACCCCA	CGGTGGCGCC	TACCGACGAC	GTGACCAACG	ACCGTTCCTA	GGTGTGACG	CTGGGTTTCA	TCCCTGTGGA	CGGTGAGGAT	TGACCGCATGA
	TCTTTGGGGT	GGCACCGCGG	ATGGGTGCTG	CATGTGTCTC	TGGCCAGGCT	CGCANACTGC	GACCGCCNAGT	AGGAGACCTT	GGTCTGGACA	GGGCGCCCTA
18601	CGTACAAAGC	GGGTTTCACC	CTAGCTGTGG	GTGATNACCG	TGTTCTGGAC	ATGGCTTCCA	CGTACTTTGA	CATCCCGCGC	GTGCTGGACA	GGGCGCGGAT
	GCATGTTCCG	CGCCAAATGG	GATCGACACC	CATNATTTGC	ACACGACCTG	TACCGAAGGT	GCAATGAAGT	GTAGCGCGCG	CAGGACCTGT	CCCCCGGAT
18701	TTTTAAAGCC	TACTCTGGCA	CTGGCTTACA	CGCCCTGGCT	CCCAAGGCTG	CCCCAAATCC	TTGGGAATGG	GATGAAGCTG	CTACTGCTCT	TGAAATNAN
	AAATTTCCGG	ATGAGACCGT	GACGGATGTT	GGGGACCGA	GGTTCCTCAC	GGGTTTATGG	AGCGCTTACC	CTACTTCGAC	GATGACGAGA	ACTTTATTTT
18801	CTAGAAGAG	AGGACGATGA	CAACGAAGAC	GAAGTACAGC	AGCAAGCTGA	GGAGCAAAA	ACTCACGATAT	TTGGGACGGC	GCTTTATTTCT	GGTATNANTA
	GATCTTCTTC	TCTGTCTACT	GTTCCTTCTG	CITCATCTGC	TGTTTGGACT	CGTCTTTTTT	TGAGTGCATA	AACCCGTCGG	CGGAATAAGA	CCATATTTAT
18901	TTACAAAGGA	GGTATTTCAA	ATAGGTGTGG	AAAGTCAAC	ACCTNATAT	GGCGATAAAA	CATTTCAACC	TGAACCTCAA	ATAGGAGNAT	CTCAGTGTGA
	AATGTTTCCCT	CCCTAAGTT	TATCCACAGC	TTCCAGTTTG	TGTAATTTATA	CGCTATTTT	GTNAGTTTGG	ACTTGGAGTT	TATCTCTCTTA	GAGTCCCAT
19001	CGAAACAGNA	ATTAATCATG	CAGCTGGGAG	AGTCCCTAAA	NAGACTACCC	CAATGMAACC	ATTTTACGCT	TCATATGCCAA	AACCCACAAA	TGAAATGTT
	GGTTTGTCTT	TAAATTAGTAC	GTCGACCTTC	TCAGGATTTT	TTCTGATGGG	GTTACTTTGG	TACMATGCCA	AGTATACGTT	TTGGGTGTTT	ACTTTTACCT
19101	GGGCAAGCCA	TTCTTTGTAA	GCACCAAAAT	GGAAAGCTAG	AAAGTCAAGT	GGAAATGCCA	TTTTTCTCAA	CTACTGAGGC	AGCCCGCAGGC	AAATGTTGAT
	CCCGTTCCGT	AAGAACATTT	CGTTGTTTGA	CGTTTCTATC	TTTCTGATTC	CTTTTACGTT	AAATGAGGTT	GATGACTCCG	TCGGGCTCCG	TTACCACTAT
19201	ACTTGAATCC	TAAAGTGGTA	TTGTACNGTG	AAGATGTAGA	TATAGAAACC	CCACACACTC	ATATTTCTTA	CATGCCCACT	ATTNAGGAG	GTAACTCACAT
	TGAACTGAGG	ATTCACCAT	NACATGTAC	TTCTACATCT	ATATCTTTGG	GTTCTGTGG	TATNAGGANT	GTACGGGTGA	TAAATCTCTTC	CATTTGAGTG

Figure 15L

pMRKAD5gag MER682

19301	AGAACTAATG GGCACAAAT CTATGECGAA CAGGCTAAT TACATTGCTT TTACGAA'AA TTTTATATGTT CTAAATGTAAT ACAACAGCNC GGGTAATATG
19401	TCTTGATTAC CCGTTGTTA GATACGGTT GTCCGTAATA ATGTAACGAA ATGCTCTGTT AANATAACCA GATTACATAA TGTTGTCGTG CCCATTATATC
19501	GGTGTTCTGG CCGGCCAAGC ATGCTAGTTG AATGCTGTTG TAGATTITGA ACACAGAAAC ACAGAGCTTT CATACCAGCT TTGCTTTCAT TCCATATGTTG
19601	CCACAGAGCC GCGCGTTGG TAGCGTCAAC TTACAGCAAC ATCTAAAGCT TCTGTCTTTG TGCTCTGANA GTATGOTCOA AAGCAACTA AGCTAACCA
19701	ATAGAACCA GACTTTTCT ATGCTGAAATC AGCTTTTGA CAGCTATGAT CCAGATCTTA GAATATATGA AATCATGGA ACTGAGATG AACTTTCANA
19801	TATCTTGGTC CATGAANA GATGATTA TAGACCTTAG TCCGACAACT GTCCATACTA GTCTTACAAT GTCTTACAAT CTTAATTAAC TTTAGTACCT TGACTTCTAC TTGAATATTT
19901	TTACTGCTTT CCAGTGGAG GTGATTA GACAGACT CTACCAAGG TAAACCTAA AACAGCTCAG AACAGCTCAG GAAATGGAAT GGGAANAAGA TGCTACAGAA
20001	AATGACGAA GTGACCCCTC CACACTAAT ATGCTCTGA GAATGTTCC ATTTGGATT TTGTCCAGTC TTTTACCTA CCTTTTCT AGATGTCTT
20101	TTTTACGATA AATGAAAT AAGATTGGA AATATTTTG CCATGMAAT CAATCTAAT GCCMACTGT GGAGAAATTT CCTGTACTCC AACATAGCC
20201	AAAAGTCTAT TTTTACTTTA TTCTCAACCT TTATTAAAC GTTACCTTTA GTTAGATTTA CCGTTGGACA CCTCTTTANA GGACATGAGG TTGTATCTG
20301	TGTATTTGCC CGACAAAGCTA AAGTACAGTC CTTCACAGT AANAATTTCT GATAACCCAA ACACCTACGA ACACCATGAG AAGCGAGTGG TGGCTCCCTG
20401	ACATAAACGG GCTGTTGGAT TTCAATGTCAG GAAGTTGCA TTTTAAAGA CTATTGGTT GTTGATGCT TGTGATGCT GATGTACTG TCCCTCACCC ACCGAGGGT
20501	GCTAGTGGAC TGCTACATTA ACCTTGGAGC AGCTGCTCC CTTGACTATA TGACACAGCT CAACCCATTT ACCCACTTT ACCCACTTT GCAATGCTTG CCTGCGCTAC
20601	CGATCACCTG ACATGTAAAT TGGACCTCG TCGACCCAGG GAACTGATAT ACCTGTTGCA GTTGGSTANA TTGGTGGTG CATTACGACC CGTTACGACC GGACGGGATG
20701	CGCTCAATGT TGCTGGGCAA TGGTGGCTAT GTGCCCTTCC ACATCCAGCT GCTCAGNAG TTCTTTGCCA TTAAAAACCT CCTTCTCTCG CCGGGCTCAT
20801	GCGAGTTACA ACGACCCGTT ACCAGCGATA CACGGGAAG TGTAAGTCCA CCGAGTCTTC AAGMAACGGT AATTTTGGGA GGAAAGAGGAC GGCCCGAGTA
20901	ACACCTACGA GTGGAACTTC AGGAAGGATG TTAACTGGT TCTGTAGAGC TCCCTAGGNA ATGACCTAAG GTTGAACGA GGTGAACGA AGTTGATAT
21001	TGTGATGCT CACCTTGAA GCTTCTCTAC AATGTACCA AATGCTCTCG AGGATCTCTT TACTGGATTG CCAACTGCTT CCNACTGCTT CCGTCTGTAAT TCANACTATC
21101	CATTGCTCTT TACGCCACCT TCTTCTCCAT TCTTCTCCAT GGGCCACAA CCGCTTACGA CATGCTTAGA CATGCTTAGA AAGCACCA CTGCTGTTGTT GAAATTTGCT
21201	GTAAACGGA ATGCGGTGGA AGAAGGGTA CCGGGTGTG TGXCGAGGT GCGMACTCCG GTACGAAATCT TGGCTGTTGTT TGGCTGTTGTT GCGGCTGTTG CCGGCTGTTG
21301	TATCTCTCC CCGCCACAT GCTCTACCT ATACCGCCA CCGTCCCAT CTGCCCCATA TCCATCTCTT TCCATCTCTT AGGTAGGGA AGGTAGGGA
21401	ATAGAGAGGC GCGGTTGTA CGAGATGGA TATGGGGT TCGATGCTT GACCGCTTAT ACACCTACTC TGTGGATGAG ACCGAGATAT GCGATGATC TACCTTGTGA
21501	CCTTCACCG GGAATTTCTGA TTCTTTGGG ATGATGAGC CATTAATTT GACCTTCTCT TCCAGCTGCTT CCGCTTCTTA CCCCCAACGA GTTTTGAATTT
21601	ATGAGTTTG CACACCTTTA AGAAGTTGG CATTACCTTT CATTATGGA GTATGAGAGC CTGAGMAGAC AGTCCAGCCG ACCGTTACTG GCGGCTGCTT CAUACTTTAA
21701	AATGGATTG GTGTTGAAAT TCTTCCACCG GTATGGA GATTACCTTT GTTGCCAGT GTTACATGAC CAAGACTTG TTCTTGTAC AATGCTAGC TAACTATAGC ATTGCTTACC
21801	AAGGCTCAG TTGACGGGA GGGTTACAA GGTGCTCAGT CAAGCTCTA CATTTGACCT AAGGACCATG GTGGATGAG GTGGATGATA CTAAATACAA
21901	TTGCGGATC AACTGCCCC CTCAATCTTG CAAGGCTCA CATTTGACCT AAGGACCATG GTGGATGAG GTGGATGATA GTGGATGATA GTGGATGATA
22001	AAGGCTTCTA TATCCAGAG AGCTACAAG ACCGATGTA CTCTCTCTT AGAANCTTCC AGCCCATGAG TCGGGTACTC GGCAGTCCAC CACTACTAT GATTATATGT
22101	TCCGAGAT ATAGGCTCTC TCGATGTTCC TCGGTTACAT GAGGAGAA TCTTTGAGG TCGGGTACTC TCGGGTACTC TCGGGTACTC TCGGGTACTC TCGGGTACTC
22201	GGACTACCAA CAGGTGGCA TCCATACCA ACACAAAC TCTGATTTG TTGGCTACTT TGGCTACTT TGGCTACTT TGGCTACTT TGGCTACTT TGGCTACTT
22301	CCTGATGTT GTCCACCGT AGGATGTGTT TGTGTTGTT AGACCTAAC AGCGATGGA ACAGGCTG ACAGGCTG ACAGGCTG ACAGGCTG ACAGGCTG
22401	TTCCCTATC CCGTTATAG CAGACCTGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA GTTGACGA
22501	AAGGGATAG CCGATATCC GTTCTGGCT CMACTGCTT AATGGTCTT TTTCANAGA ACGCTAGCTT GGGAAACCG GGTAGGCTAG AGCTCATTTA

Figure 15M

pMRKAd59ag MER6R2

21001 TTATGTCCAT GGGGACATC ACAGACCTGG GCCAATACCT TCCTACATC AACTCCGCC ACAGGCTAGA CATGACTTTT GAGTGGATC CCATGGACGA
 ANTACAGGTA CCGCGGTGAG TGCTCGACC CGTTTCTGGA AGAGATGCTG TGGAGTCGG TGGCGATCT GTACGMAAA CTCCACCTAG GTTACCTGCT
 GCTCACCTTT CTTTATGTTT TGTTCAGAT CTTTACGCTG GTCCGTGTC ACCATCCGA CCGGCTGTC ATCGAATCCG TGTACCTGG CAGCCCTT
 CCGGTGGAA GAATACAAA ACAATCTTA GAACTGCAC GAACTGCAC TGTCTGGCT TGTCTGGCT TAGCTTTGGC AC/ TGGACGC GTGGGGAN ;
 21201 TCGCCCGCA AGCCACAC ACATAGAGC AAGCAATC AAGCAATC GCGCCCATGG GCTCCATGA GCAGGAATG AAGCCATG TCATAGATCT
 AGCCGCCGT TCGGTGTTG TATTTCTTCG TCTGTTGTAG TTGTTGTGA CCGCGGTACC CAGGCTACT GTCTCTGAC TTTCGGTAAC AGTTCTTAGA
 TGGTTGTGG CCATATTTT TGGCACCTA TACAGCGC TTTCAGGCT TTGTTCTCC ACACAGCTC GCTCGGCCA TAGTCATATC GCGCGCTG
 ACCAACACCC GGTATATAA ACCCGTGGT ACTGTTCCG AAGTCTCCA AAGTCTCCA TGTCTTGGAG CCGACCGGT ATCAGTTATG CCGGCCAGTG
 GAGACTGGG GGTACACTG GATGCTTT GCTTGAAAC GCTAGCTA AACTGCTAC CTCTTGGCTT TCTTGACCAG TCTGACCCG CCACTCAAGC
 CTCGACCCC CCGATGTGAC CTACCGGAA CCGACCTTG GGTGATTT TGTACGATG GAGAACTG GAAACCGAA AAGTCTGAC AAGTCTGAC
 AAGTTTACCA GTTGTAGTAC GATGCTTC GAGTCTTC TCGCGCTAG CCGCTTCT TCTTCCCGG ACCCTGTAT AAGTCTGAA AAGTCTGAC
 TCCAAATGCT CAACTCATG CTCAGTGAG CCGCGCATC ACGCGCATC GCGGTAAAG AAGTCTGAA TGGGACATA TTGCGACCTT TTGCGGTTGG
 GCGGCCAAC TCGCGCGCT GTGACTAT CTCTGCAAG TTCTCCAG CTTTCCCA CTGCGCCAA CTGCGCATG ATCACAACC CACTCATGAA
 CCGCGGTG AGCCCGGGA CACTGATA CAGTGTAC AAGAGTGC GGAACGCTT GACCGGCTT TGAGGCTACC TAGTGTGGG GTGCTACTT
 21701 CTTATTACCG GGTACCCAA CTCCATGCTC AACATGCCC AGTACAGCC CACCTGCTG CCGAACCGG AACAGCTTA CAGCTTCTG GAGCGCAC
 GAATATGCG CCGATGGTT GAGTACGAG TTCTACGGG TCCATGCTG GGTGACGCA CCGTGTGTC TTGCGAGAT TTGCGAGAC CTCGCGGTG
 CCGCTACTT CCGCAGCCAC AGTCCGACA TTAGGAGCG CACTTCTTT TGTACCTGA AATCATGTA AATCATGTA TGTATCTG CCGCTGCTA
 CCGGATGAA GCGGTGGTG TCAGCGCTT AATCTCGG GTGAGGAAA ACAGTGAAT TTTGTACAT TTTTATTACA TGTATCTG GAAAGTTAT
 AGGCAATGC TTTTATTTG ACATCTCGG GTGATTTAT GTGCGTCTG CCGGTTTAA AATCAAGG GGTCTGCTG CCGCTGCTA
 TCGTTTACG AATAAACA TGTAGAGCC CACTAATAA TGTGTGGG AAGTCTGAG CCGGCAATT TTGATTTCC CCGAGACCG CCGTACGGAT
 TCGCCCACTG CAGGAGAC GTTGCTATC TGTGTTTAG TGTCTCAT TGTCTCAT AACTCAGG AACTCAGG CCGGAGCTC GGTGAAGTT TCACTCCAC
 ACGGCTGAC COTCCCTGT CAACCTATG ACCACMATC ACTAGTGAA TTTGAGTGG TGTGTTAG CCGCTGCTG CCGCTGCTG GATACACAG
 22101 GGCTGCGAC CATACCCAC GCGTTTACCA GGTGCGGCG CCAATCTTTG AGTCTGCTG AGTCTGCTG CCGAGTTG CCGAGTTG GATACACAG
 CCGACGCGTG GTAGTGTG CCGAATCTG CCGACCGCG CCAATAGAAC TTACAGCTCA ACCCGGAGG CCGGACCGG CCGGACCGG CTATGTTG
 GTTGCAGCAG TGGACACTA TCAGCGCGG GTGGTACAG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG
 CAAGTGTG ACCTTGTAT AGTCCCGCC CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG CAGTCTTCTG
 GCGAACCGAG TCACTTTG TAGCTGCTT CCGAATAGG CCGAATAGG CCGAATAGG CCGAATAGG CCGAATAGG CCGAATAGG
 CCGTGTGCTC AGTTGAACC ATCGACGGA GGTCTTCTG CCGACCGG CCGACCGG CCGACCGG CCGACCGG CCGACCGG
 CCGTGTGCTC GTTAGGATC AGGCTTCA TAAAGCTT GATCTGCTA AAGCTTCA TTTGCTGTA TTTGCTGTA TTTGCTGTA TTTGCTGTA
 GCGAGACCG CAATCTATG TCGGAGCTT ATTTGGA ATTTGGA ATTTGGA ATTTGGA ATTTGGA ATTTGGA ATTTGGA ATTTGGA
 22501 GCGGGAAC TGATGCGCG GACAGCGCG GTGCTTACG CAGCAGCTG CAGCAGCTG CAGCAGCTG CAGCAGCTG CAGCAGCTG
 CCGCTTTTG ACTAACCGG CTGTCGCGG CAGCAGCTG CAGCAGCTG CAGCAGCTG CAGCAGCTG CAGCAGCTG CAGCAGCTG

Figure 15N

pMRKAd5gag MER682

22601	ATCTTGGGCT TGCTAGACTG CTCTTTACGC GGGGCTTCC CTTTTCCTT CTTTTCCTT ATTTCATCA CGTCTCTCTT ATTTATCATTA ATGCTTCCCT	TAGAACCGGA ACGATCTGAC GAGTAACTGG GCGGAGCGG GCGGAGCGG TAAAGTTAGT GCACGAGGGA TAAATAGTAT TACGAGAGCA	PsiI ~~~~~
22701	GTAGACACTT AGCTGCGCT TCGATTTAG CCGAGCGGTG CATTACACAC GCGGAGCGCG TGGGCTCTGT ATGCTTGTAG GTTACCCTCTG CAAATGACCTG	CATCTGTGAA TTCCAGCGGA AGCTAGACTC GCGTGGCCAC GCTTCTTCTG CCGTCTGGGC ACCCGAGCAC TACGACATC CAGTGGAGAC GTTTCCTCTA	PsiI ~~~~~
22801	CAGGTACGCC TGCGAGGATC GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT	GTCCATGGGG ACCTCTTTAG CCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT	PsiI ~~~~~
22901	CATACGGGCG CCGAGAGCTT CACTTCTTCA GCGAGTAGTT TCAAGTTTCC CTTTAGATCG TTATCCACGT GGTACTTGTG CATCAGCGCG CCGGAGCGCT	GTATGCCCGC GGTCTCGAAG GTGAGCCAGT CCGTCACTCA ACTTCAGCG GAAATCTAGC AATAGGTGCA CCATGAACAG GTAGTCCGCG GGGGCTCG	PsiI ~~~~~
23001	CCATGCCCTT CTCCACGCA GACAGGATCG GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT	GGTACGGGAA GAGGTGCGT CTGTGCTAGC GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT GCGGATCATT	PsiI ~~~~~
23101	CCGATACCA CCGGCGACTG GGTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT	GGCGTATGCT GCGGCGACTG CCGGCGACTG TCAAGTCTG GCGGCGACTG TCAAGTCTG GCGGCGACTG TCAAGTCTG GCGGCGACTG TCAAGTCTG	PsiI ~~~~~
23201	ACCATTTGTA GCGGCGACTG TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT	GGGATGAT GCGGCGACTG TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT TCTCTCTTCT	PsiI ~~~~~
23301	TCTTGGGCGC ATGCGGACTG TCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	AGAACCCGCG TTACCGGTTT AGGCGGCGG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
23401	CTCGATACGC CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	GAGCTATGCG GCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
23501	GCACCGGCGC CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	CGTGGCGGCG GCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
23601	AGAGGAGCAG CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	TCTTCTCTGTC GCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
23701	GGAGGAGGGA GTGATTTATG AGCAGGACCC AGGTTTGTGA AGGTTTGTGA AGGTTTGTGA AGGTTTGTGA AGGTTTGTGA AGGTTTGTGA AGGTTTGTGA	CCCTCTCTCTT CACTAATAGC TCGTCTCTGG TCCANNAACAT TCGTCTCTGG TCCANNAACAT TCGTCTCTGG TCCANNAACAT TCGTCTCTGG TCCANNAACAT	PsiI ~~~~~
23801	GCAGAGGCA ACCAGGAGCA AGTGGGCGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG	CGTCTCTCTT TCGTCTCTGG TCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG GCGGAGGAGG	PsiI ~~~~~
23901	GCGGCGGCTT CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	CGGCGGCTTAT CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
24001	ACCGGCGGCTT CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	TGGGCGGCTT CCGGCGACTG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	PsiI ~~~~~
24101	TTTTTCCAAA ACTGCAAGAT ACCCTATCC TCGGAGGCGA ACCGAGGCG TCGGAGGCGA ACCGAGGCG TCGGAGGCGA ACCGAGGCG TCGGAGGCGA	AAAAAGGTTT TGACGTTCTA TGGGATAGG ACCGAGGCGT TCGGAGGCGA ACCGAGGCG TCGGAGGCGA ACCGAGGCG TCGGAGGCGA ACCGAGGCG	PsiI ~~~~~

Figure 150

pMRKAd5gag MER6R2

24201 CCTCGCTCAA CGAAGTCCCA AAAATCTTTG AGGTTCTTGG AGCGGAGGAG AAGCGGCGGG CAAGCTCTCT GCAACAGGAA AACAGCGGAA ATGAAAGTCA
 GGAGCGAGTT GCTTCACGGT TTTTAGAACC TCCAGAAC TGGCTGCTC TTGCTGCGCC GTTGTGCTT TTGTGCTT TGTGCTTCTA

24301 CTCCTGAGTG TTGGTGAAC TCGAGGTGA CAGCGCTGG CTAGCGGTAC TAAATGCGAG CATCGAGTTC ACCACTTTG CTTACCGGC ACTTAACCTA
 GAGACCTCAC AACCACTTG AGCTCCACT GTTGGCGCG GATCGGCTG ATTTTTCCTC GTAGCTCCAG TGGGTGAAC GATGGGCGG TGAATTTGGT

24401 CCCCCCAAGG TCATGAGCAC AGTCATAGT GAGCTGATCG TTGGGCGTGC GAGCCCTCG GAGAGGATG CAAATTTGCA AGAACAAACA GAGGAGGGT
 GGGGGTTCC AGTACTCTTG TCAGTACTCA CTCGACTAGC ACCGTGACAG CTCCTCCCTAC GTTTAAACGT TCTTGTCTGT CTTCTCCCTA CTTCTCCCTA

24501 TACCCGCGT TGGCGAGGAG CAGCTAGGCG GCTGCTTCA AAGGTGCTG CTTGCGACT TGGAGGAGCG ACCCAACTA ATGATGGGCG CAGTCTCT
 ATGGCGCTCA ACCGCTGCTC GTGATCGCG CGACCGAGT TTGCTGCTC GACGCTGA ACCTCTCTG TCGTTTGT TACTACCGGC GTACCGAGCA

24601 TACCGTGGAG CTTGAGTGA TGCAGCGTT CTTTGCCTAC CCGGAGTGC AGCGGAGCT AGAGTAACA TTGCCTACA CTTTCTGACA GGGCTACCTA
 ATGCGACCTC GAACTCAGT GAACTCAGT ACCTGCGCAA GAACGACTG GGGCTCTAGG TCGGCTTGA TCTCTTGT AACGTGATGT GGAAGCTGT CCGGATGCT

24701 CGCCAGGCT GCAAGATCTC CAAGGTGGG CTCGTGACC CTCGTCTCTA CTTTGAATT TTGCAGGAA ACCGCTTGG GCAAAACGTG CTTCAATCCA
 GCGGTCCGA CTTTCTAGG GTTCCACCTC GAGACGTTG ACCAGAGAT GGAACCTTAA MACGTCTT TGGCGGAACC CGTTTTCAC GAAGTAAGT

24801 CGCTCAAGG CBAGGCGCG CCGGACTAG TCGCGACTG CGTTTACTA TTCTATGT ACACCTGGA GACGGCATG GGGTTTGGC AGCAGTCTT
 GCGAGTTCC GCTCCGCGG GCGTGTATG AGCGCTGAC GCAATGAT AAAGTACGA TGTGACCGT TGTCCGTAC CCGCAACCG TCGTACGAA

24901 GAGGAGTGC AACCTCAAG AGCTGCAGAA ACTGCTAAG CMAACTTGA AGACCTATG GACGCTTC GACGCTTC AACGAGCT CCGTGGCGC GCACCTGCT
 CTTCTCACG TTGAGGTTCC TCGAGCTCTT TGACGATTC GTTTTGAAT TCTTGATAC CTGCGGAG TGTCTGCGA GGCACCGCG CGTGGACCGC

25001 GACATCATTT TCCCCGACG CTTGCTTAA ACCCTGCAAC AGCTCTGCC AGCTCAAGCA TGTTCAGAA TGTTCAGAA CTTTAGGAA TTTATCTTA
 CTGTAGTAA AGGGCTTGC GACCTGATTT TGGGACGTTG TCCGACAGCG TCTGAAGTGG TCAGTTCTGT ACAAGCTTT GAATCTCTG AATATGATV

25101 AGGCTCAGG AATCTTGGC GCCACCTCT GTGCACTTC TAGGCACTTT GTGCCCATTA AGTACCGGA ATGCCCTCG CCGCTTGGG GCGACTTCTA
 TCGGAGTCC TTAGAACGG CGGTGCAGCA CAGTGAAG ATCGCTGAA CACGGTAA TCAATGCGCT TACGGAGGC GCGGAACCC CGCTGACGAT

25201 CCTCTGCGG CTAGCCACT ACCTTGCCTA CCACCTGAC ATATGGAAG AGTGAAGCG TTACGCTCTA CTGAGTGT CTTGTGCTG CAACTATCTC
 GGAAGACGTC GATCGTTGA TGGAGCGAT GGTGAGACTG TATTACCTC TGCACCTGCC ACTGCCAGT GACCTCACAG TGACAGCGAC GTTGTATATC

25301 ACCCGGACC GCTCCCTGGT TTGCATTCG CAGCTGCTTA AGCAAGTCA AATTATCGTT ACCTTTGAGC TGCAGGCTC CTGCGCTGAC GAAAGTCTG
 TGGGGCTGG CGAGGAGCA AACGTTAAG GTGAGCGAT TGTCTTCACT TTAATAGCCA TGGAACTCG ACCTCCAGG GAGCGGACTG CTTTTCAGG

25401 CGGCTCGGG GTTGAACCTC ACTCCGCGC TGTGAGCTC GCTTACCTT GCAAAATTTG TACCTGAGGA CTACCAGCC CACGAGATTA GGTTCCTACGA
 GCGAGGCCC CACTTTGAG TGGGCCCCG ACACCTGCG TCCGATGGA GGTTTTAAAC ATGACTCTT GATGCTGCG GTGCTCTAAT CCAGATATCT

25501 AGACCAATCC CCGCCGCTA ATCGGAGCT TACCGCTTC GATTAATGCG AGCTGACAT TCTTGGCAA TTGCAAGCA TCACNAAGC CCGCCAGAG
 TCTGTTAGG GCGGGGGAT TACCGCTCG ATGGCGAGG CAGTAATGCG TCCGCTTGA ACACCGGTT ACGTTCTGT AGTTGTTCTG GAGCTTCTT

25601 TTTCTGCTAC GAAAGGAGG GCGGTTTAC TTGACCTCC AGTCCGCTCA GAGCTCAG CCAATCCCC CCGCGCCCA GCGCTATCAG CAACTAGCT
 AAGACGATG CTTTCCCTGC CCGCCAAATG AACTGCGG TCAAGCGCT CCTGAGTTG GGTAGGAG GCGGCTGCT CCGGATAGTC GTCTGCGCT

Figure 15P

PMRKA159ag MER682

~~~~~  
 PsII  
 ~~~~~

25701 GGGCCCTTGC TTCCAGGAT GGCACCCAAA AAGAAGCTTC ATATTATGCC GGTACCCACG GACGAGGAGG ATACTGGGA CAGTCAGGCA GAGGAGGTTT
 CCGGGAAAG AGGGTCTTA CCGGAGGTTT TTCTTCGAGG TCCAAATXAGG CTAATXAGG CTTCTCTCC CTTATGACCT GTACAGTCCGT CTCTCCAU

~~~~~  
 Ecoli  
 ~~~~~

25801 TGGACGAGGA GGAGGAGGAC ATGATGGNAG ACTGGAGAG ACTAGAGGAG GAAGCTTCCG AGGTGGAAGA GGTGTGAGC GAACACCCGT CACCTTCG
 ACCTGCCTCT CTCTCTCTTG TACTACCTTC TGACCTCTTC GGATCTGCTC CTTCGAAGGC TCCAGCTTCT CCACAGTCTG CTTTGTGGCA GTGGGAGCA

25901 CGCATTTCCC TCGCCGGGCG CCCAGAAATC GGTTCCTTAG GGTTCGGAAC TCCAGCATGG CTACAACTTC CGCTCTCAG GCGCGGCGG CACTGCGGT TCGCGGACCC
 GCGTAAGGCG AGCGGCGCG GGTTCCTTAG GGTTCGGAAC TCCAGCATGG GGTTCGGAAC TCCAGCATGG CGCTCTCAG GCGCGGCGG CACTGCGGT TCGCGGACCC

26001 AACGTTAGAT GGGACACAC TGGACCCAG GGTTCCTTAG GGTTCGGAAC TCCAGCATGG GGTTCGGAAC TCCAGCATGG CGCTCTCAG GCGCGGCGG CACTGCGGT
 TTGGCATCTA CCTGTGCTG ACCTGTGCTC GGTTCGGAAC TCCAGCATGG GGTTCGGAAC TCCAGCATGG CGCTCTCAG GCGCGGCGG CACTGCGGT TCGCGGACCC

26101 GCGGACACAA GAAGGCAATA GTTCTTGT TGCAGACTG TGGGAGCAAC ATCTCTCTTC CCGGCGGCTT TCTTCTCTAC CATTACGCGG TCGCTTCTC
 CGCCCGTGTT CTTCGGGTAT CAGCGACGA AGTTCCTGAC ACCCCCGTTG TAGAGGAGGC GGGCGGCGNA AGAGAGATG GTAGTCCCGC ACGTGAAGG

26201 CGGTAAACAT CTGCATTACT ACCGTCACT CTACAGCCCA TACTGACCG GCGGAGCGG CAGCNACAG CAGGCGGCTA CAGAGCCGGA GCGGACCCGA
 GGCATTGTAG GACGTATGA TGGCAGTAGA GATGTGGGT ATGACGTTGG CCGCTCTCCG GTCTGTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT

26301 TAGCAAGACT CTGACAAAGC CCAAGAAATC CACAGCGCG GTCGCTCTG CAGGAGGAG GTCGCTCTG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT
 ATCGTCTGA GACTGTCTG GGTCTCTTAG GTCTCTCTG TGTCTCTG CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT

26401 CTAGAAACA GATTTTTC CACTCTCTAT ACAGAGGAG GGTCTCTG CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT
 GATCTTTGT CCAAAAGG GTGACATGA CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT

26501 CCGCAGCTG CCGTATAC CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT
 GCGGCTGAC GACATAGT TTTGCTTC TAGTGAAGC CCGTCTCTG CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT

26601 CTAGTTTCT GCGCTTCTC AATTTAAGC GCGMAACTA GGTCTCTC CAGGAGGAG GTCGCTCTG CAGGAGGAG GTCGCTCTG TCGCGGCTT GTCTTCTGT CCGCTGGCT
 GATCAAGCG CCGGAGAG GATTTATC GGTCTCTC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC TCGCGGCTT GTCTTCTGT CCGCTGGCT

26701 GAATTTCCA CCGCTTAT GTCGAGTAC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC TCGCGGCTT GTCTTCTGT CCGCTGGCT
 CTTTAAGGCT GCGGATGA CAGCTCAATG GTCGCTCTC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC CAGGAGGAG GTCGCTCTC TCGCGGCTT GTCTTCTGT CCGCTGGCT

~~~~~  
 Ecoli  
 ~~~~~

26801 GACCTTACAT GATATCCCG GTCAACGGA TACGCGGCA CCGAAACCGA ATTCTCTCTG AACAGCGGC TATTACCAC ACCTCTCTA ATAACCTTA
 CTGGGCTGA CTATAGGCG CAGTTGCTT ATGCGCGGT ATGCGCGGT GGTCTCTCT TAAAGAGAC TTCTCTCTG ATATGCTG TGTGGAGCAT TATTTGAAAT

26901 TCCCGGTAGT TCGCCCGCTG CCGTGTGT CCGGAAAT CCGCTCTCTA CCGCTCTCTA CCGCTCTCTA CCGCTCTCTA CCGCTCTCTA CCGCTCTCTA
 AGGGCATCA ACCGCGGAC GCGGACCAT GGTCTCTCTA GGTCTCTCTA GGTCTCTCTA GGTCTCTCTA GGTCTCTCTA GGTCTCTCTA GGTCTCTCTA

27001 TCAGGCGGC AGCTTGGCG GCGCTTCTG CAGGAGGAG CAGGAGGAG CAGGAGGAG CAGGAGGAG CAGGAGGAG CAGGAGGAG CAGGAGGAG
 AGTCCCGCG TCGNACGCG GCGGAAAGCA GTCTCTCTA GTCTCTCTA GTCTCTCTA GTCTCTCTA GTCTCTCTA GTCTCTCTA GTCTCTCTA GTCTCTCTA

27101 ACGAGTCTGT GAGCTCTCTG CTTGCTCTC GTCTCTCTC GATCTCTCTA ATCTCTCTA ATCTCTCTA ATCTCTCTA ATCTCTCTA ATCTCTCTA
 TGTCTAGCA CTCGAGGAG GAGCTCTCTC CAGCTCTCTC CAGCTCTCTC CAGCTCTCTC CAGCTCTCTC CAGCTCTCTC CAGCTCTCTC CAGCTCTCTC

~~~~~  
 PsII  
 ~~~~~

27201 TCTGCAGACC TGTCTCTCT AGCTCGCTC TGGAGGCAAT GGAATCTCT GGAATCTCT GGAATCTCT GGAATCTCT GGAATCTCT GGAATCTCT
 AGACGCTCTG AGCAGGAGC TCGCGCGAG ACCTCGTA CCGTCTCTA CCGTCTCTA CCGTCTCTA CCGTCTCTA CCGTCTCTA CCGTCTCTA

Figure 150

pMRKd5qag MER682

27301 CCTCCCGGCC ACTATCCGGA TCAATTATTT CCTAATTTTG ACGGTGTAA GCACTCGCG GAGGCTACG ACTGAATGTT AAGTGGAG AGAGACCAAC
 GAGGCGCCGG TGATAGGCCT AGTTANATNA GATTGAAAC TGATGATTTT CTTGAGCCGC CTGCGGATGC TTACTTTACA TTACCTCTC CGTCTCTTTT
 27401 TCGCGCTGAA ACACCTGGTC CACTGTGCTT GCCCAAGTC GTTTCCTCC CACTTCGGTG AGTTTACTA CTTTGAATTT CCGAGGATC ATATCGNGTG
 ACGCGACTT TGTGACCCAG GTGACAGCG CGGTGTTAC GAAATGCGG CTGAGGCCAC TCANAAAGAT TCANAACTTAC GGGCTCTAG TATAGCTCTT
 27501 CCGGCGCAC GCGGTCCGCG TTACCGCCA GGGAGACTT GCGCTAGTC TGATTGCGA GTTACCAG CGCCCCCTGC TAGTTGAGCG GGCACGGTA
 GCGCGCGTG CCGCAGGCG NATGGCGGT CCTCTCGA CCGGCATCG ACTAAGCCT CAAATGGTC CAAATGGTC ATCAACTCG CCTGTCCCTT
 27601 CCCGTGTGTC TCACTGTGAT TTGCNACTGT CCTAACCTTG GATTACATCA AGATCTTTGT TGCCATCTCT GTGCTGATA TAATAATAC AGAATTAATA
 GGBACACANG AGTGACACTA AACGTGACA GGTATGGGAC CTANATAGT TCTAGNANCA ACGGTAGAG CAGACTCAT ATTATTTATG TCTTTAAT
 27701 ATATACTGGG GTCTCTATCG CCATCTGTA AACGCCACG TCTTCACCG CCTTACCTGG CTTTACCTGG TACTTTTAC ATCTCTCC ATCTCTCC
 TATATGACCC CGAGGATAGC GTTAGGACAT TTGCGGTGC AGATGTGGC GGTTCGGTT GGTTCGGTT GGAATGGACC ATGAAATATG TAGAGAGGTA
 27801 CTGTGATTTA CAACAGTTTC AACCCAGAG GATGAGTCT ACGAGAGAAC CTCTCCGAG TCAGCTACTC CATCAGAAAT AACCCACCC TCCTTACCTT
 GACACTAAT GTTGTCAAAG TTGGGTCTGC CTCACTCAGA TGCTCTCTTG GAGAGGCTCG AGTCGATGAG GTAGTCTTTT TTGTGGTGG AGGATGGAG
 27901 CCGGAACTT ACCAGTGGT CACCGCGCG CACCGCGCG TTACCCGCTG ACCGTAAACC AGACTTTTTC CCGACAGACC TCNATACTC TGTTTACCA
 GCGCTTTGCA TGCTCAGCA GTGCGCGCG ACCTGGTGT GATGCGGAC TGCCATTTGG TCTGAAAAAG GCGTGTCTTG AGTTATTGAG ACNAAATGGT
 28001 AACAGGAGT GAGCTTAGA AACCTTAGG GTATTAGGC AAGGCGCAG CTACTGTGTG GTTTATGAAC AATTCAGCA ACTCTACCGG CTATTTCTAAT
 TTGTCTCTCA CTGAACTTT TTGGAACTCC CATNAATCCG TTTCGCGCT CATGACACCC CAATATCTG TTAAGTTCTG TGAGATGCC GATAAGAT
 28101 TCAGGTTTCT CTAGAACTGG GGTGGGGT ATTCTCTGC TTGTGATCT CTTTATCTT ATACTAACCC TTCTCTGCT AAGCTCGCC GCTGCTT
 AGTCCAAAGA GATCTTAGC CCAACCCCA TAAGAGACAG AACACTAAGA GAATATAGAA TATGATTCGG AAGAGACGGA TTCCGAGCG TTCCGAGCAC
 28201 TGACATTTG CATTTATGT CAGCTTTTA AACGCTGGG TGCGCACCCA AGATGATTAG GTACATAATC CATGTATTAG GATCCAAATG AGTGGAAAG CAGTCCGGT
 ACCTGTAAAC GTAAATACA GTGAAATAT TTGCGACCC AGCGGTGGT TCTACTAATC CATGTATTAG GATCCAAATG AGTGGAAAG CAGTCCGGT
 28301 GGTACCACCC AAGAGTGA TTTAAGGAG CAGCTGTA AGTTACATTT CCGAGCTGAA GCTATGATG GTACCACTCT TATAAATGC ACCACAGM
 CCATGGTGGG TTTTCCACCT AATTTCTC GTTGGACAT TACNATGTA GCTTCGACTT CGATTACTCA CCGTGGTGA ATATTTTACG TGGTCTTT
 28401 ATGAAAGCT GCTTATTGC CACAAACA AATTTGCA GTATGTTGT TATGCTATTT TACAGCCAGG TGACACTACA GAGTATAATG TTACAGTTT
 TACTTTTGA CCAATAGCG GTGTTTTTGT TTTAACCTT CATACGACAA ATACGATAAA CCGTGGTCC ACTGTGATGT CTCATATTAC AATGTCANAA
 28501 CCAGGGTAAA AGTCATANA CTTTATGTA TACTTTCCA TTTTATGAA TGTCCGACAT TACCATGTAC ATGAGCAAC AGTATAAGTT GTGCCCCCA
 GGTCCCATTT TCAGTATTTT GAATATCAT ATGAAAGGT AATATCTTT ACACGCTGTA ATGTACATG TACTCGTTG TCATATTCAA CACCGTGGT
 28601 CAANAATG TGGAACAC TGCCACTTC TGCTGACTG CTATCTAAT TACAGTCTC TCTTTGCT GTACCTTACT CTATATTAA TACANAAACA
 GTTTTACAC ACCTTTGTG ACCGTGNAAG ACGAGTGC GATACGATTA ATGTCCAGAG CCAAAACCAG GCTTTGCT CATGGATGA GATATATTT ATGTTTTCT
 28701 GACCCAGCTT TATTGAGGA AAGAAATGC CTTAATTTAC TANGTACAA AGCTAATGC ACCACTAAT GCTTTACTCG CTGCTTGC AAACNATTT
 CTGCGTGA AATACCTCT TTCTTTTAC GAATTAATG ATTCATGTT TCGATTACG TGGTATGA CGAATGAGC CGAATGAGC TGTCTTAAAT TTCTTTAAAT
 28801 AAGGTTAGC ATTATAATTA GAATAGGAT TAAACCCCTT GGTCAATTT TGTCTAATC TGTCTAATC CATTCCTCG AACAAATGAC TCTATGTCG ATATGCTCA
 TTTTCAATCG TAATATTAAT CTTATCTTAA ATTTGGGGG CCAAGTAAAG ACGAGTTATG GTAAGGGAG TTGTTACTG AGATACACCC TATACAGCT

Figure 15R

pMRKad59ag MER682

28901 GCGCTACAAC CTTGAAGTCA GGCTTCCTGG ATGTCAGCAT CTTACTTTGG CCAGCACCTG TCCCGCGGAT TTGTTCCAGT CCAACTACAG CCACCCACATC
 CCGGATGTTG GAACCTCAGT CCGAUGGATC TACAGTCTGA GACTGAAACC GGTGCTGAC AGGGGCGCTA ANCAAGGTCA GGTGATGTC GCTTGGTATG
 29001 TAACAGAGAT GACCACACCA ACCACGGCGG CCGCGCTAC CCGACTTACA TCTACACCA ATACACCCCA AGTTCTGCC TTGTCATAA ACTGGGATAA
 ATTGCTCTTA CTGGTTGTGT TGGTGGCGG GCGCGGATG GCTGAATGT AGATGTTGTT TATGTTGGT TCANAGACGG AACAGTTAT TGACCCCTAT
 29101 CTTGGGCATG TGGTGGTTCT CCATAGCGCT TATGTTGTA TGTCTTATTA TTATGTTGCT CATCTGCTGC CTAAAGCGCA AACCGCGCG ACCACCCATC
 GAACCCGTAC ACCACCAAGA GGTATGCGCA ATACAAACAT ACCGATATAT TCCATAGATT GACCGCATG AACACATGT TCTTTCTCT TACAGTATGA TTAAATGAGA
 29201 TATAGTCCCA TCATTGTGCT ACACCCAAAC AATGATGMA TCCATAGATT TCCATAGATT GACCGCATG AACACATGT TCTTTCTCT TACAGTATGA TTAAATGAGA
 ATATCAGGGT AGTAACACGA TGTGGGTTTG TTACTACTTT AGGTATCTTA CCGCGCTGAC TTGTTGTACA AGAAAGAGA ATGTCATACT AATTACTCT
 29301 CATGATTCCT CGAGTTTATA TATTACTGAC CTTGTTGGG CTTTCTTGG GGTCTCCAC ATTGGCTGG ATTGCTGACA TCGAAGTGA CTGCAATTC A
 GTACTAAGGA GCTCAAAAT ATAATGACTG CGAACACCG GAAMAAACAC GACCGAGGTG TAACCGACGC CAAAGAGTGT AGCTTCATCT GACGTAAAG.T
 29401 GCGTTCACAG TCTATTGCT TTACGGATTT GTACCCCTCA CCGTCATCTG CAGCCTCATC ACTGTGCTCA TCGCCTTTAT CCACTGCAT CACTGGTCT
 CGGAAGTGT AGATAACGA AATGCTTAA CAGTGGAGT GCGAGTAGAC GTGAGTAG TACACCCAGT AGCGAATAA GGTACCGTAA CTGACCCACA
 29501 GTGTGGGCT TGCATATCT AGACACCATC CCGAGTACAG GGACAGGACT ATAGCTGAG TTCTTAGAAT TCTTTAATTA TGAATTTTAC TGTGACTTT
 CACACCGGA ACCTATAGAG TCTGTGTAG GGTCTATGTC CCGTCTCTGA TATCGACTCG AAGATCTTA AGAATTAAT ACTTTAAATG ACACCTGAATA
 29601 CTGCTGATTA TTGACACCT ATCTGGGTT TGTTCCTCGA CCGTCCAGCC TCNAGACAT ATATCATGCA GATTCACCTG TATATGGAAT ATTCGAAGT
 GACCACTAAT AACGTGGA TAGACGCA ACACGGGCT GGAGGTTCTG AGTTCTCTGA TATAGTACGT CTAAGTGAGC ATATACCTTA TAAAGTTCTA
 29701 GCTACATGA AAAAAAGAT CTTTCGAAG CCGGTTATA TCGATCATC TCTGTTATGG TGTTCCTGAG TACCATCTTA GCGCTAGCTA TATATCCCA
 CGATGTTACT TTTTGGCTA GAAAGGCTTC GGACCAATAT ACCTTAGTAG AGACAATACC ACAAGACGTC ATGCTAGAAT CCGGATCGAT ATATAGGAT
 29801 CCGTGCATTT GCGTGGACG CAATAGATGC CAGAACCCAC CCAACTTTCC CCGCGCGCG TATGCTTCCA CTGCAACAG TTGTTGCGCG CCGCTTTGTC
 GGAACGTGA CCGACCTTGC GTTATCTACG GTACTTGGTG GGTGAAAG GCGCGGGCG ATACGAAGT GACGTTGTC AACAAACGGCC GCGGAACA :
 29901 CCAGCCATC AGCTGCGCC ACCTCTGCC ACCTCCACTG AATCAGCTA CTTTAATCTA ACAGGAGGAG ATGACTGACA CCCTAGATCT AGAATGAC
 GGTGCTTNG TCGAGCGGG TGGAGAGCG TGGGGTGAC TTAGTCTGAT GAATTAGAT TGTCTCTCTC TACTGACTGT GGGATCTAGA TCTTTACCTG
 30001 GGAATTTATA CAGACGCGG CTTGCTAGAA AGACGCGGG CAGCGCGGA GCAACGCGC ATGATCAAG AGCTCCAGA CATGTTAAC TTGCACCAAT
 CTTAATAT GTCTGCTGC GGAGGATCTT TCTGCGTCCC GTGCGCGCT TACTTAGTTC TCGAGGTTCT GTACCAATTTG AACCTGGTCA
 30101 GCAAAAGGG TATCTTTTGT CTGTAAGC AGGCCAAGT CACTAGGAC AGTAATACCA CCGACACCG CTTAGCTAC AAGTTGCCA CCAAGCGTTA
 CGTTTTCCC ATAGAAACA GAGCATTTG TCGGTTTCA GTGATGCTG TCATTATGTT GCGCTGCGC GGNATGATG TTCAACCGTT GGTTCGAT
 30201 GAATTTGGT GTCATGGTG GAGAAAGCC CATTACATA ACTCAGCACT GGTAGAAAC CCAAGGCTGC ATTCACCTAC CTTGTCACAG ACCTGAGCAT
 CTTTAACCC CAGTACCACC CTCTTTTGG GTAAATGAT TGAATCTG GGCATCTTG GCTTCGAGC TAAGTGATG GACAGGTTCC TCGACTCTTA
 30301 CTCGACCC TTATTAGAC CCGTGTGGT CTCGAAGAT TTTATCCCT TAACTAATA AAAAAATA TAAAGCATCA CTTACTTAA ATCAGTTAGC
 GAGACGTGG AATATCTG GAGACGCGA GAGTTCTAG AATAGGGA ATTGATTAT TTTTATTAT ATTGCTAGT GAATGAATTT TAGTCAATCG

Figure 155

pMRKad5gag MER682

30401 AATTTCTGT CCAGTTTATT CAGCAGCACC TCTTTGCTT CCTTCCAGCT CTGTATTGCG AGCTTCTCC TGGCTGCAG CTTCTCCAC AATCTAAATG
 TTTAAGACA GGTCAATATA GTCTCTGTGG AGGACGGGA GAGAGTCCGA GACCATACAG TCGACGAGG ACCGAGTTT GAAGAGGTG TTAGATTATC
 30501 GAATGTCAAT TTTCTCTGT CCGACCCAC TATCTTCATG TGTGTGAGA TGAAGGCGC AAGACCGTCT GAAGATACCT TCAACCCCTT
 CTTACAGTCA AAGGAGACA AGGACAGGTA GTGTGTGTG ATAGAGTAC AACAAGTCT ACTTCTGCG TCTGCGAGA CTTCTATGGA AGTTGCGG A
 30601 GTATCCATAT GACACGANA CCGTCTCTCC AACTGTGCT TTTCTTACT TTTCTTACT CTCCCTTGT ATCCCCCAAT GGGTTTCAAG AGATCTCCCC TGGGGTACT :
 CATAGGTATA CTGTCTCTTT GGCAGGAGG TTGACACGGA AAGAAATGAG GAGGGAACA TNGGGGTTA CCCAAGTTC TCTCAGGCGG ACCCCATGAG
 30701 TCTTTGCTCC TATCCGAACC TCTAGTTACC TCCATGGCA TGTCTGCTT CAAATGCGC AACGCTCTT CTTCTGGAGA GGCCTGCAC CTTACCTCC:
 AGAAGCGGG ATAGGCTTGG AGATCAATGG AGGTTACCT ACGACGCGA GTTTTACCG TTCCCGGAGA GAGACCTGCT CCGGCGGTG GAATGGAG :
 30801 AATATGTAAC CACTGTGAGC CCACCTCTCA AATAAATATA AACTTGANA TATCTGACCC CCTCAGATT ACCTCAGAG CCTTACTCT
 TTTTACATTG GTGACACTCG GTTGAGAGT TTTTGTGTT CAGTTGTAT TTGACCTTT ATAGACGTG GAGGTGTCAA TGGAGTCTTC GGGATTGACA
 30901 GGTGCGGCC GCACCTCTAA TGGTGGGG CAACACATC ACCATCAAT CACAGGCCC GCTAACCGTG CACGACTCCA AACTTAGCAT TGGCACCCCA
 CCGACGGGG CGTGGAGATT ACCAGCCTCC GTTGTGTGAG TGTACGTTA GTGTCCGGG CGATTGGCAC GTGCTGAGGT TTGANTGTA ACBGTGGG :
 31001 GGACCCCTCA CAGTGTGAGA AGGAAAGCTA GGCCTGANA CATCAGGCC CCTCACACC CCGATAGCA GTACCTTAC TATCACTCC TACACCTCT
 CCTGGGAGT GTACAGTCT TCTTTTGGAT CGGACGTTT GTAGTCCGG GTAGTGGTG GAGTATGCT TGGCTATCT CATGGBAATG ATAGTACCG AGTGGCGA
 31101 TAACTACTGC CACTGGTNGC TTGGGATTG ACTTGAAGA GCTCATTTAT ACACANAATG GAAACTAGG ACTAAGTAC GGGCTCTCT TGCNTGTA
 ATTGATGAGG GTGACCATCG AACCCGTAC TGAATTTCT CCGGTAAATA TGTGTTTAC TGTGTTTAC TGTGTTTAC TGTGTTTAC TGTGTTTAC
 31201 AGACGACCTA AACCTTTGA CCGTAGCAC TGGTCCAGT GTGACTATTA ATAATACTTC CTTCMACT CTTCMACT GAGCTTGGG TTTTGATTCA
 TCTGCTGAT TTGTGNACT GGCATCGTT GGCATCGTT ACCAGTCCA CACTGATAAT TATTATGAG GAAGTTTGA TTTCAATGAC CTGGBAACCC AANAATAA
 31301 CAAGGCAATA TGCANCTAA TGTAGCAGGA GGCATGAAGA TTTATTTCTA TTTATTTCTA TTTATTTCTA TTTATTTCTA TTTATTTCTA TTTATTTCTA
 GTTCCGTTAT ACTTGAAT ACATCGTCT CTTGATTATA TTTTATANA CTTAGGCTC AACTAAGAG TTTGTCTGG GAATATGAA TTAACACTA CANAGGCTT TACTTGTTA CAGCTTCAAA
 31401 AACTAAATCT AAGACTAGGA CAGGCTCTC TTTTATANA CTTAGGCTC AACTAAGAG TTTGTCTGG GAATATGAA TTAACACTA CANAGGCTT TACTTGTTA CAGCTTCAAA
 TTGATTTAGA TTTGATCTT GTCCCGGAG AATAATTT GAGTGGGTG TTGAACCTAT AATTGATGT GTTTCGGA ATGACCAAT GTCTGAGTTT
 31501 CAATTCGAA AAGCTTGAG TTACCTAAG CACTGECAG GGTGTGATG TTGACCTAC AGCATAGCC ATTATGCG GAGATGGCT TGAATTTGT
 GTTAGGTTT TTGAACTCC AATTGGATT GTGACGGTT CCAACTACA AACTGGATG TCGGTATCG TGAATACG TAATTACG CTCTACCGA ACTTAAACCA
 31601 TCACCTAATG CACCAACAC AATCCCTC ABAACANAA TTGECATGG CTTAGAAAT CTTAGAAAT CTTAGAAAT CTTAGAAAT CTTAGAAAT
 AGTGGATTAC GTGGTTTGT TTTAGGGGAG TTTGTGTTTT AACCGTACC GGATCTTAA CTAGTTTGT TCCGATACCA AGGATTTGAT CCTTACCTG
 31701 TTAGTTTGA CAGCAGGT GGCATTACAG TGGNAACA AATAATGAT AACTAATG TTTGACCCAC ACCAGCTCCA TCTCTTACT TCTCTTACT GTAGACTAAA
 AATCAAAACT GTGCTGTCCA CCGTAATGTC ATCTTTGTT TTTATTTACTA TTGATTTGA ACACCTGTG TGGTGGAGT AGAGGATGGA CATCTGATTT
 31801 TGCAGAGAA GATGCTAANC TCACCTTGGT CTTACANAA TGTGCTGTC AATACTTGC TACAGTTGA GTTTTGTGTG TTAAGGCGAG TTTGGCTCCA
 ACCTCTCTTT CTACGATTTB AGTGAACCA GAATTGTTTT ACACCTGAG TTTATGNAAG TTTATGNAAG TTTATGNAAG TTTATGNAAG TTTATGNAAG
 31901 ATATCTGAA CAGTTCAAG TGTCTATCT ATTATAAGAT TTGAGGAA TTTAGGCTA TTTAGGCTA TTTAGGCTA TTTAGGCTA TTTAGGCTA
 TATAGACCTT GTCAAGTTC ACAGTAGAA TAATATTCTA AACTGCTTTT ACCTCAGCAT GATTTGTAA GGAAGGACCT GGTCTTATA ACCTTCAAT
 32001 GAATGGAGA TCTTACTGA GGCENACT ATACAAAGC TGTGTGATTT ATGCTTACC TATGCTTA TCCAAATCT CAGGTAAA CTGCCAAG
 CTTTACCTCT AGATGACTT CCGTCTCGA TATGTTTCTG ACACCTANA TACGATGAG ATAGTGAAT AGGTTTGA GTGCCATTTT GACGTTTTT

Figure 15T

pMRKAD59ag MFR682

32101 TAACATTGTC AGTCAAGTTT ACTTAACGG AGACAAACTT AACCTGTGTA CACTAACCAT TACACTAAC GGTACACAGG AACAGGAGA CACAACTCCA
 ATTGTAAACAG TCAGTTCAAA TGAATTGGC TGTGTTTGA TTGACANIT GGTATTGTA ATGTATTG CCATGTGCC TTGTGCTCT GTGTGAGT
 32201 AGTGCATACT CTATGTCAAT TTCAAGGGAC TGTGTGCGC ACANITAT TATGAAATA TTGTGCACAT CTCTTTACAC TTTTTCATAC ATTGCCCCAN
 TCACGTATGA GATACAGTAA AGTACCCCTG ACCAGACCGG TGTGTATATA ATTACTTTAT AACGGTGA GGAAGATAG AAAAAGTATG TAACTGCTT
 32301 AATAAGAAAT CGTTTGTTT ATGTTTCAC ATGTTTAATTT TTCAATTTTA GAAATTTCA ATCTATTTT CATTCACTAG TATAGCCCCA CCACACATA
 TTAATTTCTA GCAAAACAA TACAAAGTG CACAAATAA AAGTAAAGT CTTTAAAGT TCAATTAATA GTAACTCATC ATATCGGGGT GTGTGTGTA
 32401 GCTTATACAG ATCACCCTAC CTTAATCMA CTCACAGAAC CCTAGTATTC AACCTGCCAC CTGCTGCCA ACACACAGAG TACACAGTCC TTTTTCCTT
 CGAATATGTC TAGTGGCATG GAATTAATTT GACTGCTCTG GATCATTAAG TTGTACGGTG GAGGAGGGT TGTGTGCTC ATGTGTGAGG AAGAGGG
 32501 GCTGGCCTTA AAGAGCATCA TATCATGGGT AACAGACATA TTCTTAGGTG TTATATTTCA CACGGTTTCC TGTCGAGCCA AACGCTCATC AGTATATTT
 CGACCCGGAAT TTTTGGTAGT ATAGTACCA TTGTCTGTAT AAGATTCAC AATATNAAGT GTGCCAAGG ACAGCTCGGT TTGCGAGTAG TCACTATA
 32601 ATAAACCTCC CGGACAGTTC ACTTAAGTTC ATGTGCTGT CTACTGTG AGCCACAGGC TGCTGTCCA CTTGCTGTTG CTTAACGGGC GGTGAGGTA
 TATTTGAGGG GCCCGTGGAG TGAATTCAG TACAGCGACA GTGTGAGGAC TGGGTGTCCG AGACAGGTT GAACGCCAAC GAATTTGCCG CGCTTCTCT
 32701 AAGTCCACGC CTACATGGGG GTAGATCAT ATCTGTGCAT CAGGATAGGG CGTGTGTCT GCAGCAGGC GCGANTAAAC TGCTGCCGCC GCGCTTCTCT
 TTCAGGTGGG GATGTACCCC CATCTCAGTA TTAGCAGTA GTCTATATCC GCCACCAGA CGTGTGTCCG CCTTATTG ACCACGGGCC CGCGGAGGCA
 32801 CCTGCAGGAA TACAAATGG CAGTGTCTC CTCAGCGATG ATTGGCAGCG CCGGAGCAT AAGGCTCTT GTCTCCGG CACAGCAGCG CACCTGAA
 GGACGTCTCT ATGTTGTACC GTACCCAGAG GATGCTCTAC TAAAGCTTTC TAAAGCTTTC TCCGCTGTA TTCCGCGGAA CAGGAGGCCG GTGTGCTGC GTGGACT
 32901 TCACTTAAT CAGCAGCTA ACTGCAGCAG AGCACCACAA TATTTTTCAA ATCTCCACAG TGCAAGGCG GTATCCAA GCTCATGGCG GCGACCACAG
 AGTGAATTTA GTGTGTCTAT TGACCTGCTG TGTGTGTTT ATAAAGTTT TTAGGCTCTC AGTTCTCCCG ACATAGTTT CCGTACCGC CCTGTGCTC
 33001 AACCCACGTC GCTATCATAC CACAAGGCA GGTAGATTA GTGCGGCTC CTCATNAACA CBTGTGACAT AACATTTACC TCTTTGGA TGTGTAAAT
 TTGGGTGCAC CGTATGATG GTGTGCGGT CCATCTAAT CCAGCTGGG GAGTATTTGT CCGACCTGTA TTTGTAAATG AGAAACCGT ACAACATTAA
 33101 CACCACTTCC CGTACCATA TAACTCTG ATTAACATG GCGCCATCA CCACCATCTT AACCCAGCTG GCCAANACCT GCTCCGCCG TATACACTA
 GTGTGTGAGG GCCATGGTAT ATTTGGAGAC TAATTTGTAC CCGGTAGGT GTGTGTAGGA TTGTGTGAC CCGCGGGCG ATATGTGAG
 33201 AGGAAACCGG GACTGGAACA ATGACAGTGG AGAGCCAGG ACTGTAAAC ATGTCTGTCA TGATATCAAT GTTGGCACA CACAGGCACA
 TCCCTTGGCC CTGACCTTGT TACTGTCTCC TCTCGGCTCC TGACATGAG TACCTAGTAG TACGAGCAGT ACTATAGTTA CAACCGTGT GTGTGCTGT
 33301 CGTGCATACA CTTCTCAGG ATTACAGCT CTTCCCGCT TGAACCAT TCCAGGGA CAACCATTC CTGATCAGC GTAAATCCA CACTGCAGG
 GCACGTATGT GAGGAGTCC TAATGTTCCA GAGGCGCA ATCTGTGTAT AGGTCTCTT GTTGGGTAG GACTTAGTGG CATTTAGGT GTGACCTC
 33401 AAGACCTTCC ACCTAATCA CGTTGTGCT TGTCAAAAGT TTACATTCGG GACAGAGGG ATGATCTCC AGTATGAG CCGGGTTTC TGTCTCAA
 TTCGTGAGCG TGCATTGAGT GCACACCTA ACATTTTCAC AATGTAAAGC CTTCTCTCC TACTAGTAG TCAATCCATC GCGGCCAAG ACAGAGTTT
 33501 GAGGTAGAC GATCCTACT GTACGAGTG GCGGAGACA ACCAGATCG TGTGTGCTG AGTGTATTC CAATGGAAC GCGGAGCTA GTCTATTT
 CCTCCATCTG CTAGGATGA CATCTCTCAG CGGCTCTGT TGGCTCTAGC ACNACCACA TCACAGTAGG GTTTACCTTG CGGCTGCTAT CAGTATTAAT

Figure 15U

pMRKad5gag MER682

33601	CTGAGCAGA ACCAGGTGG GCGGTACAA ACAGATCTGC GTCTCCGCTC TGCCGCTTA GATCCCTCG TGTAGAGTT GTAGTATATC CACTCTCTTA
33701	GACITTCGTTT TGTCCACGC CCGCATGTT TGTCTAGCG CAGAGGCCAG AGCGGGGAT CTATCGAGAC CATCATATAG CATCATATAG GTGAGAGAT
33801	AAGCATCCAG GCGCCCTCG GCTTCGGGT CTATGTAAAC TCTTTCATGC GCTCTCCGC TATATACATC CACCAACCCA GANTAGCCA CACCCAGCC/
	TTCGTAGGTC CCGGGGAC CGAAGCCCA GATACATTG AGAAGTACG CGCGAGCGG ACTATTGTAG GTGGTGGGT CTTATTCGGT GTGGTTCGTT
	ACCTACACAT TCGTCTCG AGTCACACAC GCGAGGAGCG GGAAGACCTG GAAGACCAT GTTTTTTTTT TTATTCCTCA NGATTATCCA AAACCTT/AAA
	TGGATGTGTA AGCAAGACG TCAGTGTGTG CCTCTCTCG CCTCTTGTA CAAATAAAA AATAAGGTTT TCTAATAGGT TTTGGAGTTT
33901	ATGAAGATCT ATTAAGTGA CCGCTCCCG TCCGTTGGG TGTCTNACT CTACAGCCA AGACAGATA ATGCCATTG TAAGATGTTG CACATGGCT
	TACTTCTAGA TAATTCACIT GCGGAGCGG AGGCCAGCG ACCAGTTTGA GATGTGCTAT TCTTGTCTAT TACCGTAAC ATTCTACAC GTGTACCG/
34001	TCCAAAGGC AAGCGCCT CACGTCCAAG TCGAGTAA CCGTAAACCC TTCAGGTGA ATCTCTCTA TAAACATTCC AGCACCCTCA ACCATGCTCA
	AGTTTTTCG TTTGCGGGA GTGCAGTTC ACCTGCATTT CCGATTTGG AGTCCCACT TAGAGGAGAT ATTTGTNAGG TCGTGGNAGT TCGTACGGGT
34101	AAATAATTC ATCTGCCAC CTCTCTNATA TATCTCTAAG CAAATCCGA ATATTAGTC CGGCCATTGT AAAATCTGC TCCAGAGCGC CCTCCACCTT
	TTATTAAAG TAGAGCGGTG GAAGAGTTAT ATAGANTTC GTTTAGGCT TATAATTCAG CCGCGTAACA TTTTAGACG AGGTCTCGG GSAGGTGGAA
34201	CAGCCTCAG CAGCAATCA TGATTGCCAA AATTCAGTT CCTCAGAC CTGTATAAGA TTCAAAGCG TTCAAAAGCG GAACATTAC AANAATCCG CGATCCCGTA
	GTCCGAGTTC GTCCCTTAGT ACTAACGTTT TTAAGTCCA TTAAGTCTG GACATATTCT NAGTTTTGCG CTGTATATG TTTTATATGC GCTAGGGCAT
34301	GATCCCTTC CAGGCCAGC TGAACATAAT CGTCAGGTC TGCACGACC AGCGGGCA AGCGCGCC CTTCGCCGCG ACGAACCATG ACAAAAGAAC CCACACTGAT
	CCAGGGAAG GTCCGGTGG ACTTCTATTA GCACCTCCAG ACCTGCCGCT TCGCGCCGT GAAGGGGCG TCGTGGTAC TGTTTCTTGG GGTGTGACTA
34401	TATGACACC ATACTGGAG CTATGCTAAC CAGCTAGCC CCGATGTAG CTGTGTGCT GGGCGCGAT ATANATGCA AGGTGCTCT CAAATAATC/
	ATACTGTGG TATGAGCTC GATACGATG GTGCATCGG GGTACATTC GAACAACGTA CCGCGCGTA TATTTTACGT TCCAGGAGA GTTTTTTAGT
34501	GCCAAAGCT CCGCAAAA AGAAGCACA TCGTAGTCA TCTCATTCAG ATAAAGGCG GTAAGCTCG GAACACCAC AGAANAAGAC ACCATTTTTC
	CCGTTTCGA GCGGTTTT TCTTTCGCT AGCATCAGTA CCGTAGCTC TATTTCCGTC CATTCGAGGC CATTCGAGGC TGTGTGTTG TGGTAAANG
34601	TCTCAACAT GTCTGGGGT TCTGCTATA ACACAAATA AATAACANA AATCAATTA MACATTAGAA GCTGTCTTA CAACAGGAAA AACAACCTT
	AGAGTTTGA CAGACGCCA AAGACGTATT TGTGTTTTAT TTTGTNAAT TTGTAACTCT TGTAACTCT CCGACAGAA GTTGTCTCTT TGTGTGGAA
34701	ATAAGCATTA GACGACTAC GGCATGCCG GGTGACCTG AATAAAGCA GTCAACCTGA GTCAACCTGA TTTAAAGCA CCACCGACAG CTCTCTGGTC ATGTCCGGAG
	TATTCGTATT CTGCTGATG CCGGTACGG CCGATCGCA TTTTCTGAC CAGTGCNCT ANTTTTCTGT ANTTTTCTGT GGTGGCTGT CAGGAGCTC
34801	TCATATATTA AGACTGGTA AACACATCAG GTTGATTCAC ATCGGTCA GTCAANAAGC GACCGAATA GACCGGGA ATACATAACC GEAGTGTAG
	AGTATTACAT TCTGAGCCAT TTGTGTAGTC CAATACGT TAGCCAGTCA CGATTTTTCG CTGGCTTTAT CCGGCCCTCT TATGTATGG TATGTATGG
34901	AGACACATT ACAGCCCCA TAGGAGGTAT AACAAATTA ATAGAGAGA AATACATA AACACCTGA AATACCTCT GCTAGGCAA AATAGCACCC
	TCTGTTGTA TGTCCGGGT ATCTCCATA TTGTTTTAT TATCTCTCT TTTTGTGTAT TTGTGGACTT TTGTGGAGA CCGATCCCTT TTATCTGTGG
35001	TCCCGCTCA GAACAACATA CAGCGCTTC CAGCGGCG AGCGCTTACC CAGCTTACC AGTAANAAG AATACCTATT AAAAAACAC CACTCCGAT
	AGCGGAGGT CTGTGTGTAT GTCCGGAGG TGTCCGCTC GGTATTGTA GGTCCGATGG TCATTTTTTC TTTTGTATTA TTTTGTATG GTGAGCTGTG
35101	GCCACCACT CAATCAGTCA CAGTGTAAA AAGGCCAG TCCAGAGCGA GTATATATAG GACTAANA GACTAANA TGACGTAACT GTTAAAGTCC ACANAANA
	CCGTGGTGA GTTAGTCAGT GTCAATTTT TCCCGGTT CAGTCTCCT CATATATATC CTGATTTTTT ACTGCATGC CAATTCAGG TATTTTTTGT
35201	CCAGAAAC CCGACCGGA CCTACGCCA GAACGAAAG CCANAAACC CACACTCC TCAATCTC TCANATCT ACTTCGGTT TCCACGTTA CGTCACTTC
	GGGTCTTTTG GCGTCCCTT GATGCCGT CTTGCTTC GTTTGAGG AGTTAGCAG TGAAGGAAA AGGTGCAT AGGTGCAT GCAGTGAAG

Figure 15V

pMRKd5gag MER6R2

35301 CATTTTAAGA AACTACAAAT TCCCAACACA TACAACTTAC TCCGCCCTAA AACCTACGTC ACCGCCCTCG TCCCNAGCC CCGGCCACG TCACAAACAC TC
 GTAAATTTCT TTTGATGTTA AGGTTTGTT AGTTTCAATG AGCGGKATTT TCGGATCCAG TCGGCGGCGC AAGGTTGGG GCGCGGTGC AGTGTTTGAG
 ~~~~~ PstI ~~~~~  
 35401 CACCCCTCA TTATCATATT GGTTCATC CANAATAAGG TATATTATTG ATCATGTTAA TTANGAATTC GGATCTGCGA CCGGAGGCTG GATGGCTTT  
 GTGGGGAGT AATAGTATAA CCGAAGTTAG GTTTTATTCC ATATAATTAAC TACTACAAAT TACTTCTTAAG CCTAGACGCT GCGCTCCGAC CTACCCGAAAG  
 35501 CCCATTATGA TTCTTCTCGC TTCCGGCGGC ATCGGATGC CCGGTTGCA GCGCATGCTG TCCAGGACG TAGATGACGA CCATCAGGGA CAGCTTCAAG  
 GGGTAATACT AAGAGAGGCG AAGGCGCGCG TAGCCCTACG CCGGTACGAC AGTCCGCTCC GCGCCCTTGA CGAGCATCAC AAAAACTGAC GCTCAGTCT  
 35601 GCCAGCCAAA GGCAGGAAAC CGTAAANAAG CCGGTTGCT GCGTTTCTC CATAGCTCC GCGCCCTTGA CGAGCATCAC AAAAACTGAC GCTCAGTCT  
 CCGTCTGTTT CCGTCTCTTG GCATTTTCTC GCGGCAACGA CCGCAANAAG GATATCCGAG CCGGGGACT CCGGGGCTCT CCGTCTCCGA ATCTACTGCT GTGGAAGTTC  
 35701 GAGGTGGCGA AACCAGACAG GACTATAAAG ATACCAGCG ATACCAGCG TTTCCTCTC GAGCTCTCT GAGCTCTCT GAGCTCTCT TACCGGATAC  
 CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGTTCCG AAGGCGGAC CTTCCGAGGA GCACGAGGCT GGTACGGCGA ATGGCTTAT  
 35801 CTGTCCTCTT TTCTCTCTTC GGGAGGCTG GCGCTTCTC ATAGCTCAG CTGTAGGTAT CTCAGTTCTG TGTAGGTCT TCGCTCCAG CTGGCTCT  
 GACAGGCGGA AAGAGGGAAG CCGTCTGAC CCGCAAGAG TATCGAGTCC GACATCCATA GATCANAACC ACATCCAGCA AGCGAGGTT GACCCGACN  
 35901 TGCACGAAAC CCGGCTTAC CCGGACCGCT GCGCTTATC CCGTAATAT CCGTAATAT GCGTAATAT GCGTAATAT GCGTAATAT GCGTAATAT  
 ACGTCTCTTG GCGGCAAGTC GCGCTGCGA GCGGAGGTA TGTAGGCGT GCTACAGAGT TCTTGAAGT GTGCTTAC TCGCTTAC AGTATTGCT  
 36001 CACTGCTAAC AGGATTAGCA GAGGAGGTA TGTAGGCGT GCTACAGAGT TCTTGAAGT GTGCTTAC TCGCTTAC AGTATTGCT  
 GTGACCAATG TCTTAATCTG CTGCTCTCT ACATCCGCA CCGTCTCTA AGAATCTCA GCGCTTAC TCGCTTAC AGTATTGCT  
 36101 ATCTGCGCTC TCGTGAAGCC AGTACCTTC GGAANAAG GCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 TAGACGCGAG ACGACTTCTG TCAATGGAAG CCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 36201 ACGAGGAGAT TACGCGCAGA AANAAGGAT CTCAGAAGA TCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 TCGTCTCTA ATGCGCTCT TTTTCTCTA GAGTCTCTT AGAAGACTAG AAGAGACTCC CCGACTGAG GCGAGCACAT CTATTGATG CGATCTCTA  
 36301 TTTGGTCAAG AGATTATCA AAGGATCTT CACCTAGATC CTTTAAATC AATCTAATG ATATATGAT AATCTGCTC TGTAGGCTC TGTAGGCTC  
 AAGCAGTAC TCTAATAGTT TTTCTCTAGAA GTGATCTAG CCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 36401 TCAGTGAGGC ACCTATCTCA CCGTCTCTC TATTTCTCTC TATTTCTCTC ATCCATAGT GCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 AGTCACTCTG TGTATAGAT CCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 36501 TGGGCCAGT GCTGCAATGA TACCGCGAGA CCGGCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 ACGCGGCTA CCGGCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 36601 CCTGCAACTT TATCTCTCTC CATCTCTCTC ATTATTTCTT GCGGCGGAGC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 GCGCTCTGAA ATAGGCGGAG GTAGGCTCAGA TAAATTAACA CCGGCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 36701 CTACAGGCTAT CCGTCTCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC TGTAGGCTC  
 GATGCTCTGA GCACCAAGT GCGAGGCTA ACGCTATCTG ACGCTATCTG ACGCTATCTG ACGCTATCTG ACGCTATCTG ACGCTATCTG ACGCTATCTG  
 ~~~~~ PstI ~~~~~  
 36801 AAAAAAGGTT AGCTCTCTC GTCTCTCTC GGTGCTCAGA AGTAAGTTG CCGGAGTGT ATTACTCATG GTTATGCGAG CACTGCTATA TTCCTTACT
 TTTTCCGCAA TCGAGGAGC CAGGAGGCTA GCACATCTT TCATTCNACC GCGCTCACA TAGTGAGTAC CAATACCTTC GTGACGTATT AAGAGATATA
 36901 GTCATGCCAT CCGTANGAT CTTTCTCTG ACTGCTGAGT ACTNACCMA GTCATCTCTA GAATAGTCTA TCGGCGGAGC GAGTGTCTT TCGCGGCGCT
 CAGTACGGTA GGCATCTTAC GAAAGACAC TGACCACTCA TGAGTTGTT CAGTAAGACT CTTATCATAT ACGCGCTG GCTNACGAGA ACGGCGCGCA

Figure 15W

PMRKA15qag MER682

37001

CAACACGGGA TAAATACCGG CCACATAGCA GAACTTTAAA AGTGTGATC ATTGGAAMAC GTTCTTCGGG GCGAATAC TCAGGATCT TACCGCTGTT

37101

GTTCGGCCCT ATTATGGCGC GGTGTATCGT CTTCAAATTT TCACGAGTAG TAACCTTTTG CAAGAGGCC CGCTTTTGAG AGTTCTTAGA ATGGCGACCA

37201

GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACCTA TCTTACATAT TCTTACGCTT CACCAACCTT TCTGGGTGAG CAAMAACAGG AACCCAAAT

37301

CTCTAGGTCA AGCTACATTG GGTGAGCAGG TGGGTGACT AGAAGTCTTA GAAATGAAA GTGTCTGCAA AGACCCACTC GTTTTTGTCC TTCCGTTTTA

37401

GCCGCANAAA AGGGAATNAG GCGGACACCG ANATGTTGAA TACTCATACT CTTCCTTTT CAATATTAAT GMAGCATTA TCAGGTTAT TCTCTCATCA

37501

CGGCGTTTTT TCCCTTATTC CCGCTGTGEC TTTACAACTT ATGAGTATTA GAAGGAMAAA GTTATAATAA CTTCGTAAAT AGTCCCAATA ACAGAGTAC

37601

GCGGATACAT ATTGGAATGT ATTAGAANA ATAAACAAAT AGGCTTTTC CCAACATTC CCGAAMAGT GCCACCTGAC GTCTAAGAAA CCATTATTA

37701

CGCCTATGTA TAACTTACA TAAATCTTTT TATTGTTTA TCCCCAGGC GCGGTAMAG GGGCTTTTCA CGGTGGACTG CAGATCTTT GGTATAATA

37801

CATGACATTA ACCTATAAAA ATAGGCGTAT CACGAGGEC TTTCGTCTTC AAGAAATGGA TCCGAATTC TAAT (SEQ ID NO: 27)

37901

GTACTGTAAT TGGATATTTT TATCCGCATA GTGCTCCGCG AAGCAGCAG TTCTTAACT AGCCTTAMGA ATTA (SEQ ID NO: 28)

Figure 15X

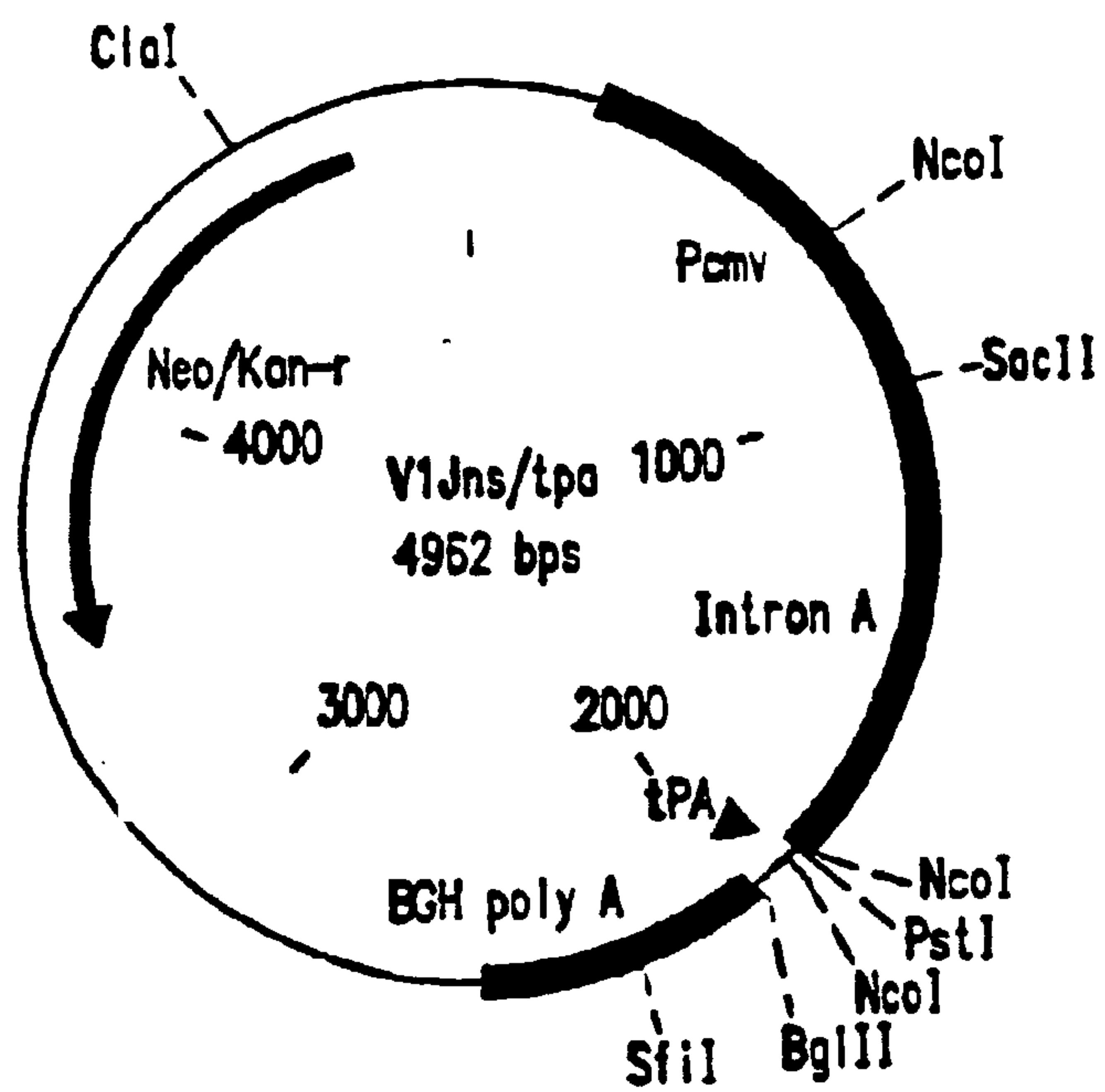
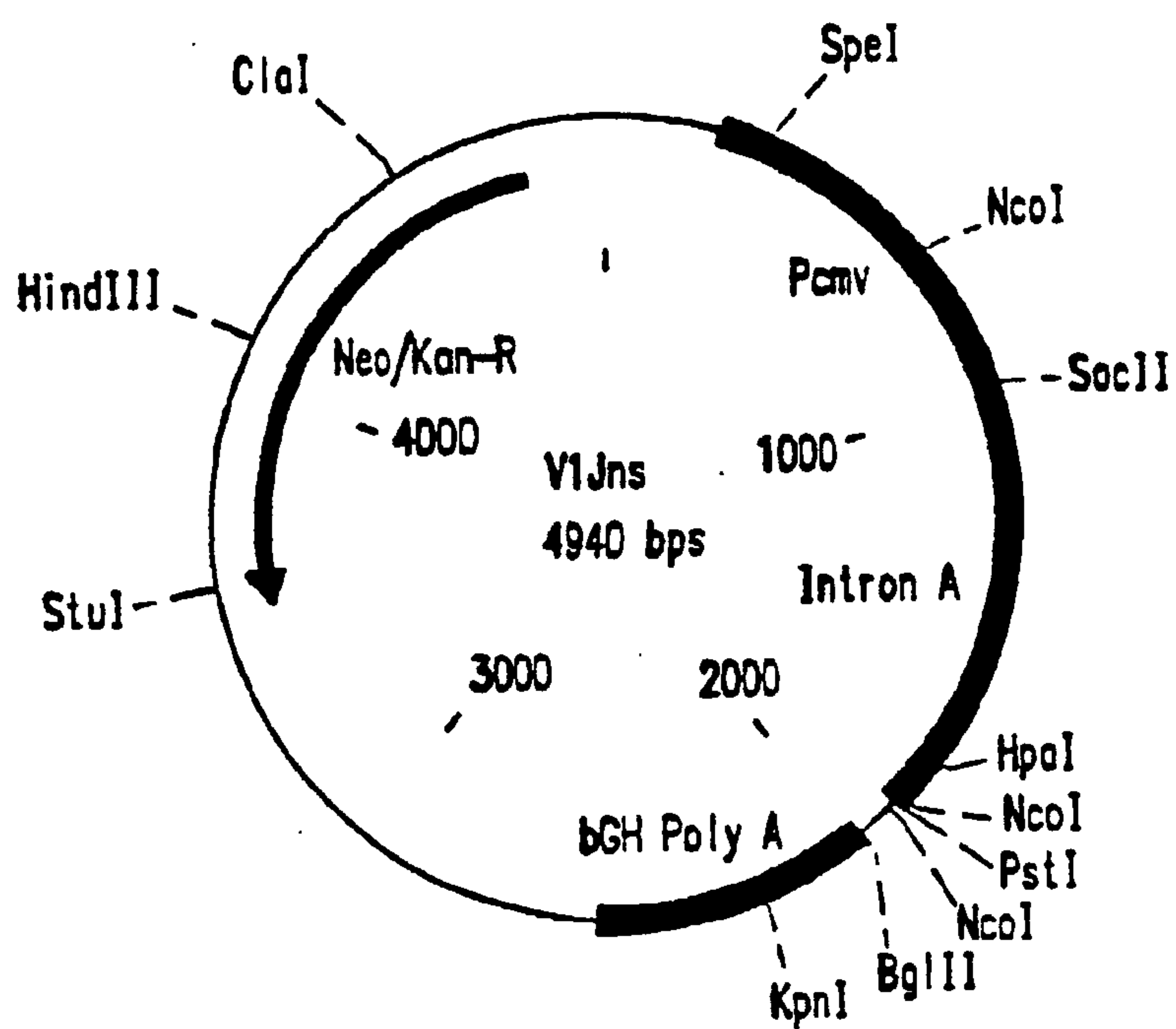


FIGURE 16

AGATCTACCATGGCCCCATCTCCCCATTGAGACTGTCCCTGTGAAGCTGAAGCCTGGCATGGATGGCCCCAAGGTGAA
 BgIII MetAlaProIleSerProIleGluThrValProValLysLeuLysProGlyMetAspGlyProLysValLys
 1 10 20

GCAGTGGCCCCTGACTGAGGAGAAGATCAAGGCCCTGCTGGAAATCTGCACTGAGATGGAGAAGGAGGGCAAATCTCCA
 sGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCysThrGluMetGluLysGluGlyLysIleSerL
 30 40 50

AGATTGGCCCCGAGAACCCCTACAACACCCCTGTGTTTGCCATCAAGAAGAAGGACTCCACCAAGTGGAGGAAGCTGGTG
 ysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSerThrLysTrpArgLysLeuVal
 60 70

GACTTCAGGGAGCTGAACAAGAGGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCTGGCCTGAAGAA
 AspPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGlyLeuLysLys
 80 90 100

GAAGAAGTCTGTGACTGTGCTGGCTGTGGGGGATGCCTACTTCTGTGCCCCCTGGATGAGGACTTCAGGAAGTACACTG
 sLysLysSerValThrValLeuAlaValGlyAspAlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrA
 110 120 130

CCTTCACCATCCCCTCCATCAACAATGAGACCCCTGGCATCAGGTACCAGTACAATGTGCTGCCCCAGGGCTGGAAGGGC
 loPheTrnIleProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrpLysGly
 140 150

TCCCCTGCCATCTTCCAGTCCCTCCATGACCAAGATCCTGGAGCCCTTCAGGAAGCAGAACCCTGACATTGTGATCTACCA
 SerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArgLysGlnAsnProAspIleValIleTyrGI
 160 170 180

GTACATGGCTGCCCTGTATGTGGGCTCTGACCTGGAGATTGGGCAGCACAGGACCAAGATTGAGGAGCTGAGGCAGCACC
 nTyrMetAlaAlaLeuTyrValGlySerAspLeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHisL
 190 200 210

TGCTGAGGTGGGGCCTGACCACCCCTGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTATGAGCTGCAC
 euLeuArgTrpGlyLeuThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyrGluLeuHis
 220 230

CCGACAAGTGGACTGTGCAGCCCATTTGTGCTGCCTGAGAAGGACTCCTGGACTGTGAATGACATCCAGAAGCTGGTGGG
 ProAspLysTrpThrValGlnProIleValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuValGI
 240 250 260

CAAGCTGAAGTGGGCTCCCAAATCTACCCTGGCATCAAGGTGAGGCAGCTGTGCAAGCTGCTGAGGGGCACCAAGGCC
 yLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLeuCysLysLeuLeuArgGlyThrLysAlaL
 270 280 290

FIGURE 17A

TGACTGAGGTGATCCCCCTGACTGAGGAGGCTGAGCTGGAGCTGGCTGAGAACAGGAGATCCTGAAGGAGCCTGTGCAT
 E^uThrGluVol^lIleProLeuThrGluGluAlaGluLeuGluLeuAlaGluAsnArgGlu^lIleLeuLysGluProVolHis
 300 310

GCGGTGTA^{CT}ATGACCCCTCCAAGGACCTGATTGCTGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAAATCTA
 GlyVol^lTyrTyrAspProSerLysAspLeu^lIleAlaGlu^lIleGlnLysGlnGlyGlnGlyGlnTrpThrTyrGln^lIleTy
 320 330 340

CCAGGAGCCCTTCAAGAACCTGAAGACTGGCAAGTATGCCAGGATGAGGGGGCCCCACACCAATGATGTGAAGCAGCTGA
 rGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThrAsnAspVol^lLysGlnLeu^T
 350 360 370

CTCAGGCTGTGCAGAAGATCACCCTGAGTCCATTGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAG
 hrGluAlaVol^lGlnLys^lIleThrThrGluSer^lIleVol^lIleTrpGlyLysThrProLysPheLysLeuPro^lIleGlnLys
 380 390

GAGACCTGGGAGACCTGGTGGACTGAGTACTGGCAGGCCACCTGGATCCCTGAGTGGGAGTTTGTGAACACCCCCCCCCT
 GluThrTrpGluThrTrpTrpThrGluTyrTrpGlnAlaThrTrp^lIleProGluTrpGluPheVol^lAsnThrProProLe
 400 410 420

GGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATTGTGGGGCTGAGACCTTCTATGTGGCTGGGGCTGCCAACAGGG
 uVol^lLysLeuTrpTyrGlnLeuGluLysGluPro^lIleVol^lGlyAlaGluThrPheTyrVol^lAlaGlyAlaAlaAsnArgG
 430 440 450

AGACCAAGCTGGGCAAGGCTGGCTATGTGACCAACAGGGGCAGGCAGAGGTGGTGACCCTGACTGACACCACCAACCAG
 luThrLysLeuGlyLysAlaGlyTyrVol^lThrAsnArgGlyArgGlnLysVol^lVol^lThrLeuThrAspThrThrAsnGln
 460 470

AAGACTGCCCTCCAGGCCATCTACCTGGCCCTCCAGGACTCTGGCCTGGAGGTGAACATTGTGACTGCCTCCAGTATGC
 LysThrAlaLeuGlnAla^lIleTyrLeuAlaLeuGlnAspSerGlyLeuGluVol^lAsn^lIleVol^lThrAlaSerGlnTyrAl
 480 490 500

CCTGGGCATCATCCAGGCCCAGCCTGATCAGTCTGAGTCTGAGCTGGTGAACCAGATCATTGAGCAGCTGATCAAGAAGG
 aLeuGly^lIle^lIleGlnAlaGlnProAspGlnSerGluSerGluLeuVol^lAsnGln^lIle^lIleGluGlnLeu^lIleLysLysG
 510 520 530

AGAAGGTGTACCTGGCCTGGGTGCCTGCCACAAGGGCATTGGGGGCAATGAGCAGGTGGACAAGCTGGTGTCTGCTGGC
 luLysVol^lTyrLeuAlaTrpVol^lProAlaHisLysGly^lIleGlyGlyAsnGluGlnVol^lAspLysLeuVol^lSerAlaGly
 540 550

ATCAGGAAGGTGCTGTTCTGATGGCATTGACAAGGCCCAGGATGAGCATGAGAAGTACCACTCCAACCTGGAGGGCTAT
^lIleArgLysVol^lLeuPheLeuAspGly^lIleAspLysAlaGlnAspGluHisGluLysTyrHisSerAsnTrpArgAlaMe
 560 570 580

FIGURE 17B

GGCTCTGACTTCAACCTGCCCCCTGTGGTGGCTAAGGAGATTGTGGCTCCTGTGACAAGTCCAGCTGAAGGGGGAGG
 tAlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluA
 590 600 610

CCATGCATGGGCAGGTGGACTGCTCCCCTGGCATCTGGCAGCTGGCCTGCACCCACCTGGAGGGCAAGGTGATCCTGGTG
 lAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGlnLeuAlaCysThrHisLeuGluGlyLysValIleLeuVal
 620 630

GCTGTGCATGTGGCTCCGGCTACATTGAGGCTGAGGTGATCCCTGCTGAGACAGGCCAGGAGACTGCCTACTTCCTGCT
 AlaValHisValAlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLe
 640 650 660

GAAGCTGGCTGGCAGGTGGCCTGTGAAGACCATCCACACTGCCAATGGCTCCAACCTCACTGGGGCCACAGTGAGGGCTG
 uLysLeuAlaGlyArgTrpProValLysThrIleHisThrAlaAsnGlySerAsnPheThrGlyAlaThrValArgAlaA
 670 680 690

CCTGCTGGTGGGCTGGCATCAAGCAGGAGTTTGGCATCCCCTACAACCCCCAGTCCCAGGGGGTGGTGGCCTCCATGAAC
 lAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValAlaSerMetAsn
 700 710

AAGGAGCTGAAGAAGATCATTGGCCAGGTGAGGGACCAGGCTGAGCACCTGAAGACAGCTGTGCAGATGGCTGTGTTTAT
 LysGluLeuLysLysIleIleGlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIle
 720 730 740

CCACAACTTCAAGAGGAAGGGGGGCATCGGGGGCTACTCCGCTGGGGAGAGGATTGTGGACATCATTGCCACAGACATCC
 eHisAsnPheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArgIleValAspIleIleAlaThrAspIleG
 750 760 770

AGACCAAGGAGCTCCAGAAGCAGATCACCAAGATCCAGAACTTCAGGGTGTACTACAGGGACTCCAGGAACCCCTGTGG
 lnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArgAspSerArgAsnProLeuTrp
 780 790

AAGGGCCCTGCCAAGCTGCTGTGGAAGGGGGAGGGGGCTGTGGTGATCCAGGACAACTCTGACATCAAGGTGGTGGCCAG
 LysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAlaValValIleGlnAspAsnSerAspIleLysValValProAr
 800 810 820

GAGGAAGGCCAAGATCATCAGGGACTATGGCAAGCAGATGGCTGGGGATGACTGTGTGGCCTCCAGGCAGGATGAGGACT
 gArgLysAlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAlaSerArgGlnAspGluAspx
 830 840 850

AAAGCCCGGGCAGATC" (SEQ ID NO: 3)
 Xx BgΠI (SEQ ID NO: 4)

FIGURE 17C

FIGURE 18

| | | |
|-----|---|------|
| WT | - ATG GGT GGC AAG TGG TCA AAA CGT AGT GTG CCT GGA TGG TCT | -42 |
| | | |
| OPT | - ATG GGC GGC AAG TGG TCC AAG AGG TCC GTG CCC GGC TGG TCC | |
| | M G G K W S K R S V P G W S | -14 |
| WT | - ACT GTA AGG GAA AGA ATG AGA CGA GCT GAG CCA GCA GCA GAT | -84 |
| | | |
| OPT | - ACC GTG AGG GAG AGG ATG AGG AGG GCC GAG CCC GCC GCC GAC | |
| | T V R E R M R R A E P A A D | -28 |
| WT | - AGG GTG AGA CGA ACT GAG CCA GCA GCA GTA GGG GTG GGA GCA | -126 |
| | | |
| OPT | - AGG GTG AGG AGG ACC GAG CCC GCC GCC GTG GGC GTG GGC GCC | |
| | R V R R T E P A A V G V G A | -42 |
| WT | - GTA TCT CGA GAC CTG GAA AAA CAT GGA GCA ATC ACA AGT AGC | -168 |
| | | |
| OPT | - GTG TCC AGG GAC CTG GAG AAG CAC GGC GCC ATC ACC TCC TCC | |
| | V S R D L E K H G A I T S S | -56 |
| WT | - AAT ACA GCA GCT ACC AAT GCT GAT TGT GCC TGG CTA GAA GCA | -210 |
| | | |
| OPT | - AAC ACC GCC GCC ACC AAC GCC GAC TGC GCC TGG CTG GAG GCC | |
| | N T A A T N A D C A W L E A | -70 |
| WT | - CAA GAG GAT GAG GAA GTG GGT TTT CCA GTC AGA CCT CAG GTA | -252 |
| | | |
| OPT | - CAG GAG GAC GAG GAG GTG GGC TTC CCC GTG AGG CCC CAG GTG | |
| | Q E D E E V G F P V R P Q V | -84 |
| WT | - CCT TTA AGA CCA ATG ACT TAC AAG GGA GCT GTA GAT CTT AGC | -294 |
| | | |
| OPT | - CCC CTG AGG CCC ATG ACC TAC AAG GGC GCC GTG GAC CTG TCC | |
| | P L R P M T Y K G A V D L S | -98 |
| WT | - CAC TTT TTA AAA GAA AAG GGG GGA CTG GAA GGG CTA ATT CAC | -336 |
| | | |
| OPT | - CAC TTC CTG AAG GAG AAG GGC GGC CTG GAG GGC CTG ATC CAC | |
| | H F L K E K G G L E G L I H | -112 |
| WT | - TCA CAG AAA AGA CAA GAT ATC CTT GAT CTG TGG GTC TAC CAC | -378 |
| | | |
| OPT | - TCC CAG AAG AGG CAG GAC ATC CTG GAC CTG TGG GTG TAC CAC | |
| | S Q K R Q D I L D L W V Y H | -126 |
| WT | - ACA CAA GGC TAC TTC CCT GAT TGG CAG AAC TAC ACA CCA GGG | -420 |
| | | |
| OPT | - ACC CAG GGC TAC TTC CCC GAC TGG CAG AAC TAC ACC CCC GGC | |
| | T Q G Y F P D W Q N Y T P G | -140 |

FIGURE 19A

| | | |
|-----|---|------|
| WT | - CCA GGA ATC AGA TTT CCA TTG ACC TTT GGA TGG TGC TTC AAG | -462 |
| | | |
| OPT | - CCC GGC ATC AGG TTC CCC CTG ACC TTC GGC TGG TGC TTC AAG | |
| | P G I R F P L T F G W C F K | -154 |
| WT | - CTA GTA CCA GTT GAG CCA GAA AAG GTA GAA GAG GCC AAT GAA | -504 |
| | | |
| OPT | - CTG GTG CCC GTG GAG CCC GAG AAG GTG GAG GAG GCC AAC GAG | |
| | L V P V E P E K V E E A N E | -168 |
| WT | - GGA GAG AAC AAC TGC TTG TTA CAC CCT ATG AGC CAG CAT GGG | -546 |
| | | |
| OPT | - GGC GAG AAC AAC TGC CTG CTG CAC CCC ATG TCC CAG CAC GGC | |
| | G E N N C L L H P M S Q H G | -182 |
| WT | - ATA GAG GAC CCG GAG AAG GAA GTG TTA GAG TGG AGG TTT GAC | -588 |
| | | |
| OPT | - ATC GAG GAC CCC GAG AAG GAG GTG CTG GAG TGG AGG TTC GAC | |
| | I E D P E K E V L E W R F D | -196 |
| WT | - AGC AAG CTA GCA TTT CAT CAC GTG GCC CGA GAG CTG CAT CCG | -630 |
| | | |
| OPT | - TCC AAG CTG GCC TTC CAC CAC GTG GCC AGG GAG CTG CAC CCC | |
| | S K L A F H H V A R E L H P | -210 |
| WT | - GAG TAC TAC AAG GAC TGC TGA (SEQ ID NO:30) | -651 |
| | | |
| OPT | - GAG TAC TAC AAG GAC TGC TAA (contained within SEQ ID NO: 9) | |
| | E Y Y K D C (SEQ ID NO: 10) | -216 |

FIGURE 19B

V1Jns/nef *PstI* *BglII*
CATGGGTCCTTTCTGCAGTCACCGTCCTTGAGATCTGCCACC ATG GGC GGC AAG TGG TCC AAG AGG TCC GTG CCC
M G G K W S K R S V P

. CAC CCC GAG TAC TAC AAG GAC TGC TAA *SrfI* *BglII*
AGCCCGGGCAGATCTGCTGTGCTTCTAGTTGCCAGC (SEQ ID NO: 38)
H P E Y Y K D C * (contained within SEQ ID NO: 10)

V1Jns/nef(G2A.LLAA)

PstI *BglII*
CATGGGTCCTTTCTGCAGTCACCGTCCTTGAGATCTGCCACC ATG GCC GGC AAG TGG TCC AAG AGG TCC GTG CCC
M A G K W S K R S V P

. CAC CCC GAG TAC TAC AAG GAC TGC TAA *SrfI* *BglII*
AGCCCGGGCAGATCTGCTGTGCTTCTAGTTGCCAGC (SEQ ID NO: 39)
H P E Y Y K D C * (contained within SEQ ID NO: 14)

V1Jns/tpanef & V1Jns/tpanef(LLAA)

PstI
CATGGGTCCTTTCTGCAGTCACCGTCCTTATATCTAGATCACC ATG GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG
M D A M K R G L C C V

CTG CTG CTG TGT GGA GCA GTC TTC GTT TCG CCC AGC GAG *BglII* ATCTCC TCC AAG AGG TCC GTG CCC
L L L C G A V F V S S S K R S V P

. CAC CCC GAG TAC TAC AAG GAC TGC TAA *SrfI* *BglII*
AGCCCGGGCAGATCTGCTGTGCTTCTAGTTGCCAGC (SEQ ID NO: 40)
H P E Y Y K D C * (contained within SEQ ID NO: 16)

FIGURE 20

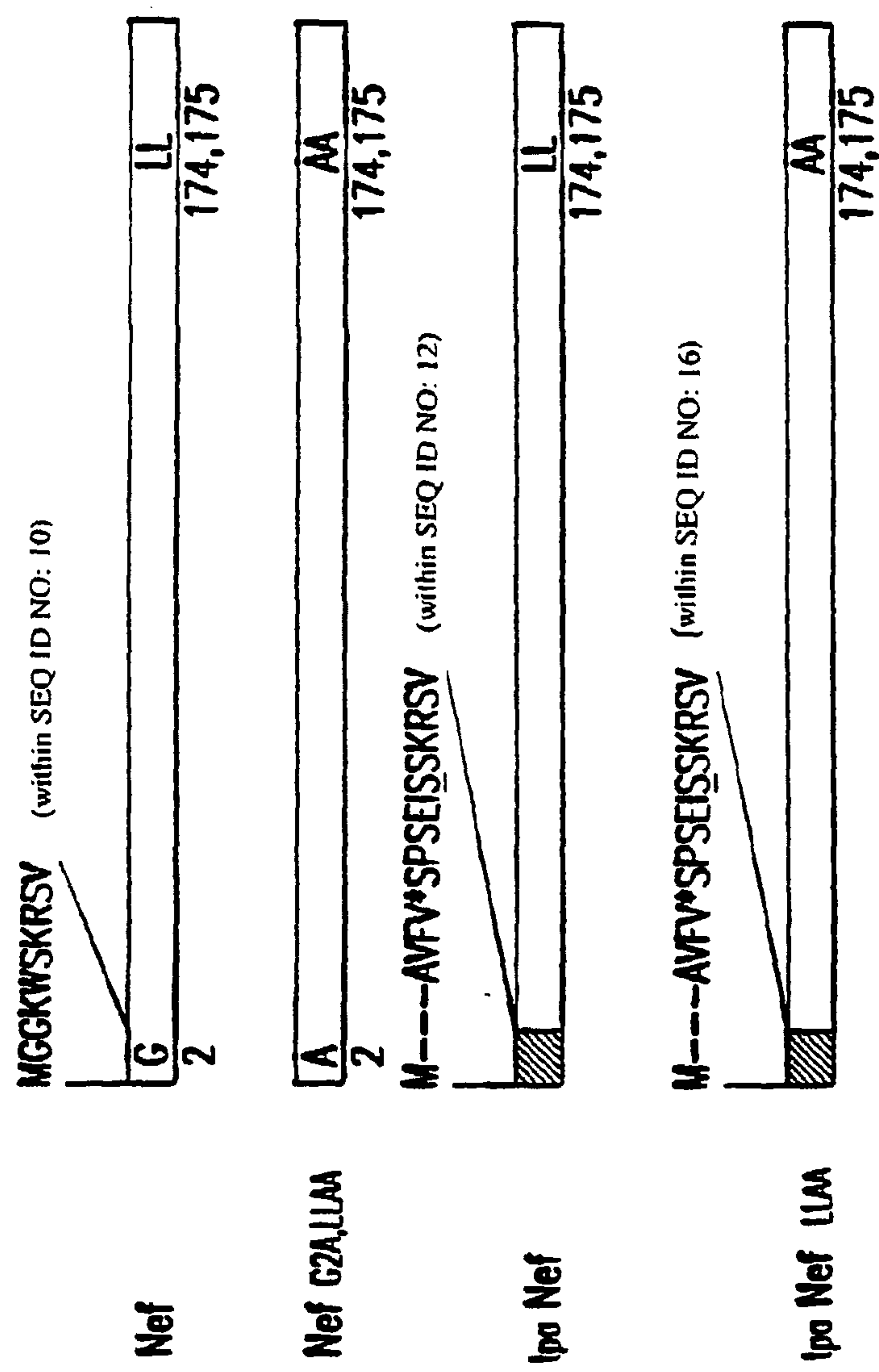


FIGURE 21

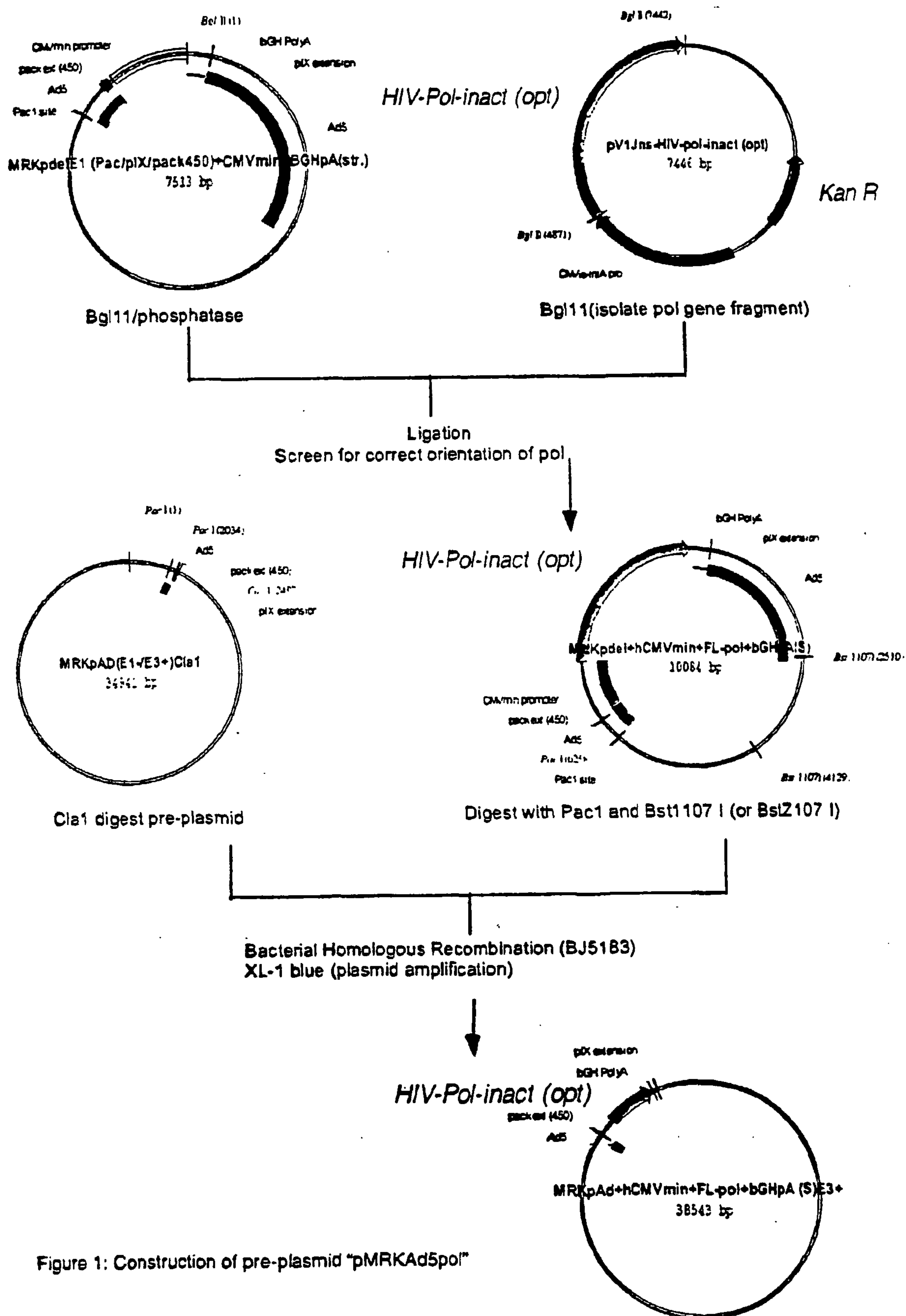


Figure 1: Construction of pre-plasmid "pMRKAd5pol"

FIGURE 22

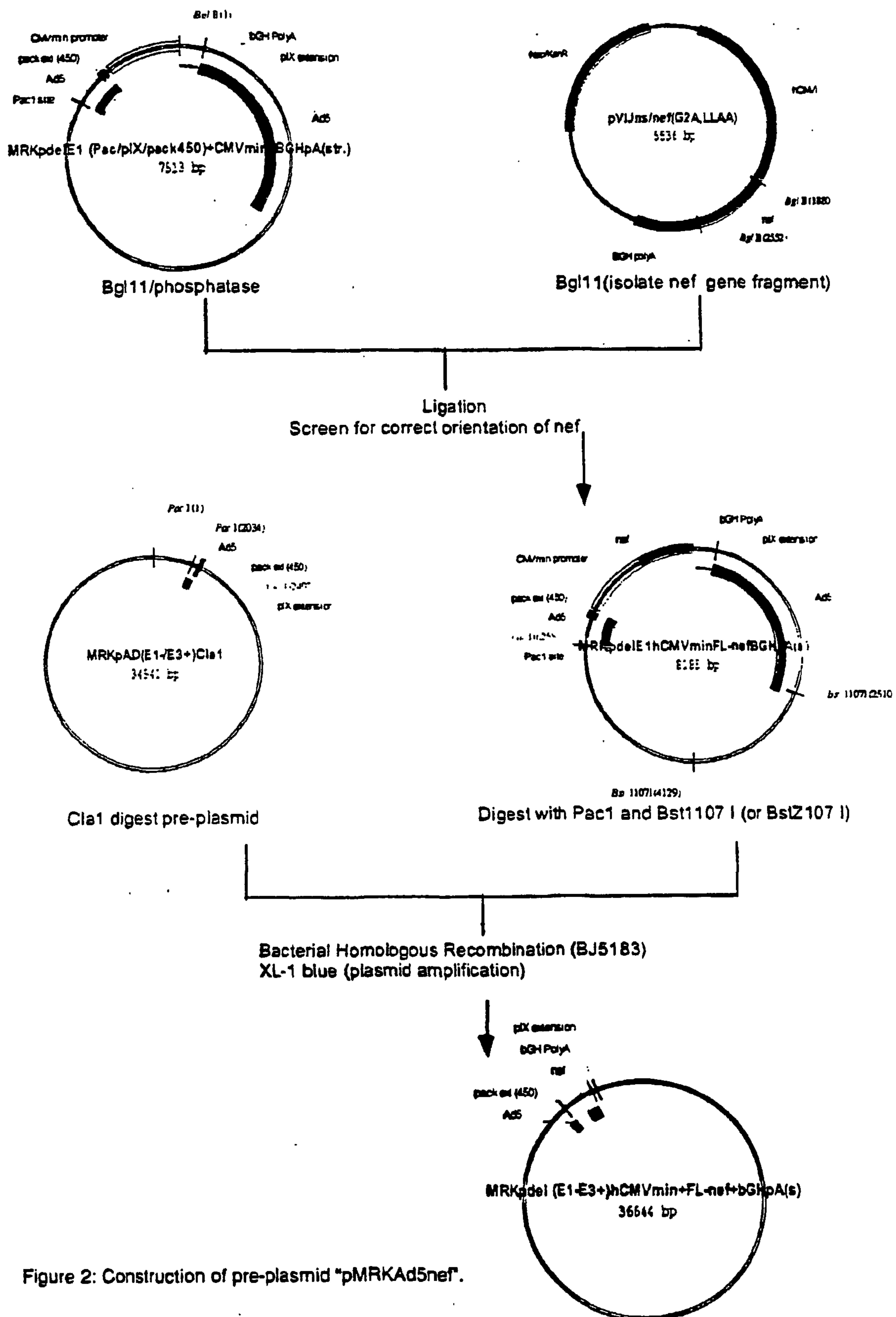
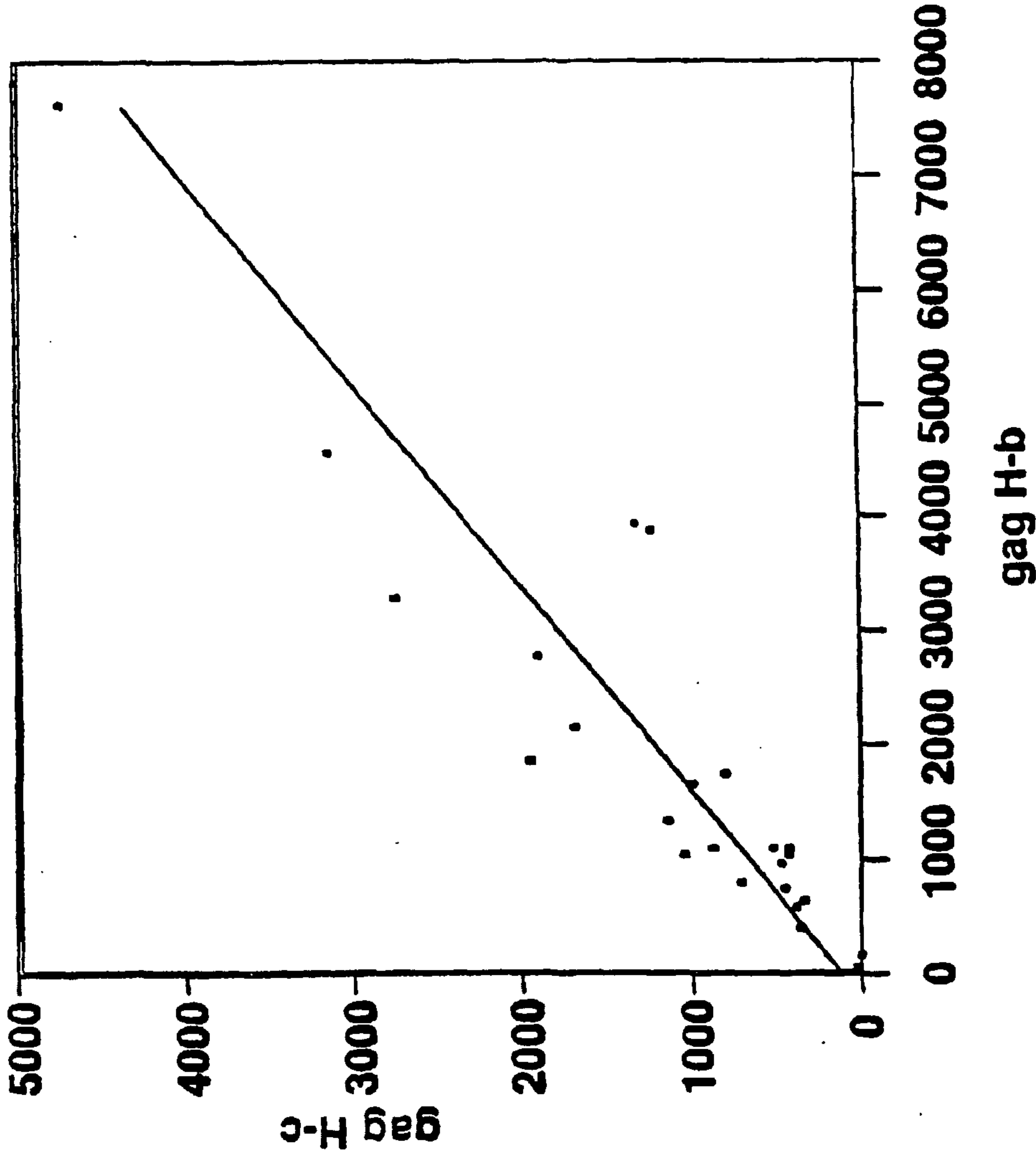


Figure 2: Construction of pre-plasmid "pMRKAd5nef".

Comparison of Clade B vs. Clade C Anti-gag T Cell Responses in Clade B HIV-Infected Subjects



Linear Fit

$$\text{gag H-c} = 111.603 + 0.55866 \text{ gag H-b}$$

Summary of Fit

| | |
|----------------------------|----------|
| RSquare | 0.816775 |
| RSquare Adj | 0.80914 |
| Root Mean Square Error | 474.9639 |
| Mean of Response | 1158.115 |
| Observations (or Sum Wgts) | 26 |

Comparison of Clade B vs. Clade C Anti-nef T Cell Responses in Clade B HIV-Infected Subjects

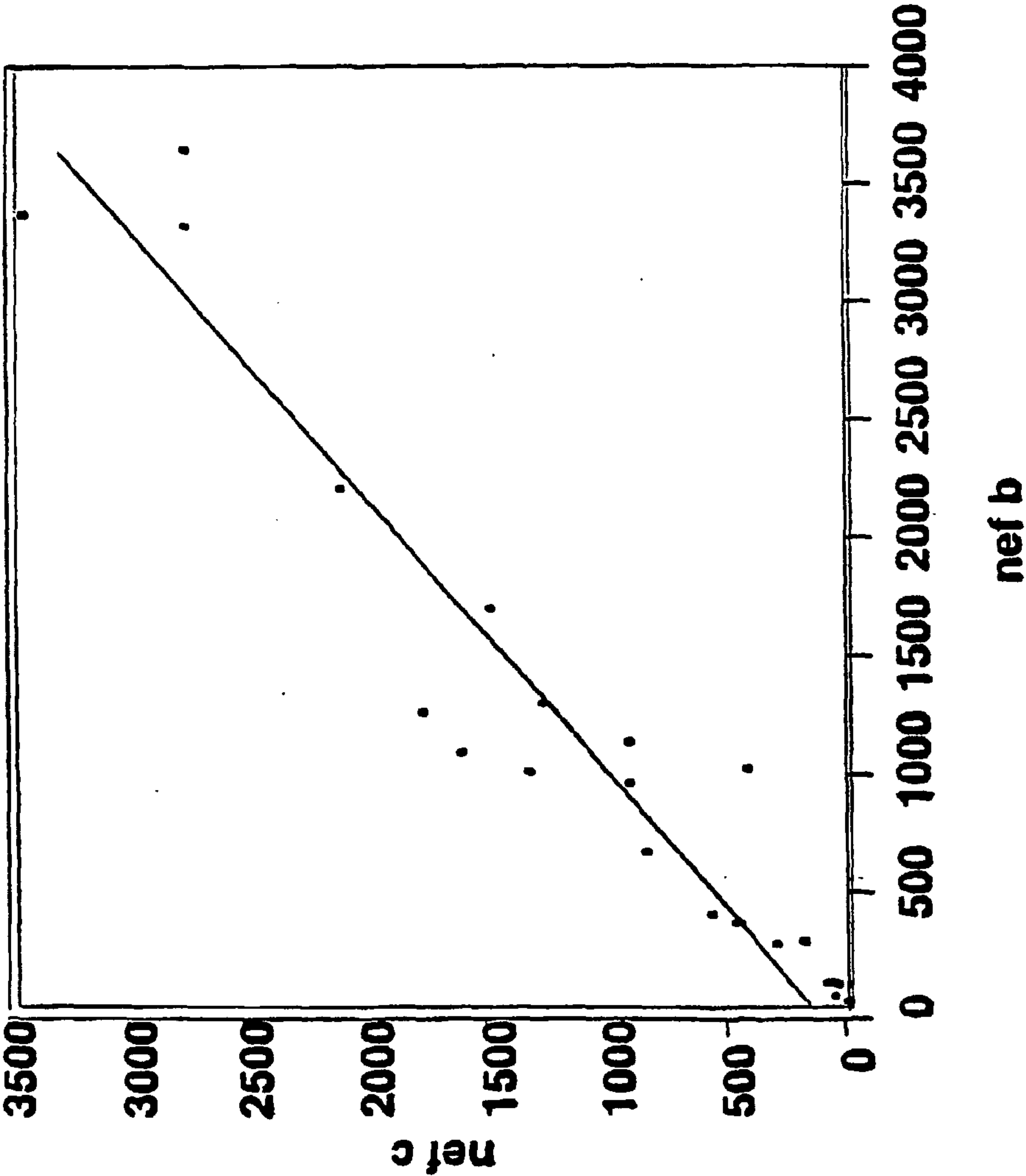


FIGURE 25

MRKAd5pol MER1062
(MRKAd5 Pre-Adenoviral Vector Containing the LA opt pol Coding Region)

```

1  CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG
   GTAGTAGTTA TTATATGGAA TAAAACCTAA CTCGGTTAT ACTATTACTC

51  GGGGTGGAGT TTGTGACGTG GCGCGGGGCG TGGGAACGGG GCGGGTGACG
   CCCACCTCA AACACTGCAC CGCGCCCCGC ACCCTTGCCC CGCCCACTGC

101 TAGTAGTGTG GCGGAAGTGT GATGTTGCAA GTGTGGCGGA ACACATGTAA
   ATCATCACAC CGCCTTCACA CTACAACGTT CACACCGCCT TGTGTACATT

151 GCGACGGATG TGGCAAAAGT GACGTTTTTG GTGTGCGCCG GTGTACACAG
   CGCTGCCTAC ACCGTTTTCA CTGCAAAAC CACACGCGGC CACATGTGTC

201 GAAGTGACAA TTTTCGCGCG GTTTTAGGCG GATGTTGTAG TAAATTTGGG
   CTTCACTGTT AAAAGCGCGC CAAAATCCGC CTACAACATC ATTTAAACCC

251 CGTAACCGAG TAAGATTTGG CCATTTTCGC GGGAAACTG AATAAGAGGA
   GCATTGGCTC ATTCTAAACC GGTAAAGCG CCCTTTTGAC TTATTCTCCT

301 AGTGAAATCT GAATAATTTT GTGTTACTCA TAGCGCGTAA TATTGTCTA
   TCACTTTAGA CTTATTAAAA CACAATGAGT ATCGCGCATT ATAAACAGAT

351 GGGCCGCGGG GACTTTGACC GTTTACGTGG AGACTCGCCC AGGTGTTTTT
   CCCGGCGCCC CTGAAACTGG CAAATGCACC TCTGAGCGGG TCCACAAAAA

401 CTCAGGTGTT TTCCGCGTTC CGGGTCAAAG TTGGCGTTTT ATTATTATAG
   GAGTCCACAA AAGGCGCAAG GCCCAGTTTC AACCGCAAAA TAATAATATC

451 GCGGCCGCGA TCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT
   CGCCGGCGCT AGSTAACGTA TGCAACATAG GTATAGTATT ATACATGTAA

501 TATATTGGCT CATGTCCAAC ATTACCGCCA TGTTGACATT GATTATTGAC
   ATATAACCGA GTACAGGTTG TAATGGCGGT ACAACTGTAA CTAATAACTG

551 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
   ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT

601 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
   ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC

651 CCCAACGACC CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
   GGGTTGCTGG GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA

701 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT
   TTGCGGTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

751 AAAGTGGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC
   TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG

801 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA
   GGATAACTGC AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT

851 CATGACCTTA TGGGACTTTC CTAATTGGCA GTACATCTAC GTATTAGTCA
   GTACTGGAAT ACCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT

```

Figure 26A

901 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT

951 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT

1001 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA
ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT

1051 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
TGTTGAGGCG GGGTAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC

1101 GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1151 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC
GGTAGGTGCG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG

1201 TCCGCGGCCG GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT
AGGCGCCGGC CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCCTCA

1251 GAGATCTACC ATGGCCCCCA TCTCCCCCAT TGAGACTGTG CCTGTGAAGC
CTCTAGATGG TACCGGGGGT AGAGGGGGTA ACTCTGACAC GGACACTTCG

1301 TGAAGCCTGG CATGGATGGC CCCAAGGTGA AGCAGTGGCC CCTGACTGAG
ACTTCGGACC GTACCTACCG GGGTTCCACT TCGTCACCGG GGACTGACTC

1351 GAGAAGATCA AGGCCCTGGT GGAAATCTGC ACTGAGATGG AGAAGGAGGG
CTCTTCTAGT TCCGGGACCA CCTTTAGACG TGACTCTACC TCTTCCTCCC

1401 CAAAATCTCC AAGATTGGCC CCGAGAACCC CTACAACACC CCTGTGTTTG
GTTTTAGAGG TTCTAACCGG GGCTCTTGGG GATGTTGTGG GGACACAAAC

1451 CCATCAAGAA GAAGGACTCC ACCAAGTGGA GGAAGCTGGT GGACTTCAGG
GGTAGTTCTT CTTCTGAGG TGGTTCACCT CCTTCGACCA CCTGAAGTCC

1501 GAGCTGAACA AGAGGACCCA GGACTTCTGG GAGGTGCAGC TGGGCATCCC
CTCGACTTGT TCTCCTGGGT CCTGAAGACC CTCCACGTGC ACCCGTAGGG

1551 CCACCCCGCT GGCCTGAAGA AGAAGAAGTC TGTGACTGTG CTGGCTGTGG
GGTGGGGCGA CCGGACTTCT TCTTCTTCAG ACACTGACAC GACCGACACC

1601 GGGATGCCTA CTTCTCTGTG CCCCTGGATG AGGACTTCAG GAAGTACACT
CCCTACGGAT GAAGAGACAC GGGGACCTAC TCCTGAAGTC CTTTCATGTGA

1651 GCCTTCACCA TCCCCTCCAT CAACAATGAG ACCCCTGGCA TCAGGTACCA
CGGAAGTGGT AGGGGAGGTA GTTGTTACTC TGGGGACCGT AGTCCATGGT

1701 GTACAATGTG CTGCCCCAGG GCTGGAAGGG CTCCCCTGCC ATCTTCCAGT
CATGTTACAC GACGGGGTCC CGACCTTCCC GAGGGGACGG TAGAAGGTCA

1751 CCTCCATGAC CAAGATCCTG GAGCCCTTCA GGAAGCAGAA CCCTGACATT
GGAGGTACTG GTTCTAGGAC CTCGGGAAGT CCTTCGTCTT GGGACTGTAA

1801 GTGATCTACC AGTACATGGC TGCCCTGTAT GTGGGCTCTG ACCTGGAGAT
CACTAGATGG TCATGTACCG ACGGGACATA CACCCGAGAC TGGACCTCTA

Figure 24B

1901 GGGGCCTGAC CACCCCTGAC AAGAAGCACC AGAAGGAGCC CCCCTTCCTG
CCCCGGACTG GTGGGGACTG TTCTTCGTGG TCTTCCTCGG GGGGAAGGAC

1951 TGGATGGGCT ATGAGCTGCA CCCCGACAAG TGGACTGTGC AGCCCATTTGT
ACCTACCCGA TACTCGACGT GGGGCTGTTC ACCTGACACG TCGGGTAACA

2001 GCTGCCTGAG AAGGACTCCT GGACTGTGAA TGACATCCAG AAGCTGGTGG
CGACGGACTC TTCCTGAGGA CCTGACACTT ACTGTAGGTC TTCGACCACC

2051 GCAAGCTGAA CTGGGCCTCC CAAATCTACC CTGGCATCAA GGTGAGGCAG
CGTTCGACTT GACCCGGAGG GTTTAGATGG GACCGTAGTT CCACTCCGTC

2101 CTGTGCAAGC TGCTGAGGGG CACCAAGGCC CTGACTGAGG TGATCCCCCT
GACACGTTCG ACGACTCCCC GTGGTTCCGG GACTGACTCC ACTAGGGGGA

2151 GACTGAGGAG GCTGAGCTGG AGCTGGCTGA GAACAGGGAG ATCCTGAAGG
CTGACTCCTC CGACTCGACC TCGACCGACT CTTGTCCCTC TAGGACTTCC

2201 AGCCTGTGCA TGGGGTGTAC TATGACCCCT CCAAGGACCT GATTGCTGAG
TCGGACACGT ACCCCACATG ATACTGGGGA GGTTCCTGGA CTAACGACTC

2251 ATCCAGAAGC AGGGCCAGGG CCAGTGGACC TACCAAATCT ACCAGGAGCC
TAGGTCTTCG TCCCGGTCCC GGTCACCTGG ATGGTTTAGA TGGTCCTCGG

2301 CTTCAAGAAC CTGAAGACTG GCAAGTATGC CAGGATGAGG GGGGCCCACA
GAAGTTCTTG GACTTCTGAC CGTTCATACG GTCCTACTCC CCCCGGGTGT

2351 CCAATGATGT GAAGCAGCTG ACTGAGGCTG TGCAGAAGAT CACCACTGAG
GGTTACTACA CTTCGTCGAC TGACTCCGAC ACGTCTTCTA GTGGTGACTC

2401 TCCATTGTGA TCTGGGGCAA GACCCCCAAG TTCAAGCTGC CCATCCAGAA
AGGTAACACT AGACCCCGTT CTGGGGGGTTC AAGTTCGACG GGTAGGTCTT

2451 GGAGACCTGG GAGACCTGGT GGACTGAGTA CTGGCAGGCC ACCTGGATCC
CCTCTGGACC CTCTGGACCA CCTGACTCAT GACCGTCCGG TGGACCTAGG

2501 CTGAGTGGGA GTTTGTGAAC ACCCCCCCCC TGGTGAAGCT GTGGTACCAG
GACTCACCTT CAAACACTTG TGGGGGGGGG ACCACTTCGA CACCATGGTC

2551 CTGGAGAAGG AGCCCATTTGT GGGGGCTGAG ACCTTCTATG TGGCTGGGGC
GACCTCTTCC TCGGGTAACA CCCCCGACTC TGGAAGATAC ACCGACCCCG

2601 TGCCAACAGG GAGACCAAGC TGGGCAAGGC TGGCTATGTG ACCAACAGGG
ACGGTTGTCC CTCTGGTTCG ACCCGTTCGG ACCGATACAC TGGTTGTCCC

2651 GCAGGCAGAA GGTGGTGACC CTGACTGACA CCACCAACCA GAAGACTGCC
CGTCCGTCTT CCACCACTGG GACTGACTGT GGTGGTTGGT CTTCTGACGG

2701 CTCCAGGCCA TCTACCTGGC CCTCCAGGAC TCTGGCCTGG AGGTGAACAT
GAGGTCCGGT AGATGGACCG GGAGGTCCTG AGACCGGACC TCCACTTGTA

2751 TGTGACTGCC TCCCAGTATG CCCTGGGCAT CATCCAGGCC CAGCCTGATC
ACACTGACGG AGGGTCATAC GGGACCCGTA GTAGGTCCGG GTCGGACTAG

Figure 26 C

2851 GAGAAGGTGT ACCTGGCCTG GGTGCCTGCC CACAAGGGCA TTGGGGGCAA
CTCTTCCACA TGGACCGGAC CCACGGACGG GTGTTCCCGT AACCCCGTT

2901 TGAGCAGGTG GACAAGCTGG TGTCTGCTGG CATCAGGAAG GTGCTGTTCC
ACTCGTCCAC CTGTTGACC ACAGACGACC GTAGTCCTTC CACGACAAGG

2951 TGGATGGCAT TGACAAGGCC CAGGATGAGC ATGAGAAGTA CCACTCCAAC
ACCTACCGTA ACTGTTCCGG GTCCTACTCG TACTCTTCAT GGTGAGGTTG

3001 TGGAGGGCTA TGGCCTCTGA CTTCAACCTG CCCCTGTGG TGGCTAAGGA
ACCTCCCGAT ACCGGAGACT GAAGTTGGAC GGGGGACACC ACCGATTCTT

3051 GATTGTGGCC TCCTGTGACA AGTGCCAGCT GAAGGGGGAG GCCATGCATG
CTAACACCGG AGGACACTGT TCACGGTCGA CTTCCCCCTC CGGTACGTAC

3101 GGCAGGTGGA CTGCTCCCCT GGCATCTGGC AGCTGGCCTG CACCCACCTG
CCGTCCACCT GACGAGGGGA CCGTAGACCG TCGACCGGAC GTGGGTGGAC

3151 GAGGGCAAGG TGATCCTGGT GGCTGTGCAT GTGGCCTCCG GCTACATTGA
CTCCCGTTCC ACTAGGACCA CCGACACGTA CACCGGAGGC CGATGTAAC

3201 GGCTGAGGTG ATCCCTGCTG AGACAGGCCA GGAGACTGCC TACTTCCTGC
CCGACTCCAC TAGGGACGAC TCTGTCCGGT CCTCTGACGG ATGAAGGACG

3251 TGAAGCTGGC TGGCAGGTGG CCTGTGAAGA CCATCCACAC TGCCAATGGC
ACTTCGACCG ACCGTCCACC GGACACTTCT GGTAGGTGTG ACGGTTACCG

3301 TCCAACCTCA CTGGGGCCAC AGTGAGGGCT GCCTGCTGGT GGGCTGGCAT
AGGTTGAAGT GACCCCGGTG TCACTCCCGA CGGACGACCA CCCGACCGTA

3351 CAAGCAGGAG TTTGGCATCC CCTACAACCC CCAGTCCCAG GGGGTGGTGG
GTTTCGTCTC AAACCGTAGG GGATGTTGGG GGTGAGGGTC CCCCACCACC

3401 CCTCCATGAA CAAGGAGCTG AAGAAGATCA TTGGGCAGGT GAGGGACCAG
GGAGGTACTT GTTCCTCGAC TTCTTCTAGT AACCCGTCCA CTCCCTGGTC

3451 GCTGAGCACC TGAAGACAGC TGTGCAGATG GCTGTGTTCA TCCACAACCT
CGACTCGTGG ACTTCTGTG ACACGTCTAC CGACACAAGT AGGTGTTGAA

3501 CAAGAGGAAG GGGGGCATCG GGGGCTACTC CGCTGGGGAG AGGATTGTGG
GTTCTCCTTC CCCCCGTAGC CCCCAGATGAG GCGACCCCTC TCCTAACACC

3551 ACATCATTGC CACAGACATC CAGACCAAGG AGCTCCAGAA GCAGATCACC
TGTAGTAACG GTGTCTGTAG GTCTGGTTCC TCGAGGTCTT CGTCTAGTGG

3601 AAGATCCAGA ACTTCAGGGT GTACTACAGG GACTCCAGGA ACCCCCTGTG
TTCTAGGTCT TGAAGTCCCA CATGATGTCC CTGAGGTCCT TGGGGGACAC

3651 GAAGGGCCCT GCCAAGCTGC TGTGGAAGGG GGAGGGGGCT GTGGTGATCC
CTTCCCGGGA CGGTTCGACG ACACCTTCCC CCTCCCCCGA CACCACTAGG

3701 AGGACAACCTC TGACATCAAG GTGGTGCCCA GGAGGAAGGC CAAGATCATC
TCCTGTTGAG ACTGTAGTTC CACCACGGGT CCTCCTTCCG GTTCTAGTAG

7 June 26 D

3801 GGATGAGGAC TAAAGCCCGG GCAGATCTGC TGTGCCTTCT AGTTGCCAGC
CCTACTCCTG ATTTCTGGGCC CGTCTAGACG ACACGGAAGA TCAACGGTCG

3851 CATCTGTTGT TTGCCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC
GTAGACAACA AACGGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACGG

3901 ACTCCCCTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT CGCATTGTCT
TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA GCGTAACAGA

3951 GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG
CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC

4001 GGGAGGATTG GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT
CCCTCCTAAC CCTTCTGTTA TCGTCCGTAC GACCCCTACG CCACCCGAGA

4051 ATGGCCGATC GCGCGCCCGT ACTGAAATGT GTGGGCGTGG CTTAAGGGTG
TACCGGCTAG CCGCGCGGCA TGACTTTACA CACCCGCACC GAATTCCCAC

4101 GGAAAGAATA TATAAGGTGG GGGTCTTATG TAGTTTTGTA TCTGTTTTGC
CCTTTCTTAT ATATTCCACC CCCAGAATAC ATCAAAACAT AGACAAAACG

4151 AGCAGCCGCC GCCGCCATGA GCACCAACTC GTTTGATGGA AGCATTGTGA
TCGTCGGCGG CCGCGGTACT CGTGGTTGAG CAAACTACCT TCGTAACACT

4201 GCTCATATTT GACAACGCGC ATGCCCCCAT GGGCCGGGGT GCGTCAGAAT
CGAGTATAAA CTGTTGCGCG TACGGGGGTA CCCGGCCCCA CGCAGTCTTA

4251 GTGATGGGCT CCAGCATTGA TGGTCGCCCC GTCCTGCCCC CAAACTCTAC
CACTACCCGA GGTCGTAACCT ACCAGCGGGG CAGGACGGGC GTTTGAGATG

4301 TACCTTGACC TACGAGACCG TGTCTGGAAC GCCGTTGGAG ACTGCAGCCT
ATGGAAGTGG ATGCTCTGGC ACAGACCTTG CGGCAACCTC TGACGTCGGA

4351 CCGCCGCCGC TTCAGCCGCT GCAGCCACCG CCCGCGGGAT TGTGACTGAC
GGCGGCGGCG AAGTCGGCGA CGTCGGTGGC GGGCGCCCTA AACTGACTG

4401 TTTGCTTTCC TGAGCCCGCT TGCAAACAGT GCAGCTTCCC GTTCATCCGC
AAACGAAAGG ACTCGGGCGA ACGTTTGTCA CGTCGAAGGG CAAGTAGGCG

4451 CCGCGATGAC AAGTTGACGG CTCTTTTGGC ACAATTGGAT TCTTTGACCC
GGCGCTACTG TTCAACTGCC GAGAAAACCG TGTTAACCTA AGAAACTGGG

4501 GGGAAGTTAA TGTCGTTTCT CAGCAGCTGT TGGATCTGCG CCAGCAGGTT
CCCTTGAATT ACAGCAAAGA GTCGTCGACA ACCTAGACGC GGTCGTCCAA

4551 TCTGCCCTGA AGGCTTCCTC CCCTCCCAAT GCGGTTTAAA ACATAAATAA
AGACGGGACT TCCGAAGGAG GGGAGGGTTA CGCCAAATTT TGTATTTATT

4601 AAAACCAGAC TCTGTTTGGG TTTGGATCAA GCAAGTGTCT TGCTGTCTTT
TTTTGGTCTG AGACAAACCT AAACCTAGTT CGTTCACAGA ACGACAGAAA

4651 ATTTAGGGGT TTGCGCGCGC CGGTAGGCCC GGGACCAGCG GTCTCGGTGC
TAAATCCCCA AAACGCGCGC GCCATCCGGG CCCTGGTCGC CAGAGCCAGC

Figure 26E

4751 GTTCAGATAC ATGGGCATAA GCCCGTCTCT GGGGTGGAGG TAGCACCACCT
CAAGTCTATG TACCCGTATT CGGGCAGAGA CCCACCTCC ATCGTGGTGA

4801 GCAGAGCTTC ATGCTGCGGG GTGGTGTGT AGATGATCCA GTCGTAGCAG
CGTCTCGAAG TACGACGCCC CACCACAACA TCTACTAGGT CAGCATCGTC

4851 GAGCGCTGGG CGTGGTGCCT AAAAATGTCT TTCAGTAGCA AGCTGATTGC
CTCGCGACCC GCACCACGGA TTTTACAGA AAGTCATCGT TCGACTAACG

4901 CAGGGGCAGG CCCTTGGTGT AAGTGTTTAC AAAGCGGTTA AGCTGGGATG
GTCCCCGTCC GGAACCACA TTCACAAATG TTTCGCCAAT TCGACCCTAC

4951 GGTGCATACG TGGGGATATG AGATGCATCT TGGACTGTAT TTTTAGGTTG
CCACGTATGC ACCCCTATAC TCTACGTAGA ACCTGACATA AAAATCCAAC

5001 GCTATGTTCC CAGCCATATC CCTCCGGGGA TTCATGTTGT GCAGAACCAC
CGATACAAGG GTCGGTATAG GGAGGCCCT AAGTACAACA CGTCTTGGTG

5051 CAGCACAGTG TATCCGGTGC ACTTGGGAAA TTTGTCTATGT AGCTTAGAAG
GTCGTGTCAC ATAGGCCACG TGAACCCTTT AACAGTACA TCGAATCTTC

5101 GAAATGCGTG GAAGAACTTG GAGACGCCCT TGTGACCTCC AAGATTTTCC
CTTTACGCAC CTTCTTGAAC CTCTGCGGGA AACTGGAGG TTCTAAAAGG

5151 ATGCATTTCGT CCATAATGAT GGCAATGGGC CCACGGGCGG CGGCCTGGGC
TACGTAAGCA GGTATTACTA CCGTTACCCG GGTGCCCCGCC GCCGGACCCG

5201 GAAGATATTT CTGGGATCAC TAACGTCATA GTTGTGTTCC AGGATGAGAT
CTTCTATAAA GACCCTAGTG ATTGCAGTAT CAACACAAGG TCCTACTCTA

5251 CGTCATAGGC CATTTTACACA AAGCGCGGGC GGAGGGTGCC AGACTGCGGT
GCAGTATCCG GTAAAAATGT TTCGCGCCCG CCTCCCACGG TCTGACGCCA

5301 ATAATGGTTC CATCCGGCCC AGGGGCGTAG TTACCCTCAC AGATTTGCAT
TATTACCAAG GTAGGCCGGG TCCCCGCATC AATGGGAGTG TCTAAACGTA

5351 TTCCCACGCT TTGAGTTCAG ATGGGGGGAT CATGTCTACC TGCGGGGCGA
AAGGGTGCGA AACTCAAGTC TACCCCCCTA GTACAGATGG ACGCCCCGCT

5401 TGAAGAAAC GGTTCCTGGG GTAGGGGAGA TCAGCTGGGA AGAAAGCAGG
ACTTCTTTTG CCAAAGGCC CATCCCCTCT AGTCGACCCT TCTTTCGTCC

5451 TTCCTGAGCA GCTGCGACTT ACCGCAGCCG GTGGGCCCCGT AAATCACACC
AAGGACTCGT CGACGCTGAA TGGCGTCGGC CACCCGGGCA TTTAGTGTGG

5501 TATTACCGGC TGCAACTGGT AGTTAAGAGA GCTGCAGCTG CCGTCATCCC
ATAATGGCCG ACGTTGACCA TCAATTCTCT CGACGTCGAC GGCAGTAGGG

5551 TGAGCAGGGG GGCCACTTCG TTAAGCATGT CCCTGACTCG CATGTTTTCC
ACTCGTCCCC CCGGTGAAGC AATTCGTACA GGGACTGAGC GTACAAAAGG

5601 CTGACCAAAT CCGCCAGAAG GCGCTCGCCG CCCAGCGATA GCAGTTCTTG
GACTGGTTTA GCGGTCTTC CGCGAGCGGC GGTCTGCTAT CGTCAAGAAC

Figure 26 F

5701 TTTTGAGCGT TTGACCAAGC AGTTCCAGGC GGTCCCACAG CTCGGTCACC
 AAAACTCGCA AACTGGTTTCG TCAAGGTCCG CCAGGGTGTC GAGCCAGTGG
 5751 TGCTCTACGG CATCTCGATC CAGCATATCT CCTCGTTTCG CGGGTTGGGG
 ACGAGATGCC GTAGAGCTAG GTCGTATAGA GGAGCAAAGC GCCCAACCCC
 5801 CGGCTTTTCGC TGTACGGCAG TAGTCGGTGC TCGTCCAGAC GGGCCAGGGT
 GCCGAAAGCG ACATGCCGTC ATCAGCCACG AGCAGGTCTG CCCGGTCCCA
 5851 CATGTCTTTC CACGGGCGCA GGGTCCTCGT CAGCGTAGTC TGGGTCACGG
 GTACAGAAAG GTGCCCCGCT CCCAGGAGCA GTCGCATCAG ACCCAGTGCC
 5901 TGAAGGGGTG CGCTCCGGGC TGC GCGCTGG CCAGGGTGCG CTTGAGGCTG
 ACTTCCCCAC GCGAGGCCCG ACGCGCGACC GGTCCCACGC GA ACTCCGAC
 5951 GTCCTGCTGG TGCTGAAGCG CTGCCGGTCT TCGCCCTGCG CGTCGGCCAG
 CAGGACGACC ACGACTTCGC GACGGCCAGA AGCGGGACGC GCAGCCGGTC
 6001 GTAGCATTTG ACCATGGTGT CATAGTCCAG CCCCTCCGCG GCGTGGCCCT
 CATCGTAAAC TGGTACCACA GTATCAGGTC GGGGAGGCGC CGCACC GGGA
 6051 TGGCGCGCAG CTTGCCCTTG GAGGAGGCGC CGCACGAGGG GCAGTGCAGA
 ACCGCGCGTC GAACGGGAAC CTCCTCCGCG GCGTGCTCCC CGTCACGTCT
 6101 CTTTTGAGGG CGTAGAGCTT GGGCGCGAGA AATACCGATT CCGGGGAGTA
 GAAACTCCC GCATCTCGAA CCCGCGCTCT TTATGGCTAA GGCCCCTCAT
 6151 GGCATCCGCG CCGCAGGCCC CGCAGACGGT CTCGCATTCC ACGAGCCAGG
 CCGTAGGCGC GGCGTCCGGG GCGTCTGCCA GAGCGTAAGG TGCTCGGTCC
 6201 TGAGCTCTGG CCGTTCGGGG TCAAAAACCA GGTTTCCCCC ATGCTTTTTG
 ACTCGAGACC GGCAAGCCCC AGTTTTTGCT CCAAAGGGGG TACGAAAAAC
 6251 ATGCGTTTCT TACCTCTGGT TTCCATGAGC CGGTGTCCAC GCTCGGTGAC
 TACGCAAAGA ATGGAGACCA AAGGTACTCG GCCACAGGTG CGAGCCACTG
 6301 GAAAAGGCTG TCCGTGTCCC CGTATACAGA CTTGAGAGGC CTGTCCTCGA
 CTTTTCGAC AGGCACAGGG GCATATGTCT GAACTCTCCG GACAGGAGCT
 6351 GCGGTGTTCC GCGGTCCCTCC TCGTATAGAA ACTCGGACCA CTCTGAGACA
 CGCCACAAGG CGCCAGGAGG AGCATATCTT TGAGCCTGGT GAGACTCTGT
 6401 AAGGCTCGCG TCCAGGCCAG CACGAAGGAG GCTAAGTGGG AGGGGTAGCG
 TTCCGAGCGC AGGTCCGGTC GTGCTTCCTC CGATTACCC TCCCCATCGC
 6451 GTCGTTGTCC ACTAGGGGGT CCACTCGCTC CAGGGTGTGA AGACACATGT
 CAGCAACAGG TGATCCCCCA GGTGAGCGAG GTCCCACACT TCTGTGTACA
 6501 CGCCCTCTTC GGCATCAAGG AAGGTGATTG GTTTGTAGGT GTAGGCCACG
 GCGGGAGAAG CCGTAGTTCC TTCCACTAAC CAAACATCCA CATCCGGTGC
 6551 TGACCGGGTG TTCCTGAAGG GGGGCTATAA AAGGGGGTGG GGGCGCGTTC
 ACTGGCCCCAC AAGGACTTCC CCCC GATATT TTCCCCCACC CCCGCGCAAG

Figure 266

6651 AGTACTCCCT CTGAAAAGCG GGCATGACTT CTGCGCTAAG ATTGTCAGTT
TCATGAGGGA GACTTTTTCGC CCGTACTGAA GACGCGATTC TAACAGTCAA

6701 TCCAAAAACG AGGAGGATTT GATATTCACC TGGCCCGCGG TGATGCCTTT
AGGTTTTTGC TCCTCCTAAA CTATAAGTGG ACCGGGCGCC ACTACGGAAA

6751 GAGGGTGGCC GCATCCATCT GGTGAGAAAA GACAATCTTT TTGTTGTCAA
CTCCCACCGG CGTAGGTAGA CCAGTCTTTT CTGTTAGAAA AACAACAGTT

6801 GCTTGGTGGC AAACGACCCG TAGAGGGCGT TGGACAGCAA CTTGGCGATG
CGAACCACCG TTTGCTGGGC ATCTCCCGCA ACCTGTCGTT GAACCGCTAC

6851 GAGCGCAGGG TTTGGTTTTT GTCGCGATCG GCGCGCTCCT TGGCCGCGAT
CTCGCGTCCC AAACCAAAAA CAGCGCTAGC CCGCGAGGA ACCGGCGCTA

6901 GTTTAGCTGC ACGTATTCGC GCGCAACGCA CCGCCATTCG GGAAAGACGG
CAAATCGACG TGCATAAGCG CGCGTTGCGT GCGGTAAGC CCTTTCTGCC

6951 TGGTGCGCTC GTCGGGCACC AGGTGCACGC GCCAACCGCG GTTGTGCAGG
ACCACGCGAG CAGCCCGTGG TCCACGTGCG CCGTTGGCGC CAACACGTCC

7001 GTGACAAGGT CAACGCTGGT GGCTACCTCT CCGCGTAGGC GCTCGTTGGT
CACTGTTCCA GTTGCACCA CCGATGGAGA GCGCATCCG CGAGCAACCA

7051 CCAGCAGAGG CGGCCGCCCT TCGCGAGCA GAATGGCGGT AGGGGGTCTA
GGTCGTCTCC GCCGGCGGGA ACGCGCTCGT CTTACCGCCA TCCCCAGAT

7101 GCTGCGTCTC GTCCGGGGGG TCTGCGTCCA CCGTAAAGAC CCCGGGCAGC
CGACGCAGAG CAGGCCCCCC AGACGCAGGT GCCATTTCTG GGGCCCGTCG

7151 AGGCGCGCGT CGAAGTAGTC TATCTTGCAT CCTTGCAAGT CTAGCGCCTG
TCCGCGCGCA GCTTCATCAG ATAGAACGTA GGAACGTTCA GATCGCGGAC

7201 CTGCCATGCG CGGGCGGCAA GCGCGCGCTC GTATGGGTTG AGTGGGGGAC
GACGGTACGC GCCCGCGGTT CGCGCGCGAG CATACCCAAC TCACCCCTG

7251 CCCATGGCAT GGGGTGGGTG AGCGCGGAGG CGTACATGCC GCAAATGTCTG
GGGTACCGTA CCCCACCCAC TCGCGCTCC GCATGTACGG CGTTTACAGC

7301 TAAACGTAGA GGGGCTCTCT GAGTATTCCA AGATATGTAG GGTAGCATCT
ATTTGCATCT CCCCAGAGAG CTCATAAGGT TCTATACATC CCATCGTAGA

7351 TCCACCGCGG ATGCTGGCGC GCACGTAATC GTATAGTTCTG TGCGAGGGAG
AGGTGGCGCC TACGACCGCG CGTGCATTAG CATATCAAGC ACGCTCCCTC

7401 CGAGGAGGTC GGGACCGAGG TTGCTACGGG CGGGCTGCTC TGCTCGGAAG
GCTCCTCCAG CCCTGGCTCC AACGATGCCC GCCCGACGAG ACGAGCCTTC

7451 ACTATCTGCC TGAAGATGGC ATGTGAGTTG GATGATATGG TTGGACGCTG
TGATAGACGG ACTTCTACCG TACACTCAAC CTACTATACC AACCTGCGAC

7501 GAAGACGTTG AAGCTGGCGT CTGTGAGACC TACCGCGTCA CGCACGAAGG
CTTCTGCAAC TTCGACCGCA GACACTCTGG ATGGCGCAGT GCGTGCTTCC

Figure 26 H

7601 TCTAGGGCGC AGTAGTCCAG GGTTCCTTG ATGATGTCAT ACTTATCCTG
AGATCCCGCG TCATCAGGTC CCAAAGGAAC TACTACAGTA TGAATAGGAC

7651 TCCCTTTTTT TTCCACAGCT CGCGGTTGAG GACAACTCT TCGCGGTCTT
AGGGAAAAAA AAGGTGTCGA GCGCCAACTC CTGTTTGAGA AGCGCCAGAA

7701 TCCAGTACTC TTGGATCGGA AACCCGTCGG CCTCCGAACG GTAAGAGCCT
AGGTCATGAG AACCTAGCCT TTGGGCAGCC GGAGGCTTGC CATTCTCGGA

7751 AGCATGTAGA ACTGGTTGAC GGCCTGGTAG GCGCAGCATC CCTTTTCTAC
TCGTACATCT TGACCAACTG CCGGACCATC CGCGTCGTAG GGAAAAGATG

7801 GGGTAGCGCG TATGCCTGCG CGGCCTTCCG GAGCGAGGTG TGGGTGAGCG
CCCATCGCGC ATACGGACGC GCCGGAAGGC CTCGCTCCAC ACCCACTCGC

7851 CAAAGGTGTC CCTGACCATG ACTTTGAGGT ACTGGTATTT GAAGTCAGTG
GTTTCCACAG GGACTGGTAC TGAAACTCCA TGACCATAAA CTTCAGTCAC

7901 TCGTCGCATC CGCCCTGCTC CCAGAGCAAA AAGTCCGTGC GCTTTTGGGA
AGCAGCGTAG GCGGGACGAG GGTCTCGTTT TTCAGGCACG CGAAAAACCT

7951 ACGCGGATTT GGCAGGGCGA AGGTGACATC GTTGAAGAGT ATCTTTCCCG
TGCGCCTAAA CCGTCCCGCT TCCACTGTAG CAACTTCTCA TAGAAAGGGC

8001 CGCGAGGCAT AAAGTTGCGT GTGATGCGGA AGGGTCCCGG CACCTCGGAA
GCGCTCCGTA TTCAACGCA CACTACGCCT TCCCAGGGCC GTGGAGCCTT

8051 CGGTTGTTAA TTACCTGGGC GCGGAGCACG ATCTCGTCAA AGCCGTTGAT
GCCAACAATT AATGGACCCG CCGCTCGTGC TAGAGCAGTT TCGGCAACTA

8101 GTTGTGGCCC ACAATGTAAA GTTCCAAGAA GCGCGGGATG CCCTTGATGG
CAACACCGGG TGTTACATTT CAAGGTTCTT CGCGCCCTAC GGGAACTACC

8151 AAGGCAATTT TTTAAGTTCC TCGTAGGTGA GCTCTTCAGG GGAGCTGAGC
TTCCGTAAA AAATTCAAGG AGCATCCACT CGAGAAGTCC CCTCGACTCG

8201 CCGTGCTCTG AAAGGGCCCA GTCTGCAAGA TGAGGGTTGG AAGCGACGAA
GGCACGAGAC TTTCCCGGGT CAGACGTTCT ACTCCCAACC TTCGCTGCTT

8251 TGAGCTCCAC AGGTCACGGG CCATTAGCAT TTGCAGGTGG TCGCGAAAGG
ACTCGAGGTG TCCAGTGCCC GGTAATCGTA AACGTCCACC AGCGCTTTCC

8301 TCCTAAACTG GCGACCTATG GCCATTTTTT CTGGGGTGAT GCAGTAGAAG
AGGATTTGAC CGCTGGATAC CGGTAAAAAA GACCCCACTA CGTCATCTTC

8351 GTAAGCGGGT CTTGTTCCCA GCGGTCCCAT CCAAGGTTCG CGGCTAGGTC
CATTCGCCCA GAACAAGGGT CGCCAGGGTA GGTTCCAAGC GCCGATCCAG

8401 TCGCGCGGCA GTCACTAGAG GCTCATCTCC GCCGAAC TTC ATGACCAGCA
AGCGCGCCGT CAGTGATCTC CGAGTAGAGG CGGCTTGAAG TACTGGTCGT

8451 TGAAGGGCAC GAGCTGCTTC CCAAAGGCC CCATCCAAGT ATAGGTCTCT
ACTTCCCGTG CTCGACGAAG GGTTTCCGGG GGTAGGTTCA TATCCAGAGA

Figure 26I

8551 GAAGAACTGG ATCTCCCGCC ACCAATTGGA GGAGTGGCTA TTGATGTGGT
CTTCTTGACC TAGAGGGCGG TGGTTAACCT CCTCACCGAT AACTACACCA

8601 GAAAGTAGAA GTCCCTGCGA CGGGCCGAAC ACTCGTGCTG GCTTTTGTA
CTTTCATCTT CAGGGACGCT GCCCGGCTTG TGAGCACGAC CGAAAACATT

8651 AAACGTGCGC AGTACTGGCA GCGGTGCACG GGCTGTACAT CCTGCACGAG
TTTGCACGCG TCATGACCGT CGCCACGTGC CCGACATGTA GGACGTGCTC

8701 GTTGACCTGA CGACCGCGCA CAAGGAAGCA GAGTGGGAAT TTGAGCCCCT
CAACTGGACT GCTGGCGCGT GTTCCTTCGT CTCACCCTTA AACTCGGGGA

8751 CGCCTGGCGG GTTTGGCTGG TGGTCTTCTA CTTCCGGCTGC TTGTCCTTGA
GCGGACCGCC CAAACCGACC ACCAGAAGAT GAAGCCGACG AACAGGAACT

8801 CCGTCTGGCT GCTCGAGGGG AGTTACGGTG GATCGGACCA CCACGCCGCG
GGCAGACCGA CGAGCTCCCC TCAATGCCAC CTAGCCTGGT GGTGCGGCGC

8851 CGAGCCCAA GTCCAGATGT CCGCGCGCGG CGGTCCGAGC TTGATGACAA
GCTCGGGTTT CAGGTCTACA GGCGCGCGCC GCCAGCCTCG AACTACTGTT

8901 CATCGCGCAG ATGGGAGCTG TCCATGGTCT GGAGCTCCCG CGGCGTCAGG
GTAGCGCGTC TACCCTCGAC AGGTACCAGA CCTCGAGGGC GCCGCAGTCC

8951 TCAGGCGGGA GCTCCTGCAG GTTTACCTCG CATAGACGGG TCAGGGCGCG
AGTCCGCCCT CGAGGACGTC CAAATGGAGC GTATCTGCCC AGTCCCGCGC

9001 GGCTAGATCC AGGTGATACC TAATTTCCAG GGGCTGGTTG GTGGCGGCGT
CCGATCTAGG TCCACTATGG ATTAAAGGTC CCCGACCAAC CACCGCCGCA

9051 CGATGGCTTG CAAGAGGCCG CATCCCCGCG GCGCGACTAC GGTACCGCGC
GCTACCGAAC GTTCTCCGGC GTAGGGGCGC GCGCTGATG CCATGGCGCG

9101 GGCGGGCGGT GGGCCGCGGG GGTGTCCTTG GATGATGCAT CTAAAAGCGG
CCGCCCCGCA CCCGGCGCCC CCACAGGAAC CTACTACGTA GATTTTCGCC

9151 TGACGCGGGC GAGCCCCCGG AGGTAGGGGG GGCTCCGGAC CCGCCGGGAG
ACTGCGCCCCG CTCGGGGGCC TCCATCCCCC CCGAGGCCTG GCGGGCCCTC

9201 AGGGGGCAGG GGCACGTCGG CGCCGCGCGC GGGCAGGAGC TGGTGCTGCG
TCCCCGTCC CCGTGCAGCC GCGGCGCGCG CCCGTCCTCG ACCACGACGC

9251 CGCGTAGGTT GCTGGCGAAC GCGACGACGC GGCGGTTGAT CTCCTGAATC
GCGCATCCAA CGACCGCTTG CGCTGCTGCG CCGCCAACTA GAGGACTTAG

9301 TGGCGCCTCT GCGTGAAGAC GACGGGCCCC GTGAGCTTGA ACCTGAAAGA
ACCGCGGAGA CGCACTTCTG CTGCCCCGGC CACTCGAACT TGGACTTTCT

9351 GAGTTCGACA GAATCAATTT CGGTGTCGTT GACGGCGGCC TGGCGCAAAA
CTCAAGCTGT CTTAGTTAAA GCCACAGCAA CTGCCGCCGG ACCGCGTTTT

9401 TCTCCTGCAC GTCTCCTGAG TTGTCTTGAT AGGCGATCTC GGCCATGAAC
AGAGGACGTG CAGAGGACTC AACAGAACTA TCCGCTAGAG CCGGTACTTG

Figure 26 J

9501 GCGGCGGAGG TCGTTGGAAG TCGGGGCCAT GAGCTGCGAG AAGGCGTTGA
CCGCCGCTCC AGCAACCTTT ACGCCCGGTA CTCGACGCTC TTCCGCAACT

9551 GGCCTCCCTC GTTCCAGACG CGGCTGTAGA CCACGCCCCC TTCGGCATCG
CCGGAGGGAG CAAGGTCTGC GCCGACATCT GGTGCGGGGG AAGCCGTAGC

9601 CGGGCGCGCA TGACCACCTG CGCGAGATTG AGCTCCACGT GCCGGGCGAA
GCCCCGCGCT ACTGGTGGAC GCGCTCTAAC TCGAGGTGCA CGGCCCGCTT

9651 GACGGCGTAG TTTCGCAGGC GCTGAAAGAG GTAGTTGAGG GTGGTGGCGG
CTGCCGCATC AAAGCGTCCG CGACTTTCTC CATCAACTCC CACCACCGCC

9701 TGTGTTCTGC CACGAAGAAG TACATAACCC AGCGTCGCAA CGTGGATTCTG
ACACAAGACG GTGCTTCTTC ATGTATTGGG TCGCAGCGTT GCACCTAAGC

9751 TTGATATCCC CCAAGGCCTC AAGGCGCTCC ATGGCCTCGT AGAAGTCCAC
AACTATAGGG GGTTCGCGAG TTCCGCGAGG TACCGGAGCA TCTTCAGGTG

9801 GGCGAAGTTG AAAAAGTGGG AGTTGCGCGC CGACACGGTT AACTCCTCCT
CCGCTTCAAC TTTTGTGACC TCAACGCGCG GCTGTGCCAA TTGAGGAGGA

9851 CCAGAAGACG GATGAGCTCG GCGACAGTGT CGCGCACCTC GCGCTCAAAG
GGTCTTCTGC CTACTCGAGC CGCTGTCACA GCGCGTGGAG CGCGAGTTTC

9901 GCTACAGGGG CCTCTTCTTC TTCTTCAATC TCCTCTTCCA TAAGGGCCTC
CGATGTCCCC GGAGAAGAAG AAGAAGTTAG AGGAGAAGGT ATTCCCGGAG

9951 CCCTTCTTCT TCTTCTGGCG GCGGTGGGGG AGGGGGGACA CGGCGGCGAC
GGGAAGAAGA AGAAGACCGC CGCCACCCCC TCCCCCTGT GCCGCCGCTG

10001 GACGGCGCAC CGGGAGGCGG TCGACAAAGC GCTCGATCAT CTCCCCGCGG
CTGCCGCGTG GCCCTCCGCC AGCTGTTTCG CGAGCTAGTA GAGGGGCGCC

10051 CGACGGCGCA TGGTCTCGGT GACGGCGCGG CCGTTCTCGC GGGGGCGCAG
GCTGCCGCGT ACCAGAGCCA CTGCCGCGCC GGCAAGAGCG CCCCCGCGTC

10101 TTGAAGACG CCGCCCGTCA TGTCCCGGTT ATGGGTTGGC GGGGGGCTGC
AACCTTCTGC GCGGGGCGAG ACAGGGCCAA TACCAACCG CCCCCGACG

10151 CATGCGGCAG GGATACGGCG CTAACGATGC ATCTCAACAA TTGTTGTGTA
GTACGCCGTC CCTATGCCGC GATTGCTACG TAGAGTTGTT AACAAACACAT

10201 GGTACTCCGC CGCCGAGGGA CCTGAGCGAG TCCGCATCGA CCGGATCGGA
CCATGAGGCG GCGGCTCCCT GGACTCGCTC AGGCGTAGCT GGCCTAGCCT

10251 AAACCTCTCG AGAAAGGCGT CTAACGATC ACAGTCGCAA GGTAGGCTGA
TTTGGAGAGC TCTTCCGCA GATTGGTCAG TGTCAGCGTT CCATCCGACT

10301 GCACCGTGGC GGGCGGCGAG GGGCGGCGGT CGGGGTTGTT TCTGGCGGAG
CGTGGCACCG CCCGCCGTCG CCCGCCGCCA GCCCCAACA AGACCGCCTC

10351 GTGCTGCTGA TGATGTAATT AAAGTAGGCG GTCTTGAGAC GCGGATGGT
CACGACGACT ACTACATTAA TTTCATCCGC CAGAACTCTG CCGCCTACCA

Figure 26 K

10451 CGGCCATGCC CCAGGCTTCG TTTTGACATC GGCGCAGGTC TTTGTAGTAG
GCCGGTACGG GGTCCGAAGC AAAACTGTAG CCGCGTCCAG AAACATCATC

10501 TCTTGCATGA GCCTTTCTAC CGGCACTTCT TCTTCTCCTT CCTCTTGTC
AGAACGTACT CGGAAAGATG GCCGTGAAGA AGAAGAGGAA GGAGAACAGG

10551 TGCATCTCTT GCATCTATCG CTGCGGCGGC GGCGGAGTTT GGCCGTAGGT
ACGTAGAGAA CGTAGATAGC GACGCCGCCG CCGCCTCAA CCGGCATCCA

10601 GGCGCCCTCT TCCTCCCATG CGTGTGACCC CGAAGCCCCT CATCGGCTGA
CCGCGGGAGA AGGAGGGTAC GCACACTGGG GCTTCGGGGA GTAGCCGACT

10651 AGCAGGGCTA GGTCGGCGAC AACGCGCTCG GCTAATATGG CCTGCTGCAC
TCGTCCCGAT CCAGCCGCTG TTGCGCGAGC CGATTATACC GGACGACGTG

10701 CTGCGTGAGG GTAGACTGGA AGTCATCCAT GTCCACAAAG CGGTGGTATG
GACGCACTCC CATCTGACCT TCAGTAGGTA CAGGTGTTTC GCCACCATAC

10751 CGCCCGTGTT GATGGTGTA GTGCAGTTGG CCATAACGGA CCAGTTAACG
GCGGGCACAA CTACCACATT CACGTCAACC GGTATTGCCT GGTCAATTGC

10801 GTCTGGTGAC CCGGCTGCGA GAGCTCGGTG TACCTGAGAC GCGAGTAAGC
CAGACCACTG GGCCGACGCT CTCGAGCCAC ATGGACTCTG CGCTCATTCG

10851 CCTCGAGTCA AATACGTAGT CGTTGCAAGT CCGCACCAGG TACTGGTATC
GGAGCTCAGT TTATGCATCA GCAACGTTCA GGCGTGGTCC ATGACCATAG

10901 CCACCAAAAA GTGCGGCGGC GGCTGGCGGT AGAGGGGCCA GCGTAGGGTG
GGTGGTTTTT CACGCCGCCG CCGACCGCCA TCTCCCCGGT CGCATCCCAC

10951 GCCGGGGCTC CGGGGGCGAG ATCTTCCAAC ATAAGGCGAT GATATCCGTA
CGGCCCCGAG GCCCCCGCTC TAGAAGGTTG TATTCCGCTA CTATAGGCAT

11001 GATGTACCTG GACATCCAGG TGATGCCGGC GGCGGTGGTG GAGGCGCGCG
CTACATGGAC CTGTAGGTCC ACTACGGCCG CCGCCACCAC CTCCGCGCGC

11051 GAAAGTCGCG GACGCGGTTC CAGATGTTGC GCAGCGGCAA AAAGTGCTCC
CTTTCAGCGC CTGCGCCAAG GTCTACAACG CGTCGCCGTT TTTCACGAGG

11101 ATGGTCGGGA CGCTCTGGCC GGTCAGGCGC GCGCAATCGT TGACGCTCTA
TACCAGCCCT GCGAGACCGG CCAGTCCGCG CGCGTTAGCA ACTGCGAGAT

11151 GACCGTGCAA AAGGAGAGCC TGTAAGCGGG CACTCTTCCG TGGTCTGGTG
CTGGCACGTT TTCCTCTCGG ACATTGCCCC GTGAGAAGGC ACCAGACCAC

11201 GATAAATTCG CAAGGGTATC ATGGCGGACG ACCGGGGTTC GAGCCCCGTA
CTATTTAAGC GTTCCCATAG TACCGCCTGC TGGCCCCAAG CTCGGGGCAT

11251 TCCGGCCGTC CGCCGTGATC CATGCGGTTA CCGCCCGCGT GTCGAACCCA
AGGCCGGCAG GCGGCACTAG GTACGCCAAT GGCGGGCGCA CAGCTTGGGT

11301 GGTGTGCGAC GTCAGACAAC GGGGGAGTGC TCCTTTTGGC TTCCTTCCAG
CCACACGCTG CAGTCTGTTG CCCCCTCAG AGGAAAACCG AAGGAAGGTC

Figure 26L

11401 AAGCGGTTAG GCTGGAAAGC GAAAGCATTA AGTGGCTCGC TCCCTGTAGC
TTCGCCAATC CGACCTTTCG CTTTCGTAAT TCACCGAGCG AGGGACATCG

11451 CGGAGGGTTA TTTTCCAAGG GTTGAGTCGC GGGACCCCCG GTTCGAGTCT
GCCTCCCAAT AAAAGGTTCC CAACTCAGCG CCCTGGGGGC CAAGCTCAGA

11501 CGGACCGGCC GGACTGCGGC GAACGGGGGT TTGCCTCCCC GTCATGCAAG
GCCTGGCCGG CCTGACGCCG CTTGCCCCCA AACGGAGGGG CAGTACGTTC

11551 ACCCCGCTTG CAAATTCCTC CGGAAACAGG GACGAGCCCC TTTTGTGCTT
TGGGGCGAAC GTTTAAGGAG GCCTTTGTCC CTGCTCGGGG AAAAAACGAA

11601 TTCCCAGATG CATCCGGTGC TGCGGCAGAT GCGCCCCCTT CCTCAGCAGC
AAGGGTCTAC GTAGGCCACG ACGCCGTCTA CGCGGGGGGA GGAGTCGTCTG

11651 GGCAAGAGCA AGAGCAGCGG CAGACATGCA GGGCACCTTC CCCTCCTCCT
CCGTTCTCGT TCTCGTCGCC GTCTGTACGT CCCGTGGGAG GGGAGGAGGA

11701 ACCGCGTCAG GAGGGGCGAC ATCCGCGGTT GACGCGGCAG CAGATGGTGA
TGCGCGAGTC CTCCCCGCTG TAGGCGCCAA CTGCGCCGTC GTCTACCACT

11751 TTACGAACCC CCGCGGCGCC GGGCCCGGCA CTACCTGGAC TTGGAGGAGG
AATGCTTGGG GCGCGCGCGG CCCGGGCGGT GATGGACCTG AACCTCCTCC

11801 GCGAGGGCCT GCGCGGGCTA GGAGCGCCCT CTCCTGAGCG GCACCCAAGG
CGCTCCCGGA CCGCGCCGAT CCTCGCGGGA GAGGACTCGC CGTGGGTTCC

11851 GTGCAGCTGA AGCGTGATAC GCGTGAGGCG TACGTGCCGC GGCAGAACCT
CACGTCGACT TCGCACTATG CGCACTCCGC ATGCACGGCG CCGTCTTGGA

11901 GTTTCGCGAC CGCGAGGGAG AGGAGCCCGA GGAGATGCGG GATCGAAAGT
CAAAGCGCTG GCGCTCCCTC TCCTCGGGCT CCTCTACGCC CTAGCTTTCA

11951 TCCACGCAGG GCGCGAGCTG CGGCATGGCC TGAATCGCGA GCGGTTGCTG
AGGTGCGTCC CGCGCTCGAC GCCGTACCGG ACTTAGCGCT CGCCAACGAC

12001 CGCGAGGAGG ACTTTGAGCC CGACGCGCGA ACCGGGATTA GTCCCGCGCG
GCGCTCCTCC TGAAACTCGG GCTGCGCGCT TGGCCCTAAT CAGGGCGCGC

12051 CGCACACGTG GCGGCCGCGG ACCTGGTAAC CGCATACTAG CAGACGGTGA
GCGTGTGCAC CGCCGGCGGC TGGACCATTG GCGTATGCTC GTCTGCCACT

12101 ACCAGGAGAT TAACTTTCAA AAAAGCTTTA ACAACCACGT GCGTACGCTT
TGGTCCTCTA ATTGAAAGTT TTTTCGAAAT TGTGTTGGTCA CGCATGCGAA

12151 GTGGCGCGCG AGGAGGTGGC TATAGGACTG ATGCATCTGT GGGACTTTGT
CACCGCGCGC TCCTCCACCG ATATCCTGAC TACGTAGACA CCCTGAAACA

12201 AAGCGCGCTG GAGCAAAACC CAAATAGCAA GCCGCTCATG GCGCAGCTGT
TTCGCGCGAC CTCGTTTTGG GTTTATCGTT CGGCGAGTAC CGCGTCGACA

12251 TCCTTATAGT GCAGCACAGC AGGGACAACG AGGCATTAG GGATGCGCTG
AGGAATATCA CGTCGTGTCTG TCCCTGTTGC TCCGTAAGTC CCTACGCGAC

Figure 26 M

12351 CCTGCAGAGC ATAGTGGTGC AGGAGCGCAG CTTGAGCCTG GCTGACAAGG
GGACGTCTCG TATCACCACG TCCTCGCGTC GAACTCGGAC CGACTGTTCC

12401 TGGCCGCCAT CAACTATTCC ATGCTTAGCC TGGGCAAGTT TTACGCCCCG
ACCGGCGGTA GTTGATAAGG TACGAATCGG ACCCGTTCAA AATGCGGGCG

12451 AAGATATACC ATACCCCTTA CGTTCCCATTA GACAAGGAGG TAAAGATCGA
TTCTATATGG TATGGGGAAT GCAAGGGTAT CTGTTCTCTC ATTTCTAGCT

12501 GGGGTTCTAC ATGCGCATGG CGCTGAAGGT GCTTACCTTG AGCGACGACC
CCCCAAGATG TACGCGTACC GCGACTTCCA CGAATGGAAC TCGCTGCTGG

12551 TGGGCGTTTA TCGCAACGAG CGCATCCACA AGGCCGTGAG CGTGAGCCGG
ACCGCAAAT AGCGTTGCTC GCGTAGGTGT TCCGGCACTC GCACTCGGCC

12601 CGGCGCGAGC TCAGCGACCG CGAGCTGATG CACAGCCTGC AAAGGGCCCT
GCCGCGCTCG AGTCGCTGGC GCTCGACTAC GTGTCGGACG TTTCCCGGGA

12651 GGCTGGCACG GGCAGCGGCG ATAGAGAGGC CGAGTCCTAC TTTGACGCGG
CCGACCGTGC CCGTCGCCGC TATCTCTCCG GCTCAGGATG AAAGTGGCGC

12701 GCGCTGACCT GCGCTGGGCC CCAAGCCGAC GCGCCCTGGA GGCAGCTGGG
CGCGACTGGA CGCGACCCGG GGTTCGGCTG CGCGGGACCT CCGTCGACCC

12751 GCCGGACCTG GGCTGGCGGT GGCACCCGCG CGCGCTGGCA ACGTCGGCGG
CGGCCTGGAC CCGACCGCCA CCGTGGGCGC GCGCGACCGT TGCAGCCGCC

12801 CGTGGAGGAA TATGACGAGG ACGATGAGTA CGAGCCAGAG GACGGCGAGT
GCACCTCCTT ATACTGCTCC TGCTACTCAT GCTCGGTCTC CTGCCGCTCA

12851 ACTAAGCGGT GATGTTTCTG ATCAGATGAT GCAAGACGCA ACGGACCCGG
TGATTGCGCA CTACAAAGAC TAGTCTACTA CGTTCTGCGT TGCCTGGGCC

12901 CGGTGCGGGC GGCCTGTCAG AGCCAGCCGT CCGGCCTTAA CTCCACGGAC
GCCACGCCCCG CCGCGACGTC TCGGTCGGCA GGCCGGAATT GAGGTGCCTG

12951 GACTGGCGCC AGGTCATGGA CCGCATCATG TCGCTGACTG CGCGCAATCC
CTGACCGCGG TCCAGTACCT GGCGTAGTAC AGCGACTGAC GCGCGTTAGG

13001 TGACGCGTTC CGGCAGCAGC CGCAGGCCAA CCGGCTCTCC GCAATTCTGG
ACTGCGCAAG GCCGTCGTCG GCGTCCGGTT GGCCGAGAGG CGTTAAGACC

13051 AAGCGGTGGT CCCGGCGCGC GCAAACCCCA CGCACGAGAA GGTGCTGGCG
TTCGCCACCA GGGCCGCGCG CGTTTGGGGT GCGTGCTCTT CCACGACCGC

13101 ATCGTAAACG CGCTGGCCGA AAACAGGGCC ATCCGGCCCCG ACGAGGCCGG
TAGCATTTGC GCGACCGGCT TTTGTCCCGG TAGGCCGGGC TGCTCCGGCC

13151 CCTGGTCTAC GACGCGCTGC TTCAGCGCGT GGCTCGTTAC AACAGCGGCA
GGACCAGATG CTGCGCGACG AAGTCGCGCA CCGAGCAATG TTGTGCGCGT

13201 ACGTGCAGAC CAACCTGGAC CGGCTGGTGG GGGATGTGCG CGAGGCCGTG
TGCACGTCTG GTTGGACCTG GCCGACCACC CCCTACACGC GCTCCGGCAC

Figure 26 N

13301 ACTAAACGCC TTCCTGAGTA CACAGCCCGC CAACGTGCCG CGGGGACAGG
TGATTTGCGG AAGGACTCAT GTGTCGGGCG GTTGACACGC GCCCCTGTCC

13351 AGGACTACAC CAACTTTGTG AGCGCACTGC GGCTAATGGT GACTGAGACA
TCCTGATGTG GTTGAAACAC TCGCGTGACG CCGATTACCA CTGACTCTGT

13401 CCGCAAAGTG AGGTGTACCA GTCTGGGCCA GACTATTTTT TCCAGACCAG
GGCGTTTCAC TCCACATGGT CAGACCCGGT CTGATAAAAA AGGTCTGGTC

13451 TAGACAAGGC CTGCAGACCG TAAACCTGAG CCAGGCTTTC AAAA ACTTGC
ATCTGTTCCG GACGTCTGGC ATTTGGACTC GGTCCGAAAG TTTTGAACG

13501 AGGGGCTGTG GGGGGTGCGG GCTCCCACAG GCGACCGCGC GACCGTGTCT
TCCCCGACAC CCCCCACGCC CGAGGGTGTC CGCTGGCGCG CTGGCACAGA

13551 AGCTTGCTGA CGCCCAACTC GCGCCTGTTG CTGCTGCTAA TAGCGCCCTT
TCGAACGACT GCGGGTTGAG CGCGGACAAC GACGACGATT ATCGCGGGAA

13601 CACGGACAGT GGCAGCGTGT CCCGGGACAC ATACCTAGGT CACTTGCTGA
GTGCCTGTCA CCGTCGCACA GGGCCCTGTG TATGGATCCA GTGAACGACT

13651 CACTGTACCG CGAGGCCATA GGTCAGGCGC ATGTGGACGA GCATACTTTC
GTGACATGGC GCTCCGGTAT CCAGTCCGCG TACACCTGCT CGTATGAAAG

13701 CAGGAGATTA CAAGTGTCAG CCGCGCGCTG GGGCAGGAGG ACACGGGCAG
GTCCTCTAAT GTTCACAGTC GGCGCGCGAC CCCGTCCTCC TGTGCCCGTC

13751 CCTGGAGGCA ACCCTAAACT ACCTGCTGAC CAACCGGCGG CAGAAGATCC
GGACCTCCGT TGGGATTTGA TGGACGACTG GTTGGCCGCC GTCTTCTAGG

13801 CCTCGTTGCA CAGTTTAAAC AGCGAGGAGG AGCGCATTTT GCGCTACGTG
GGAGCAACGT GTCAAATTTG TCGCTCCTCC TCGCGTAAAA CGCGATGCAC

13851 CAGCAGAGCG TGAGCCTTAA CCTGATGCGC GACGGGGTAA CGCCCAGCGT
GTCGTCTCGC ACTCGGAATT GGACTACGCG CTGCCCCATT GCGGGTCGCA

13901 GGCGCTGGAC ATGACCGCGC GCAACATGGA ACCGGGCATG TATGCCTCAA
CCGCGACCTG TACTGGCGCG CGTTGTACCT TGGCCCGTAC ATACGGAGTT

13951 ACCGGCCGTT TATCAACCGC CTAATGGACT ACTTGCAATG CGCGGCCGCC
TGGCCGGCAA ATAGTTGGCG GATTACCTGA TGAACGTAGC GCGCCGGCGG

14001 GTGAACCCCG AGTATTTTAC CAATGCCATC TTGAACCCGC ACTGGCTACC
CACTTGGGGC TCATAAAGTG GTTACGGTAG AACTTGGGCG TGACCGATGG

14051 GCCCCCTGGT TTCTACACCG GGGGATTCGA GGTGCCCCGAG GGTAACGATG
CGGGGGACCA AAGATGTGGC CCCCTAAGCT CCACGGGCTC CCATTGCTAC

14101 GATTCCTCTG GGACGACATA GACGACAGCG TGTTTTCCCC GCAACCGCAG
CTAAGGAGAC CCTGCTGTAT CTGCTGTCGC ACAAAGGGG CGTTGGCGTC

14151 ACCCTGCTAG AGTTGCAACA GCGCGAGCAG GCAGAGGCGG CGCTGCGAAA
TGGGACGATC TCAACGTTGT CGCGCTCGTC CGTCTCCGCC GCGACGCTTT

Figure 260

14251 CGCGGTCAGA TGCTAGTAGC CCATTTCCAA GCTTGATAGG GTCTCTTACC
GCGCCAGTCT ACGATCATCG GGTAAAGGTT CGAACTATCC CAGAGAATGG

14301 AGCACTCGCA CCACCCGCCC GCGCCTGCTG GCGGAGGAGG AGTACCTAAA
TCGTGAGCGT GGTGGGCGGG CGCGGACGAC CCGCTCCTCC TCATGGATTT

14351 CAACTCGCTG CTGCAGCCGC AGCGCGAAAA AAACCTGCCT CCGGCATTTC
GTTGAGCGAC GACGTCGGCG TCGCGCTTTT TTTGGACGGA GGCCGTAAAG

14401 CCAACAACGG GATAGAGAGC CTAGTGGACA AGATGAGTAG ATGGAAGACG
GGTTGTTGCC CTATCTCTCG GATCACCTGT TCTACTCATC TACCTTCTGC

14451 TACGCGCAGG AGCACAGGGA CGTGCCAGGC CCGCGCCCGC CCACCCGTCC
ATGCGCGTCC TCGTGTCCTT GCACGGTCCG GCGCGGGGCG GGTGGGCAGC

14501 TCAAAGGCAC GACCGTCAGC GGGGTCTGGT GTGGGAGGAC GATGACTCGG
AGTTTCCGTG CTGGCAGTCG CCCCAGACCA CACCCTCCTG CTA CTGAGCC

14551 CAGACGACAG CAGCGTCCTG GATTTGGGAG GGAGTGGCAA CCCGTTTGCG
GTCTGCTGTC GTCGCAGGAC CTAAACCCTC CCTCACCGTT GGGCAAACGC

14601 CACCTTCGCC CCAGGCTGGG GAGAATGTTT TAAAAA AAAAGCATGA
GTGGAAGCGG GGTCCGACCC CTCTTACAAA ATTTTTTTTT TTTTCGTACT

14651 TGCAAAATAA AAAACTCACC AAGGCCATGG CACCGAGCGT TGGTTTTCTT
ACGTTTTATT TTTTGAGTGG TTCCGGTACC GTGGCTCGCA ACCAAAAGAA

14701 GTATTCCCCT TAGTATGCGG CGCGCGGCGA TGTATGAGGA AGGTCCTCCT
CATAAGGGGA ATCATA CGCC GCGCGCCGCT ACATACTCCT TCCAGGAGGA

14751 CCCTCCTACG AGAGTGTGGT GAGCGCGGCG CCAGTGGCGG CGGCGCTGGG
GGGAGGATGC TCTCACACCA CTCGCGCCGC GGTACCGCC GCCGCGACCC

14801 TTCTCCCTTC GATGCTCCCC TGGACCCGCC GTTTGTGCCT CCGCGGTACC
AAGAGGGAAG CTACGAGGGG ACCTGGGCGG CAAACACGGA GCGGCCATGG

14851 TGCGGCCTAC CGGGGGGAGA AACAGCATCC GTTACTCTGA GTTGGCACCC
ACGCCGGATG GCCCCCTCT TTGTCGTAGG CAATGAGACT CAACCGTGGG

14901 CTATTGACA CCACCCGTGT GTACCTGGTG GACAACAAGT CAACGGATGT
GATAAGCTGT GGTGGGCACA CATGGACCAC CTGTTGTTCA GTTGCCTACA

14951 GGCATCCCTG AACTACCAGA ACGACCACAG CAACTTTCTG ACCACGGTCA
CCGTAGGGAC TTGATGGTCT TGCTGGTGTC GTTGAAAGAC TGGTGCCAGT

15001 TTCAAAACAA TGA CTACAGC CCGGGGGAGG CAAGCACACA GACCATCAAT
AAGTTTTGTT ACTGATGTG GGGCCCCCTCC GTTCGTGTGT CTGGTAGTTA

15051 CTTGACGACC GGTCGCACTG GGGCGGCGAC CTGAAAACCA TCCTGCATAC
GAACTGCTGG CCAGCGTGAC CCCGCCGCTG GACTTTTGGT AGGACGTATG

15101 CAACATGCCA AATGTGAACG AGTTCATGTT TACCAATAAG TTTAAGGCGC
GTTGTACGGT TTACACTTGC TCAAGTACAA ATGGTTATTC AAATTCCGCG

Figure 26 P

15151 GGGTGATGGT GTCGCGCTTG CCTACTAAGG ACAATCAGGT GGAGCTGAAA
CCCCTACCA CAGCGCGAAC GGATGATTCC TGTTAGTCCA CCTCGACTTT

15201 TACGAGTGGG TGGAGTTCAC GCTGCCCCGAG GGCAACTACT CCGAGACCAT
ATGCTCACCC ACCTCAAGTG CGACGGGCTC CCGTTGATGA GGCTCTGGTA

15251 GACCATAGAC CTTATGAACA ACGCGATCGT GGAGCACTAC TTGAAAGTGG
CTGGTATCTG GAATACTTGT TGCCTAGCA CCTCGTGATG AACTTTCACC

15301 GCAGACAGAA CGGGGTCTTG GAAAGCGACA TCGGGGTAAA GTTTGACACC
CGTCTGTCTT GCCCAAGAC CTTTCGCTGT AGCCCCATTT CAAACTGTGG

15351 CGCAACTTCA GACTGGGGTT TGACCCCGTC ACTGGTCTTG TCATGCCTGG
GCGTTGAAGT CTGACCCCAA ACTGGGGCAG TGACCAGAAC AGTACGGACC

15401 GGTATATACA AACGAAGCCT TCCATCCAGA CATCATTTTG CTGCCAGGAT
CCATATATGT TTGCTTCGGA AGGTAGGTCT GTAGTAAAC GACGGTCCTA

15451 GCGGGGTGGA CTTACCCAC AGCCGCCTGA GCAACTTGTT GGGCATCCGC
CGCCCCACCT GAAGTGGGTG TCGGCGGACT CGTTGAACAA CCCGTAGGCG

15501 AAGCGGCAAC CCTTCCAGGA GGGCTTTAGG ATCACCTACG ATGATCTGGA
TTCGCCGTTG GGAAGGTCCT CCCGAAATCC TAGTGGATGC TACTAGACCT

15551 GGGTGGTAAC ATTCCCGCAC TGTTGGATGT GGACGCCTAC CAGGCGAGCT
CCCACCATTG TAAGGGCGTG ACAACCTACA CCTGCGGATG GTCCGCTCGA

15601 TGAAAGATGA CACCGAACAG GCGGGGGGTG GCGCAGGCGG CAGCAACAGC
ACTTTCTACT GTGGCTTGTC CCGCCCCAC CGCGTCCGCC GTCGTTGTG

15651 AGTGGCAGCG GCGCGGAAGA GAACTCCAAC GCGGCAGCCG CGGCAATGCA
TCACCGTCGC CGCGCCTTCT CTTGAGGTTG CGCCGTCGGC GCCGTTACGT

15701 GCCGGTGGAG GACATGAACG ATCATGCCAT TCGCGGCGAC ACCTTTGCCA
CGGCCACCTC CTGTACTTGC TAGTACGGTA AGCGCCGCTG TGGAAACGGT

15751 CACGGGCTGA GGAGAAGCGC GCTGAGGCCG AAGCAGCGGC CGAAGCTGCC
GTGCCCCGACT CCTCTTCGCG CGACTCCGGC TTCGTCGCC GCTTCGACGG

15801 GCCCCCGCTG CGCAACCCGA GGTCGAGAAG CCTCAGAAGA AACCGGTGAT
CGGGGGCGAC GCGTTGGGCT CCAGCTCTTC GGAGTCTTCT TTGGCCACTA

15851 CAAACCCCTG ACAGAGGACA GCAAGAAACG CAGTTACAAC CTAATAAGCA
GTTTGGGGAC TGTCTCCTGT CGTTCTTTGC GTCAATGTTG GATTATTCGT

15901 ATGACAGCAC CTTACCCAG TACCGCAGCT GGTACCTTGC ATACAACTAC
TACTGTCGTG GAAGTGGGTC ATGGCGTCGA CCATGGAACG TATGTTGATG

15951 GGCGACCCTC AGACCGGAAT CCGCTCATGG ACCCTGCTTT GCACTCCTGA
CCGCTGGGAG TCTGGCCTTA GGCGAGTACC TGGGACGAAA CGTGAGGACT

16001 CGTAACCTGC GGCTCGGAGC AGGTCTACTG GTCGTTGCCA GACATGATGC
GCATTGGACG CCGAGCCTCG TCCAGATGAC CAGCAACGGT CTGTACTACG

16051 AAGACCCCGT GACCTTCCGC TCCACGCGCC AGATCAGCAA CTTTCCGGTG
TTCTGGGGCA CTGGAAGGCG AGGTGCGCGG TCTAGTCGTT GAAAGGCCAC

Figure 26 Q

16151 GGCCGTCTAC TCCCAACTCA TCCGCCAGTT TACCTCTCTG ACCCACGTGT
CCGGCAGATG AGGGTTGAGT AGGCGGTCAA ATGGAGAGAC TGGGTGCACA

16201 TCAATCGCTT TCCCGAGAAC CAGATTTTGG CGCGCCCGCC AGCCCCCACC
AGTTAGCGAA AGGGCTCTTG GTCTAAAACC GCGCGGGCGG TCGGGGGTGG

16251 ATCACCACCG TCAGTGAAAA CGTTCCTGCT CTCACAGATC ACGGGACGCT
TAGTGGTGGC AGTCACTTTT GCAAGGACGA GAGTGTCTAG TGCCCTGCGA

16301 ACCGCTGCGC AACAGCATCG GAGGAGTCCA GCGAGTGACC ATTACTGACG
TGGCGACGCG TTGTCGTAGC CTCCTCAGGT CGCTCACTGG TAATGACTGC

16351 CCAGACGCCG CACCTGCCCC TACGTTTACA AGGCCCTGGG CATAGTCTCG
GGTCTGCGGC GTGGACGGGG ATGCAAATGT TCCGGGACCC GTATCAGAGC

16401 CCGCGCGTCC TATCGAGCCG CACTTTTTGA GCAAGCATGT CCATCCTTAT
GGCGCGCAGG ATAGCTCGGC GTGAAAACT CGTTCGTACA GGTAGGAATA

16451 ATCGCCCAGC AATAACACAG GCTGGGGCCT GCGCTTCCCA AGCAAGATGT
TAGCGGGTCG TTATTGTGTC CGACCCCGGA CGCGAAGGGT TCGTTCTACA

16501 TTGGCGGGGC CAAGAAGCGC TCCGACCAAC ACCCAGTGCG CGTGCGCGGG
AACCGCCCCG GTTCTTCGCG AGGCTGGTTG TGGGTCACGC GCACGCGCCC

16551 CACTACCGCG CGCCCTGGGG CGCGCACAAA CGCGGCCGCA CTGGGCGCAC
GTGATGGCGC GCGGGACCCC GCGCGTGTTT GCGCCGGCGT GACCCGCGTG

16601 CACCGTCGAT GACGCCATCG ACGCGGTGGT GGAGGAGGCG CGCAACTACA
GTGGCAGCTA CTGCGGTAGC TGCGCCACCA CCTCCTCCGC GCGTTGATGT

16651 CGCCCACGCC GCCACCAGTG TCCACAGTGG ACGCGGCCAT TCAGACCGTG
GCGGGTGCGG CGGTGGTCAC AGGTGTCACC TGCGCCGGTA AGTCTGGCAC

16701 GTGCGCGGAG CCCGGCGCTA TGCTAAAATG AAGAGACGGC GGAGGCGCGT
CACGCGCCTC GGGCCGCGAT ACGATTTTAC TTCTCTGCCG CCTCCGCGCA

16751 AGCACGTCGC CACCGCCGCC GACCCGGCAC TGCCGCCCAA CGCGCGGCGG
TCGTGCAGCG GTGGCGGCGG CTGGGCGGTG ACGGCGGGTT GCGCGCCGCC

16801 CGGCCCTGCT TAACCGCGCA CGTCGCACCG GCCGACGGGC GGCCATGCGG
GCCGGGACGA ATTGGCGCGT GCAGCGTGGC CGGCTGCCCC CCGGTACGCC

16851 GCCGCTCGAA GGCTGGCCGC GGGTATTGTC ACTGTGCCCC CCAGGTCCAG
CGGCGAGCTT CCGACCGGCG CCCATAACAG TGACACGGGG GGTCCAGGTC

16901 GCGACGAGCG GCCGCCGAG CAGCCGCGGC CATTAGTGCT ATGACTCAGG
CGCTGCTCGC CGGCGGCGTC GTCGGCGCCG GTAATCACGA TACTGAGTCC

16951 GTCGCAGGGG CAACGTGTAT TGGGTGCGCG ACTCGGTTAG CGGCCTGCGC
CAGCGTCCCC GTTGACACATA ACCCACGCGC TGAGCCAATC GCCGGACGCG

17001 GTGCCCCTGC GCACCCGCCC CCCGCGCAAC TAGATTGCAA GAAAAAATA
CACGGGCACG CGTGGGCGGG GGGCGCGTTG ATCTAACGTT CTTTTTTGAT

Figure 26 R

17101 CTATGTCCAA GCGCAAAATC AAAGAAGAGA TGCTCCAGGT CATCGCGCCG
GATACAGGTT CGCGTTTTAG TTTCTTCTCT ACGAGGTCCA GTAGCGCGGC

17151 GAGATCTATG GCGGGGGCTT CTTCCTTCTC GTCCTAATGT TCGGGGCTTT

17201 GCTAAAGCGG GTCAAAAAGA AAAAGAAAGA TGATGATGAT GAACTTGACG
CGATTTTCGCC CAGTTTTTCT TTTTCTTTCT ACTACTACTA CTTGAACTGC

17251 ACGAGGTGGA ACTGCTGCAC GCTACCGCGC CCAGGCGACG GGTACAGTGG
TGCTCCACCT TGACGACGTG CGATGGCGCG GGTCCGCTGC CCATGTCACC

17301 AAAGGTCGAC GCGTAAAACG TGTTTTGCGA CCCGGCACCA CCGTAGTCTT
TTTCCAGCTG CGCATTTTGC ACAAACGCT GGGCCGTGGT GGCATCAGAA

17351 TACGCCCCGGT GAGCGCTCCA CCCGCACCTA CAAGCGCGTG TATGATGAGG
ATGCGGGCCA CTCGCGAGGT GGGCGTGGAT GTTCGCGCAC ATACTACTCC

17401 TGTACGGCGA CGAGGACCTG CTTGAGCAGG CCAACGAGCG CCTCGGGGAG
ACATGCCGCT GCTCCTGGAC GAACTCGTCC GGTGCTCGC GGAGCCCCTC

17451 TTTGCCTACG GAAAGCGGCA TAAGGACATG CTGGCGTTGC CGCTGGACGA
AAACGGATGC CTTTCGCCGT ATTCTGTAC GACCGCAACG GCGACCTGCT

17501 GGGCAACCCA ACACCTAGCC TAAAGCCCGT AACACTGCAG CAGGTGCTGC
CCCGTTGGGT TGTGGATCGG ATTTCCGGCA TTGTGACGTC GTCCACGACG

17551 CCGCGCTTGC ACCGTCCGAA GAAAAGCGCG GCCTAAAGCG CGAGTCTGGT
GGCGCGAACG TGGCAGGCTT CTTTTCGCGC CGGATTTCGC GCTCAGACCA

17601 GACTTGGCAC CCACCGTGCA GCTGATGGTA CCCAAGCGCC AGCGACTGGA
CTGAACCGTG GGTGGCACGT CGACTACCAT GGGTTCGCGG TCGCTGACCT

17651 AGATGTCTTG GAAAAAATGA CCGTGGAACC TGGGCTGGAG CCCGAGGTCC
TCTACAGAAC CTTTTTTACT GGCACCTTGG ACCCGACCTC GGGCTCCAGG

17701 GCGTGCGGCC AATCAAGCAG GTGGCGCCGG GACTGGGCGT GCAGACCGTG
CGCACGCCGG TTAGTTCGTC CACCGCGGCC CTGACCCGCA CGTCTGGCAC

17751 GACGTTTCAGA TACCCACTAC CAGTAGCACC AGTATTGCCA CCGCCACAGA
CTGCAAGTCT ATGGGTGATG GTCATCGTGG TCATAACGGT GCGGGTGTCT

17801 GGGCATGGAG ACACAAACGT CCCCAGTTGC CTCAGCGGTG GCGGATGCCG
CCCGTACCTC TGTGTTTGCA GGGGCCAACG GAGTCGCCAC CGCCTACGGC

17851 CGGTGCAGGC GGTCGCTGCG GCCGCGTCCA AGACCTCTAC GGAGGTGCAA
GCCACGTCCG CCAGCGACGC CGGCGCAGGT TCTGGAGATG CCTCCACGTT

17901 ACGGACCCGT GGATGTTTTCG CGTTTCAGCC CCCCAGCGCC CGCGCCGTTT
TGCCTGGGCA CCTACAAAGC GCAAAGTCGG GGGGCCGCGG GCGCGGCAAG

17951 GAGGAAGTAC GGCGCCGCCA GCGCGCTACT GCCCGAATAT GCCCTACATC
CTCCTTCATG CCGCGGCGGT CGCGCGATGA CGGGCTTATA CGGGATGTAG

Figure 26S

18051 AGACGAGCAA CTACCCGACG CCGAACCACC ACTGGAACCC GCCGCCGCCG
TCTGCTCGTT GATGGGCTGC GGCTTGGTGG TGACCTTGGG CGGCGGGCGG

18101 TCGCCGTCGC CAGCCCGTGC TGGCCCCGAT TTCCGTGCGC AGGGTGGCTC
AGCGGCAGCG GTCGGGCACG ACCGGGGCTA AAGGCACGCG TCCCACCGAG

18151 GCGAAGGAGG CAGGACCCTG GTGCTGCCAA CAGCGCGCTA CCACCCCAGC
CGCTTCCTCC GTCCTGGGAC CACGACGGTT GTCGCGCGAT GGTGGGGTGC

18201 ATCGTTTAAA AGCCGGTCTT TGTGGTTCTT GCAGATATGG CCCTCACCTG
TAGCAAATTT TCGGCCAGAA ACACCAAGAA CGTCTATACC GGGAGTGGAC

18251 CCGCCTCCGT TTCCCGGTGC CGGGATTCCG AGGAAGAATG CACCGTAGGA
GGCGGAGGCA AAGGGCCACG GCCCTAAGGC TCCTTCTTAC GTGGCATCCT

18301 GGGGCATGGC CGGCCACGGC CTGACGGGCG GCATGCGTCG TGCGCACCAC
CCCCGTACCG GCCGGTGCCG GACTGCCCGC CGTACGCAGC ACGCGTGGTG

18351 CGGCGGCGGC GCGCGTCGCA CCGTCGCATG CGCGGCGGTA TCCTGCCCCCT
GCCGCCGCCG CCGGCAGCGT GGCAGCGTAC GCGCCGCCAT AGGACGGGGA

18401 CCTTATTCCA CTGATCGCCG CGGCGATTGG CGCCGTGCCC GGAATTGCAT
GGAATAAGGT GACTAGCGGC GCCGCTAACC GCGGCACGGG CCTTAACGTA

18451 CCGTGGCCTT GCAGGCGCAG AGACACTGAT TAAAAACAAG TTGCATGTGG
GGCACCGGAA CGTCCGCGTC TCTGTGACTA ATTTTGTTC AACGTACACC

18501 AAAAATCAAA ATAAAAAGTC TGGACTCTCA CGCTCGCTTG GTCCTGTAAC
TTTTTAGTTT TATTTTTCAG ACCTGAGAGT GCGAGCGAAC CAGGACATTG

18551 TATTTGTAG AATGGAAGAC ATCAACTTTG CGTCTCTGGC CCCGCGACAC
ATAAACATC TTACCTTCTG TAGTTGAAAC GCAGAGACCG GGGCGCTGTG

18601 GGCTCGCGCC CGTTCATGGG AAAGTGGCAA GATATCGGCA CCAGCAATAT
CCGAGCGCGG GCAAGTACCC TTTGACCGTT CTATAGCCGT GGTCGTTATA

18651 GAGCGGTGGC GCCTTCAGCT GGGGCTCGCT GTGGAGCGGC ATTAAAAATT
CTCGCCACCG CGGAAGTCGA CCCCAGCGCA CACCTCGCCG TAATTTTTAA

18701 TCGGTTCCAC CGTTAAGAAC TATGGCAGCA AGGCCTGGAA CAGCAGCACA
AGCCAAGGTG GCAATTCTTG ATACCGTCGT TCCGGACCTT GTCGTCGTGT

18751 GGCCAGATGC TGAGGGATAA GTTGAAAGAG CAAAATTTCC AACAAAAGGT
CCGGTCTACG ACTCCCTATT CAACTTTCTC GTTTTAAAGG TTGTTTTCCA

18801 GGTAAGATGC CTGGCCTCTG GCATTAGCGG GGTGGTGGAC CTGGCCAACC
CCATCTACCG GACCGGAGAC CGTAATCGCC CCACCACCTG GACCGGTGG

18851 AGGCAGTGCA AAATAAGATT AACAGTAAGC TTGATCCCCG CCCTCCCGTA
TCCGTCACGT TTTATTCTAA TTGTCATTCT AACTAGGGGC GGGAGGGCAT

18901 GAGGAGCCTC CACCGGCCGT GGAGACAGTG TCTCCAGAGG GGCGTGGCGA
CTCCTCGGAG GTGGCCGGCA CCTCTGTCAC AGAGGTCTCC CCGCACCGCT

Figure 26 T

19001 AGCCTCCCTC GTACGAGGAG GCACTAAAGC AAGGCCTGCC CACCACCCGT
TCGGAGGGAG CATGCTCCTC CGTGATTTCG TTCCGGACGG GTGGTGGGCA

19051 CCCATCGCGC CCATGGCTAC CGGAGTGCTG GGCCAGCACA CACCCGTAAC
GGGTAGCGCG GGTACCGATG GCCTCACGAC CCGGTCGTGT GTGGGCATTG

19101 GCTGGACCTG CCTCCCCCG CCGACACCCA GCAGAAACCT GTGCTGCCAG
CGACCTGGAC GGAGGGGGGC GGCTGTGGGT CGTCTTTGGA CACGACGGTC

19151 GCCCGACCGC CGTTGTTGTA ACCCGTCCTA GCCGCGCGTC CCTGCGCCGC
CGGGCTGGCG GCAACAACAT TGGGCAGGAT CGGCGCGCAG GGACGCGGCG

19201 GCCGCCAGCG GTCCGCGATC GTTGCGGCCC GTAGCCAGTG GCAACTGGCA
CGGCGGTCGC CAGGCGCTAG CAACGCCGGG CATCGGTCAC CGTTGACCGT

19251 AAGCACACTG AACAGCATCG TGGGTCTGGG GGTGCAATCC CTGAAGCGCC
TTCGTGTGAC TTGTCGTAGC ACCCAGACCC CCACGTTAGG GACTTCGCGG

19301 GACGATGCTT CTGATAGCTA ACGTGTCGTA TGTGTGTCAT GTATGCGTCC
CTGCTACGAA GACTATCGAT TGCACAGCAT ACACACAGTA CATA CGCAGG

19351 ATGTCGCCGC CAGAGGAGCT GCTGAGCCGC CGCGCGCCCG CTTTCCAAGA
TACAGCGGCG GTCTCCTCGA CGACTCGGCG GCGCGCGGGC GAAAGGTTCT

19401 TGGCTACCCC TTCGATGATG CCGCAGTGGT CTTACATGCA CATCTCGGGC
ACCGATGGGG AAGCTACTAC GGCCTCACCA GAATGTACGT GTAGAGCCCC

19451 CAGGACGCCT CGGAGTACCT GAGCCCCGGG CTGGTGCAGT TTGCCCCGCG
GTCCTGCGGA GCCTCATGGA CTCGGGGGCC GACCACGTCA AACGGGCGCG

19501 CACCGAGACG TACTTCAGCC TGAATAACAA GTTTAGAAAC CCCACGGTGG
GTGGCTCTGC ATGAAGTCGG ACTTATTGTT CAAATCTTTG GGGTGCCACC

19551 CGCCTACGCA CGACGTGACC ACAGACCGGT CCCAGCGTTT GACGCTGCGG
GCGGATGCGT GCTGCACTGG TGTCTGGCCA GGGTCGCAA CTGCGACGCC

19601 TTCATCCCTG TGGACCGTGA GGATACTGCG TACTCGTACA AGGCGCGGTT
AAGTAGGGAC ACCTGGCACT CCTATGACGC ATGAGCATGT TCCGCGCCAA

19651 CACCCTAGCT GTGGGTGATA ACCGTGTGCT GGACATGGCT TCCACGTACT
GTGGGATCGA CACCCACTAT TGGCACACGA CCTGTACCGA AGGTGCATGA

19701 TTGACATCCG CGGCGTGCTG GACAGGGGCC CTACTTTTAA GCCCTACTCT
AACTGTAGGC GCCGCACGAC CTGTCCCCGG GATGAAAATT CGGGATGAGA

19751 GGCCTGCCT ACAACGCCCT GGCTCCCAAG GGTGCCCCAA ATCCTTGCGA
CCGTGACGGA TGTTCGCGGA CCGAGGGTTC CCACGGGGTT TAGGAACGCT

19801 ATGGGATGAA GCTGCTACTG CTCTTGAAAT AAACCTAGAA GAAGAGGACG
TACCCTACTT CGACGATGAC GAGAACTTTA TTTGGATCTT CTTCTCCTGC

19851 ATGACAACGA AGACGAAGTA GACGAGCAAG CTGAGCAGCA AAAA ACTCAC
TACTGTTGCT TCTGCTTCAT CTGCTCGTTC GACTCGTCGT TTTT TGAGTG

Figure 26 U

```

19951 TCAAATAGGT GTCGAAGGTC AAACACCTAA ATATGCCGAT AAAACATTTT
      AGTTTATCCA CAGCTTCCAG TTTGTGGATT TATACGGCTA TTTTGTAAAG

20001 AACCTGAACC TCAAATAGGA GAATCTCAGT GGTACGAAAC AGAAATTAAT
      TTGGACTTGG AGTTTATCCT CTTAGAGTCA CCATGCTTTG TCTTTAATTA

20051 CATGCAGCTG GGAGAGTCCT AAAAAAGACT ACCCCAATGA AACCATGTTA
      GTACGTCGAC CCTCTCAGGA TTTTTTCTGA TGGGGTTACT TTGGTACAAT

20101 CGGTTCATAT GCAAAACCCA CAAATGAAAA TGGAGGGCAA GGCATTCTTG
      GCCAAGTATA CGTTTTGGGT GTTTACTTTT ACCTCCCGTT CCGTAAGAAC

20151 TAAAGCAACA AAATGGAAAG CTAGAAAGTC AAGTGGAAAT GCAATTTTTT
      ATTCGTTGT TTTACCTTTC GATCTTTCAG TTCACCTTTA CGTTAAAAAG

20201 TCAACTACTG AGGCAGCCGC AGGCAATGGT GATAACTTGA CTCCTAAAGT
      AGTTGATGAC TCCGTCGGCG TCCGTTACCA CTATTGAACT GAGGATTTCA

20251 GGTATTGTAC AGTGAAGATG TAGATATAGA AACCCCAGAC ACTCATATTT
      CCATAACATG TCACTTCTAC ATCTATATCT TTGGGGTCTG TGAGTATAAA

20301 CTTACATGCC CACTATTAAG GAAGGTAAct CACGAGAACT AATGGGCCAA
      GAATGTACGG GTGATAATTC CTTCCATTGA GTGCTCTTGA TTACCCGGTT

20351 CAATCTATGC CCAACAGGCC TAATTACATT GCTTTTAGGG ACAATTTTAT
      GTTAGATACG GGTGTGCCG ATTAATGTAA CGAAAATCCC TGTTAAAATA

20401 TGGTCTAATG TATTACAACA GCACGGGTAA TATGGGTGTT CTGGCGGGCC
      ACCAGATTAC ATAATGTTGT CGTGCCCAT ATACCCACAA GACCGCCCGG

20451 AAGCATCGCA GTTGAATGCT GTTGTAGATT TGCAAGACAG AAACACAGAG
      TTCGTAGCGT CAACTTACGA CAACATCTAA ACGTTCTGTC TTTGTGTCTC

20501 CTTTCATACC AGCTTTTGCT TGATTCCATT GGTGATAGAA CCAGGTACTT
      GAAAGTATGG TCGAAAACGA ACTAAGGTAA CCACTATCTT GGTCCATGAA

20551 TTCTATGTGG AATCAGGCTG TTGACAGCTA TGATCCAGAT GTTAGAATTA
      AAGATACACC TTAGTCCGAC AACTGTCGAT ACTAGGTCTA CAATCTTAAT

20601 TTGAAAATCA TGGAAGTGA GATGAACCTC CAAATTACTG CTTTCCACTG
      AACTTTTAGT ACCTTGACTT CTACTTGAAG GTTTAATGAC GAAAGGTGAC

20651 GGAGGTGTGA TTAATACAGA GACTCTTACC AAGGTAAAAC CTAAAACAGG
      CCTCCACACT AATTATGTCT CTGAGAAATG TTCCATTTTG GATTTTGTCC

20701 TCAGGAAAAT GGATGGGAAA AAGATGCTAC AGAATTTTCA GATAAAAATG
      AGTCCTTTTA CCTACCCTTT TTCTACGATG TCTTAAAAGT CTATTTTAC

20751 AAATAAGAGT TGGAATAAAT TTTGCCATGG AAATCAATCT AAATGCCAAC
      TTTATTCTCA ACCTTTATTA AAACGGTACC TTTAGTTAGA TTTACGGTTG

20801 CTGTGGAGAA ATTCCTGTGA CTCCAACATA GCGCTGTATT TGCCCGACAA
      GACACCTCTT TAAAGGACAT GAGGTGTAT CGCGACATAA ACGGGCTGTT

```

Figure 26 v

20901 ACGACTACAT GAACAAGCGA GTGGTGGCTC CCGGGCTAGT GGACTGCTAC
TGCTGATGTA CTTGTTCGCT CACCACCGAG GGCCCGATCA CCTGACGATG

20951 ATTAACCTTG GAGCACGCTG GTCCCTTGAC TATATGGACA ACGTCAACCC
TAATTGGAAC CTCGTGCGAC CAGGGAAC TG ATATACCTGT TGCAGTTGGG

21001 ATTTAACCAC CACCGCAATG CTGGCCTGCG CTACCGCTCA ATGTTGCTGG
TAAATTGGTG GTGGCGTTAC GACCGGACGC GATGGCGAGT TACAACGACC

21051 GCAATGGTCG CTATGTGCCC TTCCACATCC AGGTGCCTCA GAAGTTCTTT
CGTTACCAGC GATACACGGG AAGGTGTAGG TCCACGGAGT CTTCAAGAAA

21101 GCCATTAAAA ACCTCCTTCT CCTGCCGGGC TCATACACCT ACGAGTGGAA
CGGTAATTTT TGGAGGAAGA GGACGGCCCG AGTATGTGGA TGCTCACCTT

21151 CTTCAGGAAG GATGTTAACA TGGTTCTGCA GAGCTCCCTA GGAAATGACC
GAAGTCCTTC CTACAATTGT ACCAAGACGT CTCGAGGGAT CCTTTACTGG

21201 TAAGGGTTGA CGGAGCCAGC ATTAAGTTTG ATAGCATTTG CCTTTACGCC
ATTCCCAACT GCCTCGGTCG TAATTCAAAC TATCGTAAAC GGAAATGCGG

21251 ACCTTCTTCC CCATGGCCCA CAACACCGCC TCCACGCTTG AGGCCATGCT
TGGAAGAAGG GGTACCGGGT GTTGTGGCGG AGGTGCGAAC TCCGGTACGA

21301 TAGAAACGAC ACCAACGACC AGTCCTTTAA CGACTATCTC TCCGCCGCCA
ATCTTTGCTG TGGTTGCTGG TCAGGAAATT GCTGATAGAG AGGCGGCGGT

21351 ACATGCTCTA CCCTATACCC GCCAACGCTA CCAACGTGCC CATATCCATC
TGTACGAGAT GGGATATGGG CGGTTGCGAT GGTGTCACGG GTATAGGTAG

21401 CCCTCCCGCA ACTGGGCGGC TTTCCGCGGC TGGGCCTTCA CGCGCCTTAA
GGGAGGGCGT TGACCCGCCG AAAGGCGCCG ACCCGGAAGT GCGCGGAATT

21451 GACTAAGGAA ACCCCATCAC TGGGCTCGGG CTACGACCCT TATTACACCT
CTGATTCCTT TGGGGTAGTG ACCCGAGCCC GATGCTGGGA ATAATGTGGA

21501 ACTCTGGCTC TATACCCTAC CTAGATGGAA CCTTTTACCT CAACCACACC
TGAGACCGAG ATATGGGATG GATCTACCTT GGAAAATGGA GTTGGTGTGG

21551 TTTAAGAAGG TGGCCATTAC CTTTGACTCT TCTGTCAGCT GGCCTGGCAA
AAATTCTTCC ACCGGTAATG GAAACTGAGA AGACAGTCGA CCGGACCGTT

21601 TGACCGCCTG CTTACCCCCA ACGAGTTTGA AATTAAGCGC TCAGTTGACG
ACTGGCGGAC GAATGGGGGT TGCTCAAAC TTAATTCGCG AGTCAACTGC

21651 GGGAGGGTTA CAACGTTGCC CAGTGTAACA TGACCAAAGA CTGGTTCCTG
CCCTCCCAAT GTTGCAACGG GTCACATTGT ACTGGTTTCT GACCAAGGAC

21701 GTACAAATGC TAGCTAACTA TAACATTGGC TACCAGGGCT TCTATATCCC
CATGTTTACG ATCGATTGAT ATTGTAACCG ATGGTCCCGA AGATATAGGG

21751 AGAGAGCTAC AAGGACCGCA TGTACTCCTT CTTTAGAAAC TTCCAGCCCA
TCTCTCGATG TTCCTGGCGT ACATGAGGAA GAAATCTTTG AAGGTCGGGT

Figure 26 W

21851 GGCATCCTAC ACCAACACAA CAACTCTGGA TTTGTTGGCT ACCTTGCCCC
CCGTAGGATG TGGTTGTGTT GTTGAGACCT AAACAACCGA TGGAACGGGG

21901 CACCATGCGC GAAGGACAGG CCTACCCTGC TAACTTCCCC TATCCGCTTA
GTGGTACGCG CTCCTGTCC GGATGGGACG ATTGAAGGGG ATAGGCGAAT

21951 TAGGCAAGAC CGCAGTTGAC AGCATTACCC AGAAAAAGTT TCTTTGCGAT
ATCCGTTCTG GCGTCAACTG TCGTAATGGG TCTTTTTCAA AGAAACGCTA

22001 CGCACCCCTTT GGCGCATCCC ATTCTCCAGT AACTTTATGT CCATGGGCGC
GCGTGGGAAA CCGCGTAGGG TAAGAGGTCA TTGAAATACA GGTACCCGCG

22051 ACTCACAGAC CTGGGCCAAA ACCTTCTCTA CGCCAACTCC GCCACGCGC
TGAGTGTCTG GACCCGTTT TGAAGAGAT GCGGTTGAGG CGGGTGCGCG

22101 TAGACATGAC TTTTGAGGTG GATCCCATGG ACGAGCCCAC CCTTCTTTAT
ATCTGTACTG AAAACTCCAC CTAGGGTACC TGCTCGGGTG GGAAGAAATA

22151 GTTTTGTTTG AAGTCTTTGA CGTGGTCCGT GTGCACCAGC CGCACCGCGG
CAAACAAAC TTCAGAACT GCACCAGGCA CACGTGGTCG GCGTGGCGCC

22201 CGTCATCGAA ACCGTGTACC TGCGCACGCC CTTCTCGGCC GGCAACGCCA
GCAGTAGCTT TGGCACATGG ACGCGTGCGG GAAGAGCCGG CCGTTGCGGT

22251 CAACATAAAG AAGCAAGCAA CATCAACAAC AGCTGCCGCC ATGGGCTCCA
GTTGTATTTC TTCGTTCGTT GTAGTTGTTG TCGACGGCGG TACCCGAGGT

22301 GTGAGCAGGA ACTGAAAGCC ATTGTCAAAG ATCTTG GTTG TGGGCCATAT
CACTCGTCCT TGACTTTCGG TAACAGTTTC TAGAACCAAC ACCCGGTATA

22351 TTTTGGGCA CCTATGACAA GCGCTTTCCA GGCTTTGTTT CTCCACACAA
AAAAACCCGT GGATACTGTT CGCGAAAGGT CCGAAACAA GAGGTGTGTT

22401 GCTCGCCTGC GCCATAGTCA ATACGGCCGG TCGCGAGACT GGGGGCGTAC
CGAGCGGACG CCGTATCAGT TATGCCGGCC AGCGCTCTGA CCCCCGCATG

22451 ACTGGATGGC CTTTGCCTGG AACCCGCACT CAAAAACATG CTACCTCTTT
TGACCTACCG GAAACGGACC TTGGGCGTGA GTTTTGTAC GATGGAGAAA

22501 GAGCCCTTTG GCTTTTCTGA CCAGCGACTC AAGCAGGTTT ACCAGTTTGA
CTCGGGAAAC CGAAAAGACT GGTGCTGAG TTCGTCCAA TGGTCAAAC

22551 GTACGAGTCA CTCCTGCGCC GTAGCGCCAT TGCTTCTTCC CCCGACCGCT
CATGCTCAGT GAGGACGCGG CATCGCGGTA ACGAAGAAGG GGGCTGGCGA

22601 GTATAACGCT GGAAAAGTCC ACCCAAAGCG TACAGGGGCC CAACTCGGCC
CATATTGCGA CCTTTTCAGG TGGGTTTCGC ATGTCCCCGG GTTGAGCCGG

22651 GCCTGTGGAC TATTCTGCTG CATGTTTCTC CACGCCTTTG CCAACTGGCC
CGGACACCTG ATAAGACGAC GTACAAAGAG GTGCGGAAAC GGTGACCGG

22701 CCAAACCTCC ATGGATCACA ACCCCACCAT GAACCTTATT ACCGGGGTAC
GGTTTGAGGG TACCTAGTGT TGGGGTGGTA CTTGGAATAA TGGCCCCATG

Figure 26 X

22801 CAGGAACAGC TCTACAGCTT CCTGGAGCGC CACTCGCCCT ACTTCGCGAG
GTCCTTGTCG AGATGTCGAA GGACCTCGCG GTGAGCGGGA TGAAGGCGTC

22851 CCACAGTGCG CAGATTAGGA GCGCCACTTC TTTTGTGCAC TTGAAAAACA
GGTGTACGCG GTCTAATCCT CGCGGTGAAG AAAAACAGTG AACTTTTTGT

22901 TGTAAAAATA ATGTACTAGA GACACTTTCA ATAAAGGCAA ATGCTTTTAT
ACATTTTAT TACATGATCT CTGTGAAAGT TATTTCCGTT TACGAAAATA

22951 TTGTACACTC TCGGGTGATT ATTTACCCCC ACCCTTGCCG TCTGCGCCGT
AACATGTGAG AGCCCACTAA TAAATGGGGG TGGGAACGGC AGACGCGGCA

23001 TTAAAAATCA AAGGGGTTCT GCCGCGCATC GCTATGCGCC ACTGGCAGGG
AATTTTGTAGT TTCCCAAGA CGGCGCGTAG CGATACGCGG TGACCGTCCC

23051 ACACGTTGCG ATACTGGTGT TTAGTGCTCC ACTTAAACTC AGGCACAACC
TGTGCAACGC TATGACCACA AATCACGAGG TGAATTTGAG TCCGTGTTGG

23101 ATCCGCGGCA GCTCGGTGAA GTTTTCACTC CACAGGCTGC GCACCATCAC
TAGGCGCCGT CGAGCCACTT CAAAAGTGAG GTGTCCGACG CGTGGTAGTG

23151 CAACGCGTTT AGCAGGTCGG GCGCCGATAT CTTGAAGTCG CAGTTGGGGC
GTTGCGCAA TCGTCCAGCC CGCGGCTATA GAACTTCAGC GTCAACCCCG

23201 CTCCGCCCTG CGCGCGCGAG TTGCGATACA CAGGGTTGCA GCACTGGAAC
GAGGCGGGAC GCGCGCGCTC AACGCTATGT GTCCCAACGT CGTGACCTTG

23251 ACTATCAGCG CCGGGTGGTG CACGCTGGCC AGCACGCTCT TGTCGGAGAT
TGATAGTCGC GGCCACCAC GTGCGACCGG TCGTGCGAGA ACAGCCTCTA

23301 CAGATCCGCG TCCAGGTCCT CCGCGTTGCT CAGGGCGAAC GGAGTCAACT
GTCTAGGCGC AGGTCCAGGA GGCGCAACGA GTCCCGCTTG CCTCAGTTGA

23351 TTGGTAGCTG CCTTCCCAA AAGGGCGCGT GCCCAGGCTT TGAGTTGCAC
AACCATCGAC GGAAGGGTTT TTCCCGCGCA CGGGTCCGAA ACTCAACGTG

23401 TCGCACCGTA GTGGCATCAA AAGGTGACCG TGCCCGGTCT GGGCGTTAGG
AGCGTGGCAT CACCGTAGTT TTCCACTGGC ACGGGCCAGA CCCGCAATCC

23451 ATACAGCGCC TGCATAAAAG CCTTGATCTG CTTAAAAGCC ACCTGAGCCT
TATGTCGCGG ACGTATTTTC GGAAGTAGAC GAATTTTCGG TGGACTCGGA

23501 TTGCGCCTTC AGAGAAGAAC ATGCCGCAAG ACTTGCCGGA AACTGATTG
AACGCGGAAG TCTCTTCTTG TACGGCGTTC TGAACGGCCT TTTGACTAAC

23551 GCCGGACAGG CCGCGTCGTG CACGCAGCAC CTTGCGTCGG TGTTGGAGAT
CGGCCTGTCC GGCGCAGCAC GTGCGTCGTG GAACGCAGCC ACAACCTCTA

23601 CTGCACCACA TTTCGGCCCC ACCGGTTCTT CACGATCTTG GCCTTGCTAG
GACGTGGTGT AAAGCCGGGG TGGCCAAGAA GTGCTAGAAC CGGAACGATC

23651 ACTGCTCCTT CAGCGCGCGC TGCCCGTTTT CGCTCGTCAC ATCCATTTCA
TGACGAGGAA GTCGCGCGCG ACGGGCAAAA GCGAGCAGTG TAGGTAAAGT

Figure 26 Y

WO 02/022080

PCT/US01/28861

| | | | | | |
|-------|------------|------------|------------|-------------|-------------|
| 23701 | ATCACGTGCT | CCTTATTTAT | CATAATGCTT | CCGTGTAGAC | ACTTAAGCTC |
| | TAGTGCACGA | GGAATAAATA | GTATTACGAA | GGCACATCTG | TGAATTTCGAG |
| 23751 | GCCTTCGATC | TCAGCGCAGC | GGTGCAGCCA | CAACGCGCAG | CCCGTGGGCT |
| | CGGAAGCTAG | AGTCGCGTCG | CCACGTCGGT | GTTGCGCGTC | GGGCACCCGA |
| 23801 | CGTGATGCTT | GTAGGTCACC | TCTGCAAACG | ACTGCAGGTA | CGCCTGCAGG |
| | GCACTACGAA | CATCCAGTGG | AGACGTTTGC | TGACGTCCAT | GCGGACGTCC |
| 23851 | AATCGCCCCA | TCATCGTCAC | AAAGGTCTTG | TTGCTGGTGA | AGGTCAGCTG |
| | TTAGCGGGGT | AGTAGCAGTG | TTTCCAGAAC | AACGACCACT | TCCAGTCGAC |
| 23901 | CAACCCGCGG | TGCTCCTCGT | TCAGCCAGGT | CTTGCATACG | GCCGCCAGAG |
| | GTTGGGCGCC | ACGAGGAGCA | AGTCGGTCCA | GAACGTATGC | CGGCGGTCTC |
| 23951 | CTTCCACTTG | GTCAGGCAGT | AGTTTGAAGT | TCGCCTTTAG | ATCGTTATCC |
| | GAAGGTGAAC | CAGTCCGTCA | TCAAACCTCA | AGCGGAAATC | TAGCAATAGG |
| 24001 | ACGTGGTACT | TGTCCATCAG | CGCGCGCGCA | GCCTCCATGC | CCTTCTCCCA |
| | TGCACCATGA | ACAGGTAGTC | GCGCGCGCGT | CGGAGGTACG | GGAAGAGGGT |
| 24051 | CGCAGACACG | ATCGGCACAC | TCAGCGGGTT | CATCACCGTA | ATTTCACTTT |
| | GCGTCTGTGC | TAGCCGTGTG | AGTCGCCCAA | GTAGTGGCAT | TAAAGTGAAA |
| 24101 | CCGCTTCGCT | GGGCTCTTCC | TCTTCCTCTT | GCGTCCGCAT | ACCACGCGCC |
| | GGCGAAGCGA | CCCGAGAAGG | AGAAGGAGAA | CGCAGGCGTA | TGGTGCGCGG |
| 24151 | ACTGGGTCGT | CTTCATTCAG | CCGCCGCACT | GTGCGCTTAC | CTCCTTTGCC |
| | TGACCCAGCA | GAAGTAAGTC | GGCGGCGTGA | CACGCGAATG | GAGGAAACGG |
| 24201 | ATGCTTGATT | AGCACCGGTG | GGTTGCTGAA | ACCCACCATT | TGTAGCGCCA |
| | TACGAATAA | TCGTGGCCAC | CCAACGACTT | TGGGTGGTAA | ACATCGCGGT |
| 24251 | CATCTTCTCT | TTCTTCCTCG | CTGTCCACGA | TTACCTCTGG | TGATGGCGGG |
| | GTAGAAGAGA | AAGAAGGAGC | GACAGGTGCT | AATGGAGACC | ACTACCGCCC |
| 24301 | CGCTCGGGCT | TGGGAGAAGG | GCGCTTCTTT | TTCTTCTTGG | GCGCAATGGC |
| | GCGAGCCCGA | ACCCTCTTCC | CGCGAAGAAA | AAGAAGAACC | CGCGTTACCG |
| 24351 | CAAATCCGCC | GCCGAGGTCT | ATGGCCGCGG | GCTGGGTGTG | CGCGGCACCA |
| | GTTTAGGCGG | CGGCTCCAGC | TACCGGCGCC | CGACCCACAC | GCGCCGTGGT |
| 24401 | GCGCGTCTTG | TGATGAGTCT | TCCTCGTCCT | CGGACTCGAT | ACGCCGCCTC |
| | CGCGCAGAAC | ACTACTCAGA | AGGAGCAGGA | GCCTGAGCTA | TGCGGCGGAG |
| 24451 | ATCCGCTTTT | TTGGGGGCGC | CCGGGGAGGC | GGCGGCGACG | GGGACGGGGA |
| | TAGGCGAAAA | AACCCCGCGC | GGCCCCTCCG | CCGCCGCTGC | CCCTGCCCTT |
| 24501 | CGACACGTCC | TCCATGGTTG | GGGGACGTCG | CGCCGCACCG | CGTCCGCGCT |
| | GCTGTGCAGG | AGGTACCAAC | CCCCTGCAGC | GCGGCGTGGC | GCAGGCGCGA |
| 24551 | CGGGGGTGGT | TTGCGGCTGC | TCCTCTTCCC | GA CTGGCCAT | TTCCTTCTCC |
| | GCCCCCAGCA | AAGCGCGACG | AGGAGAAGGG | CTGACCGGTA | AAGGAAGAGG |
| 24601 | TATAGGCAGA | AAAAGATCAT | GGAGTCAGTC | GAGAAGAAGG | ACAGCCTAAC |
| | ATATCCGTCT | TTTTCTAGTA | CCTCAGTCAG | CTCTTCTTCC | TGTCGGATTG |

24701 CTACCACCTT CCCCGTCGAG GCACCCCCGC TTGAGGAGGA GGAAGTGATT
GATGGTGGAA GGGGCAGCTC CGTGGGGGCG AACTCCTCCT CCTTCACTAA

24751 ATCGAGCAGG ACCCAGGTTT TGTAAGCGAA GACGACGAGG ACCGCTCAGT
TAGCTCGTCC TGGGTCCAAA ACATTGCTT CTGCTGCTCC TGGCGAGTCA

24801 ACCAACAGAG GATAAAAAGC AAGACCAGGA CAACGCAGAG GCAAACGAGG
TGGTTGTCTC CTATTTTTCG TTCTGGTCCT GTTGCGTCTC CGTTTGCTCC

24851 AACAAGTCGG GCGGGGGGAC GAAAGGCATG GCGACTACCT AGATGTGGGA
TTGTTTCAGCC CGCCCCCTG CTTTCCGTAC CGCTGATGGA TCTACACCCT

24901 GACGACGTGC TGTGTAAGCA TCTGCAGCGC CAGTGCGCCA TTATCTGCGA
CTGCTGCACG ACAACTTCGT AGACGTCGCG GTCACGCGGT AATAGACGCT

24951 CGCGTTGCAA GAGCGCAGCG ATGTGCCCCT CGCCATAGCG GATGTCAGCC
GCGCAACGTT CTCGCGTCGC TACACGGGGA GCGGTATCGC CTACAGTCGG

25001 TTGCCTACGA ACGCCACCTA TTCTCACC GC GTACCCCC CAAACGCCAA
AACGGATGCT TGCGGTGGAT AAGAGTGGCG CGCATGGGGG GTTTGCGGTT

25051 GAAAACGGCA CATGCGAGCC CAACCCGCGC CTCAACTTCT ACCCCGTATT
CTTTTGCCGT GTACGCTCGG GTTGGGCGCG GAGTTGAAGA TGGGGCATAA

25101 TGCCGTGCCA GAGGTGCTTG CCACCTATCA CATCTTTTTC CAAAACTGCA
ACGGCACGGT CTCCACGAAC GGTGGATAGT GTAGAAAAAG GTTTTGACGT

25151 AGATACCCCT ATCCTGCCGT GCCAACCGCA GCCGAGCGGA CAAGCAGCTG
TCTATGGGGA TAGGACGGCA CGGTGCGCT CGGCTCGCCT GTTCGTCGAC

25201 GCCTTGCGGC AGGGCGCTGT CATACTGAT ATCGCCTCGC TCAACGAAGT
CGGAACGCCG TCCCGCGACA GTATGGACTA TAGCGGAGCG AGTTGCTTCA

25251 GCCAAAAATC TTTGAGGGTC TTGGACGCGA CGAGAAGCGC GCGGCAAACG
CGGTTTTTAG AAACCTCCAG AACCTGCGCT GCTCTTCGCG CGCCGTTTGC

25301 CTCTGCAACA GGAAAACAGC GAAAATGAAA GTCACTCTGG AGTGTTGGTG
GAGACGTTGT CCTTTTGTCT CTTTACTTT CAGTGAGACC TCACAACCAC

25351 GAACTCGAGG GTGACAACGC GCGCCTAGCC GTACTAAAC GCAGCATCGA
CTTGAGCTCC CACTGTTGCG CGCGGATCGG CATGATTTTG CGTCGTAGCT

25401 GGTCACCCAC TTTGCCTACC CGGCACTTAA CCTACCCCCC AAGGTCATGA
CCAGTGGGTG AAACGGATGG GCCGTGAATT GGATGGGGGG TTCCAGTACT

25451 GCACAGTCAT GAGTGAGCTG ATCGTGCGCC GTGCGCAGCC CCTGGAGAGG
CGTGTCAGTA CTCCTCGAC TAGCACGCGG CACGCGTCGG GGACCTCTCC

25501 GATGCAAATT TGCAAGAACA AACAGAGGAG GGCCTACCCG CAGTTGGCGA
CTACGTTTAA ACGTTCTTGT TTGTCTCCTC CCGGATGGGC GTCAACCGCT

25551 CGAGCAGCTA GCGCGCTGGC TTCAAACGCG CGAGCCTGCC GACTTGGAGG
GCTCGTCGAT CGCGCGACCG AAGTTTGCGC GCTCGGACGG CTGAACCTCC

Figure 26 AA

25651 TGCATGCAGC GGTTCCTTGC TGACCCGGAG ATGCAGCGCA AGCTAGAGGA
ACGTACGTCG CCAAGAAACG ACTGGGCTC TACGTCGCGT TCGATCTCCT

25701 AACATTGCAC TACACCTTTC GACAGGGCTA CGTACGCCAG GCCTGCAAGA
TTGTAACGTG ATGTGGAAAG CTGTCCCGAT GCATGCGGTC CGGACGTTCT

25751 TCTCCAACGT GGAGCTCTGC AACCTGGTCT CCTACCTTGG AATTTTGCAC
AGAGGTTGCA CCTCGAGACG TTGGACCAGA GGATGGAACC TTAAAACGTG

25801 GAAAACCGCC TTGGGCAAAA CGTGCTTCAT TCCACGCTCA AGGGCGAGGC
CTTTTGGCGG AACCCGTTTT GCACGAAGTA AGGTGCGAGT TCCCGCTCCG

25851 GCGCCGCGAC TACGTCCGCG ACTGCGTTTA CTTATTTCTA TGCTACACCT
CGCGGCGCTG ATGCAGGCGC TGACGCAAAT GAATAAAGAT ACGATGTGGA

25901 GGCAGACGGC CATGGGCGTT TGGCAGCAGT GCTTGGAGGA GTGCAACCTC
CCGTCTGCCG GTACCCGCAA ACCGTCGTCA CGAACCTCCT CACGTTGGAG

25951 AAGGAGCTGC AGAAACTGCT AAAGCAAAC TTGAAGGACC TATGGACGGC
TTCCTCGACG TCTTTGACGA TTTCGTTTTG AACTTCCTGG ATACCTGCCG

26001 CTTCAACGAG CGCTCCGTGG CCGCGCACCT GCGGACATC ATTTTCCCCG
GAAGTTGCTC GCGAGGCACC GGCGCGTGGA CCGCCTGTAG TAAAAGGGGC

26051 AACGCCTGCT TAAAACCCTG CAACAGGGTC TGCCAGACTT CACCAGTCAA
TTGCGGACGA ATTTTGGGAC GTTGTCCCAG ACGGTCTGAA GTGGTCAGTT

26101 AGCATGTTGC AGAACTTTAG GAACTTTATC CTAGAGCGCT CAGGAATCTT
TCGTACAACG TCTTGAAATC CTTGAAATAG GATCTCGCGA GTCCTTAGAA

26151 GCCCGCCACC TGCTGTGCAC TTCCTAGCGA CTTTGTGCCC ATTAAGTACC
CGGGCGGTGG ACGACACGTG AAGGATCGCT GAAACACGGG TAATTCATGG

26201 GCGAATGCCC TCCGCCGCTT TGGGGCCACT GCTACCTTCT GCAGCTAGCC
CGCTTACGGG AGGCGGCGAA ACCCCGGTGA CGATGGAAGA CGTCGATCGG

26251 AACTACCTTG CCTACCACTC TGACATAATG GAAGACGTGA GCGGTGACGG
TTGATGGAAC GGATGGTGAG ACTGTATTAC CTTCTGCACT CGCCACTGCC

26301 TCTACTGGAG TGTCACGTGC GCTGCAACCT ATGCACCCCG CACCGCTCCC
AGATGACCTC ACAGTGACAG CGACGTTGGA TACGTGGGGC GTGGCGAGGG

26351 TGGTTTGCAA TTCGCAGCTG CTTAACGAAA GTCAAATTAT CGGTACCTTT
ACCAAACGTT AAGCGTCGAC GAATTGCTTT CAGTTTAATA GCCATGGAAA

26401 GAGCTGCAGG GTCCCTCGCC TGACGAAAAG TCCGCGGCTC CGGGGTGAA
CTCGACGTCC CAGGGAGCGG ACTGCTTTTC AGGCGCCGAG GCCCAACTT

26451 ACTCACTCCG GGGCTGTGGA CGTCGGCTTA CCTTCGCAA TTTGTACCTG
TGAGTGAGGC CCCGACACCT GCAGCCGAAT GGAAGCGTTT AAACATGGAC

26501 AGGACTACCA CGCCCACGAG ATTAGGTTCT ACGAAGACCA ATCCCGCCCG
TCCTGATGGT GCGGGTGCTC TAATCCAAGA TGCTTCTGGT TAGGGCGGGC

Figure 26 AB

WO 02/022080

PCT/US01/28861

| | | | | | |
|-------|------------|-------------|------------|------------|-------------|
| 26551 | CCTAATGCGG | AGCTTACCGC | CTGCGTCATT | ACCCAGGGCC | ACATTCTTGG |
| | GGATTACGCC | TCGAATGGCG | GACGCAGTAA | TGGGTCCCGG | TGTAAGAACC |
| 26601 | CCAATTGCAA | GCCATCAACA | AAGCCCGCCA | AGAGTTTCTG | CTACGAAAGG |
| | GGTTAACGTT | CGGTAGTTGT | TTCGGGCGGT | TCTCAAAGAC | GATGCTTTCC |
| 26651 | GACGGGGGGT | TTACTTGGAC | CCCCAGTCCG | GCGAGGAGCT | CAACCCAATC |
| | CTGCCCCCCA | AATGAACCTG | GGGGTCAGGC | CGCTCCTCGA | GTTGGGTTAG |
| 26701 | CCCCCGCCGC | CGCAGCCCTA | TCAGCAGCAG | CCGCGGGCCC | TTGCTTCCCA |
| | GGGGGCGGCG | GCGTCGGGAT | AGTCGTCGTC | GGCGCCCGGG | AACGAAGGGT |
| 26751 | GGATGGCACC | CAAAAAGAAG | CTGCAGCTGC | CGCCGCCACC | CACGGACGAG |
| | CCTACCGTGG | GTTTTTCTTC | GACGTCGACG | GCGGCGGTGG | GTGCCTGCTC |
| 26801 | GAGGAATACT | GGGACAGTCA | GGCAGAGGAG | GTTTTGGACG | AGGAGGAGGA |
| | CTCCTTATGA | CCCTGTCAGT | CCGTCTCCTC | CAAAACCTGC | TCCTCCTCCT |
| 26851 | GGACATGATG | GAAGACTGGG | AGAGCCTAGA | CGAGGAAGCT | TCCGAGGTCTG |
| | CCTGTACTAC | CTTCTGACCC | TCTCGGATCT | GCTCCTTCGA | AGGCTCCAGC |
| 26901 | AAGAGGTGTC | AGACGAAACA | CCGTCACCCT | CGGTCGCATT | CCCCTCGCCG |
| | TTCTCCACAG | TCTGCTTTGT | GGCAGTGGGA | GCCAGCGTAA | GGGGAGCGGC |
| 26951 | GCGCCCCAGA | AATCGGCAAC | CGGTTCCAGC | ATGGCTACAA | CCTCCGCTCC |
| | CGCGGGGTCT | TTAGCCGTTG | GCCAAGGTCG | TACCGATGTT | GGAGGCGAGG |
| 27001 | TCAGGCGCCG | CCGGCACTGC | CCGTTGCGCC | ACCCAACCGT | AGATGGGACA |
| | AGTCCGCGGC | GGCCGTGACG | GGCAAGCGGC | TGGGTGGA | TCTACCCTGT |
| 27051 | CCACTGGAAC | CAGGGCCGGT | AAGTCCAAGC | AGCCGCCGCC | GTTAGCCCAA |
| | GGTGACCTTG | GTCCCGGCCA | TTCAGGTTCG | TCGGCGGCGG | CAATCGGGTT |
| 27101 | GAGCAACAAC | AGCGCCAAGG | CTACCGCTCA | TGGCGCGGGC | ACAAGAACGC |
| | CTCGTTGTTG | TCGCGGTTCC | GATGGCGAGT | ACCGCGCCCG | TGTTCTTGCG |
| 27151 | CATAGTTGCT | TGCTTGCAAG | ACTGTGGGGG | CAACATCTCC | TTCGCCCCGC |
| | GTATCAACGA | ACGAACGTTC | TGACACCCCC | GTTGTAGAGG | AAGCGGGCGG |
| 27201 | GCTTTCTTCT | CTACCATCAC | GGCGTGGCCT | TCCCCCGTAA | CATCCTGCAT |
| | CGAAAGAAGA | GATGGTAGTG | CCGCACCGGA | AGGGGGCATT | GTAGGACGTA |
| 27251 | TACTACCGTC | ATCTCTACAG | CCCATACTGC | ACCGGCGGCA | GCGGCAGCAA |
| | ATGATGGCAG | TAGAGATGTC | GGGTATGACG | TGGCCGCCGT | CGCCGTCTGT |
| 27301 | CAGCAGCGGC | CACACAGAAG | CAAAGGCGAC | CGGATAGCAA | GACTCTGACA |
| | GTCGTCGCCG | GTGTGTCTTC | GTTTCCGCTG | GCCTATCGTT | CTGAGACTGT |
| 27351 | AAGCCCAAGA | AATCCACAGC | GGCGGCAGCA | GCAGGAGGAG | GAGCGCTGCG |
| | TTCGGGTTCT | TTAGGTGTCTG | CCGCCGTCTG | CGTCCTCCTC | CTCGCGACGC |
| 27401 | TCTGGCGCCC | AACGAACCCG | TATCGACCCG | CGAGCTTAGA | AACAGGATTT |
| | AGACCGCGGG | TTGCTTGGGC | ATAGCTGGGC | GCTCGAATCT | TTGTCCTAAA |
| 27451 | TTCCCACTCT | GTATGCTATA | TTTCAACAGA | GCAGGGGCCA | AGAACAAGAG |
| | AAGGGTGAGA | CATACGATAT | AAAGTTGTCT | CGTCCCCGGT | TCTTGTTCTC |

Figure 26 AC

27551 TCACAAAAGC GAAGATCAGC TTCGGCGCAC GCTGGAAGAC GCGGAGGCTC
 AGTGTTTTTCG CTTCTAGTCG AAGCCGCGTG CGACCTTCTG CGCCTCCGAG

27601 TCTTCAGTAA ATACTGCGCG CTGACTCTTA AGGACTAGTT TCGCGCCCTT
 AGAAGTCATT TATGACGCGC GACTGAGAAT TCCTGATCAA AGCGCGGGAA

27651 TCTCAAATTT AAGCGCGAAA ACTACGTCAT CTCCAGCGGC CACACCCGGC
 AGAGTTTAAA TTCGCGCTTT TGATGCAGTA GAGGTCGCCG GTGTGGGCCG

27701 GCCAGCACCT GTTGTCAGCG CCATTATGAG CAAGGAAATT CCCACGCCCT
 CGGTCGTGGA CAACAGTCGC GGTAATACTC GTTCCTTTAA GGGTGCGGGA

27751 ACATGTGGAG TTACCAGCCA CAAATGGGAC TTGCGGCTGG AGCTGCCCAA
 TGTACACCTC AATGGTTCGGT GTTTACCCTG AACGCCGACC TCGACGGGTT

27801 GACTACTCAA CCCGAATAAA CTACATGAGC GCGGGACCCC ACATGATATC
 CTGATGAGTT GGGCTTATTT GATGTACTCG CGCCCTGGGG TGTACTATAG

27851 CCGGGTCAAC GGAATACGCG CCCACCGAAA CCGAATTCTC CTGGAACAGG
 GGCCAGTTG CCTTATGCGC GGGTGGCTTT GGCTTAAGAG GACCTTGTCC

27901 CGGCTATTAC CACCACACCT CGTAATAACC TTAATCCCCG TAGTTGGCCC
 GCCGATAATG GTGGTGTGGA GCATTATTGG AATTAGGGGC ATCAACCGGG

27951 GCTGCCCTGG TGTACCAGGA AAGTCCCGCT CCCACCACTG TGGTACTTCC
 CGACGGGACC ACATGGTCCT TTCAGGGCGA GGGTGGTGAC ACCATGAAGG

28001 CAGAGACGCC CAGGCCGAAG TTCAGATGAC TAACTCAGGG GCGCAGCTTG
 GTCTCTGCGG GTCCGGCTTC AAGTCTACTG ATTGAGTCCC CGCGTCGAAC

28051 CGGGCGGCTT TCGTCACAGG GTGCGGTCGC CCGGGCAGGG TATAACTCAC
 GCCCGCCGAA AGCAGTGTCC CACGCCAGCG GGCCCGTCCC ATATTGAGTG

28101 CTGACAATCA GAGGGCGAGG TATTCAGCTC AACGACGAGT CGGTGAGCTC
 GACTGTTAGT CTCCCGCTCC ATAAGTCGAG TTGCTGCTCA GCCACTCGAG

28151 CTCGCTTGGT CTCCGTCCGG ACGGGACATT TCAGATCGGC GGCGCCGGCC
 GAGCGAACCA GAGGCAGGCC TGCCCTGTAA AGTCTAGCCG CCGCGGCCGG

28201 GCTCTTCATT CACGCCTCGT CAGGCAATCC TAACTCTGCA GACCTCGTCC
 CGAGAAGTAA GTGCGGAGCA GTCCGTTAGG ATTGAGACGT CTGGAGCAGG

28251 TCTGAGCCGC GCTCTGGAGG CATTGGAACT CTGCAATTTA TTGAGGAGTT
 AGACTCGGCG CGAGACCTCC GTAACCTTGA GACGTTAAAT AACTCCTCAA

28301 TGTGCCATCG GTCTACTTTA ACCCCTTCTC GGGACCTCCC GGCCACTATC
 ACACGGTAGC CAGATGAAAT TGGGGAAGAG CCCTGGAGGG CCGGTGATAG

28351 CGGATCAATT TATTCCTAAC TTTGACGCGG TAAAGGACTC GGCGGACGGC
 GCCTAGTTAA ATAAGGATTG AAACCTGCGC ATTTCTTGAG CCGCCTGCCG

28401 TACGACTGAA TGTTAAGTGG AGAGGCAGAG CAACTGCGCC TGAAACACCT
 ATGCTGACTT ACAATTCACC TCTCCGTCTC GTTGACGCGG ACTTTGTGGA

Figure 26 AD

WO 02/022080

PCT/US01/28861

28451 GGTCCACTGT CGCCGCCACA AGTGCTTTGC CCGCGACTCC GGTGAGTTTT
CCAGGTGACA GCGGCGGTGT TCACGAAACG GCGCGTGAGG CCACTCAAAA

28501 GCTACTTTGA ATTGCCCCGAG GATCATATCG AGGGCCCCGGC GCACGGCGTC
CGATGAAACT TAACGGGCTC CTAGTATAGC TCCCGGGCCG CGTGCCGCAG

28551 CGGCTTACCG CCCAGGGAGA GCTTGCCCGT AGCCTGATTC GGGAGTTTAC
GCCGAATGGC GGGTCCCTCT CGAACGGGCA TCGGACTAAG CCCTCAAATG

28601 CCAGCGCCCC CTGCTAGTTG AGCGGGACAG GGGACCCTGT GTTCTCACTG
GGTCGCGGGG GACGATCAAC TCGCCCTGTC CCCTGGGACA CAAGAGTGAC

28651 TGATTTGCAA CTGTCCTAAC CCTGGATTAC ATCAAGATCT TTGTTGCCAT
ACTAAACGTT GACAGGATTG GGACCTAATG TAGTTCTAGA AACAAACGGTA

28701 CTCTGTGCTG AGTATAATAA ATACAGAAAT TAAAATATAC TGGGGCTCCT
GAGACACGAC TCATATTATT TATGTCTTTA ATTTTATATG ACCCCGAGGA

28751 ATCGCCATCC TGTAACGCC ACCGTCTTCA CCCGCCCAAG CAAACCAAGG
TAGCGGTAGG ACATTTGCGG TGGCAGAAGT GGGCGGGTTC GTTTGGTTCC

28801 CGAACCTTAC CTGGTACTTT TAACATCTCT CCCTCTGTGA TTTACAACAG
GCTTGGAATG GACCATGAAA ATTGTAGAGA GGGAGACACT AAATGTTGTC

28851 TTTCAACCCA GACGGAGTGA GTCTACGAGA GAACCTCTCC GAGCTCAGCT
AAAGTTGGGT CTGCCTCACT CAGATGCTCT CTTGGAGAGG CTCGAGTCGA

28901 ACTCCATCAG AAAAAACACC ACCCTCCTTA CCTGCCGGGA ACGTACGAGT
TGAGGTAGTC TTTTTTGTGG TGGGAGGAAT GGACGGCCCT TGCATGCTCA

28951 GCGTCACCGG CCGCTGCACC ACACCTACCG CCTGACCGTA AACCAGACTT
CGCAGTGGCC GCGGACGTGG TGTGGATGGC GGACTGGCAT TTGGTCTGAA

29001 TTTCCGGACA GACCTCAATA ACTCTGTTTA CCAGAACAGG AGGTGAGCTT
AAAGGCCTGT CTGGAGTTAT TGAGACAAAT GGTCTTGTC TCCACTCGAA

29051 AGAAAACCCT TAGGGTATTA GGCCAAAGGC GCAGCTACTG TGGGGTTTAT
TCTTTTGGGA ATCCCATAAT CCGGTTTCCG CGTCGATGAC ACCCCAAATA

29101 GAACAATTCA AGCAACTCTA CGGGCTATTC TAATTCAGGT TTCTCTAGAA
CTTGTTAAGT TCGTTGAGAT GCCCGATAAG ATTAAGTCCA AAGAGATCTT

29151 TCGGGGTTGG GGTATTCTC TGTCTTGTA TTCTCTTTAT TCTTATACTA
AGCCCCAACC CCAATAAGAG ACAGAACT AAGAGAAATA AGAATATGAT

29201 ACGCTTCTCT GCCTAAGGCT CGCCGCCTGC TGTGTGCACA TTTGCATTTA
TGCGAAGAGA CGGATTCCGA GCGGCGGACG ACACACGTGT AAACGTAAAT

29251 TTGTCAGCTT TTAAACGCT GGGGTCGCCA CCCAAGATGA TTAGGTACAT
AACAGTCGAA AAATTTGCGA CCCCAGCGGT GGGTTCTACT AATCCATGTA

29301 AATCCTAGGT TTAATCACC TTGCGTCAGC CCACGGTACC ACCCAAAGG
TTAGGATCCA AATGAGTGGG AACGCAGTCG GGTGCCATGG TGGGTTTTCC

29351 TGGATTTTAA GGAGCCAGCC TGTAATGTTA CATTGCGAGC TGAAGCTAAT
ACCTAAAATT CCTCGGTCGG ACATTACAAT GTAAGCGTCG ACTTCGATTA

Figure 26 AE

```

29451 TCGCCACAAA AACAAAATTG GCAAGTATGC TGTTTATGCT ATTTGGCAGC
      AGCGGTGTTT TTGTTTTAAC CGTTCATACG ACAAATACGA TAAACCGTCG

29501 CAGGTGACAC TACAGAGTAT AATGTTACAG TTTTCCAGGG TAAAAGTCAT
      GTCCACTGTG ATGTCTCATA TTACAATGTC AAAAGGTCCC ATTTTCAGTA

29551 AAAACTTTTA TGTATACTTT TCCATTTTAT GAAATGTGCG ACATTACCAT
      TTTTGAAAAT ACATATGAAA AGGTAAAATA CTTTACACGC TGTAATGGTA

29601 GTACATGAGC AAACAGTATA AGTTGTGGCC CCCACAAAAT TGTGTGGAAA
      CATGTACTCG TTTGTCATAT TCAACACCGG GGGTGTTTTA ACACACCTTT

29651 ACACTGGCAC TTTCTGCTGC ACTGCTATGC TAATTACAGT GCTCGCTTTG
      TGTGACCGTG AAAGACGACG TGACGATACG ATTAATGTCA CGAGCGAAAC

29701 GTCTGTACCC TACTCTATAT TAAATACAAA AGCAGACGCA GCTTTATTGA
      CAGACATGGG ATGAGATATA ATTTATGTTT TCGTCTGCGT CGAAATAACT

29751 GGAAAAGAAA ATGCCTTAAT TACTAAGTT ACAAAGCTAA TGTCACCACT
      CCTTTTCTTT TACGGAATTA AATGATTCAA TGTTTCGATT ACAGTGGTGA

29801 AACTGCTTTA CTCGCTGCTT GCAAAACAAA TTCAAAAAGT TAGCATTATA
      TTGACGAAAT GAGCGACGAA CGTTTTGTTT AAGTTTTTCA ATCGTAATAT

29851 ATTAGAATAG GATTTAAACC CCCC GGTCAT TTCCTGCTCA ATACCATTCC
      TAATCTTATC CTAAATTTGG GGGGCCAGTA AAGGACGAGT TATGGTAAGG

29901 CCTGAACAAT TGACTCTATG TGGGATATGC TCCAGCGCTA CAACCTTGAA
      GGACTTGTTA ACTGAGATAC ACCCTATACG AGGTCGCGAT GTTGGAACCT

29951 GTCAGGCTTC CTGGATGTCA GCATCTGACT TTGGCCAGCA CCTGTCCCGC
      CAGTCCGAAG GACCTACAGT CGTAGACTGA AACC GGTCGT GGACAGGGCG

30001 GGATTTGTTC CAGTCCAAC TACAGCGACCC ACCCTAACAG AGATGACCAA
      CCTAAACAAG GTCAGGTTGA TGTCGCTGGG TGGGATTGTC TCTACTGGTT

30051 CACAACCAAC GCGGCCGCGG CTACCGGACT TACATCTACC ACAAATACAC
      GTGTTGGTTG CGCCGGCGGC GATGGCCTGA ATGTAGATGG TGTTTATGTG

30101 CCCAAGTTTC TGCCTTTGTC AATAACTGGG ATAACCTGGG CATGTGGTGG
      GGGTTCAAAG ACGGAAACAG TTATTGACCC TATTGAACCC GTACACCACC

30151 TTCTCCATAG CGCTTATGTT TGTATGCCTT ATTATTATGT GGCTCATCTG
      AAGAGGTATC GCGAATACAA ACATACGGAA TAATAATACA CCGAGTAGAC

30201 CTGCCTAAAG CGCAAACGCG CCCGACCACC CATCTATAGT CCCATCATTG
      GACGGATTTT GCGTTTGCGC GGGCTGGTGG GTAGATATCA GGGTAGTAAC

30251 TGCTACACCC AAACAATGAT GGAATCCATA GATTGGACGG ACTGAAACAC
      ACGATGTGGG TTTGT TACTA CCTTAGGTAT CTAACCTGCC TGACTTTGTG

30301 ATGTTCTTTT CTCTTACAGT ATGATTAAAT GAGACATGAT TCCTCGAGTT
      TACAAGAAAA GAGAATGTCA TACTAATTTA CTCTGTACTA AGGAGCTCAA

```

Figure 26 AF

30401 TCGGGTTTCT CACATCGAAG TAGACTGCAT TCCAGCCTTC ACAGTCTATT
ACGCCAAAGA GTGTAGCTTC ATCTGACGTA AGGTCGGAAG TGTCAGATAA

30451 TGCTTTACGG ATTTGTCACC CTCACGCTCA TCTGCAGCCT CATCACTGTG
ACGAAATGCC TAAACAGTGG GAGTGCGAGT AGACGTCGGA GTAGTGACAC

30501 GTCATCGCCT TTATCCAGTG CATTGACTGG GTCTGTGTGC GCTTTGCATA
CAGTAGCGGA AATAGGTCAC GTAAGTACC CAGACACACG CGAAACGTAT

30551 TCTCAGACAC CATCCCCAGT ACAGGGACAG GACTATAGCT GAGCTTCTTA
AGAGTCTGTG GTAGGGGTCA TGTCCCTGTC CTGATATCGA CTCGAAGAAT

30601 GAATTCTTTA ATTATGAAAT TTAGTGTGAC TTTTCTGCTG ATTATTTGCA
CTTAAGAAAT TAATACTTTA AATGACACTG AAAAGACGAC TAATAAACGT

30651 CCCTATCTGC GTTTTGTTC CCGACCTCCA AGCCTCAAAG ACATATATCA
GGGATAGACG CAAAACAAGG GGCTGGAGGT TCGGAGTTTC TGTATATAGT

30701 TGCAGATTCA CTCGTATATG GAATATTCCA AGTTGCTACA ATGAAAAAAG
ACGTCTAAGT GAGCATATAC CTTATAAGGT TCAACGATGT TACTTTTTTC

30751 CGATCTTTCC GAAGCCTGGT TATATGCAAT CATCTCTGTT ATGGTGTTCCT
GCTAGAAAGG CTTCGGACCA ATATACGTTA GTAGAGACAA TACCACAAGA

30801 GCAGTACCAT CTTAGCCCTA GCTATATATC CCTACCTTGA CATTGGCTGG
CGTCATGGTA GAATCGGGAT CGATATATAG GGATGGAACT GTAACCGACC

30851 AACGCAATAG ATGCCATGAA CCACCCAAC TCCCCGCGC CCGCTATGCT
TTGCGTTATC TACGGTACTT GGTGGGTTGA AAGGGGCGCG GCGGATACGA

30901 TCCACTGCAA CAAGTTGTTG CCGGCGGCTT TGTCCCAGCC AATCAGCCTC
AGGTGACGTT GTTCAACAAC GGCCGCCGAA ACAGGGTCGG TTAGTCGGAG

30951 GCCCACCTTC TCCCACCCCC ACTGAAATCA GCTACTTTAA TCTAACAGGA
CGGGTGGAAG AGGGTGGGGG TGACTTTAGT CGATGAAATT AGATTGTCCT

31001 GGAGATGACT GACACCCTAG ATCTAGAAAT GGACGGAATT ATTACAGAGC
CCTCTACTGA CTGTGGGATC TAGATCTTTA CCTGCCTTAA TAATGTCTCG

31051 AGCGCCTGCT AGAAAGACGC AGGGCAGCGG CCGAGCAACA GCGCATGAAT
TCGCGGACGA TCTTTCTGCG TCCCGTCGCC GGCTCGTTGT CGCGTACTTA

31101 CAAGAGCTCC AAGACATGGT TAACTTGCAC CAGTGCAAAA GGGGTATCTT
GTTCTCGAGG TTCTGTACCA ATTGAACGTG GTCACGTTTT CCCCATAGAA

31151 TTGTCTCGTA AAGCAGGCCA AAGTCACCTA CGACAGTAAT ACCACCGGAC
AACAGAGCAT TTCGTCCGGT TTCAGTGGAT GCTGTCATTA TGGTGGCCTG

31201 ACCGCCTTAG CTACAAGTTG CCAACCAAGC GTCAGAAATT GGTGGTCATG
TGGCGGAATC GATGTTCAAC GGTGTTTCG CAGTCTTTAA CCACCAGTAC

31251 GTGGGAGAAA AGCCCATTA CATAACTCAG CACTCGGTAG AAACCGAAGG
CACCTCTTT TCGGGTAATG GTATTGAGTC GTGAGCCATC TTTGGCTTCC

Figure 26 AG

31351 AGACCCTGTG CGGTCTCAAA GATCTTATTC CCTTTAACTA ATAAAAAAAA
TCTGGGACAC GCCAGAGTTT CTAGAATAAG GGAAATTGAT TATTTTTTTT

31401 ATAATAAAGC ATCACTTACT TAAAATCAGT TAGCAAATTT CTGTCCAGTT
TATTATTTTCG TAGTGAATGA ATTTTAGTCA ATCGTTTAAA GACAGGTCAA

31451 TATTCAGCAG CACCTCCTTG CCCTCCTCCC AGCTCTGGTA TTGCAGCTTC
ATAAGTCGTC GTGGAGGAAC GGGAGGAGGG TCGAGACCAT AACGTCGAAG

31501 CTCCTGGCTG CAAACTTTCT CCACAATCTA AATGGAATGT CAGTTTCCTC
GAGGACCGAC GTTTGAAAGA GGTGTTAGAT TTACCTTACA GTCAAAGGAG

31551 CTGTTCTGT CCATCCGCAC CCACTATCTT CATGTTGTTG CAGATGAAGC
GACAAGGACA GGTAGGCGTG GGTGATAGAA GTACAACAAC GTCTACTTCG

31601 GCGCAAGACC GTCTGAAGAT ACCTTCAACC CCGTGTATCC ATATGACACG
CGCGTTCTGG CAGACTTCTA TGGAAAGTTGG GGCACATAGG TATACTGTGC

31651 GAAACCGGTC CTCCAACGTG GCCTTTTCTT ACTCCTCCCT TTGTATCCCC
CTTTGGCCAG GAGGTTGACA CGGAAAAGAA TGAGGAGGGA AACATAGGGG

31701 CAATGGGTTT CAAGAGAGTC CCCCTGGGGT ACTCTCTTTG CGCCTATCCG
GTTACCCAAA GTTCTCTCAG GGGGACCCCA TGAGAGAAAC GCGGATAGGC

31751 AACCTCTAGT TACCTCCAAT GGCATGCTTG CGCTCAAAT GGGCAACGGC
TTGGAGATCA ATGGAGGTTA CCGTACGAAC GCGAGTTTAA CCCGTTGCCG

31801 CTCTCTCTGG ACGAGGCCGG CAACCTTACC TCCCAAATG TAACCACTGT
GAGAGAGACC TGCTCCGGCC GTTGGAATGG AGGGTTTTAC ATTGGTGACA

31851 GAGCCACCT CTCAAAAAA CCAAGTCAAA CATAAACCTG GAAATATCTG
CTCGGGTGGA GAGTTTTTTT GGTTCAGTTT GTATTTGGAC CTTTATAGAC

31901 CACCCCTCAC AGTTACCTCA GAAGCCCTAA CTGTGGCTGC CGCCGCACCT
GTGGGGAGTG TCAATGGAGT CTTGCGGATT GACACCGACG GCGGCGTGGA

31951 CTAATGGTCG CGGGCAACAC ACTCACCATG CAATCACAGG CCCCCTAAC
GATTACCAGC GCCCCTTGTC TGAGTGGTAC GTTAGTGTCC GGGGCGATTG

32001 CGTGCACGAC TCCAAACTTA GCATTGCCAC CCAAGGACCC CTCACAGTGT
GCACGTGCTG AGGTTTGAAT CGTAACGGTG GGTTCCTGGG GAGTGTCACA

32051 CAGAAGGAAA GCTAGCCCTG CAAACATCAG GCCCCCTCAC CACCACCGAT
GTCTTCCTTT CGATCGGGAC GTTTGTAGTC CGGGGGAGTG GTGGTGGCTA

32101 AGCAGTACCC TTAATATCAC TGCCTCACCC CCTCTAACTA CTGCCACTGG
TCGTCATGGG AATGATAGTG ACGGAGTGGG GGAGATTGAT GACGGTGACC

32151 TAGCTTGGGC ATTGACTTGA AAGAGCCCAT TTATACACAA AATGGAAAAC
ATCGAACCCG TAACTGAACT TTCTCGGGTA AATATGTGTT TTACCTTTTG

32201 TAGGACTAAA GTACGGGGCT CCTTTGCATG TAACAGACGA CCTAAACACT
ATCCTGATTT CATGCCCGA GGAAACGTAC ATTGTCTGCT GGATTTGTGA

Figure 26 A H

32301 AACTAAAGTT ACTGGAGCCT TGGGTTTTGA TTCACAAGGC AATATGCAAC
 TTGATTTCAA TGACCTCGGA ACCCAAACCT AAGTGTTCCG TTATACGTTG
 32351 TTAATGTAGC AGGAGGACTA AGGATTGATT CTCAAAACAG ACGCCTTATA
 AATTACATCG TCCTCCTGAT TCCTAACTAA GAGTTTTGTC TCGGGAATAT
 32401 CTTGATGTTA GTTATCCGTT TGATGCTCAA AACCAACTAA ATCTAAGACT
 GAACTACAAT CAATAGGCAA ACTACGAGTT TTGGTTGATT TAGATTCTGA
 32451 AGGACAGGGC CCTCTTTTTA TAAACTCAGC CCACAACCTG GATATTAACCT
 TCCTGTCCCG GGAGAAAAAT ATTTGAGTCG GGTGTTGAAC CTATAATTGA
 32501 ACAACAAAGG CCTTTACTTG TTTACAGCTT CAAACAATTC CAAAAAGCTT
 TGTTGTTTCC GGAAATGAAC AAATGTCGAA GTTTGTTAAG GTTTTTTCGAA
 32551 GAGGTTAACC TAAGCACTGC CAAGGGGTTG ATGTTTGACG CTACAGCCAT
 CTCCAATTGG ATTCGTGACG GTTCCCCAAC TACAACTGC GATGTCGGTA
 32601 AGCCATTAAT GCAGGAGATG GGCTTGAATT TGGTTCACCT AATGCACCAA
 TCGGTAATTA CGTCCTCTAC CCGAACTTAA ACCAAGTGGA TTACGTGGTT
 32651 ACACAAATCC CCTCAAAACA AAAATTGGCC ATGGCCTAGA ATTTGATTCA
 TGTGTTTAGG GGAGTTTTGT TTTTAACCGG TACCGGATCT TAAACTAAGT
 32701 AACAAAGGCTA TGGTTCCTAA ACTAGGAACT GGCCTTAGTT TTGACAGCAC
 TTGTTCGGAT ACCAAGGATT TGATCCTTGA CCGGAATCAA AACTGTCGTG
 32751 AGGTGCCATT ACAGTAGGAA ACAAAAATAA TGATAAGCTA ACTTTGTGGA
 TCCACGGTAA TGTCATCCTT TGTTTTTATT ACTATTCGAT TGAAACACCT
 32801 CCACACCAGC TCCATCTCCT AACTGTAGAC TAAATGCAGA GAAAGATGCT
 GGTGTGGTCG AGGTAGAGGA TTGACATCTG ATTTACGTCT CTTTCTACGA
 32851 AAACCTCACTT TGGTCTTAAC AAAATGTGGC AGTCAAATAC TTGCTACAGT
 TTTGAGTGAA ACCAGAATTG TTTTACACCG TCAGTTTATG AACGATGTCA
 32901 TTCAGTTTTG GCTGTAAAG GCAGTTTGGC TCCAATATCT GGAACAGTTC
 AAGTCAAAAC CGACAATTTC CGTCAAACCG AGGTTATAGA CTTGTCAAG
 32951 AAAGTGCTCA TCTTATTATA AGATTTGACG AAAATGGAGT GCTACTAAAC
 TTTCACGAGT AGAATAATAT TCTAACTGC TTTTACCTCA CGATGATTG
 33001 AATTCCTTCC TGGACCCAGA ATATTGGAAC TTTAGAAATG GAGATCTTAC
 TTAAGGAAGG ACCTGGGTCT TATAACCTG AAATCTTTAC CTCTAGAATG
 33051 TGAAGGCACA GCCTATACAA ACGCTGTTGG ATTTATGCCT AACCTATCAG
 ACTTCCGTGT CGGATATGTT TGCACAAACC TAAATACGGA TTGGATAGTC
 33101 CTTATCCAAA ATCTCACGGT AAAACTGCCA AAAGTAACAT TGTCAGTCAA
 GAATAGGTTT TAGAGTGCCA TTTTGACGGT TTTCATTGTA ACAGTCAGTT
 33151 GTTTACTTAA ACGGAGACAA AACTAAACCT GTAACACTAA CCATTACACT
 CAAATGAATT TGCCTCTGTT TTGATTTGGA CATTGTGATT GGTAATGTGA

Figure 26 AI

33251 CATTTTCATG GGACTGGTCT GGCCACAAC T ACATTAATGA AATATTTGCC
GTAAAAGTAC CCTGACCAGA CCGGTGTTGA TGTAATTACT TTATAAACGG

33301 ACATCCTCTT ACACTTTTTC ATACATTGCC CAAGAATAAA GAATCGTTTG
TG TAGGAGAA TGTGAAAAAG TATGTAACGG GTTCTTATTT CTTAGCAAAC

33351 TGTTATGTTT CAACGTGTTT ATTTTTC AAT TGCAGAAAAT TTCAAGTCAT
ACAATACAAA GTTGCACAAA TAAAAAGTTA ACGTCTTTTA AAGTTCAGTA

33401 TTTTCATTCA GTAGTATAGC CCCACCACCA CATAGCTTAT ACAGATCACC
AAAAGTAAGT CATCATATCG GGGTGGTGGT GTATCGAATA TGTCTAGTGG

33451 GTACCTTAAT CAAACTCACA GAACCCTAGT ATTCAACCTG CCACCTCCCT
CATGGAATTA GTTTGAGTGT CTTGGGATCA TAAGTTGGAC GGTGGAGGGA

33501 CCCAACACAC AGAGTACACA GTCCTTTCTC CCCGGCTGGC CTTAAAAAGC
GGGTTGTGTG TCTCATGTGT CAGGAAAGAG GGGCCGACCG GAATTTTTCG

33551 ATCATATCAT GGGTAACAGA CATATTCTTA GGTGTTATAT TCCACACGGT
TAGTATAGTA CCCATTGTCT GTATAAGAAT CCACAATATA AGGTGTGCCA

33601 TTCCTGTCGA GCCAAACGCT CATCAGTGAT ATTAATAAAC TCCCCGGGCA
AAGGACAGCT CGGTTTGCGA GTAGTACTA TAATTATTTG AGGGGCCCGT

33651 GCTCACTTAA GTTCATGTCT CTGTCCAGCT GCTGAGCCAC AGGCTGCTGT
CGAGTGAATT CAAGTACAGC GACAGGTCGA CGACTCGGTG TCCGACGACA

33701 CCAACTTGCG GTTGCTTAAC GGGCGGCGAA GGAGAAGTCC ACGCCTACAT
GGTTGAACGC CAACGAATTG CCCGCCGCTT CCTCTTCAGG TGCGGATGTA

33751 GGGGGTAGAG TCATAATCGT GCATCAGGAT AGGGCGGTGG TGCTGCAGCA
CCCCCATCTC AGTATTAGCA CGTAGTCCTA TCCCGCCACC ACGACGTCGT

33801 GCGCGCGAAT AAAGTCTGCT GCAGCGCGCT CCGTCCTGCA GGAATACAAC
CGCGCGCTTA TTTGACGACG GCGGCGGCGA GGCAGGACGT CCTTATGTTG

33851 ATGGCAGTGG TCTCCTCAGC GATGATTCGC ACCGCCCGCA GCATAAGGCG
TACCGTCACC AGAGGAGTCG CTACTAAGCG TGGCGGGCGT CGTATTCCGC

33901 CCTTGTCCTC CGGGCACAGC AGCGCACCTT GATCTCACTT AAATCAGCAC
GGAACAGGAG GCCCGTGTCTG TCGCGTGGA CTAGAGTGAA TTTAGTCGTG

33951 AGTAACTGCA GCACAGCACC ACAATATTGT TCAAAATCCC ACAGTGCAAG
TCATTGACGT CGTGTCGTGG TGTTATAACA AGTTTTAGGG TGTCACGTTC

34001 GCGCTGTATC CAAAGCTCAT GGCGGGGACC ACAGAACCCA CGTGGCCATC
CGCGACATAG GTTTCGAGTA CCGCCCCTGG TGTCTTGGGT GCACCGGTAG

34051 ATACCACAAG CGCAGGTAGA TTAAGTGGCG ACCCCTCATA AACACGCTGG
TATGGTGTTT GCGTCCATCT AATTCACCGC TGGGGAGTAT TTGTGCGACC

34101 ACATAAACAT TACCTCTTTT GGCATGTTGT AATTCACCAC CTCCCGGTAC
TGTATTTGTA ATGGAGAAAA CCGTACAACA TTAAGTGGTG GAGGGCCATG

Figure 26 AJ

34201 GCTGGCCAAA ACCTGCCCGC CGGCTATACA CTGCAGGGAA CCGGGACTGG
CGACCGGTTT TGGACGGGCG GCCGATATGT GACGTCCCTT GGCCCTGACC

34251 AACAAATGACA GTGGAGAGCC CAGGACTCGT AACCATGGAT CATCATGCTC
TTGTTACTGT CACCTCTCGG GTCCTGAGCA TTGGTACCTA GTAGTACGAG

34301 GTCATGATAT CAATGTTGGC ACAACACAGG CACACGTGCA TACACTTCCT
CAGTACTATA GTTACAACCG TGTTGTGTCC GTGTGCACGT ATGTGAAGGA

34351 CAGGATTACA AGCTCCTCCC GCGTTAGAAC CATATCCCAG GGAACAACCC
GTCCTAATGT TCGAGGAGGG CGCAATCTTG GTATAGGGTC CCTTGTTGGG

34401 ATTCCTGAAT CAGCGTAAAT CCCACACTGC AGGGAAGACC TCGCACGTAA
TAAGGACTTA GTCGCATTTA GGGTGTGACG TCCCTTCTGG AGCGTGCATT

34451 CTCACGTTGT GCATTGTCAA AGTGTTACAT TCGGGCAGCA GCGGATGATC
GAGTGCAACA CGTAACAGTT TCACAATGTA AGCCCGTCGT CGCCTACTAG

34501 CTCCAGTATG GTAGCGCGGG TTTCTGTCTC AAAAGGAGGT AGACGATCCC
GAGGTCATAC CATCGCGCCC AAAGACAGAG TTTTCCTCCA TCTGCTAGGG

34551 TACTGTACGG AGTGCGCCGA GACAACCGAG ATCGTGTGTTGG TCGTAGTGTC
ATGACATGCC TCACGCGGCT CTGTTGGCTC TAGCACAACC AGCATCACAG

34601 ATGCCAAATG GAACGCCGGA CGTAGTCATA TTTCTGAAG CAAAACCAGG
TACGGTTTAC CTTGCGGCCT GCATCAGTAT AAAGGACTTC GTTTTGGTCC

34651 TGCGGGCGTG ACAAACAGAT CTGCGTCTCC GGTCTCGCCG CTTAGATCGC
ACGCCCCGAC TGTTTGTCTA GACGCAGAGG CCAGAGCGGC GAATCTAGCG

34701 TCTGTGTAGT AGTTGTAGTA TATCCACTCT CTCAAAGCAT CCAGGCGCCC
AGACACATCA TCAACATCAT ATAGGTGAGA GAGTTTCGTA GGTCCGCGGG

34751 CCTGGCTTCG GGTTCCTATGT AAACCTCCTC ATGCGCCGCT GCCCTGATAA
GGACCGAAGC CCAAGATACA TTTGAGGAAG TACGCGGCGA CGGGACTATT

34801 CATCCACCAC CGCAGAATAA GCCACACCCA GCCAACCTAC ACATTCTGTTT
GTAGGTGGTG GCGTCTTATT CGGTGTGGGT CGGTTGGATG TGTAAGCAAG

34851 TGCGAGTCAC ACACGGGAGG AGCGGGAAGA GCTGGAAGAA CCATGTTTTT
ACGCTCAGTG TGTGCCCTCC TCGCCCTTCT CGACCTTCTT GGTACAAAAA

34901 TTTTTTATTC CAAAAGATTA TCCAAAACCT CAAAATGAAG ATCTATTAAG
AAAAATAAG GTTTTCTAAT AGGTTTTGGA GTTTTACTTC TAGATAATTC

34951 TGAACGCGCT CCCCTCCGGT GGCGTGGTCA AACTCTACAG CCAAAGAACA
ACTTGCGCGA GGGGAGGCCA CCGCACCAGT TTGAGATGTC GGTTCCTTGT

35001 GATAATGGCA TTTGTAAGAT GTTGCACAAT GGCTTCCAAA AGGCAAACGG
CTATTACCGT AAACATTCTA CAACGTGTTA CCGAAGGTTT TCCGTTTGCC

35051 CCCTCACGTC CAAGTGGACG TAAAGGCTAA ACCCTTCAGG GTGAATCTCC
GGGAGTGCAG GTTCACCTGC ATTTCCGATT TGGGAAGTCC CACTTAGAGG

Figure 26 AK

35151 CCACCTTCTC AATATATCTC TAAGCAAATC CCGAATATTA AGTCCGGCCA
GGTGGAAGAG TTATATAGAG ATTCGTTTAG GGCTTATAAT TCAGGCCGGT

35201 TTGTAAAAAT CTGCTCCAGA GCGCCCTCCA CCTTCAGCCT CAAGCAGCGA
AACATTTTTA GACGAGGTCT CGCGGGAGGT GGAAGTCGGA GTTCGTGCGT

35251 ATCATGATTG CAAAAATTCA GGTCCTCAC AGACCTGTAT AAGATTCAAA
TAGTACTAAC GTTTTTAAGT CCAAGGAGTG TCTGGACATA TTCTAAGTTT

35301 AGCGGAACAT TAACAAAAAT ACCGCGATCC CGTAGGTCCC TTCGCAGGGC
TCGCCTTGTA ATTGTTTTTA TGGCGCTAGG GCATCCAGGG AAGCGTCCCC

35351 CAGCTGAACA TAATCGTGCA GGTCGTGACG GACCAGCGCG GCCACTTCCC
GTCGACTTGT ATTAGCACGT CCAGACGTGC CTGGTCGCGC CGGTGAAGGG

35401 CGCCAGGAAC CATGACAAAA GAACCCACAC TGATTATGAC ACGCATACTC
GCGGTCCTTG GTACTGTTTT CTTGGGTGTG ACTAATACTG TCGGTATGAG

35451 GGAGCTATGC TAACCAGCGT AGCCCCGATG TAAGCTTGTT GCATGGGCGG
CCTCGATACG ATTGGTCGCA TCGGGGCTAC ATTCGAACAA CGTACCCGCC

35501 CGATATAAAA TGCAAGGTGC TGCTCAAAAA ATCAGGCAAA GCCTCGCGCA
GCTATATTTT ACGTTCCACG ACGAGTTTTT TAGTCCGTTT CGGAGCGCGT

35551 AAAAAGAAAG CACATCGTAG TCATGCTCAT GCAGATAAAG GCAGGTAAGC
TTTTTCTTTC GTGTAGCATC AGTACGAGTA CGTCTATTTC CGTCCATTTC

35601 TCCGGAACCA CCACAGAAAA AGACACCATT TTTCTCTCAA ACATGTCTGC
AGGCCTTGGT GGTGTCTTTT TCTGTGGTAA AAAGAGAGTT TGTACAGACG

35651 GGGTTTCTGC ATAAACACAA AATAAAATAA CAAAAAACA TTAAACATT
CCCAAAGACG TATTTGTGTT TTATTTTATT GTTTTTTTGT AAATTTGTAA

35701 AGAAGCCTGT CTTACAACAG GAAAAACAAC CCTTATAAGC ATAAGACGGA
TCTTCGGACA GAATGTTGTC CTTTTTGTTG GGAATATTCG TATTCTGCCT

35751 CTACGGCCAT GCCGGCGTGA CCGTAAAAAA ACTGGTCACC GTGATTAAAA
GATGCCGGTA CGGCCGCACT GGCATTTTTT TGACCAGTGG CACTAATTTT

35801 AGCACCACCG ACAGCTCCTC GGTCATGTCC GGAGTCATAA TGTAAGACTC
TCGTGGTGGC TGTCGAGGAG CCAGTACAGG CCTCAGTATT ACATTCTGAG

35851 GGTAACACA TCAGGTTGAT TCACATCGGT CAGTGCTAAA AAGCGACCGA
CCATTTGTGT AGTCCAATA AGTGTAGCCA GTCACGATTT TTCGCTGGCT

35901 AATAGCCCGG GGAATACAT ACCCGCAGGC GTAGAGACAA CATTACAGCC
TTATCGGGCC CCCTTATGTA TGGGCGTCCG CATCTCTGTT GTAATGTCGG

35951 CCCATAGGAG GTATAACAAA ATTAATAGGA GAGAAAAACA CATAAACACC
GGGTATCCTC CATATTGTTT TAATTATCCT CTCTTTTTGT GTATTTGTGG

36001 TGAAAAACCC TCCTGCCTAG GCAAATAGC ACCCTCCCGC TCCAGAACAA
ACTTTTTGGG AGGACGGATC CGTTTTATCG TGGGAGGGCG AGGTCTTGTT

Figure 26 AL

```

36101 AAAGAAAACC TATTAAAAAA ACACCACTCG ACACGGCACC AGCTCAATCA
      TTTCTTTTGG ATAATTTTTT TGTGGTGAGC TGTGCCGTGG TCGAGTTAGT

36151 GTCACAGTGT AAAAAAGGGC CAAGTGCAGA GCGAGTATAT ATAGGACTAA
      CAGTGTCACA TTTTTTCCCC GTTCACGTCT CGCTCATATA TATCCTGATT

36201 AAAATGACGT AACGGTTAAA GTCCACAAAA AACACCCAGA AAACCGCACG
      TTTTACTGCA TTGCCAATTT CAGGTGTTTT TTGTGGGTCT TTTGGCGTGC

36251 CGAACCTACG CCCAGAAACG AAAGCCAAAA AACCCACAAC TTCCTCAAAT
      GCTTGGATGC GGGTCTTTGC TTTCGGTTTT TTGGGTGTTG AAGGAGTTTA

36301 CGTCACTTCC GTTTTCCCAC GTTACGTCAC TTCCCATTTT AAGAAAATA
      GCAGTGAAGG CAAAAGGGTG CAATGCAGTG AAGGGTAAAA TTCTTTTGAT

36351 CAATTCCCAA CACATACAAG TTA CTCCGCC CTAAACCTA CGTCACCCGC
      GTTAAGGGTT GTGTATGTTT AATGAGGCGG GATTTTGGAT GCAGTGGGCG

36401 CCCGTTCCCA CGCCCCGCGC CACGTCACAA ACTCCACCCC CTCATTATCA
      GGGCAAGGGT GCGGGGCGCG GTGCAGTGTT TGAGGTGGGG GAGTAATAGT

                                     PacI
                                     ~~~~~~

36451 TATTGGCTTC AATCCAAAAT AAGGTATATT ATTGATGATG TTAATTAAGA
      ATAACCGAAG TTAGGTTTTA TTCCATATAA TAACTACTAC AATTAATTCT

36501 ATTCGGATCT GCGACGCGAG GCTGGATGGC CTTCCCCATT ATGATTCTTC
      TAAGCCTAGA CGCTGCGCTC CGACCTACCG GAAGGGGTAA TACTAAGAAG

36551 TCGCTTCCGG CGGCATCGGG ATGCCCCGGT TGCAGGCCAT GCTGTCCAGG
      AGCGAAGGCC GCCGTAGCCC TACGGGCGCA ACGTCCGGTA CGACAGGTCC

36601 CAGGTAGATG ACGACCATCA GGGACAGCTT CAAGGCCAGC AAAAGGCCAG
      GTCCATCTAC TGCTGGTAGT CCCTGTCGAA GTTCCGGTCG TTTTCCGGTC

36651 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
      CTTGGCATT TTTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG

36701 CTGACGAGCA TCACAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
      GACTGCTCGT AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC

36751 ACAGGACTAT AAAGATACCA GGC GTTTCCC CCTGGAAGCT CCCTCGTGCG
      TGTCCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC

36801 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
      GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG

36851 CTTGCGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT
      GAAGCCCTTC GCACCGCGAA AGAGTATCGA GTGCGACATC CATAGAGTCA

36901 TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
      AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA

```

Figure 26 AM

37001 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA

37051 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC
TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG

37101 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA
ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT

37151 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA
TCGGTCAATG GAAGCCTTTT TCTCAACCAT CGAGAAGTAG GCCGTTTGT

37201 ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG
TGGTGGCGAC CATCGCCACC AAAAAACAA ACGTTCGTCG TCTAATGCGC

37251 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
GTCTTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC

37301 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA
TGCGAGTCAC CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT

37351 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATCAATCTA AAGTATATAT
AGTTTTTCCT AGAAGTGGAT CTAGGAAAAT TTAGTTAGAT TTCATATATA

37401 GAGTAAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT
CTCATTTGAA CCAGACTGTC AATGGTTACG AATTAGTCAC TCCGTGGATA

37451 CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCTGA CTCCCCGTCTG
GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

37501 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT

37551 ATGATACCGC GAGACCCACG CTCACCGGCT CCAGATTTAT CAGCAATAAA
TACTATGGCG CTCTGGGTGC GAGTGGCCGA GGTCTAAATA GTCGTTATTT

37601 CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG
GGTCGGTCGG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC

37651 CCTCCATCCA GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG
GGAGGTAGGT CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC

37701 CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG GCATCGTGGT
GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC CGTAGCACCA

37751 GTCACGCTCG TCGTTTGGA TGGCTTCATT CAGCTCCGGT TCCCAACGAT
CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA

37801 CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAGC GGTTAGCTCC
GTTCCGCTCA ATGTACTAGG GGGTACAACA CGTTTTTTTCG CCAATCGAGG

37851 TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT
AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA

Figure 26 A N

37951 GATGCTTTTC TGTGACTGGT GAGTACTCAA CCAAGTCATT CTGAGAATAG
CTACGAAAAG ACACTGACCA CTCATGAGTT GGTTCAAGTAA GACTCTTATC

38001 TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAACAC GGGATAATAC
ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTGTG CCCTATTATG

38051 CGCGCCACAT AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT
GCGCGGTGTA TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA

38101 CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC CAGTTCGATG
GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG GTCAAGCTAC

38151 TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG
ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC

38201 CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA
GCAAAGACCC ACTCGTTTTT GTCCTTCCGT TTTACGGCGT TTTTTCCTT

38251 TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

38301 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT

38351 ATGTATTTAG AAAAATAAAC AAATAGGGGT TCCGCGCACA TTTCCCCGAA
TACATAAATC TTTTATTTG TTTATCCCCA AGGCGCGTGT AAAGGGGCTT

38401 AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT
TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA

38451 AAAAATAGGC GTATCACGAG GCCCTTTCGT CTTCAAGAAT TGGATCCGAA
TTTTTATCCG CATAGTGCTC CGGGAAAGCA GAAGTTCTTA ACCTAGGCTT

PacI

~~~~~

38501 TTCTTAATTT CTTAATTAA (SEQ ID NO:32)  
AAGAATTAAA GAATTAATT (SEQ ID NO:33)

*Figure 26 A0*

1 CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG  
GTAGTAGTTA TTATATGGAA TAAAACCTAA CTTCGGTTAT ACTATTACTC

51 GGGGTGGAGT TTGTGACGTG GCGCGGGGCG TGGGAACGGG GCGGGTGACG  
CCCCACCTCA AACACTGCAC CGCGCCCCGC ACCCTTGCCC CGCCCACTGC

101 TAGTAGTGTG GCGGAAGTGT GATGTTGCAA GTGTGGCGGA ACACATGTAA  
ATCATCACAC CGCCTTCACA CTACAACGTT CACACCGCCT TGTGTACATT

151 GCGACGGATG TGGCAAAAGT GACGTTTTTG GTGTGCGCCG GTGTACACAG  
CGCTGCCTAC ACCGTTTTCA CTGCAAAAAC CACACGCGGC CACATGTGTC

201 GAAGTGACAA TTTTCGCGCG GTTTTAGGCG GATGTTGTAG TAAATTTGGG  
CTTCACTGTT AAAAGCGCGC CAAAATCCGC CTACAACATC ATTTAAACCC

251 CGTAACCGAG TAAGATTTGG CCATTTTCGC GGGAAACTG AATAAGAGGA  
GCATTGGCTC ATTCTAAACC GGTAAAAGCG CCCTTTTGAC TTATTCTCCT

301 AGTGAAATCT GAATAATTTT GTGTTACTCA TAGCGCGTAA TATTTGTCTA  
TCACTTTAGA CTTATTAAAA CACAATGAGT ATCGCGCATT ATAAACAGAT

351 GGGCCGCGGG GACTTTGACC GTTTACGTGG AGACTCGCCC AGGTGTTTTT  
CCCGGCGCCC CTGAAACTGG CAAATGCACC TCTGAGCGGG TCCACAAAAA

401 CTCAGGTGTT TTCCGCGTTC CGGGTCAAAG TTGGCGTTTT ATTATTATAG  
GAGTCCACAA AAGGCGCAAG GCCCAGTTTC AACCGCAAAA TAATAATATC

451 GCGGCCGCGA TCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT  
CGCCGGCGCT AGGTAACGTA TGCAACATAG GTATAGTATT ATACATGTAA

501 TATATTGGCT CATGTCCAAC ATTACCGCCA TGTTGACATT GATTATTGAC  
ATATAACCGA GTACAGGTG TAATGGCGGT ACAACTGTAA CTAATAACTG

551 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA  
ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT

601 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG  
ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC

651 CCCAACGACC CCCGCCCAT TACGTCAATA ATGACGTATG TTCCCATAGT  
GGGTTGCTGG GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA

701 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT  
TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

751 AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG

801 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA  
GGATAACTGC AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT

Figure 27A

901 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA  
AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT

951 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA  
ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT

1001 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA  
ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT

1051 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG  
TGTTGAGGCG GGGTAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC

1101 GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1151 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC  
GGTAGGTGCG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG

1201 TCCGCGGCCG GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT  
AGGCGCCGGC CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA

1251 GAGATCTGCC ACCATGGCCG GCAAGTGGTC CAAGAGGTCC GTGCCCCGGCT  
CTCTAGACGG TGGTACCGGC CGTTCACCAG GTTCTCCAGG CACGGGCCGA

1301 GGTCCACCGT GAGGGAGAGG ATGAGGAGGG CCGAGCCCGC CGCCGACAGG  
CCAGGTGGCA CTCCCTCTCC TACTCCTCCC GGCTCGGGCG GCGGCTGTCC

1351 GTGAGGAGGA CCGAGCCCGC CGCAGTGGGC GTGGGCGCCG TGTCCAGGGA  
CACTCCTCCT GGCTCGGGCG GCGTCACCCG CACCCGCGGC ACAGGTCCCT

1401 CCTGGAGAAG CACGGCGCCA TCACCTCCTC CAACACCGCC GCCACCAACG  
GGACCTCTTC GTGCCGCGGT AGTGGAGGAG GTTGTGGCGG CGGTGGTTGC

1451 CCGACTGCGC CTGGCTGGAG GCCCAGGAGG ACGAGGAGGT GGGCTTCCCC  
GGCTGACGCG GACCGACCTC CGGGTCCTCC TGCTCCTCCA CCCGAAGGGG

1501 GTGAGGCCCC AGGTGCCCCCT GAGGCCCATG ACCTACAAGG GCGCCGTGGA  
CACTCCGGGG TCCACGGGGA CTCCGGGTAC TGGATGTTCC CGCGGCACCT

1551 CCTGTCCAC TTCCTGAAGG AGAAGGGCGG CCTGGAGGGC CTGATCCACT  
GGACAGGGTG AAGGACTTCC TCTTCCCGCC GGACCTCCCG GACTAGGTGA

1601 CCCAGAAGAG GCAGGACATC CTGGACCTGT GGGTGTACCA CACCCAGGGC  
GGGTCTTCTC CGTCCTGTAG GACCTGGACA CCCACATGGT GTGGGTCCCG

1651 TACTTCCCCG ACTGGCAGAA CTACACCCCC GGCCCCGGCA TCAGGTTCCT  
ATGAAGGGGC TGACCGTCTT GATGTGGGGG CCGGGGCCGT AGTCCAAGGG

1701 CCTGACCTTC GGCTGGTGCT TCAAGCTGGT GCCCGTGGAG CCCGAGAAGG  
GGA CTGGAAG CCGACCACGA AGTTCGACCA CGGGCACCTC GGGCTCTTCC

1751 TGGAGGAGGC CAACGAGGGC GAGAACAAC TCGCCGCCCA CCCCATGTCC  
ACCTCCTCCG GTTGCTCCCG CTCTTGTTGA CGCGGCGGGT GGGGTACAGG

Figure 27B

1851 CTCCAAGCTG GCCTTCCACC ACGTGGCCAG GGAGCTGCAC CCCGAGTACT  
GAGGTTCGAC CGGAAGGTGG TGCACCGGTC CCTCGACGTG GGGCTCATGA

1901 ACAAGGACTG CTAAAGCCCC GGCAGATCTG CTGTGCCTTC TAGTTGCCAG  
TGTTCCCTGAC GATTTTCGGGC CCGTCTAGAC GACACGGAAG ATCAACGGTC

1951 CCATCTGTTG TTTGCCCCCTC CCCCCTGCCT TCCTTGACCC TGGAAGGTGC  
GGTAGACAAC AAACGGGGAG GGGGCACGGA AGGAACCTGGG ACCTTCCACG

2001 CACTCCCACT GTCCTTTCCT AATAAAATGA GGAAATTGCA TCGCATTGTC  
GTGAGGGTGA CAGGAAAGGA TTATTTTACT CCTTTAACGT AGCGTAACAG

2051 TGAGTAGGTG TCATTCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG  
ACTCATCCAC AGTAAGATAA GACCCCCCACC CCCACCCCGT CCTGTCGTTC

2101 GGGGAGGATT GGGAAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGGCTC  
CCCCTCCTAA CCCTTCTGTT ATCGTCCGTA CGACCCCTAC GCCACCCGAG

2151 TATGGCCGAT CGGCGCGCCG TACTGAAATG TGTGGGCGTG GCTTAAGGGT  
ATACCGGCTA GCCGCGCGGC ATGACTTTAC ACACCCGCAC CGAATTCCCA

2201 GGGAAAGAAT ATATAAGGTG GGGGTCTTAT GTAGTTTTGT ATCTGTTTTG  
CCCTTTCTTA TATATTCCAC CCCAGAATA CATCAAAACA TAGACAAAAC

2251 CAGCAGCCGC CGCCGCCATG AGCACCAACT CGTTTGATGG AAGCATTGTG  
GTCGTCGGCG GCGGCGGTAC TCGTGGTTGA GCAAACCTACC TTCGTAACAC

2301 AGCTCATATT TGACAACGCG CATGCCCCCA TGGGCCGGGG TCGGTCAGAA  
TCGAGTATAA ACTGTTGCGC GTACGGGGGT ACCCGGCCCC ACGCAGTCTT

2351 TGTGATGGGC TCCAGCATTG ATGGTCGCCC CGTCCTGCCC GCAAACCTCTA  
ACACTACCCG AGGTCGTAAC TACCAGCGGG GCAGGACGGG CGTTTGAGAT

2401 CTACCTTGAC CTACGAGACC GTGTCTGGAA CGCCGTTGGA GACTGCAGCC  
GATGGAACCTG GATGCTCTGG CACAGACCTT GCGGCAACCT CTGACGTCGG

2451 TCCGCCGCCG CTTAGCCGC TGCAGCCACC GCCCGCGGGA TTGTGACTGA  
AGGCGGCGGC GAAGTCGGCG ACGTCGGTGG CGGGCGCCCT AACACTGACT

2501 CTTTGCTTTC CTGAGCCCGC TTGCAAACAG TGCAGCTTCC CGTTCATCCG  
GAAACGAAAG GACTCGGGCG AACGTTTGTC ACGTCGAAGG GCAAGTAGGC

2551 CCCGCGATGA CAAGTTGACG GCTCTTTTGG CACAATTGGA TTCTTTGACC  
GGGCGCTACT GTTCAACTGC CGAGAAAACC GTGTTAACCT AAGAACTGG

2601 CGGGAACTTA ATGTCGTTTC TCAGCAGCTG TTGGATCTGC GCCAGCAGGT  
GCCCTTGAAT TACAGCAAAG AGTCGTCGAC AACCTAGACG CGGTCGTCCA

2651 TTCTGCCCTG AAGGCTTCCT CCCCTCCCAA TGCGGTTTAA AACATAAATA  
AAGACGGGAC TTCCGAAGGA GGGGAGGGTT ACGCCAAATT TTGTATTTAT

2701 AAAAACCAGA CTCTGTTTGG ATTTGGATCA AGCAAGTGTC TTGCTGTCTT  
TTTTTGGTCT GAGACAAACC TAAACCTAGT TCGTTCACAG AACGACAGAA

Figure 27C

2751 TATTTAGGGG TTTTGCGCGC GCGGTAGGCC CGGGACCAGC GGTCTCGGTC  
ATAAATCCCC AAAACGCGCG CGCCATCCGG GCCCTGGTCG CCAGAGCCAG

2801 GTTGAGGGTC CTGTGTATTT TTTCCAGGAC GTGGTAAAGG TGA CTCTGGA  
CAACTCCCAG GACACATAAA AAAGGTCCTG CACCATTTC ACTGAGACCT

2851 TG TTCAGATA CATGGGCATA AGCCCGTCTC TGGGGTGGAG GTAGCACCAC  
ACAAGTCTAT GTACCCGTAT TCGGGCAGAG ACCCCACCTC CATCGTGGTG

2901 TGCAGAGCTT CATGCTGCGG GGTGGTGTG TAGATGATCC AGTCGTAGCA  
ACGTCTCGAA GTACGACGCC CCACCACAAC ATCTACTAGG TCAGCATCGT

2951 GGAGCGCTGG GCGTGGTGCC TAAAAATGTC TTTCAGTAGC AAGCTGATTG  
CCTCGCGACC CGCACCACGG ATTTTACAG AAAGTCATCG TTCGACTAAC

3001 CCAGGGGCAG GCCCTTGGTG TAAGTGTTTA CAAAGCGGTT AAGCTGGGAT  
GGTCCCCGTC CGGGAACCAC ATTCACAAAT GTTTCGCCAA TTCGACCCTA

3051 GGGTGCATAC GTGGGGATAT GAGATGCATC TTGGACTGTA TTTT TAGGT  
CCCACGTATG CACCCCTATA CTCTACGTAG AACCTGACAT AAAAATCCAA

3101 GGCTATGTTT CCAGCCATAT CCCTCCGGGG ATTCATGTTG TGCAGAACCA  
CCGATACAAG GGTGCGGTATA GGGAGGCCCC TAAGTACAAC ACGTCTTGGT

3151 CCAGCACAGT GTATCCGGTG CACTTGGGAA ATTTGTCATG TAGCTTAGAA  
GGTCGTGTCA CATAGGCCAC GTGAACCCTT TAAACAGTAC ATCGAATCTT

3201 GGAAATGCGT GGAAGAACTT GGAGACGCCC TTGTGACCTC CAAGATTTTC  
CCTTTACGCA CCTTCTTGAA CCTCTGCGGG AACACTGGAG GTTCTAAAAG

3251 CATGCATTCTG TCCATAATGA TGGCAATGGG CCCACGGGCG GCGGCCTGGG  
GTACGTAAGC AGGTATTACT ACCGTTACCC GGGTGCCCGC CGCCGGACCC

3301 CGAAGATATT TCTGGGATCA CTAACGTCAT AGTTGTGTTT CAGGATGAGA  
GCTTCTATAA AGACCCTAGT GATTGCAGTA TCAACACAAG GTCCTACTCT

3351 TCGTCATAGG CCATTTTAC AAAGCGCGGG CGGAGGGTGC CAGACTGCGG  
AGCAGTATCC GGTAAAAATG TTTCGCGCCC GCCTCCCACG GTCTGACGCC

3401 TATAATGGTT CCATCCGGCC CAGGGGCGTA GTTACCCTCA CAGATTTGCA  
ATATTACCAA GGTAGGCCGG GTCCCCGCAT CAATGGGAGT GTCTAAACGT

3451 TTTCCACGCT TTTGAGTTCA GATGGGGGGA TCATGTCTAC CTGCGGGGCG  
AAAGGGTGCG AACTCAAGT CTACCCCTT AGTACAGATG GACGCCCCGC

3501 ATGAAGAAAA CGGTTTCCGG GGTAGGGGAG ATCAGCTGGG AAGAAAGCAG  
TACTTCTTTT GCCAAAGGCC CCATCCCCTC TAGTCGACCC TTCTTTCGTC

3551 GTTCCTGAGC AGCTGCGACT TACCGCAGCC GGTGGGCCCC TAAATCACAC  
CAAGGACTCG TCGACGCTGA ATGGCGTCGG CCACCCGGGC ATTTAGTGTG

3601 CTATTACCGG CTGCAACTGG TAGTTAAGAG AGCTGCAGCT GCCGTCATCC  
GATAATGGCC GACGTTGACC ATCAATTCTC TCGACGTCGA CGGCAGTAGG

3651 CTGAGCAGGG GGGCCACTTC GTTAAGCATG TCCCTGACTC GCATGTTTTT  
GACTCGTCCC CCCGGTGAAG CAATTCGTAC AGGGACTGAG CGTACAAAAG

Figure 27D

3701 CCTGACCAAA TCCGCCAGAA GCGCTCGCC GCCCAGCGAT AGCAGTTCTT  
GGACTGGTTT AGGCGGTCTT CCGCGAGCGG CGGGTCGCTA TCGTCAAGAA

3751 GCAAGGAAGC AAAGTTTTTC AACGGTTTGA GACCGTCCGC CGTAGGCATG  
CGTTCCTTCG TTTCAAAAAG TTGCCAAACT CTGGCAGGCG GCATCCGTAC

3801 CTTTTGAGCG TTTGACCAAG CAGTTCCAGG CGGTCCCACA GCTCGGTCAC  
GAAAACTCGC AAACCTGGTTC GTCAAGGTCC GCCAGGGTGT CGAGCCAGTG

3851 CTGCTCTACG GCATCTCGAT CCAGCATATC TCCTCGTTTC GCGGGTTGGG  
GACGAGATGC CGTAGAGCTA GGTCGTATAG AGGAGCAAAG CGCCCAACCC

3901 GCGGCTTTTCG CTGTACGGCA GTAGTCGGTG CTCGTCCAGA CGGGCCAGGG  
CGCCGAAAGC GACATGCCGT CATCAGCCAC GAGCAGGTCT GCCCGGTCCC

3951 TCATGTCTTT CCACGGGCGC AGGGTCCTCG TCAGCGTAGT CTGGGTCACG  
AGTACAGAAA GGTGCCCCGC TCCCAGGAGC AGTCGCATCA GACCCAGTGC

4001 GTGAAGGGGT GCGCTCCGGG CTGCGCGCTG GCCAGGGTGC GCTTGAGGCT  
CACTTCCCCA CGCGAGGCCC GACGCGCGAC CGGTCCCACG CGAACTCCGA

4051 GGTCCTGCTG GTGCTGAAGC GCTGCCGGTC TTCGCCCTGC GCGTCGGCCA  
CCAGGACGAC CACGACTTCG CGACGGCCAG AAGCGGGACG CGCAGCCGGT

4101 GGTAGCATTT GACCATGGTG TCATAGTCCA GCCCTCCGC GCGGTGGCCC  
CCATCGTAAA CTGGTACCAC AGTATCAGGT CGGGGAGGCG CCGCACCGGG

4151 TTGGCGCGCA GCTTGCCCTT GGAGGAGGCG CCGCACGAGG GGCAGTGCAG  
AACCGCGCGT CGAACGGGAA CCTCCTCCGC GCGTGCTCC CCGTCACGTC

4201 ACTTTTGAGG GCGTAGAGCT TGGGCGCGAG AAATACCGAT TCCGGGGAGT  
TGAAAACCTC CGCATCTCGA ACCCGCGCTC TTTATGGCTA AGGCCCTCA

4251 AGGCATCCGC GCCGCAGGCC CCGCAGACGG TCTCGCATTC CACGAGCCAG  
TCCGTAGGCG CGGCGTCCGG GCGTCTGCC AGAGCGTAAG GTGCTCGGTC

4301 GTGAGCTCTG GCCGTTCGGG GTCAAAAACC AGGTTTCCCC CATGCTTTTT  
CACTCGAGAC CGGCAAGCCC CAGTTTTTGG TCCAAAGGGG GTACGAAAAA

4351 GATGCGTTTC TTACCTCTGG TTTCCATGAG CCGGTGTCCA CGCTCGGTGA  
CTACGCAAAG AATGGAGACC AAAGGTACTC GGCCACAGGT GCGAGCCACT

4401 CGAAAAGGCT GTCCGTGTCC CCGTATACAG ACTTGAGAGG CCTGTCCTCG  
GCTTTTCCGA CAGGCACAGG GGCATATGTC TGAACCTCTC GGACAGGAGC

4451 AGCGGTGTTC CGCGGTCCTC CTCGTATAGA AACTCGGACC ACTCTGAGAC  
TCGCCACAAG GCGCCAGGAG GAGCATATCT TTGAGCCTGG TGAGACTCTG

4501 AAAGGCTCGC GTCCAGGCCA GCACGAAGGA GGCTAAGTGG GAGGGGTAGC  
TTTCCGAGCG CAGGTCCGGT CGTGCTTCCT CCGATTCACC CTCCCCATCG

4551 GGTCGTTGTC CACTAGGGGG TCCACTCGCT CCAGGGTGTG AAGACACATG  
CCAGCAACAG GTGATCCCCC AGGTGAGCGA GGTCCCACAC TTCTGTGTAC

4601 TCGCCCTCTT CGGCATCAAG GAAGGTGATT GGTTTGTAGG TGTAGGCCAC  
AGCGGGAGAA GCCGTAGTTC CTTCCACTAA CCAAACATCC ACATCCGGTG

Figure 27E

```

4701  CGTCCTCACT CTCTTCGCA TCGCTGTCTG CGAGGGCCAG CTGTTGGGGT
      GCAGGAGTGA GAGAAGGCGT AGCGACAGAC GCTCCCGGTC GACAACCCCA

4751  GAGTACTCCC TCTGAAAAGC GGGCATGACT TCTGCGCTAA GATTGTCAGT
      CTCATGAGGG AGACTTTTTCG CCCGTACTGA AGACGCGATT CTAACAGTCA

4801  TTCCAAAAC GAGGAGGATT TGATATTAC CTGGCCCGCG GTGATGCCTT
      AAGGTTTTTG CTCCTCCTAA ACTATAAGTG GACCGGGCGC CACTACGGAA

4851  TGAGGGTGGC CGCATCCATC TGGTCAGAAA AGACAATCTT TTTGTTGTCA
      ACTCCCACCG GCGTAGGTAG ACCAGTCTTT TCTGTTAGAA AAACAACAGT

4901  AGCTTGGTGG CAAACGACCC GTAGAGGGCG TTGGACAGCA ACTTGGCGAT
      TCGAACCACC GTTTGCTGGG CATCTCCCGC AACCTGTCGT TGAACCGCTA

4951  GGAGCGCAGG GTTTGGTTTT TGTCGCGATC GCGCGCTCC TTGGCCGCGA
      CCTCGCGTCC CAAACCAAAA ACAGCGCTAG CCGCGCGAGG AACC GGCGCT

5001  TGTTTAGCTG CACGTATTCG CGCGCAACGC ACCGCCATTC GGGAAAGACG
      ACAAATCGAC GTGCATAAGC GCGCGTTGCG TGGCGGTAAG CCCTTTCTGC

5051  GTGGTGCCTG CGTCGGGCAC CAGGTGCACG CGCCAACCGC GGTGTGTCAG
      CACCACGCGA GCAGCCCGTG GTCCACGTGC GCGGTTGGCG CCAACACGTC

5101  GGTGACAAGG TCAACGCTGG TGGCTACCTC TCCGCGTAGG CGCTCGTTGG
      CCACTGTTCC AGTTGCGACC ACCGATGGAG AGGCGCATCC GCGAGCAACC

5151  TCCAGCAGAG GCGGCCGCCC TTGCGCGAGC AGAATGGCGG TAGGGGGTCT
      AGGTCGTCTC CGCCGGCGGG AACGCGCTCG TCTTACCGCC ATCCCCCAGA

5201  AGCTGCGTCT CGTCCGGGGG GTCTGCGTCC ACGGTAAAGA CCCC GGCGCAG
      TCGACGCAGA GCAGGCCCCC CAGACGCAGG TGCCATTTCT GGGGCCCGTC

5251  CAGGCGCGCG TCGAAGTAGT CTATCTTGCA TCCTTGCAAG TCTAGCGCCT
      GTCCGCGCGC AGCTTCATCA GATAGAACGT AGGAACGTTC AGATCGCGGA

5301  GCTGCCATGC GCGGGCGGCA AGCGCGCGCT CGTATGGGTT GAGTGGGGGA
      CGACGGTACG CGCCCGCCGT TCGCGCGCGA GCATACCCAA CTCACCCCTT

5351  CCCCATGGCA TGGGGTGGGT GAGCGCGGAG GCGTACATGC CGCAAATGTC
      GGGGTACCGT ACCCCACCCA CTCGCGCCTC CGCATGTACG GCGTTTACAG

5401  GTAAACGTAG AGGGGCTCTC TGAGTATTCC AAGATATGTA GGGTAGCATC
      CATTTGCATC TCCCCGAGAG ACTCATAAGG TTCTATACAT CCCATCGTAG

5451  TTCCACCGCG GATGCTGGCG CGCACGTAAT CGTATAGTTC GTGCGAGGGA
      AAGGTGGCGC CTACGACCGC GCGTGCTATTA GCATATCAAG CACGCTCCCT

5501  GCGAGGAGGT CGGGACCGAG GTTGCTACGG GCGGGCTGCT CTGCTCGGAA
      CGCTCCTCCA GCCCTGGCTC CAACGATGCC CGCCCGACGA GACGAGCCTT

5551  GACTATCTGC CTGAAGATGG CATGTGAGTT GGATGATATG GTTGGACGCT
      CTGATAGACG GACTTCTACC GTACACTCAA CCTACTATAC CAACCTGCGA

```

Figure 27F

5651 GAGGCGTAGG AGTCGCGCAG CTTGTTGACC AGCTCGGCGG TGACCTGCAC  
CTCCGCATCC TCAGCGCGTC GAACAACTGG TCGAGCCGCC ACTGGACGTG

5701 GTCTAGGGCG CAGTAGTCCA GGGTTTCCTT GATGATGTCA TACTTATCCT  
CAGATCCCGC GTCATCAGGT CCCAAAGGAA CTACTACAGT ATGAATAGGA

5751 GTCCCTTTTT TTTCCACAGC TCGCGGTGA GGACAACTC TTCGCGGTCT  
CAGGGAAAAA AAAGGTGTCG AGCGCCAACCT CCTGTTTGAG AAGCGCCAGA

5801 TTCCAGTACT CTTGGATCGG AAACCCGTCG GCCTCCGAAC GGTAAGAGCC  
AAGGTCATGA GAACCTAGCC TTTGGGCAGC CGGAGGCTTG CCATTCTCGG

5851 TAGCATGTAG AACTGGTTGA CGGCCTGGTA GGCGCAGCAT CCCTTTTCTA  
ATCGTACATC TTGACCAACT GCCGGACCAT CCGCGTCGTA GGGAAAAGAT

5901 CGGGTAGCGC GTATGCCTGC GCGGCCTTCC GGAGCGAGGT GTGGGTGAGC  
GCCCATCGCG CATACGGACG CGCCGGAAGG CCTCGCTCCA CACCCACTCG

5951 GCAAAGGTGT CCCTGACCAT GACTTTGAGG TACTGGTATT TGAAGTCAGT  
CGTTTCCACA GGGACTGGTA CTGAAACTCC ATGACCATAA ACTTCAGTCA

6001 GTCGTCGCAT CCGCCCTGCT CCCAGAGCAA AAAGTCCGTG CGCTTTTGG  
CAGCAGCGTA GGCGGGACGA GGGTCTCGTT TTTCAGGCAC GCGAAAACC

6051 AACGCGGATT TGGCAGGGCG AAGGTGACAT CGTTGAAGAG TATCTTCCC  
TTGCGCCTAA ACCGTCCCGC TTCCACTGTA GCAACTTCTC ATAGAAAGGG

6101 GCGCGAGGCA TAAAGTTGCG TGTGATGCGG AAGGGTCCCG GCACCTCGGA  
CGCGCTCCGT ATTTCAACGC AACTACGCC TTCCAGGGC CGTGGAGCCT

6151 ACGGTTGTTA ATTACCTGGG CGGCGAGCAC GATCTCGTCA AAGCCGTTGA  
TGCCAACAAT TAATGGACCC GCCGCTCGTG CTAGAGCAGT TTCGGCAACT

6201 TGTTGTGGCC CACAATGTAA AGTTCCAAGA AGCGCGGGAT GCCCTTGATG  
ACAACACCGG GTGTTACATT TCAAGGTCTT TCGCGCCCTA CGGGAACCTAC

6251 GAAGGCAATT TTTTAAGTTC CTCGTAGGTG AGCTCTTCAG GGGAGCTGAG  
CTTCCGTTAA AAAATTCAAG GAGCATCCAC TCGAGAAGTC CCCTCGACTC

6301 CCCGTGCTCT GAAAGGGCCC AGTCTGCAAG ATGAGGGTTG GAAGCGACGA  
GGGCACGAGA CTTTCCCGGG TCAGACGTTT TACTCCCAAC CTTCGCTGCT

6351 ATGAGCTCCA CAGGTCACGG GCCATTAGCA TTTGCAGGTG GTCGCGAAAG  
TACTCGAGGT GTCCAGTGCC CGGTAATCGT AAACGTCCAC CAGCGCTTTC

6401 GTCCTAAACT GGCGACCTAT GGCCATTTTT TCTGGGGTGA TGCAGTAGAA  
CAGGATTTGA CCGCTGGATA CCGGTAAAAA AGACCCCACT ACGTCATCTT

6451 GGTAAGCGGG TCTTGTTCCC AGCGGTCCCA TCCAAGGTTC GCGGCTAGGT  
CCATTCGCCC AGAACAAGGG TCGCCAGGGT AGGTTCCAAG CGCCGATCCA

6501 CTCGCGCGGC AGTCACTAGA GGCTCATCTC CGCCGAACTT CATGACCAGC  
GAGCGCGCCG TCAGTGATCT CCGAGTAGAG GCGGCTTGAA GTACTGGTCG

Figure 27G

6601 TACATCGTAG GTGACAAAGA GACGCTCGGT GCGAGGATGC GAGCCGATCG  
ATGTAGCATC CACTGTTTCT CTGCGAGCCA CGCTCCTACG CTCGGCTAGC

6651 GGAAGAACTG GATCTCCCGC CACCAATTGG AGGAGTGGCT ATTGATGTGG  
CCTTCTTGAC CTAGAGGGCG GTGGTTAACC TCCTCACCGA TAACTACACC

6701 TGAAAGTAGA AGTCCCTGCG ACGGGCCGAA CACTCGTGCT GGCTTTTGTA  
ACTTTCATCT TCAGGGACGC TGCCCGGCTT GTGAGCACGA CCGAAAACAT

6751 AAAACGTGCG CAGTACTGGC AGCGGTGCAC GGGCTGTACA TCCTGCACGA  
TTTTGCACGC GTCATGACCG TCGCCACGTG CCCGACATGT AGGACGTGCT

6801 GGTTGACCTG ACGACCGCGC ACAAGGAAGC AGAGTGGGAA TTTGAGCCCC  
CCAACTGGAC TGCTGGCGCG TGTTCCTTCG TCTCACCTT AAACTCGGGG

6851 TCGCCTGGCG GGTTTGGCTG GTGGTCTTCT ACTTCGGCTG CTTGTCCTTG  
AGCGGACCGC CCAAACCGAC CACCAGAAGA TGAAGCCGAC GAACAGGAAC

6901 ACCGTCTGGC TGCTCGAGGG GAGTTACGGT GGATCGGACC ACCACGCCGC  
TGGCAGACCG ACGAGCTCCC CTCAATGCCA CCTAGCCTCG TGGTGCGGCG

6951 GCGAGCCCAA AGTCCAGATG TCCGCGCGCG GCGGTCGGAG CTTGATGACA  
CGCTCGGGTT TCAGGTCTAC AGGCGCGCGC CGCCAGCCTC GAACTACTGT

7001 ACATCGCGCA GATGGGAGCT GTCCATGGTC TGGAGCTCCC GCGGCGTCAG  
TGTAGCGCGT CTACCCTCGA CAGGTACCAG ACCTCGAGGG CGCCGCAGTC

7051 GTCAGGCGGG AGCTCCTGCA GGTTTACCTC GCATAGACGG GTCAGGGCGC  
CAGTCCGCCC TCGAGGACGT CCAAATGGAG CGTATCTGCC CAGTCCCGCG

7101 GGGCTAGATC CAGGTGATAC CTAATTTCCA GGGGCTGGTT GGTGGCGGCG  
CCCGATCTAG GTCCACTATG GATTAAAGGT CCCCACCAA CCACCGCCGC

7151 TCGATGGCTT GCAAGAGGCC GCATCCCCGC GGCGCGACTA CGGTACCGCG  
AGCTACCGAA CGTTCTCCGG CGTAGGGGCG CCGCGCTGAT GCCATGGCGC

7201 CGGCGGGCGG TGGGCCGCGG GGGTGTCTT GGATGATGCA TCTAAAAGCG  
GCCGCCCCGC ACCCGGCGCC CCCACAGGAA CCTACTACGT AGATTTTCGC

7251 GTGACGCGGG CGAGCCCCCG GAGGTAGGGG GGGCTCCGGA CCCGCCGGGA  
CACTGCGCCC GCTCGGGGGC CTCCATCCCC CCCGAGGCCT GGGCGGCCCT

7301 GAGGGGGCAG GGGCACGTCG GCGCCGCGCG CGGGCAGGAG CTGGTGCTGC  
CTCCCCCGTC CCCGTGCAGC CGCGGCGCGC GCCCGTCCTC GACCACGACG

7351 GCGCGTAGGT TGCTGGCGAA CGCGACGACG CGGCGGTTGA TCTCCTGAAT  
CGCGCATCCA ACGACCGCTT GCGCTGCTGC GCCGCCAACT AGAGGACTTA

7401 CTGGCGCCTC TGCGTGAAGA CGACGGGCCC GGTGAGCTTG AACCTGAAAG  
GACCGCGGAG ACGCACTTCT GCTGCCCCGG CCACTCGAAC TTGGACTTTC

7451 AGAGTTCGAC AGAATCAATT TCGGTGTCGT TGACGGCGGC CTGGCGCAAA  
TCTCAAGCTG TCTTAGTTAA AGCCACAGCA ACTGCCGCCG GACCGCGTTT

Figure 27H

7551 CTGCTCGATC TCTTCCTCCT GGAGATCTCC GCGTCCGGCT CGCTCCACGG  
GACGAGCTAG AGAAGGAGGA CCTCTAGAGG CGCAGGCCGA GCGAGGTGCC

7601 TGGCGGCGAG GTCGTTGGAA ATGCGGGCCA TGAGCTGCGA GAAGGCGTTG  
ACCGCCGCTC CAGCAACCTT TACGCCCGGT ACTCGACGCT CTTCCGCAAC

7651 AGGCCTCCCT CGTTCCAGAC GCGGCTGTAG ACCACGCCCC CTTCGGCATC  
TCCGGAGGGA GCAAGGTCTG CGCCGACATC TGGTGCGGGG GAAGCCGTAG

7701 GCGGGCGCGC ATGACCACCT GCGCGAGATT GAGCTCCACG TGCCGGGCGA  
CGCCCGCGCG TACTGGTGGA CGCGCTCTAA CTCGAGGTGC ACGGCCCGCT

7751 AGACGGCGTA GTTTCGCAGG CGCTGAAAGA GGTAGTTGAG GGTGGTGGCG  
TCTGCCGCAT CAAAGCGTCC GCGACTTTCT CCATCAACTC CCACCACCGC

7801 GTGTGTTCTG CCACGAAGAA GTACATAACC CAGCGTCGCA ACGTGGATTC  
CACACAAGAC GGTGCTTCTT CATGTATTGG GTCGCAGCGT TGCACCTAAG

7851 GTTGATATCC CCCAAGGCCT CAAGGCGCTC CATGGCCTCG TAGAAGTCCA  
CAACTATAGG GGGTTCGGA GTTCCGCGAG GTACCGGAGC ATCTTCAGGT

7901 CGGCGAAGTT GAAAACTGG GAGTTGCGCG CCGACACGGT TAACTCCTCC  
GCCGCTTCAA CTTTTTGACC CTCAACGCGC GGCTGTGCCA ATTGAGGAGG

7951 TCCAGAAGAC GGATGAGCTC GGCGACAGTG TCGCGCACCT CGCGCTCAAA  
AGGTCTTCTG CCTACTCGAG CCGCTGTCAC AGCGCGTGGA GCGCGAGTTT

8001 GGCTACAGGG GCCTCTTCTT CTTCTTCAAT CTCCTCTTCC ATAAGGGCCT  
CCGATGTCCC CGGAGAAGAA GAAGAAGTTA GAGGAGAAGG TATTCCCGGA

8051 CCCCTTCTTC TTCTTCTGGC GGCGGTGGGG GAGGGGGGAC ACGGCGGCGA  
GGGGAAGAAG AAGAAGACCG CCGCCACCCC CTCCCCCTG TGCCGCCGCT

8101 CGACGGCGCA CCGGGAGGCG GTCGACAAAG CGCTCGATCA TCTCCCCGCG  
GCTGCCGCGT GGCCCTCCGC CAGCTGTTTC GCGAGCTAGT AGAGGGGCGC

8151 GCGACGGCGC ATGGTCTCGG TGACGGCGCG GCCGTTCTCG CGGGGGCGCA  
CGCTGCCGCG TACCAGAGCC ACTGCCGCGC CGGCAAGAGC GCCCCGCGT

8201 GTTGGAAGAC GCCGCCGTC ATGTCCCGGT TATGGGTTGG CGGGGGGCTG  
CAACCTTCTG CGGCGGGCAG TACAGGGCCA ATACCCAACC GCCCCCGAC

8251 CCATGCGGCA GGGATACGGC GCTAACGATG CATCTCAACA ATTGTTGTGT  
GGTACGCCGT CCCTATGCCG CGATTGCTAC GTAGAGTTGT TAACAACACA

8301 AGGTACTCCG CCGCCGAGGG ACCTGAGCGA GTCCGCATCG ACCGGATCGG  
TCCATGAGGC GCGGGCTCCC TGGACTCGCT CAGGCGTAGC TGGCCTAGCC

8351 AAAACCTCTC GAGAAAGGCG TCTAACCAGT CACAGTCGCA AGGTAGGCTG  
TTTTGGAGAG CTCTTTCCGC AGATTGGTCA GTGTCAGCGT TCCATCCGAC

8401 AGCACCGTGG CGGGCGGCAG CGGGCGGCGG TCGGGGTTGT TTCTGGCGGA  
TCGTGGCACC GCCCGCCGTC GCCCGCCGCC AGCCCCAACA AAGACCGCCT

Figure 27I

8501 TCGACAGAAG CACCATGTCC TTGGGTCCGG CCTGCTGAAT GCGCAGGCGG  
AGCTGTCTTC GTGGTACAGG AACCCAGGCC GGACGACTTA CGCGTCCGCC

8551 TCGGCCATGC CCCAGGCTTC GTTTTGACAT CGGCGCAGGT CTTTGTAGTA  
AGCCGGTACG GGGTCCGAAG CAAAACTGTA GCCGCGTCCA GAAACATCAT

8601 GTCTTG CATG AGCCTTTCTA CCGGCACTTC TTCTTCTCCT TCCTCTTGTC  
CAGAACGTAC TCGGAAAGAT GGCCGTGAAG AAGAAGAGGA AGGAGAACAG

8651 CTGCATCTCT TGCATCTATC GCTGCGGCGG CGGCGGAGTT TGGCCGTAGG  
GACGTAGAGA ACGTAGATAG CGACGCCGCC GCCGCCTCAA ACCGGCATCC

8701 TGGCGCCCTC TTCCTCCCAT GCGTGTGACC CCGAAGCCCC TCATCGGCTG  
ACCGCGGGAG AAGGAGGGTA CGCACACTGG GGCTTCGGGG AGTAGCCGAC

8751 AAGCAGGGCT AGGTCCGGCGA CAACGCGCTC GGCTAATATG GCCTGCTGCA  
TTCGTCCCGA TCCAGCCGCT GTTGCGCGAG CCGATTATAC CGGACGACGT

8801 CCTGCGTGAG GG TAGACTGG AAGTCATCCA TGTCCACAAA GCGGTGGTAT  
GGACGCACTC CCATCTGACC TTCAGTAGGT ACAGGTGTTT CGCCACCATA

8851 GCGCCCGTGT TGATGGTGTA AGTGCAGTTG GCCATAACGG ACCAGTTAAC  
CGCGGGCACA ACTACCACAT TCACGTCAAC CGGTATTGCC TGGTCAATTG

8901 GGTCTGGTGA CCCGGCTGCG AGAGCTCGGT GTACCTGAGA CGCGAGTAAG  
CCAGACCACT GGGCCGACGC TCTCGAGCCA CATGGACTCT GCGCTCATTC

8951 CCCTCGAGTC AAATACGTAG TCGTTGCAAG TCCGCACCAG GTACTGGTAT  
GGGAGCTCAG TTTATGCATC AGCAACGTTC AGGCGTGGTC CATGACCATA

9001 CCCACCAAAA AGTGCGGCGG CGGCTGGCGG TAGAGGGGCC AGCGTAGGGT  
GGGTGGTTTT TCACGCCGCC GCCGACCGCC ATCTCCCCGG TCGCATCCCA

9051 GGCCGGGGCT CCGGGGGCGA GATCTTCCAA CATAAGGCGA TGATATCCGT  
CCGGCCCCGA GGCCCCCGCT CTAGAAGGTT GTATTCCGCT ACTATAGGCA

9101 AGATGTACCT GGACATCCAG GTGATGCCGG CGGCGGTGGT GGAGGCGCGC  
TCTACATGGA CCTGTAGGTC CACTACGGCC GCCGCCACCA CCTCCGCGCG

9151 GGAAAGTCGC GGACGCGGTT CCAGATGTTG CGCAGCGGCA AAAAGTGCTC  
CCTTTCAGCG CCTGCGCCAA GGTCTACAAC GCGTCGCCGT TTTTCACGAG

9201 CATGGTCGGG ACGCTCTGGC CGGTCAGGCG CGCGCAATCG TTGACGCTCT  
GTACCAGCCC TGCAGAGACCG GCCAGTCCGC GCGCGTTAGC AACTGCGAGA

9251 AGACCGTGCA AAAGGAGAGC CTGTAAGCGG GCACTCTTCC GTGGTCTGGT  
TCTGGCACGT TTTCTCTCG GACATTCGCC CGTGAGAAGG CACCAGACCA

9301 GGATAAATTC GCAAGGGTAT CATGGCGGAC GACCGGGGTT CGAGCCCCGT  
CCTATTTAAG CGTTCCCATATA GTACCGCCTG CTGGCCCCAA GCTCGGGGCA

9351 ATCCGGCCGT CCGCCGTGAT CCATGCGGTT ACCGCCCGCG TGTCGAACCC  
TAGGCCGGCA GCGGGCACTA GGTACGCCAA TGGCGGGCGC ACAGCTTGGG

Figure 27J

9451 GGC GCGGCGG CTGCTGCGCT AGCTTTTTTTG GCCACTGGCC GCGCGCAGCG  
CCGCGCCGCC GACGACGCGA TCGAAAAAAC CGGTGACCGG CGCGCGTCGC

9501 TAAGCGGTTA GGCTGGAAAG CGAAAGCATT AAGTGGCTCG CTCCCTGTAG  
ATTCGCCAAT CCGACCTTTC GCTTTCGTAA TTCACCGAGC GAGGGACATC

9551 CCGGAGGGTT ATTTTCCAAG GGTTGAGTCG CGGGACCCCC GTTTCGAGTC  
GGCCTCCCAA TAAAAGGTTC CCAACTCAGC GCCCTGGGGG CCAAGCTCAG

9601 TCGGACCGGC CGGACTGCGG CGAACGGGGG TTTGCCTCCC CGTCATGCAA  
AGCCTGGCCG GCCTGACGCC GCTTGCCCCC AAACGGAGGG GCAGTACGTT

9651 GACCCCGCTT GCAAATTCCT CCGGAAACAG GGACGAGCCC CTTTTTTGCT  
CTGGGGCGAA CGTTTAAGGA GGCCTTTGTC CCTGCTCGGG GAAAAAACGA

9701 TTTCCAGAT GCATCCGGTG CTGCGGCAGA TCGCCCCCCC TCCTCAGCAG  
AAAGGGTCTA CGTAGGCCAC GACGCCGTCT ACGCGGGGGG AGGAGTCGTC

9751 CGGCAAGAGC AAGAGCAGCG GCAGACATGC AGGGCACCCCT CCCCTCCTCC  
GCCGTTCTCG TTCTCGTCGC CGTCTGTACG TCCCGTGGGA GGGGAGGAGG

9801 TACCGCGTCA GGAGGGGCGA CATCCGCGGT TGACGCGGCA GCAGATGGTG  
ATGGCGCAGT CCTCCCCGCT GTAGGCGCCA ACTGCGCCGT CGTCTACCAC

9851 ATTACGAACC CCCGCGGCGC CGGGCCCCGGC ACTACCTGGA CTTGGAGGAG  
TAATGCTTGG GGGCGCCGCG GCCCGGGCCG TGATGGACCT GAACCTCCTC

9901 GGCGAGGGCC TGGCGCGGCT AGGAGCGCCC TCTCCTGAGC GGCACCCAAG  
CCGCTCCCGG ACCGCGCCGA TCCTCGCGGG AGAGGACTCG CCGTGGGTTC

9951 GGTGCAGCTG AAGCGTGATA CGCGTGAGGC GTACGTGCCG CGGCAGAACC  
CCACGTCGAC TTCGCACTAT GCGCACTCCG CATGCACGGC GCCGTCTTGG

10001 TGTTTCGCGA CCGCGAGGGA GAGGAGCCCC AGGAGATGCG GGATCGAAAG  
ACAAAGCGCT GGCCTCCCT CTCCTCGGGC TCCTCTACGC CCTAGCTTTC

10051 TTCCACGCAG GGCGCGAGCT GCGGCATGGC CTGAATCGCG AGCGGTTGCT  
AAGGTGCGTC CCGCGCTCGA CGCCGTACCG GACTTAGCGC TCGCCAACGA

10101 GCGCGAGGAG GACTTTGAGC CCGACGCGCG AACCGGGATT AGTCCCGCGC  
CGCGCTCCTC CTGAAACTCG GGCTGCGCGC TTGGCCCTAA TCAGGGCGCG

10151 GCGCACACGT GGCGGCCGCC GACCTGGTAA CCGCATACGA GCAGACGGTG  
CGCGTGTGCA CCGCCGGCGG CTGGACCATT GGCGTATGCT CGTCTGCCAC

10201 AACCAGGAGA TTAAC TTCA AAAAGCTTT AACAACCACG TCGGTACGCT  
TTGGTCCTCT AATTGAAAGT TTTTTCGAAA TTGTTGGTGC ACGCATGCGA

10251 TGTGGCGCGC GAGGAGGTGG CTATAGGACT GATGCATCTG TGGGACTTTG  
ACACCGCGCG CTCCTCCACC GATATCCTGA CTACGTAGAC ACCCTGAAAC

10301 TAAGCGCGCT GGAGCAAAAC CCAAATAGCA AGCCGCTCAT GGCGCAGCTG  
ATTCGCGCGA CCTCGTTTTG GGT TTATCGT TCGGCGAGTA CCGCGTCGAC

Figure 27K

10401 GCTAAACATA GTAGAGCCCG AGGGCCGCTG GCTGCTCGAT TTGATAAACA  
CGATTTGTAT CATCTCGGGC TCCCGGCGAC CGACGAGCTA AACTATTTGT

10451 TCCTGCAGAG CATAGTGGTG CAGGAGCGCA GCTTGAGCCT GGCTGACAAG  
AGGACGTCTC GTATCACCAC GTCCTCGCGT CGAACTCGGA CCGACTGTTC

10501 GTGGCCGCCA TCAACTATTC CATGCTTAGC CTGGGCAAGT TTTACGCCCG  
CACCGGCGGT AGTTGATAAG GTACGAATCG GACCCGTTCA AAATGCGGGC

10551 CAAGATATAC CATACCCCTT ACGTTCCCAT AGACAAGGAG GTAAAGATCG  
GTTCTATATG GTATGGGGAA TGCAAGGGTA TCTGTTCTC CATTCTAGC

10601 AGGGGTTCTA CATGCGCATG GCGCTGAAGG TGCTTACCTT GAGCGACGAC  
TCCCCAAGAT GTACGCGTAC CGCGACTTCC ACGAATGGAA CTCGCTGCTG

10651 CTGGGCGTTT ATCGCAACGA GCGCATCCAC AAGGCCGTGA GCGTGAGCCG  
GACCCGCAA TAGCGTTGCT CGCGTAGGTG TTCCGGCACT CGCACTCGGC

10701 GCGGCGCGAG CTCAGCGACC GCGAGCTGAT GCACAGCCTG CAAAGGGCCC  
CGCCGCGCTC GAGTCGCTGG CGCTCGACTA CGTGTCGGAC GTTTCCCGGG

10751 TGGCTGGCAC GGGCAGCGGC GATAGAGAGG CCGAGTCCTA CTTTGACGCG  
ACCGACCGTG CCCGTCGCCG CTATCTCTCC GGCTCAGGAT GAAACTGCGC

10801 GGCGCTGACC TGCCTGGGGC CCAAGCCGA CGCGCCCTGG AGGCAGCTGG  
CCGCGACTGG ACGCGACCCG GGGTTCGGCT GCGCGGGACC TCCGTCGACC

10851 GGCCGGACCT GGGCTGGCGG TGGCACCCGC GCGCGCTGGC AACGTCGGCG  
CCGGCCTGGA CCCGACCGCC ACCGTGGGCG CGCGCGACCG TTGCAGCCGC

10901 GCGTGGAGGA ATATGACGAG GACGATGAGT ACGAGCCAGA GGACGGCGAG  
CGCACCTCCT TATACTGCTC CTGCTACTCA TGCTCGGTCT CCTGCCGCTC

10951 TACTAAGCGG TGATGTTTCT GATCAGATGA TGCAAGACGC AACGGACCCG  
ATGATTCGCC ACTACAAAGA CTAGTCTACT ACGTTCTGCG TTGCCTGGGC

11001 GCGGTGCGGG CGGCGCTGCA GAGCCAGCCG TCCGGCCTTA ACTCCACGGA  
CGCCACGCCC GCCGCGACGT CTCGGTCGGC AGGCCGGAAT TGAGGTGCCT

11051 CGACTGGCGC CAGGTCATGG ACCGCATCAT GTCGCTGACT GCGCGCAATC  
GCTGACCGCG GTCCAGTACC TGGCGTAGTA CAGCGACTGA CGCGCGTTAG

11101 CTGACGCGTT CCGGCAGCAG CCGCAGGCCA ACCGGCTCTC CGCAATTCTG  
GACTGCGCAA GGCCGTCGTC GGCGTCCGGT TGGCCGAGAG GCGTTAAGAC

11151 GAAGCGGTGG TCCCGGCGCG CGCAAACCCC ACGCACGAGA AGGTGCTGGC  
CTTCGCCACC AGGGCCGCGC GCGTTTGGGG TGCGTGCTCT TCCACGACCG

11201 GATCGTAAAC GCGCTGGCCG AAAACAGGGC CATCCGGCCC GACGAGGCCG  
CTAGCATTTG CGCGACCGGC TTTTGTCCCG GTAGGCCGGG CTGCTCCGGC

11251 GCCTGGTCTA CGACGCGCTG CTTGAGCGCG TGGCTCGTTA CAACAGCGGC  
CGGACCAGAT GCTGCGCGAC GAAGTCGCGC ACCGAGCAAT GTTGTCGCCG

Figure 27L

11351 GGCGCAGCGT GAGCGCGCGC AGCAGCAGGG CAACCTGGGC TCCATGGTTG  
CCGCGTCGCA CTCGCGCGCG TCGTCGTCCC GTTGGACCCG AGGTACCAAC

11401 CACTAAACGC CTTCTGAGT ACACAGCCCG CCAACGTGCC GCGGGGACAG  
GTGATTTGCG GAAGGACTCA TGTGTCGGGC GGTTCACCGG CGCCCCTGTC

11451 GAGGACTACA CCAACTTTGT GAGCGCACTG CGGCTAATGG TGACTGAGAC  
CTCCTGATGT GGTGAAACA CTCGCGTGAC GCCGATTACC ACTGACTCTG

11501 ACCGCAAAGT GAGGTGTACC AGTCTGGGCC AGACTATTTT TTCCAGACCA  
TGGCGTTTCA CTCCACATGG TCAGACCCCG TCTGATAAAA AAGGTCTGGT

11551 GTAGACAAGG CCTGCAGACC GTAAACCTGA GCCAGGCTTT CAAAACTTG  
CATCTGTTCC GGACGTCTGG CATTTGGACT CGGTCCGAAA GTTTTTGAAC

11601 CAGGGGCTGT GGGGGGTGCG GGCTCCCACA GGCGACCGCG CGACCGTGTC  
GTCCCCGACA CCCCCACGC CCGAGGGTGT CCGCTGGCGC GCTGGCACAG

11651 TAGCTTGCTG ACGCCCAACT CGCGCCTGTT GCTGCTGCTA ATAGCGCCCT  
ATCGAACGAC TGCGGGTTGA GCGCGGACAA CGACGACGAT TATCGCGGGA

11701 TCACGGACAG TGGCAGCGTG TCCCGGGACA CATACTAGG TCACTTGCTG  
AGTGCCCTGTC ACCGTCGCAC AGGGCCCTGT GTATGGATCC AGTGAACGAC

11751 AACTGTACC GCGAGGCCAT AGGTCAGGCG CATGTGGACG AGCATACTTT  
TGTGACATGG CGCTCCGGTA TCCAGTCCGC GTACACCTGC TCGTATGAAA

11801 CCAGGAGATT ACAAGTGTCA GCCGCGCGCT GGGGCAGGAG GACACGGGCA  
GGTCCTCTAA TGTTACAGT CCGCGCGCGA CCCCCTCCTC CTGTGCCCGT

11851 GCCTGGAGGC AACCCTAAAC TACCTGCTGA CCAACCGGCG GCAGAAGATC  
CGGACCTCCG TTGGGATTTG ATGGACGACT GGTGGCCCGC CGTCTTCTAG

11901 CCCTCGTTGC ACAGTTTAAA CAGCGAGGAG GAGCGCATTT TCGCTACGT  
GGGAGCAACG TGTCAAATTT GTCGCTCCTC CTCGCGTAAA ACGCGATGCA

11951 GCAGCAGAGC GTGAGCCTTA ACCTGATGCG CGACGGGGTA ACGCCCAGCG  
CGTCGTCTCG CACTCGGAAT TGGACTACGC GCTGCCCCAT TCGGGTCTGC

12001 TGGCGCTGGA CATGACCGCG CGCAACATGG AACC GGCGCAT GTATGCCTCA  
ACCGCGACCT GTACTGGCGC GCGTTGTACC TTGGCCCGTA CATA CGGAGT

12051 AACC GGCGCAT TTATCAACCG CCTAATGGAC TACTTGATC GCGCGGCCGC  
TTGGCCGGCA AATAGTTGGC GGATTACCTG ATGAACGTAG CGCGCCGGCG

12101 CGTGAACCCC GAGTATTTCA CCAATGCCAT CTTGAACCCG CACTGGCTAC  
GCACTTGGGG CTCATAAAGT GGTTACGGTA GAACTTGGGC GTGACCGATG

12151 CGCCCCCTGG TTTCTACACC GGGGGATTCT AGGTGCCCCG GGGTAACGAT  
GCGGGGGACC AAAGATGTGG CCCCCTAAGC TCCACGGGCT CCCATTGCTA

12201 GGATTCCTCT GGGACGACAT AGACGACAGC GTGTTTTCCC CGCAACCGCA  
CCTAAGGAGA CCCTGCTGTA TCTGCTGTCT CACAAAAGGG GCGTTGGCGT

Figure 27 M

12301 AGGAAAGCTT CCGCAGGCCA AGCAGCTTGT CCGATCTAGG CGCTGCGGCC  
TCCTTTCGAA GGCGTCCGGT TCGTCGAACA GGCTAGATCC GCGACGCCGG

12351 CCGCGGTCAG ATGCTAGTAG CCCATTTCCA AGCTTGATAG GGTCTCTTAC  
GGCGCCAGTC TACGATCATC GGGTAAAGGT TCGAACTATC CCAGAGAATG

12401 CAGCACTCGC ACCACCCGCC CGCGCCTGCT GGGCGAGGAG GAGTACCTAA  
GTCGTGAGCG TGGTGGGCGG GCGCGGACGA CCCGCTCCTC CTCATGGATT

12451 ACAACTCGCT GCTGCAGCCG CAGCGCGAAA AAAACCTGCC TCCGGCATT  
TGTTGAGCGA CGACGTCGGC GTCGCGCTTT TTTTGGACGG AGGCCGTAAA

12501 CCCAACAACG GGATAGAGAG CCTAGTGGAC AAGATGAGTA GATGGAAGAC  
GGGTTGTTGC CCTATCTCTC GGATCACCTG TTCTACTCAT CTACCTTCTG

12551 GTACGCGCAG GAGCACAGGG ACGTGCCAGG CCCGCGCCCG CCCACCCGTC  
CATGCGCGTC CTCGTGTCCC TGCACGGTCC GGGCGCGGGC GGGTGGGCAG

12601 GTCAAAGGCA CGACCGTCAG CGGGGTCTGG TGTGGGAGGA CGATGACTCG  
CAGTTTCCGT GCTGGCAGTC GCCCCAGACC ACACCCTCCT GCTACTGAGC

12651 GCAGACGACA GCAGCGTCCT GGATTTGGGA GGGAGTGGCA ACCCGTTTGC  
CGTCTGCTGT CGTCGCAGGA CCTAAACCCT CCCTCACCGT TGGGCAAACG

12701 GCACCTTCGC CCCAGGCTGG GGAGAATGTT TTAAAAAAA AAAAAGCATG  
CGTGGAAGCG GGGTCCGACC CCTCTTACAA AATTTTTTTT TTTTTCGTAC

12751 ATGCAAAATA AAAAATCAC CAAGGCCATG GCACCGAGCG TTGGTTTTCT  
TACGTTTTAT TTTTGTAGTG GTTCCGGTAC CGTGGCTCGC AACC AAAAGA

12801 TGTATTCCCC TTAGTATGCG GCGCGCGGCG ATGTATGAGG AAGGTCCTCC  
ACATAAGGGG AATCATACGC CGCGCGCCGC TACATACTCC TTCCAGGAGG

12851 TCCCTCCTAC GAGAGTGTGG TGAGCGCGGC GCCAGTGGCG GCGGCGCTGG  
AGGGAGGATG CTCTCACACC ACTCGCGCCG CGGTCACCGC CGCCGCGACC

12901 GTTCTCCCTT CGATGCTCCC CTGGACCCGC CGTTTGTGCC TCCGCGGTAC  
CAAGAGGGAA GCTACGAGGG GACCTGGGCG GCAAACACGG AGGCGCCATG

12951 CTGCGGCCTA CCGGGGGGAG AAACAGCATC CGTTACTCTG AGTTGGCACC  
GACGCCGGAT GGCCCCCTC TTTGTCTAG GCAATGAGAC TCAACCGTGG

13001 CCTATTGAC ACCACCCGTG TGTACCTGGT GGACAACAAG TCAACGGATG  
GGATAAGCTG TGGTGGGCAC ACATGGACCA CCTGTTGTTC AGTTGCCTAC

13051 TGGCATCCCT GAACTACCAG AACGACCACA GCAACTTTCT GACCACGGTC  
ACCGTAGGGA CTTGATGGTC TTGCTGGTGT CGTTGAAAGA CTGGTGCCAG

13101 ATTCAAACA ATGACTACAG CCCGGGGGAG GCAAGCACAC AGACCATCAA  
TAAGTTTTGT TACTGATGTC GGGCCCCCTC CGTTCGTGTG TCTGGTAGTT

13151 TCTTGACGAC CGGTCGCACT GGGGCGGCGA CCTGAAAACC ATCCTGCATA  
AGAACTGCTG GCCAGCGTGA CCCC GCCGCT GGACTTTTGG TAGGACGTAT

Figure 27N

13251 CGGGTGATGG TGTCGCGCTT GCCTACTAAG GACAATCAGG TGGAGCTGAA  
GCCCCACTACC ACAGCGCGAA CGGATGATTC CTGTTAGTCC ACCTCGACTT

13301 ATACGAGTGG GTGGAGTTCA CGCTGCCCCG GGGCAACTAC TCCGAGACCA  
TATGCTCACC CACCTCAAGT GCGACGGGCT CCCGTTGATG AGGCTCTGGT

13351 TGACCATAGA CCTTATGAAC AACGCGATCG TGGAGCACTA CTTGAAAGTG  
ACTGGTATCT GGAATACTTG TTGCGCTAGC ACCTCGTGAT GAACTTTCAC

13401 GGCAGACAGA ACGGGGTTCCT GGAAAGCGAC ATCGGGGTAA AGTTTGACAC  
CCGTCTGTCT TGCCCCAAGA CCTTTCGCTG TAGCCCCATT TCAAACGTGT

13451 CCGCAACTTC AGACTGGGGT TTGACCCCGT CACTGGTCTT GTCATGCCTG  
GGCGTTGAAG TCTGACCCCA AACTGGGGCA GTGACCAGAA CAGTACGGAC

13501 GGGTATATAC AAACGAAGCC TTCCATCCAG ACATCATTTT GCTGCCAGGA  
CCCATATATG TTTGCTTCGG AAGGTAGGTC TGTAATAAAA CGACGGTCCT

13551 TGCGGGGTGG ACTTCACCCA CAGCCGCCTG AGCAACTTGT TGGGCATCCG  
ACGCCCCACC TGAAGTGGGT GTCGGCGGAC TCGTTGAACA ACCCGTAGGC

13601 CAAGCGGCAA CCCTTCCAGG AGGGCTTTAG GATCACCTAC GATGATCTGG  
GTTCGCCGTT GGGAAGGTCC TCCCGAAATC CTAGTGGATG CTACTAGACC

13651 AGGGTGGTAA CATTCCCGCA CTGTTGGATG TGGACGCCTA CCAGGCGAGC  
TCCCACCATT GTAAGGGCGT GACAACCTAC ACCTGCGGAT GGTCCGCTCG

13701 TTGAAAGATG ACACCGAACA GGGCGGGGGT GGCGCAGGCG GCAGCAACAG  
AACTTTCTAC TGTGGCTTGT CCCGCCCCCA CCGCGTCCGC CGTCGTTGTC

13751 CAGTGGCAGC GGCGCGGAAG AGAACTCCAA CGCGGCAGCC GCGGCAATGC  
GTCACCGTCG CCGCGCCTTC TCTTGAGGTT GCGCCGTCGG CGCCGTTACG

13801 AGCCGGTGGA GGACATGAAC GATCATGCCA TTCGCGGCGA CACCTTTGCC  
TCGGCCACCT CCTGTACTTG CTAGTACGGT AAGCGCCGCT GTGGAAACGG

13851 ACACGGGCTG AGGAGAAGCG CGCTGAGGCC GAAGCAGCGG CCGAAGCTGC  
TGTGCCCGAC TCCTCTTCGC GCGACTCCGG CTTCGTGCGC GGCTTCGACG

13901 CGCCCCCGCT GCGCAACCCG AGGTCGAGAA GCCTCAGAAG AAACCGGTGA  
GCGGGGGCGA CGCGTTGGGC TCCAGCTCTT CGGAGTCTTC TTTGGCCACT

13951 TCAAACCCCT GACAGAGGAC AGCAAGAAAC GCAGTTACAA CCTAATAAGC  
AGTTTGGGGA CTGTCTCCTG TCGTTCTTTG CGTCAATGTT GGATTATTTCG

14001 AATGACAGCA CCTTCACCCA GTACCGCAGC TGGTACCTTG CATACAACTA  
TTACTGTCGT GGAAGTGGGT CATGGCGTCG ACCATGGAAC GTATGTTGAT

14051 CGGCGACCCT CAGACCGGAA TCCGCTCATG GACCCTGCTT TGCACCTCTG  
GCCGCTGGGA GTCTGGCCTT AGGCGAGTAC CTGGGACGAA ACGTGAGGAC

14101 ACGTAACCTG CGGCTCGGAG CAGGTCTACT GGTCGTTGCC AGACATGATG  
TGCATTGGAC GCCGAGCCTC GTCCAGATGA CCAGCAACGG TCTGTACTAC

Figure 270

14201 GGTGGGCGCC GAGCTGTTGC CCGTGCCTTC CAAGAGCTTC TACAACGACC  
CCACCCGCGG CTCGACAACG GGCACGTGAG GTTCTCGAAG ATGTTGCTGG

14251 AGGCCGTCTA CTCCCACTC ATCCGCCAGT TTACCTCTCT GACCCACGTG  
TCCGGCAGAT GAGGGTTGAG TAGGCGGTCA AATGGAGAGA CTGGGTGCAC

14301 TTCAATCGCT TTCCCGAGAA CCAGATTTTG GCGCGCCCGC CAGCCCCCAC  
AAGTTAGCGA AAGGGCTCTT GGTCTAAAAC CGCGCGGGCG GTCGGGGGTG

14351 CATCACCACC GTCAGTGAAA ACGTTCCTGC TCTCACAGAT CACGGGACGC  
GTAGTGGTGG CAGTCACTTT TGCAAGGACG AGAGTGTCTA GTGCCCTGCG

14401 TACCGCTGCG CAACAGCATC GGAGGAGTCC AGCGAGTGAC CATTACTGAC  
ATGGCGACGC GTTGTCTGAG CCTCCTCAGG TCGCTCACTG GTAATGACTG

14451 GCCAGACGCC GCACCTGCCC CTACGTTTAC AAGGCCCTGG GCATAGTCTC  
CGGTCTGCGG CGTGGACGGG GATGCAAATG TTCCGGGACC CGTATCAGAG

14501 GCCGCGCGTC CTATCGAGCC GCACTTTTTG AGCAAGCATG TCCATCCTTA  
CGGCGCGCAG GATAGCTCGG CGTGAAAAAC TCGTTCGTAC AGGTAGGAAT

14551 TATCGCCCAG CAATAACACA GGCTGGGGCC TGCCTTCCC AAGCAAGATG  
ATAGCGGGTC GTTATTGTGT CCGACCCCGG ACGCGAAGGG TTCGTTCTAC

14601 TTTGGCGGGG CCAAGAAGCG CTCCGACCAA CACCCAGTGC GCGTGCGCGG  
AAACCGCCCC GGTCTTCTGC GAGGCTGGTT GTGGGTCACG CGCACGCGCC

14651 GCACTACCGC GCGCCCTGGG GCGCGCACAA ACGCGGCCGC ACTGGGCGCA  
CGTGATGGCG CGCGGGACCC CGCGCGTGTT TGCGCCGGCG TGACCCGCGT

14701 CCACCGTCGA TGACGCCATC GACGCGGTGG TGGAGGAGGC GCGCAACTAC  
GGTGGCAGCT ACTGCGGTAG CTGCGCCACC ACCTCCTCCG CGCGTTGATG

14751 ACGCCACGC CGCCACCAGT GTCCACAGTG GACGCGGCCA TTCAGACCGT  
TGCGGGTGCG GCGGTGGTCA CAGGTGTCAC CTGCGCCGGT AAGTCTGGCA

14801 GGTGCGCGGA GCCCGGCGCT ATGCTAAAAT GAAGAGACGG CGGAGGCGCG  
CCACGCGCCT CGGGCCGCGA TACGATTTTA CTTCTCTGCC GCCTCCGCGC

14851 TAGCACGTCG CCACCGCCGC CGACCCGGCA CTGCCGCCCA ACGCGCGGCG  
ATCGTGACG GGTGGCGGCG GCTGGGCGGT GACGGCGGGT TGCGCGCCGC

14901 GCGGCCCTGC TTAACCGCGC ACGTCGCACC GGCCGACGGG CGGCCATGCG  
CGCCGGGACG AATTGGCGCG TGCAGCGTGG CCGGCTGCCC GCCGGTACGC

14951 GGCCGCTCGA AGGCTGGCCG CGGGTATTGT CACTGTGCCC CCCAGGTCCA  
CCGGCGAGCT TCCGACCGGC GCCCATAACA GTGACACGGG GGGTCCAGGT

15001 GGCGACGAGC GGCCGCCGCA GCAGCCGCGG CCATTAGTGC TATGACTCAG  
CCGCTGCTCG CCGGCGGCGT CGTCGGCGCC GGTAATCACG ATACTGAGTC

15051 GGTCGCAGGG GCAACGTGTA TTGGGTGCGC GACTCGGTTA GCGGCCTGCG  
CCAGCGTCCC CGTTGCACAT AACCACGCG CTGAGCCAAT CGCCGGACGC

Figure 27P

15151 ACTTAGACTC GTACTGTTGT ATGTATCCAG CGGCGGCGGC GCGCAACGAA  
TGAATCTGAG CATGACAACA TACATAGGTC GCCGCCGCCG CGCGTTGCTT

15201 GCTATGTCCA AGCGCAAAAT CAAAGAAGAG ATGCTCCAGG TCATCGCGCC  
CGATACAGGT TCGCGTTTTA GTTTCTTCTC TACGAGGTCC AGTAGCGCGG

15251 GGAGATCTAT GGCCCCCGA AGAAGGAAGA GCAGGATTAC AAGCCCCGAA  
CCTCTAGATA CCGGGGGGCT TCTTCCTTCT CGTCCTAATG TTCGGGGCTT

15301 AGCTAAAGCG GGTCAAAAAG AAAAAGAAAG ATGATGATGA TGAAGTTGAC  
TCGATTTTCG CCAGTTTTTC TTTTCTTTC TACTACTACT ACTTGAAGTG

15351 GACGAGGTGG AACTGCTGCA CGCTACCGCG CCCAGGCGAC GGGTACAGTG  
CTGCTCCACC TTGACGACGT GCGATGGCGC GGGTCCGCTG CCCATGTCAC

15401 GAAAGGTCGA CGCGTAAAC GTGTTTTGCG ACCCGGCACC ACCGTAGTCT  
CTTTCCAGCT GCGCATTTTG CACAAAACGC TGGGCCGTGG TGGCATCAGA

15451 TTACGCCCGG TGAGCGCTCC ACCCGCACCT ACAAGCGCGT GTATGATGAG  
AATGCGGGCC ACTCGCGAGG TGGGCGTGGA TGTTCGCGCA CATACTACTC

15501 GTGTACGGCG ACGAGGACCT GCTTGAGCAG GCCAACGAGC GCCTCGGGGA  
CACATGCCGC TGCTCCTGGA CGAACTCGTC CGGTTGCTCG CGGAGCCCCT

15551 GTTTGCCTAC GGAAAGCGGC ATAAGGACAT GCTGGCGTTG CCGCTGGACG  
CAAACGGATG CCTTTCGCCG TATTCCTGTA CGACCGCAAC GGCGACCTGC

15601 AGGGCAACCC AACACCTAGC CTAAAGCCCC TAACACTGCA GCAGGTGCTG  
TCCCGTTGGG TTGTGGATCG GATTTCCGGC ATTGTGACGT CGTCCACGAC

15651 CCCGCGCTTG CACCGTCCGA AGAAAAGCGC GGCCTAAAGC GCGAGTCTGG  
GGGCGCGAAC GTGGCAGGCT TCTTTTCGCG CCGGATTTCG CGCTCAGACC

15701 TGACTTGGCA CCCACCGTGC AGCTGATGGT ACCCAAGCGC CAGCGACTGG  
ACTGAACCGT GGGTGGCAGC TCGACTACCA TGGGTTTCGCG GTCGCTGACC

15751 AAGATGTCTT GGAAAAAATG ACCGTGGAAC CTGGGCTGGA GCCCGAGGTC  
TTCTACAGAA CCTTTTTTAC TGGCACCTTG GACCCGACCT CGGGCTCCAG

15801 CGCGTGCGGC CAATCAAGCA GGTGGCGCCG GGAAGTGGCG TGCAGACCGT  
GCGCACGCCG GTTAGTTCGT CCACCGCGGC CCTGACCCGC ACGTCTGGCA

15851 GGACGTTTCAG ATACCCACTA CCAGTAGCAC CAGTATTGCC ACCGCCACAG  
CCTGCAAGTC TATGGGTGAT GGTTCATCGTG GTCATAACGG TGGCGGTGTC

15901 AGGGCATGGA GACACAAACG TCCCCGGTTG CCTCAGCGGT GGCGGATGCC  
TCCCGTACCT CTGTGTTTGC AGGGGCCAAC GGAGTCGCCA CCGCCTACGG

15951 GCGGTGCAGG CGGTCGCTGC GGCCGCGTCC AAGACCTCTA CGGAGGTGCA  
CGCCACGTCC GCCAGCGACG CCGGCGCAGG TTCTGGAGAT GCCTCCACGT

16001 AACGGACCCG TGGATGTTTC GCGTTTCAGC CCCCCGGCGC CCGCGCCGTT  
TTGCCTGGGC ACCTACAAAG CGCAAAGTCG GGGGGCCGCG GGC GCGGCAA

Figure 27Q

16051	CGAGGAAGTA	CGGCGCCGCC	AGCGCGCTAC	TGCCCCGAATA	TGCCCTACAT
	GCTCCTTCAT	GCCGCGGCGG	TCGCGCGATG	ACGGGCTTAT	ACGGGATGTA
16101	CCTTCCATTG	CGCCTACCCC	CGGCTATCGT	GGCTACACCT	ACCGCCCCAG
	GGAAGGTAAC	GCGGATGGGG	GCCGATAGCA	CCGATGTGGA	TGGCGGGGTC
16151	AAGACGAGCA	ACTACCCGAC	GCCGAACCAC	CACTGGAACC	CGCCGCCGCC
	TTCTGCTCGT	TGATGGGCTG	CGGCTTGGTG	GTGACCTTGG	GCGGCGGCGG
16201	GTCGCCGTCG	CCAGCCCGTG	CTGGCCCCGA	TTTCCGTGCG	CAGGGTGGCT
	CAGCGGCAGC	GGTCGGGCAC	GACCGGGGCT	AAAGGCACGC	GTCCCACCGA
16251	CGCGAAGGAG	GCAGGACCCT	GGTGCTGCCA	ACAGCGCGCT	ACCACCCCAG
	GCGCTTCCTC	CGTCCTGGGA	CCACGACGGT	TGTCGCGCGA	TGGTGGGGTC
16301	CATCGTTTAA	AAGCCGGTCT	TTGTGGTTCT	TGCAGATATG	GCCCTCACCT
	GTAGCAAATT	TTCGGCCAGA	AACACCAAGA	ACGTCTATAC	CGGGAGTGGA
16351	GCCGCCTCCG	TTTCCCGGTG	CCGGGATTCC	GAGGAAGAAT	GCACCGTAGG
	CGGCGGAGGC	AAAGGGCCAC	GGCCCTAAGG	CTCCTTCTTA	CGTGGCATCC
16401	AGGGGCATGG	CCGGCCACGG	CCTGACGGGC	GGCATGCGTC	GTGCGCACCA
	TCCCCGTACC	GGCCGGTGCC	GGACTGCCCC	CCGTACGCAG	CACGCGTGGT
16451	CCGGCGGCGG	CGCGCGTCGC	ACCGTCGCAT	GCGCGGCGGT	ATCCTGCCCC
	GGCCGCCGCC	GCGCGCAGCG	TGGCAGCGTA	CGCGCCGCCA	TAGGACGGGG
16501	TCCTTATTCC	ACTGATCGCC	GCGGCGATTG	GCGCCGTGCC	CGGAATTGCA
	AGGAATAAGG	TGACTAGCGG	CGCCGCTAAC	CGCGGCACGG	GCCTTAACGT
16551	TCCGTGGCCT	TGCAGGCGCA	GAGACACTGA	TTAAAAACAA	GTTGCATGTG
	AGGCACCGGA	ACGTCCGCGT	CTCTGTGACT	AATTTTGTGTT	CAACGTACAC
16601	GAAAAATCAA	AATAAAAAGT	CTGGACTCTC	ACGCTCGCTT	GGTCCTGTAA
	CTTTTTAGTT	TTATTTTTC A	GACCTGAGAG	TGCGAGCGAA	CCAGGACATT
16651	CTATTTTGTA	GAATGGAAGA	CATCAACTTT	GCGTCTCTGG	CCCCGCGACA
	GATAAAACAT	CTTACCTTCT	GTAGTTGAAA	CGCAGAGACC	GGGGCGCTGT
16701	CGGCTCGCGC	CCGTTTCATGG	GAAACTGGCA	AGATATCGGC	ACCAGCAATA
	GCCGAGCGCG	GGCAAGTACC	CTTTGACCGT	TCTATAGCCG	TGGTCGTTAT
16751	TGAGCGGTGG	CGCCTTCAGC	TGGGGCTCGC	TGTGGAGCGG	CATTAAAAAT
	ACTCGCCACC	GCGGAAGTCG	ACCCCGAGCG	ACACCTCGCC	GTAATTTTTA
16801	TTCGGTTCCA	CCGTTAAGAA	CTATGGCAGC	AAGGCCTGGA	ACAGCAGCAC
	AAGCCAAGGT	GGCAATTCTT	GATACCGTCG	TTCCGGACCT	TGTCGTCGTG
16851	AGGCCAGATG	CTGAGGGATA	AGTTGAAAGA	GCAAAATTTC	CAACAAAAGG
	TCCGGTCTAC	GACTCCCTAT	TCAACTTTCT	CGTTTTAAAG	GTTGTTTTCC
16901	TGGTAGATGG	CCTGGCCTCT	GGCATTAGCG	GGGTGGTGGA	CCTGGCCAAC
	ACCATCTACC	GGACCGGAGA	CCGTAATCGC	CCCACCACCT	GGACCGGTTG
16951	CAGGCAGTGC	AAAATAAGAT	TAACAGTAAG	CTTGATCCCC	GCCCTCCCGT
	GTCCGTCACG	TTTTATTCTA	ATTGTCATTC	GAAGTAGGGG	CGGGAGGGCA

Figure 27R

17051 AAAAGCGTCC GCGCCCCGAC AGGGAAGAAA CTCTGGTGAC GCAAATAGAC  
TTTTCGCAGG CGCGGGGCTG TCCCTTCTTT GAGACCACTG CGTTTATCTG

17101 GAGCCTCCCT CGTACGAGGA GGCATAAAG CAAGGCCTGC CCACCACCCG  
CTCGGAGGGA GCATGCTCCT CCGTGATTTC GTTCCGGACG GGTGGTGGGC

17151 TCCCATCGCG CCCATGGCTA CCGGAGTGCT GGGCCAGCAC ACACCCGTAA  
AGGGTAGCGC GGGTACCGAT GGCCTCACGA CCCGGTCGTG TGTGGGCATT

17201 CGCTGGACCT GCCTCCCCCC GCCGACACCC AGCAGAAACC TGTGCTGCCA  
GCGACCTGGA CGGAGGGGGG CGGCTGTGGG TCGTCTTTGG ACACGACGGT

17251 GGCCCGACCG CCGTTGTTGT AACCCGTCCT AGCCGCGCGT CCCTGCGCCG  
CCGGGCTGGC GGCAACAACA TTGGGCAGGA TCGGCGCGCA GGGACGCGGC

17301 CGCCGCCAGC GGTCCGCGAT CGTTGCGGCC CGTAGCCAGT GGCAACTGGC  
GCGGCGGTCTG CCAGGCGCTA GCAACGCCGG GCATCGGTCA CCGTTGACCG

17351 AAAGCACACT GAACAGCATC GTGGGTCTGG GGGTGCAATC CCTGAAGCGC  
TTTCGTGTGA CTTGTCTAGT CACCCAGACC CCCACGTTAG GGACTTCGCG

17401 CGACGATGCT TCTGATAGCT AACGTGTCGT ATGTGTGTCA TGTATGCGTC  
GCTGCTACGA AGACTATCGA TTGCACAGCA TACACACAGT ACATACGCAG

17451 CATGTCGCCG CCAGAGGAGC TGCTGAGCCG CCGCGCGCCC GCTTTCCAAG  
GTACAGCGGC GGTCTCCTCG ACGACTCGGC GCGCGCGGG CGAAAGGTTC

17501 ATGGCTACCC CTTGATGAT GCCGCAGTGG TCTTACATGC ACATCTCGGG  
TACCGATGGG GAAGCTACTA CGGCGTCACC AGAATGTACG TGTAAGAGCCC

17551 CCAGGACGCC TCGGAGTACC TGAGCCCCGG GCTGGTGCA G TTTGCCCGCG  
GGTCTGCGG AGCCTCATGG ACTCGGGGCC CGACCACGTC AAACGGGCGC

17601 CCACCGAGAC GTA CTT CAGC CTGAATAACA AGTTTAGAAA CCCCACGGTG  
GGTGGCTCTG CATGAAGTCG GACTTATTGT TCAAATCTTT GGGGTGCCAC

17651 GCGCCTACGC ACGACGTGAC CACAGACCGG TCCCAGCGTT TGACGCTGCG  
CGCGGATGCG TGCTGCACTG GTGTCTGGCC AGGGTCGCAA ACTGCGACGC

17701 GTTCATCCCT GTGGACCGTG AGGATACTGC GTACTCGTAC AAGGCGCGGT  
CAAGTAGGGA CACCTGGCAC TCCTATGACG CATGAGCATG TTCCGCGCCA

17751 TCACCCTAGC TGTGGGTGAT AACCGTGTGC TGGACATGGC TTCCACGTAC  
AGTGGGATCG ACACCCACTA TTGGCACACG ACCTGTACCG AAGGTGCATG

17801 TTTGACATCC GCGGCGTGCT GGACAGGGGC CCTACTTTTA AGCCCTACTC  
AAACTGTAGG CGCCGCACGA CCTGTCCCCG GGATGAAAAT TCGGGATGAG

17851 TGGCACTGCC TACAACGCC TGGCTCCCAA GGGTGCCCCA AATCCTTGCG  
ACCGTGACGG ATGTTGCGGG ACCGAGGGTT CCCACGGGGT TTAGGAACGC

17901 AATGGGATGA AGCTGCTACT GCTCTTGAAA TAAACCTAGA AGAAGAGGAC  
TTACCCTACT TCGACGATGA CGAGAACTTT ATTTGGATCT TCTTCTCCTG

Figure 275

17951 GATGACAACG AAGACGAAGT AGACGAGCAA GCTGAGCAGC AAAAAACTCA  
CTACTGTTGC TTCTGCTTCA TCTGCTCGTT CGACTCGTCG TTTTTTGAGT

18001 CGTATTTGGG CAGGCGCCTT ATTCTGGTAT AAATATTACA AAGGAGGGTA  
GCATAAACCC GTCCGCGGAA TAAGACCATA TTTATAATGT TTCCTCCCAT

18051 TTCAAATAGG TGTCGAAGGT CAAACACCTA AATATGCCGA TAAAACATTT  
AAGTTTATCC ACAGCTTCCA GTTTGTGGAT TTATACGGCT ATTTTGTAAT

18101 CAACCTGAAC CTCAAATAGG AGAATCTCAG TGGTACGAAA CAGAAATTAA  
GTTGGACTTG GAGTTTATCC TCTTAGAGTC ACCATGCTTT GTCTTTAATT

18151 TCATGCAGCT GGGAGAGTCC TAAAAAAGAC TACCCCAATG AAACCATGTT  
AGTACGTCGA CCCTCTCAGG ATTTTTTCTG ATGGGGTTAC TTTGGTACAA

18201 ACGGTTTATA TGCAAAACCC ACAAATGAAA ATGGAGGGCA AGGCATTCTT  
TGCCAAGTAT ACGTTTTGGG TGTTTACTTT TACCTCCCGT TCCGTAAGAA

18251 GTAAAGCAAC AAAATGGAAA GCTAGAAAGT CAAGTGGAAA TGCAATTTTT  
CATTTCTGTTG TTTTACCTTT CGATCTTTCA GTTCACCTTT ACGTTAAAAA

18301 CTCAACTACT GAGGCAGCCG CAGGCAATGG TGATAACTTG ACTCCTAAAG  
GAGTTGATGA CTCCGTCGGC GTCCGTTACC ACTATTGAAC TGAGGATTTC

18351 TGGTATTGTA CAGTGAAGAT GTAGATATAG AAACCCCAAG CACTCATATT  
ACCATAACAT GTCACCTTCTA CATCTATATC TTTGGGGTCT GTGAGTATAA

18401 TCTTACATGC CCACTATTAA GGAAGGTAAC TCACGAGAAC TAATGGGCCA  
AGAATGTACG GGTGATAATT CCTTCCATTG AGTGCTCTTG ATTACCCGGT

18451 ACAATCTATG CCCAACAGGC CTAATTACAT TGCTTTTAGG GACAATTTTA  
TGTTAGATAC GGGTTGTCCG GATTAATGTA ACGAAAATCC CTGTTAAAT

18501 TTGGTCTAAT GTATTACAAC AGCACGGGTA ATATGGGTGT TCTGGCGGGC  
AACCAGATTA CATAATGTTG TCGTGCCCAT TATACCCACA AGACCGCCCG

18551 CAAGCATCGC AGTTGAATGC TGTGTAGAT TTGCAAGACA GAAACACAGA  
GTTTCGTAGC TCAACTTACG ACAACATCTA AACGTTCTGT CTTTGTGTCT

18601 GCTTTCATAC CAGCTTTTGC TTGATTCCAT TGGTGATAGA ACCAGGTACT  
CGAAAGTATG GTCGAAAACG AACTAAGGTA ACCACTATCT TGGTCCATGA

18651 TTTCTATGTG GAATCAGGCT GTTGACAGCT ATGATCCAGA TGTTAGAATT  
AAAGATACAC CTTAGTCCGA CAACTGTCGA TACTAGGTCT ACAATCTTAA

18701 ATTGAAAATC ATGGAAGTGA AGATGAACTT CCAAATTACT GCTTCCACT  
TAACCTTTAG TACCTTGACT TCTACTTGAA GGTTTAATGA CGAAAGGTGA

18751 GGGAGGTGTG ATTAATACAG AGACTCTTAC CAAGGTAAAA CCTAAAACAG  
CCCTCCACAC TAATTATGTC TCTGAGAATG GTTCCATTTT GGATTTTGTC

18801 GTCAGGAAAA TGGATGGGAA AAAGATGCTA CAGAATTTTC AGATAAAAAT  
CAGTCCTTTT ACCTACCCTT TTTCTACGAT GTCTTAAAAG TCTATTTTAA

18851 GAAATAAGAG TTGGAAATAA TTTTGCCATG GAAATCAATC TAAATGCCAA  
CTTTATTCTC AACCTTTATT AAAACGGTAC CTTTAGTTAG ATTTACGGTT

Figure 27T

18951 AGCTAAAGTA CAGTCCTTCC AACGTAAAAA TTTCTGATAA CCCAAACACC  
TCGATTTTCAT GTCAGGAAGG TTGCATTTTT AAAGACTATT GGGTTTGTGG

19001 TACGACTACA TGAACAAGCG AGTGGTGGCT CCCGGGCTAG TGGACTGCTA  
ATGCTGATGT ACTTGTTTCG CACCAACCGA GGGCCCGATC ACCTGACGAT

19051 CATTAACCTT GGAGCACGCT GGTCCCTTGA CTATATGGAC AACGTCAACC  
GTAATTGGAA CCTCGTGCGA CCAGGGAACCT GATATACCTG TTGCAGTTGG

19101 CATTTAACCA CCACCGCAAT GCTGGCCTGC GCTACCGCTC AATGTTGCTG  
GTAAATTGGT GGTGGCGTTA CGACCGGACG CGATGGCGAG TTACAACGAC

19151 GGCAATGGTC GCTATGTGCC CTTCCACATC CAGGTGCCTC AGAAGTTCTT  
CCGTTACCAG CGATACACGG GAAGGTGTAG GTCCACGGAG TCTTCAAGAA

19201 TGCCATTAAA AACCTCCTTC TCCTGCCGGG CTCATACACC TACGAGTGGA  
ACGGTAATTT TTGGAGGAAG AGGACGGCCC GAGTATGTGG ATGCTCACCT

19251 ACTTCAGGAA GGATGTTAAC ATGGTTCTGC AGAGCTCCCT AGGAAATGAC  
TGAAGTCCTT CCTACAATTG TACCAAGACG TCTCGAGGGA TCCTTTACTG

19301 CTAAGGGTTG ACGGAGCCAG CATTAAGTTT GATAGCATTT GCCTTTACGC  
GATTCCCAAC TGCCTCGGTC GTAATTCAA CTATCGTAAA CGGAAATGCG

19351 CACCTTCTTC CCCATGGCCC ACAACACCGC CTCCACGCTT GAGGCCATGC  
GTGGAAGAAG GGGTACCGGG TGTGTGGCG GAGGTGCGAA CTCCGGTACG

19401 TTAGAAACGA CACCAACGAC CAGTCCTTTA ACGACTATCT CTCCGCCGCC  
AATCTTTGCT GTGGTTGCTG GTCAGGAAAT TGCTGATAGA GAGGCGGCGG

19451 AACATGCTCT ACCCTATACC CGCCAACGCT ACCAACGTGC CCATATCCAT  
TTGTACGAGA TGGGATATGG GCGGTTGCGA TGGTTGCACG GGTATAGGTA

19501 CCCCTCCCGC AACTGGGCGG CTTTCCGCGG CTGGGCCTTC ACGCGCCTTA  
GGGGAGGGCG TTGACCCGCC GAAAGGCGCC GACCCGGAAG TGCGCGGAAT

19551 AGACTAAGGA AACCCCATCA CTGGGCTCGG GCTACGACCC TTATTACACC  
TCTGATTCCT TTGGGGTAGT GACCCGAGCC CGATGCTGGG AATAATGTGG

19601 TACTCTGGCT CTATACCCTA CCTAGATGGA ACCTTTTACC TCAACCACAC  
ATGAGACCGA GATATGGGAT GGATCTACCT TGGAAAATGG AGTTGGTGTG

19651 CTTTAAGAAG GTGGCCATTA CTTTGTACTC TTCTGTCAGC TGGCCTGGCA  
GAAATTCTTC CACCGGTAAT GGAAACTGAG AAGACAGTCG ACCGGACCGT

19701 ATGACCGCCT GCTTACCCCC AACGAGTTTG AAATTAAGCG CTCAGTTGAC  
TACTGGCGGA CGAATGGGGG TTGCTCAAAC TTTAATTTCG GAGTCAACTG

19751 GGGGAGGGTT ACAACGTTGC CCAGTGTAAC ATGACCAAAG ACTGGTTCCT  
CCCCTCCCAA TGTGCAACG GGTCACATTG TACTGGTTTC TGACCAAGGA

19801 GGTACAAATG CTAGCTAACT ATAACATTGG CTACCAGGGC TTCTATATCC  
CCATGTTTAC GATCGATTGA TATTGTAACC GATGGTCCCG AAGATATAGG

Figure 274

19851	CAGAGAGCTA GTCTCTCGAT	CAAGGACCGC GTTCTTGGCG	ATGTACTCCT TACATGAGGA	TCTTTAGAAA AGAAATCTTT	CTTCCAGCCC GAAGGTCGGG
19901	ATGAGCCGTC TACTCGGCAG	AGGTGGTGGA TCCACCACCT	TGATACTAAA ACTATGATTT	TACAAGGACT ATGTTCTCTGA	ACCAACAGGT TGGTTGTCCA
19951	GGGCATCCTA CCCGTAGGAT	CACCAACACA GTGGTTGTGT	ACAACTCTGG TGTTGAGACC	ATTTGTTGGC TAAACAACCG	TACCTTGCCC ATGGAACGGG
20001	CCACCATGCG GGTGGTACGC	CGAAGGACAG GCTTCCTGTC	GCCTACCCTG CGGATGGGAC	CTAACTTCCC GATTGAAGGG	CTATCCGCTT GATAGGCGAA
20051	ATAGGCAAGA TATCCGTTCT	CCGCAGTTGA GGCGTCAACT	CAGCATTACC GTCGTAATGG	CAGAAAAAGT GTCCTTTTCA	TTCTTTGCGA AAGAAACGCT
20101	TCGCACCCTT AGCGTGGGAA	TGGCGCATCC ACCGCGTAGG	CATTCTCCAG GTAAGAGGTC	TAACTTTATG ATTGAAATAC	TCCATGGGCG AGGTACCCGC
20151	CACTCACAGA GTGAGTGTCT	CCTGGGCCAA GGACCCGGTT	AACCTTCTCT TTGGAAGAGA	ACGCCAACTC TGCGGTTGAG	CGCCCACGCG GCGGGTGCGC
20201	CTAGACATGA GATCTGTACT	CTTTTGAGGT GAAAACTCCA	GGATCCCATG CCTAGGGTAC	GACGAGCCCA CTGCTCGGGT	CCCTTCTTTA GGGAAGAAAT
20251	TGTTTTGTTT ACAAAACAAA	GAAGTCTTTG CTTCAGAAAC	ACGTGGTCCG TGCACCAGGC	TGTGCACCAG ACACGTGGTC	CCGCACCGCG GGCGTGGCGC
20301	GCGTCATCGA CGCAGTAGCT	AACCGTGTAC TTGGCACATG	CTGCGCACGC GACGCGTGCG	CCTTCTCGGC GGAAGAGCCG	CGGCAACGCC GCCGTTGCGG
20351	ACAACATAAA TGTTGTATTT	GAAGCAAGCA CTTCGTTCGT	ACATCAACAA TGTAGTTGTT	CAGCTGCCGC GTCGACGGCG	CATGGGCTCC GTACCCGAGG
20401	AGTGAGCAGG TCACTCGTCC	AACTGAAAGC TTGACTTTTCG	CATTGTCAAA GTAACAGTTT	GATCTTGGTT CTAGAACCAA	GTGGGCCATA CACCCGGTAT
20451	TTTTTTGGGC AAAAAACCCG	ACCTATGACA TGGATACTGT	AGCGCTTTCC TCGCGAAAGG	AGGCTTTGTT TCCGAAACAA	TCTCCACACA AGAGGTGTGT
20501	AGCTCGCCTG TCGAGCGGAC	CGCCATAGTC GCGGTATCAG	AATACGGCCG TTATGCCGGC	GTCGCGAGAC CAGCGCTCTG	TGGGGGCGTA ACCCCGCAT
20551	CACTGGATGG GTGACCTACC	CCTTTGCCTG GGAAACGGAC	GAACCCGCAC CTTGGGCGTG	TCAAAAACAT AGTTTTTGTA	GCTACCTCTT CGATGGAGAA
20601	TGAGCCCTTT ACTCGGGAAA	GGCTTTTCTG CCGAAAAGAC	ACCAGCGACT TGGTCGCTGA	CAAGCAGGTT GTTCGTCCAA	TACCAGTTTG ATGGTCAAAC
20651	AGTACGAGTC TCATGCTCAG	ACTCCTGCGC TGAGGACGCG	CGTAGCGCCA GCATCGCGGT	TTGCTTCTTC AACGAAGAAG	CCCCGACCGC GGGGCTGGCG
20701	TGTATAACGC ACATATTGCG	TGGAAAAGTC ACCTTTTCAG	CACCCAAAGC GTGGGTTTCG	GTACAGGGGC CATGTCCCCG	CCAACTCGGC GGTTGAGCCG
20751	CGCCTGTGGA GCGGACACCT	CTATTCTGCT GATAAGACGA	GCATGTTTCT CGTACAAAGA	CCACGCCTTT GGTGCGGAAA	GCCAACTGGC CGGTTGACCG

Figure 27 V.

20851 CCCAACTCCA TGCTCAACAG TCCCCAGGTA CAGCCCACCC TCGGTCGCAA  
GGGTTGAGGT ACGAGTTGTC AGGGGTCCAT GTCGGGTGGG ACGCAGCGTT

20901 CCAGGAACAG CTCTACAGCT TCCTGGAGCG CCACTCGCCC TACTTCCGCA  
GGTCCTTGTC GAGATGTCGA AGGACCTCGC GGTGAGCGGG ATGAAGGCGT

20951 GCCACAGTGC GCAGATTAGG AGCGCCACTT CTTTTTGTC AATTGAAAAAC  
CGGTGTCACG CGTCTAATCC TCGCGGTGAA GAAAAACAGT GAACTTTTTG

21001 ATGTAAAAAT AATGTACTAG AGACACTTTC AATAAAGGCA AATGCTTTTA  
TACATTTTTA TTACATGATC TCTGTGAAAG TTATTTCCGT TTACGAAAAT

21051 TTTGTACACT CTCGGGTGAT TATTTACCCC CACCCTTGCC GTCTGCGCCG  
AAACATGTGA GAGCCCACTA ATAAATGGGG GTGGGAACGG CAGACGCGGC

21101 TTTAAAAATC AAAGGGGTTC TGCCGCGCAT CGCTATGCGC CACTGGCAGG  
AAATTTTGTAG TTTCCCAAG ACGGCGCGTA GCGATACGCG GTGACCGTCC

21151 GACACGTTGC GATACTGGTG TTTAGTGCTC CACTTAAACT CAGGCACAAC  
CTGTGCAACG CTATGACCAC AAATCACGAG GTGAATTTGA GTCCGTGTTG

21201 CATCCGCGGC AGCTCGGTGA AGTTTTCACT CCACAGGCTG CGCACCATCA  
GTAGGCGCCG TCGAGCCACT TCAAAAGTGA GGTGTCCGAC GCGTGGTAGT

21251 CCAACGCGTT TAGCAGGTCG GCGCCGATA TCTTGAAGTC GCAGTTGGGG  
GGTTGCGCAA ATCGTCCAGC CCGCGGCTAT AGAACTTCAG CGTCAACCCC

21301 CCTCCGCCCT GCGCGCGCGA GTTGCGATAC ACAGGGTTGC AGCACTGGAA  
GGAGGCGGGA CGCGCGCGCT CAACGCTATG TGTCCCAACG TCGTGACCTT

21351 CACTATCAGC GCCGGGTGGT GCACGCTGGC CAGCACGCTC TTGTCGGAGA  
GTGATAGTCG CGGCCACCA CGTGCGACCG GTCGTGCGAG AACAGCCTCT

21401 TCAGATCCGC GTCCAGGTCC TCCGCGTTGC TCAGGGCGAA CGGAGTCAAC  
AGTCTAGGCG CAGGTCCAGG AGGCGCAACG AGTCCCGCTT GCCTCAGTTG

21451 TTTGGTAGCT GCCTTCCCAA AAAGGGCGCG TGCCCAGGCT TTGAGTTGCA  
AAACCATCGA CGGAAGGGTT TTTCCCGCGC ACGGGTCCGA AACTCAACGT

21501 CTCGCACCGT AGTGGCATCA AAAGGTGACC GTGCCCCGTC TGGGCGTTAG  
GAGCGTGGCA TCACCGTAGT TTTCCACTGG CACGGGCCAG ACCCGCAATC

21551 GATACAGCGC CTGCATAAAA GCCTTGATCT GCTTAAAAGC CACCTGAGCC  
CTATGTCGCG GACGTATTTT CGGAAGTAGA CGAATTTTCG GTGGACTCGG

21601 TTTGCGCCTT CAGAGAAGAA CATGCCGCAA GACTTGCCGG AAAACTGATT  
AAACGCGGAA GTCTCTTCTT GTACGGCGTT CTGAACGGCC TTTTGACTAA

21651 GGCCGGACAG GCCGCGTCGT GCACGCAGCA CCTTGCGTCG GTGTTGGAGA  
CCGGCCTGTC CGGCGCAGCA CGTGCGTCGT GGAACGCAGC CACAACCTCT

21701 TCTGCACCAC ATTTCCGCCC CACCGGTTCT TCACGATCTT GGCCTTGCTA  
AGACGTGGTG TAAAGCCGGG GTGGCCAAGA AGTGCTAGAA CCGGAACGAT

Figure 27 W

21801 AATCACGTGC TCCTTATTTA TCATAATGCT TCCGTGTAGA CACTTAAGCT  
TTAGTGCACG AGGAATAAAT AGTATTACGA AGGCACATCT GTGAATTCGA

21851 CGCCTTCGAT CTCAGCGCAG CGGTGCAGCC ACAACGCGCA GCCCGTGGGC  
GCGGAAGCTA GAGTCGCGTC GCCACGTCGG TGTTGCGCGT CGGGCACCCG

21901 TCGTGATGCT TGTAGGTCAC CTCTGCAAAC GACTGCAGGT ACGCCTGCAG  
AGCACTACGA ACATCCAGTG GAGACGTTTG CTGACGTCCA TGCGGACGTC

21951 GAATCGCCCC ATCATCGTCA CAAAGGTCTT GTTGCTGGTG AAGGTCAGCT  
CTTAGCGGGG TAGTAGCAGT GTTTCCAGAA CAACGACCAC TTCCAGTCGA

22001 GCAACCCGCG GTGCTCCTCG TTCAGCCAGG TCTTGCATAC GGCCGCCAGA  
CGTTGGGCGC CACGAGGAGC AAGTCGGTCC AGAACGTATG CCGGCGGTCT

22051 GCTTCCACTT GGTCAGGCAG TAGTTTGAAG TTCGCCTTTA GATCGTTATC  
CGAAGGTGAA CCAGTCCGTC ATCAAAC TTCGCGGAAAT CTAGCAATAG

22101 CACGTGGTAC TTGTCCATCA GCGCGCGCGC AGCCTCCATG CCCTTCTCCC  
GTGCACCATG AACAGGTAGT CGCGCGCGCG TCGGAGGTAC GGGAAGAGGG

22151 ACGCAGACAC GATCGGCACA CTCAGCGGGT TCATCACCGT AATTTCACTT  
TGCGTCTGTG CTAGCCGTGT GAGTCGCCCA AGTAGTGGCA TTAAAGTGAA

22201 TCCGCTTCGC TGGGCTCTTC CTCTTCCTCT TCGGTCCGCA TACCACGCGC  
AGGCGAAGCG ACCCGAGAAG GAGAAGGAGA ACGCAGGCGT ATGGTGCGCG

22251 CACTGGGTCTG TCTTCATTCA GCCGCCGCAC TGTGCGCTTA CCTCCTTTGC  
GTGACCCAGC AGAAGTAAGT CGGCGGCGTG ACACGCGAAT GGAGGAAACG

22301 CATGCTTGAT TAGCACCGGT GGGTTGCTGA AACCCACCAT TTGTAGCGCC  
GTACGAACTA ATCGTGGCCA CCCAACGACT TTGGGTGGTA AACATCGCGG

22351 ACATCTTCTC TTTCTTCCTC GCTGTCCACG ATTACCTCTG GTGATGGCGG  
TGTAGAAGAG AAAGAAGGAG CGACAGGTGC TAATGGAGAC CACTACCGCC

22401 GCGCTCGGGC TTGGGAGAAG GCGCTTCTT TTTCTTCTTG GGCGCAATGG  
CGCGAGCCCG AACCTCTTC CCGCGAAGAA AAAGAAGAAC CCGCGTTACC

22451 CCAAATCCGC CGCCGAGGTC GATGGCCGCG GGCTGGGTGT GCGCGGCACC  
GGTTTAGGCG GCGGCTCCAG CTACCGGCGC CCGACCCACA CGCGCCGTGG

22501 AGCGCGTCTT GTGATGAGTC TTCCTCGTCC TCGGACTCGA TACGCCGCCT  
TCGCGCAGAA CACTACTCAG AAGGAGCAGG AGCCTGAGCT ATGCGGCGGA

22551 CATCCGCTTT TTTGGGGGCG CCCGGGGAGG CGGCGGCGAC GGGGACGGGG  
GTAGGCGAAA AAACCCCGC GGGCCCCTCC GCCGCCGCTG CCCCTGCCCC

22601 ACGACACGTC CTCCATGGTT GGGGGACGTC GCGCCGCACC GCGTCCGCGC  
TGCTGTGCAG GAGGTACCAA CCCCTGCAG CGCGGCGTGG CGCAGGCGCG

22651 TCGGGGGTGG TTTCGCGCTG CTCCTCTTCC CGACTGGCCA TTTCCTTCTC  
AGCCCCACC AAAGCGCGAC GAGGAGAAGG GCTGACCGGT AAAGGAAGAG

Figure 27 X

22751 CCGCCCCCTC TGAGTTCGCC ACCACCGCCT CCACCGATGC CGCCAACGCG  
GGCGGGGGAG ACTCAAGCGG TGGTGGCGGA GGTGGCTACG GCGGTTGCGC

22801 CCTACCACCT TCCCCGTCGA GGCACCCCCG CTTGAGGAGG AGGAAGTGAT  
GGATGGTGGA AGGGGCAGCT CCGTGGGGGC GAACTCCTCC TCCTTCACTA

22851 TATCGAGCAG GACCCAGGTT TTGTAAGCGA AGACGACGAG GACCGCTCAG  
ATAGCTCGTC CTGGGTCCAA AACATTCGCT TCTGCTGCTC CTGGCGAGTC

22901 TACCAACAGA GGATAAAAAG CAAGACCAGG ACAACGCAGA GGCAAACGAG  
ATGGTTGTCT CCTATTTTTC GTTCTGGTCC TGTTCGCTCT CCGTTTGCTC

22951 GAACAAGTCG GGCGGGGGGA CGAAAGGCAT GCGGACTACC TAGATGTGGG  
CTTGTTTCAGC CCGCCCCCCT GCTTTCCGTA CCGCTGATGG ATCTACACCC

23001 AGACGACGTG CTGTTGAAGC ATCTGCAGCG CCAGTGCGCC ATTATCTGCG  
TCTGCTGCAC GACAACCTCG TAGACGTCGC GGTACGCGG TAATAGACGC

23051 ACGCGTTGCA AGAGCGCAGC GATGTGCCCC TCGCCATAGC GGATGTCAGC  
TGCGCAACGT TCTCGCGTCG CTACACGGGG AGCGGTATCG CCTACAGTCG

23101 CTTGCCTACG AACGCCACCT ATTCTCACC GCGGTACCCC CCAAACGCCA  
GAACGGATGC TTGCGGTGGA TAAGAGTGGC GCGCATGGGG GGTTTGCGGT

23151 AGAAAACGGC ACATGCGAGC CCAACCCGCG CCTCAACTTC TACCCCGTAT  
TCTTTTGCCG TGTACGCTCG GGTGGGCGC GGAGTTGAAG ATGGGGCATA

23201 TTGCCGTGCC AGAGGTGCTT GCCACCTATC ACATCTTTTT CCAAACCTGC  
AACGGCACGG TCTCCACGAA CGGTGGATAG TGTAGAAAAA GGTTTGTACG

23251 AAGATACCCC TATCCTGCCG TGCCAACCGC AGCCGAGCGG ACAAGCAGCT  
TTCTATGGGG ATAGGACGGC ACGGTTGGCG TCGGCTCGCC TGTTCGTCGA

23301 GGCCTTGCGG CAGGGCGCTG TCATACCTGA TATCGCCTCG CTCAACGAAG  
CCGGAACGCC GTCCCGCGAC AGTATGGACT ATAGCGGAGC GAGTTGCTTC

23351 TGCCAAAAAT CTTTGAGGGT CTTGGACGCG ACGAGAAGCG CGCGGCAAAC  
ACGGTTTTTA GAAACTCCCA GAACCTGCGC TGCTCTTCGC GCGCCGTTTG

23401 GCTCTGCAAC AGGAAAACAG CGAAAATGAA AGTCACTCTG GAGTGTTGGT  
CGAGACGTTG TCCTTTTGTC GCTTTTACTT TCAGTGAGAC CTCACAACCA

23451 GGAACCTCGAG GGTGACAACG CGCGCCTAGC CGTACTAAAA CGCAGCATCG  
CCTTGAGCTC CCACTGTTGC GCGCGGATCG GCATGATTTT GCGTCGTAGC

23501 AGGTCACCCA CTTTGCCTAC CCGGCACTTA ACCTACCCCC CAAGGTCATG  
TCCAGTGGGT GAAACGGATG GGCCGTGAAT TGGATGGGGG GTTCCAGTAC

23551 AGCACAGTCA TGAGTGAGCT GATCGTGCGC CGTGCGCAGC CCCTGGAGAG  
TCGTGTCAGT ACTCACTCGA CTAGCACGCG GCACGCGTCG GGGACCTCTC

23601 GGATGCAAAT TTGCAAGAAC AAACAGAGGA GGGCCTACCC GCAGTTGGCG  
CCTACGTTTA AACGTTCTTG TTTGTCTCCT CCCGGATGGG CGTCAACCGC

Figure 27 Y

23701 GAGCGACGCA AACTAATGAT GGCCGCAGTG CTCGTTACCG TGGAGCTTGA  
CTCGCTGCGT TTGATTACTA CCGGCGTCAC GAGCAATGGC ACCTCGAACT

23751 GTGCATGCAG CGGTTCTTTG CTGACCCGGA GATGCAGCGC AAGCTAGAGG  
CACGTACGTC GCCAAGAAAC GACTGGGCCT CTACGTCGCG TTCGATCTCC

23801 AAACATTGCA CTACACCTTT CGACAGGGCT ACGTACGCCA GGCCTGCAAG  
TTTGTAACGT GATGTGGAAA GCTGTCCCGA TGCATGCGGT CCGGACGTTC

23851 ATCTCCAACG TGGAGCTCTG CAACCTGGTC TCCTACCTTG GAATTTTGCA  
TAGAGGTTGC ACCTCGAGAC GTTGGACCAG AGGATGGAAC CTTAAAACGT

23901 CGAAAACCGC CTTGGGCAAA ACGTGCTTCA TTCCACGCTC AAGGGCGAGG  
GCTTTTGGCG GAACCCGTTT TGCACGAAGT AAGGTGCGAG TTCCCGCTCC

23951 CGCGCCGCGA CTACGTCCGC GACTGCGTTT ACTTATTTCT ATGCTACACC  
GCGCGGCGCT GATGCAGGCG CTGACGCAAA TGAATAAAGA TACGATGTGG

24001 TGGCAGACGG CCATGGGCGT TTGGCAGCAG TGCTTGGAGG AGTGCAACCT  
ACCGTCTGCC GGTACCCGCA AACCGTCGTC ACGAACCTCC TCACGTTGGA

24051 CAAGGAGCTG CAGAACTGC TAAAGCAAAA CTTGAAGGAC CTATGGACGG  
GTTCTCGAC GTCTTTGACG ATTTCTGTTT GAACTTCCTG GATACCTGCC

24101 CCTTCAACGA GCGCTCCGTG GCCGCGCACC TGGCGGACAT CATTTTCCCC  
GGAAGTTGCT CGCGAGGCAC CGGCGCGTGG ACCGCCTGTA GTAAAAGGGG

24151 GAACGCCTGC TTAAAACCTT GCAACAGGGT CTGCCAGACT TCACCAGTCA  
CTTGCGGACG AATTTTGGGA CGTTGTCCCA GACGGTCTGA AGTGGTCAGT

24201 AAGCATGTTG CAGAACTTTA GGAACTTTAT CCTAGAGCGC TCAGGAATCT  
TTCGTACAAC GTCTTGAAAT CCTTGAAATA GGATCTCGCG AGTCCTTAGA

24251 TGCCCGCCAC CTGCTGTGCA CTTCTAGCG ACTTTGTGCC CATTAAGTAC  
ACGGGCGGTG GACGACACGT GAAGGATCGC TGAAACACGG GTAATTCATG

24301 CGCGAATGCC CTCCGCCGCT TTGGGGCCAC TGCTACCTTC TGCAGCTAGC  
GCGCTTACGG GAGGCGGCGA AACCCCGGTG ACGATGGAAG ACGTCGATCG

24351 CAACTACCTT GCCTACCACT CTGACATAAT GGAAGACGTG AGCGGTGACG  
GTTGATGGAA CGGATGGTGA GACTGTATTA CCTTCTGCAC TCGCCACTGC

24401 GTCTACTGGA GTGTCACTGT CGCTGCAACC TATGCACCCC GCACCGCTCC  
CAGATGACCT CACAGTGACA GCGACGTTGG ATACGTGGGG CGTGGCGAGG

24451 CTGGTTTGCA ATTCGCAGCT GCTTAACGAA AGTCAAATTA TCGGTACCTT  
GACCAAACGT TAAGCGTCGA CGAATTGCTT TCAGTTTAAT AGCCATGGAA

24501 TGAGCTGCAG GGTCCCTCGC CTGACGAAAA GTCCGCGGCT CCGGGGTTGA  
ACTCGACGTC CCAGGGAGCG GACTGCTTTT CAGGCGCCGA GGCCCCAACT

24551 AACTCACTCC GGGGCTGTGG ACGTCGGCTT ACCTTCGCAA ATTTGTACCT  
TTGAGTGAGG CCCCACACCC TGCAGCCGAA TGGAAGCGTT TAAACATGGA

Figure 272

24601 GAGGACTACC ACGCCCACGA GATTAGGTTC TACGAAGACC AATCCCGCCC  
CTCCTGATGG TGCGGGTGCT CTAATCCAAG ATGCTTCTGG TTAGGGCGGG

24651 GCCTAATGCG GAGCTTACCG CCTGCGTCAT TACCCAGGGC CACATTCTTG  
CGGATTACGC CTCGAATGGC GGACGCAGTA ATGGGTCCCC GTGTAAGAAC

24701 GCCAATTGCA AGCCATCAAC AAAGCCCGCC AAGAGTTTCT GCTACGAAAG  
CGGTTAACGT TCGGTAGTTG TTTCGGGCGG TTCTCAAAGA CGATGCTTTC

24751 GGACGGGGGG TTTACTTGGA CCCCCAGTCC GGCGAGGAGC TCAACCCAAT  
CCTGCCCCCC AAATGAACCT GGGGGTCAGG CCGCTCCTCG AGTTGGGTTA

24801 CCCCCGCGG CCGCAGCCCT ATCAGCAGCA GCCGCGGGCC CTTGCTTCCC  
GGGGGGCGGC GGCGTCGGGA TAGTCGTCGT CGGCGCCCCG GAACGAAGGG

24851 AGGATGGCAC CCAAAAAGAA GCTGCAGCTG CCGCCGCCAC CCACGGACGA  
TCCTACCGTG GGTTTTTCTT CGACGTCGAC GGCGGCGGTG GGTGCCTGCT

24901 GGAGGAATAC TGGGACAGTC AGGCAGAGGA GGTTTTGGAC GAGGAGGAGG  
CCTCCTTATG ACCCTGTCAG TCCGTCTCCT CCAAACCTG CTCCTCCTCC

24951 AGGACATGAT GGAAGACTGG GAGAGCCTAG ACGAGGAAGC TTCCGAGGTC  
TCCTGTACTA CCTTCTGACC CTCTCGGATC TGCTCCTTCG AAGGCTCCAG

25001 GAAGAGGTGT CAGACGAAAC ACCGTCACCC TCGGTGCGAT TCCCCTCGCC  
CTTCTCCACA GTCTGCTTTG TGGCAGTGGG AGCCAGCGTA AGGGGAGCGG

25051 GGCGCCCCAG AAATCGGCAA CCGGTTCCAG CATGGCTACA ACCTCCGCTC  
CCGCGGGGTC TTTAGCCGTT GGCCAAGGTC GTACCGATGT TGGAGGCGAG

25101 CTCAGGCGCC GCCGGCACTG CCCGTTCGCC GACCCAACCG TAGATGGGAC  
GAGTCCGCGG CGGCCGTGAC GGGCAAGCGG CTGGGTTGGC ATCTACCCTG

25151 ACCACTGGAA CCAGGGCCGG TAAGTCCAAG CAGCCGCCGC CGTTAGCCCA  
TGGTGACCTT GGTCCCGGCC ATTCAGGTTC GTCGGCGGCG GCAATCGGGT

25201 AGAGCAACAA CAGCGCCAAG GCTACCGCTC ATGGCGCGGG CACAAGAACG  
TCTCGTTGTT GTCGCGGTTT CGATGGCGAG TACCGCGCCC GTGTTCTTGC

25251 CCATAGTTGC TTGCTTGCAA GACTGTGGGG GCAACATCTC CTTGCCCCGC  
GGTATCAACG AACGAACGTT CTGACACCCC CGTTGTAGAG GAAGCGGGCG

25301 CGCTTTCTTC TCTACCATCA CGGCGTGGCC TTCCCCCGTA ACATCCTGCA  
GCGAAAGAAG AGATGGTAGT GCCGCACCGG AAGGGGGCAT TGTAGGACGT

25351 TTACTACCGT CATCTCTACA GCCATACTG CACCGGCGGC AGCGGCAGCA  
AATGATGGCA GTAGAGATGT CGGGTATGAC GTGGCCGCCG TCGCCGTCGT

25401 ACAGCAGCGG CCACACAGAA GCAAAGGCGA CCGGATAGCA AGACTCTGAC  
TGTCGTGCGC GGTGTGTCTT CGTTTCCGCT GGCCTATCGT TCTGAGACTG

25451 AAAGCCCAAG AAATCCACAG CGGCGGCAGC AGCAGGAGGA GGAGCGCTGC  
TTTCGGGTTC TTTAGGTGTC GCCGCCGTCG TCGTCCTCCT CCTCGCGACG

25501 GTCTGGCGCC CAACGAACCC GTATCGACCC GCGAGCTTAG AAACAGGATT  
CAGACCGCGG GTTGCTTGGG CATAGCTGGG CGCTCGAATC TTTGTCCTAA

Figure 27 AA

25551 TTTCCCACTC TGTATGCTAT ATTTCAACAG AGCAGGGGCC AAGAACAAGA  
AAAGGGTGAG ACATACGATA TAAAGTTGTC TCGTCCCCCG TTCTTGTTCT

25601 GCTGAAAATA AAAAACAGGT CTCTGCGATC CCTCACCCGC AGCTGCCTGT  
CGACTTTTAT TTTTGTCCA GAGACGCTAG GGAGTGGGCG TCGACGGACA

25651 ATCACAAAAG CGAAGATCAG CTTCGGCGCA CGCTGGAAGA CGCGGAGGCT  
TAGTGTTTTC GCTTCTAGTC GAAGCCGCGT GCGACCTTCT GCGCCTCCGA

25701 CTCTTCAGTA AATACTGCGC GCTGACTCTT AAGGACTAGT TTCGCGCCCT  
GAGAAGTCAT TTATGACGCG CACTGAGAA TTCCTGATCA AAGCGCGGGA

25751 TTCTCAAATT TAAGCGCGAA AACTACGTCA TCTCCAGCGG CCACACCCGG  
AAGAGTTTAA ATTCGCGCTT TTGATGCAGT AGAGGTCGCC GGTGTGGGCC

25801 CGCCAGCACC TGTTGTCAGC GCCATTATGA GCAAGGAAAT TCCCACGCCC  
GCGGTCGTGG ACAACAGTCG CGGTAATACT CGTTCCTTTA AGGGTGCGGG

25851 TACATGTGGA GTTACCAGCC ACAAATGGGA CTTGCGGCTG GAGCTGCCCA  
ATGTACACCT CAATGGTCGG TGTTTACCCT GAACGCCGAC CTCGACGGGT

25901 AGACTACTCA ACCCGAATAA ACTACATGAG CGCGGGACCC CACATGATAT  
TCTGATGAGT TGGGCTTATT TGATGTACTC GCGCCCTGGG GTGTACTATA

25951 CCCGGGTCAA CGGAATACGC GCCCACCGAA ACCGAATTCT CCTGGAACAG  
GGCCCCAGTT GCCTTATGCG CGGGTGGCTT TGGCTTAAGA GGACCTTGTC

26001 GCGGCTATTA CCACCACACC TCGTAATAAC CTTAATCCCC GTAGTTGGCC  
CGCCGATAAT GGTGGTGTGG AGCATTATTG GAATTAGGGG CATCAACCGG

26051 CGCTGCCCTG GTGTACCAGG AAAGTCCCGC TCCCACCACT GTGGTACTTC  
GCGACGGGAC CACATGGTCC TTTCAGGGCG AGGGTGGTGA CACCATGAAG

26101 CCAGAGACGC CCAGGCCGAA GTTCAGATGA CTAATCAGG GCGCAGCTT  
GGTCTCTGCG GGTCCGGCTT CAAGTCTACT GATTGAGTCC CCGCGTCGAA

26151 GCGGGCGGCT TTCGTCACAG GGTGCGGTCG CCCGGGCAGG GTATAACTCA  
CGCCCGCCGA AAGCAGTGTC CCACGCCAGC GGGCCCGTCC CATATTGAGT

26201 CCTGACAATC AGAGGGCGAG GTATTCAGCT CAACGACGAG TCGGTGAGCT  
GGACTGTTAG TCTCCCGCTC CATAAGTCGA GTTGCTGCTC AGCCACTCGA

26251 CCTCGCTTGG TCTCCGTCCG GACGGGACAT TTCAGATCGG CGGCGCCGGC  
GGAGCGAACC AGAGGCAGGC CTGCCCTGTA AAGTCTAGCC GCCGCGGCCG

26301 CGCTCTTCAT TCACGCCTCG TCAGGCAATC CTAATCTGC AGACCTCGTC  
GCGAGAAGTA AGTGCGGAGC AGTCCGTTAG GATTGAGACG TCTGGAGCAG

26351 CTCTGAGCCG CGCTCTGGAG GCATTGGAAC TCTGCAATTT ATTGAGGAGT  
GAGACTCGGC GCGAGACCTC CGTAACCTTG AGACGTTAA TAACTCCTCA

26401 TTGTGCCATC GGTCTACTTT AACCCTTCT CGGGACCTCC CGGCCACTAT  
AACACGGTAG CCAGATGAAA TTGGGGAAGA GCCCTGGAGG GCCGGTGATA

26451 CCGGATCAAT TTATTCCTAA CTTTGACGCG GTAAAGGACT CGGCGGACGG  
GGCCTAGTTA AATAAGGATT GAAACTGCGC CATTCCTGA GCCGCCTGCC

Figure 27 AB

WO 02/022080

PCT/US01/28861

26501 CTACGACTGA ATGTTAAGTG GAGAGGCAGA GCAACTGCGC CTGAAACACC  
GATGCTGACT TACAATTCAC CTCTCCGTCT CGTTGACGCG GACTTTGTGG

26551 TGGTCCACTG TCGCCGCCAC AAGTGCTTTG CCCGCGACTC CGGTGAGTTT  
ACCAGGTGAC AGCGGCGGTG TTCACGAAAC GGGCGCTGAG GCCACTCAAA

26601 TGCTACTTTG AATTGCCCCG GGATCATATC GAGGGCCCCG CGCACGGCGT  
ACGATGAAAC TTAACGGGCT CCTAGTATAG CTCCCAGGGC GCGTGCCGCA

26651 CCGGCTTACC GCCCAGGGAG AGCTTGCCCCG TAGCCTGATT CGGGAGTTTA  
GGCCGAATGG CGGGTCCCTC TCGAACGGGC ATCGGACTAA GCCCTCAAAT

26701 CCCAGCGCCC CCTGCTAGTT GAGCGGGACA GGGGACCCTG TGTTCCTCACT  
GGGTCGCGGG GGACGATCAA CTCGCCCTGT CCCCTGGGAC ACAAGAGTGA

26751 GTGATTTGCA ACTGTCCTAA CCCTGGATTA CATCAAGATC TTTGTTGCCA  
CACTAAACGT TGACAGGATT GGGACCTAAT GTAGTTCTAG AAACAACGGT

26801 TCTCTGTGCT GAGTATAATA AATACAGAAA TTAAAATATA CTGGGGCTCC  
AGAGACACGA CTCATATTAT TTATGTCTTT AATTTTATAT GACCCCGAGG

26851 TATCGCCATC CTGTAAACGC CACCGTCTTC ACCCGCCCAA GCAAACCAAG  
ATAGCGGTAG GACATTTGCG GTGGCAGAAG TGGGCGGGTT CGTTTGGTTC

26901 GCGAACCTTA CCTGGTACTT TTAACATCTC TCCCTCTGTG ATTTACAACA  
CGCTTGGAAT GGACCATGAA AATTGTAGAG AGGGAGACAC TAAATGTTGT

26951 GTTTCAACCC AGACGGAGTG AGTCTACGAG AGAACCTCTC CGAGCTCAGC  
CAAAGTTGGG TCTGCCTCAC TCAGATGCTC TCTTGAGAG GCTCGAGTCG

27001 TACTCCATCA GAAAAACAC CACCCTCCTT ACCTGCCGGG AACGTACGAG  
ATGAGGTAGT CTTTTTTGTG GTGGGAGGAA TGGACGGCCC TTGCATGCTC

27051 TCGGTCACCG GCCGCTGCAC CACACCTACC GCCTGACCGT AAACCAGACT  
ACGCAGTGGC CGGCGACGTG GTGTGGATGG CGGACTGGCA TTTGGTCTGA

27101 TTTTCCGGAC AGACCTCAAT AACTCTGTTT ACCAGAACAG GAGGTGAGCT  
AAAAGGCCTG TCTGGAGTTA TTGAGACAAA TGGTCTTGTC CTCCACTCGA

27151 TAGAAAACCC TTAGGGTATT AGGCCAAAGG CGCAGCTACT GTGGGGTTTA  
ATCTTTTGGG AATCCCATAA TCCGGTTTCC GCGTCGATGA CACCCCAAAT

27201 TGAACAATTC AAGCAACTCT ACGGGCTATT CTAATTCAGG TTTCTCTAGA  
ACTTGTTAAG TTCGTTGAGA TGCCCGATAA GATTAAGTCC AAAGAGATCT

27251 ATCGGGGTTG GGGTTATTCT CTGTCTTGTC ATTCTCTTTA TTCTTATACT  
TAGCCCCAAC CCCAATAAGA GACAGAACAC TAAGAGAAAT AAGAATATGA

27301 AACGCTTCTC TGCCTAAGGC TCGCCGCCTG CTGTGTGCAC ATTTGCATTT  
TTGCGAAGAG ACGGATTCCG AGCGGCGGAC GACACACGTG TAAACGTAAA

27351 ATTGTCAGCT TTTTAAACGC TGGGGTCGCC ACCCAAGATG ATTAGGTACA  
TAACAGTCGA AAAATTTGCG ACCCCAGCGG TGGGTCTAC TAATCCATGT

27401 TAATCCTAGG TTTACTCACC CTTGCGTCAG CCCACGGTAC CACCCAAAAG  
ATTAGGATCC AAATGAGTGG GAACGCAGTC GGGTGCCATG GTGGGTTTTC

*Figure 27 AC*

WO 02/022080

PCT/US01/28861

27451 GTGGATTTTA AGGAGCCAGC CTGTAATGTT ACATTCGCAG CTGAAGCTAA  
CACCTAAAAT TCCTCGGTCTG GACATTACAA TGTAAGCGTC GACTTCGATT

27501 TGAGTGCACC ACTCTTATAA AATGCACCAC AGAACATGAA AAGCTGCTTA  
ACTCACGTGG TGAGAATATT TTACGTGGTG TCTTGTACTT TTCGACGAAT

27551 TTCGCCACAA AAACAAAATT GGCAAGTATG CTGTTTATGC TATTTGGCAG  
AAGCGGTGTT TTTGTTTTAA CCGTTCATAC GACAAATACG ATAAACCGTC

27601 CCAGGTGACA CTACAGAGTA TAATGTTACA GTTTTCCAGG GTAAAAGTCA  
GGTCCACTGT GATGTCTCAT ATTACAATGT CAAAAGGTCC CATTTTCAGT

27651 TAAAACTTTT ATGTATACTT TTCCATTTTA TGAAATGTGC GACATTACCA  
ATTTTGAAAA TACATATGAA AAGGTAAAAT ACTTTACACG CTGTAATGGT

27701 TGTACATGAG CAAACAGTAT AAGTTGTGGC CCCACAAAA TTGTGTGGAA  
ACATGTACTC GTTTGTCATA TTCAACACCG GGGGTGTTTT AACACACCTT

27751 AACACTGGCA CTTTCTGCTG CACTGCTATG CTAATTACAG TGCTCGCTTT  
TTGTGACCGT GAAAGACGAC GTGACGATAC GATTAATGTC ACGAGCGAAA

27801 GGTCTGTACC CTACTCTATA TTAAATACAA AAGCAGACGC AGCTTTATTG  
CCAGACATGG GATGAGATAT AATTTATGTT TTCGTCTGCG TCGAAATAAC

27851 AGGAAAAGAA AATGCCTTAA TTTACTAAGT TACAAAGCTA ATGTCACCAC  
TCCTTTTCTT TTACGGAATT AAATGATTCA ATGTTTCGAT TACAGTGGTG

27901 TAACTGCTTT ACTCGCTGCT TGCAAAACAA ATTCAAAAAG TTAGCATTAT  
ATTGACGAAA TGAGCGACGA ACGTTTTGTT TAAGTTTTTC AATCGTAATA

27951 AATTAGAATA GGATTTAAAC CCCCCGGTCA TTTCTGCTC AATACCATT  
TTAATCTTAT CCTAAATTG GGGGGCCAGT AAAGGACGAG TTATGGTAAG

28001 CCCTGAACAA TTGACTCTAT GTGGGATATG CTCCAGCGCT ACAACCTTGA  
GGGACTTGTT AACTGAGATA CACCCTATAC GAGGTCGCGA TGTGGAAC

28051 AGTCAGGCTT CCTGGATGTC AGCATCTGAC TTTGGCCAGC ACCTGTCCCG  
TCAGTCCGAA GGACCTACAG TCGTAGACTG AAACCGGTCG TGGACAGGGC

28101 CGGATTTGTT CCAGTCCAAC TACAGCGACC CACCCTAACA GAGATGACCA  
GCCTAAACAA GGTCAGGTTG ATGTCGCTGG GTGGGATTGT CTCTACTGGT

28151 ACACAACCAA CGCGGCCGCC GCTACCGGAC TTACATCTAC CACAAATACA  
TGTGTTGGTT GCGCCGGCGG CGATGGCCTG AATGTAGATG GTGTTTATGT

28201 CCCCAGTTT CTGCCTTTGT CAATAACTGG GATAACTTGG GCATGTGGTG  
GGGGTTCAA GACGGAAACA GTTATTGACC CTATTGAACC CGTACACCAC

28251 GTTCTCCATA GCGCTTATGT TTGTATGCCT TATTATTATG TGGCTCATCT  
CAAGAGGTAT CGCGAATACA AACATACGGA ATAATAATAC ACCGAGTAGA

28301 GCTGCCTAAA GCGCAAACGC GCCCGACCAC CCATCTATAG TCCCATCATT  
CGACGGATTT CGCGTTTGCG CGGGCTGGTG GGTAGATATC AGGGTAGTAA

28351 GTGCTACACC CAAACAATGA TGAATCCAT AGATTGGACG GACTGAAACA  
CACGATGTGG GTTTGTTACT ACCTTAGGTA TCTAACCTGC CTGACTTTGT

Figure 27A D

28451 TTTTATATTA CTGACCCTTG TTGCGCTTTT TTGTGCGTGC TCCACATTGG  
AAAATATAAT GACTGGGAAC AACGCGAAAA AACACGCACG AGGTGTAACC

28501 CTGCGGTTTC TCACATCGAA GTAGACTGCA TTCCAGCCTT CACAGTCTAT  
GACGCCAAAG AGTGTAGCTT CATCTGACGT AAGGTCGGAA GTGTCAGATA

28551 TTGCTTTACG GATTTGTCAC CCTCACGCTC ATCTGCAGCC TCATCACTGT  
AACGAAATGC CTAAACAGTG GGAGTGCGAG TAGACGTCGG AGTAGTGACA

28601 GGTCATCGCC TTTATCCAGT GCATTGACTG GGTCTGTGTG CGCTTTGCAT  
CCAGTAGCGG AAATAGGTCA CGTAACTGAC CCAGACACAC GCGAAACGTA

28651 ATCTCAGACA CCATCCCCAG TACAGGGACA GGACTATAGC TGAGCTTCTT  
TAGAGTCTGT GGTAGGGGTC ATGTCCCTGT CCTGATATCG ACTCGAAGAA

28701 AGAATTCTTT AATTATGAAA TTTACTGTGA CTTTTCTGCT GATTATTTGC  
TCTTAAGAAA TTAATACTTT AAATGACACT GAAAAGACGA CTAATAACG

28751 ACCCTATCTG CGTTTTGTTC CCCGACCTCC AAGCCTCAAA GACATATATC  
TGGGATAGAC GCAAAACAAG GGGCTGGAGG TTCGGAGTTT CTGTATATAG

28801 ATGCAGATTC ACTCGTATAT GGAATATTCC AAGTTGCTAC AATGAAAAAA  
TACGTCTAAG TGAGCATATA CCTTATAAGG TTCAACGATG TTACTTTTTT

28851 GCGATCTTTC CGAAGCCTGG TTATATGCAA TCATCTCTGT TATGGTGTTC  
CGCTAGAAAG GCTTCGGACC AATATACGTT AGTAGAGACA ATACCACAAG

28901 TGCAGTACCA TCTTAGCCCT AGCTATATAT CCCTACCTTG ACATTGGCTG  
ACGTCATGGT AGAATCGGGA TCGATATATA GGGATGGAAC TGTAACCGAC

28951 GAACGCAATA GATGCCATGA ACCACCCAAC TTTCCCCGCG CCCGCTATGC  
CTTGCGTTAT CTACGGTACT TGGTGGGTTG AAAGGGGCGC GGGCGATACG

29001 TTCCACTGCA ACAAGTTGTT GCCGGCGGCT TTGTCCCAGC CAATCAGCCT  
AAGGTGACGT TGTTCAACAA CGGCCGCCGA AACAGGGTCG GTTAGTCGGA

29051 CGCCACCTT CTCCCACCCC CACTGAAATC AGCTACTTTA ATCTAACAGG  
GCGGGTGGAA GAGGGTGGGG GTGACTTTAG TCGATGAAAT TAGATTGTCC

29101 AGGAGATGAC TGACACCCTA GATCTAGAAA TGGACGGAAT TATTACAGAG  
TCCTCTACTG ACTGTGGGAT CTAGATCTTT ACCTGCCTTA ATAATGTCTC

29151 CAGCGCCTGC TAGAAAGACG CAGGGCAGCG GCCGAGCAAC AGCGCATGAA  
GTCGCGGACG ATCTTTCTGC GTCCCGTCGC CGGCTCGTTG TCGCGTACTT

29201 TCAAGAGCTC CAAGACATGG TTAAC TTGCA CCAGTGCAAA AGGGGTATCT  
AGTTCTCGAG GTTCTGTACC AATTGAACGT GGTCACGTTT TCCCCATAGA

29251 TTTGTCTCGT AAAGCAGGCC AAAGTCACCT ACGACAGTAA TACCACCGGA  
AAACAGAGCA TTTCGTCCGG TTTCAGTGGA TGCTGTCATT ATGGTGGCCT

29301 CACCGCCTTA GCTACAAGTT GCCAACCAAG CGTCAGAAAT TGGTGGTCAT  
GTGGCGGAAT CGATGTTCAA CGGTTGGTTC GCAGTCTTTA ACCACCAGTA

Figure 27 A E

29401 GCTGCATTCA CTCACCTTGT CAAGGACCTG AGGATCTCTG CACCCTTATT  
CGACGTAAGT GAGTGGAACA GTTCCTGGAC TCCTAGAGAC GTGGGAATAA

29451 AAGACCCTGT GCGGTCTCAA AGATCTTATT CCCTTTAACT AATAAAAAAA  
TTCTGGGACA CGCCAGAGTT TCTAGAATAA GGGAAATTGA TTATTTTTTT

29501 AATAATAAAG CATCACTTAC TTAAAATCAG TTAGCAAATT TCTGTCCAGT  
TTATTATTTC GTAGTGAATG AATTTTAGTC AATCGTTTAA AGACAGGTCA

29551 TTATTCAGCA GCACCTCCTT GCCCTCCTCC CAGCTCTGGT ATTGCAGCTT  
AATAAGTCGT CGTGGAGGAA CGGGAGGAGG GTCGAGACCA TAACGTCGAA

29601 CCTCCTGGCT GCAAACCTTC TCCACAATCT AAATGGAATG TCAGTTTCCT  
GGAGGACCGA CGTTTGAAAG AGGTGTTAGA TTTACCTTAC AGTCAAAGGA

29651 CCTGTTCTCTG TCCATCCGCA CCCACTATCT TCATGTTGTT GCAGATGAAG  
GGACAAGGAC AGGTAGGCGT GGGTGATAGA AGTACAACAA CGTCTACTTC

29701 CGCGCAAGAC CGTCTGAAGA TACCTTCAAC CCCGTGTATC CATATGACAC  
GCGCGTTCTG GCAGACTTCT ATGGAAGTTG GGGCACATAG GTATACTGTG

29751 GGAAACCGGT CCTCCAACCTG TGCCTTTTCT TACTCCTCCC TTTGTATCCC  
CCTTTGGCCA GGAGGTTGAC ACGGAAAAGA ATGAGGAGGG AACATAGGG

29801 CCAATGGGTT TCAAGAGAGT CCCCCTGGGG TACTCTCTTT GCGCCTATCC  
GGTTACCCAA AGTTCTCTCA GGGGGACCCC ATGAGAGAAA CGCGGATAGG

29851 GAACCTCTAG TTACCTCCAA TGGCATGCTT GCGCTCAAAA TGGGCAACGG  
CTTGAGATC AATGGAGGTT ACCGTACGAA CGCGAGTTTT ACCCGTTGCC

29901 CCTCTCTCTG GACGAGGCCG GCAACCTTAC CTCCCAAAT GTAACCACTG  
GGAGAGAGAC CTGCTCCGGC CGTTGGAATG GAGGGTTTTA CATTGGTGAC

29951 TGAGCCCACC TCTCAAAAAA ACCAAGTCAA ACATAAACCT GGAAATATCT  
ACTCGGGTGG AGAGTTTTTT TGGTTCAGTT TGTATTTGGA CCTTTATAGA

30001 GCACCCCTCA CAGTTACCTC AGAAGCCCTA ACTGTGGCTG CCGCCGCACC  
CGTGGGGAGT GTCAATGGAG TCTTCGGGAT TGACACCGAC GCGGGCGTGG

30051 TCTAATGGTC GCGGGCAACA CACTCACCAT GCAATCACAG GCCCCGCTAA  
AGATTACCAG CGCCCGTTGT GTGAGTGGTA CGTTAGTGTC CGGGGCGATT

30101 CCGTGCACGA CTCCAAACTT AGCATTGCCA CCAAGGACC CCTCACAGTG  
GGCACGTGCT GAGGTTTGAA TCGTAACGGT GGGTTCCTGG GGAGTGTCAC

30151 TCAGAAGGAA AGCTAGCCCT GCAAACATCA GGCCCCCTCA CCACCACCGA  
AGTCTTCCTT TCGATCGGGA CGTTTGTAGT CCGGGGGAGT GGTGGTGGCT

30201 TAGCAGTACC CTTACTATCA CTGCCTCACC CCCTCTAACT ACTGCCACTG  
ATCGTCATGG GAATGATAGT GACGGAGTGG GGGAGATTGA TGACGGTGAC

30251 GTAGCTTGGG CATTGACTTG AAAGAGCCCA TTTATACACA AAATGGAAAA  
CATCGAACCC GTAACCTGAAC TTTCTCGGGT AAATATGTGT TTTACCTTTT

Figure 27 AF

30351 TTTGACCGTA GCAACTGGTC CAGGTGTGAC TATTAATAAT ACTTCCTTGC  
AAACTGGCAT CGTTGACCAG GTCCACACTG ATAATTATTA TGAAGGAACG

30401 AAATAAAGT TACTGGAGCC TTGGGTTTTG ATTCACAAGG CAATATGCAA  
TTTGATTTC A TGACCTCGG AACCCAAAAC TAAGTGTTCC GTTATACGTT

30451 CTTAATGTAG CAGGAGGACT AAGGATTGAT TCTCAAACA GACGCCTTAT  
GAATTACATC GTCCTCCTGA TTCCTAACTA AGAGTTTTGT CTGCGGAATA

30501 ACTTGATGTT AGTTATCCGT TTGATGCTCA AAACCAACTA AATCTAAGAC  
TGAATAACA TCAATAGGCA AACTACGAGT TTTGGTTGAT TTAGATTCTG

30551 TAGGACAGGG CCCTCTTTTT ATAACTCAG CCCACAACCT GGATATTAAC  
ATCCTGTCCC GGGAGAAAAA TATTTGAGTC GGGTGTTGAA CCTATAATTG

30601 TACAACAAG GCCTTTACTT GTTACAGCT TCAAACAATT CAAAAAGCT  
ATGTTGTTTC CGGAAATGAA CAAATGTCGA AGTTTGTTAA GGTTTTTCGA

30651 TGAGGTTAAC CTAAGCACTG CCAAGGGGTT GATGTTTGAC GCTACAGCCA  
ACTCCAATTG GATTCGTGAC GGTTCGCCAA CTACAACTG CGATGTCGGT

30701 TAGCCATTAA TGCAGGAGAT GGGCTTGAAT TTGGTTCACC TAATGCACCA  
ATCGGTAATT ACGTCCTCTA CCCGAACCTA AACCAAGTGG ATTACGTGGT

30751 AACACAAATC CCCTCAAAAC AAAAATTGGC CATGGCCTAG AATTTGATTC  
TTGTGTTTAG GGGAGTTTTG TTTTAAACCG GTACCGGATC TTAACTAAG

30801 AAACAAGGCT ATGGTTCCTA AACTAGGAAC TGGCCTTAGT TTTGACAGCA  
TTTGTTCCGA TACCAAGGAT TTGATCCTTG ACCGGAATCA AACTGTCTG

30851 CAGGTGCCAT TACAGTAGGA AACAAAAATA ATGATAAGCT AACTTTGTGG  
GTCCACGGTA ATGTCATCCT TTGTTTTTAT TACTATTCGA TTGAAACACC

30901 ACCACACCAG CTCCATCTCC TAACTGTAGA CTAAATGCAG AGAAAGATGC  
TGGTGTGGTC GAGGTAGAGG ATTGACATCT GATTACGTC TCTTTCTACG

30951 TAAACTCACT TTGGTCTTAA CAAAATGTGG CAGTCAAATA CTTGCTACAG  
ATTTGAGTGA AACCAGAATT GTTTTACACC GTCAGTTTAT GAACGATGTC

31001 TTTCAGTTTT GGCTGTTAAA GGCAGTTTGG CTCCAATATC TGGAACAGTT  
AAAGTCAAAA CCGACAATTT CCGTCAAACC GAGGTTATAG ACCTTGTCAG

31051 CAAAGTGCTC ATCTTATTAT AAGATTTGAC GAAAATGGAG TGCTACTAAA  
GTTTCACGAG TAGAATAATA TTCTAACTG CTTTACCTC ACGATGATTT

31101 CAATTCCTTC CTGGACCCAG AATATTGGAA CTTTAGAAAT GGAGATCTTA  
GTTAAGGAAG GACCTGGGTC TTATAACCTT GAAATCTTTA CCTCTAGAAT

31151 CTGAAGGCAC AGCCTATACA AACGCTGTTG GATTTATGCC TAACCTATCA  
GACTTCCGTG TCGGATATGT TTGCGACAAC CTAAATACGG ATTGGATAGT

31201 GCTTATCCAA AATCTCACGG TAAACTGCC AAAAGTAACA TTGTCAGTCA  
CGAATAGGTT TTAGAGTGCC ATTTTGACGG TTTTCATTGT AACAGTCAGT

Figure 27 AG

WO 02/022080

PCT/US01/28861

```

31251 AGTTTACTTA AACGGAGACA AACTAAACC TGTAACACTA ACCATTACAC
      TCAAATGAAT TTGCCTCTGT TTTGATTTGG ACATTGTGAT TGGTAATGTG

31301 TAAACGGTAC ACAGGAAACA GGAGACACAA CTCCAAGTGC ATACTCTATG
      ATTTGCCATG TGTCCCTTTGT CCTCTGTGTT GAGG TTCACG TATGAGATAC

31351 TCATTTTTCAT GGGACTGGTC TGGCCACAAC TACATTAATG AAATATTTGC
      AGTAAAAGTA CCCTGACCAG ACCGGTGTG ATGTAATTAC TTTATAAACC

31401 CACATCCTCT TACACTTTTT CATAATTGC CCAAGAATAA AGAATCGTTT
      GTGTAGGAGA ATGTGAAAAA GTATGTAACG GGTTCCTATT TCTTAGCAAA

31451 GTGTTATGTT TCAACGTGTT TATTTTTCAG TTGCAGAAAA TTTCAAGTCA
      CACAATACAA AGTTGCACAA ATAAAAAGTT AACGTCTTTT AAAGTTCAGT

31501 TTTTTCATTC AGTAGTATAG CCCACCACC ACATAGCTTA TACAGATCAC
      AAAAAGTAAG TCATCATATC GGGGTGGTGG TGTATCGAAT ATGTCTAGTG

31551 CGTACCTTAA TCAAACCTAC AGAACCTAG TATTCAACCT GCCACCTCCC
      GCATGGAATT AGTTTGAGTG TCTTGGGATC ATAAGTTGGA CGGTGGAGGG

31601 TCCCAACACA CAGAGTACAC AGTCCTTTCT CCCCGGCTGG CCTTAAAAAG
      AGGGTTGTGT GTCTCATGTG TCAGGAAAGA GGGGCCGACC GGAATTTTTC

31651 CATCATATCA TGGGTAACAG ACATATTCTT AGGTGTTATA TTCCACACGG
      GTAGTATAGT ACCCATTGTC TGTATAAGAA TCCACAATAT AAGGTGTGCC

31701 TTTCTGTGCG AGCCAAACGC TCATCAGTGA TATTAATAAA CTCCCCGGGC
      AAAGGACAGC TCGGTTTGCG AGTAGTCACT ATAATTATTT GAGGGGCCCG

31751 AGCTCACTTA AGTTCATGTC GCTGTCCAGC TGCTGAGCCA CAGGCTGCTG
      TCGAGTGAAT TCAAGTACAG CGACAGGTCG ACGACTCGGT GTCCGACGAC

31801 TCCAACTTGC GGTGCTTAA CGGGCGGCGA AGGAGAAGTC CACGCCTACA
      AGGTTGAACG CCAACGAATT GCCCGCCGCT TCCTCTTCAG GTGCGGATGT

31851 TGGGGGTAGA GTCATAATCG TGCATCAGGA TAGGGCGGTG GTGCTGCAGC
      ACCCCCATCT CAGTATTAGC ACGTAGTCCT ATCCCGCCAC CACGACGTGC

31901 AGCGCGCGAA TAAACTGCTG CCGCCGCCGC TCCGTCCTGC AGGAATACAA
      TCGCGCGCTT ATTTGACGAC GCGGCGGCG AGGCAGGACG TCCTTATGTT

31951 CATGGCAGTG GTCTCCTCAG CGATGATTCG CACCGCCCGC AGCATAAGGC
      GTACCGTCAC CAGAGGAGTC GCTACTAAGC GTGGCGGGCG TCGTATTCCG

32001 GCCTTGTCCT CCGGGCACAG CAGCGCACCC TGATCTCACT TAAATCAGCA
      CGGAACAGGA GGCCCGTGTC GTCGCGTGGG ACTAGAGTGA ATTTAGTCGT

32051 CAGTAACTGC AGCACAGCAC CACAATATTG TTCAAATCC CACAGTGCAA
      GTCATTGACG TCGTGTCGTG GTGTTATAAC AAGTTT TAGG GTGTCACGTT

32101 GGCGCTGTAT CCAAAGCTCA TGGCGGGGAC CACAGAACCC ACGTGGCCAT
      CCGCGACATA GGTTCGAGT ACCGCCCTG GTGTCTTGGG TGCACCGGTA

32151 CATACCACAA GCGCAGGTAG ATTAAGTGGC GACCCCTCAT AAACACGCTG
      GTATGGTGTT CGCGTCCATC TAATTCACCG CTGGGGAGTA TTTGTGCGAC

```

Figure 27AH

32251 CCATATAAAC CTCTGATTAA ACATGGCGCC ATCCACCACC ATCCTAAACC  
GGTATATTTG GAGACTAATT TGTACCGCGG TAGGTGGTGG TAGGATTTGG

32301 AGCTGGCCAA AACCTGCCCC CCGGCTATAC ACTGCAGGGA ACCGGGACTG  
TCGACCGGTT TTGGACGGGC GGCCGATATG TGACGTCCCT TGGCCCTGAC

32351 GAACAATGAC AGTGGAGAGC CCAGGACTCG TAACCATGGA TCATCATGCT  
CTTGTTACTG TCACCTCTCG GGTCTGAGC ATTGGTACCT AGTAGTACGA

32401 CGTCATGATA TCAATGTTGG CACAACACAG GCACACGTGC ATACACTTCC  
GCAGTACTAT AGTTACAACC GTGTTGTGTC CGTGTGCACG TATGTGAAGG

32451 TCAGGATTAC AAGCTCCTCC CGCGTTAGAA CCATATCCCA GGGAACAACC  
AGTCCTAATG TTCGAGGAGG GCGCAATCTT GGTATAGGGT CCCTTGTTGG

32501 CATTCCTGAA TCAGCGTAAA TCCCACACTG CAGGGAAGAC CTCGCACGTA  
GTAAGGACTT AGTCGCATTT AGGGTGTGAC GTCCCTTCTG GAGCGTGCAT

32551 ACTCACGTTG TGCATTGTCA AAGTGTTACA TTCGGGCAGC AGCGGATGAT  
TGAGTGCAAC ACGTAACAGT TTCACAATGT AAGCCCGTCG TCGCCTACTA

32601 CCTCCAGTAT GGTAGCGCGG GTTTCTGTCT CAAAAGGAGG TAGACGATCC  
GGAGGTCATA CCATCGCGCC CAAAGACAGA GTTTTCCTCC ATCTGCTAGG

32651 CTA CTGTACG GAGTGCGCC AGACAACCGA GATCGTGTTG GTCGTAGTGT  
GATGACATGC CTCACGCGGC TCTGTTGGCT CTAGCACAA CAGCATCACA

32701 CATGCCAAAT GGAACGCCGG ACGTAGTCAT ATTCCTGAA GCAAACCAG  
GTACGGTTTA CCTTGCGGCC TGCATCAGTA TAAAGGACTT CGTTTTGGTC

32751 GTGCGGGCGT GACAAACAGA TCTGCGTCTC CGGTCTCGCC GCTTAGATCG  
CACGCCCGCA CTGTTTGTCT AGACGCAGAG GCCAGAGCGG CGAATCTAGC

32801 CTCTGTGTAG TAGTTGTAGT ATATCCACTC TCTCAAAGCA TCCAGGCGCC  
GAGACACATC ATCAACATCA TATAGGTGAG AGAGTTTCGT AGGTCCGCGG

32851 CCCTGGCTTC GGGTTCTATG TAACTCCTT CATGCGCCGC TGCCCTGATA  
GGGACCGAAG CCCAAGATAC ATTTGAGGAA GTACGCGGCG ACGGGACTAT

32901 ACATCCACCA CCGCAGAATA AGCCACACCC AGCCAACCTA CACATTCGTT  
TG TAGGTGGT GCGTCTTAT TCGGTGTGGG TCGGTTGGAT GTGTAAGCAA

32951 CTGCGAGTCA CACACGGGAG GAGCGGGAAG AGCTGGAAGA ACCATGTTTT  
GACGCTCAGT GTGTGCCCTC CTCGCCCTT TCGACCTTCT TGGTACAAAA

33001 TTTTTTTATT CCAAAGATT ATCCAAACC TCAAATGAA GATCTATTAA  
AAAAAATAA GGTTTTCTAA TAGGTTTTGG AGTTTTACTT CTAGATAATT

33051 GTGAACGCGC TCCCCTCCGG TGGCGTGGTC AA ACTCTACA GCCAAAGAAC  
CACTTGCGCG AGGGGAGGCC ACCGCACCAG TTTGAGATGT CGGTTTCTTG

33101 AGATAATGGC ATTTGTAAGA TGTTGCACAA TGGCTTCCAA AAGGCAAACG  
TCTATTACCG TAAACATTCT ACAACGTGTT ACCGAAGGTT TTCCGTTTGC

Figure 27 AI

33201 CTCTATAAAC ATTCCAGCAC CTTCAACCAT GCCCAAATAA TTCTCATCTC  
GAGATATTTG TAAGGTCGTG GAAGTTGGTA CGGGTTTATT AAGAGTAGAG

33251 GCCACCTTCT CAATATATCT CTAAGCAAAT CCCGAATATT AAGTCCGGCC  
CGGTGGAAGA GTTATATAGA GATTCGTTTA GGGCTTATAA TTCAGGCCGG

33301 ATTGTAAAAA TCTGCTCCAG AGCGCCCTCC ACCTTCAGCC TCAAGCAGCG  
TAACATTTTT AGACGAGGTC TCGCGGGAGG TGGAAAGTCGG AGTTCGTCGC

33351 AATCATGATT GCAAAAATTC AGGTTCCCTCA CAGACCTGTA TAAGATTCAA  
TTAGTACTAA CGTTTTTAAG TCCAAGGAGT GTCTGGACAT ATTCTAAGTT

33401 AAGCGGAACA TTAACAAAAA TACCGCGATC CCGTAGGTCC CTTCGCAGGG  
TTCGCCTTGT AATTGTTTTT ATGGCGCTAG GGCATCCAGG GAAGCGTCCC

33451 CCAGCTGAAC ATAATCGTGC AGGTCTGCAC GGACCAGCGC GGCCACTTCC  
GGTCGACTTG TATTAGCAGC TCCAGACGTG CCTGGTCGCG CCGGTGAAGG

33501 CCGCCAGGAA CCATGACAAA AGAACCACACA CTGATTATGA CACGCATACT  
GGCGGTCCTT GGTACTGTTT TCTTGGGTGT GACTAATACT GTGCGTATGA

33551 CGGAGCTATG CTAACCAGCG TAGCCCCGAT GTAAGCTTGT TGCATGGGCG  
GCCTCGATAC GATTGGTCGC ATCGGGGCTA CATTCAACA ACGTACCCGC

33601 GCGATATAAA ATGCAAGGTG CTGCTCAAAA AATCAGGCAA AGCCTCGCGC  
CGCTATATTT TACGTTCCAC GACGAGTTTT TTAGTCCGTT TCGGAGCGCG

33651 AAAAAAGAAA GCACATCGTA GTCATGCTCA TGCAGATAAA GGCAGGTAAG  
TTTTTCTTT CGTGTAGCAT CAGTACGAGT ACGTCTATTT CCGTCCATTC

33701 CTCCGGAACC ACCACAGAAA AAGACACCAT TTTTCTCTCA AACATGTCTG  
GAGGCCTTGG TGGTGTCTTT TTCTGTGGTA AAAAGAGAGT TTGTACAGAC

33751 CGGGTTTCTG CATAAACACA AAATAAAATA ACAAAAAAAC ATTTAAACAT  
GCCCAAAGAC GTATTTGTGT TTTATTTTAT TGTTTTTTTG TAAATTTGTA

33801 TAGAAGCCTG TCTTACAACA GGAAAAACAA CCCTTATAAG CATAAGACGG  
ATCTTCGGAC AGAATGTTGT CCTTTTTGTT GGAATATTC GTATTCTGCC

33851 ACTACGGCCA TGCCGGCGTG ACCGTAAAAA AACTGGTCAC CGTGATTAAA  
TGATGCCGGT ACGGCCGCAC TGGCATTTTT TTGACCAGTG GCACTAATTT

33901 AAGCACCACC GACAGCTCCT CGGTCATGTC CGGAGTCATA ATGTAAGACT  
TTCGTGGTGG CTGTCGAGGA GCCAGTACAG GCCTCAGTAT TACATTCTGA

33951 CGGTAAACAC ATCAGGTTGA TTCACATCGG TCAGTGCTAA AAAGCGACCG  
GCCATTTGTG TAGTCCAAC T AAGTGTAGCC AGTCACGATT TTTCGCTGGC

34001 AAATAGCCCG GGGGAATACA TACCCGCAGG CGTAGAGACA ACATTACAGC  
TTTATCGGGC CCCCTTATGT ATGGGCGTCC GCATCTCTGT TGTAATGTGC

34051 CCCCATAGGA GGTATAACAA AATTAATAGG AGAGAAAAAC ACATAAACAC  
GGGGTATCCT CCATATTGTT TTAATTATCC TCTCTTTTGT TGTATTTGTG

Figure 27A J

34151 ACATACAGCG CTTCCACAGC GGCAGCCATA ACAGTCAGCC TTACCAGTAA  
TGTATGTCGC GAAGGTGTCG CCGTCGGTAT TGTCAGTCGG AATGGTCATT

34201 AAAAGAAAAC CTATTAAAA AACACCACTC GACACGGCAC CAGCTCAATC  
TTTTCTTTTG GATAATTTTT TTGTGGTGAG CTGTGCCGTG GTCGAGTTAG

34251 AGTCACAGTG TAAAAAAGGG CCAAGTGCAG AGCGAGTATA TATAGGACTA  
TCAGTGTCAC ATTTTTTCCC GGTTACAGTC TCGCTCATAT ATATCCTGAT

34301 AAAAATGACG TAACGGTTAA AGTCCACAAA AAACACCCAG AAAACCGCAC  
TTTTTACTGC ATTGCCAATT TCAGGTGTTT TTTGTGGGTC TTTTGGCGTG

34351 GCGAACCTAC GCCCAGAAAC GAAAGCCAAA AAACCCACAA CTTCTCTCAA  
CGCTTGGATG CGGGTCTTTG CTTTCGGTTT TTTGGGTGTT GAAGGAGTTT

34401 TCGTCACTTC CGTTTTCCCA CGTTACGTCA CTTCCCATTT TAAGAAAAC  
AGCAGTGAAG GCAAAAGGGT GCAATGCAGT GAAGGGTAAA ATTCTTTTGA

34451 ACAATTCCCA ACACATACAA GTTACTCCGC CCTAAAACCT ACGTCACCCG  
TGTTAAGGGT TGTGTATGTT CAATGAGGCG GGATTTTGGA TGCAGTGGGC

34501 CCCC GTTCCC ACGCCCCGCG CCACGTCACA AACTCCACCC CCTCATTATC  
GGGGCAAGGG TCGGGGGCGC GGTGCAGTGT TTGAGGTGGG GGAGTAATAG

PacI  
-----

34551 ATATTGGCTT CAATCCAAAA TAAGGTATAT TATTGATGAT GTTAATTAAG  
TATAACCGAA GTTAGGTTTT ATTCCATATA ATAAC TACTA CAATTAATTC

34601 AATTCGGATC TGCGACGCGA GGCTGGATGG CCTTCCCAT TATGATTCTT  
TTAAGCCTAG ACGCTGCGCT CCGACCTACC GGAAGGGGTA ATACTAAGAA

34651 CTCGCTTCCG GCGGCATCGG GATGCCCCGCG TTGCAGGCCA TGCTGTCCAG  
GAGCGAAGGC CGCCGTAGCC CTACGGGCGC AACGTCCGGT ACGACAGGTC

34701 GCAGGTAGAT GACGACCATC AGGGACAGCT TCAAGGCCAG CAAAAGGCCA  
CGTCCATCTA CTGCTGGTAG TCCCTGTCGA AGTTCCGGTC GTTTTCCGGT

34751 GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC  
CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG

34801 CCTGACGAGC ATCACAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC  
GGACTGCTCG TAGTGTTTTT AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG

34851 GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC  
CTGTCCTGAT ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG

34901 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC  
CGAGAGGACA AGGCTGGGAC GCGGAATGGC CTATGGACAG GCGGAAAGAG

34951 CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG  
GGAAGCCCTT CGCACC GCGA AAGAGTATCG AGTGCGACAT CCATAGAGTC

35001 TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG  
AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC

Figure 27 AK

AAGTCGGGCT GCGGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG

35101 CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT  
GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTGAC CATTGTCCTA

35151 TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC  
ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG

35201 CTAACTACGG CTACACTAGA AGGACAGTAT TTGGTATCTG CGCTCTGCTG  
GATTGATGCC GATGTGATCT TCCTGTCATA AACCATAGAC GCGAGACGAC

35251 AAGCCAGTTA CCTTCGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA  
TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT

35301 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC  
TTGGTGCGA CCATCGCCAC CAAAAAACA AACGTTTCGTC GTCTAATGCG

35351 GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT  
CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA ACTAGAAAAG ATGCCCCAGA

35401 GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT  
CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA

35451 ATCAAAAAGG ATCTTCACCT AGATCCTTTT AAATCAATCT AAAGTATATA  
TAGTTTTTCC TAGAAGTGGA TCTAGGAAA TTTAGTTAGA TTTCATATAT

35501 TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA  
ACTCATTTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA CTCCGTGGAT

35551 TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC  
AGAGTCGCTA GACAGATAAA GCAAGTAGGT ATCAACGGAC TGAGGGGCAG

35601 GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC  
CACATCTATT GATGCTATGC CCTCCC GAAT GGTAGACCGG GGTCACGACG

35651 AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA  
TTACTATGGC GCTCTGGGTG CGAGTGGCCG AGGTCTAAAT AGTCGTTATT

35701 ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC  
TGGTCGGTCG GCCTTCCCGG CTCGCGTCTT CACCAGGACG TTGAAATAGG

35751 GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC  
CGGAGGTAGG TCAGATAATT AACAACGGCC CTTGATCTC ATTCATCAAG

35801 GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG  
CGGTCAATTA TCAAACGCGT TGCAACAACG GTAACGATGT CCGTAGCACC

35851 TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA  
ACAGTGCGAG CAGCAAACCA TACCGAAGTA AGTCGAGGCC AAGGGTTGCT

35901 TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAG CGGTTAGCTC  
AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTTC GCCAATCGAG

35951 CTTGCGTCCT CCGATCGTTG TCAGAAGTAA GTTGCCGCA GTGTTATCAC  
GAAGCCAGGA GGCTAGCAAC AGTCTTCATT CAACCGGCGT CACAATAGTG

Figure 2 AL

36051 AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA  
TCTACGAAAA GACACTGACC ACTCATGAGT TGGTTCAGTA AGACTCTTAT

36101 GTGTATGCGG CGACCGAGTT GCTCTTGCCC GCGTCAACA CGGGATAATA  
CACATACGCC GCTGGCTCAA CGAGAACGGG CCGCAGTTGT GCCCTATTAT

36151 CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT  
GGCGCGGTGT ATCGTCTTGA AATTTTCACG AGTAGTAACC TTTTGCAAGA

36201 TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT  
AGCCCCGCTT TTGAGAGTTC CTAGAATGGC GACAACTCTA GGTCAAGCTA

36251 GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA  
CATTGGGTGA GCACGTGGGT TGA CTAGTAAG TCGTAGAAAA TGAAAGTGGT

36301 GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA  
CGCAAAGACC CACTCGTTTT TGTCTTCCG TTTTACGGCG TTTTTCCT

36351 ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTCAATA  
TATTCCCGCT GTGCCTTTAC AACTTATGAG TATGAGAAGG AAAAAGTTAT

36401 TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG  
AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC

36451 AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA  
TTACATAAAT CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT

36501 AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA CATTAAACCTA  
TTTCACGGTG GACTGCAGAT TCTTTGGTAA TAATAGTACT GTAATTGGAT

36551 TAAAAATAGG CGTATCACGA GGCCCTTTTCG TCTTCAAGAA TTGGATCCGA  
ATTTTTATCC GCATAGTGCT CCGGGAAAGC AGAAGTTCTT AACCTAGGCT

PacI

~~~~~

36601 ATTCTTAATT TCTTAATTAA (SEQ ID NO:34)
TAAGAATTAA AGAATTAATT (SEQ ID NO:35)

Figure 27A M

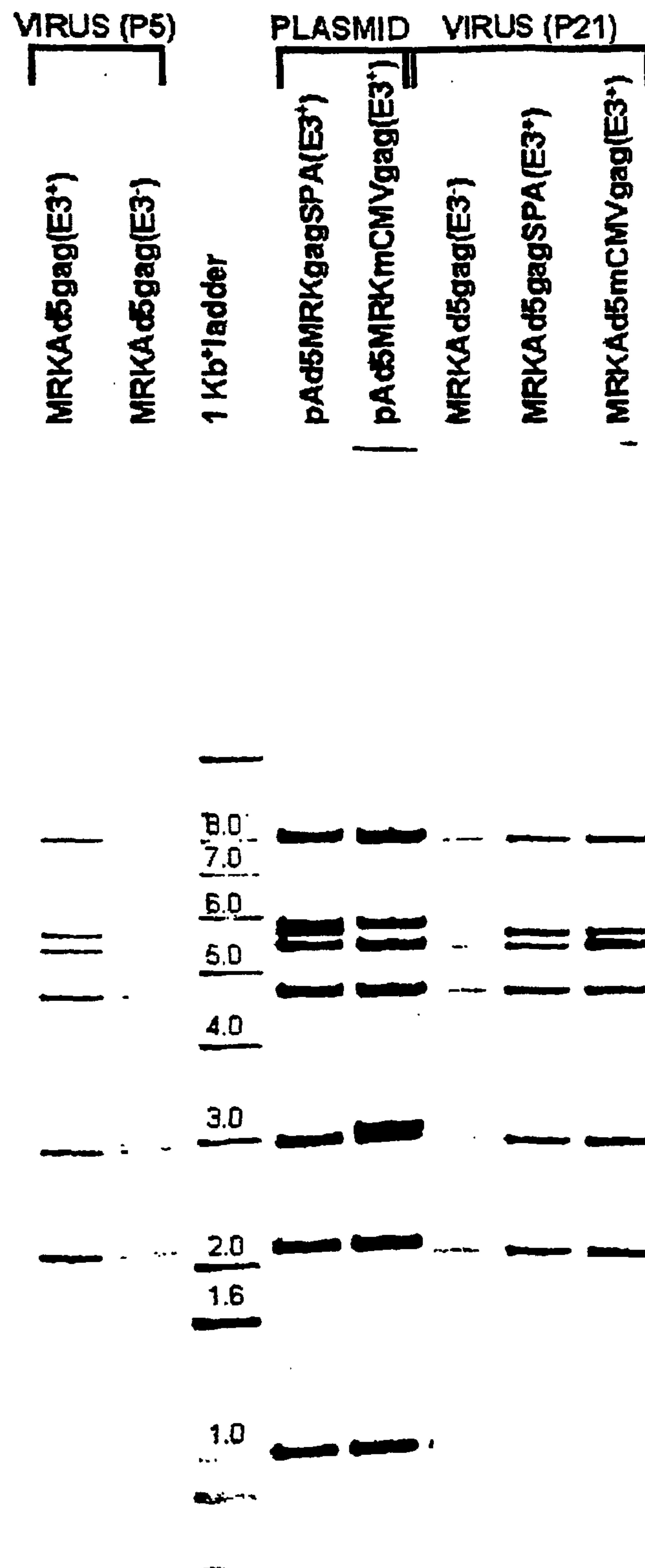


FIGURE 28

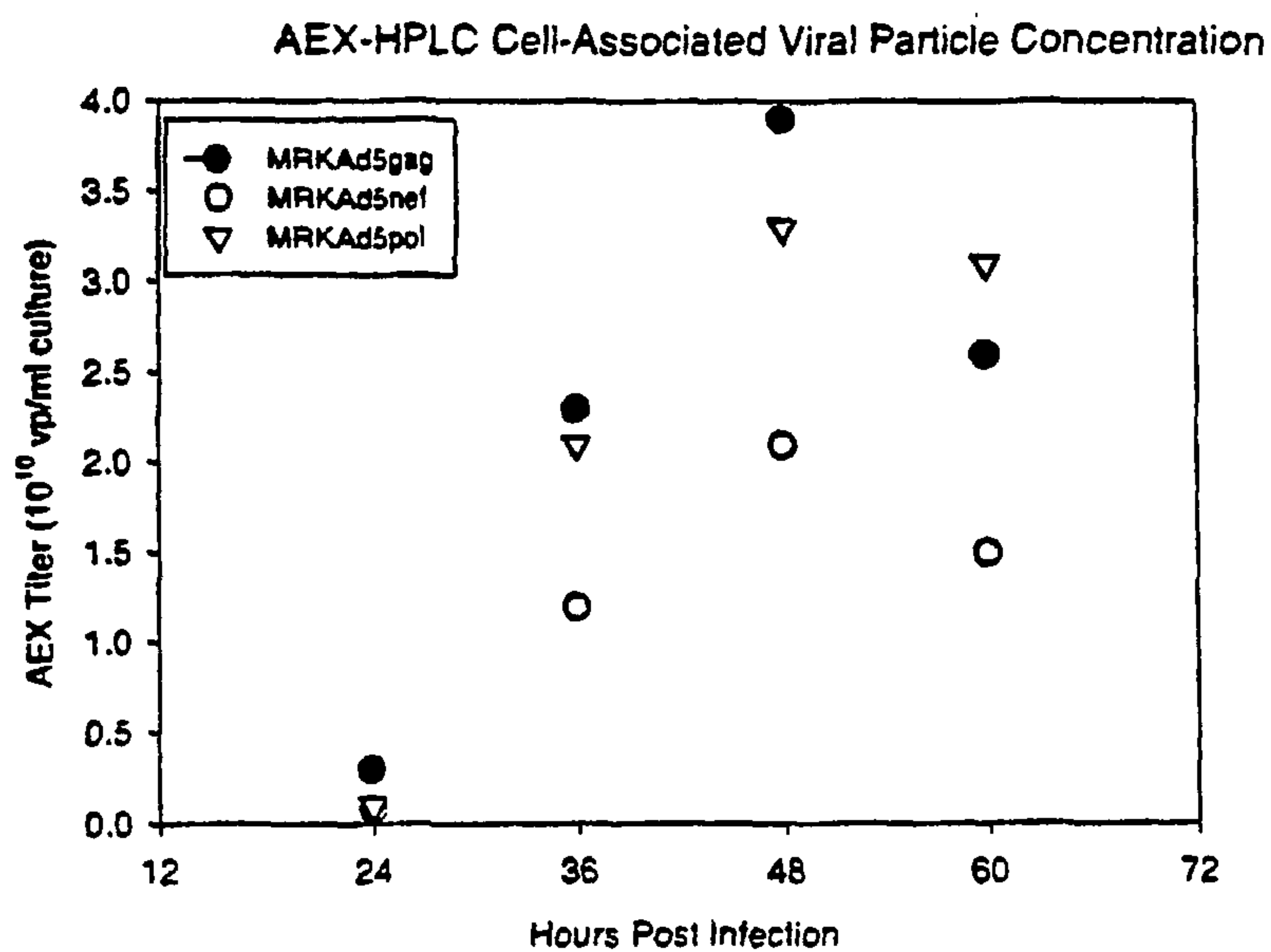


FIGURE 29A

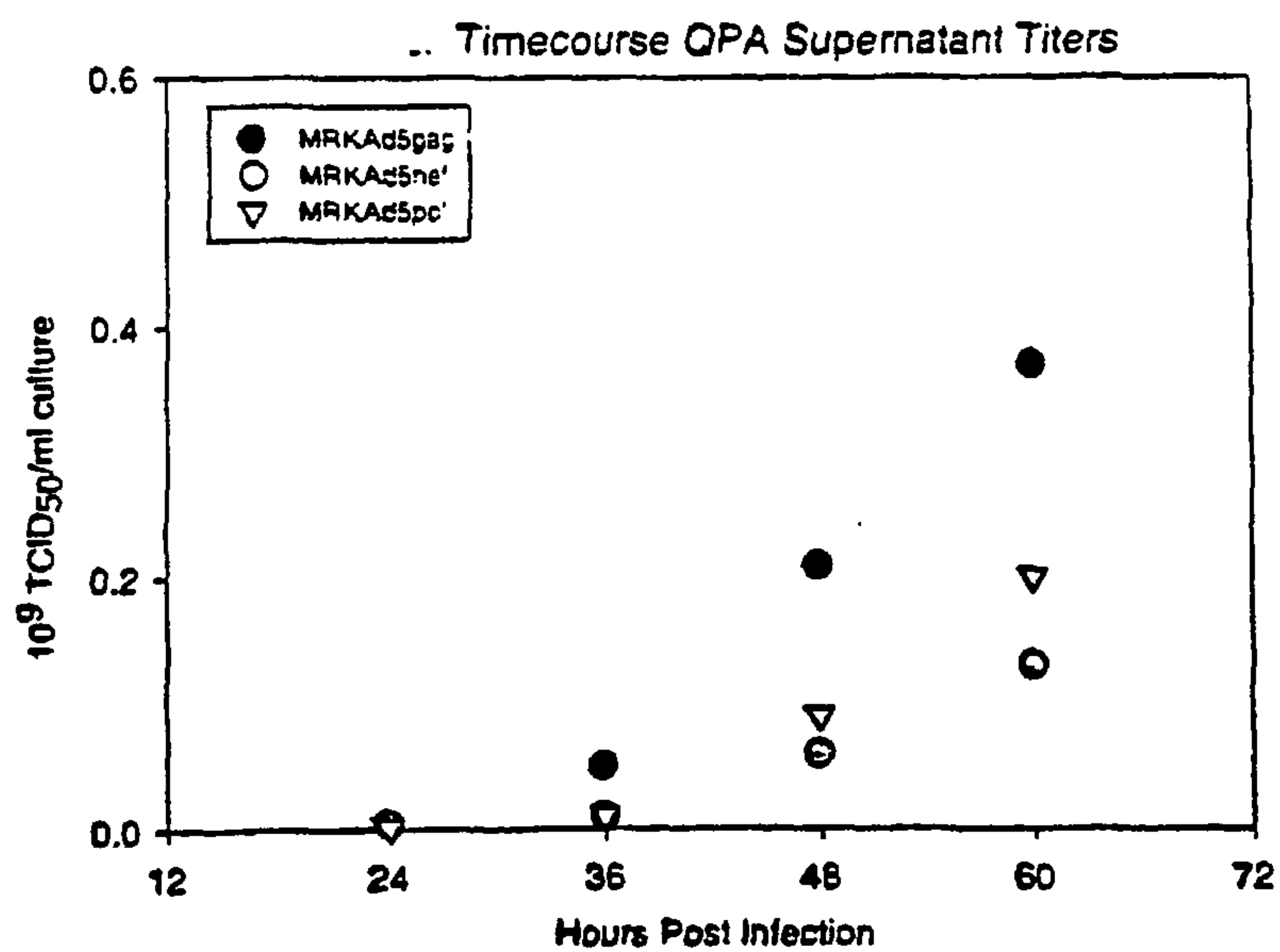


FIGURE 29B

| | |
|---|-----|
| atg gat gca atg aag aga ggg ctc tgc tgt gtg ctg ctg ctg tgt gga
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15 | 48 |
| gca gtc ttc gtt tgc ccc agc gag atc tcc att gtg tgg gcc tcc agg
Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ile Val Trp Ala Ser Arg
20 25 30 | 96 |
| gag ctg gag agg ttt gct gtg aac cct ggc ctg ctg gag acc tct gag
Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser Glu
35 40 45 | 144 |
| ggg tgc agg cag atc ctg ggc cag ctc cag ccc tcc ctg caa aca ggc
Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly
50 55 60 | 192 |
| tct gag gag ctg agg tcc ctg tac aac aca gtg gct acc ctg tac tgt
Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys
65 70 75 80 | 240 |
| gtg cac cag aag att gat gtg aag gac acc aag gag gcc ctg gag aag
Val His Gln Lys Ile Asp Val Lys Asp Thr Lys Glu Ala Leu Glu Lys
85 90 95 | 288 |
| att gag gag gag cag aac aag tcc aag aag aag gcc cag cag gct gct
Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala
100 105 110 | 336 |
| gct ggc aca ggc aac tcc agc cag gtg tcc cag aac tac ccc att gtg
Ala Gly Thr Gly Asn Ser Ser Gln Val Ser Gln Asn Tyr Pro Ile Val
115 120 125 | 384 |
| cag aac ctc cag ggc cag atg gtg cac cag gcc atc tcc ccc cgg acc
Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Ile Ser Pro Arg Thr
130 135 140 | 432 |
| ctg aat gcc tgg gtg aag gtg gtg gag gag aag gcc ttc tcc cct gag
Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala Phe Ser Pro Glu
145 150 155 160 | 480 |
| gtg atc ccc atg ttc tct gcc ctg tct gag ggt gcc acc ccc cag gac
Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp
165 170 175 | 528 |
| ctg aac acc atg ctg aac aca gtg ggg ggc cat cag gct gcc atg cag
Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala Ala Met Gln
180 185 190 | 576 |
| atg ctg aag gag acc atc aat gag gag gct gct gag tgg gac agg ctg
Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Leu
195 200 205 | 624 |
| cat cct gtg cac gct ggc ccc att gcc ccc ggc cag atg agg gag ccc
His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro
210 215 220 | 672 |
| agg ggc tct gac att gct ggc acc acc tcc acc ctc cag gag cag att
Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile
225 230 235 240 | 720 |
| ggc tgg atg acc aac aac ccc ccc atc cct gtg ggg gaa atc tac aag
Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys
245 250 255 | 768 |

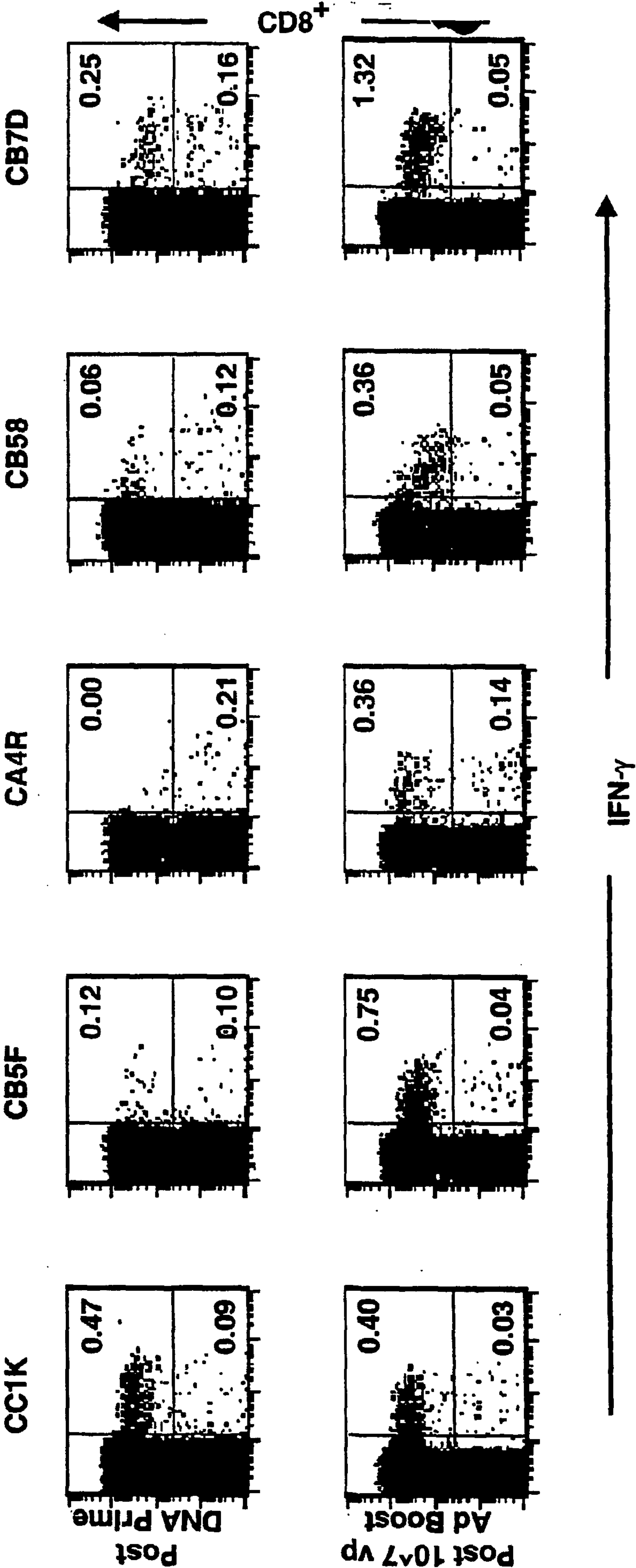
Figure 30A'

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|------|------|
| agg | tgg | atc | atc | ctg | ggc | ctg | aac | aag | att | gtg | agg | atg | tac | tcc | ccc | 816 |
| Arg | Trp | Ile | Ile | Leu | Gly | Leu | Asn | Lys | Ile | Val | Arg | Met | Tyr | Ser | Pro | |
| | | | 260 | | | | | 265 | | | | | 270 | | | |
| acc | tcc | atc | ctg | gac | atc | agg | cag | ggc | ccc | aag | gag | ccc | ttc | agg | gac | 864 |
| Thr | Ser | Ile | Leu | Asp | Ile | Arg | Gln | Gly | Pro | Lys | Glu | Pro | Phe | Arg | Asp | |
| | | 275 | | | | | 280 | | | | | 285 | | | | |
| tat | gtg | gac | agg | ttc | tac | aag | acc | ctg | agg | gct | gag | cag | gcc | tcc | cag | 912 |
| Tyr | Val | Asp | Arg | Phe | Tyr | Lys | Thr | Leu | Arg | Ala | Glu | Gln | Ala | Ser | Gln | |
| | 290 | | | | | 295 | | | | | 300 | | | | | |
| gag | gtg | aag | aac | tgg | atg | aca | gag | acc | ctg | ctg | gtg | cag | aat | gcc | aac | 960 |
| Glu | Val | Lys | Asn | Trp | Met | Thr | Glu | Thr | Leu | Leu | Val | Gln | Asn | Ala | Asn | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | |
| cct | gac | tgc | aag | acc | atc | ctg | aag | gcc | ctg | ggc | cct | gct | gcc | acc | ctg | 1008 |
| Pro | Asp | Cys | Lys | Thr | Ile | Leu | Lys | Ala | Leu | Gly | Pro | Ala | Ala | Thr | Leu | |
| | | | | 325 | | | | | 330 | | | | | 335 | | |
| gag | gag | atg | atg | aca | gcc | tgc | cag | ggg | gtg | ggg | ggc | cct | ggt | cac | aag | 1056 |
| Glu | Glu | Met | Met | Thr | Ala | Cys | Gln | Gly | Val | Gly | Gly | Pro | Gly | His | Lys | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |
| gcc | agg | gtg | ctg | gct | gag | gcc | atg | tcc | cag | gtg | acc | aac | tcc | gcc | acc | 1104 |
| Ala | Arg | Val | Leu | Ala | Glu | Ala | Met | Ser | Gln | Val | Thr | Asn | Ser | Ala | Thr | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| atc | atg | atg | cag | agg | ggc | aac | ttc | agg | aac | cag | agg | aag | aca | gtg | aag | 1152 |
| Ile | Met | Met | Gln | Arg | Gly | Asn | Phe | Arg | Asn | Gln | Arg | Lys | Thr | Val | Lys | |
| | 370 | | | | | 375 | | | | | 380 | | | | | |
| tgc | ttc | aac | tgt | ggc | aag | gtg | ggc | cac | att | gcc | aag | aac | tgt | agg | gcc | 1200 |
| Cys | Phe | Asn | Cys | Gly | Lys | Val | Gly | His | Ile | Ala | Lys | Asn | Cys | Arg | Ala | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| ccc | agg | aag | aag | ggc | tgc | tgg | aag | tgt | ggc | aag | gag | ggc | cac | cag | atg | 1248 |
| Pro | Arg | Lys | Lys | Gly | Cys | Trp | Lys | Cys | Gly | Lys | Glu | Gly | His | Gln | Met | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| aag | gac | tgc | aat | gag | agg | cag | gcc | aac | ttc | ctg | ggc | aaa | atc | tgg | ccc | 1296 |
| Lys | Asp | Cys | Asn | Glu | Arg | Gln | Ala | Asn | Phe | Leu | Gly | Lys | Ile | Trp | Pro | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| tcc | cac | aag | ggc | agg | cct | ggc | aac | ttc | ctc | cag | tcc | agg | cct | gag | ccc | 1344 |
| Ser | His | Lys | Gly | Arg | Pro | Gly | Asn | Phe | Leu | Gln | Ser | Arg | Pro | Glu | Pro | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| aca | gcc | cct | ccc | gag | gag | tcc | ttc | agg | ttt | ggg | gag | gag | aag | acc | acc | 1392 |
| Thr | Ala | Pro | Pro | Glu | Glu | Ser | Phe | Arg | Phe | Gly | Glu | Glu | Lys | Thr | Thr | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| ccc | agc | cag | aag | cag | gag | ccc | att | gac | aag | gag | ctg | tac | ccc | ctg | gcc | 1440 |
| Pro | Ser | Gln | Lys | Gln | Glu | Pro | Ile | Asp | Lys | Glu | Leu | Tyr | Pro | Leu | Ala | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| tcc | ctg | agg | tcc | ctg | ttt | ggc | aac | gac | ccc | tcc | tcc | cag | taa | (SID NO:36) | 1482 | |
| Ser | Leu | Arg | Ser | Leu | Phe | Gly | Asn | Asp | Pro | Ser | Ser | Gln | * | (SID NO:37) | | |
| | | | | 485 | | | | | 490 | | | | | | | |

Figure 30 B

Figure 31

IFN- γ Secretion against Gag 20-aa pool from CD3⁺ T cells of Monkey PBMCs



Comparison of Single-Modality Adenovirus Immunization with DNA+Adjuvant Prime/Adenovirus Boost

Immunizations

Ad Prime/Boost DNA-CRL1005 Prime/Ad Boost

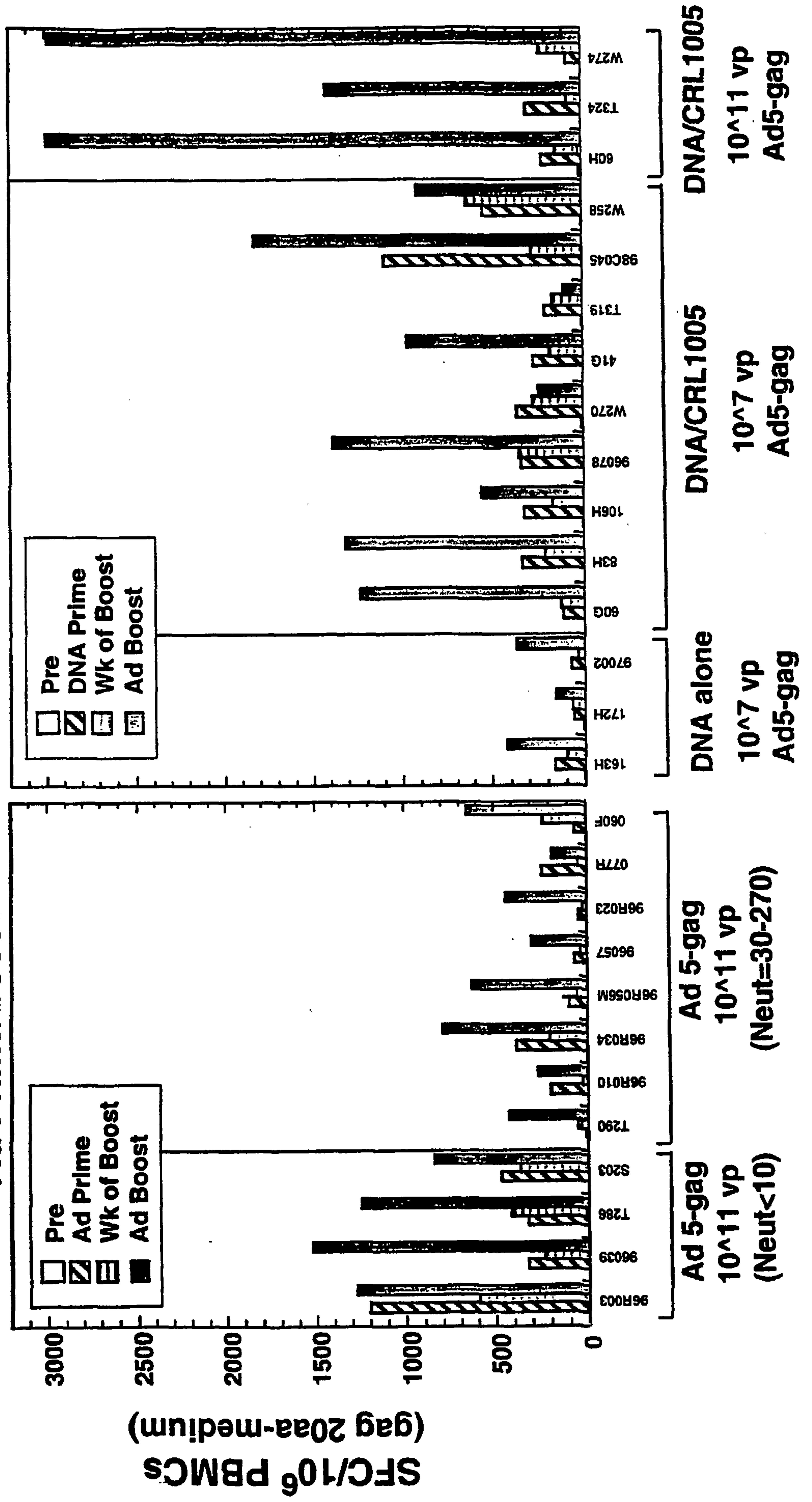


FIGURE 33A

ATGGGTGCTA GGGCTTCTGT GCTGTCTGGT GGTGAGCTGG ACAAGTGGGA GAAGATCAGG
 CTGAGGCCTG GTGGCAAGAA GAAGTACAAG CTAAAGCACA TTGTGTGGGC CTCCAGGGAG
 CTGGAGAGGT TTGCTGTGAA CCCTGGCCTG CTGGAGACCT CTGAGGGGTG CAGGCAGATC
 CTGGGCCAGC TCCAGCCCTC CCTGCAAACA GGCTCTGAGG AGCTGAGGTC CCTGTACAAC
 ACAGTGGCTA CCCTGTACTG TGTGCACCAG AAGATTGATG TGAAGGACAC CAAGGAGGCC
 CTGGAGAAGA TTGAGGAGGA GCAGAACAAG TCCAAGAAGA AGGCCAGCA GGCTGCTGCT
 GGCACAGGCA ACTCCAGCCA GGTGTCCCAG AACTACCCCA TTGTGCAGAA CCTCCAGGGC
 CAGATGGTGC ACCAGGCCAT CTCCCCCGG ACCCTGAATG CCTGGGTGAA GGTGGTGGAG
 GAGAAGGCCT TCTCCCTGA GGTGATCCCC ATGTTCTCTG CCCTGTCTGA GGGTGCCACC
 CCCCAGGACC TGAACACCAT GCTGAACACA GTGGGGGGCC ATCAGGCTGC CATGCAGATG
 CTGAAGGAGA CCATCAATGA GGAGGCTGCT GAGTGGGACA GGCTGCATCC TGTGCACGCT
 GGCCCCATTG CCCCCGGCCA GATGAGGGAG CCCAGGGGCT CTGACATTGC TGGCACCACC
 TCCACCCTCC AGGAGCAGAT TGGCTGGATG ACCAACAACC CCCCATCCC TGTGGGGGAA
 ATCTACAAGA GGTGGATCAT CCTGGGCCTG AACAAAGATTG TGAGGATGTA CTCCCCCACC
 TCCATCCTGG ACATCAGGCA GGGCCCCAAG GAGCCCTTCA GGGACTATGT GGACAGGTTC
 TACAAGACCC TGAGGGCTGA GCAGGCCTCC CAGGAGGTGA AGAACTGGAT GACAGAGACC
 CTGCTGGTGC AGAATGCCAA CCCTGACTGC AAGACCATCC TGAAGGCCCT GGGCCCTGCT
 GCCACCCTGG AGGAGATGAT GACAGCCTGC CAGGGGGTGG GGGGCCCTGG TCACAAGGCC
 AGGGTGCTGG CTGAGGCCAT GTCCCAGGTG ACCAACTCCG CCACCATCAT GATGCAGAGG
 GGCAACTTCA GGAACCAGAG GAAGACAGTG AAGTGCTTCA ACTGTGGCAA GGTGGGCCAC
 ATTGCCAAGA ACTGTAGGGC CCCCAGGAAG AAGGGCTGCT GGAAGTGTGG CAAGGAGGGC
 CACCAGATGA AGGACTGCAA TGAGAGGCAG GCCAACTTCC TGGGCAAAT CTGGCCCTCC
 CACAAGGGCA GGCCTGGCAA CTTCTCCAG TCCAGGCCTG AGCCCACAGC CCCTCCCGAG
 GAGTCCTTCA GGTTTGGGGA GGAGAAGACC ACCCCCAGCC AGAAGCAGGA GCCCATTGAC
 AAGGAGCTGT ACCCCCTGGC CTCCCTGAGG TCCCTGTTTG GCAACGACCC CTCCTCCAG
 ATGGCTCCCA TCTCCCCCAT TGAGACTGTG CCTGTGAAGC TGAAGCCTGG CATGGATGGC
 CCAAGGTGA AGCAGTGGCC CCTGACTGAG GAGAAGATCA AGGCCCTGGT GGAAATCTGC
 ACTGAGATGG AGAAGGAGGG CAAAATCTCC AAGATTGGCC CCGAGAACCC CTACAACACC
 CCTGTGTTTG CCATCAAGAA GAAGGACTCC ACCAAGTGA GGAAGCTGGT GGAATTCAGG
 GAGCTGAACA AGAGGACCCA GGACTTCTGG GAGGTGCAGC TGGGCATCCC CCACCCCGCT
 GGCCTGAAGA AGAAGAAGTC TGTGACTGTG CTGGCTGTGG GGGATGCCTA CTTCTCTGTG
 CCCCTGGATG AGGACTTCAG GAAGTACACT GCCTTCACCA TCCCCTCCAT CAACAATGAG
 ACCCCTGGCA TCAGGTACCA GTACAATGTG CTGCCCCAGG GCTGGAAGGG CTCCCCTGCC
 ATCTTCCAGT CCTCCATGAC CAAGATCCTG GAGCCCTTCA GGAAGCAGAA CCCTGACATT
 GTGATCTACC AGTACATGGC TGCCCTGTAT GTGGGCTCTG ACCTGGAGAT TGGGCAGCAC
 AGGACCAAGA TTGAGGAGCT GAGGCAGCAC CTGCTGAGGT GGGGCTGAC CACCCCTGAC
 AAGAAGCACC AGAAGGAGCC CCCCTTCCTG TGGATGGGCT ATGAGCTGCA CCCCAGACAAG
 TGGACTGTGC AGCCCATTTG GCTGCCTGAG AAGGACTCCT GGAAGTGTGAA TGACATCCAG
 AAGCTGGTGG GCAAGCTGAA CTGGGCCTCC CAAATCTACC CTGGCATCAA GGTGAGGCAG
 CTGTGCAAGC TGCTGAGGGG CACCAAGGCC CTGACTGAGG TGATCCCCCT GACTGAGGAG
 GCTGAGCTGG AGCTGGCTGA GAACAGGGAG ATCCTGAAGG AGCCTGTGCA TGGGGTGTAC

FIGURE 33B

TATGACCCCT CCAAGGACCT GATTGCTGAG ATCCAGAAGC AGGGCCAGGG CCAGTGGACC
TACCAAATCT ACCAGGAGCC CTTCAAGAAC CTGAAGACTG GCAAGTATGC CAGGATGAGG
GGGGCCCACA CCAATGATGT GAAGCAGCTG ACTGAGGCTG TGCAGAAGAT CACCACTGAG
TCCATTGTGA TCTGGGGCAA GACCCCCAAG TTCAAGCTGC CCATCCAGAA GGAGACCTGG
GAGACCTGGT GGA CTGAGTA CTGGCAGGCC ACCTGGATCC CTGAGTGGGA GTTTGTGAAC
ACCCCCCCCC TGGTGAAGCT GTGGTACCAG CTGGAGAAGG AGCCCATTTGT GGGGGCTGAG
ACCTTCTATG TGGCTGGGGC TGCCAACAGG GAGACCAAGC TGGGCAAGGC TGGCTATGTG
ACCAACAGGG GCAGGCAGAA GGTGGTGACC CTGACTGACA CCACCAACCA GAAGACTGCC
CTCCAGGCCA TCTACCTGGC CCTCCAGGAC TCTGGCCTGG AGGTGAACAT TGTGACTGCC
TCCCAGTATG CCCTGGGCAT CATCCAGGCC CAGCCTGATC AGTCTGAGTC TGAGCTGGTG
AACCAGATCA TTGAGCAGCT GATCAAGAAG GAGAAGGTGT ACCTGGCCTG GGTGCCTGCC
CACAAGGGCA TTGGGGGCAA TGAGCAGGTG GACAAGCTGG TGTCTGCTGG CATCAGGAAG
GTGCTGTTCC TGGATGGCAT TGACAAGGCC CAGGATGAGC ATGAGAAGTA CCACTCCAAC
TGGAGGGCTA TGGCCTCTGA CTTCAACCTG CCCCTGTGG TGGCTAAGGA GATTGTGGCC
TCCTGTGACA AGTGCCAGCT GAAGGGGGAG GCCATGCATG GGCAGGTGGA CTGCTCCCCT
GGCATCTGGC AGCTGGCCTG CACCCACCTG GAGGGCAAGG TGATCCTGGT GGCTGTGCAT
GTGGCCTCCG GCTACATTGA GGCTGAGGTG ATCCCTGCTG AGACAGGCCA GGAGACTGCC
TACTTCCTGC TGAAGCTGGC TGGCAGGTGG CCTGTGAAGA CCATCCACAC TGCCAATGGC
TCCAAC TTCA CTGGGGCCAC AGTGAGGGCT GCCTGCTGGT GGGCTGGCAT CAAGCAGGAG
TTTGGCATCC CCTACAACCC CCAGTCCCAG GGGGTGGTGG CCTCCATGAA CAAGGAGCTG
AAGAAGATCA TTGGGCAGGT GAGGGACCAG GCTGAGCACC TGAAGACAGC TGTGCAGATG
GCTGTGTTCA TCCACAACCT CAAGAGGAAG GGGGGCATCG GGGGCTACTC CGCTGGGGAG
AGGATTGTGG ACATCATTGC CACAGACATC CAGACCAAGG AGCTCCAGAA GCAGATCACC
AAGATCCAGA ACTTCAGGGT GTACTACAGG GACTCCAGGA ACCCCCTGTG GAAGGGCCCT
GCCAAGCTGC TGTGGAAGGG GGAGGGGGCT GTGGTGATCC AGGACAACTC TGACATCAAG
GTGGTGCCCA GGAGGAAGGC CAAGATCATC AGGGACTATG GCAAGCAGAT GGCTGGGGAT
GACTGTGTGG CCTCCAGGCA GGATGAGGAC TAA

SEQ ID NO: 38

FIGURE 34A

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp Glu Lys
 Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile Val Trp
 Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser
 Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly Ser
 Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln
 Lys Ile Asp Val Lys Asp Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln
 Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Ala Gly Thr Gly Asn Ser Ser
 Gln Val Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys
 Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr
 Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala Ala Met
 Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Leu His
 Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser
 Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn
 Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu
 Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala
 Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val Gln
 Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr
 Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro Gly His Lys Ala
 Arg Val Leu Ala Glu Ala Met Ser Gln Val Thr Asn Ser Ala Thr Ile Met Met
 Gln Arg Gly Asn Phe Arg Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly
 Lys Val Gly His Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp
 Lys Cys Gly Lys Glu Gly His Gln Met Lys Asp Cys Asn Glu Arg Gln Ala Asn
 Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln
 Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Phe Gly Glu Glu
 Lys Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp Lys Glu Leu Tyr Pro Leu
 Ala Ser Leu Arg Ser Leu Phe Gly Asn Asp Pro Ser Ser Gln Met Ala Pro Ile
 Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys
 Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys
 Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr
 Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu
 Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu
 Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala
 Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr
 Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr
 Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met
 Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln
 Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr
 Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp
 Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro

FIGURE 34B

Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val
Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro
Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr
Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu
Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu
Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr
Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile
Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp
Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe
Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile
Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu
Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr
Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp
Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile
Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln
Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu
Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn
Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile
Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val
Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro
Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp
Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val
Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn
Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile
Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val
Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln
Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu
Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln
Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
SEQ ID NO: 39