

(19) World Intellectual Property
Organization
International Bureau



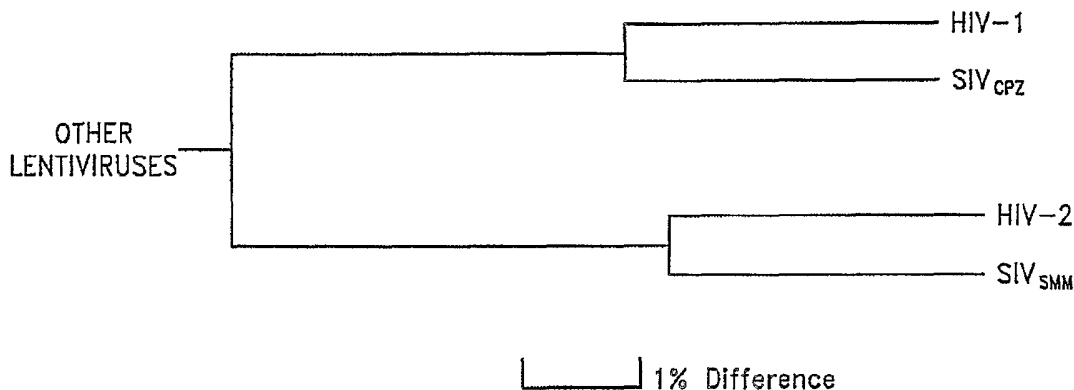
(43) International Publication Date
14 October 2004 (14.10.2004)

PCT

(10) International Publication Number
WO 2004/087201 A2

- (51) International Patent Classification⁷: **A61K 39/12**, 39/275, 39/21, 39/285
- (74) Agent: **DELANEY, Karoline, A.**; Knobbe, Martens, Olson and Bear, LLP, 2040 Main Street, Fourteenth Floor, Irvine, CA 92614 (US).
- (21) International Application Number: PCT/US2004/009906
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 29 March 2004 (29.03.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/459,175 28 March 2003 (28.03.2003) US
- (71) Applicant (*for all designated States except US*): **THE GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; National Institutes of Health, Office of Technology Transfer, 6011 Executive Blvd., Suite 325, Rockville, MD 20852-3804 (US).
- (84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **MOSS, Bernard** [US/US]; 10301 Dickens Ave., Bethesda, MD 20814 (US). **EARL, Patricia, L.** [US/US]; 3520 Woodbine Street, Chevy Chase, MD 20815 (US). **WYATT, Linda** [US/US]; 11201 Huntover Drive, Rockville, MD 20852 (US). **ELLER, Leigh Anne** [US/US]; 11404 Flowerton Place, Germantown, MD 20876 (US). **VANCOTT, Thomas** [US/US]; 19108 Mt Airey Rd., Brookeville, MD 20833 (US). **HARRIS, Matthew, Edward** [US/US]; 12742 Avenida de Espuela, Poway, CA 92064 (US).
- Declaration under Rule 4.17:**
— *of inventorship (Rule 4.17(iv)) for US only*
- Published:**
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: MVA EXPRESSING MODIFIED HIV ENVELOPE, GAG, AND POL GENES



(57) Abstract: The invention provides modified virus Ankara (MVA), a replication-deficient strain of vaccinia virus, expressing human immunodeficiency virus (HIV) env, gag, and pol genes.

MVA EXPRESSING MODIFIED HIV ENVELOPE, GAG, AND POL GENES

Field of the Invention

The invention provides modified vaccinia Ankara (MVA), a replication-deficient
5 strain of vaccinia virus, expressing human immunodeficiency virus (HIV) *env*, *gag*, and *pol*
genes.

Background of the Invention

Cellular immunity plays an important role in the control of immunodeficiency virus
infections (P.J. Goulder *et al.* 1999 *AIDS* 13:S121). Recently, a DNA vaccine designed to
10 enhance cellular immunity by cytokine augmentation successfully contained a highly
virulent immunodeficiency virus challenge (D.H. Barouch *et al.* 2000 *Science* 290:486).
Another promising approach to raising cellular immunity is DNA priming followed by
recombinant poxvirus boosters (H.L. Robinson *et al.* 2000 *AIDS Rev* 2:105). This
heterologous prime/boost regimen induces 10- to 100-fold higher frequencies of T cells
15 than priming and boosting with DNA or recombinant poxvirus vaccines alone. Previously,
investigators showed that boosting a DNA-primed response with a poxvirus was superior to
boosting with DNA or protein for the control of a non-pathogenic immunodeficiency virus
(H.L. Robinson *et al.* 1999 *Nat Med* 5:526). There is a need for the control of a pathogenic
immunodeficiency virus.

Summary of the Invention

Here we report that DNA priming followed by a recombinant modified vaccinia
Ankara (rMVA) booster has controlled a highly pathogenic immunodeficiency virus
challenge in a rhesus macaque model. Both the DNA and rMVA components of the
vaccine expressed multiple immunodeficiency virus proteins. Two DNA inoculations at 0
25 and 8 weeks and a single rMVA booster at 24 weeks effectively controlled an intrarectal
challenge administered seven months after the booster. These findings are envisioned as
indicating that a relatively simple multiprotein DNA/MVA vaccine can help to control the
acquired immune deficiency syndrome (AIDS) epidemic. We also report that inoculations
of rMVA induce good immune responses even without DNA priming.

Brief Description of the Drawings

30 **Figure 1.** Phylogenetic relationships of HIV-1 and HIV-2 based on identity of *pol*
gene sequences. SIV_{cpz} and SIV_{smm} are subhuman primate lentiviruses recovered from a
chimpanzee and sooty mangabey monkey, respectively.

Figure 2. Phylogenetic relationships of HIV-1 groups M, N and O with four different SIV_{cpz} isolates based on full-length *pol* gene sequences. The bar indicates a genetic distance of 0.1 (10% nucleotide divergence) and the *asterisk* positions group N HIV-1 isolates based on *env* sequences.

5 **Figure 3.** Tropic and biologic properties of HIV-1 isolates.

Figure 4. HIV-encoded proteins. The location of the HIV genes, the sizes of primary translation products (in some cases polyproteins), and the processed mature viral proteins are indicated.

Figure 5. Schematic representation of a mature HIV-1 virion.

10 **Figure 6.** Linear representation of the HIV-1 Env glycoprotein. The *arrow* indicates the site of gp160 cleavage to gp120 and gp41. In gp120, *cross-hatched* areas represent variable domains (V₁ to V₅) and *open boxes* depict conserved sequences (C₁ to C₅). In the gp41 ectodomain, several domains are indicated: the N-terminal fusion peptide, and the two ectodomain helices (N- and C-helix). The membrane-spanning domain is represented by a *black box*. In the gp41 cytoplasmic domain, the Tyr-X-X-Leu (YXXL) endocytosis motif (SEQ ID NO: 9) and two predicted helical domains (helix-1 and -2) are shown. Amino acid numbers are indicated.

Figure 7. Temporal frequencies of Gag-specific T cells. (A) Gag-specific CD8 T cell responses raised by DNA priming and rMVA booster immunizations. The schematic presents mean Gag-CM9-tetramer data generated in the high-dose i.d. DNA-immunized animals. (B) Gag-specific IFN- γ ELISPOTs in A*01 (open bars) and non-A*01 (filled bars) macaques at various times before challenge and at two weeks after challenge. Three pools of 10 to 13 Gag peptides (22-mers overlapping by 12) were used for the analyses. The numbers above data bars represent the arithmetic mean $\pm$ SD for the ELISPOTs within each group. The numbers at the top of the graphs designate individual animals. *, data not available; #, <20 ELISPOTs per 1×10^6 peripheral blood mononuclear cells (PBMC). Temporal data for Gag-CM9-Mamu-A*01 tetramer-specific T cells can be found in Figure 12.

30 **Figure 8.** Temporal viral loads, CD4 counts, and survival after challenge of vaccinated and control animals. (A) Geometric mean viral loads and (B) geometric mean CD4 counts. (C) Survival curve for vaccinated and control animals. The dotted line represents all 24 vaccinated animals. (D) Viral loads and (E) CD4 counts for individual animals in the vaccine and control groups. The key to animal numbers is presented in (E).

Assays for the first 12 weeks after challenge had a detection level of 1000 copies of RNA per milliliter of plasma. Animals with loads below 1000 were scored with a load of 500. For weeks 16 and 20, the detection level was 300 copies of RNA per milliliter. Animals with levels of virus below 300 were scored at 300.

5 **Figure 9.** Postchallenge T cell responses in vaccine and control groups. (A) Temporal tetramer⁺ cells (dashed line) and viral loads (solid line). (B) Intracellular cytokine assays for IFN- γ production in response to stimulation with the Gag-CM9 peptide at two weeks after challenge. This *ex vivo* assay allows evaluation of the functional status of the peak postchallenge tetramer⁺ cells displayed in Figure 7A. (C) Proliferation assay at
10 12 weeks after challenge. Gag-Pol-Env (open bars) and Gag-Pol (hatched bars) produced by transient transfections were used for stimulation. Supernatants from mock-transfected cultures served as control antigen. Stimulation indices are the growth of cultures in the presence of viral antigens divided by the growth of cultures in the presence of mock antigen.

15 **Figure 10.** Lymph node histomorphology at 12 weeks after challenge. (A) Typical lymph node from a vaccinated macaque showing evidence of follicular hyperplasia characterized by the presence of numerous secondary follicles with expanded germinal centers and discrete dark and light zones. (B) Typical lymph node from an infected control animal showing follicular depletion and paracortical lymphocellular atrophy. (C) A
20 representative lymph node from an age-matched, uninfected macaque displaying nonreactive germinal centers. (D) The percentage of the total lymph node area occupied by germinal centers was measured to give a non-specific indicator of follicular hyperplasia. Data for uninfected controls are for four age-matched rhesus macaques.

25 **Figure 11.** Temporal antibody responses. Micrograms of total Gag (A) or Env (B) antibody were determined with ELISAs. The titers of neutralizing antibody for SHIV-89.6 (C) and SHIV-89.6P (D) were determined with MT-2 cell killing and neutral red staining (D.C. Montefiori *et al.* 1988 *J Clin Microbiol* 26:231). Titers are the reciprocal of the serum dilution giving 50% neutralization of the indicated viruses grown in human PBMC. Symbols for animals are the same as in Figure 8.

30 **Figure 12.** Gag-CM9-Mamu-A*01 tetramer-specific T cells in Mamu-A*01 vaccinated and control macaques at various times before challenge and at two weeks after challenge. The number at the upper right corner of each plot represents the frequency of

tetramer-specific CD8 T cells as a % of total CD8 T cells. The numbers above each column of FACS data designate individual animals.

Figure 13. Map of plasmid transfer vector pLW-48.

Figure 14 A-I. Sequences of plasmid transfer vector pLW-48, Psy II promoter (which controls ADA envelope expression), ADA envelope truncated, PmH5 promoter (which controls HXB2 gag pol expression), and HXB2 *gag pol* (with safety mutations, Δ integrase).

Figure 15. Plasmid transfer vector pLW-48 and making MVA recombinant virus MVA/HIV 48.

Figure 16. A clade B gag pol.

Figure 17. Sequence of new Psyn II promoter.

Figure 18. pLAS-1 and pLAS-2.

Figure 19. pLAS-1/UGDgag.

Figure 20. pLAS-2/UGDenv.

Figure 21. pLAS-2/UGDrev env.

Figure 22. Schematic for recombinant MVA production.

Figure 23. Overview of making recombinant MVA/UGD viruses.

Figure 24. Immunoprecipitation analysis.

Figure 25. Functional analysis of expressed proteins. **A.** Virus-like particle assay. **B.** Env fusion assay.

Figure 26. MVA/UGD induced HIV env- and gag-specific antibody responses. **A.** HIV p24-specific serum IgG responses. **B.** HIV env-specific serum IgG responses. **C.** MVA/UGD induced HIV env- and gag-specific antibody responses (study 1).

Figure 27. MVA/UGD induced gag-specific intracellular IFN- γ production.

Figure 28A. MVA/UGD induced gag-specific IFN- γ ELISPOT.

Figure 28B. MVA/UGD induced pol-specific IFN- γ ELISPOT.

Figure 29. MVA/UGD induced gag-specific tetramer staining.

Figure 30. MVA/UGD induced gag-specific antibody responses (study 2).

Figure 31A & B. MVA/UGD induced gag- and pol-specific intracellular IFN- γ production (study 2).

Figure 32 A, B, & C. MVA/UGD induced gag- and pol-specific IFN- γ ELISPOT (study 2).

Figure 33. MVA/UGD induced gag-specific tetramer staining (study 2).

Figure 34. MVA/UGD induced gag-specific cytotoxic T cell killing.

Figure 35. Immunoprecipitation analysis of cell lysates (MVA/KEA).

Figure 36. Gag particle assay (MVA/KEA).

Figure 37. Fusion assay (MVA/KEA).

5 **Figure 38.** MVA/KEA induced HIV-1 env-specific antibody responses.

Figure 39. MVA/KEA induced gag-specific intracellular IFN- γ production.

Figure 40. MVA/KEA induced gag-specific tetramer staining.

Figure 41. MVA/KEA induced gag-specific IFN- γ ELISPOT.

Figure 42. Immunoprecipitation analysis of cell lysates (MVA/TZC).

10 **Figure 43.** Fusion assay (MVA/TZC).

Brief Description of the Appendices

Appendix 1. DNA sequences of gagpol and env genes from Ugandan HIV-1 clade D isolates.

Appendix 2. DNA sequences of MVA shuttle plasmids, pLAS-1 and pLAS-2.

15 **Appendix 3.** DNA sequences of gagpol and env genes from Kenyan HIV-1 clade A isolates.

Appendix 4. DNA sequences of gagpol and env genes from Tanzanian HIV-1 clade C isolates.

20 **Appendix 5.** American Type Culture Collection, Budapest Treaty Deposit Form BP/1, questions 1-8, regarding MVA 1974/NIH Clone 1.

Appendix 6. American Tissue Type Collection, Additional Information Required When Depositing A Virus for Patent Purposes, regarding MVA 1974/NIH Clone 1.

Deposit of Microorganism

25 The following microorganism has been deposited in accordance with the terms of the Budapest Treaty with the American Type Culture Collection (ATCC), Manassas, VA, on the date indicated:

Microorganism	Accession No.	Date
MVA 1974/NIH Clone 1	PTA-5095	March 27, 2003

30 MVA 1974/NIH Clone 1 was deposited as ATCC Accession No. PTA-5095 on March 27, 2003 with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209, USA. This deposit was made under the provisions of

the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from date of deposit. The deposit will be made available by ATCC under the terms of the Budapest Treaty, and
5 subject to an agreement between Applicant and ATCC which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto
10 according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14). Availability of the deposited strain is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

Detailed Description of the Preferred Embodiment

15 Recombinant MVA Virus

Vaccinia virus, a member of the genus Orthopoxvirus in the family of Poxviridae, was used as live vaccine to immunize against the human smallpox disease. Successful worldwide vaccination with vaccinia virus culminated in the eradication of variola virus, the causative agent of the smallpox ("The global eradication of smallpox. Final report of
20 the global commission for the certification of smallpox eradication". History of Public Health, No. 4, Geneva: World Health Organization, 1980). Since that WHO declaration, vaccination has been universally discontinued except for people at high risk of poxvirus infections (e.g. laboratory workers).

More recently, vaccinia viruses have also been used to engineer viral vectors for
25 recombinant gene expression and for the potential use as recombinant live vaccines (Mackett, M. *et al.* 1982 *PNAS USA* **79**:7415-7419; Smith, G.L. *et al.* 1984 *Biotech Genet Engin Rev* **2**:383-407). This entails DNA sequences (genes) which code for foreign antigens being introduced, with the aid of DNA recombination techniques, into the genome of the vaccinia viruses. If the gene is integrated at a site in the viral DNA which is non-
30 essential for the life cycle of the virus, it is possible for the newly produced recombinant vaccinia virus to be infectious, that is to say able to infect foreign cells and thus to express the integrated DNA sequence (EP Patent Applications No. 83,286 and No. 110,385). The recombinant vaccinia viruses prepared in this way can be used, on the one hand, as live

vaccines for the prophylaxis of infectious diseases, on the other hand, for the preparation of heterologous proteins in eukaryotic cells.

For vector applications health risks would be lessened by the use of a highly attenuated vaccinia virus strain. Several such strains of vaccinia virus were especially developed to avoid undesired side effects of smallpox vaccination. Thus, the modified vaccinia Ankara (MVA) has been generated by long-term serial passages of the Ankara strain of vaccinia virus (CVA) on chicken embryo fibroblasts (for review see Mayr, A. *et al.* 1975 *Infection* 3:6-14; Swiss Patent No. 568,392). The MVA virus is publicly available from American Type Culture Collection as ATCC No. VR-1508. MVA is distinguished by its great attenuation, that is to say by diminished virulence and ability to replicate in primate cells while maintaining good immunogenicity. The MVA virus has been analyzed to determine alterations in the genome relative to the parental CVA strain. Six major deletions of genomic DNA (deletion I, II, III, IV, V, and VI) totaling 31,000 base pairs have been identified (Meyer, H. *et al.* 1991 *J Gen Virol* 72:1031-1038). The resulting MVA virus became severely host cell restricted to avian cells.

Furthermore, MVA is characterized by its extreme attenuation. When tested in a variety of animal models, MVA was proven to be avirulent even in immunosuppressed animals. More importantly, the excellent properties of the MVA strain have been demonstrated in extensive clinical trials (Mayr A. *et al.* 1978 *Zentralbl Bakteriol [B]* 167:375-390; Stickl *et al.* 1974 *Dtsch Med Wschr* 99:2386-2392). During these studies in over 120,000 humans, including high-risk patients, no side effects were associated with the use of MVA vaccine.

MVA replication in human cells was found to be blocked late in infection preventing the assembly to mature infectious virions. Nevertheless, MVA was able to express viral and recombinant genes at high levels even in non-permissive cells and was proposed to serve as an efficient and exceptionally safe gene expression vector (Sutter, G. and Moss, B. 1992 *PNAS USA* 89:10847-10851). Additionally, novel vaccinia vector vaccines were established on the basis of MVA having foreign DNA sequences inserted at the site of deletion III within the MVA genome (Sutter, G. *et al.* 1994 *Vaccine* 12:1032-1040).

The recombinant MVA vaccinia viruses can be prepared as set out hereinafter. A DNA-construct which contains a DNA-sequence which codes for a foreign polypeptide flanked by MVA DNA sequences adjacent to a naturally occurring deletion, *e.g.* deletion

III, or other non-essential sites, within the MVA genome, is introduced into cells infected with MVA, to allow homologous recombination. Once the DNA-construct has been introduced into the eukaryotic cell and the foreign DNA has recombined with the viral DNA, it is possible to isolate the desired recombinant vaccinia virus in a manner known per se, preferably with the aid of a marker. The DNA-construct to be inserted can be linear or circular. A plasmid or polymerase chain reaction product is preferred. The DNA-construct contains sequences flanking the left and the right side of a naturally occurring deletion, e.g. deletion III, within the MVA genome. The foreign DNA sequence is inserted between the sequences flanking the naturally occurring deletion. For the expression of a DNA sequence or gene, it is necessary for regulatory sequences, which are required for the transcription of the gene, to be present on the DNA. Such regulatory sequences (called promoters) are known to those skilled in the art, and include for example those of the vaccinia 11 kDa gene as are described in EP-A-198,328, and those of the 7.5 kDa gene (EP-A-110,385). The DNA-construct can be introduced into the MVA infected cells by transfection, for example by means of calcium phosphate precipitation (Graham *et al.* 1973 *Virol* **52**:456-467; Wigler *et al.* 1979 *Cell* **16**:777-785), by means of electroporation (Neumann *et al.* 1982 *EMBO J* **1**:841-845), by microinjection (Graessmann *et al.* 1983 *Meth Enzymol* **101**:482-492), by means of liposomes (Straubinger *et al.* 1983 *Meth Enzymol* **101**:512-527), by means of spheroplasts (Schaffner 1980 *PNAS USA* **77**:2163-2167) or by other methods known to those skilled in the art.

HIVs and Their Replication

The etiological agent of acquired immune deficiency syndrome (AIDS) is recognized to be a retrovirus exhibiting characteristics typical of the lentivirus genus, referred to as human immunodeficiency virus (HIV). The phylogenetic relationships of the human lentiviruses are shown in **Figure 1**. HIV-2 is more closely related to SIV_{smm}, a virus isolated from sooty mangabey monkeys in the wild, than to HIV-1. It is currently believed that HIV-2 represents a zoonotic transmission of SIV_{smm} to man. A series of lentiviral isolates from captive chimpanzees, designated SIV_{cpz}, are close genetic relatives of HIV-1.

The earliest phylogenetic analyses of HIV-1 isolates focused on samples from Europe/North America and Africa; discrete clusters of viruses were identified from these two areas of the world. Distinct genetic subtypes or clades of HIV-1 were subsequently defined and classified into three groups: M (major); O (outlier); and N (non-M or O) (**Fig.**

2). The M group of HIV-1, which includes over 95% of the global virus isolates, consists of at least eight discrete clades (A, B, C, D, F, G, H, and J), based on the sequence of complete viral genomes. Members of HIV-1 group O have been recovered from individuals living in Cameroon, Gabon, and Equatorial Guinea; their genomes share less than 50% identity in nucleotide sequence with group M viruses. The more recently discovered group N HIV-1 strains have been identified in infected Cameroonians, fail to react serologically in standard whole-virus enzyme-linked immunosorbent assay (ELISA), yet are readily detectable by conventional Western blot analysis.

Most current knowledge about HIV-1 genetic variation comes from studies of group M viruses of diverse geographic origin. Data collected during the past decade indicate that the HIV-1 population present within an infected individual can vary from 6% to 10% in nucleotide sequence. HIV-1 isolates within a clade may exhibit nucleotide distances of 15% in *gag* and up to 30% in *gp120* coding sequences. Interclade genetic variation may range between 30% and 40% depending on the gene analyzed.

All of the HIV-1 group M subtypes can be found in Africa. Clade A viruses are genetically the most divergent and were the most common HIV-1 subtype in Africa early in the epidemic. With the rapid spread of HIV-1 to southern Africa during the mid to late 1990s, clade C viruses have become the dominant subtype and now account for 48% of HIV-1 infections worldwide. Clade B viruses, the most intensively studied HIV-1 subtype, remain the most prevalent isolates in Europe and North America.

High rates of genetic recombination are a hallmark of retroviruses. It was initially believed that simultaneous infections by genetically diverse virus strains were not likely to be established in individuals at risk for HIV-1. By 1995, however, it became apparent that a significant fraction of the HIV-1 group M global diversity included interclade viral recombinants. It is now appreciated that HIV-1 recombinants will be found in geographic areas such as Africa, South America, and Southeast Asia, where multiple HIV-1 subtypes coexist and may account for more than 10% of circulating HIV-1 strains. Molecularly, the genomes of these recombinant viruses resemble patchwork mosaics, with juxtaposed diverse HIV-1 subtype segments, reflecting the multiple crossover events contributing to their generation. Most HIV-1 recombinants have arisen in Africa and a majority contains segments originally derived from clade A viruses. In Thailand, for example, the composition of the predominant circulating strain consists of a clade A *gag* plus *pol* gene segment and a clade E *env* gene. Because the clade E *env* gene in Thai HIV-1 strains is

closely related to the clade E *env* present in virus isolates from the Central African Republic, it is believed that the original recombination event occurred in Africa, with the subsequent introduction of a descendent virus into Thailand. Interestingly, no full-length HIV-1 subtype E isolate (*i.e.*, with subtype E *gag*, *pol*, and *env* genes) has been reported to date.

The discovery that α and β chemokine receptors function as coreceptors for virus fusion and entry into susceptible CD4⁺ cells has led to a revised classification scheme for HIV-1 (**Fig. 3**). Isolates can now be grouped on the basis of chemokine receptor utilization in fusion assays in which HIV-1 gp120 and CD4⁺ coreceptor proteins are expressed in separate cells. As indicated in Figure 3, HIV-1 isolates using the CXCR4 receptor (now designated X4 viruses) are usually T cell line (TCL)-tropic syncytium inducing (SI) strains, whereas those exclusively utilizing the CCR5 receptor (R5 viruses) are predominantly macrophage (M)-tropic and non-syncytium inducing (NSI). The dual-tropic R5/X4 strains, which may comprise the majority of patient isolates and exhibit a continuum of tropic phenotypes, are frequently SI.

As is the case for all replication-competent retroviruses, the three primary HIV-1 translation products, all encoding structural proteins, are initially synthesized as polyprotein precursors, which are subsequently processed by viral or cellular proteases into mature particle-associated proteins (**Fig. 4**). The 55-kd Gag precursor Pr55^{Gag} is cleaved into the matrix (MA), capsid (CA), nucleocapsid (NC), and p6 proteins. Autocatalysis of the 160-kd Gag-Pol polyprotein, Pr160^{Gag-Pol}, gives rise to the protease (PR), the heterodimeric reverse transcriptase (RT), and the integrase (IN) proteins, whereas proteolytic digestion by a cellular enzyme(s) converts the glycosylated 160-kd Env precursor gp160 to the gp120 surface (SU) and gp41 transmembrane (TM) cleavage products. The remaining six HIV-1 encoded proteins (Vif, Vpr, Tat, Rev, Vpu, and Nef) are the primary translation products of spliced mRNAs.

Gag

The Gag proteins of HIV, like those of other retroviruses, are necessary and sufficient for the formation of noninfectious, virus-like particles. Retroviral Gag proteins are generally synthesized as polyprotein precursors; the HIV-1 Gag precursor has been named, based on its apparent molecular mass, Pr55^{Gag}. As noted previously, the mRNA for Pr55^{Gag} is the unspliced 9.2-kb transcript (**Fig. 4**) that requires Rev for its expression in the cytoplasm. When the *pol* ORF is present, the viral protease (PR) cleaves Pr55^{Gag} during or

shortly after budding from the cell to generate the mature Gag proteins p17 (MA), p24 (CA), p7 (NC), and p6 (see Fig. 4). In the virion, MA is localized immediately inside the lipid bilayer of the viral envelope, CA forms the outer portion of the cone-shaped core structure in the center of the particle, and NC is present in the core in a ribonucleoprotein complex with the viral RNA genome (Fig. 5).

The HIV Pr55^{Gag} precursor oligomerizes following its translation and is targeted to the plasma membrane, where particles of sufficient size and density to be visible by EM are assembled. Formation of virus-like particles by Pr55^{Gag} is a self-assembly process, with critical Gag-Gag interactions taking place between multiple domains along the Gag precursor. The assembly of virus-like particles does not require the participation of genomic RNA (although the presence of nucleic acid appears to be essential), *pol*-encoded enzymes, or Env glycoproteins, but the production of infectious virions requires the encapsidation of the viral RNA genome and the incorporation of the Env glycoproteins and the Gag-Pol polyprotein precursor Pr160^{Gag-Pol}.

Pol

Downstream of *gag* lies the most highly conserved region of the HIV genome, the *pol* gene, which encodes three enzymes: PR, RT, and IN (see Fig. 4). RT and IN are required, respectively, for reverse transcription of the viral RNA genome to a double-stranded DNA copy, and for the integration of the viral DNA into the host cell chromosome. PR plays a critical role late in the life cycle by mediating the production of mature, infectious virions. The *pol* gene products are derived by enzymatic cleavage of a 160-kd Gag-Pol fusion protein, referred to as Pr160^{Gag-Pol}. This fusion protein is produced by ribosomal frameshifting during translation of Pr55^{Gag} (see Fig. 4). The frame-shifting mechanism for Gag-Pol expression, also utilized by many other retroviruses, ensures that the *pol*-derived proteins are expressed at a low level, approximately 5% to 10% that of Gag. Like Pr55^{Gag}, the N-terminus of Pr160^{Gag-Pol} is myristylated and targeted to the plasma membrane.

Protease

Early pulse-chase studies performed with avian retroviruses clearly indicated that retroviral Gag proteins are initially synthesized as polyprotein precursors that are cleaved to generate smaller products. Subsequent studies demonstrated that the processing function is provided by a viral rather than a cellular enzyme, and that proteolytic digestion of the Gag and Gag-Pol precursors is essential for virus infectivity. Sequence analysis of retroviral

PRs indicated that they are related to cellular "aspartic" proteases such as pepsin and renin. Like these cellular enzymes, retroviral PRs use two apposed Asp residues at the active site to coordinate a water molecule that catalyzes the hydrolysis of a peptide bond in the target protein. Unlike the cellular aspartic proteases, which function as pseudodimers (using two
5 folds within the same molecule to generate the active site), retroviral PRs function as true dimers. X-ray crystallographic data from HIV-1 PR indicate that the two monomers are held together in part by a four-stranded antiparallel β -sheet derived from both N- and C-terminal ends of each monomer. The substrate-binding site is located within a cleft formed between the two monomers. Like their cellular homologs, the HIV PR dimer contains
10 flexible "flaps" that overhang the binding site and may stabilize the substrate within the cleft; the active-site Asp residues lie in the center of the dimer. Interestingly, although some limited amino acid homology is observed surrounding active-site residues, the primary sequences of retroviral PRs are highly divergent, yet their structures are remarkably similar.

15 Reverse Transcriptase

By definition, retroviruses possess the ability to convert their single-stranded RNA genomes into double-stranded DNA during the early stages of the infection process. The enzyme that catalyzes this reaction is RT, in conjunction with its associated RNaseH activity. Retroviral RTs have three enzymatic activities: (a) RNA-directed DNA
20 polymerization (for minus-strand DNA synthesis), (b) RNaseH activity (for the degradation of the tRNA primer and genomic RNA present in DNA-RNA hybrid intermediates), and (c) DNA-directed DNA polymerization (for second- or plus-strand DNA synthesis).

The mature HIV-1 RT holoenzyme is a heterodimer of 66 and 51 kd subunits. The 51-kd subunit (p51) is derived from the 66-kd (p66) subunit by proteolytic removal of the
25 C-terminal 15-kd RNaseH domain of p66 by PR (see **Fig. 4**). The crystal structure of HIV-1 RT reveals a highly asymmetric folding in which the orientations of the p66 and p51 subunits differ substantially. The p66 subunit can be visualized as a right hand, with the polymerase active site within the palm, and a deep template-binding cleft formed by the palm, fingers, and thumb subdomains. The polymerase domain is linked to RNaseH by the
30 connection subdomain. The active site, located in the palm, contains three critical Asp residues (110, 185, and 186) in close proximity, and two coordinated Mg^{2+} ions. Mutation of these Asp residues abolishes RT polymerizing activity. The orientation of the three active-site Asp residues is similar to that observed in other DNA polymerases (*e.g.*, the

Klenow fragment of *E. coli* DNA polI). The p51 subunit appears to be rigid and does not form a polymerizing cleft; Asp 110, 185, and 186 of this subunit are buried within the molecule. Approximately 18 base pairs of the primer-template duplex lie in the nucleic acid binding cleft, stretching from the polymerase active site to the RNaseH domain.

5 In the RT-primer-template-dNTP structure, the presence of a dideoxynucleotide at the 3' end of the primer allows visualization of the catalytic complex trapped just prior to attack on the incoming dNTP. Comparison with previously obtained structures suggests a model whereby the fingers close in to trap the template and dNTP prior to nucleophilic attack of the 3'-OH of the primer on the incoming dNTP. After the addition of the
10 incoming dNTP to the growing chain, it has been proposed that the fingers adopt a more open configuration, thereby releasing the pyrophosphate and enabling RT to bind the next dNTP. The structure of the HIV-1 RNaseH has also been determined by x-ray crystallography; this domain displays a global folding similar to that of *E. coli* RNaseH.

Integrase

15 A distinguishing feature of retrovirus replication is the insertion of a DNA copy of the viral genome into the host cell chromosome following reverse transcription. The integrated viral DNA (the provirus) serves as the template for the synthesis of viral RNAs and is maintained as part of the host cell genome for the lifetime of the infected cell. Retroviral mutants deficient in the ability to integrate generally fail to establish a
20 productive infection.

The integration of viral DNA is catalyzed by integrase, a 32-kd protein generated by PR-mediated cleavage of the C-terminal portion of the HIV-1 Gag-Pol polyprotein (see Fig. 4).

Retroviral IN proteins are composed of three structurally and functionally distinct
25 domains: an N-terminal, zinc-finger-containing domain, a core domain, and a relatively nonconserved C-terminal domain. Because of its low solubility, it has not yet been possible to crystallize the entire 288-amino-acid HIV-1 IN protein. However, the structure of all three domains has been solved independently by x-ray crystallography or NMR methods. The crystal structure of the core domain of the avian sarcoma virus IN has also
30 been determined. The N-terminal domain (residues 1 to 55), whose structure was solved by NMR spectroscopy, is composed of four helices with a zinc coordinated by amino acids His-12, His-16, Cys-40, and Cys-43. The structure of the N-terminal domain is reminiscent of helical DNA binding proteins that contain a so-called helix-turn-helix motif; however, in

the HIV-1 structure this motif contributes to dimer formation. Initially, poor solubility hampered efforts to solve the structure of the core domain. However, attempts at crystallography were successful when it was observed that a Phe-to-Lys change at IN residue 185 greatly increased solubility without disrupting *in vitro* catalytic activity. Each monomer of the HIV-1 IN core domain (IN residues 50 to 212) is composed of a five-stranded β -sheet flanked by helices; this structure bears striking resemblance to other polynucleotidyl transferases including RNaseH and the bacteriophage MuA transposase. Three highly conserved residues are found in analogous positions in other polynucleotidyl transferases; in HIV-1 IN these are Asp-64, Asp-116 and Glu-152, the so-called D,D-35-E motif. Mutations at these positions block HIV IN function both *in vivo* and *in vitro*. The close proximity of these three amino acids in the crystal structure of both avian sarcoma virus and HIV-1 core domains supports the hypothesis that these residues play a central role in catalysis of the polynucleotidyl transfer reaction that is at the heart of the integration process. The C-terminal domain, whose structure has been solved by NMR methods, adopts a five-stranded β -barrel folding topology reminiscent of a Src homology 3 (SH3) domain. Recently, the x-ray structures of SIV and Rous sarcoma virus IN protein fragments encompassing both the core and C-terminal domains have been solved.

Env

The HIV Env glycoproteins play a major role in the virus life cycle. They contain the determinants that interact with the CD4 receptor and coreceptor, and they catalyze the fusion reaction between the lipid bilayer of the viral envelope and the host cell plasma membrane. In addition, the HIV Env glycoproteins contain epitopes that elicit immune responses that are important from both diagnostic and vaccine development perspectives.

The HIV Env glycoprotein is synthesized from the singly spliced 4.3-kb Vpu/Env bicistronic mRNA (see **Fig. 4**); translation occurs on ribosomes associated with the rough endoplasmic reticulum (ER). The 160-kd polyprotein precursor (gp160) is an integral membrane protein that is anchored to cell membranes by a hydrophobic stop-transfer signal in the domain destined to be the mature TM Env glycoprotein, gp41 (**Fig. 6**). The gp160 is cotranslationally glycosylated, forms disulfide bonds, and undergoes oligomerization in the ER. The predominant oligomeric form appears to be a trimer, although dimers and tetramers are also observed. The gp160 is transported to the Golgi, where, like other retroviral envelope precursor proteins, it is proteolytically cleaved by cellular enzymes to the mature SU glycoprotein gp120 and TM glycoprotein gp41 (see **Fig. 6**). The cellular

enzyme responsible for cleavage of retroviral Env precursors following a highly conserved Lys/Arg-X-Lys/Arg-Arg motif is furin or a furin-like protease, although other enzymes may also catalyze gp160 processing. Cleavage of gp160 is required for Env-induced fusion activity and virus infectivity. Subsequent to gp160 cleavage, gp120 and gp41 form a noncovalent association that is critical for transport of the Env complex from the Golgi to the cell surface. The gp120-gp41 interaction is fairly weak, and a substantial amount of gp120 is shed from the surface of Env-expressing cells.

The HIV Env glycoprotein complex, in particular the SU (gp120) domain, is very heavily glycosylated; approximately half the molecular mass of gp160 is composed of oligosaccharide side chains. During transport of Env from its site of synthesis in the ER to the plasma membrane, many of the side chains are modified by the addition of complex sugars. The numerous oligosaccharide side chains form what could be imagined as a sugar cloud obscuring much of gp120 from host immune recognition. As shown in **Figure 6**, gp120 contains interspersed conserved (C₁ to C₅) and variable (V₁ to V₅) domains. The Cys residues present in the gp120s of different isolates are highly conserved and form disulfide bonds that link the first four variable regions in large loops.

A primary function of viral Env glycoproteins is to promote a membrane fusion reaction between the lipid bilayers of the viral envelope and host cell membranes. This membrane fusion event enables the viral core to gain entry into the host cell cytoplasm. A number of regions in both gp120 and gp41 have been implicated, directly or indirectly, in Env-mediated membrane fusion. Studies of the HA₂ hemagglutinin protein of the orthomyxoviruses and the F protein of the paramyxoviruses indicated that a highly hydrophobic domain at the N-terminus of these proteins, referred to as the fusion peptide, plays a critical role in membrane fusion. Mutational analyses demonstrated that an analogous domain was located at the N-terminus of the HIV-1, HIV-2, and SIV TM glycoproteins (see **Fig. 6**). Nonhydrophobic substitutions within this region of gp41 greatly reduced or blocked syncytium formation and resulted in the production of noninfectious progeny virions.

C-terminal to the gp41 fusion peptide are two amphipathic helical domains (see **Fig. 6**) which play a central role in membrane fusion. Mutations in the N-terminal helix (referred to as the N-helix), which contains a Leu zipper-like heptad repeat motif, impair infectivity and membrane fusion activity, and peptides derived from these sequences exhibit potent antiviral activity in culture. The structure of the ectodomain of HIV-1 and

SIV gp41, the two helical motifs in particular, has been the focus of structural analyses in recent years. Structures were determined by x-ray crystallography or NMR spectroscopy either for fusion proteins containing the helical domains, a mixture of peptides derived from the N- and C-helices, or in the case of the SIV structure, the intact gp41 ectodomain
5 sequence from residue 27 to 149. These studies obtained fundamentally similar trimeric structures, in which the two helical domains pack in an antiparallel fashion to generate a six-helix bundle. The N-helices form a coiled-coil in the center of the bundle, with the C-helices packing into hydrophobic grooves on the outside.

In the steps leading to membrane fusion CD4 binding induces conformation
10 changes in Env that facilitate coreceptor binding. Following the formation of a ternary gp120/CD4/coreceptor complex, gp41 adopts a hypothetical conformation that allows the fusion peptide to insert into the target lipid bilayer. The formation of the gp41 six-helix bundle (which involves antiparallel interactions between the gp41 N- and C-helices) brings the viral and cellular membranes together and membrane fusion takes place.

15 Use of Recombinant MVA Virus To Boost CD+8 Cell Immune Response

The present invention relates to generation of a CD8⁺ T cell immune response against an antigen and also eliciting an antibody response. More particularly, the present invention relates to "prime and boost" immunization regimes in which the immune response induced by administration of a priming composition is boosted by administration
20 of a boosting composition. The present invention is based on inventors' experimental demonstration that effective boosting can be achieved using modified vaccinia Ankara (MVA) vectors, following priming with any of a variety of different types of priming compositions including recombinant MVA itself.

A major protective component of the immune response against a number of
25 pathogens is mediated by T lymphocytes of the CD8⁺ type, also known as cytotoxic T lymphocytes (CTL). An important function of CD8⁺ cells is secretion of gamma interferon (IFN γ), and this provides a measure of CD8⁺ T cell immune response. A second component of the immune response is antibody directed to the proteins of the pathogen.

The present invention employs MVA which, as the experiments described below
30 show, has been found to be an effective means for providing a boost to a CD8⁺ T cell immune response primed to antigen using any of a variety of different priming compositions and also eliciting an antibody response.

Remarkably, the experimental work described below demonstrates that use of embodiments of the present invention allows for recombinant MVA virus expressing an HIV antigen to boost a CD8⁺ T cell immune response primed by a DNA vaccine and also eliciting an antibody response. The MVA was found to induce a CD8⁺ T cell response
5 after intradermal, intramuscular or mucosal immunization. Recombinant MVA has also been shown to prime an immune response that is boosted by one or more inoculations of recombinant MVA.

Non-human primates immunized with plasmid DNA and boosted with the MVA were effectively protected against intramucosal challenge with live virus. Advantageously,
10 the inventors found that a vaccination regime used intradermal, intramuscular or mucosal immunization for both prime and boost can be employed, constituting a general immunization regime suitable for inducing CD8⁺ T cells and also eliciting an antibody response, *e.g.* in humans.

The present invention in various aspects and embodiments employs an MVA vector
15 encoding an HIV antigen for boosting a CD8⁺ T cell immune response to the antigen primed by previous administration of nucleic acid encoding the antigen and also eliciting an antibody response.

A general aspect of the present invention provides for the use of an MVA vector for boosting a CD8⁺ T cell immune response to an HIV antigen and also eliciting an antibody
20 response.

One aspect of the present invention provides a method of boosting a CD8⁺ T cell immune response to an HIV antigen in an individual, and also eliciting an antibody response, the method including provision in the individual of an MVA vector including nucleic acid encoding the antigen operably linked to regulatory sequences for production of
25 antigen in the individual by expression from the nucleic acid, whereby a CD8⁺ T cell immune response to the antigen previously primed in the individual is boosted.

An immune response to an HIV antigen may be primed by immunization with plasmid DNA or by infection with an infectious agent.

A further aspect of the invention provides a method of inducing a CD8⁺ T cell
30 immune response to an HIV antigen in an individual, and also eliciting an antibody response, the method comprising administering to the individual a priming composition comprising nucleic acid encoding the antigen and then administering a boosting composition which comprises an MVA vector including nucleic acid encoding the antigen

operably linked to regulatory sequences for production of antigen in the individual by expression from the nucleic acid.

A further aspect provides for use of an MVA vector, as disclosed, in the manufacture of a medicament for administration to a mammal to boost a CD8⁺ T cell
5 immune response to an HIV antigen, and also eliciting an antibody response. Such a medicament is 'generally for administration following prior administration of a priming composition comprising nucleic acid encoding the antigen.

The priming composition may comprise any viral vector, such as a vaccinia virus vector such as a replication-deficient strain such as modified vaccinia Ankara (MVA) or
10 NYVAC (Tartaglia *et al.* 1992 *Virology* **118**:217-232), an avipox vector such as fowlpox or canarypox, *e.g.* the strain known as ALVAC (Paoletti *et al.* 1994 *Dev Biol Stand* **82**:65-69), or an adenovirus vector or a vesicular stomatitis virus vector or an alphavirus vector.

The priming composition may comprise DNA encoding the antigen, such DNA preferably being in the form of a circular plasmid that is not capable of replicating in
15 mammalian cells. Any selectable marker should not be resistance to an antibiotic used clinically, so for example Kanamycin resistance is preferred to Ampicillin resistance. Antigen expression should be driven by a promoter which is active in mammalian cells, for instance the cytomegalovirus immediate early (CMV IE) promoter.

In particular embodiments of the various aspects of the present invention,
20 administration of a priming composition is followed by boosting with a boosting composition, or first and second boosting compositions, the first and second boosting compositions being the same or different from one another. Still further boosting compositions may be employed without departing from the present invention. In one embodiment, a triple immunization regime employs DNA, then adenovirus as a first
25 boosting composition, then MVA as a second boosting composition, optionally followed by a further (third) boosting composition or subsequent boosting administration of one or other or both of the same or different vectors. Another option is DNA then MVA then adenovirus, optionally followed by subsequent boosting administration of one or other or both of the same or different vectors.

30 The antigen to be encoded in respective priming and boosting compositions (however many boosting compositions are employed) need not be identical, but should share at least one CD8⁺ T cell epitope. The antigen may correspond to a complete antigen, or a fragment thereof. Peptide epitopes or artificial strings of epitopes may be employed,

more efficiently cutting out unnecessary protein sequence in the antigen and encoding sequence in the vector or vectors. One or more additional epitopes may be included, for instance epitopes which are recognized by T helper cells, especially epitopes recognized in individuals of different HLA types.

5 An HIV antigen of the invention to be encoded by a recombinant MVA virus includes polypeptides having immunogenic activity elicited by an amino acid sequence of an HIV Env, Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, or Nef amino acid sequence as at least one CD8⁺ T cell epitope. This amino acid sequence substantially corresponds to at least one 10-900 amino acid fragment and/or consensus sequence of a known HIV Env or Pol; or at least
10 one 10-450 amino acid fragment and/or consensus sequence of a known HIV Gag; or at least one 10-100 amino acid fragment and/or consensus sequence of a known HIV Vif, Vpr, Tat, Rev, Vpu, or Nef.

 Although a full length Env precursor sequence is presented for use in the present invention, Env is optionally deleted of subsequences. For example, regions of the gp120
15 surface and gp41 transmembrane cleavage products can be deleted.

 Although a full length Gag precursor sequence is presented for use in the present invention, Gag is optionally deleted of subsequences. For example, regions of the matrix protein (p17), regions of the capsid protein (p24), regions of the nucleocapsid protein (p7), and regions of p6 (the C-terminal peptide of the Gag polyprotein) can be deleted.

20 Although a full length Pol precursor sequence is presented for use in the present invention, Pol is optionally deleted of subsequences. For example, regions of the protease protein (p10), regions of the reverse transcriptase protein (p66/p51), and regions of the integrase protein (p32) can be deleted.

 Such an HIV Env, Gag, or Pol can have overall identity of at least 50% to a known
25 Env, Gag, or Pol protein amino acid sequence, such as 50-99% identity, or any range or value therein, while eliciting an immunogenic response against at least one strain of an HIV.

 Percent identify can be determined, for example, by comparing sequence information using the GAP computer program, version 6.0, available from the University
30 of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (*J Mol Biol* 1970 48:443), as revised by Smith and Waterman (*Adv Appl Math* 1981 2:482). Briefly, the GAP program defines identity as the number of aligned symbols (*i.e.*, nucleotides or amino acids) which are

identical, divided by the total number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program include: (1) a unitary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov and Burgess (*Nucl Acids Res* 1986 14:6745), as described by Schwartz
5 and Dayhoff (eds., *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington, D.C. 1979, pp. 353-358); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

In a preferred embodiment, an Env of the present invention is a variant form of at least one HIV envelope protein. Preferably, the Env is composed of gp120 and the
10 membrane-spanning and ectodomain of gp41 but lacks part or all of the cytoplasmic domain of gp41.

Known HIV sequences are readily available from commercial and institutional HIV sequence databases, such as GENBANK, or as published compilations, such as Myers *et al.* eds., *Human Retroviruses and AIDS, A Compilation and Analysis of Nucleic Acid and*
15 *Amino Acid Sequences*, Vol. I and II, Theoretical Biology and Biophysics, Los Alamos, N. Mex. (1993), or <http://hiv-web.lanl.gov/>.

Substitutions or insertions of an HIV Env, Gag, or Pol to obtain an additional HIV Env, Gag, or Pol, encoded by a nucleic acid for use in a recombinant MVA virus of the present invention, can include substitutions or insertions of at least one amino acid residue
20 (*e.g.*, 1-25 amino acids). Alternatively, at least one amino acid (*e.g.*, 1-25 amino acids) can be deleted from an HIV Env, Gag, or Pol sequence. Preferably, such substitutions, insertions or deletions are identified based on safety features, expression levels, immunogenicity and compatibility with high replication rates of MVA.

Amino acid sequence variations in an HIV Env, Gag, or Pol of the present invention
25 can be prepared *e.g.*, by mutations in the DNA. Such HIV Env, Gag, or Pol include, for example, deletions, insertions or substitutions of nucleotides coding for different amino acid residues within the amino acid sequence. Obviously, mutations that will be made in nucleic acid encoding an HIV Env, Gag, or Pol must not place the sequence out of reading frame and preferably will not create complementary domains that could produce secondary
30 mRNA structures.

HIV Env, Gag, or Pol-encoding nucleic acid of the present invention can also be prepared by amplification or site-directed mutagenesis of nucleotides in DNA or RNA encoding an HIV Env, Gag, or Pol and thereafter synthesizing or reverse transcribing the

encoding DNA to produce DNA or RNA encoding an HIV Env, Gag, or Pol, based on the teaching and guidance presented herein.

Recombinant MVA viruses expressing HIV Env, Gag, or Pol of the present invention, include a finite set of HIV Env, Gag, or Pol-encoding sequences as substitution
5 nucleotides that can be routinely obtained by one of ordinary skill in the art, without undue experimentation, based on the teachings and guidance presented herein. For a detailed description of protein chemistry and structure, see Schulz, G.E. *et al.*, 1978 *Principles of Protein Structure*, Springer-Verlag, New York, N.Y., and Creighton, T.E., 1983 *Proteins: Structure and Molecular Properties*, W. H. Freeman & Co., San Francisco, CA. For a
10 presentation of nucleotide sequence substitutions, such as codon preferences, see Ausubel *et al.* eds. *Current Protocols in Molecular Biology*, Greene Publishing Assoc., New York, N.Y. 1994 at §§ A.1.1-A.1.24, and Sambrook, J. *et al.* 1989 *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. at Appendices C and D.

15 Thus, one of ordinary skill in the art, given the teachings and guidance presented herein, will know how to substitute other amino acid residues in other positions of an HIV *env*, *gag*, or *pol* DNA or RNA to obtain alternative HIV Env, Gag, or Pol, including substitutional, deletional or insertional variants.

Within the MVA vector, regulatory sequences for expression of the encoded antigen
20 will include a natural, modified or synthetic poxvirus promoter. By "promoter" is meant a sequence of nucleotides from which transcription may be initiated of DNA operably linked downstream (*i.e.* in the 3' direction on the sense strand of double-stranded DNA). "Operably linked" means joined as part of the same nucleic acid molecule, suitably positioned and oriented for transcription to be initiated from the promoter. DNA operably
25 linked to a promoter is "under transcriptional initiation regulation" of the promoter. Other regulatory sequences including terminator fragments, polyadenylation sequences, marker genes and other sequences may be included as appropriate, in accordance with the knowledge and practice of the ordinary person skilled in the art: see, for example, Moss, B. (2001). Poxviridae: the viruses and their replication. In Fields Virology, D.M. Knipe, and
30 P.M. Howley, eds. (Philadelphia, Lippincott Williams & Wilkins), pp. 2849-2883. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into

cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology*, 1998 Ausubel *et al.* eds., John Wiley & Sons.

Promoters for use in aspects and embodiments of the present invention must be compatible with poxvirus expression systems and include natural, modified and synthetic
5 sequences.

Either or both of the priming and boosting compositions may include an adjuvant, such as granulocyte macrophage-colony stimulating factor (GM-CSF) or encoding nucleic acid therefor.

Administration of the boosting composition is generally about 1 to 6 months after
10 administration of the priming composition, preferably about 1 to 3 months.

Preferably, administration of priming composition, boosting composition, or both priming and boosting compositions, is intradermal, intramuscular or mucosal immunization.

Administration of MVA vaccines may be achieved by using a needle to inject a
15 suspension of the virus. An alternative is the use of a needleless injection device to administer a virus suspension (using, *e.g.*, Biojector™ needleless injector) or a resuspended freeze-dried powder containing the vaccine, providing for manufacturing individually prepared doses that do not need cold storage. This would be a great advantage for a vaccine that is needed in rural areas of Africa.

20 MVA is a virus with an excellent safety record in human immunizations. The generation of recombinant viruses can be accomplished simply, and they can be manufactured reproducibly in large quantities. Intradermal, intramuscular or mucosal administration of recombinant MVA virus is therefore highly suitable for prophylactic or therapeutic vaccination of humans against AIDS which can be controlled by a CD8⁺ T cell
25 response.

The individual may have AIDS such that delivery of the antigen and generation of a CD8⁺ T cell immune response to the antigen is of benefit or has a therapeutically beneficial effect.

Most likely, administration will have prophylactic aim to generate an immune
30 response against HIV or AIDS before infection or development of symptoms.

Components to be administered in accordance with the present invention may be formulated in pharmaceutical compositions. These compositions may comprise a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well

known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may depend on the route of administration, *e.g.* intravenous, cutaneous or subcutaneous, nasal, intramuscular, intraperitoneal routes.

5 As noted, administration is preferably intradermal, intramuscular or mucosal.

Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous, subcutaneous, intramuscular or mucosal injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilizers, buffers, antioxidants and/or other additives may be included as required.

15 A slow-release formulation may be employed.

Following production of MVA particles and optional formulation of such particles into compositions, the particles may be administered to an individual, particularly human or other primate. Administration may be to another mammal, *e.g.* rodent such as mouse, rat or hamster, guinea pig, rabbit, sheep, goat, pig, horse, cow, donkey, dog or cat.

20 Administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, *e.g.* decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors, or in a veterinary context a veterinarian, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in *Remington's Pharmaceutical Sciences*, 16th edition, 25 1980, Osol, A. (ed.).

30 In one preferred regimen, DNA is administered at a dose of 250 μ g to 2.5 mg/injection, followed by MVA at a dose of 10^6 to 10^9 infectious virus particles/injection.

A composition may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated.

Delivery to a non-human mammal need not be for a therapeutic purpose, but may be for use in an experimental context, for instance in investigation of mechanisms of immune responses to an antigen of interest, *e.g.* protection against HIV or AIDS.

Further aspects and embodiments of the present invention will be apparent to those of ordinary skill in the art, in view of the above disclosure and following experimental exemplification, included by way of illustration and not limitation, and with reference to the attached figures.

EXAMPLE 1

Control of a Mucosal Challenge and Prevention of AIDS by a Multiprotein DNA/MVA Vaccine

Here we tested DNA priming and poxvirus boosting for the ability to protect against a highly pathogenic mucosal challenge. The 89.6 chimera of simian and human immunodeficiency viruses (SHIV-89.6) was used for the construction of immunogens and its highly pathogenic derivative, SHIV-89.6P, for challenge (G.B. Karlsson *et al.* 1997 *J Virol* **71**:4218). SHIV-89.6 and SHIV-89.6P do not generate cross-neutralizing antibody (D.C. Montefiori *et al.* 1998 *J Virol* **72**:3427) and allowed us to address the ability of vaccine-raised T cells and non-neutralizing antibodies to control an immunodeficiency virus challenge. Modified vaccinia Ankara (MVA) was used for the construction of the recombinant poxvirus. MVA has been highly effective at boosting DNA-primed CD8 T cells and enjoys the safety feature of not replicating efficiently in human or monkey cells (H.L. Robinson *et al.* 2000 *AIDS Reviews* **2**:105).

To ensure a broad immune response both the DNA and recombinant MVA (rMVA) components of the vaccine expressed multiple immunodeficiency virus proteins. The DNA prime (DNA/89.6) expressed simian immunodeficiency virus (SIV) Gag, Pol, Vif, Vpx, and Vpr and human immunodeficiency virus-1 (HIV-1) Env, Tat, and Rev from a single transcript (R.J. Gorelick *et al.* 1999 *Virology* **253**:259; M.M. Sauter *et al.* 1996 *J Cell Biol* **132**:795).

Molecularly cloned SHIV-89.6 sequences were cloned into the vector pGA2 using ClaI and RsrII sites. This cloning deleted both long terminal repeats (LTRs) and *nef*. The SHIV-89.6 sequences also were internally mutated for a 12-base pair region encoding the first four amino acids of the second zinc finger in nucleocapsid. This mutation renders

SHIV viruses noninfectious (R.J. Gorelick *et al.* 1999 *Virology* **253**:259). A mutation in gp41 converted the tyrosine at position 710 to cysteine to achieve better expression of Env on the plasma membrane of DNA-expressing cells (M.M. Sauter *et al.* 1996 *J Cell Biol* **132**:795). pGA2 uses the CMV immediate early promoter without intron A and the bovine growth hormone polyadenylation sequence to express vaccine inserts. Vaccine DNA was produced by Althea (San Diego, CA). In transient transfections of 293T cells, DNA/89.6 produced about 300 ng of Gag and 85 ng of Env per 1×10^6 cells.

The rMVA booster (MVA/89.6) expressed SIV Gag, Pol, and HIV-1 Env under the control of vaccinia virus early/late promoters.

The MVA double recombinant virus expressed both the HIV 89.6 Env and the SIV 239 Gag-Pol, which were inserted into deletion II and deletion III of MVA, respectively. The 89.6 Env protein was truncated for the COOH-terminal 115 amino acids of gp41. The modified H5 promoter controlled the expression of both foreign genes.

Vaccination was accomplished by priming with DNA at 0 and 8 weeks and boosting with rMVA at 24 weeks (Fig. 7A).

I.d. and i.m. DNA immunizations were delivered in phosphate-buffered saline (PBS) with a needleless jet injector (Bioject, Portland, OR) to deliver five i.d. 100- μ l injections to each outer thigh for the 2.5-mg dose of DNA or one i.d. 100- μ l injection to the right outer thigh for the 250- μ g dose of plasmid. I.m. deliveries of DNA were done with one 0.5-ml injection of DNA in PBS to each outer thigh for the 2.5-mg dose and one 100- μ l injection to the right outer thigh for the 250- μ g dose. 1×10^8 pfu of MVA/89.6 was administered both i.d. and i.m. with a needle. One 100- μ l dose was delivered to each outer thigh for the i.d. dose and one 500- μ l dose to each outer thigh for the i.m. dose. Control animals received 2.5 mg of the pGA2 vector without vaccine insert with the Bioject device to deliver five 100- μ l doses i.d. to each outer thigh. The control MVA booster immunization consisted of 2×10^8 pfu of MVA without an insert delivered i.d. and i.m. as described for MVA/89.6.

Four groups of six rhesus macaques each were primed with either 2.5 mg (high-dose) or 250 μ g (low-dose) of DNA by intradermal (i.d.) or intramuscular (i.m.) routes using a needleless jet injection device (Bioject, Portland, OR) (T.M. Allen *et al.* 2000 *J Immunol* **164**:4968).

Young adult rhesus macaques from the Yerkes breeding colony were cared for under guidelines established by the Animal Welfare Act and the NIH "Guide for the Care

and Use of Laboratory Animals” with protocols approved by the Emory University Institutional Animal Care and Use Committee. Macaques were typed for the *Mamu-A*01* allele with polymerase chain reaction (PCR) analyses (M.A. Egan *et al.* 2000 *J Virol* 74:7485; I. Ourmanov *et al.* 2000 *J Virol* 74:2740). Two or more animals containing at least one *Mamu-A*01* allele were assigned to each group. Animal numbers are as follows:

1, RBr-5*; 2, RIm-5*; 3, RQf-5*; 4, RZe-5; 5, ROm-5; 6, RDm-5; 7, RAj-5*; 8, Rji-5*; 9, RAl-5*; 10, RDe-5*; 11, RAi-5; 12, RPr-5; 13, RKw-4*; 14, RWz-5*; 15, RGo-5; 16, RLp-4; 17, RWd-6; 18, RAf-5; 19, RPb-5*; 20, Rli-5*; 21, RIq-5; 22, RSp-4; 23, RSn-5; 24, RGd-6; 25, RMb-5*; 26, RGy-5*; 27, RUf-4; and 28, RPm-5. Animals with the *A*01* allele are indicated with asterisks.

Gene gun deliveries of DNA were not used because these had primed non-protective immune responses in a 1996 - 98 trial (H.L. Robinson *et al.* 1999 *Nat Med* 5:526). The MVA/89.6 booster immunization (2×10^8 plaque-forming units, pfu) was injected with a needle both i.d. and i.m. A control group included two mock immunized animals and two naive animals. The challenge was given at 7 months after the rMVA booster to test for the generation of long-term immunity. Because most HIV-1 infections are transmitted across mucosal surfaces, an intrarectal challenge was administered.

DNA priming followed by rMVA boosting generated high frequencies of virus-specific T cells that peaked at one week following the rMVA booster (Fig. 7). The frequencies of T cells recognizing the Gag-CM9 epitope were assessed by means of Mamu-A*01 tetramers, and the frequencies of T cells recognizing epitopes throughout Gag were assessed with pools of overlapping peptides and an enzyme-linked immunospot (ELISPOT) assay (C.A. Power *et al.* 1999 *J Immunol Methods* 227:99).

For tetramer analyses, about 1×10^6 peripheral blood mononuclear cells (PBMC) were surface-stained with antibodies to CD3 conjugated to fluorescein isothiocyanate (FITC) (FN-18; Biosource International, Camarillo, CA), CD8 conjugated to peridinin chlorophyll protein (PerCP) (SK1; Becton Dickinson, San Jose, CA), and Gag-CM9 (CTPYDINQM)-*Mamu-A*01* tetramer (SEQ ID NO: 6) conjugated to allophycocyanin (APC), in a volume of 100 μ l at 8° to 10°C for 30 min. Cells were washed twice with cold PBS containing 2% fetal bovine serum (FBS), fixed with 1% paraformaldehyde in PBS, and analyzed within 24 hrs on a FACScaliber (Becton Dickinson, San Jose, CA). Cells were initially gated on lymphocyte populations with forward scatter and side scatter and then on CD3 cells. The CD3 cells were then analyzed for CD8 and tetramer-binding cells.

About 150,000 lymphocytes were acquired for each sample. Data were analyzed using FloJo software (Tree Star, San Carlos, CA).

For interferon- γ (IFN- γ) ELISPOTs, MULTISCREEN 96 well filtration plates (Millipore Inc. Bedford, MA) were coated overnight with antibody to human IFN- γ (Clone B27, Pharmingen, San Diego, CA) at a concentration of 2 μ g/ml in sodium bicarbonate buffer (pH 9.6) at 8° to 10°C. Plates were washed two times with RPMI medium and then blocked for 1 hour with complete medium (RPMI containing 10% FBS) at 37°C. Plates were washed five more times with plain RPMI medium, and cells were seeded in duplicate in 100 μ l complete medium at numbers ranging from 2×10^4 to 5×10^5 cells per well. Peptide pools were added to each well to a final concentration of 2 μ g/ml of each peptide in a volume of 100 μ l in complete medium. Cells were cultured at 37°C for about 36 hrs under 5% CO₂. Plates were washed six times with wash buffer (PBS with 0.05% Tween-20) and then incubated with 1 μ g of biotinylated antibody to human IFN- γ per milliliter (clone 7-86-1; Diapharma Group, West Chester, OH) diluted in wash buffer containing 2% FBS. Plates were incubated for 2 hrs at 37°C and washed six times with wash buffer. Avidin-horseradish peroxidase (Vector Laboratories, Burlingame, CA) was added to each well and incubated for 30 to 60 min at 37°C. Plates were washed six times with wash buffer and spots were developed using stable DAB as substrate (Research Genetics, Huntsville, AL). Spots were counted with a stereo dissecting microscope. An ovalbumin peptide (SIINFEKL) (SEQ ID NO: 7) was included as a control in each analysis. Background spots for the ovalbumin peptide were generally <5 for 5×10^5 PBMCs. This background when normalized for 1×10^6 PBMC was <10. Only ELISPOT counts of twice the background (≥ 20) were considered significant. The frequencies of ELISPOTs are approximate because different dilutions of cells have different efficiencies of spot formation in the absence of feeder cells (C.A. Power *et al.* 1999 *J Immunol Methods* 227: 99). The same dilution of cells was used for all animals at a given time point, but different dilutions were used to detect memory and acute responses.

Gag-CM9 tetramer analyses were restricted to macaques that expressed the *Mamu-A*01* histocompatibility type, whereas ELISPOT responses did not depend on a specific histocompatibility type. As expected, the DNA immunizations raised low levels of memory cells that expanded to high frequencies within 1 week of the rMVA booster (Fig. 7 and 12). In *Mamu-A*01* macaques, CD8 cells specific to the Gag-CM9 epitope expanded to frequencies as high as 19% of total CD8 T cells (Fig. 12). This peak of specific cells

underwent a 10- to 100-fold contraction into the DNA/MVA memory pool (**Fig. 7A and 12**). ELISPOTs for three pools of Gag peptides also underwent a major expansion (frequencies up to 4000 spots for 1×10^6 PBMC) before contracting from 5- to 20-fold into the DNA/MVA memory response (**Fig. 7B**). The frequencies of ELISPOTs were the same
5 in macaques with and without the *A*01* histocompatibility type ($P > 0.2$).

Simple linear regression was used to estimate correlations between postbooster and postchallenge ELISPOT responses, between memory and postchallenge ELISPOT responses, and between logarithmically transformed viral loads and ELISPOT frequencies. Comparisons between vaccine and control groups and *A*01* and non *A*01* macaques were
10 performed by means of two-sample *t* tests with logarithmically transformed viral load and ELISPOT responses. Two-way analyses of variance were used to examine the effects of dose and route of administration on peak DNA/MVA ELISPOTs, on memory DNA/MVA ELISPOTs, and on logarithmically transformed Gag antibody data.

At both peak and memory phases of the vaccine response, the rank order for the
15 height of the ELISPOTs in the vaccine groups was 2.5 mg i.d. > 2.5 mg i.m. > 250 μ g i.d. > 250 μ g i.m. (**Fig. 7B**). The IFN- γ ELISPOTs included both CD4 and CD8 cells. Gag-CM9-specific CD8 cells had good lytic activity after restimulation with peptide.

The highly pathogenic SHIV-89.6P challenge was administered intrarectally at 7 months after the rMVA booster, when vaccine-raised T cells were in memory (**Fig. 7**).

20 The challenge stock (5.7×10^9 copies of viral RNA per milliliter) was produced by one intravenous followed by one intrarectal passage in rhesus macaques of the original SHIV-89.6P stock (G.B. Karlsson *et al.* 1997 *J Virol* 71:4218). Lymphoid cells were harvested from the intrarectally infected animal at peak viremia, CD8-depleted, and mitogen-stimulated for stock production. Before intrarectal challenge, fasted animals were
25 anesthetized (ketamine, 10 mg/kg) and placed on their stomach with the pelvic region slightly elevated. A feeding tube (8Fr (2.7 mm) x 16 inches (41 cm); Sherwood Medical, St. Louis, MO) was inserted into the rectum for a distance of 15 to 20 cm. Following insertion of the feeding tube, a syringe containing 20 intrarectal infectious doses in 2 ml of RPMI-1640 plus 10% FBS was attached to the tube and the inoculum was slowly injected
30 into the rectum. After delivery of the inoculum, the feeding tube was flushed with 3.0 ml of RPMI without FBS and then slowly withdrawn. Animals were left in place, with pelvic regions slightly elevated, for a period of ten minutes after the challenge.

The challenge infected all of the vaccinated and control animals (Fig. 8). However, by 2 weeks after challenge, titers of plasma viral RNA were at least 10-fold lower in the vaccine groups (geometric means of 1×10^7 to 5×10^7) than in the control animals (geometric mean of 4×10^8) (Fig. 8A) (S. Staprans *et al.* in: *Viral Genome Methods* K. Adolph, ed. CRC Press, Boca Raton, FL, 1996 pp. 167-184; R. Hofmann-Lehmann *et al.* 2000 *AIDS Res Hum Retroviruses* 16:1247).

For the determination of SHIV copy number, viral RNA from 150 μ l of ACD anticoagulated plasma was directly extracted with the QIAamp Viral RNA kit (Qiagen), eluted in 60 μ l of AVE buffer, and frozen at -80°C until SHIV RNA quantitation was performed. Five microliters of purified plasma RNA was reverse-transcribed in a final 20- μ l volume containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 4 mM MgCl_2 , 1 mM each deoxynucleotide triphosphate (dNTP), 2.5 μ M random hexamers, 20 units MultiScribe RT, and 8 units ribonuclease inhibitor. Reactions were incubated at 25°C for 10 min, followed by incubation at 42°C for 20 min, and inactivation of reverse transcriptase at 99°C for 5 min. The reaction mix was adjusted to a final volume of 50 μ l containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 4 mM MgCl_2 , 0.4 mM each dNTP, 0.2 μ M forward primer, 0.2 μ M reverse primer, 0.1 μ M probe, and 5 units AmpliTaq Gold DNA polymerase (all reagents from PerkinElmer Applied Biosystems, Foster City, CA). The primer sequences within a conserved portion of the SIV *gag* gene are the same as those described previously (S. Staprans *et al.* in: *Viral Genome Methods* K. Adolph, ed. CRC Press, Boca Raton, FL, 1996 pp. 167-184). A PerkinElmer Applied Biosystems 7700 Sequence Detection System was used with the PCR profile: 95°C for 10 min, followed by 40 cycles at 93°C for 30 s, and 59.5°C for 1 min. PCR product accumulation was monitored with the 7700 sequence detector and a probe to an internal conserved *gag* gene sequence: 6FAM-CTGTCTGCGTCATTTGGTGC-Tamra (SEQ ID NO: 8), where FAM and Tamra denote the reporter and quencher dyes. SHIV RNA copy number was determined by comparison with an external standard curve consisting of virion-derived SIVmac239 RNA quantified by the SIV bDNA method (Bayer Diagnostics, Emeryville, CA). All specimens were extracted and amplified in duplicate, with the mean result reported. With a 0.15-ml plasma input, the assay has a sensitivity of 10^3 RNA copies per milliliter of plasma and a linear dynamic range of 10^3 to 10^8 RNA copies ($R^2 = 0.995$). The intraassay coefficient of variation was $<20\%$ for samples containing $>10^4$ SHIV RNA copies per milliliter, and $<25\%$ for samples containing 10^3 to 10^4 SHIV RNA copies per milliliter. To more

accurately quantitate low SHIV RNA copy number in vaccinated animals at weeks 16 and 20, we made the following modifications to increase the sensitivity of the SHIV RNA assay: (i) Virions from ≤ 1 ml of plasma were concentrated by centrifugation at 23,000g at 10°C for 150 min before viral RNA extraction, and (ii) a one-step reverse transcriptase PCR method was used (R. Hofmann-Lehmann *et al.* 2000 *AIDS Res Hum Retroviruses* 16:1247). These changes provided a reliable quantification limit of 300 SHIV RNA copies per milliliter, and gave SHIV RNA values that were highly correlated to those obtained by the first method used ($r = 0.91$, $P < 0.0001$).

By 8 weeks after challenge, both high-dose DNA-primed groups and the low-dose i.d. DNA-primed group had reduced their geometric mean loads to about 1000 copies of viral RNA per milliliter. At this time, the low-dose i.m. DNA-primed group had a geometric mean of 6×10^3 copies of viral RNA and the nonvaccinated controls had a geometric mean of 2×10^6 . By 20 weeks after challenge, even the low-dose i.m. group had reduced its geometric mean copies of viral RNA to 1000. Among the 24 vaccinated animals, only one animal, animal number 22 in the low-dose i.m. group, had intermittent viral loads above 1×10^4 copies per milliliter (Fig 8D).

By 5 weeks after challenge, all of the nonvaccinated controls had undergone a profound depletion of CD4 cells (Fig 8B). All of the vaccinated animals maintained their CD4 cells, with the exception of animal 22 in the low dose i.m. group (see above), which underwent a slow CD4 decline (Fig. 8E). By 23 weeks after challenge, three of the four control animals had succumbed to AIDS (Fig. 8C). These animals had variable degrees of enterocolitis with diarrhea, cryptosporidiosis, colicystitis, enteric campylobacter infection, splenomegaly, lymphadenopathy, and SIV-associated giant cell pneumonia. In contrast, all 24 vaccinated animals maintained their health.

Containment of the viral challenge was associated with a burst of antiviral T cells (Fig. 7 and 9A). At one week after challenge, the frequency of tetramer⁺ cells in the peripheral blood had decreased, potentially reflecting the recruitment of specific T cells to the site of infection (Fig. 9A). However, by two weeks after challenge, tetramer⁺ cells in the peripheral blood had expanded to frequencies as high as, or higher than, after the rMVA booster (Fig. 7 and 9A). The majority of the tetramer⁺ cells produced IFN- γ in response to a 6-hour peptide stimulation (Fig. 9B) (S.L. Waldrop *et al.* 1997 *J Clin Invest* 99:1739) and did not have the “stunned” IFN- γ negative phenotype sometimes observed in viral infections (F. Lechner *et al.* 2000 *J Exp Med* 191:1499).

For intracellular cytokine assays, about 1×10^6 PBMC were stimulated for 1 hour at 37°C in 5 ml polypropylene tubes with 100 μg of Gag-CM9 peptide (CTPYDINQM) (SEQ ID NO: 6) per milliliter in a volume of 100 μl RPMI containing 0.1% bovine serum albumin (BSA) and 1 μg of antibody to human CD28 and 1 μg of antibody to human CD49d (Pharmingen, San Diego, CA) per milliliter. Then, 900 μl of RPMI containing 10% FBS and monensin (10 $\mu\text{g}/\text{ml}$) was added, and the cells were cultured for an additional 5 hrs at 37°C at an angle of 5° under 5% CO_2 . Cells were surface stained with antibodies to CD8 conjugated to PerCP (clone SK1, Becton Dickinson) at 8° to 10°C for 30 min, washed twice with cold PBS containing 2% FBS, and fixed and permeabilized with Cytofix/Cytoperm solution (Pharmingen). Cells were then incubated with antibodies to human CD3 (clone FN-18; Biosource International, Camarillo, CA) and IFN- γ (Clone B27; Pharmingen) conjugated to FITC and phycoerythrin, respectively, in Perm wash solution (Pharmingen) for 30 min at 4°C . Cells were washed twice with Perm wash, once with plain PBS, and resuspended in 1% paraformaldehyde in PBS. About 150,000 lymphocytes were acquired on the FACScaliber and analyzed with FloJo software.

The postchallenge burst of T cells contracted concomitant with the decline of the viral load. By 12 weeks after challenge, virus-specific T cells were present at about one-tenth of their peak height (Figs. 7A and 9A). In contrast to the vigorous secondary response in the vaccinated animals, the naive animals mounted a modest primary response (Figs. 7B and 9A). Tetramer⁺ cells peaked at less than 1% of total CD8 cells (Fig. 9A), and IFN- γ -producing ELISPOTs were present at a mean frequency of about 300 as opposed to the much higher frequencies of 1000 to 6000 in the vaccine groups (Fig. 7B) ($P < 0.05$).

The tetramer⁺ cells in the control group, like those in the vaccine group, produced IFN- γ after peptide stimulation (Fig. 9B). By 12 weeks after challenge, three of the four controls had undetectable levels of IFN- γ -producing ELISPOTs. This rapid loss of antiviral T cells in the presence of high viral loads may reflect the lack of CD4 help.

T cell proliferative responses demonstrated that virus-specific CD4 cells had survived the challenge and were available to support the antiviral immune response (Fig. 9C).

About 0.2 million PBMC were stimulated in triplicate for 5 days with the indicated antigen in 200 μl of RPMI at 37°C under 5% CO_2 . Supernatants from 293T cells transfected with DNA expressing either SHIV-89.6 Gag and Pol or SHIV-89.6 Gag, Pol and Env were used directly as antigens (final concentration of ~ 0.5 μg of p27 Gag per

milliliter). Supernatants from mock DNA (vector alone)-transfected cells served as negative controls. On day six, cells were pulsed with 1 μ Ci of tritiated thymidine per well for 16 to 20 hours. Cells were harvested with an automated cell harvester (TOMTEC, Harvester 96, Model 1010, Hamden, CT) and counted with a Wallac 1450 MICROBETA
5 Scintillation counter (Gaithersburg, MD). Stimulation indices are the counts of tritiated thymidine incorporated in PBMC stimulated with 89.6 antigens divided by the counts of tritiated thymidine incorporated by the same PBMC stimulated with mock antigen.

At 12 weeks after challenge, mean stimulation indices for Gag-Pol-Env or Gag-Pol proteins ranged from 35 to 14 in the vaccine groups but were undetectable in the control
10 group. Consistent with the proliferation assays, intracellular cytokine assays demonstrated the presence of virus-specific CD4 cells in vaccinated but not control animals. The overall rank order of the vaccine groups for the magnitude of the proliferative response was 2.5 mg i.d. > 2.5 mg i.m. > 250 μ g i.d. > 250 μ g i.m.

At 12 weeks after challenge, lymph nodes from the vaccinated animals were
15 morphologically intact and responding to the infection, whereas those from the infected controls had been functionally destroyed (**Fig. 10**). Nodes from vaccinated animals contained large numbers of reactive secondary follicles with expanded germinal centers and discrete dark and light zones (**Fig. 10A**). By contrast, lymph nodes from the non-vaccinated control animals showed follicular and paracortical depletion (**Fig. 10B**), while
20 those from unvaccinated and unchallenged animals displayed normal numbers of minimally reactive germinal centers (**Fig. 10C**). Germinal centers occupied < 0.05% of total lymph node area in the infected controls, 2% of the lymph node area in the uninfected controls, and up to 18% of the lymph node area in the vaccinated groups (**Fig. 10D**). More vigorous immune reactivity in the low-dose than the high-dose DNA-primed animals was suggested
25 by more extensive germinal centers in the low dose group (**Fig. 10D**). At 12 weeks after challenge, *in situ* hybridization for viral RNA revealed rare virus-expressing cells in lymph nodes from 3 of the 24 vaccinated macaques, whereas virus-expressing cells were readily detected in lymph nodes from each of the infected control animals. In the controls, which had undergone a profound depletion in CD4 T cells, the cytomorphology of infected lymph
30 node cells was consistent with a macrophage phenotype.

The prime/boost strategy raised low levels of antibody to Gag and undetectable levels of antibody to Env (**Fig. 11**). Postchallenge, antibodies to both Env and Gag

underwent anamnestic responses with total Gag antibody reaching heights approaching 1 mg/ml and total Env antibody reaching heights of up to 100 µg/ml.

Enzyme-linked immunosorbent assays (ELISAs) for total antibody to Gag used bacterially produced SIV gag p27 to coat wells (2 µg per milliliter in bicarbonate buffer). ELISAs for antibody to Env antibody used 89.6 Env produced in transiently transfected 293T cells and captured with sheep antibody against Env (catalog number 6205; International Enzymes, Fairbrook CA). Standard curves for Gag and Env ELISAs were produced with serum from a SHIV-89.6-infected macaque with known amounts of immunoglobulin G (IgG) specific for Gag or Env. Bound antibody was detected with peroxidase-conjugated goat antibody to macaque IgG (catalog # YNGMOIGGFCP; Accurate Chemical, Westbury, NY) and TMB substrate (Catalog # T3405; Sigma, St. Louis, MO). Sera were assayed at threefold dilutions in duplicate wells. Dilutions of test sera were performed in whey buffer (4% whey and 0.1% tween 20 in 1X PBS). Blocking buffer consisted of whey buffer plus 0.5% nonfat dry milk. Reactions were stopped with 2M H₂SO₄ and the optical density read at 450 nm. Standard curves were fitted and sample concentrations were interpolated as µg of antibody per ml of serum using SOFTmax 2.3 software (Molecular Devices, Sunnyvale, CA).

By 2 weeks after challenge, neutralizing antibodies for the 89.6 immunogen, but not the SHIV-89.6P challenge, were present in the high-dose DNA-primed groups (geometric mean titers of 352 in the i.d. and 303 in the i.m. groups) (Fig. 11C) (D.C. Montefiori *et al.* 1988 *J Clin Microbiol* 26:231). By 5 weeks after challenge, neutralizing antibody to 89.6P had been generated (geometric mean titers of 200 in the high-dose i.d. and 126 in the high-dose i.m. group) (Fig. 11D) and neutralizing antibody to 89.6 had started to decline. By 16 to 20 weeks after challenge, antibodies to Gag and Env had fallen in most animals.

Our results demonstrate that a multiprotein DNA/MVA vaccine can raise a memory immune response capable of controlling a highly virulent mucosal immunodeficiency virus challenge. Our levels of viral control were more favorable than have been achieved using only DNA (M.A. Egan *et al.* 2000 *J Virol* 74:7485) or rMVA vaccines (I. Ourmanov *et al.* 2000 *J Virol* 74:2740) and were comparable to those obtained for DNA immunizations adjuvanted with interleukin-2 (D.H. Barouch *et al.* 2000 *Science* 290:486). All of these previous studies have used more than three vaccine inoculations, none have used mucosal challenges, and most have challenged at peak effector responses and not allowed a prolonged post vaccination period to test for "long term" efficacy.

The dose of DNA had statistically significant effects on both cellular and humoral responses ($P < 0.05$), whereas the route of DNA administration affected only humoral responses. Intradermal DNA delivery was about 10 times more effective than i.m. inoculations for generating antibody to Gag ($P = 0.02$). Neither route nor dose of DNA appeared to have a significant effect on protection. At 20 weeks after challenge, the high-dose DNA-primed animals had slightly lower geometric mean levels of viral RNA (7×10^2 and 5×10^2) than the low-dose DNA-primed animals (9×10^2 and 1×10^3).

The DNA/MVA vaccine controlled the infection, rapidly reducing viral loads to near or below 1000 copies of viral RNA per milliliter of blood. Containment, rather than prevention of infection, affords the opportunity to establish a chronic infection (H.L. Robinson *et al.* 1999 *Nat Med* 5:526). By rapidly reducing viral loads, a multiprotein DNA/MVA vaccine will extend the prospect for long-term non-progression and limit HIV transmission (J.W. Mellors *et al.* 1996 *Science* 272:1167; T.C. Quinn *et al.* 2000 *N Engl J Med* 342:921).

EXAMPLE 2

MVA Expressing Modified HIV Env, Gag, and Pol Genes

This disclosure describes the construction of a modified vaccinia Ankara (MVA) recombinant virus, MVA/HIV clade B recombinant virus expressing the HIV strain ADA env and the HXB2 gag pol (MVA/HIV ADA env + HXB2 gag pol). For amplification, the lab name of MVA/HIV 48 will be used, which denotes the plasmid from which the construct comes.

The HIV *gag-pol* genes were derived from the Clade B infectious HXB2 virus. The *gag-pol* gene was truncated so that most of the integrase coding sequences were removed and amino acids 185, 266, and 478 were mutated to inactivate reverse transcriptase, inhibit strand transfer activity, and inhibit the RNaseH activity, respectively. The Clade B CCR5 tropic envelope gene was derived from the primary ADA isolate; TTTTNT (SEQ ID NO: 14) sequences were mutated without changing coding capacity to prevent premature transcription termination and the cytoplasmic tail was truncated in order to improve surface expression, immunogenicity, and stability of the MVA vector. The HIV genes were inserted into a plasmid transfer vector so that *gag-pol* gene was regulated by the modified H5 early/late vaccinia virus promoter and the *env* gene was regulated by the newly designed early/late Psyn II promoter to provide similar high levels of expression. A self-deleting GUS reporter gene was included to allow detection and isolation of the

recombinant virus. The HIV genes were flanked by MVA sequences to allow homologous recombination into the deletion 3 site so that the recombinant MVA would remain TK positive for stability and high expression in resting cells. The recombinant MVA was isolated and shown to express abundant amounts of *gag-pol-env* and to process *gag*.

5 Production of HIV-like particles was demonstrated by centrifugation and by electron microscopy. The presence of *env* in the HIV-like particles was demonstrated by immunoelectron microscopy.

Table of Sequences

Description	SEQ ID NO	FIG. NO
pLW-48 map	N/A	13
pLW-48 sequence	1	14
Psyn II promoter	2	14
ADA envelope truncated	3	14
PmH5 promoter	4	14
HXB2 <i>gag pol</i>	5	14

10 Plasmid Transfer Vector

The plasmid transfer vector used to make the MVA recombinant virus, pLW-48, (**Figure 15**) by homologous recombination was constructed as follows:

1. From the commercially obtained plasmid, pGem-4Z (Promega), flanking areas on either side of deletion III, designated flank 1 and flank 2, containing 926 and 520 base pairs
 15 respectively, were amplified by PCR from the MVA stains of vaccinia virus. Within these flanks, a promoter, the mH5, which had been modified from the originally published sequence by changing two bases that had been shown by previously published work to increase the expression of the cloned gene, was added.

2. A clade B *gag pol* (**Figure 16**) was truncated so that the integrase was removed
 20 and was cloned into the plasmid so that it was controlled by the mH5 promoter. This gene contained the complete HXB2 sequence of the *gag*. The *pol* gene has reverse transcriptase safety mutations in amino acid 185 within the active site of RT, in amino acid 266 which inhibits strand transfer activity, and at amino acid 478 which inhibits the RNaseH activity. In addition, the integrase gene was deleted past EcoRI site.

25 3. A direct repeat of 280 basepairs, corresponding to the last 280 base pairs of MVA flank 1, was added after flank 1.

4. The p11 promoter and GUS reporter gene were added between the two direct repeats of flank 1 so that this screening marker could initially be used for obtaining the recombinant virus, yet deleted out in the final recombinant virus (Scheifflinger, F. *et al.* 1998 *Arch Virol* **143**:467-474; Carroll, M.W. and B. Moss 1995 *BioTechniques* **19**:352-355).

5. A new promoter, Psyn II, was designed to allow for increased expression of the ADA env. The sequence of this new early/late promoter is given in **Figure 17**.

6. A truncated version of the ADA envelope with a silent 5TNT mutation was obtained by PCR and inserted in the plasmid under the control of the Psyn II promoter. The envelope was truncated in the cytoplasmic tail of the gp41 gene, deleting 115 amino acids of the cytoplasmic tail. This truncation was shown to increase the amount of envelope protein on the surface of infected cells and enhance immunogenicity of the envelope protein in mice, and stability of the recombinant virus in tissue culture.

Recombinant MVA Construction

1. MVA virus, which may be obtained from ATCC Number VR-1508, was plaque purified three times by terminal dilutions in chicken embryo fibroblasts (CEF), which were made from 9 day old SPF Premium SPAFAS fertile chicken eggs, distributed by B and E Eggs, Stevens, PA.

2. Secondary CEF cells were infected at an MOI of 0.05 of MVA and transfected with 2 µg of pLW-48, the plasmid described above. Following a two-day incubation at 37°C, the virus was harvested, frozen and thawed 3x, and plated out on CEF plates.

3. At 4 days, those foci of infection that stained blue after addition of X-gluc substrate, indicating that recombination had occurred between the plasmid and the infecting virus, were picked and inoculated on CEF plates. Again, those foci that stained blue were picked.

4. These GUS containing foci were plated out in triplicate and analyzed for GUS staining (which we wanted to now delete) and ADA envelope expression. Individual foci were picked from the 3rd replicate plates of those samples that had about equal numbers of mixed populations of GUS staining and nonstaining foci as well as mostly envelope staining foci.

5. These foci were again plated out in triplicate, and analyzed the same way. After 5 passages, a virus was derived which expressed the envelope protein but which had

deleted the GUS gene because of the double repeat. By immunostaining, this virus also expressed the gag pol protein.

Characterization of MVA Recombinant Virus, MVA/HIV 48

1. Aliquots of MVA/HIV 48 infected cell lysates were analyzed by radioimmunoprecipitation and immunostaining with monoclonal antibodies for expression of both the envelope and gag pol protein. In both of these tests, each of these proteins was detected.

2. The recombinant virus was shown to produce gag particles in the supernatant of infected cells by pelleting the ³⁵S-labeled particles on a 20% sucrose cushion.

3. Gag particles were also visualized both outside and budding from cells as well as within vacuoles of cells in the electron microscope in thin sections. These gag particles had envelope protein on their surface.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer, and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Summary

In summary, we have made a recombinant MVA virus, MVA/HIV 48, which has high expression of the ADA truncated envelope and the HXB2 gag pol. The MVA recombinant virus is made using a transiently expressed GUS marker that is deleted in the final virus. High expression of the ADA envelope is possible because of a new hybrid early/late promoter, Psyn II. In addition, the envelope has been truncated because we have

shown truncation of the envelope enhances the amount of protein on the surface of the infected cells, and hence enhances immunogenicity; stability of the recombinant is also enhanced. The MVA recombinant makes gag particles which have been shown by pelleting the particles through sucrose and analyzing by PAGE. Gag particles with
5 envelope protein on the surface have also been visualized in the electron microscope.

EXAMPLE 3

Additional Modified or Synthetic Promoters Designed for Gene Expression in MVA Or Other Poxviruses

Additional modified or synthetic promoters were designed for gene expression in
10 MVA or other poxviruses. Promoters were modified to allow expression at early and late times after infection and to reduce possibility of homologous recombination between identical sequences when multiple promoters are used in same MVA vector. Promoters are placed upstream of protein coding sequence.

m7.5 promoter (SEQ ID NO: 10):

15 CGCTTTTATAGTAAGTTTTTCACCCATAAATAATAAATACAATAATTAATTTCT
CGTAAAAATTGAAAAACTATTCTAATTTATTGCACGGT

Psyn II promoter (SEQ ID NO: 2):

TAAAAAATGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTTCT
ATGCTATAAATAATAAATA

Psyn III promoter (SEQ ID NO: 11):

20 TAAAAATTGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTTCT
ATACTATAAATAATAAATA

Psyn IV promoter (SEQ ID NO: 12):

25 TAAAAATTGAAAAACTATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTTCT
ATACTATAAATAATAAATA

Psyn V promoter (SEQ ID NO: 13):

AAAAAATGATAAAGTAGGTTTCAGTTTTATTGCTGGTTTAAAATCACGCTTTCGA
GTAAAAACTACGAATATAAAT

EXAMPLE 4

Tables A-F

Table A: MVA/48 immunization – guinea pigs.

Groups of guinea pigs were immunized at days 0 and 30 with 1×10^8 infectious units of MVA/48 by either the intramuscular (IM) or intradermal (ID) route. As a control another group was immunized IM with the same dose of non-recombinant MVA. Sera taken before as well as after each immunization were analyzed for neutralizing activity against HIV-1-MN. Titers are the reciprocal serum dilution at which 50% of MT-2 cells were protected from virus-induced killing. Significant neutralizing activity was observed in all animals after the second immunization with MVA/48 (day 49).

Table B: Frequencies of HIV-1 gag-specific T cells following immunization of mice with MVA/48.

Groups of BalbC mice were immunized at days 0 and 21 with 1×10^7 infectious units of MVA/48 by one of three routes: intraperitoneal (IP), intradermal (ID), or intramuscular (IM). A control group was immunized with non-recombinant MVA. At 5 weeks after the last immunization, splenocytes were prepared and stimulated *in vitro* with an immunodominant peptide from HIV-1 p24 for 7 days. The cells were then mixed either with peptide-pulsed P815 cells or with soluble peptide. Gamma interferon-producing cells were enumerated in an ELISPOT assay. A value of >500 was assigned to wells containing too many spots to count. Strong T cell responses have been reported in mice immunized IP with other viruses. In this experiment, IP immunization of mice with MVA/48 elicited very strong HIV-1 gag-specific T cell responses.

Table C: DNA prime and MVA/48 boost – total ELISPOTS per animal.

Ten rhesus macaques were primed (weeks 0 and 8) with a DNA vaccine expressing HIV-1 antigens including Ada envelope and HXB2 gagpol. At week 24 the animals were boosted intramuscularly with 1×10^8 infectious units of MVA/48. Fresh peripheral blood mononuclear cells (PBMC) were analyzed for production of gamma interferon in an ELISPOT assay as follows: PBMC were incubated for 30-36 hours in the presence of pools of overlapping peptides corresponding to the individual HIV-1 antigens in the vaccines. The total number of gamma interferon-producing cells from each animal is shown in the table. T cell responses to DNA vaccination were limited (weeks 2-20). However, boosting with MVA/48 resulted in very strong HIV-1-specific T cell responses in all animals (week 25).

Table D: Antibody response following immunization of macaques with MVA/SHIV KB9.

Groups of rhesus macaques were immunized with 2×10^8 infectious units of MVA/SHIV-KB9 at weeks 0 and 4 by one of several routes: Tonsilar, intradermal (ID), or intramuscular (IM). Another group was immunized with non-recombinant MVA using the same routes. Serum samples from 2 weeks after the second immunization were analyzed for binding to KB9 envelope protein by ELISA and for neutralization of SHIV-89.6P and SHIV-89.6. In the ELISA assay, soluble KB9 envelope protein was captured in 96 well plates using an antibody to the C-terminus of gp120. Serial dilutions of sera were analyzed and used to determine the endpoint titers. Neutralization of SHIV-89.6P and SHIV-89.6 was determined in an MT-2 cell assay. Titers are the reciprocal serum dilution at which 50% of the cells were protected from virus-induced killing. In *in vitro* neutralization assays, SHIV-89.6P and SHIV-89.6 are heterologous, *i.e.* sera from animals infected with one of the viruses do not neutralize the other virus. Thus, two immunizations with MVA/SHIV-KB9 elicited good ELISA binding antibodies in all animals and neutralizing antibodies to the homologous virus (SHIV-89.6P) in some animals. In addition, heterologous neutralizing antibodies were observed in a subset of animals.

Table E: Frequencies of gag CM-9-specific CD3/CD8 T cells following immunization of macaques with MVA/SHIV-KB9.

Groups of MamuA*01 positive rhesus macaques were immunized with 2×10^8 infectious units of MVA/SHIV-KB9 at weeks 0 and 4 by one of several routes: tonsilar, intradermal (ID), or intramuscular IM). Another group was immunized with non-recombinant MVA. The frequencies of CD3+/CD8+ T cells that bound tetrameric complex containing the SIV gag-specific peptide CM9 were determined by flow cytometry at various times after each immunization. Time intervals were as follows: 1a, 1b, and 1d were one, two, and four weeks after the first immunization, respectively; 2a, 2b, 2c, and 2d were one, two, three, and twelve weeks after the second immunization, respectively. Values above background are shown in bold face. Strong SIV gag-specific responses were observed after a single immunization with MVA/SHIV-KB9 in all immunized animals. Boosting was observed in most animals following the second immunization. In addition, measurable tetramer binding was still found twelve weeks after the second immunization.

Table F: Frequencies of specific T cells following immunization of macaques with MVA/SHIV KB9.

Groups of macaques were immunized with MVA/SHIV-KB9 as described above. MVA/SHIV-KB9 expresses 5 genes from the chimeric virus, SHIV-89.6P: envelope, gag,

polymerase, tat, and nef. Thus, the frequencies of T cells specific for each of the 5 antigens were analyzed using pools of peptides corresponding to each individual protein. Fresh PBMC were stimulated with pools of peptides for 30-36 hours *in vitro*. Gamma interferon-producing cells were enumerated in an ELISPOT assay. The total number of cells specific for each antigen is given as "total # spots". In addition, the number of responding animals and average # of spots per group is shown. PBMC were analyzed at one week after the first immunization (1a) and one week after the second immunization (2a). Another group of 7 animals was immunized with non-recombinant MVA. In these animals, no spots above background levels were detected. Thus, a single immunization with MVA/SHIV-KB9 elicited strong SHIV-specific T cell responses in all animals. Gag and envelope responses were the strongest; most animals had responses to gag, all animals had responses to envelope. The Elispot responses were also observed after the second immunization with MVA/SHIV-KB9, albeit at lower levels. At both times, the rank order of responses was: tonsillar > ID > IM. We show good immune response to nef and some immune response to tat.

TABLE A. MVA/48 immunization – guinea pigs HIV-MN neutralizing antibody - reciprocal titer

Animal #	Group	Route	day 0	Day 4 MVA #1	day 30	day 3 MVA#2	day 49
885	MVA	I.M.	<20	I.M.	31	I.M.	24
891	"	"	<20	"	85	"	<20
882	MVA/48	I.M.	<20	I.M.	<20	I.M.	5,524
883	"	"	<20	"	68	"	691
886	"	"	<20	"	<20	"	4,249
890	"	"	<20	"	180	"	89
879	MVA/48	I.D.	<20	I.D.	<20	I.D.	817
881	"	"	<20	"	<20	"	234
888	"	"	<20	"	24	"	112
889	"	"	<20	"	22	"	376

TABLE B. Frequencies of HIV-gag-specific T cells following immunization of mice with MVA/48

<u>Group</u>	<u>P815 cells + gag peptide</u>		<u>gag peptide</u>		<u>no stimulation</u>	
MVA control	0	2	0	4	1	2
MVA/48 (IP)	>500	>500	>500	>500	8	8
MVA/48 (ID)	12	5	49	33	4	2
MVA/48 (IM)	22	18	66	49	12	8

TABLE C. DNA prime and MVA/48 boost. Total ELISPOTS per Animal

<u>Animal #</u>	<u>WEEKS</u>						
	<u>-2</u>	<u>2</u>	<u>6</u>	<u>10²</u>	<u>14²</u>	<u>20²</u>	<u>25²</u>
<u>RLW</u>	4	731*	<	47	43	50	3905
<u>RVI</u>	5	997*	<	<	<	8	205
<u>Roa</u>	< ¹	<	1	<	<	<	245
<u>RHc</u>	<	<	<	<	<	<	535
<u>Ryl</u>	<	<	<	<	<	<	4130
<u>RQk</u>	<	46	<	<	<	<	630
<u>RDr</u>	<	<	<	14	<	<	1965
<u>RZc</u>	<	5	<	58	<	<	925
<u>RSf</u>	<	118	<	<	<	20	5570
<u>Ras</u>	<	69	<	<	<	<	1435
Total	9	1966	1	119	43	78	19545
Geo Mean	4.5	105.3	1.0	33.7	43.0	20.0	1147.7

5

DNA primes were at 0 and 8 weeks and MVA/48 boost was at 24 weeks

¹ < = Background (2x the number of ELISPOTs in the unstimulated control + 10)

² Costimulatory antibodies were added to the ELISPOT incubations

10 * Animals from this bleed date exhibited higher than usual ELISPOTs.

**TABLE D. Antibody response following immunization of macaques with MVA/SHIV
KB9**

Animal #	Route	KB9 env ELISA titer	KB9 env elisa		SHIV- 89.6 Nab titer	SHIV- 89.6P Nab titer	SHIV- 89.6 # pos animals	SHIV- 89.6P # pos animals
			average	std dev.				
598	tonsil	25,600	31,086	20,383	<20	<20	3	2
601	"	51,200			<20	<20		
606	"	25,600			<20	<20		
642	"	51,200			75	31		
646	"	51,200			61	48		
653	"	6,400			<20	<20		
654	"	6,400			22	<20		
602	i.d.	25,600	18,800	15,341	38	<20	2	4
604	"	12,800			<20	262		
608	"	3,200			20	66		
637	"	12,800			<20	35		
638	"	51,200			<20	<20		
645	"	25,600			<20	<20		
647	"	12,800			32	162		
650	"	6,400			<20	<20		
599	i.m.	6,400	17,000	16,516	<20	<20	0	3
600	"	6,400			<20	29		
609	"	6,400			<20	<20		
639	"	51,200			<20	85		
640	"	12,800			<20	<20		
641	"	25,600			<20	41		
649	"	1,600			<20	<20		
651	"	25,600			20	<20		
603	Control	<100	<100		<20	<20	0	0
605	"	<100			<20	<20		
607	"	<100			<20	<20		
643	"	<100			<20	<20		
644	"	<100			<20	<20		
648	"	<100			<20	<20		
652	"	<100			<20	<20		

TABLE E. Frequencies of gag CM9-specific CD3/CD8 T cells following immunization of macaques with MVA/SHIV KB9

Animal #	Route	Virus	pre-bleed	1a	1b	1d	2a	2b	2c	2d
598	Tonsil	MVA/K B9	0.018	0.41	0.79	0.25	2.64	1.13	0.51	0.21
601	"	"	0.071	0.34	0.38	0.27	0.83	0.7	0.36	0.039
646	"	"	0.022	0.68	0.76	0.43	1.12	0.91	0.53	0.15
653	"	"	0.041	0.69	0.85	0.53	0.68	0.49	0.47	0.3
648	"	MVA		0.033	0.039		0.022	0.058	0.033	0.013
602	i.d.	MVA/K B9	0.019	0.17	0.92	0.5	0.95	0.59	0.5	0.2
604	"	"	0.013	0.11	0.38	0.32	0.44	0.38	0.19	0.25
650	"	"	0.095	0.17	0.6	0.23	2.87	1.12	0.9	0.16
647	"	"	0.032	0.22	0.38	0.14	0.84	0.91	0.34	0.17
652	"	MVA		0.041	0.038	0.059	0.025	0.022	0.026	0.055
599	i.m.	MVA/K B9		0.081	0.31	0.082		0.12	0.054	0.11
600	"	"	0.034	0.15	0.41	0.17	0.29	0.27	0.16	0.049
649	"	"	0.00486	0.35	1.34	0.56	2.42	0.77	0.69	0.22
651	"	"	0.049	0.12	0.69	0.25	1.01	0.32	0.24	0.22
603	"	MVA		0.024	0.087	0.073		0.082	0.027	0.17

TABLE F. Frequencies of specific T cells following immunization of macaques with MVA/SHIV KB9

Study groups	Gag specific			Tat specific			Nef specific			Env specific			Total	
	# responding animals	Total # spots	average # spots	# responding animals	total # spots	average # spots	# responding animals	total # spots	average # spots	# Responding animals	Total # spots	Average # spots	# Responding animals	Total
tonsil 1a	4/6	1325	221	0/6	0	0	3/6	195	33	6/6	8760	1460	6/6	6/6
tonsil 2a	5/6	1405	234	0/6	0	0	1/6	560	93	6/6	4485	748	6/6	6/6
i.d. 1a	7/7	1335	191	0/7	0	0	2/7	215	31	7/7	7320	1046	7/7	7/7
i.d. 2a	4/7	755	108	0/7	0	0	1/7	55	8	7/7	2700	386	7/7	7/7
i.m. 1a	7/7	925	132	1/7	60	9	3/7	180	26	7/7	5490	784	7/7	7/7
i.m. 2a	4/7	250	36	0/7	0	0	0/7	0	0	6/7	2205	315	6/7	6/7

EXAMPLE 5

Construction and Pre-clinical Immunogenicity of a Recombinant MVA Vaccine Expressing Subtype D HIV-1 Env, Gag and Pol for Use in Uganda.

Recombinant MVA vaccines have been successful in generating SIV and SHIV specific humoral and CD8 T cell responses in non-human primates and, alone or in combination with DNA vaccines, have provided protection in rhesus macaques from disease after pathogenic SHIV challenge. An overall program goal is to conduct clinical vaccine trials in Africa using vaccines that induce both neutralizing antibody and CD8 T cell specific responses and that are based upon representative full-length HIV-1 sequences isolated from the target vaccine cohorts. The predominant incident and prevalent HIV-1 subtype in Uganda is subtype D. Several R5 subtype D HIV-1 strains were selected and used to prepare recombinant MVA vaccines expressing env, gag, protease and RT. Initially, multiple env and gag/pol clones from 3 pure Ugandan subtype D isolates were selected. These sequences were separately cloned into pCR2.1 and tested for expression levels *in-vitro* by immunoprecipitation, and for envelope function as assessed by envelope-mediated fusion with CD4 and CCR5 or CXCR4 expressing cells. Based on these *in-vitro* analyses, several R5 subtype D env and gag/pol sequences were selected and cloned into MVA shuttle plasmids containing GFP and the modified H5 promoter for sequential cloning into deletions II and III, respectively, of MVA. The parent MVA used was a 1974 stock chosen to eliminate FDA concerns regarding potential BSE contamination. Several recombinant MVA (rMVA-UGD) expressing subtype D env and gag/pol were prepared in primary CEF cells using gamma-irradiated FBS from a USDA approved source and selected using GFP expression. These rMVA-UGD were further plaque-purified and amplified to titers sufficient for *in-vivo* immunogenicity studies. Pre-clinical humoral and cellular immunogenicity of the various rMVA-UGD were then assessed in BALB/c mice.

MVA Expressing Altered HIV-1 Envelope, Gag, and Polymerase Genes from Ugandan Clade D

This example describes the construction of 5 recombinant Modified Vaccinia Virus Ankara (MVA) viruses expressing envelope (*env*) and *gagpol* genes from HIV-1 clade D isolates from Uganda.

Sequences from Ugandan Clade D:

Env and *gagpol* genes from three Ugandan clade D HIV-1 isolates were used:

HIV-1 Isolate name	GenBank Accession #	Lab designation (LVD)
99UGA03349	AF484518	B
99UGA07412	AF484477	C
98UG57128	AF484502	D

Env and *gagpol* genes were PCR amplified from Ugandan HIV-1 clade D isolates by short term co-cultures on normal human PBMC (Harris *et al.* 2002 *AIDS Research and Human Retroviruses* 18:1281) using the oligonucleotides shown in **Table G** and cloned into pCR2.1-TOPO (Invitrogen). (HIV-I infected individuals contain a population or quasi-species of related but distinct viruses. Upon co-culture, multiple viruses can emerge such that the sequences of individual amplified genes from the co-culture may differ from the sequence of the full genome.) The resulting amplified *env* genes have a C-terminal deletion of 115 amino acids that was previously shown to enhance expression and yield a more stable recombinant virus. The resulting *gagpol* genes have a deletion of the entire integrase and RNase H portions of the genes. Within both the *env* and *gagpol* genes, several mutations were made by site-directed mutagenesis (Quik Change from Stratagene). In the *env* genes, silent mutations were made to eliminate two naturally occurring vaccinia virus early transcription termination signals (TTTTTNT, SEQ ID NO: 14) (Earl *et. al* 1990 *J.Virol* 64:2451). In the *pol* genes, two mutations were made in separate locations in the active site of reverse transcriptase to abolish enzymatic activity (Larder *et. al* 1987 *Nature* 327:716)(see Tables H(i) & (ii) for details on changes made to *env* and *gagpol* genes).

PCR2.1-TOPO plasmids containing the amplified genes were first characterized with respect to the orientation of the gene. Clones in which the gene was oriented properly with respect to the T7 promoter were chosen and protein expression was analyzed as previously described (Earl *et al.* 1997 *J Virol* 71:2674). Briefly, BS-C-1 cells were infected with vTF7-3 (Fuerst *et al.* 1986 *PNAS USA* 83:8122), a recombinant vaccinia virus expressing T7 RNA polymerase, transfected with a plasmid, and metabolically labeled. Cell lysates were subjected to immunoprecipitation with serum pooled from several HIV-1 clade D-infected individuals. Proteins were analyzed by SDS-polyacrylamide gel electrophoresis and visualized by autoradiography. One *env* and one *gagpol* DNA clone from each clade D isolate was chosen for construction of recombinant MVA viruses. DNA

sequencing was performed to confirm the integrity of each gene. Sequences are presented in Appendix 1.

Cloning into shuttle plasmids:

Two MVA shuttle plasmids, pLAS-1 and pLAS-2 (**Figure 18**), were used for construction of recombinant MVA viruses. DNA sequences of both plasmids are presented in Appendix 2. In both plasmids, the foreign gene is driven by the modified H5 promoter. In addition, both plasmids contain a cassette with the gene for green fluorescent protein (GFP) driven by the vaccinia p11 promoter. This cassette is flanked by direct repeats that will readily recombine to eliminate GFP during virus propagation. Thus, GFP is used as a positive screening marker in early rounds of plaque purification, and as a negative screening marker in final recombinant virus selection (**Figure 22**). MVA flanking sequences in pLAS-1 and pLAS-2 direct recombination into deletion III (Del III) and deletion II (Del II) of MVA, respectively.

Gagpol genes from 2 isolates (99UGA03349 and 99UGA07412) were cloned separately into pLAS-1 for insertion into Del III of MVA. The plasmids were named pLAS-1/UGD/Bgag and pLAS-1/UGD/Cgag (**Figure 19 & Table I**). When the *env* gene is cloned into the NotI restriction site, a short open reading frame precedes the *env* open reading frame. This open reading frame is out of frame with *env* and terminates at approximately nucleotide 75 in the *env* gene.

Env genes from three isolates (99UGA03349, 99UGA07412, 98UG57128) were cloned separately into MVA shuttle plasmid pLAS-2, for insertion into Del II of MVA. Plasmids were named pLAS-2/UGD/Benv, pLAS-2/UGD/Cenv, and pLAS-2/UGD/Denv (**Figure 20 & Table I**). When the *env* gene is cloned into the NotI restriction site, a short open reading frame precedes the *env* open reading frame. This open reading frame is out of frame with *env* and terminates at approximately nucleotide 75 in the *env* gene.

Foreign genes in pLAS-1 and pLAS-2 recombine into the vaccinia genome in the same orientation as the surrounding vaccinia genes. To test the effect of reversing the orientation of the *env* gene on the level of gene expression and stability of viruses, two of the *env* genes and their promoters were excised from pLAS-2 with restriction endonucleases BspE1 and EcoRV; sticky ends were filled in with Klenow enzyme; and the fragments were then reinserted into pLAS-2 the opposite orientation (**Figure 21**). Plasmids were named pLAS-2/UGD/revCenv and pLAS-2/UGD/revDenv (**Table I**).

Recombinant MVA construction:

Parent MVA: MVA 1974/NIH Clone 1 was used as the parent for all recombinant viruses. It was derived from a stock of MVA at passage 572, prepared on 2/22/1974 in the laboratory of A. Mayr in Germany. After receipt in the Laboratory of Viral Diseases, this stock was passaged a total of 6 times in chicken embryo fibroblast (CEF) cells, including 3 clonal purifications. Amplification was performed on the final, clonally purified virus. All CEF cells were derived from specific pathogen-free (SPAFAS) eggs.

Recombinant viruses expressing gagpol: CEF cells were infected with MVA 1974/NIH Clone 1 and transfected with either pLAS-1/UGD/Bgag or pLAS-1/UGD/Cgag for insertion into Del III. Two to three rounds of plaque purification were performed based on GFP expression. Further rounds of plaque purification were performed by picking plaques based on lack of GFP expression and concomitant positive gag expression as measured by immunostaining using a monoclonal antibody to HIV-1 p24 (183-H12-5C; obtained from the NIH AIDS Research and Reference Reagent Program) (**Figure 22**). Recombinant gagpol-expressing viruses were amplified and characterized for gag expression by immunoprecipitation as described above. The two viruses were named MVA/UGD/Bgag and MVA/UGD/Cgag. These viruses were then used as the parent in making gagpol/env recombinant viruses (see below).

Recombinant viruses expressing gagpol and env: Recombinant viruses, MVA/UGD/Bgag and MVA/UGD/Cgag were used as parent viruses for insertion of *env* genes. Thus, CEF cells were infected with either MVA/UGD/Bgag or MVA/UGD/Cgag and subsequently transfected with one of the pLAS-2-env-containing plasmids described above (**Figure 23 & Table I**). As above, the first two rounds of plaque purification were performed based on GFP expression. In subsequent rounds of purification, plaques were selected based on loss of GFP expression and positive gag and env expression as measured by immunostaining in duplicate cultures (**Figure 22**). A total of 5 gagpol/env-expressing viruses (MVA/UGD-1 through -5) were amplified and characterized (**Table J**).

Characterization of recombinant MVA/UGD viruses:

The 5 MVA/UGD viruses have been characterized for gene expression and function. Immunoprecipitation of env and gag proteins is shown in **Figure 24**. BS-C-1 cells were infected with individual recombinant viruses at a multiplicity of infection of 10, metabolically labeled, and lysates were subjected to immunoprecipitation with a pool of sera from HIV-1 clade D infected individuals. Viruses expressing gagpol only

(MVA/UGD/Bgag and Cgag) were included, as was non-recombinant MVA as a negative control and MVA/CMDR as a positive control. The latter virus expresses gagpol/env from a Clade E HIV-1 isolate. All viruses produced high levels of gag protein and efficient processing into p24 was observed. In addition, all env-expressing viruses produced high levels of env protein (gp160).

Figure 25 demonstrates that the gag and env proteins produced by the MVA/UGD viruses are functional. Virus-like particles were obtained by centrifugation of the supernatant of infected cells through a sucrose cushion (Karacostas *et al.* 1993 *Virology* 193:661). Pelleted material was then separated by SDS-polyacrylamide gel electrophoresis and analyzed by autoradiography (Panel A). As seen, p55 and p24 gag proteins were found in the pellet indicating that virus-like particles were formed. Panel B shows results of an assay in which env-expressing cells (infected with MVA/UGD virus) were mixed with cells expressing CD4 and co-receptor (X4 or R5) (Nussbaum, Broder, & Berger 1994 *J. Virol* 68:5411). Fusion was measured by beta-galactosidase activity in cell lysates. As shown, all five MVA/UGD viruses induced fusion with CD4/R5-expressing cells.

Immunogenicity of recombinant MVA/UGD viruses (study 1):

Groups of Balb/c mice were immunized with individual MVA/UGD viruses, non-recombinant MVA (negative control), or MVA/CMDR (positive control - expressing clade E gagpol/env) at weeks 0 and 3. The dose was 10^7 infectious units per immunization and the route was intraperitoneal. Humoral and cell mediated responses were measured and are shown in **Figures 26-28**.

Antibody responses after 2 immunizations are shown in **Figure 26**. Reciprocal endpoint ELISA titers to p24 at various times after immunization are shown in Panel A. All UGD viruses elicited gag-specific antibodies after 2 immunizations. Env-specific responses are shown in Panel B. In this experiment, pooled sera from groups of mice were used to immunoprecipitate metabolically labeled, autologous gp160 proteins. As seen, sera from mice immunized with MVA/UGD-1, -3, and -4 reacted with gp160 (the other viruses were not tested in this assay). Reciprocal endpoint titers to gp140 env at various times after immunization are shown in Panel C. All UGD viruses elicited env-specific antibodies after 2 immunizations.

T cell responses were measured with several assays. First, gag and pol peptide-specific intracellular interferon gamma (IFN- γ) responses were measured by intracellular cytokine staining. Splenocytes were collected 3 weeks after immunization, stimulated in-

vitro for 7 days, and then cultured overnight with peptide-pulsed P815 cells. Brefeldin A was added to prevent secretion of INF- γ . CD3 positive, CD8 positive, IFN- γ positive cells were enumerated by flow cytometry. Analyses were performed after one and two immunizations. Both gag- and pol-specific responses were observed after two immunizations (**Figure 27**) (samples from animals immunized with MVA/UGD-2, and 5 were not assayed). Second, gag- and pol-specific INF- γ responses were measured by ELISPOT (**Figure 28 A & B**). Briefly, splenocytes from immunized mice were mixed with gag or pol peptide-pulsed P815 cells in 96-well nitrocellulose plates coated with anti-IFN- γ antibody. After overnight incubation, spots were visualized by sequential incubation with anti-IFN- γ biotin antibody, streptavidin-HRP, and AEC substrate. Spots were enumerated using a Zeiss ELISPOT reader. Gag peptide-specific responses were found after one immunization and were boosted in most groups after the second immunization. Pol peptide-specific responses were found in several groups after two immunizations. Third, gag peptide-specific responses were measured by tetramer staining (H-2Kd gag LAI tetramer: AMQMLKETI, SEQ ID NO: 63) (**Figure 29**). Splenocytes were stimulated *in vitro* with either gag peptide or MVA/CM240gagpol, a recombinant virus expressing a clade E gagpol. CD3 positive, CD8 positive, tetramer positive cells were enumerated by flow cytometry. Positive tetramer staining was observed with cells from several groups of mice.

Immunogenicity of recombinant MVA/UGD viruses (study 2):

A second mouse immunogenicity study was performed to confirm the humoral and cellular immunogenicity of MVA/UGD-3 and MVA/UGD-4. BALB/c mice (10 per group) were administered intraperitoneal immunizations of 10^7 infectious units of MVA at weeks 0 and 3. Five mice per group were sacrificed two weeks after the 1st and 2nd immunizations and spleens were removed for assessment of cellular immunogenicity. Sera were collected from each mouse at weeks -1, 0, 1, 2, 3, 4 and 5. Splenocytes and sera were pooled together by group. HIV gag-specific serum IgG responses were detected from all MVA/UGD-immunized groups after the 2nd immunization (**Figure 30**). These gag-specific responses were predominantly of subclass IgG2a for both MVA/UGD-3 and MVA/UGD-4 demonstrating a Th1-type response (**Table Q**).

HIV-specific cell-mediated immunity was assessed by four separate assays: (1) intracellular IFN- γ staining by flow cytometry (ICS), (2) IFN- γ secretion by ELISPOT, (3) gag-peptide specific tetramer staining and (4) cytotoxic T lymphocyte (CTL) killing. (1)

ICS: Splenocytes were collected two weeks after the 1st immunization, stimulated for 7 days with MVA-infected P815 cells and then incubated overnight with P815 cells pulsed with a gag or pol peptide previously shown to be target of CD8 T cells in BALB/c mice (Casimiro *et al.* 2002 *J Virol* 76:185). Brefeldin A was included in the overnight incubation to prevent cytokine secretion. Both HIV gag- and pol-specific responses were detected for the MVA/UGD-immunized, but not control immunized, mice as evidenced by the production of intracellular INF- γ after peptide stimulation (**Figure 31A**). For example, 8.5% and 5% of splenocyte lymphocytes from MVA/UGD-4-immunized mice were positive for gag and pol, respectively. Similar results were obtained for the MVA/UGD immunized mice after the 2nd immunization (**Figure 31B**). (2) IFN- γ ELISPOT: HIV gag-specific IFN- γ responses were detected by ELISPOT without prior *in-vitro* stimulation after both the 1st (**Figure 32A**) and 2nd immunization (**Figure 32B**) with a boost detected after the 2nd immunization. HIV gag-specific responses were stronger than the pol-specific responses. HIV pol-specific responses were detectable after a 7-day *in-vitro* stimulation with P815 cells pulsed with pol peptide (**Figure 32C**). (3) Tetramer staining: HIV gag peptide-specific responses were measured by tetramer staining (H-2Kd gag LAI tetramer: AMQMLKETI, SEQ ID NO: 63) (**Figure 33**). Splenocytes were stained pre or post a 7-day *in-vitro* stimulation with P815 cells either pulsed with gag peptide or infected with MVA/CM240gag/pol, a recombinant virus expressing a subtype E gagpol. CD3 positive/CD8 positive/tetramer positive cells were enumerated by flow cytometry. Positive tetramer staining was observed for all of the MVA/UGD immunized groups both pre-IVS and post-IVS with MVA-infected P815 cells. (4) CTL: Splenocytes removed 2 weeks after the 1st immunization were stimulated *in-vitro* with MVA/CME (a recombinant MVA expressing env and gagpol from a subtype E HIV-1 isolate) infected P815 cells for 7 days and tested for the ability to lyse P815 cells pulsed with gag peptide. Splenocytes from all MVA/UGD, but not MVA control, immunized mice efficiently lysed peptide-pulsed P815 cells at E:T ratios of 20:1 (**Figure 34**).

EXAMPLE 6

Construction and Pre-clinical Immunogenicity of a Recombinant MVA Vaccine

Expressing Subtype A HIV-1 Env, Gag and Pol for Use in Kenya.

As part of the overall program goal to conduct clinical vaccine trials in Africa, HIV-1 sequences from Kenya were selected. The predominant incident and prevalent HIV-1 subtype in Kenya is subtype A. Several R5 subtype A HIV-1 strains were selected and

used to prepare recombinant MVA vaccines expressing env and gagpol (gag, protease and RT). One gagpol and two env clones from pure Kenyan subtype A isolates were selected.

HIV-1 Isolate name	GenBank Accession #	Publication
00KE-KER2008	AF457052	AIDS 16:1809 (2002)
00KE-KNH1144	AF457066	"
00KE-KNH1207	AF457068	"

All the steps described in EXAMPLE 5 for construction of subtype D recombinant MVA viruses were followed for the selected subtype A clones. These include:

- 5 PCR amplification of truncated genes and cloning into pCR2.1.
- Testing for *in vitro* expression.
- Testing for env and gag function.
- Site directed mutagenesis in env to eliminate vaccinia virus early transcription termination sites.
- 10 Site directed mutagenesis in pol to inactivate enzymatic activity.
- Cloning into MVA shuttle plasmids pLAS-1 (gagpol) and pLAS-2 (env).
- Recombination of gagpol into MVA 1974/NIH Clone 1 using primary CEF cells.
- Recombination of env into the recombinant virus expressing gag to produce a single virus expressing both gagpol and env.
- 15 Recombination of env into MVA 1974/NIH Clone 1 to produce a virus expressing env only.

In addition to the mutations described in EXAMPLE 5 and utilized with the Ugandan subtype D env genes, two other mutations were introduced into one of the Kenyan clade A env genes (KNH1144). First, the tyr at position 717 was mutated using site directed mutagenesis to either ala or ser. This mutation has been shown to increase cell surface expression of env proteins (Rowell *et al.* 1995 *J Immunol* **155**:473; LaBranche *et al.* 1995 *J Virol* **69**:5217). Second, the env protein was further truncated at the C-terminus just prior to the transmembrane domain yielding a soluble, secreted form of the protein. Published studies have shown that immunization with this form of the env protein results in enhanced antibody production as compared to membrane bound env.

The specifics of deletions and mutations in env and gagpol genes for KEA isolates are given in **Table K**. Plasmids and viruses expressing KEA env and gagpol genes are given in **Table L** and **M**.

Characterization of recombinant MVA/KEA viruses:

The MVA/KEA viruses were characterized for gene expression and function. Immunoprecipitation of env and gag proteins is shown in **Figure 35**. BSC-1 cells were infected with individual recombinant viruses at a multiplicity of infection of 10, metabolically labeled, and lysates were subjected to immunoprecipitation with a pool of antibodies including: monoclonal antibody T24 (env), monoclonal antibody 183-H12-5C (gag), and pooled HIV-1+ sera. MVA/UGD and WR/vEJW-1 were included as positive controls for env and gag expression, respectively. All MVA/KEA viruses express high levels of env and/or gag, as expected.

Virus-like particles were obtained by centrifugation of the supernatant of infected cells through a sucrose cushion (Karacostas *et al.* 1993 *Virology* **193**:661). Gag p24 protein was found in the pelleted material indicating the formation of virus-like particles (**Figure 36**).

A cell-cell fusion assay was used to assess the function of expressed, membrane bound env. In this assay env-expressing cells (infected with MVA/KEA virus) were mixed with cells expressing CD4 and co-receptor (X4 or R5) (Nussbaum, *et al.* 1994 *J Virol* **68**:5411). Fusion was measured by beta-galactosidase activity in cell lysates. All viruses expressing membrane bound env induced fusion with CD4/R5-expressing cells (**Figure 37**).

Immunogenicity of env and gag in mice:

The 5 recombinant MVA viruses expressing env, gag and pol from Kenyan subtype A HIV-1 isolates (MVA/KEA-1 through MVA/KEA-5), the 3 viruses expressing subtype A env alone (MVA/KEA-6 through MVA/KEA-8) and the MVA expressing subtype A gag/pol (MVA/KEA-9) were evaluated in an *in-vivo* mouse immunogenicity study designed to measure the humoral and cellular immunogenicity of these vaccines. BALB/c mice (10 per group) were administered intraperitoneal immunizations of 10^7 infectious units of individual MVA/KEA viruses at weeks 0 and 3. Five mice per group were sacrificed two weeks after the 1st and 2nd immunizations and spleens were removed for assessment of cellular immunogenicity. Sera were collected from each mouse at weeks -1, 0, 1, 2, 3, 4 and 5. Splenocytes and sera were pooled together by group. HIV env-specific serum IgG responses were detected from all MVA/KEA-immunized groups after the 2nd immunization (**Figure 38**). While env-specific responses were detected in all groups except for the control group, they were strongest in MVA/KEA-3, 4, 5, 6 and 8.

HIV-specific cell-mediated immunity was assessed by three assays: (1) intracellular IFN- γ staining by flow cytometry (ICS), (2) IFN- γ secretion by ELISPOT, and (3) gag-peptide specific tetramer staining. (1) ICS: Splenocytes were collected two weeks after the 2nd immunization, stimulated for 7 days with MVA-infected P815 cells and then incubated overnight with P815 cells pulsed with a gag or pol peptide previously shown to be target of CD8 T cells in BALB/c mice (Casimiro *et al.* 2002 *J Virol* 76:185). Brefeldin A was included during the overnight incubation to prevent cytokine secretion. HIV gag-specific responses were detected for each of the groups immunized with MVA/KEA viruses expressing env, gag and pol or gag and pol, but not control immunized, mice as evidenced by the production of intracellular IFN- γ after peptide stimulation (**Figure 39**). Splenocytes positive for HIV gag ranged from 7% to 22% for the MVA/KEA-immunized mice. (2) IFN- γ ELISPOT: HIV gag-specific IFN- γ responses from splenocytes without prior *in-vitro* stimulation were detected in all groups receiving MVA/KEA viruses expressing gag after the 1st immunization (**Figure 40**). (3) Tetramer staining: HIV gag peptide-specific responses were measured by tetramer staining (H-2Kd gag LAI tetramer: AMQMLKETI, SEQ ID NO: 63) (**Figure 41**). Splenocytes were stained pre-IVS. CD3 positive CD8 positive, tetramer positive cells were enumerated by flow cytometry. Similar to the above ICS and IFN- γ ELISPOT results, positive tetramer staining was observed for all of the groups immunized with MVA-KEA expressing gag.

EXAMPLE 7

Construction of a Recombinant MVA Vaccine Expressing Subtype C HIV-1 Env, Gag, and Pol for Use in Tanzania.

As part of the overall program goal to conduct clinical vaccine trials in Africa, HIV-1 sequences from Tanzania were selected. The predominant incident and prevalent HIV-1 subtype in Tanzania is subtype C. Several R5 subtype C HIV-1 strains were selected and used to prepare recombinant MVA vaccines expressing env, gag, protease and RT. One gagpol and one env clone from pure Tanzanian subtype C isolates were selected.

HIV-1 Isolate name	GenBank Accession #	Publication
00TZA-125	AY253304	AIDS Res & Hum Retrov, in press
00TZA-246	AY253308	"

Steps described in EXAMPLE 5 were followed for the selected subtype C clones. These include:

PCR amplification of truncated genes and cloning into pCR2.1.

Testing for *in vitro* expression.

Testing for env and gag function.

Site directed mutagenesis in env to eliminate vaccinia virus early transcription termination sites.

Site directed mutagenesis in pol to inactivate enzymatic activity.

5 Cloning into MVA shuttle plasmids pLAS-1 (gagpol) and pLAS-2 (env).

Recombination of gagpol into MVA 1974/NIH Clone 1 using primary CEF cells.

Recombination of env into the recombinant virus expressing gag to produce a single virus expressing both gagpol and env.

10 Recombination of env into MVA 1974/NIH Clone 1 to produce a virus expressing env only.

The specifics of deletions and mutations in env and gagpol genes for TZC isolates is given in **Table N**. Plasmids and viruses expressing TZC env and gagpol genes are given in **Table O** and **P**.

Characterization of recombinant MVA/TZA viruses:

15 The KEA viruses were characterized for gene expression. Immunoprecipitation of env and gag proteins is shown in **Figure 42**. BSC-1 cells were infected with individual recombinant viruses at a multiplicity of infection of 10, metabolically labeled, and lysates were subjected to immunoprecipitation with a pool of sera from HIV-1 infected individuals. MVA/CMDR and MVA were included as positive and negative controls, respectively. The MVA/KEA virus expresses high levels of env and gag.

20 Virus-like particles were obtained by centrifugation of the supernatant of infected cells through a sucrose cushion (Karacostas *et al.* 1993 *Virology* 193:661). Gag p24 protein was found in the pelleted material indicating the formation of virus-like particles (**Figure 43**).

EXAMPLE 8

Stability of expression of HIV-1 genes in recombinant MVA viruses.

25 The stability of the inserted env and gagpol genes in recombinant viruses from each of the subtypes was tested after serial passage in CEF cells. Viruses were grown in CEF cells using a procedure that mimics that used for expansion of virus for large scale vaccine production, i.e. infection at low multiplicity, growth for 3 days at 37C, harvesting of virus from cell lysates. After repeated passage, the virus stocks were tested for stability of the inserts using a 3-day immunostaining protocol. In this protocol, CEF cell monolayers were infected at a low multiplicity allowing for visualization of individual viral foci. After 3

days, the monolayers were fixed and then stained with monoclonal antibodies specific for either env or gag. Staining and non-staining foci were enumerated and results are shown in **Table R**. Very few non-staining foci were detected after 10-11 passages of each virus indicating that the inserted genes are stable after repeated passage in culture.

Table G. Oligonucleotides used for PCR amplification of env and gagpol genes

Clade D Ugandan HIV-1 isolates	env	
	5' primer	3' primer
99UGA03349	GCGCCCCGGGTCGACGCGGCCGCGCCATGAGAGTG AGGGGATACAGAGGAAC (SEQ ID NO: 15)	GCGCCCCGGGCGCGCCGAGAAAAATTAGCCTTGCTC TCCACCTTCTTCTATTC (SEQ ID NO: 16)
99UGA07412	GCGCCCCGGGTCGACGCGGCCGCGCCATGAGAGTG AGGAGACAGTGAGGAATTAT (SEQ ID NO: 17)	GCGCCCCGGGCGCGCCGAGAAAAATTAGCCTTGCTC TCCACCTTCTTCTATTC (SEQ ID NO: 18)
98UG57128	GCGCCCCGGGTCGACGCGGCCGCGCCATGAGAGTG AGGGGATAGAGAGGAATTAT (SEQ ID NO: 19)	GCGCCCCGGGCGCGCCGAGAAAAATTAGCCTTGCTC TCCACCTTCTCCTC (SEQ ID NO: 20)

Clade D Ugandan HIV-1 isolates	gagpol	
	5' primer	3' primer
99UGA03349	GCGCCCCGGGCCATGGGTGCGAGAGCGTCAGTAT TAAGC (SEQ ID NO: 21)	GCGCCCCGGGAGAAAAATTAGAGGTTTCTGCTCCT ACTATGGGTTCT (SEQ ID NO: 22)
99UGA07412	GCGCCCCGGGCCATGGGTGCGAGAGCGTCAGTGT TAAGT (SEQ ID NO: 23)	GCGCCCCGGGAGAAAAATTAGAGGTTTCTGCTCCT ACTATGGGTTCT (SEQ ID NO: 24)

Table H(i) . Deletions/mutations in env genes

Clade D Ugandan HIV-1 isolates	envelope							
	C-terminal truncation				TTTTTNT silent mutations			
	DNA sequence at C-terminus	# of last nucleotide	amino acid sequence at C- terminus	# of last amino acid	amino acid nucleotide	#	amino acid nucleotide	#
99UGA0334	GGAATAGAAAGAAAGGTGGA GAGCAAGGC (SEQ ID NO: 25)	2178	GIEEGGEQG (SEQ ID NO: 26)	726	phe to phe T to C	168 504	phe to phe T to C	374 1122
99UGA07412	GGAATAGAAAGAAAGGTGGA GAGCAAGGC (SEQ ID NO: 27)	2211	GIEEGGEQG (SEQ ID NO: 28)	737	phe to phe T to C	173 519	phe to phe T to C	384 1152
98UG57128	GGAACAGAAAGGAGAAAGGTGGA GAGCAAGGC (SEQ ID NO: 29)	2262	GTEGEGGEQG (SEQ ID NO: 30)	754	phe to phe T to C	182 546	phe to phe T to C	391 1173

Table H(ii). Deletions/mutations in gagpol genes

Clade D Ugandan HIV-1 isolates	gagpol					pol		
	C-terminal truncation					RT active site mutations		
	DNA sequence at C-terminus	# of last nucleotide	amino acid sequence at C- terminus	# of last amino acid	amino acid nucleotide	amino acid nucleotide	#	amino acid nucleotide
99UGA03349	AAGGAACCCCATAGTAGGAGCA GAAACCTTC (SEQ ID NO: 31)	3065	KEPIVGAETF (SEQ ID NO: 32)	1022	asp to glu T to G	asp to his GAT to CAC	692 2075	767 2298 & 2300
99UGA07412	AAGGAACCCCATAGTAGGAGCA GAAACTTTC (SEQ ID NO: 33)	3077	KEPIVGAETF (SEQ ID NO: 34)	1026	asp to glu T to G	asp to his GAT to CAC	696 2087	771 2310 & 2312

Table I. UGD Plasmids

<u>Plasmid</u>	<u>parent plasmid</u>	<u>HIV isolate</u>	<u>Insertion site in MVA</u>	<u>Direction in relation to vaccinia</u>
pLAS-1/UGD/Bgag	pLAS-1	99UGA03349	Del III	same
pLAS-1/UGD/Cgag	pLAS-1	99UGA07412	Del III	same
pLAS-2/UGD/Benv	pLAS-2	99UGA03349	Del II	same
pLAS-2/UGD/Cenv	pLAS-2	99UGA07412	Del II	same
pLAS-2/UGD/Denv	pLAS-2	98UG57128	Del II	same
pLAS-2/UGD/revCenv	pLAS-2	99UGA07412	Del II	reverse
pLAS-2/UGD/revDenv	pLAS-2	98UG57128	Del II	reverse

Table J. UGD Viruses

<u>Virus</u>	<u>gag/pol HIV isolate</u>	<u>gagpol Direction in relation to vaccinia</u>	<u>env HIV isolate</u>	<u>Env Direction in relation to vaccinia</u>
MVA/UGD-1	99UGA07412	same	99UGA07412	Same
MVA/UGD-2	99UGA03349	same	98UG57128	Reverse
MVA/UGD-3	99UGA07412	same	99UGA03349	Same
MVA/UGD-4	99UGA03349	same	99UGA07412	Reverse
MVA/UGD-5	99UGA03349	same	98UG57128	Same
MVA/UGD-Bgag	99UGA03349	same	-	-
MVA/UGD-Cgag	99UGA07412	same	-	-

Table K(i). Deletions/mutations in env genes (KEA)

Clade A Kenyan HIV-1 isolates	envelope							
	C-terminal truncation				TTTTTNT mutations		Tyrosine mutations	
	DNA sequence at C-terminus	# of last nucleotide	amino acid sequence at C-terminus	# of last amino acid	amino acid nucleotide	#	amino acid nucleotide	#
00KE-KNH1207	AGAAATCGAAGGAGAAAGGTGG AGAGCAAGAC (SEQ ID NO: 35)	2209	GRIEGEGGEQD (SEQ ID NO: 36)	736	phe to phe T to C	378 1134		
00KE-KNH1144	AGAAATCGAAGGAGAAAGGTGG AGAGCAAGAC (SEQ ID NO: 37)	2239	GRIEGEGGEQD (SEQ ID NO: 38)	746	phe to phe T to C	389 1167		
00KE-KNH1144	AGAAATCGAAGGAGAAAGGTGG AGAGCAAGAC (SEQ ID NO: 39)	2239	GRIEGEGGEQD (SEQ ID NO: 40)	746			tyr to ala TAC to GCG	717 2149 to 2151
00KE-KNH1144	AGAAATCGAAGGAGAAAGGTGG AGAGCAAGAC (SEQ ID NO: 41)	2239	GRIEGEGGEQD (SEQ ID NO: 42)	746			tyr to ser TAC to AGC	717 2149 to 2151
00KE-KNH1144	GACATATCAAAATTGGCTGTGG TATATAAGA (SEQ ID NO: 43)	2064	DISNWLWYIR (SEQ ID NO: 44)	688				

Table K(ii) Mutations/deletions in gagpol genes (KEA)

Clade A Kenyan HIV-1 isolates	gagpol				pol			
	C-terminal truncation				RT active site mutations			
	DNA sequence at C-terminus	# of last nucleotide in gagpol	amino acid sequence at C-terminus	# of last amino acid in pol	amino acid nucleotide	#amino acid in pol nucleotide # in gagpol	amino acid nucleotide	#amino acid in pol nucleotide # in gagpol
00KE-KER2008	GACCCCATAGCAGGA GCAGAGACTTTC (SEQ ID NO: 45)	3074	KDPIAGAETF (SEQ ID NO: 46)	595	asp to glu T to G	265 2084	asp to his GAT to CAC	340 2307 to 2309

Table L. KEA Plasmids

Plasmid	parent plasmid	HIV isolate	Insertion site in MVA	Direction in relation to vaccinia
pLAS-1/KER2008gag	pLAS-1	00KE-KER2008	Del III	same
pLAS-2/KNH1144env	pLAS-2	00KE-KNH1144	Del II	same
pLAS-2/KNH1207env	pLAS-2	00KE-KNH1207	Del II	same
pLAS-2/KNH1144gp140	pLAS-2	00KE-KNH1144	Del II	same
pLAS-2/KNH1144(Y/A)	pLAS-2	00KE-KNH1144	Del II	same
pLAS-2/KNH1144(Y/S)	pLAS-2	00KE-KNH1144	Del II	same

Table M. KEA Viruses

Virus	gag/pol HIV isolate	gag/pol Direction in relation to vaccinia	env HIV isolate	env Direction in relation to vaccinia
MVA/KEA-1	00KE-KER2008	same	00KE-KNH1144	same
MVA/KEA-2	00KE-KER2008	same	00KE-KNH1207	same
MVA/KEA-3	00KE-KER2008	same	00KE-KNH1144 (gp140)	same
MVA/KEA-4	00KE-KER2008	same	00KE-KNH1144 (Y/A)	same
MVA/KEA-5	00KE-KER2008	same	00KE-KNH1144 (Y/S)	same
MVA/KEA-6	-		00KE-KNH1144	same
MVA/KEA-7	-		00KE-KNH1207	same
MVA/KEA-8	-		00KE-KNH1144 (gp140)	same
MVA/KEA-9	00KE-KER2008	same	-	

Table N(i). Mutations/deletions in env genes (TZC)

Clade C Tanzanian HIV-1 isolates	envelope				
	C-terminal truncation			TTTTTNT silent mutations	
	DNA sequence at C-terminus	# of last nucleotide	amino acid sequence at C-terminus	# of last amino acid	amino acid nucleotide #
00TZA-125	GGAATCGAAGAAGGTGGAGAGCAA GAC (SEQ ID NO: 47)	2223	GTEEGGEQD (SEQ ID NO: 48)	741	phe to phe T to C 396 1161

Table N(ii). Mutations/deletions in gagpol (TZC)

Clade C Tanzanian HIV-1 isolates	gagpol			pol		
	C-terminal truncation			RT active site mutations		
	DNA sequence at C-terminus	# of last nucleotide of gagpol	amino acid sequence at C-terminus of pol	# of last amino acid in pol	amino acid nucleotide # in gagpol	amino acid nucleotide # in gagpol
00TZA-246	AAAGAACCCATAGTAGGAGC AGAAACTTTCT (SEQ ID NO: 49)	3083	KEPIVGAETF (SEQ ID NO: 50)	601	asp to glu T to G 271 2093	asp to his GAT to CAC 346 2316 & 2318

Table O. TZC Plasmids

Plasmid	parent plasmid	HIV isolate	Insertion site in MVA	Direction in relation to vaccinia
pLAS-1/TZC246gag	pLAS-1	00TZA-246	Del III	same
pLAS-2/TZC125env	pLAS-2	00TZA-125	Del II	same

Table P. TZC Viruses

Virus	gag/pol HIV isolate	gag/pol Direction in relation to vaccinia	env HIV isolate	env Direction in relation to vaccinia
MVA/TZC-1	00TZA-246	same	00TZA-125	same
MVA/TZC-2	00TZA-246	same	-	
MVA/TZC-3	-		00TZA-125	same

Table Q. HIV-1 gag IgG2a/IgG1 ratios

Group		Week1	Ratio G2a/G1	Week2	Ratio G2a/G1	Week3	Ratio G2a/G1
MVA/CMDR	G1	34,100	2.5	34,100	2.5	29,900	2
	G2a	85,300		85,300		59,700	
MVA control	G1	<100	-	<100	-	<100	-
	G2a	<100		<100		<100	
MVA/UGD-4	G1	34,100	1.6	25,600	1.6	19,200	1.8
	G2a	55,500		41,600		35,200	
MVA/UGD-3	G1	3,200	3.3	3,200	5.3	3,200	5.3
	G2a	10,600		17,100		17,100	
MVA/UGD-1	G1	9,600	0.7	4,800	1.3	4,800	1.3
	G2a	6,400		6,400		6,400	
MVA/UGD gag	G1	4,800	6.2	3,200	9.3	2,400	12.4
	G2a	29,900		29,900		29,900	

Table R. Stability of inserts in rMVA viruses

Virus	Gag			Env		
	total foci #	non-staining foci #	%	total foci #	non-staining foci #	%
After 10 passages MVA/UGD-4	270	3	1.1	270	3	1.1
After 11 passages						
MVA/KEA-1	360	2	0.56	310	0	0
MVA/KEA-2	230	3	1.3	280	0	0
MVA/KEA-3	340	0	0	340	0	0
MVA/KEA-4	210	0	0	210	0	0
MVA/KEA-5	220	1	0.45	210	0	0
After 11 passages MVA/TZC	600	5	0.83	800	1	0.13

Appendix 1 - DNA sequences of gagpol and env genes from Ugandan HIV-1 clade D isolates:

99UGA03349 gagpol (SEQ ID NO: 51):

5 ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGAAAAATTAGATGAATGGGAAAAAATTCGGTTA
 CGGCCAGGGGAAAAACAAAAATATAGATTAAACATTTAGTATGGGCAAGCAGGGAGCTAGAA
 CGATTTGCACTTAATCCTGGTCTTTTAGAAACATCAGAAGGCTGTAGACAAATAATAGAACAGC
 10 TACAACCATCTATTCAGACAGGATCAGAGGAACTTAAATCATTACATAATACAGTAGTAACCT
 CTATTGTGTACATGAAAGGATAAAGGTAGCAGATACCAAGGAAGCTTTAGATAAGATAAAGGA
 AGAACAAACCAAAAGTAAGAAAAAAGCACAGCAAGCAACAGCTGACAGCAGCCAGGTCAGCC
 AAAATTATCCTATAGTACAAAACCTACAGGGGCAAATGGTACACCAGTCCTTATCACCTAGGAC
 TTTGAATGCATGGGTAAAAGTAATAGAAGAGAAGGCTTTCAGCCCAGAAGTAATACCCATGTTT
 TCAGCATTATCAGAAGGAGCCACCCCAACAGATTTAAACACCATGCTAAACACAGTGGGGGGA
 15 CATCAAGCAGCCATGCAAATGTTAAAAGAGACTATCAATGAGGAAGCTGCAGAATGGGATAGG
 CTACATCCAGTGCCTGCAGGGCCTGTTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAT
 ATAGCAGGAACTACCAGTACCCTTCAGGAACAAATAGGATGGATGACAAGCAATCCACCTATCC
 CAGTAGGAGAAATCTATAAAAGATGGATAATCCTAGGATTAAATAAAATAGTAAGAATGTATA
 GCCCTGTGAGCATTGTTGGACATAAGACAAGGACCAAGGAACCTTTAGAGACTATGTAGATCG
 20 GTTCTATAAACTCTACGAGCCGAGCAAGCTTCACAGGATGTAAAAAATTGGATGACTGAAACC
 TTGTTAGTCCAAAATGCGAATCCAGATTGTAAAACTATCTTAAAAGCATTGGGACCAGCGGCTA
 CATTAGAAGAAATGATGACAGCATGTGAGGGAGTGGGGGGACCCAGTCATAAAGCAAGAGTTT
 TGGCTGAGGCAATGAGCCAAGCATCAAAACAAATGCTGTTATAATGATGCAGAGGGGCAATTT
 CAAGGGCAAGAAAATCATTAAGTGTTCAACTGTGGCAAAGAAGGACACCTAGCAAAAAATTG
 25 TAGGGCTCCTAGGAAAAGAGGCTGTTGGAAATGTGGAAAGGAAGGGCACCAAATGAAAGATTG
 TAATGAAAGACAGGCTAATTTTTTAGGGAGAATTTGGCCTTCCCACAAGGGGAGGCCAGGGAAT
 TTCCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGCAGAGAGCTTCGGGTTTGGGGAAGAG
 ATAACACCCTCCAGAAAACAGGAGGGGAAAGAGGAGCTGTATCCTTCAGCCTCCCTCAATCAC
 TCTTTGGCAACGACCCCTAGTCACAATAAAAAATAGGGGGACAGCTAAAGGAAGCTCTATTAGAT
 30 ACAGGAGCAGATGATACAGTAGTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAATG
 ATAGGGGGAATTGGGGGCTTATCAAAGTAAGACAGTATGATCAAATACTCGTAGAAATCTATG
 GATATAAGGCTACAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATTT
 GTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGTCCTATTGAACTGTACCAGTAAAA
 TAAAGTCAGGGATGGATGGTCCAAGAGTTAAACAATGGCCATTGACAGAAGAGAAAAATAAAG
 35 CACTAATAGAAATTTGTACAGAAATGGAAAAAGGAAGGAAAACTTTCAAGAATTGGACCTGAAA
 ATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGACAGTACTAAGTGGAGAAAATTAGT
 AGATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAACTAGGAATACCACAT
 CCTGCAGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGAGGTGGGTGATGCATATTTTTCAG
 TTCCCTTATATGAAGACTTTAGAAAATACACTGCATTCCACCATACCTAGTATAAACAATGAGAC
 40 ACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCGGCAATATTC
 CAAAGTAGCATGACAAAAATTTTGAACCTTTTAGAAAAACAAAATCCAGAAGTGGTTATCTACC
 AATACATGCACGATTTGTATGTAGGATCTGACTTAGAAAATAGGGCAGCATAGAATAAAAAATAGA
 GGAATTAAGGGGACACCTATTGAAGTGGGGATTACCACACCAGACAAAAATCATCAGAAGGA
 ACCTCCATTTCTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAAAAC
 45 TGCCAGAAAAAGAAAGCTGGACTGTCAATGATCTGCAGAAGTTAGTGGGGAAATTAAATTGGG
 CAAGTCAAATTTATTGAGGAATTAAAGTAAGACAATTATGCAAATGCCTTAGGGGAACCAAAGC
 ACTGACAGAAGTAGTACCACTGACAGAAGAAGCAGAATTAGAACTGGCAGAAAACAGGGAACT
 TCTAAAAGAAACAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAG
 50 AAACAAGGGCAAGACCAATGGACATATCAAATTTATCAAGAACAATATAAAAAATTTGAAAACA
 GGAAAGTATGCAAAGAGGAGGAGTACCCACACTAATGATGTAAACAATTAACAGAGGCAGTG
 CAAAAAATAGCCCAAGAATGTATAGTGATATGGGGAAAGACTCCTAAATTCAGACTACCCATAC
 AAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAGGCCACCTGGATTCTGAGTGGG
 AGTTTGTCAATACCCCTCCCTTGGTTAAATTATGGTACCAGTTAGAGAAGGAACCCATAGTAGG
 AGCAGAAACCTTCTAA

55

99UGA07412 gagpol (SEQ ID NO: 52):

ATGGGTGCGAGAGCGTCAGTGTTAAGTGGGGGAAAATTAGATGAATGGGAAAGAATTCGGTTA
CGGCCAGGGGGAAAACAAAAGATATAAACTAAACATATAGTATGGGCAAGCAGGGAGCTAGAG
5 CGATTTGCACTTAATCCTGGCCTTTTAGAAACATCAGAAGGCTGTAAACAAATATTGGGACAGC
TACAACCAGCTATTCAGACAGGATCAGAAGAACTTAAATCATTATATAATACAGTAGCAACCCT
CTATTGTGTACATGAGAGGCTAAAGGTAACAGACACCAAGGAAGCTTTAGACAAAATAGAGGA
AGAACAAACCAAAAAGTAAGAAAAAAGCACAGCAAGCAACAGCTGACACAAAAACAGCAGCC
AGGTCAGCCAAAATTATCCTATAGTACAAAACCTACAGGGGCAAATGGTACACCAGGCTATATC
10 ACCTAGAACGTTGAACGCATGGGTAAAAGTAATAGAGGAGAAGGCTTTAGCCCAAGTAAT
ACCCATGTTTTTTCAGCATTATCAGAAGGAGGCCACCCACAAGATTTAAACACCATGCTAAACACA
GTGGGGGGACATCAGGCAGCCATGCAGATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAA
TGGGATAGGTTACATCCAGTACATGCAGGGCCTATTGCACCAGGACAAATGAGAGAACCAACA
GGAAGTGATATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACCAGCAAT
15 CCACCTATCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTAGGATTAAATAAAAATAGTAA
GGATGTATAGCCCTGTCAGTATTTTGGACATAAAACAAGGGCCAAAGGAACCCCTTTAGAGACTA
TGTAGATCGGTTCTATAAACTCTAAGGGCCGAGCAAGCTTCACAGGAGGTAAAAGGTTGGATG
ACCGAAACCTTGTTGGTCCAAAATGCAAACCCAGATTGTAAAACCATCTTAAAAGCATTGGGAC
CAGCGGCTACATTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGTCATAAAG
20 CAAGAGTTTTGGCTGAGGCAATGAGTCAAGTCTCAACAAATACTGCTATAATGATGCAGAGAGG
CAATTTTAAAGGGCCCAAAGAAAAGCATTAAAGTGTTTTAACTGTGGCAAAGAAGGTCACACAGCA
AAAAACTGTAGAGCTCCTAGGAAAAGGGGCTGTTGGAAATGTGGAAGGGAAGGACATCAAATG
AAAGATTGCACTGAAAGACAGGCTAATTTTTTAGGGAAAAATTTGGCCTTCCACAAGGGAAGGC
CAGGGAATTTCTTCAGAACAGACCAGAGCCAACAGCCCCACCAGAAGAAAGCTTCGGGTTTGG
25 GGAAGAGATAACACCCTCTCAGAAACAGGAGAAGAAGGACAAGGAGCTGTATCCTGTAGCTTC
CCTCAAATCACTCTTTGGCAACGACCCCTTGTCACAATAAAGATAGGGGGACAGCTAAAGGAAG
CTCTACTAGATACAGGAGCAGATGATACAGTATTAGAAAGTAAGACAGTATGAGCAAAATACTTG
AACCAAAAATGATAGGGGGAATTGGAGGCTTTATCAAAGTAAGACAGTATGAGCAAAATACTTG
TAGAAATCTGTGGACAGAAAGCTATAGGTACAGTATTAGTAGGGCCTACACCTGTCAACATAAT
30 TGGAAGAAATTTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCAATTAGCCCTATTGAAACTG
TACCAGTAAAATTAAGGCCAGGGATGGACGGTCCAAAAGTTAAACAATGGCCATTGACAGAAG
AAAAGGTAAAAGCACTAATAGAAATTTGTACAGAAATGGAAGGAAGGAAAAATTTCAAGAA
TTGGACCTGAAAATCCATACAATACTCCAATATTTGCCATAAAGAAAAAGGACAGTACTAAGTG
GAGAAAATTAGTAGATTTTCAGGGAACCTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAACTA
35 GGAATACCACATCCTGCGGGGCTAAAAAAGAAAAAATCAGTAACAGTACTGGAGGTGGGTGAT
GCATATTTTTTCAGTTCCCTTATATGAAGATTTTAGAAAAATATACTGCATTACCATACCTAGTAT
AAACAATGAAACACCAGGAATTAGATATCAGTACAATGTGCTTCCACAAGGGTGGAAAGGATC
ACCAGCAATATTCCAAAAGTAGCATGACAAAAATCTTAGAACCTTTTAGAAAAACAAAATCCAGAA
ATGGTTATCTATCAATACATGCACGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATA
40 GAATAAAAATAGAAGAATTAAGGGGACACCTGTTGAAGTGGGGATTTACCACACCAGACAAAA
AGCATCAGAAAGAACCTCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGT
ACAGTCTATAAACTGCCAGAACAAGAAAGCTGGACTGTCAATGATATACAGAAGTTAGTGGG
AAAATTAATTTGGCAAGCCAGATTTATCCAGGAATTAAGGTAAAGACAATTATGCAAATGCATT
AGGGGTACCAAAAGCACTGACAGAAGTAGTACCAGTACAGAGAAGAAGCAGAATTAGAACTGGCA
45 GAAAACAGGGAAATTCTAAGAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAAGACTTA
ATAGCAGAGATACAGAAACAAGGGCAAGACCAGTGGACATACCAAATTTATCAAGAACAATAT
AAAAATCTGAAAACAGGAAAGTATGCAAAAGTGAGGGGTACCCACACTAATGATGTAAAACAA
TTAACAGAGGCAGTACAAAAAATAACCCAAGAATGTATAGTGATATGGGGAAAGCCTCCTAAA
TTTAGACTACCCATACAAAAAGAAACATGGGAAATATGGTGGACAGAGTATTGGCAGGCCACCT
50 GGATTCTGAGTGGGAGTTTGTCAATACCCCTCCTTTAGTTAAATTATGGTACCAATTAGAGAAG
GAACCCATAGTAGGAGCAGAACTTTCTAA

99UGA03349 envelope (SEQ ID NO: 53):

ATGAGAGTGAGGGGGATACAGAGGAACTATCAAACTTGTGGAGATGGGGCACCTTGCTCCTTG
GGATGTTGATGATATGTAAGGCTACAGAACAGTTGTGGGTCACAGTTTACTATGGGGTACCTGT
5 GTGGAAAGAAGCAACCACTACTCTATTTTGTGCATCAGATGCTAAATCATATAAAGAAGAAGCA
CATAATATCTGGGCTACACATGCCTGTGTACCAACAGACCCCAACCCACGAGAATTAATAATAG
AAAATGTCACAGAAAACCTTTAACATGTGGAAAAATAACATGGTGGAGCAGATGCATGAGGATA
TAATCAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAATTAACCCCACTCTGTGTCACTTTA
AATGCACTGAATGGAGGAAGAATAACACTATCAATGCCACCAGAATAGAAATGAAAACTGC
10 TCTTCAATCTAACACAGAAAATAAGAGATAGGAAAAAGCAAGTGCATGCACTTTTCTATAAAC
TTGATGTGGTACCAATAGATGATAATAATAGTACTAATACCAGCTATAGGTAAATAAATTGTAA
TACCTCAGCCATTACACAGGCGTGTCCAAAGGTAACCTTTGAGCCAATTCCCATACATTATTGTG
CCCCAGCTGGATATGCGATTCTAAAATGTAACAATAAGAAGTTCAATGGGACAGGTCCATGCGA
TAATGTCAGTACAGTACAGTGTACACATGGAATTAGGCCAGTAGTATCCACTCAATTGTTGTTGA
15 ATGGCAGTCTAGCAGAAGAAGACATAATAATTAGATCTGAGAATCTCACAAATAATGCTAAAAT
CATAATAGTACAGCTTAATGAGTCTGTAAACAATTAATTGCACAAGGCCCTACAACAATACAAGA
AGAGGTGTACATATAGGACCAGGGCGAGCATACTATACAACAGACATAATAGGAGATATAAGA
CAAGCACATTGTAAACATTAGTGGAGCAGAATGGAATAAGACTTTACATCGGGTAGCTAAAAAAT
TAAGAGACCTATTTAAAAAGACAACAATAATTTTAAACCGTCCTCCGGAGGGGACCCAGAAAT
20 TACAACACACAGCTTTAATTGTAGAGGGGAATTCTTCTACTGCAATACAACAAGACTGTTTAAT
AGCATATGGGGAAATAATAGTACAGGAGTTGATGAGAGTATAACACTCCCATGCAGAATAAAA
CAAATTATAAACATGTGGCAGGGAGTAGGAAAAGCAATGTATGCCCCCTCCCATGAAGGACTAA
TCAGCTGCTCATCAAATATTACAGGATTACTGTTGACAAGAGATGGTGGTGAAGTAACAGTAG
TCAGAATGAGACCTTCAGACCTGGAGGGGGAGATATGAGAGACAATTGGAGAAGTGAATTATA
25 TAAATATAAAGTAGTAAGAATTGAACCATTAGGTCTAGCACCTCCAAGGCAAAAAGAAGAGT
AGTAGAAAAGAGAGAAAAGAGCAATAGGACTAGGAGCTATGTTCCCTGGGTCTTGGGAGCAGC
AGGAAGCACGATGGGCGCAGCGTCACTGACGCTGACGGTACAGGCCAGACAGCTATTGTCTGGT
ATAGTGCAACAGCAAAACAATTTGCTGAAGGCTATAGAGGCGCAACAGCACCTGTTGCAACTCA
CAGTCTGGGGCGTTAAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGCTACCTAAGGGATCA
30 ACAGCTCCTAGGAATTTGGGGTTGCTCTGGAAAACACATTTGCACCACCAATGTGCCCTGGAAC
TCTAGCTGGAGTAATAAACTCTAAAATCAATTTGGGATAACATGACCTGGATGGAGTGGGAAA
GAGAAATTGACAATTACACAGGGATAATATACAATTTACTTGAAGAATCGCAAACCCAGCAAG
AAAGAAATGAACAAGACCTATTGAAATTGGACCAATGGGCAAGTTTGTGGAATTGGTTTAGCAT
AACAAAATGGCTGTGGTATATAAAAAATTTTATAATGATAGTAGGAGGCTTGATAGGCTTAAGG
35 ATAGTTTTTGCTGTGCTTTCTATAGTAAATAGAGTTAGGCAGGGATATTCACCTCTGTGTTTTCA
GACCTCCTCCCAGCCCCGCGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGA
GCAAGGCTAA

99UGA07412 envelope (SEQ IS NO: 54):

ATGAGAGTGAGGGAGACAGTGAGGAATTATCAGCACTTGTGGAGATGGGGCATCATGCTCCTTG
GGATGTTAATGATATGTAGTGCTGCAGACCAGCTGTGGGTACAGTGTATTATGGGGTACCTGT
5 GTGGAAAGAAGCAACCACTACTCTATTTTGTGCATCAGATGCTAAAGCACATAAAGCAGAGGCA
CATAATATCTGGGCTACACATGCCTGTGTACCAACAGACCCCAATCCACGAGAAATAATACTAG
GAAATGTCACAGAAAACCTTAAACATGTGGAAGAATAACATGGTAGAGCAGATGCATGAGGATA
TAATCAGTTTATGGGATCAAAGTCTAAAACCATGTGTAAAATTAACCCCACTCTGTGTTACTTTA
AACTGCACTACATATTGGAATGGAACCTTTACAGGGGAATGAACTAAAGGGAAGAATAGAAGT
10 GACATAATGACATGCTCTTTCAATATAACCACAGAAATAAGAGGTAGAAAGAAGCAAGAACT
GCACTTTTCTATAAACTTGATGTGGTACCACTAGAGGATAAGGATAGTAATAAGACTACCACT
ATAGCAGCTATAGATTAATAAATTGCAATACCTCAGTCGTGACACAGGCGTGTCCAAAAGTAAC
CTTTGAGCCAATTCCCATAACATTATTGTGCCCCAGCTGGATTGCGATTCTGAAATGTAATAATA
AGACGTTCAATGGAACGGGTCCATGCAAAAATGTCAGCACAGTACAGTGTACACATGGAATTAG
15 GCCAGTAGTGCAACTCAACTGTTGTTGAATGGCAGTCTAGCAGAAGAAGAGATAATAATTAGA
TCTGAAAATATCACAAATAATGCAAAAACCATAATAGTACAGCTTAATGAGTCTGTAACAATTG
ATTGCATAAGGCCCAACAACAATACAAGAAAAAGTATACGCATAGGACCAGGGCAAGCACTCT
ATACAACAGACATAATAGGGAATATAAGACAAGCACATTGTAATGTTAGTAAAGTAAAAATGGG
GAAGAATGTTAAAAAGGGTAGCTGAAAAATTAAAAAGACCTTCTTAACCAGACAAAGAACATAA
20 CTTTTGAACCATCCTCAGGAGGGGACCCAGAAATTACAACACACAGCTTTAATTGTGGAGGGGA
ATTCTTCTACTGCAATACATCAGGACTATTTAATGGGAGTCTGCTTAATGAGCAGTTTAATGAGA
CATCAATGATACTCTCACACTCCAATGCAGAATAAAACAAATTATAAACATGTGGCAAGGAGT
AGGAAAAGCAATGTATGCCCCCTCCATTGCAGGACCAATCAGCTGTTTCATCAAAATATTACAGGA
CTATTGTTGACAAGAGATGGTGGTAATACTGGTAATGATTCAGAGATCTTCAGACCTGGAGGGG
25 GAGATATGAGAGACAATTGGAGAAGTGAATTATACAAATATAAAGTAGTAAGAATTGAACCAA
TGGGTCTAGCACCCACCAGGGCAAAAAGAAGAGTGGTGGAAAGAGAAAAAAGAGCAATAGGA
CTGGGAGCTATGTTCCCTTGGGTTCTTGGGAGCGGCAGGAAGCACGATGGGCGCAGCGTCACTGA
CGCTGACGGTACAGGCCAGACAGTTATTGTCTGGTATAGTGCAACAGCAAAACAATTTGCTGAG
AGCTATAGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATTAAACAGCTCCAGGCA
30 AGAGTCCTGGCTATGGAAAGCTACCTAAAGGATCAACAGCTCCTAGGAATTTGGGGTTGCTCTG
GAAAACACATTTGCACCACTACTGTGCCCTGGAACCTACCTGGAGTAATAGATCTGTAGAGGA
GATTTGGAATAATATGACCTGGATGCAGTGGGAAAGAGAAATTGAGAATTACACAGGTTTAATA
TACACCTTAATTGAAGAATCGCAAACCCAGCAAGAAAAGAATGAACAAGAACTATTGCAATTG
GATAAATGGGCAAGTTTGTGGAATTGGTTTAGTATAACAAAATGGCTGTGGTATATAAAAATAT
35 TCATAATGATAGTAGGAGGCTTAATAGGTTTAAGAATAGTTTTTGTGTGCTTTCTTTAGTAAAT
AGAGTTAGGCAGGGATATTCACCTCTGTCTTTTCAGACCTCCTCCAGCCCCGAGGGGACCCG
ACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGCAAGGCTAA

98UG57128 envelope (SEQ ID NO: 55):

ATGAGAGTGAGGGGGATAGAGAGGAATTATCAGCACTTATGGTGGAGATGGGGCACCATGCTC
CTTGGGATATTGATGATATGTAGTGCTGCAGAACAAATTGTGGGTCACAGTTTATTATGGGGTACC
5 TGTGTGGAAAGAAGCAACCACTACTCTATTTTGTGCATCAGATGCTAAAGCATATAAAGCAGAG
GCACACAATATCTGGGCTACACATGCCTGTGTACCAACAGACCCCAACCCACAAGAAATAGTAC
TAGAAAATGTACAGAAAACTTTAACATGTGGAAAAATAGCATGGTGGAGCAGATGCATGAGG
ATGTAATCAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAATTAACCCCACTCTGTGTCACT
TTAAACTGCACTAATGCCACTGCCACTAATGCCACTGCCACTAGTCAAAAATAGCACTGATGGTA
10 GTAATAAACTGTTAACACAGACACAGGAATGAAAACTGCTCTTTCAATGTAACCACAGATCT
AAAAGATAAGAAGAGGCAAGACTATGCACCTTTCTATAAACTTGATGTGGTACGAATAGATGAT
AAGAATACCAATGGTACTAATAACCAACTATAGATTAATAAATTGTAATACCTCAGCCATTACAC
AAGCGTGTCCAAAGATAACCTTTGAGCCAATTCCCATACATTATTGTGCCCCAGCTGGATATGCG
ATTCTAAAATGTAATAATAAGACATTCAATGGGACGGGTCCATGCAAAAACGTCAGCACAGTAC
15 AGTGTACACATGGGATTAGGCCAGTAGTGTCAACTCAACTGTTGTTGAATGGCAGTCTAGCAGA
GGAAGAGATAGTAATTAGATCTGAAAACCTCACAAATAATGCTAAAATTATAATAGTACAGCTT
AATGAAGCTGTAACAATTAATTGCACAAGACCCTCCAACAATACAAGACGAAGTGTACATATAG
GACCAGGGCAAGCAATCTATTCAACAGGACAAATAATAGGAGATATAAGAAAAGCACATTGTA
ATATTAGTAGAAAAGAATGGAATAGCACCTTACAACAGGTAACATAAAAAATTAGGAAGCCTGTT
20 TAACACAACAAAAATAATTTTAAATGCATCCTCGGGAGGGGACCCAGAAATTACAACACACAGC
TTTAATTGTAACGGGGAATTCTTCTACTGCAATACAGCAGGACTGTTTAATAGTACATGGAACA
GGACAAATAGTGAATGGATAAATAGTAAATGGACAAATAAGACAGAAGATGTAATATCACAC
TTCAATGCAGAATAAAACAAATTATAAACATGTGGCAGGGAGTAGGAAAAGCAATGTATGCCC
CTCCCGTTAGTGGAATAATCCGATGTTTCATCAAATATTACAGGACTGTTGCTGACAAGAGATGG
25 TGGTGGTGCAGATAATAATAGGCAGAATGAGACCTTCAGACCTGGGGGAGGAGATATGAGAGA
CAATTGGAGAAGTGAATTATACAAATATAAAGTAGTAAGAATTGAACCACTAGGTATAGCACCC
ACCAAGGCAAGGAGAAGAGTGGTGGAAAGAGAAAAAAGAGCAATAGGACTGGGAGCCTTGTT
CCTTGGGTTCTTGGGAACAGCAGGAAGCACGATGGGCGCAGTGTCAATGACGCTGACGGTACAG
GCCAGACAAGTATTGTCTGGTATAGTGCAACAGCAAAACAATCTGCTGAGGGCTATAGAGGCGC
30 AACAGCATCTGTTGCAACTCACAGTCTGGGGCATTAAACAGCTCCAGGCAAGAATCCTGGCTGT
GGAAAGCTACCTAAAGGATCAACAGCTCCTAGGAATTTGGGGTTGCTCTGGAAAACACATTTGC
ACCACTAATGTGCCCTGGAACCTCTAGCTGGAGTAATAAATCTCTAAATTATATTTGGAATAACAT
GACCTGGATGGAGTGGGAAAAGGAAATTGACAATTACACAGAATTAATATACAGCTTAATTGA
AGTATCGCAAATCCAGCAAGAAAAGAATGAACAAGAAGTATTGAAATTGGACAGTTGGGCAAG
35 TTTGTGGAATTGGTTTAGCATAACAAAATGGCTGTGGTATATAAAAAATATTCATAATGATAGTA
GGAGGCTTGATAGGCTTAAGAATAGTTTTTGTCTGTGCTTTCTTTAGTAAATAGAGTTAGGCAGGG
ATACTCACCTCTGTCGTTTCAGACCCTTATCCCAGCCTCGAGGGGACCCGACAGGCCCGAAGGA
ACAGAAGGAGAAGGTGGAGAGCAAGGCTAA

Appendix 2 - DNA sequences of MVA shuttle plasmids:

pLAS-1 (SEQ ID NO: 56):

5 GAATTCGTTGGTGGTCGCCATGGATGGTGTATTGTATACTGTCTAAACGCGTTAGTAAAACATG
 GCGAGGAAATAAATCATATAAAAAATGATTTTCATGATTAAACCATGTTGTGAAAAAGTCAAGAA
 CGTTCACATTGGCGGACAATCTAAAAACAATACAGTGATTGCAGATTTGCCATATATGGATAAT
 GCGGTATCCGATGTATGCAATTCAGTGTATAAAAAGAATGTATCAAGAATATCCAGATTTGCTA
 10 ATTTGATAAAGATAGATGACGATGACAAGACTCCTACTGGTGTATATAATTATTTTAAACCTAA
 AGATGCCATTCTGTATTATATCCATAGGAAAGGATAGAGATGTTTGTGAAGTATTAATCTCAT
 CTGATAAAGCGTGTGCGTGTATAGAGTTAAATTCATATAAAGTAGCCATTCTTCCCATGGATGTT
 TCCTTTTTTACCAAAGGAAATGCATCATTGATTATTCTCCTGTTTGATTCTCTATCGATGCGGCA
 CCTCTCTTAAGAAGTGTAACCGATAATAATGTTATTATATCTAGACACCAGCGTCTACATGACGA
 GCTTCCGAGTTCCAATTGGTTCAAGTTTTACATAAGTATAAAGTCCGACTATTGTTCTATATTAT
 15 ATATGGTTGTTGATGGATCTGTGATGCATGCAATAGCTGATAATAGAACTTACGCAAATATTAG
 CAAAAATATATTAGACAATACTACAATTAACGATGAGTGTAGATGCTGTTATTTTGAACCACAG
 ATTAGGATTCTTGATAGAGATGAGATGCTCAATGGATCATCGTGTGATATGAACAGACATTGTA
 TTATGATGAATTTACCTGATGTAGGCGAATTTGGATCTAGTATGTTGGGAAATATGAACCTGAC
 ATGATTAAGATTGCTCTTTCGGTGGCTGGGTACCAGGCGCGCCTTTCATTTTGTTTTTTCTATGC
 20 TATAAATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGA
 CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGG
 CAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTG
 ACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACT
 TCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGG
 25 CAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCT
 GAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACAA
 CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGAT
 CCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATC
 GGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAG
 30 ACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCT
 CGGCATGCACGAGCTGTACAAGTAAGAGCTCGGTTGTTGATGGATCTGTGATGCATGCAATAGC
 TGATAATAGAAGTACGCAAATATTAGCAAAAATATATTAGACAATACTACAATTAACGATGAG
 TGTAGATGCTGTTATTTTGAACCACAGATTAGGATTCTTGATAGAGATGAGATGCTCAATGGATC
 ATCGTGTGATATGAACAGACATTGATTATGATGAATTTACCTGATGTAGGCGAATTTGGATCTA
 35 GTATGTTGGGAAATATGAACCTGACATGATTAAAGATTGCTCTTTCGGTGGCTGGCGGCCGCTC
 GAGGCCGCTGGTACCCAACCTAAAAATTGAAAATAAAATACAAAGGTTCTTGAGGGTTGTGTTAA
 ATTGAAAGCGAGAAATAATCATAAATAAGCCCGGGGATCCTCTAGAGTCGACCTGCAGGGAAA
 GTTTTATAGGTAGTTGATAGAACAAAATACATAATTTTGTAAAAATAAATCACTTTTTATACTAA
 TATGACACGATTACCAATACTTTTGTACTAATATCATTAGTATACGCTACACCTTTTCCTCAGAC
 40 ATCTAAAAAATAGGTGATGATGCAACTTTATCATGTAATCGAAATAATACAAATGACTACGTT
 GTTATGAGTGCTTGGTATAAGGAGCCCAATTCCATTATTCTTTTAGCTGCTAAAAGCGACGCTCT
 GTATTTTGATAATTATACCAAGGATAAAATATCTTACGACTCTCCATACGATGATCTAGTTACAA
 CTATCACAATTAATCATTGACTGCTAGAGATGCCGGTACTTATGTATGTGCATTCTTTATGACA
 TCGCCTACAAATGACACTGATAAAGTAGATTATGAAGAATACTCCACAGAGTTGATTGTAAATA
 45 CAGATAGTGAATCGACTATAGACATAATACTATCTGGATCTACACATTCACCAGAACTAGTTA
 AGCTTGTCTCCCTATAGTGAGTCGTATTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
 GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCC
 TGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCGAGTC
 GGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCT
 50 ATTGGGCGCTCTTCCGCTTCCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGC
 GGTATCAGCTCACTCAAAGGCGGTAATACGTTATCCACAGAATCAGGGGATAACGCAGGAAA
 GAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGCTGGCGTT
 TTTCGATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGA
 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGT
 55 TCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTC
 ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAC
 GAACCCCCCGTTACGCCCCGACCGCTGCGCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGGT
 AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
 AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT

GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
5 ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACGCGAG
ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTA
10 AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGGCATTGCTACAGGCATCGTGTTGTACG
CTCGTCGTTTGGTATGGCTTCATTAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCC
GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAG
ATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
15 GTTGCTCTTGCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCT
CATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCACTGATCTTCAGCATCTTTTACTTTACCCAGCGTTTCTGGG
TGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTG
AATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGG
20 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA
GTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCA
CGAGGCCCTTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACTCTGACACATGCAGCTCCCG
GAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCA
GCGGGTGTGGCGGGTGTGCGGGCTGGCTTAAGTATGCGGCATCAGAGCAGATTGTAAGTACTGAGAG
25 TGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC
ATTCGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACG
CCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGGTTTTCCAG
TCACGACGTTGTAAACGACGGCCAGTGAATTGGATTAGGTGACACTATA

pLAS -2 (SEQ ID NO: 57):

CCTCCTGAAAACTGGAATTTAATACACCATTTGTGTTTCATCATCAGACATGATATTACTGGATT
TATATTGTTTATGGGTAAAGGTAGAATCTCCTTAATATGGGTACGGTGTAAAGGAATCATTATTTTA
5 TTTATATTGATGGGTACGTGAAATCTGAATTTTCTTAATAAAATATTATTTTTATTAAATGTGTATA
TGTTGTTTTGCGATAGCCATGTATCTACTAATCAGATCTATTAGAGATATTATTAATTCTGGTGC
AATATGACAAAAATTATACACTAATTAGCGTCTCGTTTCAGACATGGATCTGTCACGAATTAATA
CTTGGAAGTCTAAGCAGCTGAAAAGCTTTCTCTCTAGCAAAGATGCATTTAAGGCGGATGTCCA
TGGACATAGTGCCTTGTATTATGCAATAGCTGATAAATAACGTGCGTCTAGTATGTACGTTGTTGA
10 ACGCTGGAGCATTGAAAAATCTTCTAGAGAATGAATTTCCATTACATCAGGCAGCCACATTGGA
AGATACCAAATAGTAAAGATTTTGCTATTTCAGTGGACTGGATGATTTCGAGGTACCAGGCGCGC
CCTTTCATTTTGTTTTTTCTATGCTATAAATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGT
GGTGCCCATCCTGGTCGAGCTGGACGCGCAGCTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGC
15 CCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGCAGTGTTCAGCCGCTACCCC
GACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC
CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCA
CAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGG
20 CATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCAC
TACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCA
CCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGT
GACCGCCGCGGGATCACTCTCGGCATGCACGAGCTGTACAAGTAAGAGCTCGCTTCTCTCTA
GCAAAGATGCATTTAAGGCGGATGTCCATGGACATAGTGCCTTGTATTATGCAATGCTGATAA
25 TAACGTGCGTCTAGTATGTACGTTGTTGAACGCTGGAGCATTGAAAAATCTTCTAGAGAATGAA
TTTCCATTACATCAGGCAGCCACATTGGAAGATACCAAAATAGTAAAGATTTTGCTATTTCAGTG
GACTGGATGATTCTCCGGATGGTACCCAACCTAAAAATTGAAAAATAATACAAAGGTTCTTGAG
GGTTGTGTTAAATTGAAAGCGAGAAATAATCATAAATAAGCCCGGGGATCCTCTAGAGTCGACC
TGACAGCATGCTCGAGCGGCCCGCCAGTGTGATGGATATCTGCAGAATTTCGGCTTGGGGGGCTGC
30 AGGTGGATGCGATCATGACGTCCTCTGCAATGGATAACAATGAACCTAAAGTACTAGAAATGGT
ATATGATGCTACAATTTTACCCGAAGGTAGTAGCATGGATTGTATAAACAGACACATCAATATG
TGTATACAACGCACCTATAGTTCTAGTATAATTGCCATATTGGATAGATTCTAATGATGAACAA
GGATGAACTAAATAATACACAGTGTATATAATTAAAGAATTTATGACATACGAACAAATGGCG
ATTGACCATTATGGAGAATATGTAAACGCTATTCTATATCAAATTCGTAAAAGACCTAATCAAC
35 ATCACACCATTAATCTGTTTAAAAAATAAAAAAGAACCCGGTATGACACTTTTAAAGTGGATCC
CGTAGAATTCGTAAAAAAGTTATCGGATTTGTATCTATCTTGAACAAATATAAACCGGTTTATA
GTTACGTCCTGTACGAGAACGTCCTGTACGATGAGTTCAAATGTTTCATTGACTACGTGGAACT
AAGTATTTCTAAAATTAATGATGCATTAATTTTTGTATTGATTCTCAATCCTAAAACTAAAAATA
TGAATAAGTATTAACATAGCGGTGTACTAATTGATTTAACATAAAAAATAGTTGTTAACTAAT
40 CATGAGGACTCTACTTATTAGATATATTCTTTGGAGAAATGACAACGATCAAAACCGGCATGCA
AGCTTGTCTCCCTATAGTGAGTCGTATTAGAGCTTGGCGTAATCATGGTCTATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGGCC
TGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCGAGTC
GGGAAACCTGTGTCGAGCTGCATTAAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCT
45 ATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGC
GGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA
GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCGATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGT
50 TCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGTTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCGGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
55 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
AACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATA
ATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA

TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG
ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA
5 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTA
AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGGCATTGCTACAGGCATCGTGGTGTACAG
CTCGTCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCC
GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTGATGCCATCCGTAAG
10 ATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
GTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCT
CATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTCTGGG
TGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAAATGTTG
AATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
15 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA
GTGCCACCTGACGTCTAAGAAACCATTTATTATCATGACATTAACTATAAAAAATAGGCGTATCA
CGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCG
GAGACGGTCACAGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCA
GCGGGTGTTGGCGGGTGTGCGGGCTGGCTTAAGTATGCGGCATCAGAGCAGATTGTAAGTGTGAGAG
20 TGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCC
ATTCGCCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACG
CCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTCCCAG
TCACGACGTTGTAAAACGACGGCCAGTGAATTGGATTTAGGTGACACTATAGAATACGAATTC

Appendix 3-- DNA sequences of gagpol and env genes from Kenyan HIV-1 clade A isolates:

KER2008 gagpol (SEQ ID NO: 58):

5 ATGGGTGCGAGAGCGTCAGTATTAAGTGGGGGAAAAATTAGATGCATGGGAGAAAAATTCGGTTA
AGGCCAGGGGAAAGAAAAAATATAGACTGAAACACTTAGTATGGGCAAGCAGGGAGCTGGA
AAAATTCGTACTTAACCCTAGCCTTTTAGAAACTTCAGAAGGATGTCAGCAAATAATGAACCA
10 ATACAACCAGCTCTTCAGACAGGAACAGAAGAACTTAGATCATTATTTAATGCAGTAGCAACCC
TCTATTGTGTACATCAACGGATAGAGGTAAAAGACACCAAGGAAGCTTTAGATAAAGTAGAGG
AAATACAAAACAAGAGCAAGCAAAAAGACACAACAGGCAGCAGCTGATACAGGAAACAACAGC
AAGGTCAGCCATAATTACCCTATAGTGCAAAATGCACAAGGGCAAATGATACATCAGTCCTTAT
CACCAAGGACTTTGAATGCATGGGTAAAGGTAATAGAAGAAAGGGGTTTCAGCCCAGAAAGTAA
15 TACCCATGTTCTCAGCATTATCAGAAGGAGCCATCCCACAAGATTTAAATATGATGCTGAACAT
AGTGGGGGGACACCAGGCAGCTATGCAAATGTTAAAAGAAACTATCAATGAGGAAGCTGCAGA
ATGGGACAGGTTACATCCAGCACAGGCAGGGCCTATTCCACCAGGCCAGATAAGAGACCCAAAG
GGGAAGTGACATAGCAGGAACCTAGTACCCCTCAGGAACAAATAACATGGATGACAAACAA
CCCACCTATCCCAGTGGGAGACATCTATAAAAGATGGATAATCCTAGGATTAAATAAAATAGTA
AGAATGTATAGCCCTGTTAGCATTTTAGATATAAAACAGGGGCCAAAAGAACCCTTCAGAGACT
20 ATGTAGATAGGTTCTTTAAAGTTCTCAGAGCCGAACAAGCTACACAGGAAGTAAAAGGCTGGAT
GACAGAGACCCTGCTGTTCAAAATGCAAATCCAGATTGTAAGTCCATTTTAAGAGCATTAGGA
ACAGGGGCTCATTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGAGGACCCGGCCATAAA
GCAAGGGTTTATGCTGAGGCAATGAGTCAAGCACAAACAGGCAAATGTAATGATGCAGAGGGGC
AGCTTTAAGGGGCAGAAAAGAATTAAGTGCTTCAACTGTGGCAAAGAGGGACACCTAGCCAGA
25 AATTGCAGAGCCCCTAGGAAAAAAGGCTGTTGGAAGTGTGGGAAAGAAGGACACCAAATGAAA
GATTGCAATGAGAGACAGGCTAATTTTTTAGGGAAAAATTTGGCCTTCCAGCAAGGGGAGGCCAG
GAAATTTTCCCCAGAGCAGACCGGAGCCAACAGCCCCACCAGCAGAGATCTTTGGGATGGGGG
AAGAGATAACCTCCCCTCCGAAGCAGGAGCAGAAAGAGAGGGAACAAACCCCAACCTTTGTTT
CCCTCAAATCACTCTTTGGCAACGACCCGTTGTACAGTAAAAGTAGGAGGAGAAATGAGAGAA
30 GCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGATATAAATTTGCCAGGAAAAATGGA
AACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAATATGATCAGGTATCTAT
AGAAATTTGTGGAaaaaaAGGCTATAGGTACGGTATTAGTAGGACCTACACCTGTCAACATAATT
GGAAGAAATATGTTGACTCAGATTGGTTGTACCTTAAATTTTCCAATTAGTCCTATTGAGACTGT
ACCAAGTAACATTAAGGCCAGGAATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGA
35 GAAAAATAAAGCATTGACAGAAATTTGTAAAGAGATGGAAGGAAGGAAAAATTTCAAAAAT
TGGGCCTGAAAATCCATACAATACTCCAATTTTGCAATAAAGAAAAAGATAGCACTAAATGG
AGGAAATTAGTAGATTTTCAGAGAGCTCAATAAAAAGAACACAAGACTTTTGGGAAGTTCAATTAG
GGATACCGCATCCAGCGGCCCTAAAAAAGAAAAAATCAGTAACAGTACTAGAGGTGGGGGATG
CATATTTTTCAGTTCCCCTAGATAAAAACTTTAGAAAGTATACTGCATTTACCATACCTAGTTTA
40 AATAATGAAAACACCAGGAATCAGGTATCAGTACAATGTGCTTCCACAAGGATGGAAGGATCA
CCAGCAATATTCCAGTGCAGTATGACAAAAATCTTAGAGCCCTTTAGATCAAAAAATCCAGAAA
TAATTATCTATCAATACATGCACGACTTGTATGTAGGATCAGATTTAGAAATAGGGCAGCATAG
AGCAAAAATAGAAGAATTAAGAGCTCATCTACTGAGCTGGGGATTTACTACACCAGACAAAAA
GCATCAGAAAGAACCTCCATTCTTTGGATGGGATATGAGCTCCATCCTGACAAGTGGACAGTC
45 CAGCCTATAGAGCTGCCAGAAAAAGAAAGCTGGACTGTCAATGATATACAGAAATTAGTGGGA
AAACTAAATTGGGCCAGTCAAATTTATCCAGGAATTAAGTAAAGCAATTGTGTAAACTTCTCA
GGGGAGCCAAAGCCCTAACAGATATAGTAACACTGACTGAGGAAGCAGAATTAGAATTAGCAG
AGAACAGGGAGATTCTAAAAGACCCTGTGCATGGGGTATATTATGACCCATCAAAAGACTTAAT
AGCAGAAATACAGAAACAAGGGCAAGACCAATGGACATACCAAATTTATCAGGAGCCATTTAA
50 AAATCTAAAAACAGGAAAAATATGCAAGAAAAAGGTCTGCTCACACTAATGATGTAAGACAATT
AGCAGAAAGTAGTGCAAGAAAGTGGTCATGGAAAGCATAGTAATATGGGGAAAGACTCCTAAATT
TAAACTACCCATACAAAAAGAGACATGGGAGACATGGTGGATGGACTATTGGCAAGCTACCTG
GATTCCTGAGTGGGAGTTTGTCAATACCCCTCCCCTAGTAAAATTATGGTACCAGTTAGAGAAA
GACCCCATAGCAGGAGCAGAGACTTTCTAA

KNH1144 envelope (SEQ ID NO: 59):

ATGAGAGTGATGGGGATACAGATGAATTGTCAGCACTTATTGAGATGGGGAACTATGATCTTGG
GATTGATAATAATCTGTAATGCTGTAAACAGCAACTTGTGGGTTACTGTCTATTATGGGGTACCT
5 GTGTGGAAAGATGCAGAGACCACCTTATTTTGTGCATCAGATGCTAAAGCATATAAAACAGAAA
AGCATAATGTCTGGGCTACACATGCCTGTGTGCCCACAGACCCCAACCCACAAGAAATACCTTT
GGAAAATGTGACAGAAGAGTTTAAACATGTGGAAAAATAAAATGGTAGAACAAATGCATACAGA
TATAATCAGTCTATGGGACCAAAGCCTACAGCCATGTGTAAAGTTAACCCTCTCTGCATTACTT
10 TAAACTGTACAGATGTTACTAATGTTACAGATGTTAGTGGTACGAGGGGCAACATCACCATCAT
GAAAGAGATGGAGGGAGAAAATAAAAACTGTTCTTTCAATATGACCACAGAAAATAAGGGATAA
GAAACAGAAAAGTATATTCACTCTTTTATAGACTTGATGTAGTACCAATAAATCAGGGTAATAGT
AGTAGTAAAAACAGTAGTGAGTATAGATTAATAAGTTGTAATACCTCAGCCATTACACAAGCTT
GCCCAAAGGTAAGCTTTGAGCCAATTCCCATACATTATTGTGCCCCAGCTGGTTTTTGCGATCCTG
AAGTGTAGGGATAAGGAGTTCAATGGAACAGGGGAATGCAAGAATGTCAGCACAGTCCAATGC
15 ACACATGGAATCAAGCCAGTAGTATCAACTCAACTACTGTTAAATGGCAGTCTAGCAGAAGAAA
AGGTAAAAATCAGAACTGAAAATATCACAAACAATGCCAAAACCTATAGTAGTACAACCTTGTCG
AGCCTGTGAGAATTAATTGTACTAGACCTAATAACAATACAAGAGAGAGTGTGCGTATAGGGCC
AGGACAAGCATTCTTTGCAACAGGTGACATAATAGGGGATATAAGACAAGCACATTGTAATGTC
AGTAGATCACAATGGAATAAGACTTTACAACAGGTAGCTGAACAATTAAGAGAACACTTTAAA
20 AAAAAACAATAATATTTAACAGTTCCTCAGGAGGGGATCTAGAAATCACAACACATAGTTTCA
ATTGTGGAGGAGAATTCTTCTATTGTAATACATCAGGTCTGTTCAATAGCACCTGGAATAACCAGC
ATGTCAGGGTCAAGTAACACGGAGACAAATGACACTATAACTCTCCAATGCAGAATAAAGCAA
ATTATAAATATGTGGCAGAGAACAGGACAAGCAATATATGCCCCCTCCCATCCAGGGAGTGATAA
25 GGTGTGAATCAAACATCACAGGACTACTGTTAACAAGAGATGGTGGGGAGGAGAAGAACAGTA
CAAATGAAATCTTCAGACCTGGAGGAGGAGATATGAGGGACAACCTGGAGAAGTGAATTATATA
AGTATAAAGTAGTAAAAAATTGAACCACTAGGAGTAGCACCACCAGGGCAAGGAGAAGAGTGG
TGGGAAGAGAAAAAAGAGCAGTTGGAATAGGAGCTGTTTTCTTGGGTTCTTAGGAGCAGCAG
GAAGCACTATGGGCGCGGCGTCAATAACGCTGACGGTACAGGCCAGGCAATTATTGTCTGGCAT
AGTGCAGCAGCAGAGCAATTTGCTGAGGGCTATAGAGGCTCAACAACATATGTTGAAACTCACG
30 GTCTGGGGCATTAAACAGCTCCAGGCAAGAGTCTTGTGTGGAAAGATACCTAAGGGATCAAC
AGCTCCTAGGAATTTGGGGCTGCTCTGGAAAACCTCATCTGCACCACTAATGTGCCCTGGAACTCT
AGTTGGAGTAATAAATCTCAGGATGAAATATGGAACAACATGACCTGGCTGCAATGGGATAAA
GAAATTAGCAATTACATAAACCTAATATATAGTCTAATTGAAGAATCGCAAAACCAGCAGGAAA
AGAATGAACAAGACTTATTGGCATTGGGCAAGTGGGCAAACTGTGTGGACTTGGTTTGACATATC
35 AAATTGGCTGTGGTATATAAGAATATTTATAATGATAGTAGGAGGCTTAATAGGATTAAGAATA
GTTTTTGCTGTGCTTGCTGTAATAAAGAGAGTTAGGCAGGGATACTCACCTGTGTCATTTAGAT
CCATGCCCCAAACCCAGGGGGTCTCGACAGGCCCGGAAGAATCGAAGGAGAAGGTGGAGAGCA
AGACTAA

KNH1207 envelope (SEQ ID NO: 60)

ATGAGAGTGATGGGGATACAGATGAATTGTCAAAGCTTGTGGAGATGGGGAACTATGATCTTGG
GAATGTTAATGATTTGTAGTGTTGCAGGAAACTTGTGGGTACTGTCTACTATGGGGTACCTGTG
5 TGGAAAGAGGCAGACACCACCTTATTTTGTGTATCAAATGCTAGAGCATATGATACAGAAGTGC
ATAATGTCTGGGCTACACATGCCTGTGTACCTACGGACCCCAACCCACAAGAAATAGATTTGGA
GAATGTGACAGAAGAGTTTAACATGTGGAATAAATGCTAAAGTTAACCCCTCTCTGCGTTACTTTAG
AATTAGTCTATGGGACCAAAGCCTAAAACCATGTGTAAAGTTAACCCCTCTCTGCGTTACTTTAG
ATTGTGGCTATAATGTAACCAACTGAATTTTACCAGTAACATGAAAGGAGACATAACAAACTG
10 CTCTTACAATATGACCACAGAAATAAGGGATAGGAAACAGAAAGTGATTCACTTTTCTATAGG
CTTGATATAGTACCAATTAATGAAGAAAAGAATAATAGCAGGGAGACTAGTCCGTATAGATTAA
TAAATTGTAATACCTCAGCCATTACACAAGCTTGTCTTAAGGTATCTTTTGAACCAATTCCCATA
CATTATTGTGCCCCAGCCGGTTTTGCGATTCTAAAATGTAAGGATGCAGAGTTCAATGGAACAG
GGCCATGCAAGAATGTCAGCACAGTACAATGTACACATGGAATCAGGCCAGTAATATCAACTCA
15 ACTGCTGTAAATGGCAGTTTAGCAGAGAATGGGACAAAGATTAGATCTGAAAATATCACAAAC
AATGCCAAAACCATAATAGTACAACCTAACGAGACCGGTACAAATTAATTGTACCAGACCTAGCA
ACAATACAAGAAAAAGTGACGTATAGGACCAGGACAAGCATTCTATACAACAGGTGATATAA
CAGGGGATATAAGACAAGCATATTGTAATGTCAAGTACAGAAGTGGGAACAAGCATTAAGG
GGGTAGTTATACAATTAAGAAAACACTTTAACAAAACAATAATCTTTAACAGTTCCTCAGGAGG
20 GGATTTAGAAATTACAACACATAGTTTTAATTGTGGAGGAGAATTCTTCTATTGTGATACATCAG
GCCTGTTTAATAGCACCTGGAACACGAACACCACCGAGCCAAACAACACAACGTCAAATGGCA
CTATCATTCTCCAATGCAGAATAAAGCAAATTATAAATCTGTGGCAGAGAACCGGACAAGCAAT
GTATGCCCCCTCCCATCCAAGGGGTAATAAGGTGTGATTCCAACATTACAGGACTACTATTAACA
AGAGATGGTGGAGTAGTTGATAGTATAAATGAAACCGAAATCTTCAGACCTGGAGGAGGAGAT
25 ATGAGGGACAATTGGAGAAGTGAATTATATAAGTATAAAGTAGTAAAAATTGAACCACTAGGA
GTAGCACCCACCGGGGCAAAGAGAAGAGTGGTGGAGAGAGAAAAAAGAGCAGTTGGCATAGG
AGCTGTATTCAATTGGGTTCTTAGGAGCAGCAGGAAGCACTATGGGCGCGCGCTCAATAACGCTG
ACGGTACAGGCCAGACAATTATTGTCTGGCATAAGTGAACAGCAAAGCAATTTGCTGAGGGCTA
TAGAGGCTCAACAGCATATGTTGAGACTCACGGTCTGGGGCATTAAAGCAGCTCCAGGCAAGAGT
30 CCTGGCTGTGGAAGATACCTAAGGGATCAACAGCTCCTAGGAATTTGGGGCTGCTCTGGAAAA
CTCATCTGCACCACTAATGTGCCCTGGAACCTCTAGTTGGAGTAATAAATCTCAGGAGGAAATAT
GGGGTAACATGACCTGGCTGCAATGGGATAAAGAAATTAGCAATTACACACAAACAATATATA
ACCTACTTGAAGAATCGCAGAACCAGCAGGAAAAGAATGAACAAGACTTATTGGCATTGGACA
AGTGGGCAAATTTGCGGACTTGGTTTGACATAACAAATTGGCTGTGGTATATAAAAAATGTTTAT
35 AATGATAGTAGGAGGCTTAATAGGATTAAGAATAGTTTTTGTCTGTGCTTTCTGTAATAAATAGA
GTTAGGCAGGGATACTCACCTCTGTCTGTTTCAGACCCATATCCCGAGCCCAAGGGGTCTCGATA
GGCCCCGAAGAATCGAAGGAGAAGGTGGAGAGCAAGACTAA

Appendix 4 – DNA sequences of gagpol and env genes from Tanzanian HIV-1 clade C isolates:

TZA-246 gagpol (SEQ ID NO: 61):

5 ATGGGTGCGAGAGCGTCAATATTAAGAGGGGGAAAATTAGATCGATGGGAAAAAATTAGGTTA
AGGCCAGGGGGAAAAGAAAAGCTATATGATAAAACACCTTAGTATGGGCAAGCAGGGAGCTGGAA
AGATTTGCACTTAACCCCTAGCCTTTTAGAGACATCAGAAGGCTGTAAACAAATAATGAAACAGC
10 TACAACCAGCTCTTCAGACAGGAACAGAAGAACTTAAATCATTATTCAATGCAATAGCAGTTCT
CTATTGTGTACATGAAGGGATAGATGTAAAAGACACCAAGGAAGCCTTAGACAAGATAGAGGA
AGAACAGAACAAAAGTCAGCAAAAAACACAGCAGGCAGAAGCAGCTGGCGGAAAAGTCAGTC
AAAATTATCCTATAGTGCAGAATCTCCAAGGACAAATGGTACACCAGTCCATATCACCTAGAAC
TTTGAATGCATGGGTAAAAGTAATAGAGGAAAAGGCTTTTAGCCCAGAGGTAATACCCATGTTT
15 ACAGCATTATCAGAAGGAGCCACCCACAAAGATTTAAACACCATGCTAAATACAGTGGGGGGA
CATCAAGCAGCCATGCAAAATGTTAAAAGATACCATCAATGAGGAGGCTGCAGAATGGGATAGG
ATACATCCAGTACATGCAGGGCCTACTGCACCAGGCCAAATGAGAGAACCAAGGGGAAGTGAC
ATAGCAGGAACCTACTAGTACCCTTCAGGAACAAATAGCATGGATGACAGCTAACCCACCTGTTC
CAGTGGGAGAAATCTACAAAAGATGGTAATACTGGGTTTAAATAAAATAGTAAGAATGTATA
GCCCTGTGAGCATTTTGGACATAAAACAAGGGCCAAAGGAACCCCTTAGAGACTATGTAGATCG
20 GTTCTTTAAAACCTTTAAGAGCTGAACAGGCTACACAAGATGTAAAAAATTGGATGACAGACACC
TTGTTGGTCCAAAATGCGAAGCCAGATTGTAAGACCATTTTAAAGAGCATTAGGACCAGGGGCTA
CATTAGAAGAAATGATGACAGCATGTCAAGGAGTGGGAGGACCTGGCCACAAAGCCAGAGTTT
TGGCTGAGGCAATGAGCCAAGCAAACACACACATAATGATGCAGAGAAGCAATTTTAAAGGCT
CTAAAAGAATTGTTAAATGTTTCAACTGTGGCAAGGAAGGGCACATAGCCAGAAATTGCAGGGC
25 CCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGACTGTACTGA
GAGGCAGGCTAATTTTTTAGGGAAAATTTGGCCTTCCCACAAGGGGAGGCCAGGGAATTTCTT
CAGAACAGGTCAGAGCCAACAGCCCCACCAACGAACAGGCCAGAGCCAACAGCTCCACCAGCA
GAGAGCTTCAGGTTTCGAGGAAGCAACCCCTGCTCCGAAGCAGGAGCTGAAAGACAGGGAACCT
TTAATTTCCCTCAAATCACTCTTTGGCAGCGACCCCTCGTCTCAATAAAAGTAGGGGGTCAAACA
30 AAGGAGGCTCTTTTAGACACAGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGCCAGGA
AAATGGAAACCCAAAATGATAGGAGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAG
ATAGTTATAGAAATTTGTGGAAAAAAGGCTATAGGTACAGTATTAGTAGGACCCACCCCTGTCA
ACATAATTGGAAGAAATATGTTGACTCAGCTTGGATGCACACTAAATTTTCCAATTAGTCCATT
GAACTGTACCAGTAAAGTTAAAGCCAGGAATGGATGGCCAAAGGTTAAACAATGGCCATTG
35 ACAGAAGAAAAATAAAGGCATTAAACAGCAATTTGTGAAGAAATGGAGAAGGAAGGAAAAAT
TACAAAGATTGGGCCTGAAAAATCCATATAACACTCCAGTATTTGCCATAAAAAAGAAGGACAGT
ACTAAGTGGAGAAAATTAGTAGATTTAGGGAAACGCAATAAAAGAACTCAAGATTTTTGGGAA
GTTCAATTAGGCATACCACACCCAGCAGGGTTAAAAAAGAAAAAATCAGTGACAGTACTGGAG
GTGGGGGATGCATACTTCTCAGTTCCTTTAGATGAAGGCTTCAGGAAATATACTGCATTACCAT
40 ACCTAGTATAAACAATGAAACACCAGGAATTAGATATCAATACAATGTGCTTCCACAGGGATGG
AAAGGATCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCACAAA
ATCCAGAAATAGTCATCTATCAATATATGCACGACTTATATGTAGGATCTGACTTAGAAATAGG
GCAACATAGAGCAAAAATAGAGGAATTAAGAGAACATCTATTAAAGTGGGGATTTACCACACC
AGACAAGAAACATCAGAAAGAACCCCATTTCTTTGGATGGGGTATGAACTCCATCCTGACAAA
45 TGGACAGTACAGCCTATAACGCTGCCAGAAAAGGAAAGCTGGACTGTCAATGATATACAGAAG
TTAGTGGGAAAACCTAACTGGGCAAGTCAGATTTATGCAGGGATTAAAGTAAGGCAACTGTATA
AACTCCTTAGGGGAGCCAAAGCACTAACAGACATAGTACCACTAACTGAAGAGGCAGAAATTAG
AATTGGCAGAGAACAGGGAAATCTAAAAAGAACAGTACATGGGGTATATTATGACCCATCAA
AAGACTTGATAGCTGAAATACAGAAACAAGGGCATGACCAATGGACATATCAAATTTACCAAG
50 AACCATTCAAAAATCTGAAAACAGGGGAAGTATGCAAAAATGAGGAGTGCCACACTAATGATG
TAAACAATTAACAGAGGCAGTGCAAAAAATAGCCATGGAAGGCATGAATATGGGGAAAGA
CTCCTAAATTTAGACTGCCATTCAAAGGAAACATGGGAAACATGGTGGACAGACTATTGGCA
AGCCACCTGGATTCTGAGTGGGAGTTTGTTAATACCCCTCCCTAGTAAAAATTATGGTACCAGC
TGGAGAAAGAACCCATAGTAGGAGCAGAACTTTC

TZA-125 envelope (SEQ ID NO: 62):

ATGAGAGTGAAGGGGATATTGAGGAATTGGCAACACAGGTGGATATGGATCTGGATCATCTTAG
GCTTTTGGATGCTAATGATTTGTAATGGGAACTTGTGGGTCACTGTCTACTATGGGGTACCTGTG
5 TGGAAAGAAGCAAAATGCTCCTCTATTTTGTGCATCAGATGCTAAAGCATATGAGAAAGAAGTGC
ATAATGTCTGGGCTACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTACTGAGTGGT
AAATGTAAACAGAAAAATTTTAACATGTGGAAAAATGACATGGTAGATCAGATGCATGAGGATAT
AATCAGTTTATGGGATGAAAGCCTAAAGCCATGTGTAAAGTTGACCCCACTCTGTGTCACCTTA
AACTGTACTAATGCTAATATTAATAATGATACTGTTGCTAATAGTGGTACTTTTAAGGTTGATAA
10 TAGTAGTAATGTAGTAAAAAATTGCTCTTTCAATATAACCACAGAAATAAGAGATAAGAAGAAA
AAAGAATATTCAATTGTTTTATAGACTTGATATATTACCACTTGATAACTCTAGTGAGTCTAAGAA
CTATAGTGAGTATGTATTATAAATTGTAATGCCTCAACCGTAACACAAGCCTGTCCAAAGGTCT
CTTTTGACCCAATTCCTATACATTATTGTGCTCCAGCTGGTTATGCGATTCTAAAGTGTAAGAT
AAGACATTCAATGGAACAGGACCATGCAGTAATGTCAGCACAGTACTATGTACACATGGAATTA
15 AGCCAGTGGTATCAACTCAATTACTGTTAAATGGTAGCCTAGCAGAAGAAGGGATAGTAATTAG
ATCTGAAAATCTGACAAACAATGCCAAAACAACAATAGTACAGCTTAATGAACCTGTAGAAATT
ATGTGTGTAAGACCCGGCAATAATACAAGAAAAAGTGTGAGGATAGGACCAGGACAAACATTC
TATGCAACAGGAGGCATAATAGGAGATATAAGACAAGCACATTGTAACATTAGTAGAAGTGAT
TGGAATAAAACTTTACAAGAGGTAGGTAAAAAATTACGAGAATACTTCCACAATAAAACAATA
20 AGATTTAAACCGGCGTGTAGGAGGGGACCTGGAAATTACAACACATAGCTTTAATTGTAGAG
GAGAATTCTTCTATTGCAATACATCAGAACTGTTTACAGGTGAATATAATGGTACTGAGTATAA
GAATACTTCAAATTCAAATCCTAACATCACACTCCCATGTAGAATAAAACAATTTGTAAACATG
TGGCAGAGGGTAGGACGAGCAATGTATGCCCTCCTATTGAAGGAAACATAACATGTAACCTCAA
GTATCACAGGACTACTATTGACATGGGATGGAGGAAACAATACTAATGGCACAGAGACATTTAG
25 ACCTGGAGGAGGAGATATGAGGGATAATTGGAGAAGTGAATTATATAAATATAAAGTGGTAGA
AATTAAACCATTAGGAATAGCACCCACTAGTGCAAAAAGGAGAGTGGTGGAGAGAGAGAAAAAG
AGCAGTGGGAATAGGAGCTTTGTTCCCTTGGGTTCTTAGGAGCAGCAGGAAGCACTATGGGCGCA
GCATCAATAACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAACAGCAAAGCA
ATTTGCTGAGGGCCATAGAGGCGCAACAGCATATGTTGCAACTCACAGTCTGGGGCATTAACA
30 GCTCCAGACAAGAGTCCTGGCTATAGAAAGATACCTAAAGGATCAACAGCTCCTAGGGATTTGG
GGCTGCTCTGAAAACTCATCTGCACCACTGCTGTGCCTTGGAACTAGTTGGAGTAATAAAAA
CTGAACAGGACATTTGGAATCTAACCTGGATGCAGTGGGATAGAGAAGTTAGTAATTACACAGA
CATAATATACAGGTTGCTTGAAGACTCACAAATCCAGCAGGAAAACAATGAAAAGGATTTACTA
GCATTGGACAGTTGGAAAAATCTGTGGAATTGGTTTGACATAACAAATTGGTTGTGGTATATAA
35 GAACATTCATAATGATAGTAGGAGGCTTGATAGGCTTAAGGATAATTTTTGCTGTAATTTCTATA
GTGAATAGAGTTAGGCAGGGATACTCACCTTTGTCATTCAGACCCCTACCCCAACCCCGAGGG
GACCAGAAAGGCTCGGAGGAATCGAAGAAGAAGGTGGAGAGCAAGACTAA

Appendix 5

Budapest Treaty Deposits

ATCC™

American Type Culture Collection

10801 University Blvd., Manassas, VA 20110-2209
 Phone (703) 368-2700; fax (703) 368-2745; e-mail: apotted-ed@atcc.org

TO DEPOSIT OR TO CONVERT A DEPOSIT TO MEET THE REQUIREMENTS OF THE BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

ALL QUESTIONS MUST BE COMPLETED IN ENGLISH. PLEASE USE ONE FORM FOR EACH STRAIN DEPOSITED.

1. Name of deposit. If microorganism, give complete scientific name including genus and species and source of material; if virus, give name, whether plant or animal, and source including geographic location; if cell line, give tissue and species, geographical source of isolation, and any known hazards associated (HIV, EBV, etc.); if genetic materials, give name of organism from which vector, clone or library is derived, source of the DNA insert identified by species (e.g. human, mouse) or scientific name, and give name of gene and identity of the host organism; if consortia or mixed culture, each component of the mixture must be identified; if seeds, embryos, insect eggs, etc., give common name, scientific name, and geographical source.

Vaccinia: animal/Viral Family: Poxviridae, Genus: Orthopoxvirus

2. Strain designation (i.e., number, symbols, etc.) attenuated strain Modified Vaccinia Ankara designated JANA 1574M11 Clone 1 (Serial passage in chick embryo fibroblasts (CEF) of CVA (Ankara) strain in the laboratory of Dr. Adrian Alpar (Harvested 2/22/74 at passage 172) and plaque-purified 3x at the National Institutes of Health with constant measures). The strain designation must correspond with the vial labels.

3. Is this an original deposit under the Budapest Treaty? Yes
4. Is this a request for a conversion of a deposit already at ATCC to meet the requirements of the Budapest Treaty? If so please indicate ATCC designation. No
5. Is this deposit a mixture of microorganisms or cells? No
6. Provide details necessary to cultivate, test for viability, and store deposit. If mixture, provide description of components and a method to check presence. If plasmid, provide name of host and antibiotic resistance. Cultured in CEF or HEp-21 cells in E-MEM with 2% fetal calf serum. CPE of strains and rounded up cells appeared by 3 days. Virus is stored at -80C.

7. Provide sufficient description so that ATCC may confirm deposit properties (e.g., Gram negative rod).

Grow MVA in CEF cells, and at 2-3 days, look by electron microscopy for vaccinia-like particles. To differentiate between MVA and replication competent vaccinia strains such as Copenhagen and WR, look at CPE in CEF and various human cell lines (MRC-5 and HeLa). MVA grows in CEF, and in human cells, while replication competent vaccinia grows in all three cell lines.

- a. If deposit is a cell culture, is it being cultured in the presence of a stabilizer? If so, please list the stabilizer.

MVA

- b. If deposit is hybridoma, what is the isotype of antibody produced? N/A

8. Is the strain hazardous to humans? No Animals? No Plants? No If yes, what is the recommended biosafety level for working with this strain? BSL-2 (vaccinia MVA approved for BSL-2 use where no other vaccine viruses are being manipulated - NIH Institutional Biosafety Committee). (Refer to Biosafety in Microbiological and Biomedical Laboratories, 4th ed. U.S. Dept. of Health and Human Services at www.cdc.gov/od/ohrt/BSL4BSL2.html.)

9. Availability: Prior to the issuance of a U.S. Patent, ATCC will only make a culture available as instructed by the depositor or relevant patent office. Samples must be provided to a specific investigator if a pertinent patent office or the Budapest Treaty instructs ATCC to do so. The following questions must be answered.

- a) As of date of deposit or conversion to meet the requirements of the Budapest Treaty, do you wish the Deposit to be made available to anyone who requests a culture? If yes there are no restrictions on distribution. Answering no will ensure the deposit is not available until the patent has issued. Yes No X

- b) As of date of deposit or conversion to meet the requirements of the Budapest Treaty, do you wish the deposit to be made available to requestors which satisfy patent offices in countries not signatory to the Budapest Treaty? Yes No X If yes state which countries.

Appendix 6

AMERICAN TYPE CULTURE COLLECTION

Additional Information Required
When Depositing A Virus For Patent Purposes

ATCC™

If you are depositing a virus, please provide the following information to facilitate viability testing upon receipt of virus and to allow propagation after the virus is released.

1. Cell line or lines in which virus can be grown. If the cell line is not listed in the ATCC Catalog of Cell Lines, please provide two T25 flasks for viability testing purposes.

Chick Embryo Fibroblasts (CEF); DMK-21

2. Cell line medium (include all additives,* serum, incubation temperature, etc.):

E-MEM supplemented with 10% fetal bovine serum, 1.5% glutamine (from 200mM stock), 1% Streptomycin (from 100mg/ml stock), 0.5% neomycin solution (from 40mg/ml stock) for growth of CEF cells; for infection, use 2% fetal bovine serum. Certified reagents including gamma-irradiated fetal bovine serum and genome provirus free lysate were used to clear and propagate deposited virus.

3. Virus preparation (include system in which deposited virus was prepared):

Infect CEF cells at an MOI of 1-3 of DNA 1974/NIH Clonal 1, incubate at 37°C in 5% CO₂ for three days. Scrape cells in supernatant. Pellet cells and remove supernatant and resuspend in approximately 1/10 of original volume. Freeze and thaw 2 cycles in dry ice, dimethyl sulfoxide and store at -80°C.

4. Titer of virus (include name of cells in which virus was obtained and time required to see effect):

1 x 10⁶ infectious particles. The deposited in CEF cells with concentration with approximately 100% infection at 24 hours.

5. Incubation and growth conditions (include percent confluency optimal for inoculation, amount and dilution of inoculum, temperature, time allowed to reach flooding, and any other pertinent data):

80-100% confluent CEF are infected at a MOI of 1-3. Infected cells are incubated at 37°C under 5% CO₂ for 3-5 days. For infection, CEF medium is E-MEM supplemented as above in #2 with 2% fetal bovine serum.

6. If CPE, type and when expected for vent dilation of inoculum:

CPE is an altered morphology and rounding up of cells visible by day 2/3.

7. If viability is to be determined by other than CPE, please provide protocol and any necessary specific reagents (e.g., antiserum):

N/A

8. Storage conditions for virus:

-80C

While the present invention has been described in some detail for purposes of clarity and understanding, one skilled in the art will appreciate that various changes in form and detail can be made without departing from the true scope of the invention. All figures, tables, and appendices, as well as patents, applications, and publications, referenced to
5 above, are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. A pharmaceutical composition comprising a recombinant MVA virus expressing an HIV *env*, *gag*, and *pol* gene or modified gene thereof for production of an HIV Env, Gag, and Pol antigen by expression from said recombinant MVA virus, wherein
5 said HIV *env* gene is modified to encode an HIV Env protein composed of gp120 and the membrane-spanning and ectodomain of gp41 but lacking part or all of the cytoplasmic domain of gp41, and a pharmaceutically acceptable carrier, wherein said HIV *env*, *gag*, and *pol* genes are isolatable from an individual infected with Ugandan clade D isolate 99UGA03349, 99UGA07412, or 98UG57128.
- 10 2. The pharmaceutical composition of claim 1 comprising 99UGA03349 gagpol in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.
3. The pharmaceutical composition of claim 1 comprising 99UGA07412 gagpol in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity
15 thereto.
4. The pharmaceutical composition of claim 1 comprising 99UGA03349 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.
5. The pharmaceutical composition of claim 1 comprising 99UGA07412
20 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.
6. The pharmaceutical composition of claim 1 comprising 98UG57128 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.
- 25 7. The pharmaceutical composition of claim 1, wherein said recombinant MVA virus is MVA/UGD-1 defined as comprising 99UGA07412 gagpol in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 99UGA07412 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.
- 30 8. The pharmaceutical composition of claim 1, wherein said recombinant MVA virus is MVA/UGD-2 defined as comprising 99UGA03349 gagpol in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 98UG57128

envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.

9. The pharmaceutical composition of claim 1, wherein said recombinant MVA virus is MVA/UGD-3 defined as comprising 99UGA07412 gagpol in Appendix 1 or
5 sequence having at least about 90%, 95% or 99.9% identity thereto, and 99UGA03349 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.

10. The pharmaceutical composition of claim 1, wherein said recombinant MVA virus is MVA/UGD-4 defined as comprising 99UGA03349 gagpol in Appendix 1 or
10 sequence having at least about 90%, 95% or 99.9% identity thereto, and 99UGA07412 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.

11. The pharmaceutical composition of claim 1, wherein said recombinant MVA virus is MVA/UGD-5 defined as comprising 99UGA03349 gagpol in Appendix 1 or
15 sequence having at least about 90%, 95% or 99.9% identity thereto, and 98UG57128 envelope in Appendix 1 or sequence having at least about 90%, 95% or 99.9% identity thereto.

12. The pharmaceutical composition of claim 1 wherein said recombinant MVA virus additionally expresses an additional HIV gene or modified gene thereof for
20 production of an HIV antigen by expression from said recombinant MVA virus, wherein said additional HIV gene is a member selected from the group consisting of *vif*, *vpr*, *tat*, *rev*, *vpu*, and *nef*.

13. An MVA shuttle plasmid comprising pLAS-1 of Appendix 2 or sequence having at least about 90%, 95% or 99.9% identity thereto, or pLAS-2 of Appendix 2 or
25 sequence having at least about 90%, 95% or 99.9% identity thereto.

14. A method of making a recombinant MVA virus comprising preparing the MVA shuttle plasmid of Claim 13 and combining said MVA shuttle plasmid with a MVA virus to produce said recombinant MVA virus, and isolating said recombinant MVA virus.

15. A method of boosting a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a
30 composition of claim 1, whereby a CD8⁺ T cell immune response to the antigen previously primed in the primate is boosted.

16. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a composition of claim 1, whereby a CD8⁺ T cell immune response to the antigen in the primate is induced.
- 5 17. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a priming composition comprising nucleic acid encoding said antigen and then provision in the primate of a boosting composition which comprises claim 1, whereby a CD8⁺ T cell immune response to the antigen is induced.
- 10 18. The method of claim 15, wherein the primate is a human.
19. The method of claim 15, wherein administration of the recombinant MVA virus is by needleless injection.
20. The method of claim 15, wherein the priming composition comprises plasmid DNA encoding said antigen.
- 15 21. MVA 1974/NIH Clone 1.
22. A pharmaceutical composition comprising a recombinant MVA virus expressing an HIV *env*, *gag*, and *pol* gene or modified gene thereof for production of an HIV Env, Gag, and Pol antigen by expression from said recombinant MVA virus, wherein said HIV *env* gene is modified to encode an HIV Env protein composed of gp120 and the
- 20 membrane-spanning and ectodomain of gp41 but lacking part or all of the cytoplasmic domain of gp41, and a pharmaceutically acceptable carrier, wherein said HIV *env*, *gag*, and *pol* genes are isolatable from an individual infected with Kenyan clade A isolate 00KE-KER2008, 00KE-KNH1144, or 00KE-KNH1207.
23. The pharmaceutical composition of claim 22 comprising 00KE-KER2008
- 25 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.
24. The pharmaceutical composition of claim 22 comprising 00KE-KNH1144 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.
- 30 25. The pharmaceutical composition of claim 22 comprising 00KE-KNH1207 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

26. The pharmaceutical composition of claim 22, wherein said recombinant MVA virus is MVA/KEA-1 defined as comprising 00KE-KER2008 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00KE-KNH1144 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

27. The pharmaceutical composition of claim 22, wherein said recombinant MVA virus is MVA/KEA-2 defined as comprising 00KE-KER2008 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00KE-KNH1207 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

28. The pharmaceutical composition of claim 22, wherein said recombinant MVA virus is MVA/KEA-3 defined as comprising 00KE-KER2008 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00KE-KNH1144 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

29. The pharmaceutical composition of claim 22, wherein said recombinant MVA virus is MVA/KEA-4 defined as comprising 00KE-KER2008 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00KE-KNH1144 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

30. The pharmaceutical composition of claim 22, wherein said recombinant MVA virus is MVA/KEA-5 defined as comprising 00KE-KER2008 gagpol in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00KE-KNH1144 envelope in Appendix 3 or sequence having at least about 90%, 95% or 99.9% identity thereto.

31. The pharmaceutical composition of claim 22 wherein said recombinant MVA virus additionally expresses an additional HIV gene or modified gene thereof for production of an HIV antigen by expression from said recombinant MVA virus, wherein said additional HIV gene is a member selected from the group consisting of *vif*, *vpr*, *tat*, *rev*, *vpu*, and *nef*.

32. A method of boosting a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a

composition of claim 22, whereby a CD8⁺ T cell immune response to the antigen previously primed in the primate is boosted.

33. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a composition of claim 22, whereby a CD8⁺ T cell immune response to the antigen in the primate is induced.

34. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a priming composition comprising nucleic acid encoding said antigen and then provision in the primate of a boosting composition which comprises claim 22, whereby a CD8⁺ T cell immune response to the antigen is induced.

35. The method of claim 32, wherein the primate is a human.

36. The method of claim 32, wherein administration of the recombinant MVA virus is by needleless injection.

37. The method of claim 32, wherein the priming composition comprises plasmid DNA encoding said antigen.

38. A pharmaceutical composition comprising a recombinant MVA virus expressing an HIV *env*, *gag*, and *pol* gene or modified gene thereof for production of an HIV Env, Gag, and Pol antigen by expression from said recombinant MVA virus, wherein said HIV *env* gene is modified to encode an HIV Env protein composed of gp120 and the membrane-spanning and ectodomain of gp41 but lacking part or all of the cytoplasmic domain of gp41, and a pharmaceutically acceptable carrier, wherein said HIV *env*, *gag*, and *pol* genes are isolatable from an individual infected with Tanzanian clade C isolate 00TZA-246 or 00TZA-125.

39. The pharmaceutical composition of claim 38 comprising 00TZA-246 gagpol in Appendix 4 or sequence having at least about 90%, 95% or 99.9% identity thereto.

40. The pharmaceutical composition of claim 38 comprising 00TZA-125 envelope in Appendix 4 or sequence having at least about 90%, 95% or 99.9% identity thereto.

41. The pharmaceutical composition of claim 38, wherein said recombinant MVA virus is MVA/TZC-1 defined as comprising 00TZA-246 gagpol in Appendix 4 or sequence having at least about 90%, 95% or 99.9% identity thereto, and 00TZA-125

envelope in Appendix 4 or sequence having at least about 90%, 95% or 99.9% identity thereto.

42. The pharmaceutical composition of claim 38 wherein said recombinant MVA virus additionally expresses an additional HIV gene or modified gene thereof for production of an HIV antigen by expression from said recombinant MVA virus, wherein said additional HIV gene is a member selected from the group consisting of *vif*, *vpr*, *tat*, *rev*, *vpu*, and *nef*.

43. A method of boosting a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a composition of claim 38, whereby a CD8⁺ T cell immune response to the antigen previously primed in the primate is boosted.

44. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a composition of claim 38, whereby a CD8⁺ T cell immune response to the antigen in the primate is induced.

45. A method of inducing a CD8⁺ T cell immune response to an HIV Env, Gag, or Pol antigen in a primate, the method comprising provision in the primate of a priming composition comprising nucleic acid encoding said antigen and then provision in the primate of a boosting composition which comprises claim 38, whereby a CD8⁺ T cell immune response to the antigen is induced.

46. The method of claim 43, wherein the primate is a human.

47. The method of claim 43, wherein administration of the recombinant MVA virus is by needleless injection.

48. The method of claim 43, wherein the priming composition comprises plasmid DNA encoding said antigen.

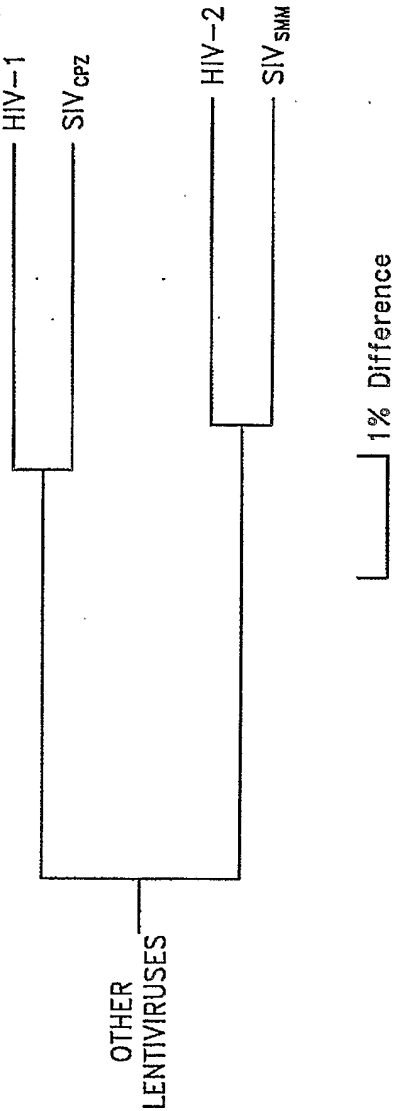


FIG. 1

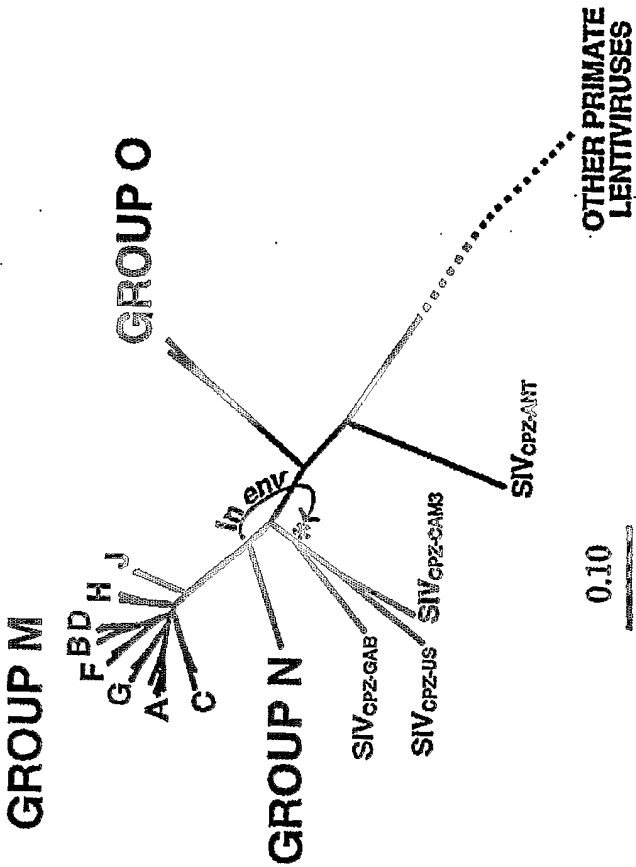


FIG. 2

Chemokine coreceptor used	PBMC replication	Macrophage replication	T-cell-line replication	REplicative phenotype	Syncytium-inducing phenotype
X4	+	-	+	Rapid/high	++
R5	+	+	-	Slow/low	-
R5/X4	+	+	+	Rapid/high	+

FIG. 3

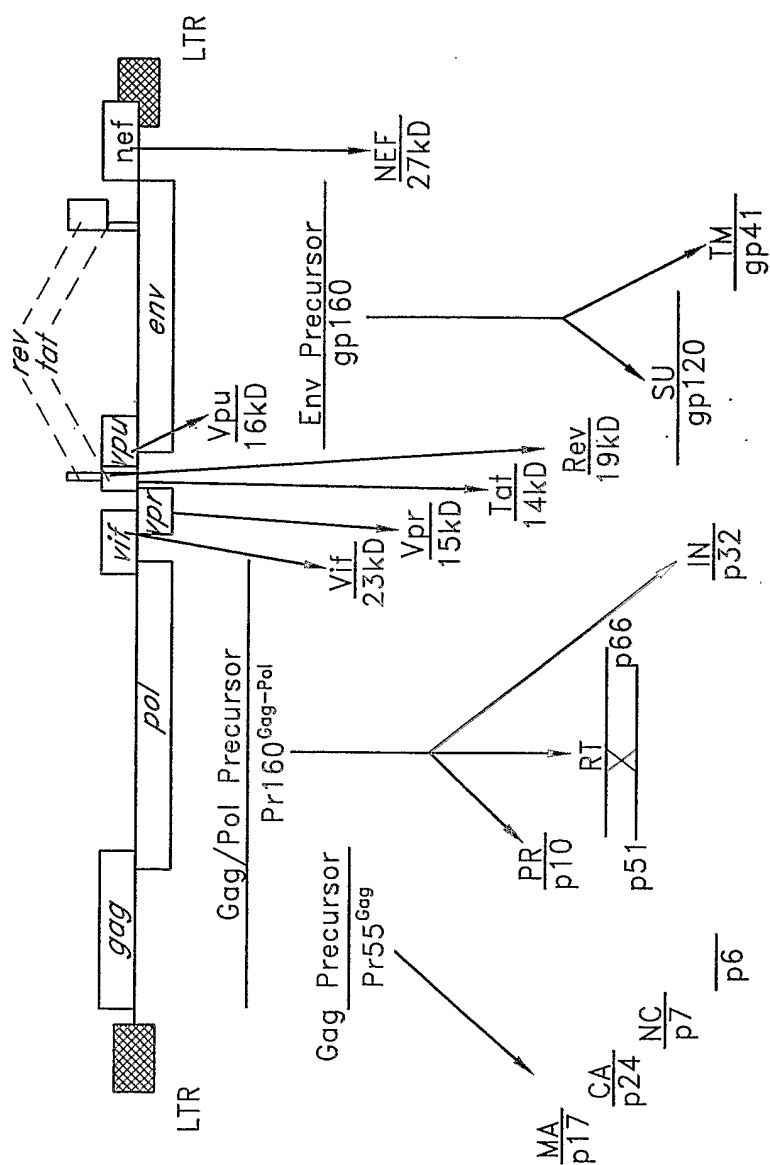


FIG. 4

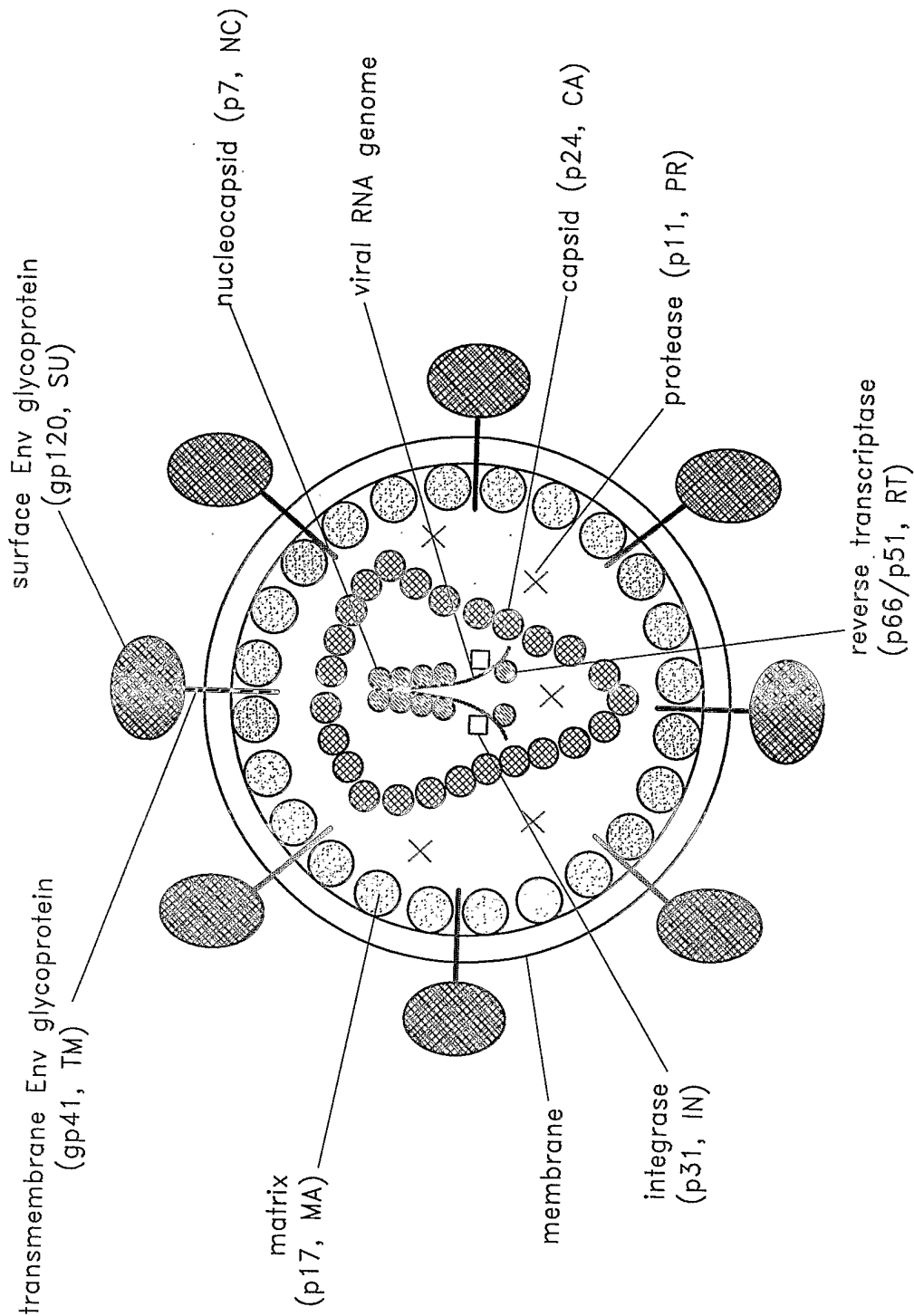


FIG. 5

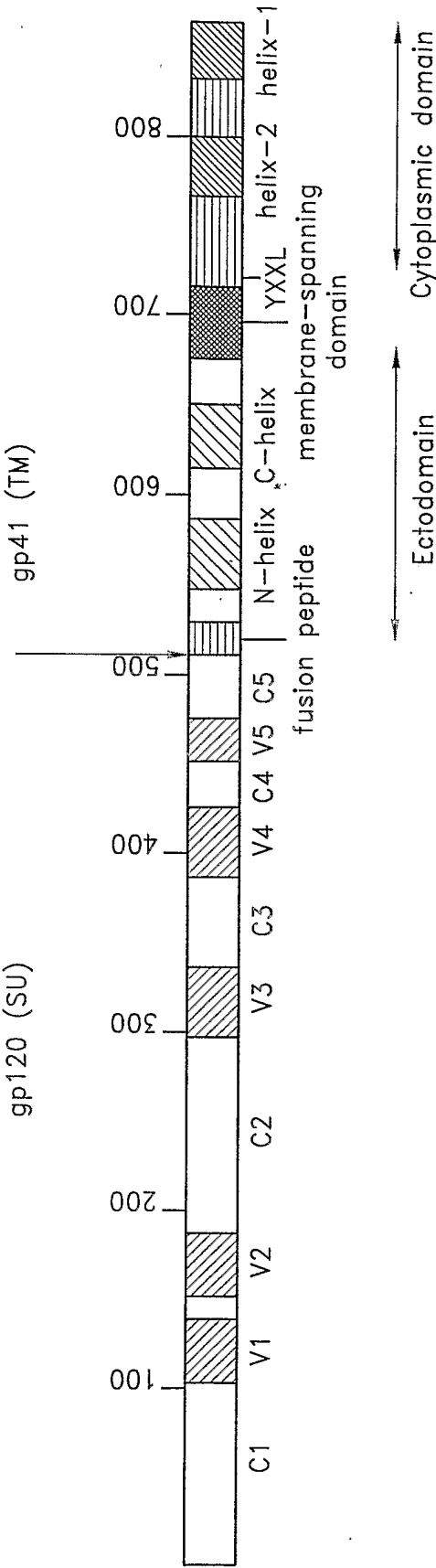


FIG. 6

7/64

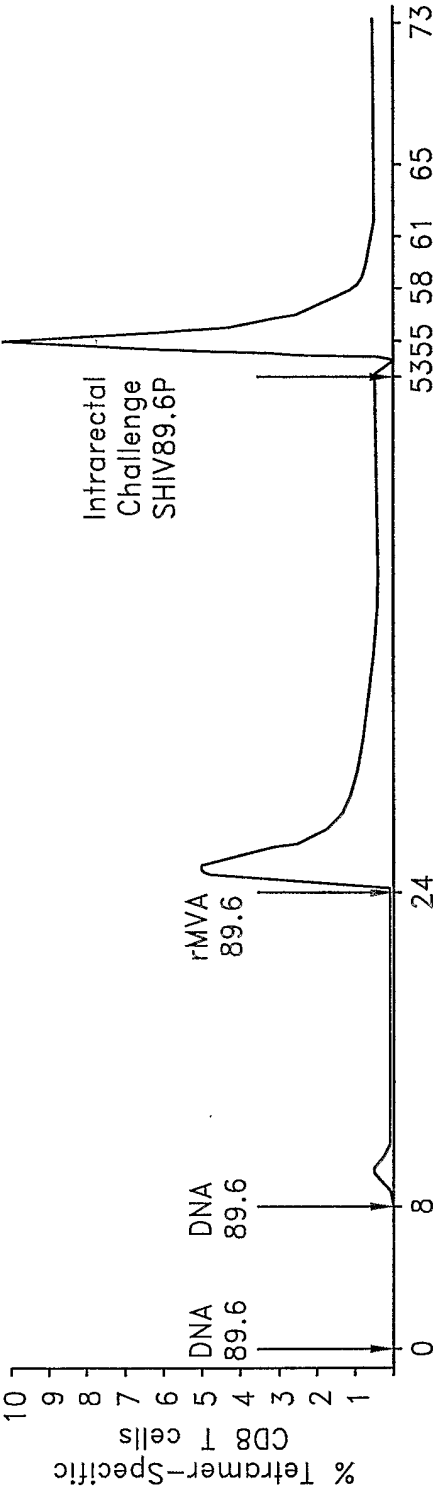


FIG. 7A

8/64

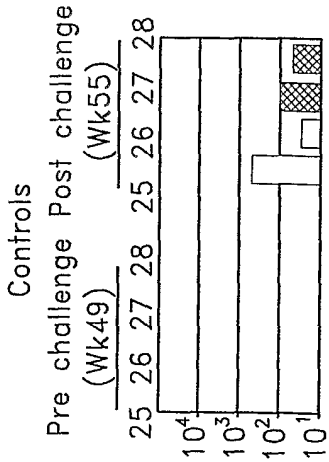
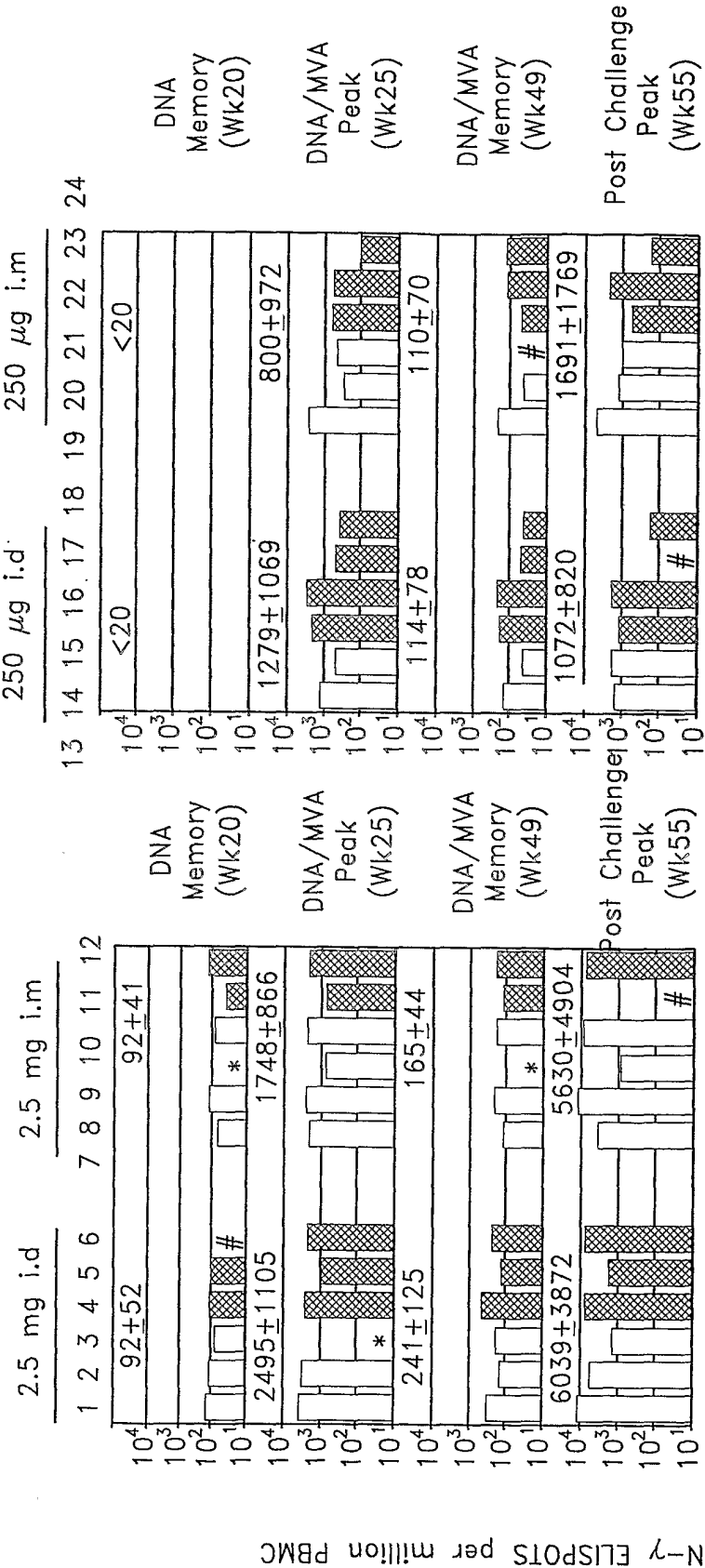


FIG. 7B

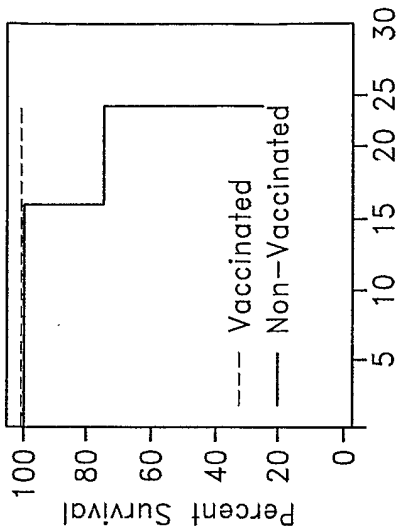


FIG. 8C

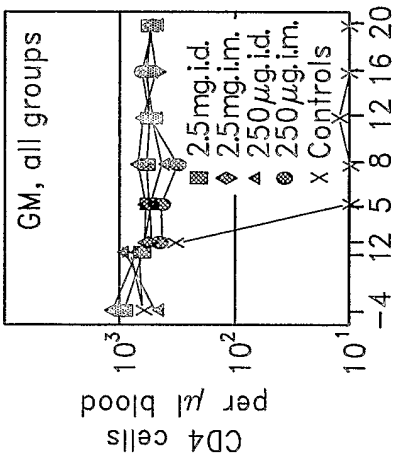


FIG. 8B

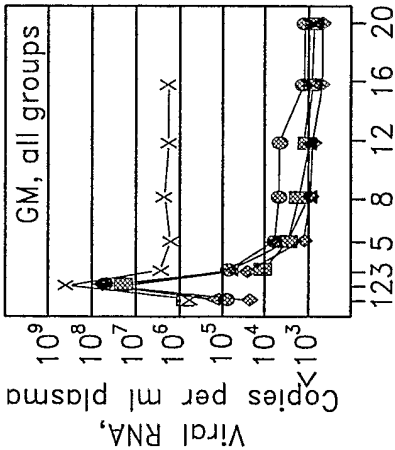


FIG. 8A

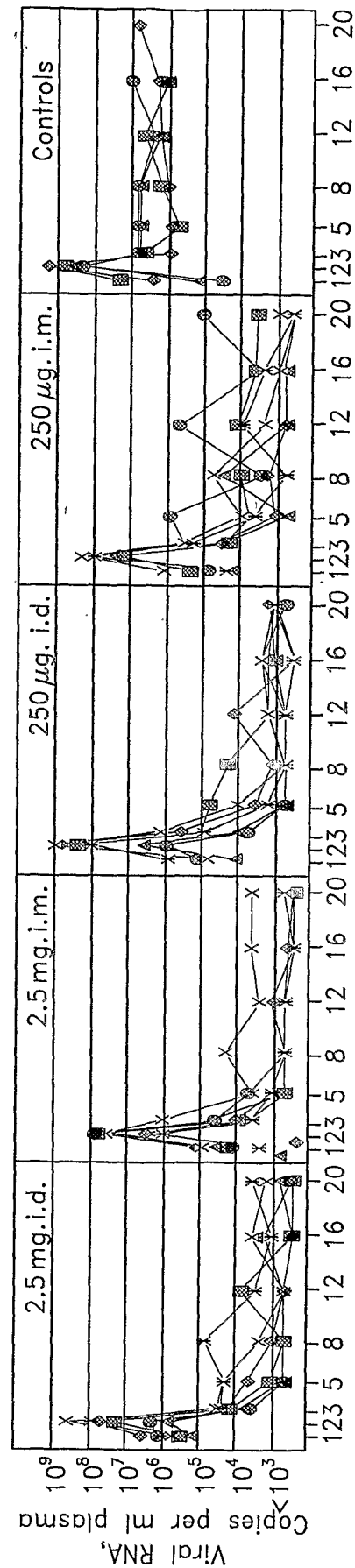


FIG. 8D

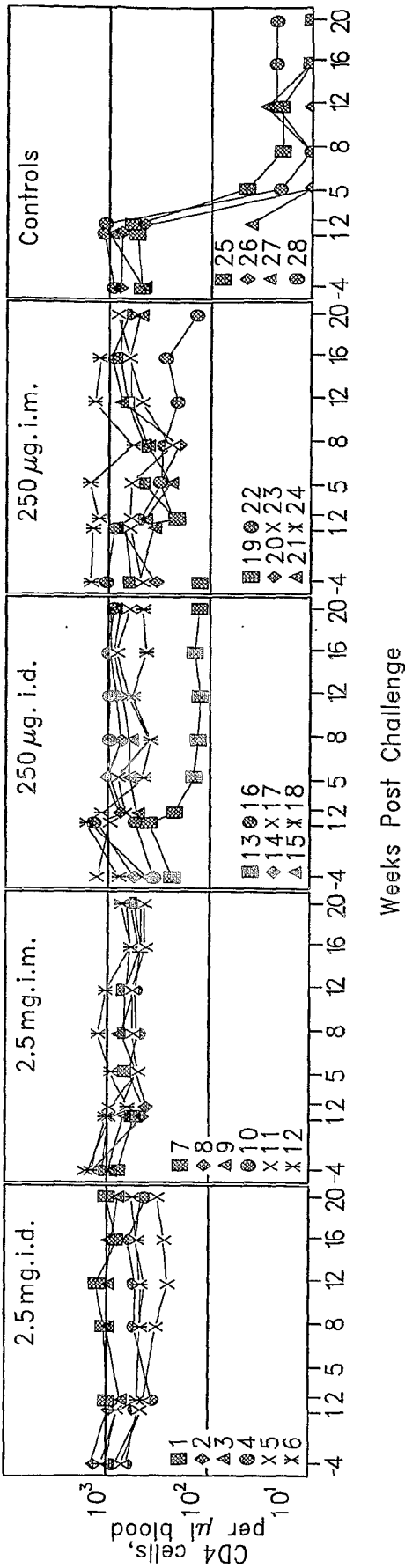


FIG. 8E

12/64

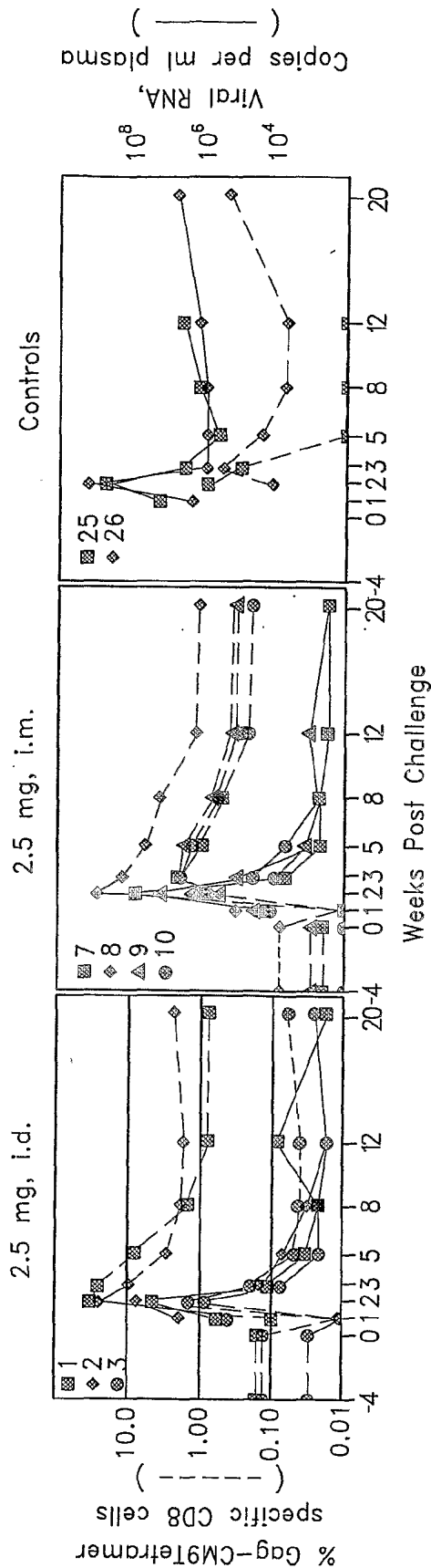


FIG. 9A

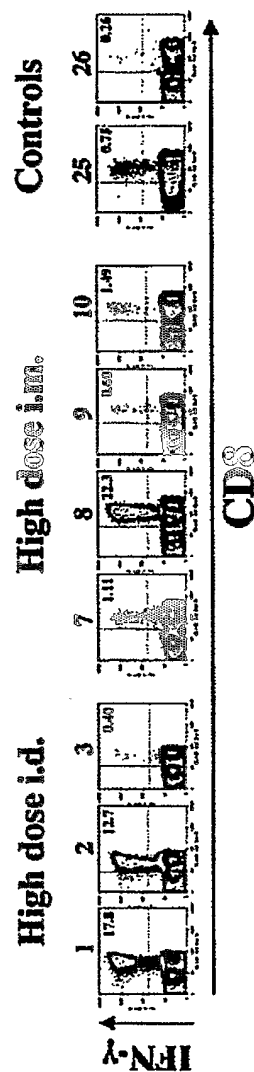


FIG. 9B

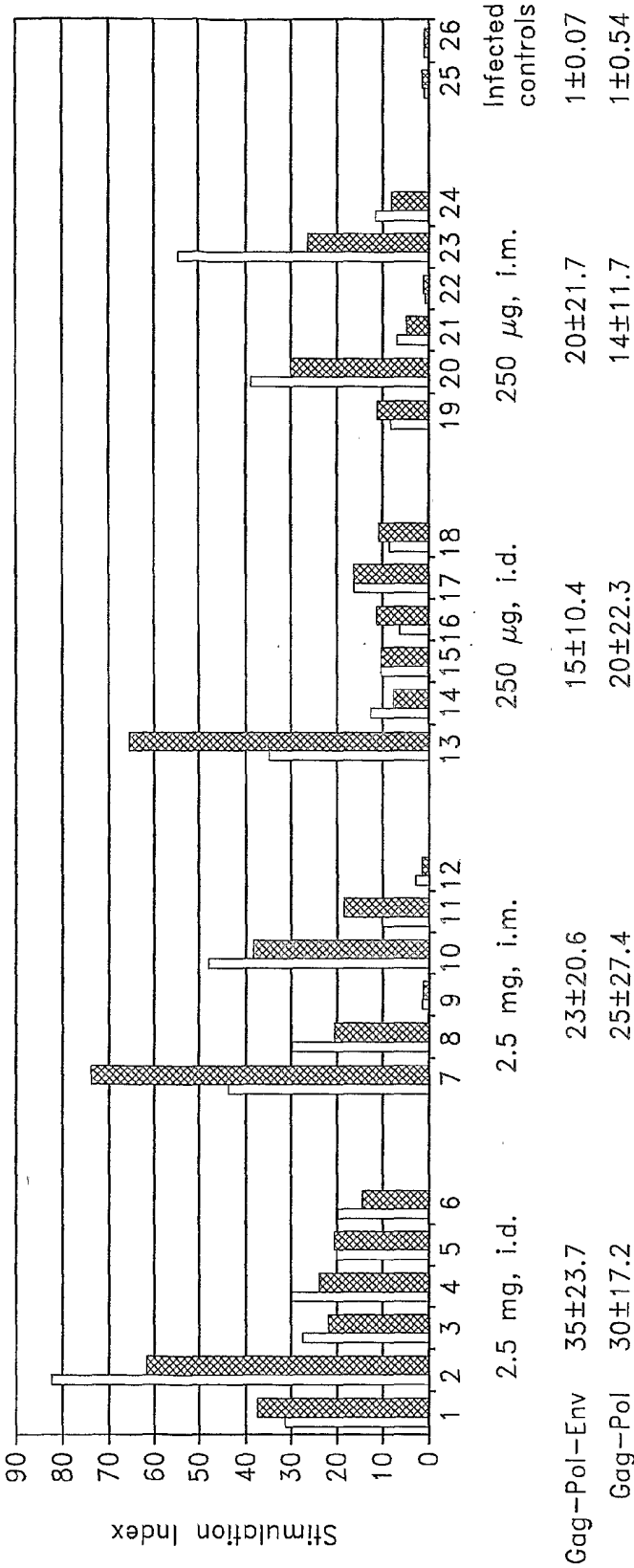


FIG. 9C

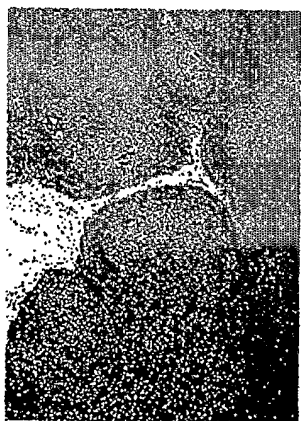
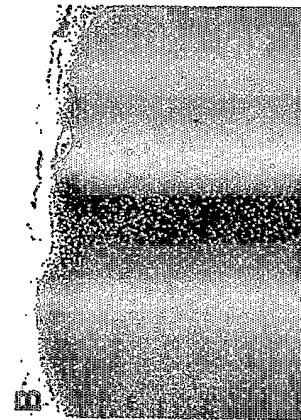
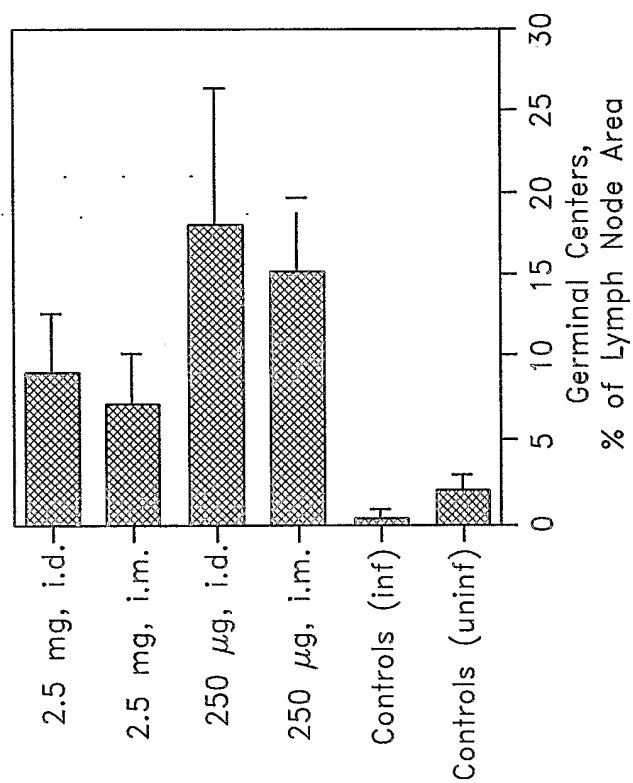


FIG. 10C

FIG. 10B

FIG. 10A

16/64

*FIG. 10D*

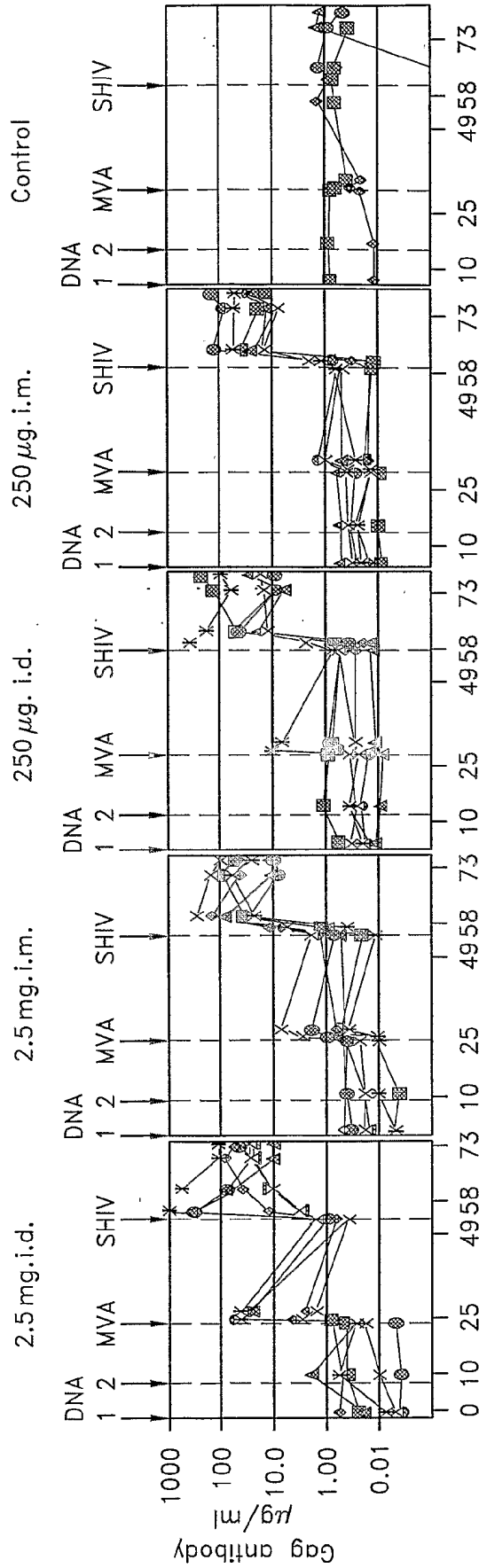


FIG. 11A

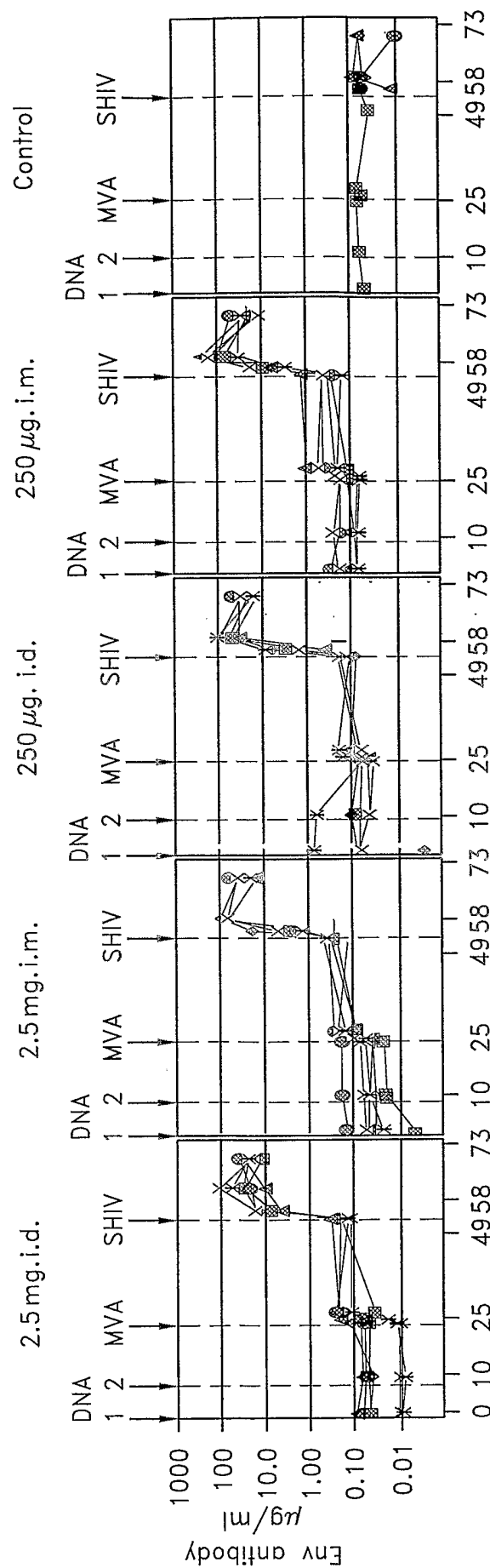


FIG. 11B

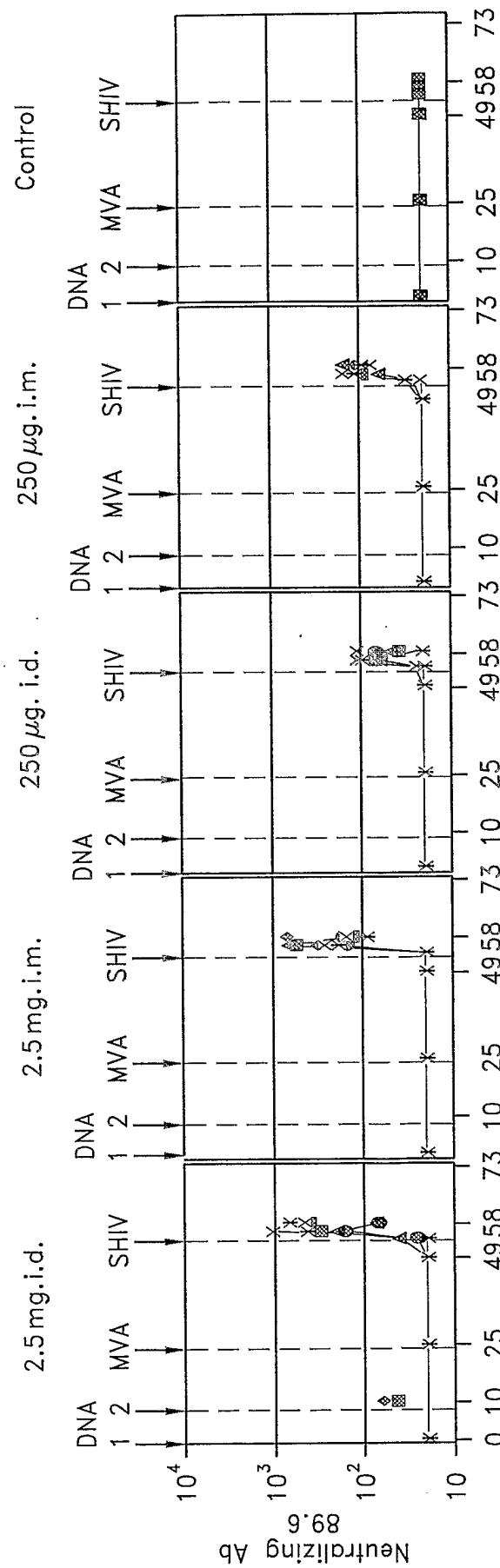


FIG. 11C

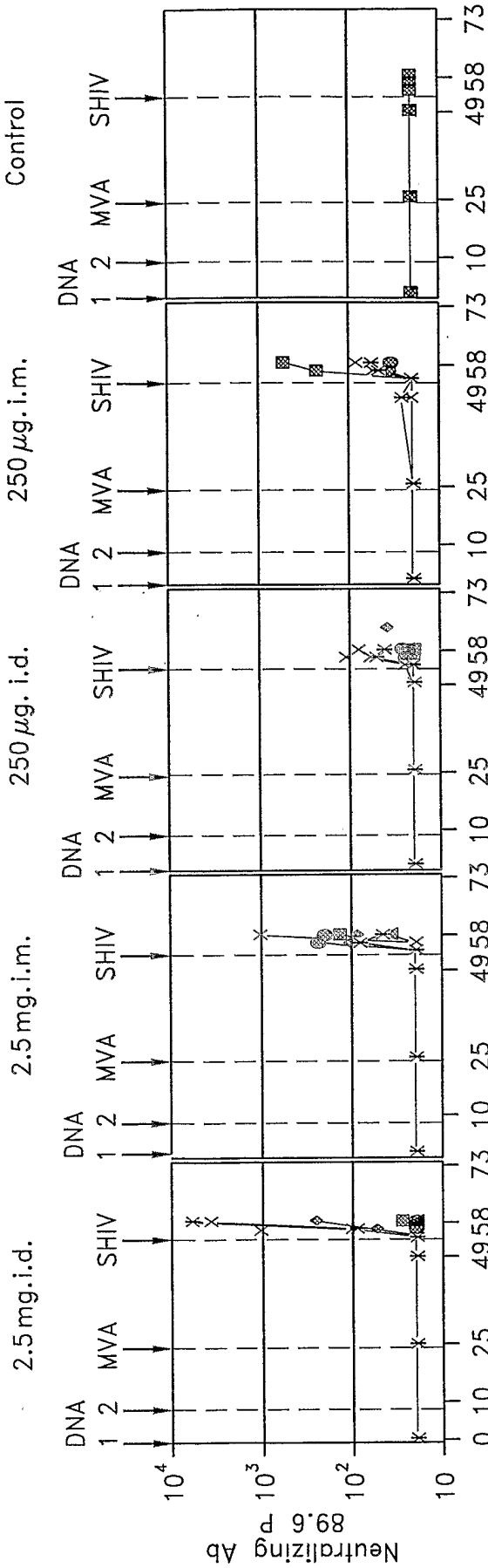


FIG. 11D

21/64

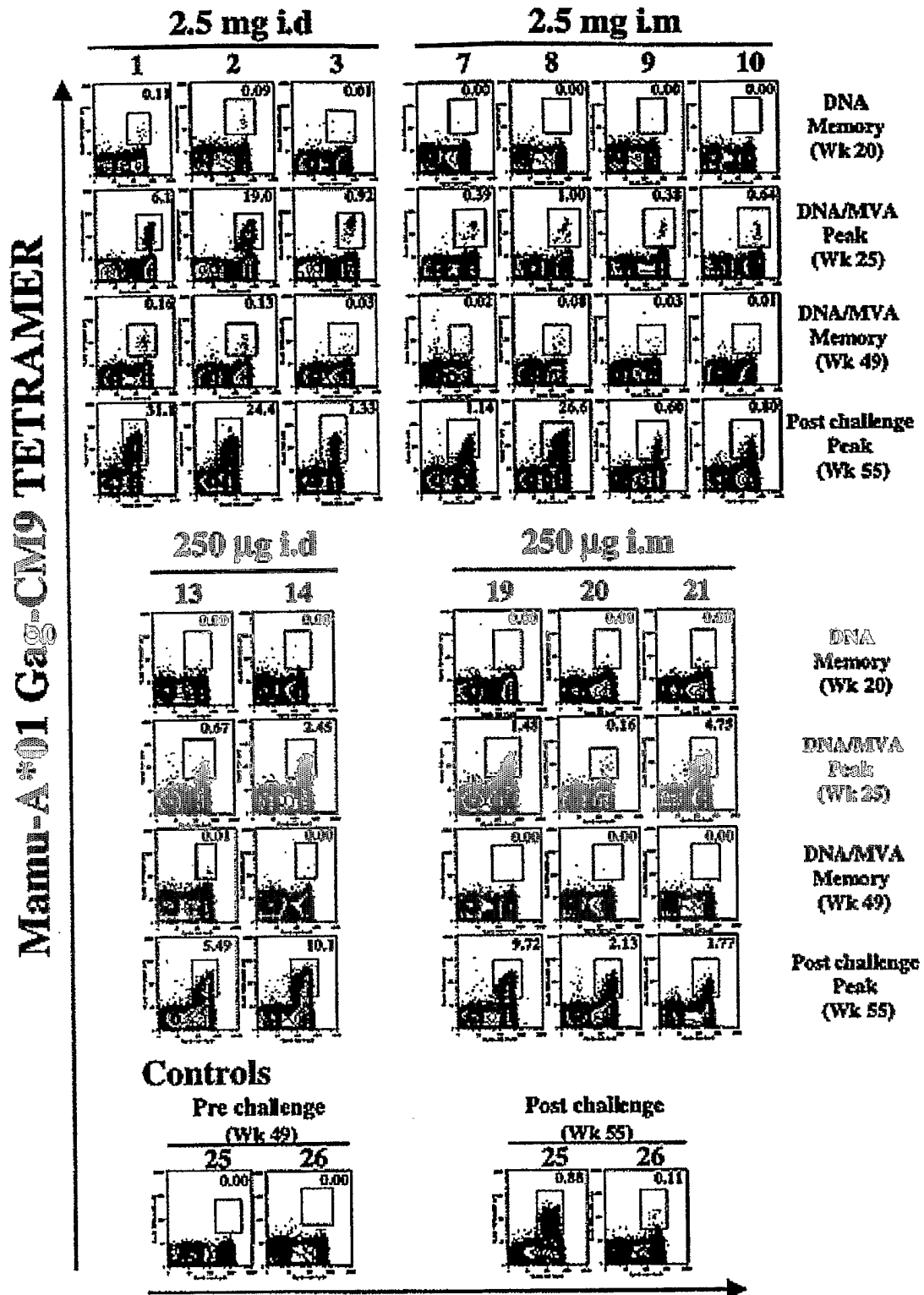
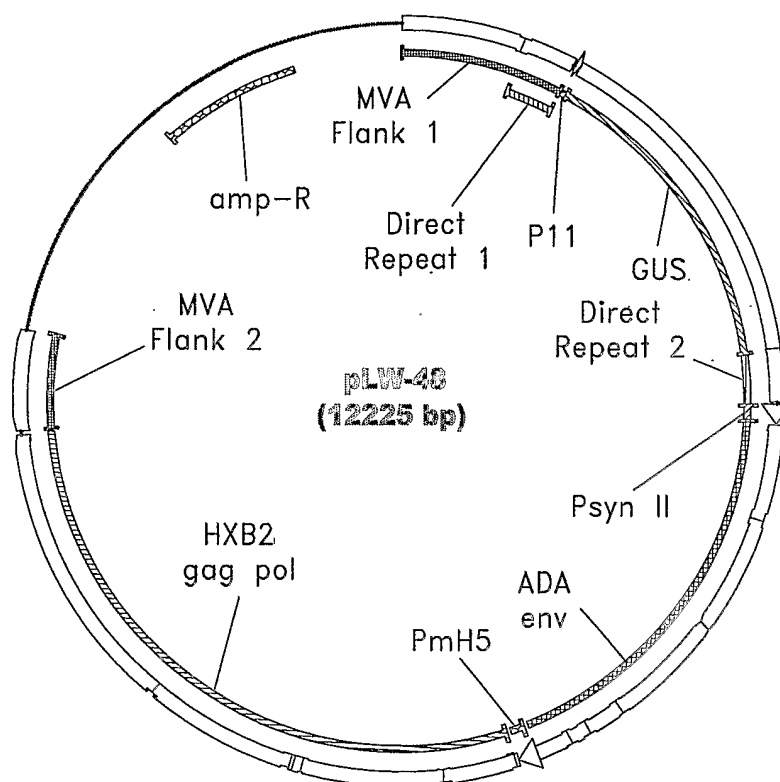


FIG. 12

22/64

*FIG. 13*

23/64

Text File of pLW-48 and the Included Individual HIV Genes and Their Promoters

Entire pLW-48 plasmid sequence:

GAATTCGTTGGTGGTCGCCATGGATGGTGTATTGTATACTGTCTAAACGCG
TTAGTAAAACATGGCGAGGAAATAAATCATATAAAAAATGATTTTCATGATTAA
ACCATGTTGTGAAAAAGTCAAGAACGTTACATTGGCGGACAATCTAAAAAC
AATACAGTGATTGCAGATTTGCCATATATGGATAATGCCGTATCCGATGTAT
GCAATTCAGTGTATAAAAAGAATGTATCAAGAATATCCAGATTTGCTAATTTG
ATAAAGATAGATGACGATGACAAGACTCCTACTGGTGTATATAATTATTTTAA
ACCTAAAGATGCCATTCTGTTATTATATCCATAGGAAAGGATAGAGATGTTT
GTGAAGTATTAATCTCATCTGATAAAGCGTGTGCGTGTATAGAGTTAAATTCA
TATAAAGTAGCCATTCTTCCCATGGATGTTTCTTTTACCAAAGGAAATGC
ATCATTGATTATTCTCCTGTTTGATTCTCTATCGATGCGGCACCTCTCTTAA
GAAGTGTAACCGATAATAATGTTATTATATCTAGACACCAGCGTCTACATGA
CGAGCTTCCGAGTTCCAATTGGTTCAAGTTTTACATAAGTATAAAGTCCGAC
TATTGTTCTATATTATATATGGTTGTTGATGGATCTGTGATGCAATAGC
TGATAATAGAACTTACGCCAAATATTAGCAAAAATATATTAGACAATACTACAA
TTAACGATGAGTGTAGATGCTGTTATTTTGAACCACAGATTAGGATTCTTGAT
AGAGATGAGATGCTCAATGGATCATCGTGTGATATGAACAGACATTGTATTA
TGATGAATTTACCTGATGTAGCGCAATTTGGATCTAGTATGTTGGGGAAATA
TGAACCTGACATGATTAAGATTGCTCTTTCGGTGGCTGGGTACCAGGCGCG
CCTTTCATTTTGTTTTTTCTATGCTATAAATGGTACGTCCTGTAGAAACCCC
AACCCGTGAAATCAAAAACTCGACGGCCTGTGGGCATTCAAGTCTGGATCG
CGAAAACTGTGGAATTGATCAGCGTTGGTGGGAAAGCGCGTTACAAGAAAG
CCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCGCCGATGCAGA
TATTCGTAATTATGCGGGCAACGTCTGGTATCAGCGCGAAGTCTTTATACCG
AAAGGTTGGGCAGGCCAGCGTATCGTGCTGCGTTTCGATGCGGTCACTCAT
TACGGCAAAGTGTGGGTCAATAATCAGGAAGTGATGGAGCATCAGGGCGG
CTATACGCCATTTGAAGCCGATGTCACGCCGTATGTTATTGCCGGGAAAAG
TGTACGTATCACCGTTTGTGTGAACAACGAACTGAACTGGCAGACTATCCC
GCCGGGAATGGTGATTACCGACGAAAACGGCAAGAAAAAGCAGTCTTACTT
CCATGATTTCTTTAACTATGCCGGAATCCATCGCAGCGTAATGCTCTACACC
ACGCCGAACACCTGGGTGGACGATATCACCGTGGTGACGCATGTCCGCGA
AGACTGTAACCACGCGTCTGTTGACTGGCAGGTGGTGGCCAATGGTGATGT
CAGCGTTGAACTGCGTGATGCGGATCAACAGGTGGTTGCAACTGGACAAG
GCACTAGCGGGACTTTGCAAGTGGTGAATCCGCACCTCTGGCAACCGGGT
GAAGGTTATCTCTATGAACTGTGCGTCACAGCCAAAAGCCAGACAGAGTGT
GATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGA
ACAGTTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCAT
GAAGATGCGGACTTGCGTGGCAAAGGATTGATAACGTGCTGATGGTGAC
GACCACGCATTAATGGACTGGATTGGGGCCAACTCCTACCGTACCTCGCAT
TACCCTTACGCTGAAGAGATGCTCGACTGGGCAGATGAACATGGCATCGTG

FIG. 14A

24/64

GTGATTGATGAACTGCTGCTGTCGGCTTTAACCTCTCTTTAGGCATTGGTT
TCGAAGCGGGCAACAAGCCGAAAGAACTGTACAGCGAAGAGGCCAGTCAAC
GGGGAAACTCAGCAAGCGCACTTACAGGCGATTAAAGAGCTGATAGCGCGT
GACAAAAACCACCCAAGCGTGCTGATGTGGAGTATTGCCAACGAACCGGAT
ACCCGTCGCAAGGTGCACGGGAATATTTGCGGCCACTGGCGGAAGCAAC
GCGTAAACTCGACCCGACGCGTCCGATCACCTGCGTCAATGTAATGTTCTG
CGACGCTCACACCGATACCATCAGCGATCTCTTTGATGTGCTGTGCCTGAA
CCGTTATTACGGATGGTATGTCCAAAGCGGCGATTGGAACGGCAGAGAA
GGTACTGGAAAAAGAACTTCTGGCCTGGCAGGAGAACTGCATCAGCCGAT
TATCATCACCGAATACGGCGTGGATACGTTAGCCGGGCTGCACTCAATGTA
CACCGACATGTGGAGTGAAGAGTATCAGTGTGCATGGCTGGATATGTATCA
CCGCGTCTTTGATCGCGTCAGCGCCGTCGTCGGTGAACAGGTATGGAATTT
CGCCGATTTTGCAGCCTCGCAAGGCATATTGCGCGTTGGCGGTAACAAGAA
AGGGATCTTCACTCGCGACCGCAAACCGAAGTCGGCGGCTTTTCTGCTGCA
AAAACGCTGGACTGGCATGAACTTCGGTGAAAAACCGCAGCAGGGAGGCA
AACAATGAGAGCTCGGTTGTTGATGGATCTGTGATGCATGCAATAGCTGATA
ATAGAACTTACGCAATATTAGCAAAAATATATTAGACAATACTACAATTAAC
GATGAGTGTAGATGCTGTTATTTTGAACCACAGATTAGGATTCTTGATAGAG
ATGAGATGCTCAATGGATCATCGTGTGATATGAACAGACATTGTATTATGAT
GAATTTACCTGATGTAGGCGAATTTGGATCTAGTATGTTGGGGAAATATGAA
CCTGACATGATTAAGATTGCTCTTTGCGTGGCTGGCGGCCCGCTCGAGTAA
AAAATGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTCTAT
GCTATAAATAATAAATAGCGGCCCGCACCATGAAAGTGAAGGGGATCAGGAA
GAATTATCAGCACCTTGTGGAAATGGGGCATCATGCTCCTTGGGATGTTGATG
ATCTGTAGTGCTGTAGAAAATTTGTGGGTCACAGTTTATTATGGGGTACCTG
TGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATA
TGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGA
CCCCAACCCACAAGAAGTAGTATTGGAAAATGTGACAGAAAAATTTTAAATG
TGGAAAAATAACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGG
ATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTACTTTAAAT
TGCACTGATTTGAGGAATGTTACTAATATCAATAATAGTAGTGAGGGAATGA
GAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGCATAAGAGATAA
GGTGAAGAAAGACTATGCACTTTTGTATAGACTTGATGTAGTACCAATAGATA
ATGATAATACTAGCTATAGGTTGATAAATTGTAATACCTCAACCATTACACAG
GCCTGTCCAAAGGTATCCTTTGAGCCAATCCCATAACATTATTGTACCCCGG
CTGGTTTTGCGATTCTAAAGTGTAAGACAAGAAGTTCAATGGAACAGGGCC
ATGTAAAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTG
TCAACTCAACTGCTGTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTA
GATCTAGTAATTTACAGACAATGCAAAAAACATAATAGTACAGTTGAAAGAA
TCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATAC
ATATAGGACCAGGAAGAGCATTTTATACAACAGGAGAAATAATAGGAGATAT
AAGACAAGCACATTGCAACATTAGTAGAACAAAATGGAATAACACTTTAAAT
CAAATAGCTACAAAATTAAGAACAATTTGGGAATAATAAAACAATAGTCTT
TAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTAAATTGT
GGAGGGGAATTCTTCTACTGTAATTCAACACAACCTGTTTAATAGTACTTGA
ATTTAATGGTACTTGGAATTTAACACAATCGAATGGTACTGAAGGAAATGA

FIG. 14B

25/64

CACTATCACACTCCCATGTAGAATAAAACAAATTATAAATATGTGGCAGGAA
GTAGGAAAAGCAATGTATGCCCCCTCCCATCAGAGGACAAATTAGATGCTCAT
CAAATATTACAGGGCTAATATTAACAAGAGATGGTGGAACTAACAGTAGTGG
GTCCGAGATCTTCAGACCTGGGGGAGGAGATATGAGGGACAATTGGAGAA
GTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACC
ACCAAGGCCAAAAAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAAC
GATAGGAGCTATGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGG
CGCAGCGTCAATAACGCTGACGGTACAGGCCAGACTATTATTGTCTGGTAT
AGTGCAACAGCAGAACAAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCT
GTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGG
CTGTGGAAAGATACCTAAGGGATCAACAGCTCCTAGGGATTTGGGGTTGCT
CTGGAAAACCTCATCTGCACCACTGCTGTGCCTTGAATGCTAGTTGGAGTA
ATAAACTCTGGATATGATTTGGGATAACATGACCTGGATGGAGTGGGAAA
GAGAAATCGAAAATTACACAGGCTTAATATACACCTTAATTGAGGAATCGCA
GAACCAACAAGAAAAGAATGAACAAGACTTATTAGCATTAGATAAGTGGGCA
AGTTTGTGGAATTGGTTTGACATATCAAATTGGCTGTGGTATGTAAAAATCTT
CATAATGATAGTAGGAGGCTTGATAGGTTAAGAATAGTTTTACTGTACTTT
CTATAGTAAATAGAGTTAGGCAGGGATACTCACCATTGTCATTTAGACCCA
CCTCCAGCCCCGAGGGGACCCGACAGGCCCGAAGGAATCGAAGAAGAAG
GTGGAGACAGAGACTAATTTTTATGCGGCCGCTGGTACCCAACCTAAAAATT
GAAAATAAATACAAAGGTTCTTGAGGGTTGTGTTAAATTGAAAGCGAGAAAT
AATCATAAATAAGCCCGGGGATCCTCTAGAGTCGACACCATGGGTGCGAGA
GCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTA
AGGCCAGGGGGGAAAGAAAAAATATAAATTAAAAACATATAGTATGGGCAAGCA
GGGAGCTAGAACGATTTCGAGTTAATCCTGGCCTGTTAGAAACATCAGAAG
GCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAG
AAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAA
AGBATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAG
CAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGACAC
AGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAA
ATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAGTAG
TAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTTCAGCATTATC
AGAAGGAGCCACCCCAAGATTTAAACACCATGCTAAACACAGTGGGGGG
ACATCAAGCAGCCATGCAAATGTTAAAGAGACCATCAATGAGGAAGCTGC
AGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCA
GATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACCTAGTACCCTTCA
GGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATT
TATAAAAGATGGATAATCCTGGGATTAATAAAATAGTAAGAATGTATAGCCC
TACCAGCATTCTGGACATAAGACAAGGACCAAAAGAACCCTTTAGAGACTAT
GTAGACCGGTTCTATAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTA
AAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTA
AGACTATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGAC
AGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTG
AAGCAATGAGCCAAGTAACAAATTCAGCTACCATAATGATGCAGAGAGGCA
ATTTTAGGAACCAAGAAAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGG
GCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAAT

FIG. 14C

26/64

GTGGAAAGGAAGGACACCAAATGAAAGATTGTA CTGAGAGACAGGCTAATT
TTTAGGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCA
GAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGG
TAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTGT
ATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAATA
AAGATAGGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGAT
ACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAG
GGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGA
AATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTC
AACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCC
CATTAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGC
CCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAAGCATTAGTAG
AAATTTGTACAGAAATGGAAAAGGAAGGGAAAAATTTCAAAAATTGGGCCTGA
GAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAAT
GGAGGAAATTAGTAGATTTTCAGAGA ACTTAATAAGAGAACTCAAGACTTCTG
GGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCA
GTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAG
ACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACC
AGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACC
AGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAAAAAACAAA
ATCCAGACATAGTTATCTATCAATACATGAACGATTTGTATGTAGGATCTGAC
TTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACATCTG
TTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCA
TTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTA
TAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAG
TGGGGAAATTGAATACCGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGC
AATTATGTAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACT
AACAGAAGAAGCAGAGCTAGA ACTTGGCAGAAAAACAGAGAGATTCTAAAAAG
ACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATA
CAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTT
AAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCCACACTAAT
GATGTAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAG
TAATATGGGGAAAGACTCCTAAATTTAACTACCCATACAAAAGGAAACATG
GGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGGA
GTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTACCAGTTAGAGAAAGAA
CCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAG
ACTAAATTAGGAAAAGCAGGATATGTTACTAACAAAGGAAGACAAAAGGTTG
TCCCCCTAACTAACACAACAATCAGAAAACCTCAGTTACAAGCAATTTATCTA
GCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACAATATG
CATTAGGAATCATTCAAGCACACCAGATAAAAGTGAATCAGAGTTAGTCAA
TCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTA
CCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGT
GCTGGAATCAGGAAAATACTATTTTATAGATGGAATAGATAAGGCCCAAGATG
AACATTAGTTTTTATGTGACCTGCAGGGAAAGTTTTATAGGTAGTTGATAG
AACAAAATACATAATTTTGTAAAAATAAATCACTTTTTATACTAATATGACACG
ATTACCAATACTTTTGTTACTAATATCATTAGTATACGCTACACCTTTTCCTCA

FIG. 14D

27/64

GACATCTAAAAAATAGGTGATGATGCAACTTTATCATGTAATCGAAATAATA
CAAATGACTACGTTGTTATGAGTGCTTGGTATAAGGAGCCCAATTCCATTAT
TCTTTTAGCTGCTAAAAGCGACGTCTTGTATTTTGATAATTATACCAAGGATA
AAATATCTTACGACTCTCCATACGATGATCTAGTTACAACATCACAATTAAA
TCATTGACTGCTAGAGATGCCGGTACTTATGTATGTGCATTCTTTATGACATC
GCCTACAAATGACACTGATAAAGTAGATTATGAAGAATACTCCACAGAGTTG
ATTGTAAATACAGATAGTGAATCGACTATAGACATAATACTATCTGGATCTAC
ACATTCACCAGAACTAGTTAAGCTTGTCTCCCTATAGTGAGTCGTATTAGA
GCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCT
CACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG
TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT
TTCGAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGCTTCCCTCGCTCAC
TGACTCGCTGCGCTCGGTCTGCTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG
TTGCTGGCGTTTTTCGATAGGCTCCGCCCCCTGACGAGCATCACAAAAAT
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG
CTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTACGCCCGACCGCTGCGCCTTATC
CGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG
CTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTG
TTCATCCATAGTTGCCTGACTCCCCGTCTGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA
CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGC
CGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTT
GGCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCA
TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT
GCAAAAAAGCGGTTAGCTCCTTCGGTCTCCTCCGATCGTTGTGAGAAGTAAGT
TGGCCGCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCA
ATACGGGATAATACCGCGCCACATAGCAGAAGTTTAAAGTGCTCATCATTG
GAAACGTTCTTCGGGGCGAAAAGTCTCAAGGATCTTACCGCTGTTGAGAT
CCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTAC

FIG. 14E

28/64

TTTCACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGCAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT
CAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGA
AAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
AATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGG
TGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTT
GGCGGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACT
GAGAGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA
ATACCGCATCAGGCGCCATTTCGCCATTAGGCTGCGCAACTGTTGGGAAGG
GCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT
GTGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGAC
GTTGTAAACGACGGCCAGTGAATTGGATTTAGGTGACACTATA

New Psyn II Promoter which controls ADA envelope expression:

TAAAAAATGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTTC
TATGCTATAAATAATAAATA

ADA envelope truncated:

ATGAAAGTGAAGGGGATCAGGAAGAATTATCAGCACTTGTGGAAATGGGGC
ATCATGCTCCTTGGGATGTTGATGATCTGTAGTGCTGTAGAAAATTTGTGGG
TCACAGTTTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTATT
TTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCC
ACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTEGAA
AATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAGATGC
ATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAATT
AACCCCACTCTGTGTTACTTTAAATTGCACTGATTTGAGGAATGTTACTAATA
TCAATAATAGTAGTGAGGGAATGAGAGGAGAAATAAAAACTGCTCTTTCAA
TATCACCACAAGCATAAGAGATAAGGTGAAGAAAGACTATGCACTTTTCTAT
AGACTTGATGTAGTACCAATAGATAATGATAATACTAGCTATAGGTTGATAAA
TTGTAATACCTCAACCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCA
ATTCCCATACATTATTGTACCCCGGCTGGTTTTGCGATTCTAAAGTGTAAG
ACAAGAAGTTCAATGGAACAGGGCCATGTAAAAATGTCAGCACAGTACAAT
GTACACATGGAATTAGGCCAGTAGTGTCAACTCAACTGCTGTAAATGGCAG
TCTAGCAGAAGAAGAGGTAGTAATTAGATCTAGTAATTTACAGACAATGCA
AAAAACATAATAGTACAGTTGAAAGAATCTGTAGAAATTAATTGTACAAGACC
CAACAACAATACAAGGAAAAGTATACATATAGGACCAGGAAGAGCATTTTAT
ACAACAGGAGAAATAATAGGAGATATAAGACAAGCACATTGCAACATTAGTA
GAACAAAATGGAATAACACTTTAAATCAAATAGCTACAAAATTTAAAGAACAA
TTTGGGAATAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGGACCCAG
AAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTCTTCTACTGTAATTCA
ACACAACCTGTTAATAGTACTTGGAATTTAATGGTACTTGGAATTAAACACA

FIG. 14F

ATCGAATGGTACTGAAGGAAATGACACTATCACACTCCCATGTAGAATAAAA
CAAATTATAAATATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCA
TCAGAGGACAAATTAGATGCTCATCAAATATTACAGGGCTAATATTAACAAG
AGATGGTGGAAC TAACAGTAGTGGGTCCGAGATCTTCAGACCTGGGGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAA
AATTGAACCATTAGGAGTAGCACCCACCAAGGCCAAAAGAAGAGTGGTGCA
GAGAGAAAAAAGAGCAGTGGGAACGATAGGAGCTATGTTCCCTTGGGTTCTT
GGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATAACGCTGACGGTAC
AGGCCAGACTATTATTGTCTGGTATAGTGCAACAGCAGAACAATTTGCTGAG
GGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAA
GCAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGATACCTAAGGGATCAACA
GCTCCTAGGGATTTGGGGTTGCTCTGGAAACTCATCTGCACCACTGCTGT
GCCTTGGAATGCTAGTTGGAGTAATAAACTCTGGATATGATTTGGGATAAC
ATGACCTGGATGGAGTGGGAAAGAGAAATCGAAAATTACACAGGCTTAATAT
ACACCTTAATTGAGGAATCGCAGAACCAACAAGAAAAGAATGAACAAGACTT
ATTAGCATTAGATAAGTGGGCAAGTTTGTGGAATTGGTTTGACATATCAAATT
GGCTGTGGTATGTAAAAATCTTCATAATGATAGTAGGAGGCTTGATAGGTTT
AAGAATAGTTTTTACTGTACTTTCTATAGTAAATAGAGTTAGGCAGGGATACT
CACCATTGTCATTTGAGACCCACCTCCCAGCCCCGAGGGGACCCGACAGG
CCCGAAGGAATCGAAGAAGAAGGTGGAGACAGAGAC

PmH5 promoter (which controls HXB2 gag pol expression):

AAAAATTGAAAATAAATACAAAGGTTCTTGAGGGTTGTGTTAAATTGAAAGC
GAGAAATAATCATAAATA

HXB2 gag pol (with safety mutations, Δ integrase):

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGGAGAATTAGATCGATGGGA
AAAAATTCGGTTAAGGCCAGGGGGGAAAGAAAAAATATAAATTAAACATATA
GTATGGGCAAGCAGGGAGCTAGAACGATTGCGAGTTAATCCTGGCCTGTTA
GAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTT
CAGACAGGATCAGAAGAAGCTTAGATCATTATATAATACAGTAGCAACCCTCT
ATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAA
GATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGC
TGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAA
CATCCAGGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCA
TGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATG
TTTTCAGCATTATCAGAAGGAGCCACCCACAAAGATTTAAACACCATGCTAA
ACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCA
ATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCT
ATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAAC
TACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATC
CCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAATAG
TAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGA
ACCTTTAGAGACTATGTAGACCGGTTCTATAAACTCTAAGAGCCGAGCAA

FIG. 14G

30/64

GCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATG
CGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACT
AGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGG
CAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACCATAAT
GATGCAGAGAGGCAATTTTAGGAACCAAAGAAAGATTGTTAAGTGTTTCAAT
TGTGGCAAAGAAGGGGCACACAGCCAGAAATTGCAGGGCCCCCTAGGAAAAA
GGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTAAGTGA
GAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCC
AGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAG
CTTCAGGTCTGGGGTAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGAT
AGACAAGGAACCTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGAC
CCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATA
CAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGA
AACCAAAAATGATAGGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGA
TCAGATACTCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTA
GGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT
GCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAAATTAAG
CCAGGAATGGATGGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAA
ATAAAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTT
CAAAAATTGGGCCTGAGAATCCATACAATACTCCAGTATTTGCCATAAAGAA
AAAAGACAGTACTAAATGGAGGAAATTAGTAGATTTTCAGAGAAGCTTAATAAG
AGAAGTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGG
TTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTC
AGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGT
ATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGG
GATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGA
GCCTTTTAAAAAACAAAATCCAGACATAGTTATCTATCAATACATGAACGATT
TGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGA
GCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACA
TCAGAAAGAACCCTCCATTCTTTGGATGGGTTATGAACTCCATCCTGATAAA
TGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAAT
GACATACAGAAGTTAGTGGGGAAATTGAATACCGCAAGTCAGATTTACCCA
GGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAA
CAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAAGTGGCAGAAAACA
GAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAGA
CTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAT
TTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGG
GGTGCCCACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAA
CCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTACCCAT
ACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTG
GATTCTGAGTGGGAGTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTAC
CAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGG
GCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAACAAA
GGAAGACAAAAGGTTGTCCCCCTAACTAACACAACAAATCAGAAAACCTCAGT
TACAAGCAATTTATCTAGCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTA
ACAGACTCACAATATGCATTAGGAATCATTCAAGCACAAACCAGATAAAAGTG

FIG. 14H

31/64

AATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGT
CTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGT
AGATAAATTAGTCAGTGCTGGAATCAGGAAAATACTATTTTATAGATGGAATA
GATAAGGCCCAAGATGAACATTAG

FIG. 14I

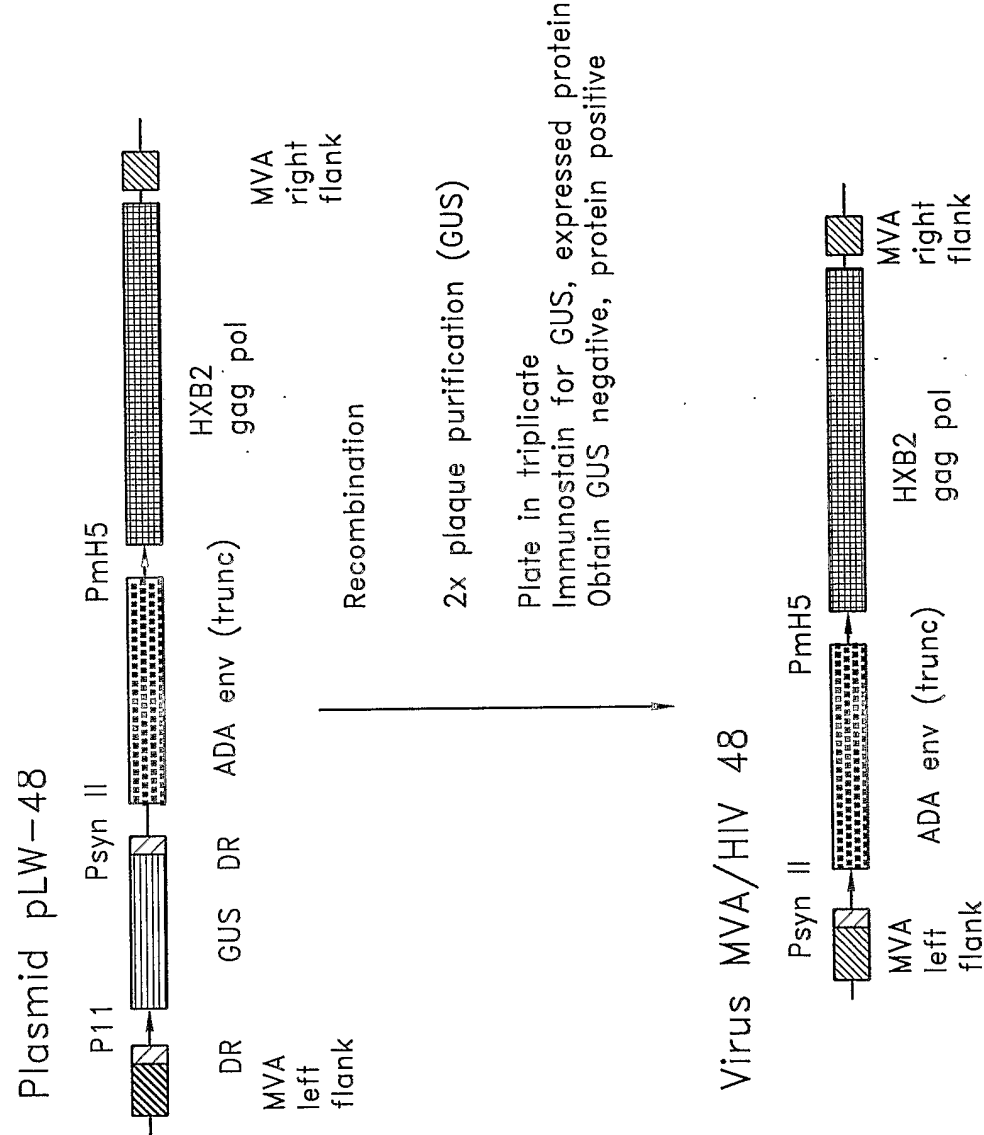


FIG. 15

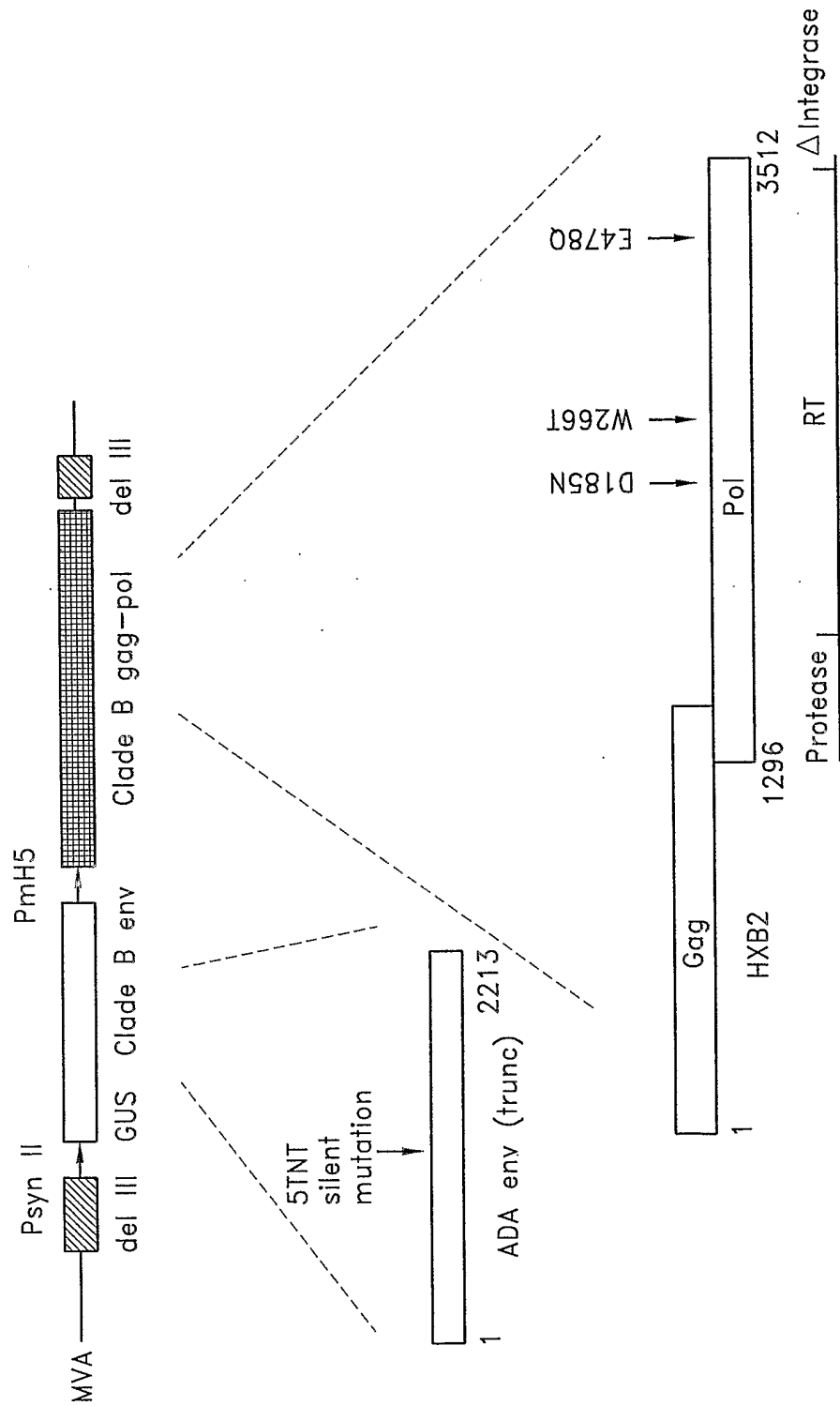


FIG. 16

34/64

Sequence of new Psyn II promoter:

Early part of promoter

Critical region Early start site

TAAAAATGAAAAATATTCTAATTTATAGGACGGT

Late part of promoter

TTTGATTTTCITTTTTTCTATGCTATAAATAATAAATA

FIG. 17

FIGURE 18

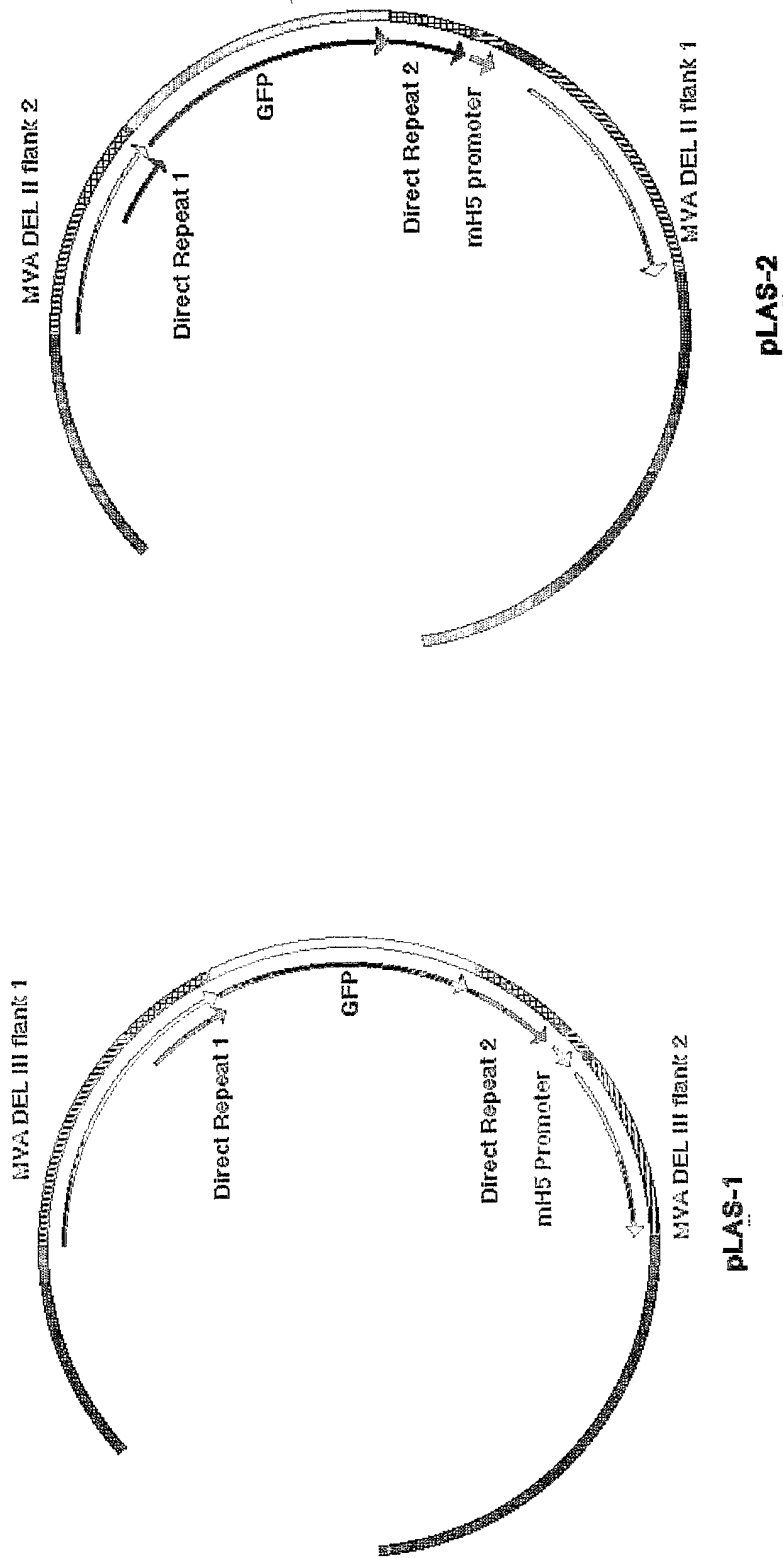


FIGURE 20

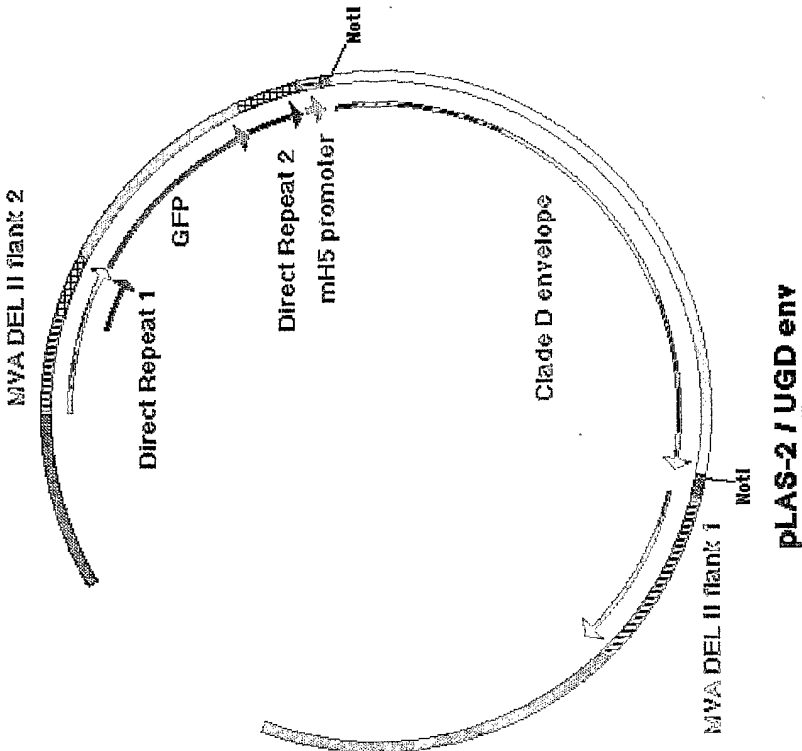


FIGURE 19

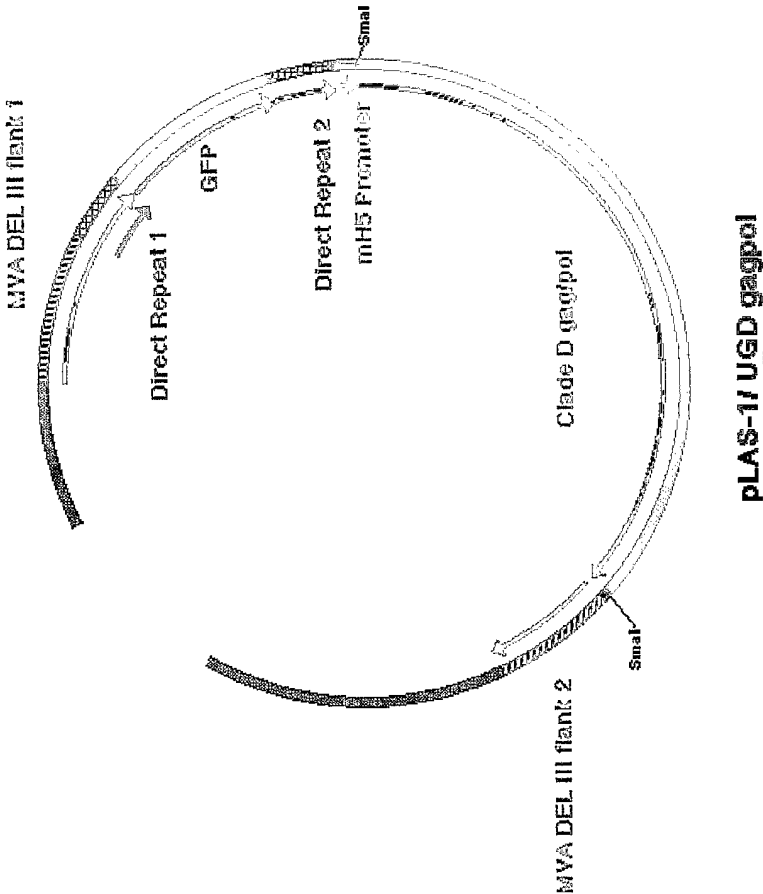


FIGURE 21

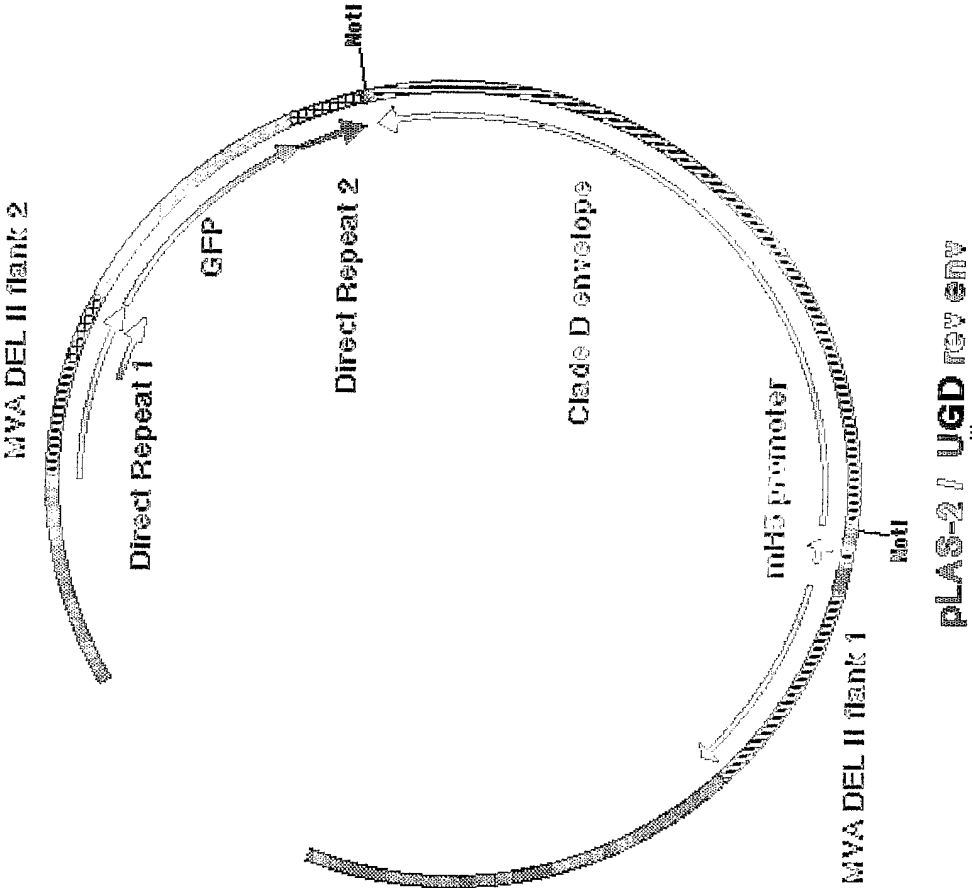


FIGURE 22

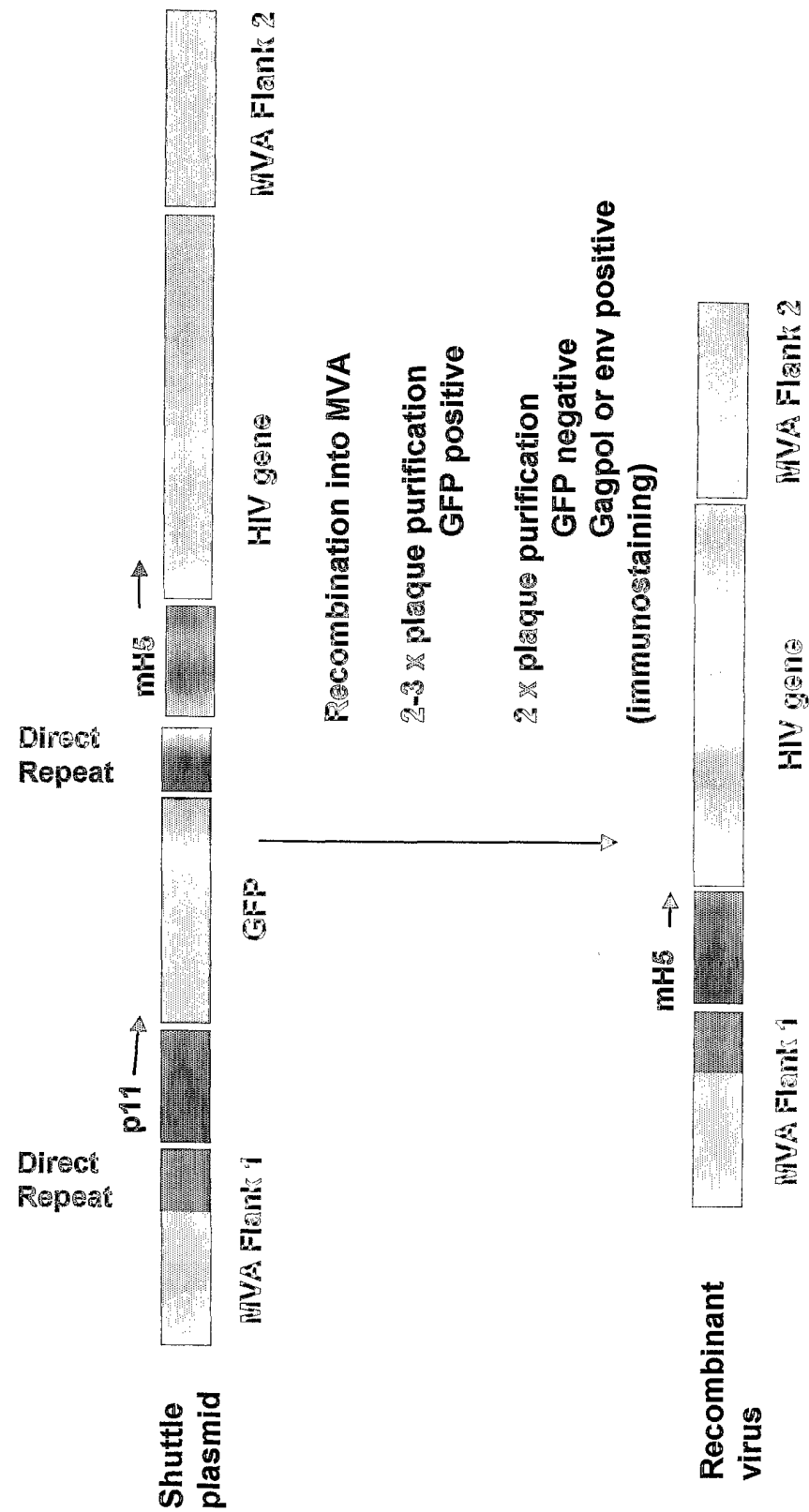


FIGURE 23

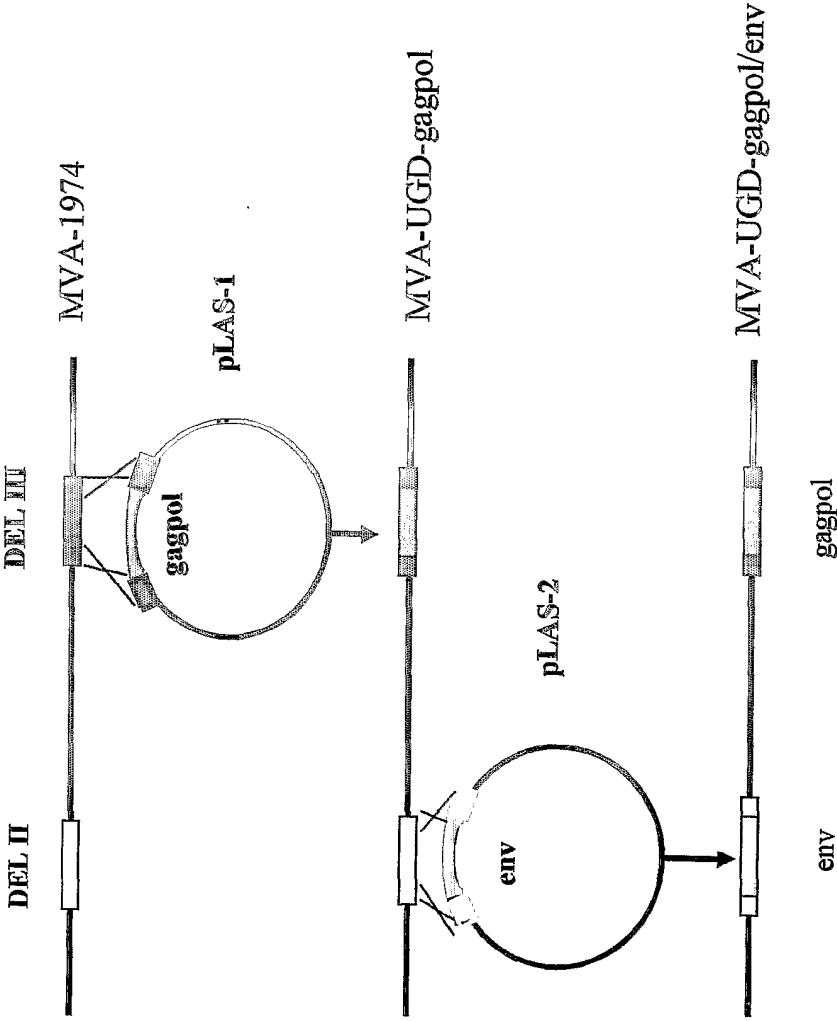


FIGURE 24

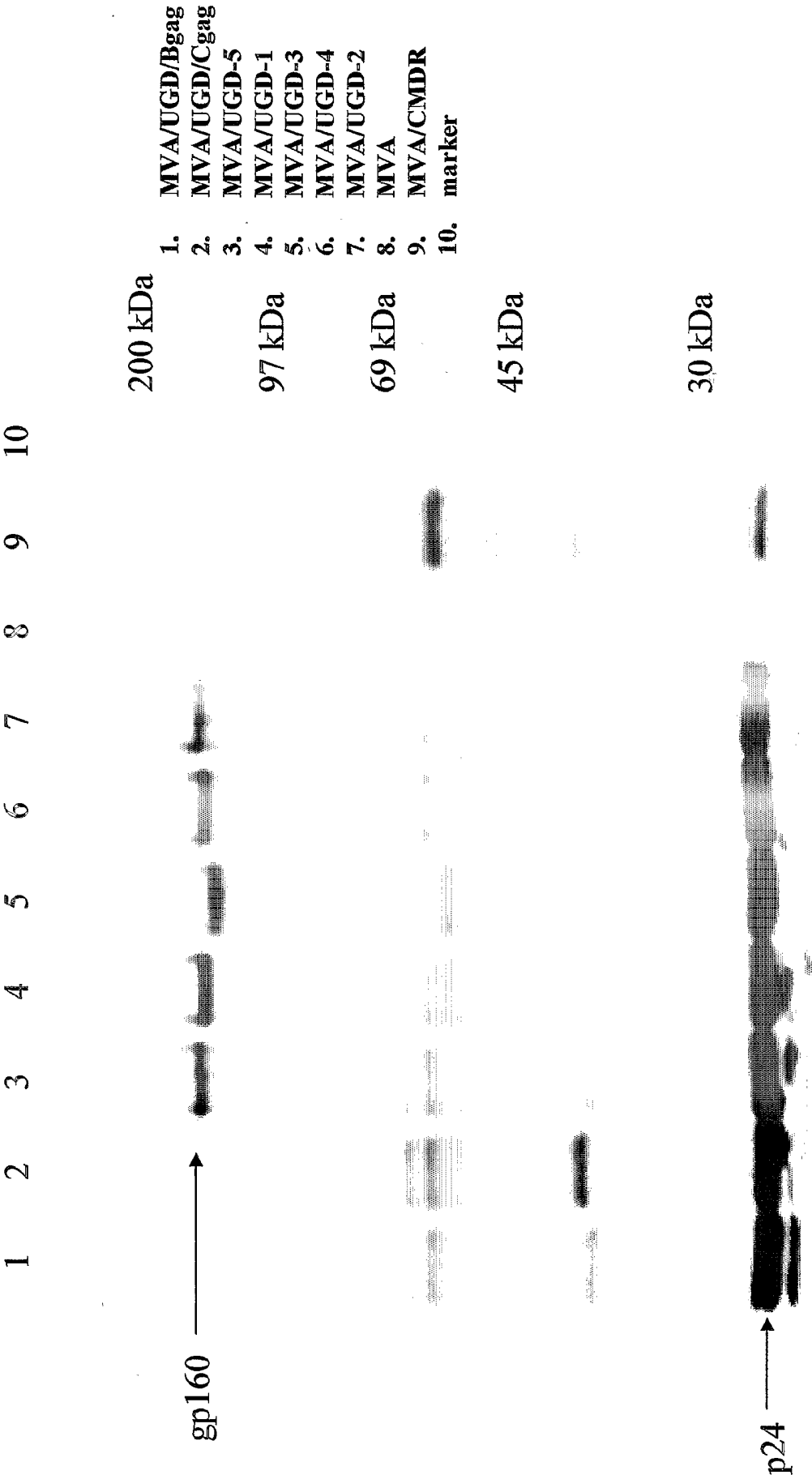
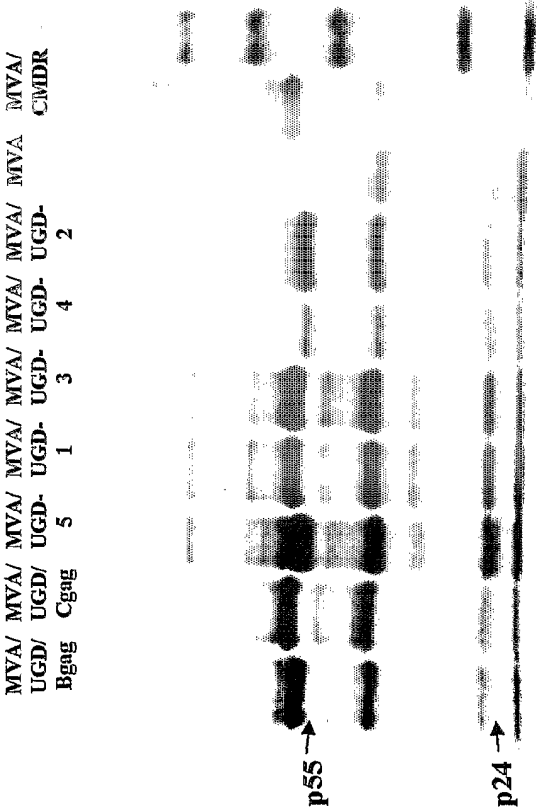


Figure 25.

A *Virus-like particle assay*



B *Env Fusion Assay*

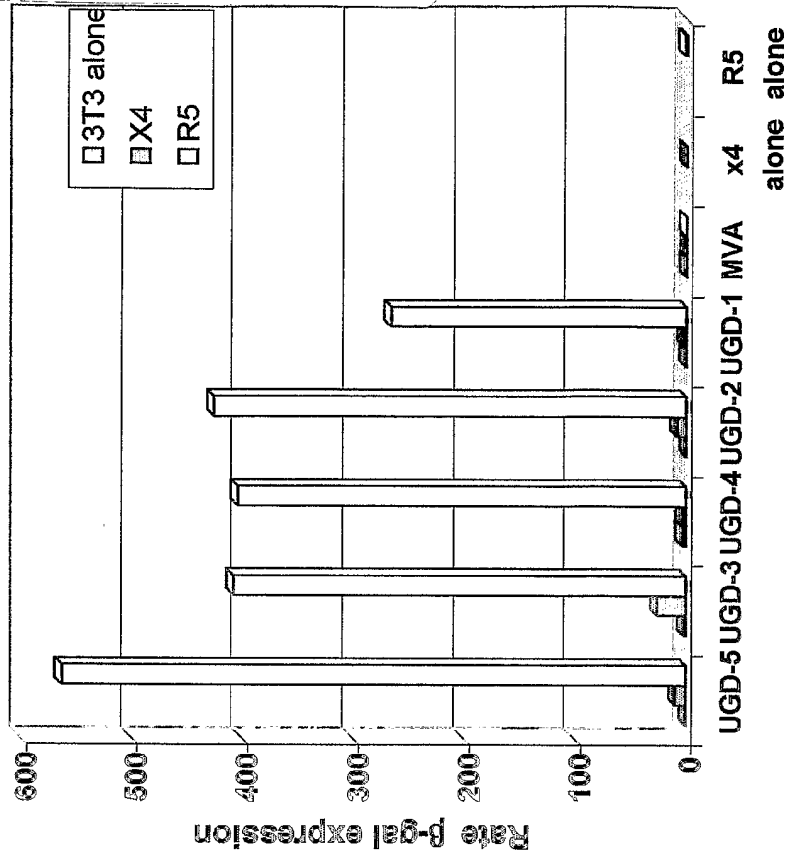
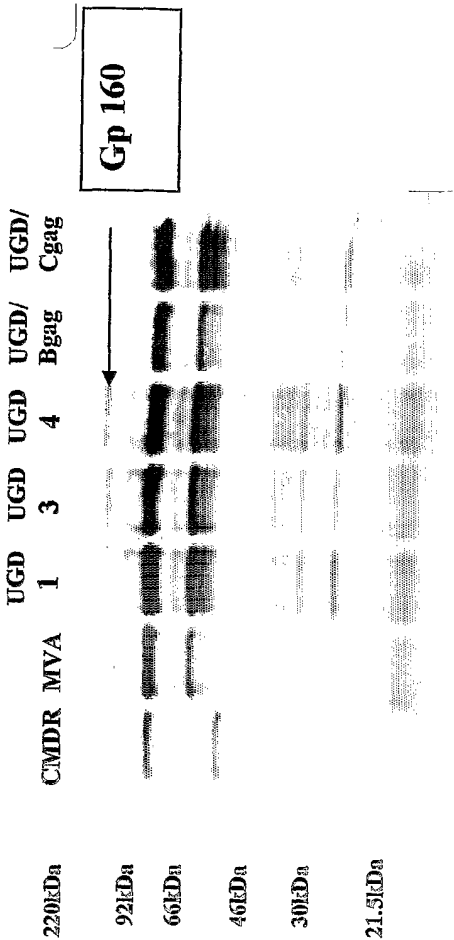
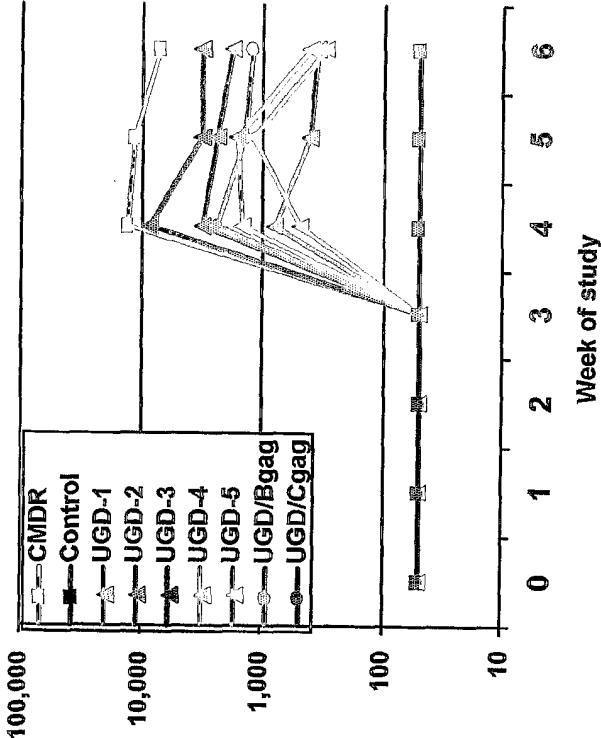


Figure 26.

A. HIV p24-specific serum IgG responses B. HIV env-specific serum IgG responses



Immunoprecipitation w/ autologous sera

Figure 26C.

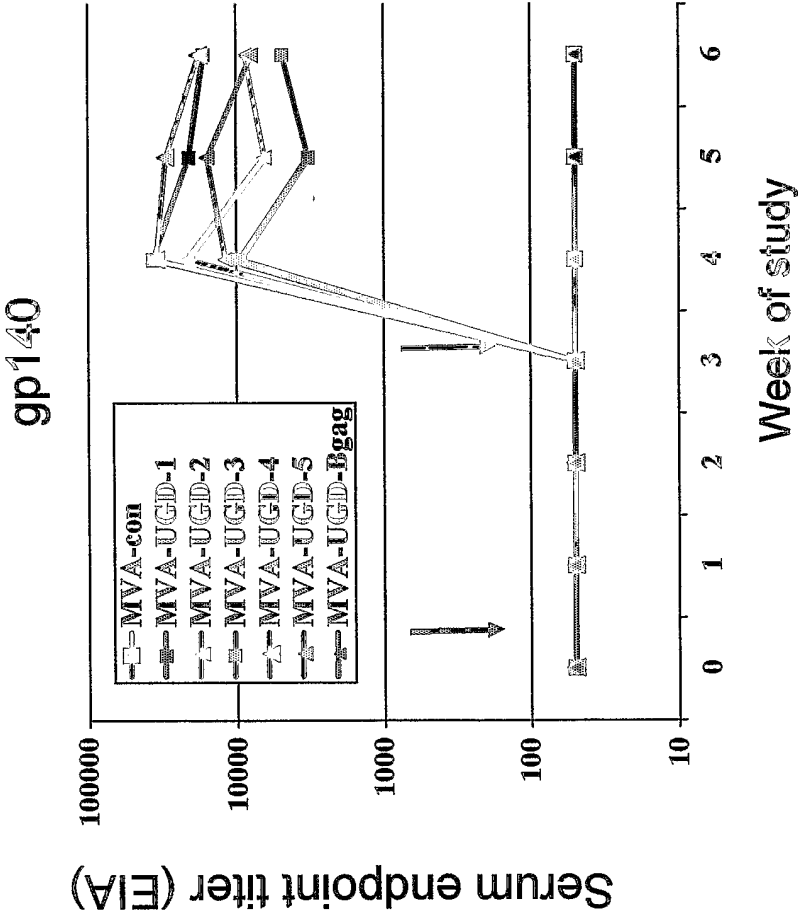


Figure 27.

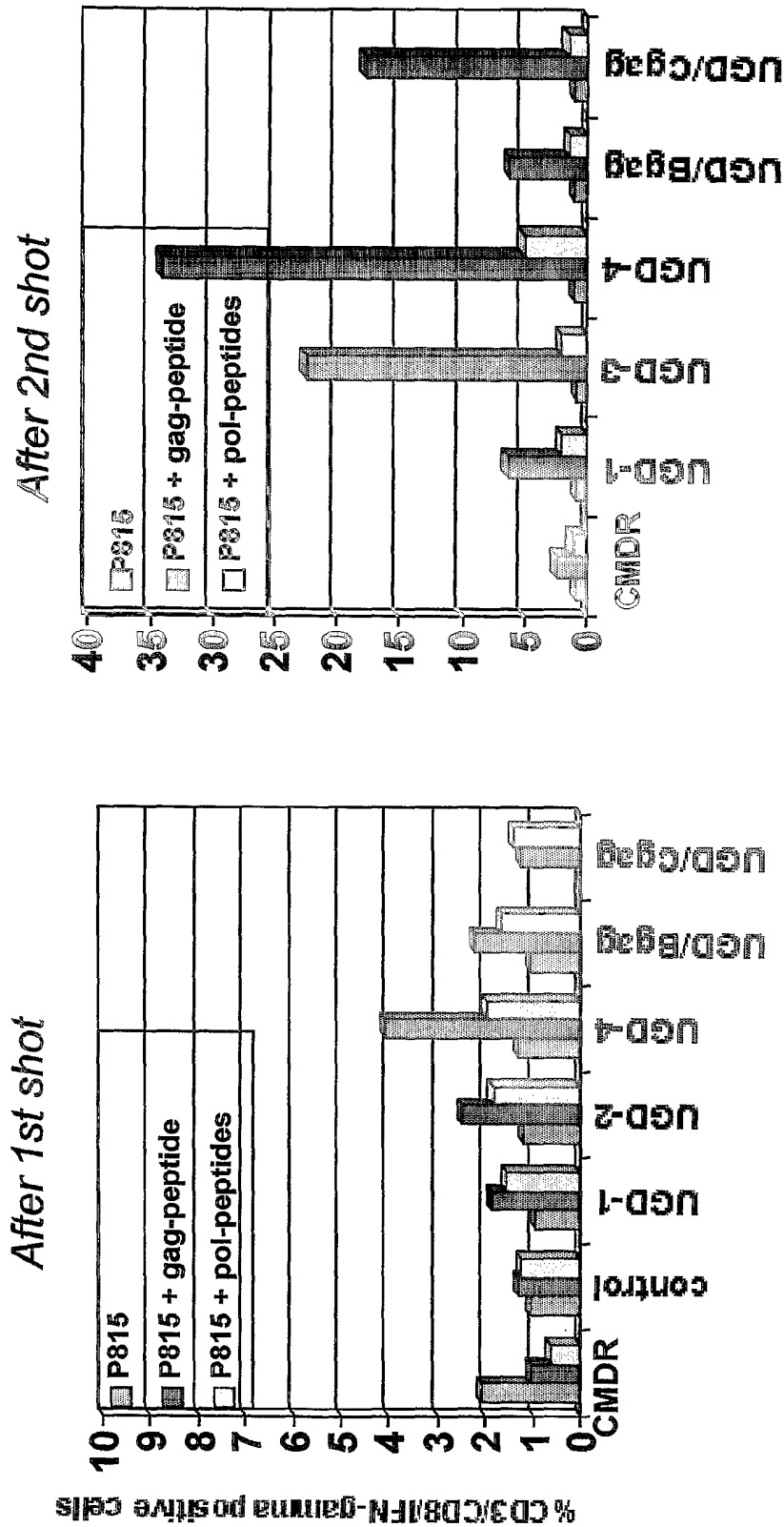


Figure 28A.

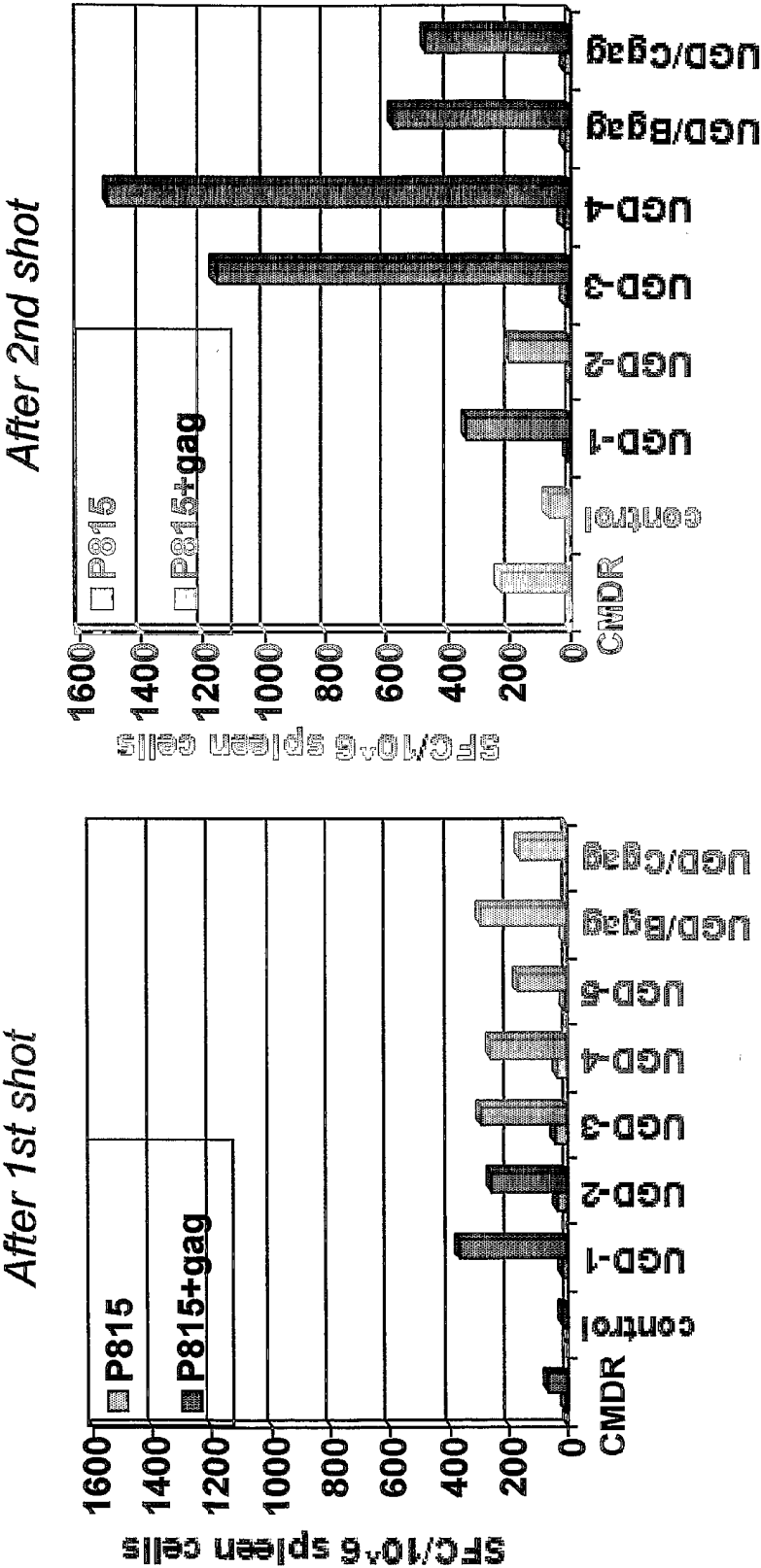


Figure 28B.

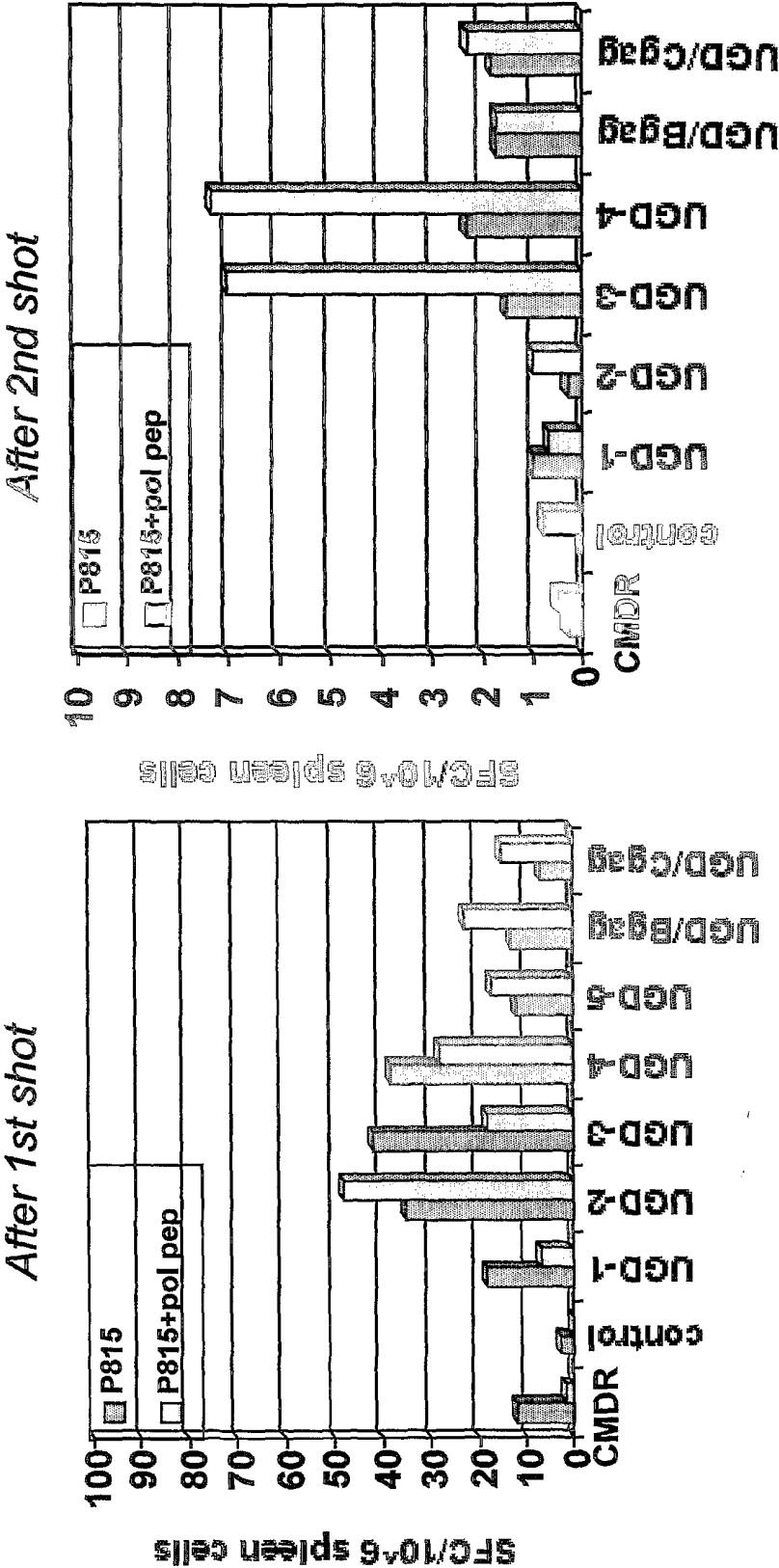


Figure 29.

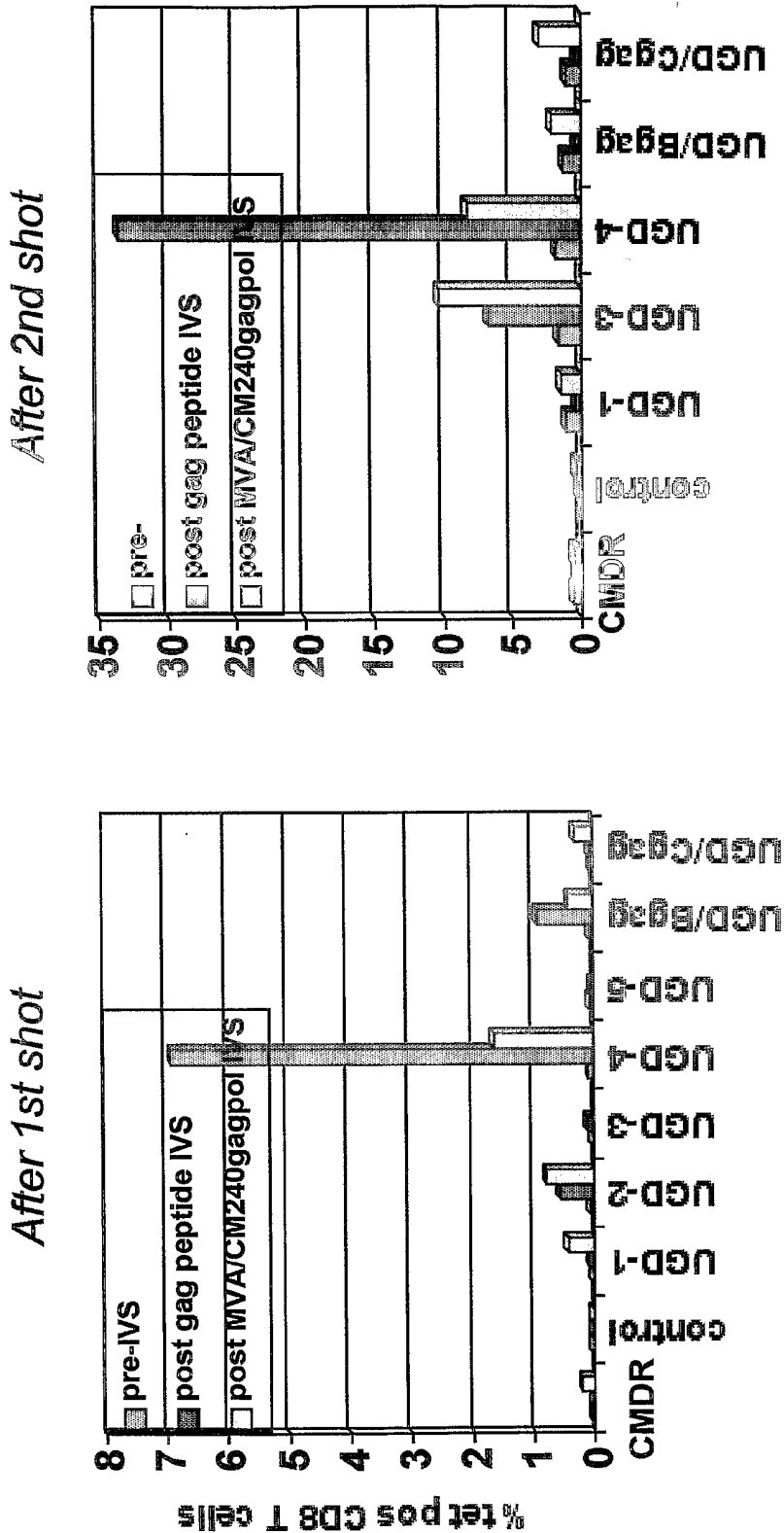


Figure 30.

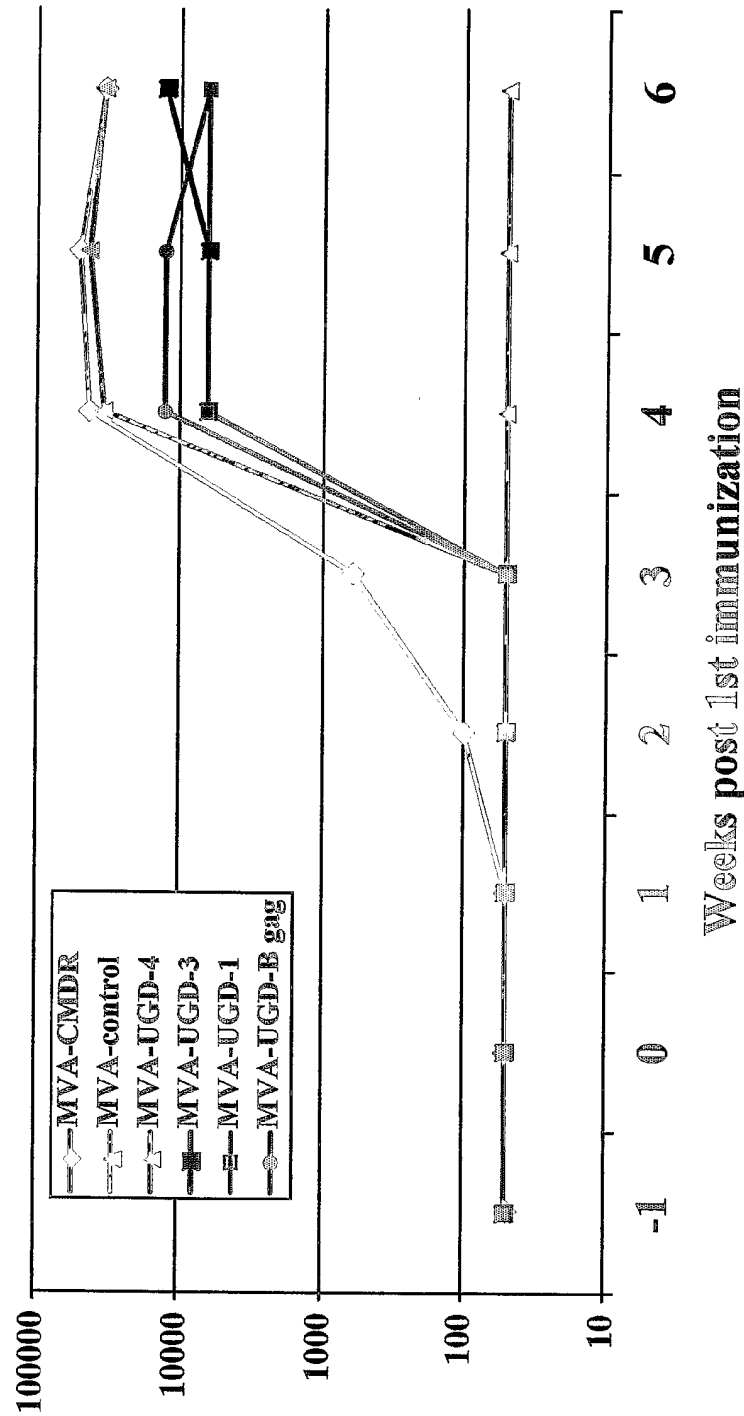


Figure 31A.

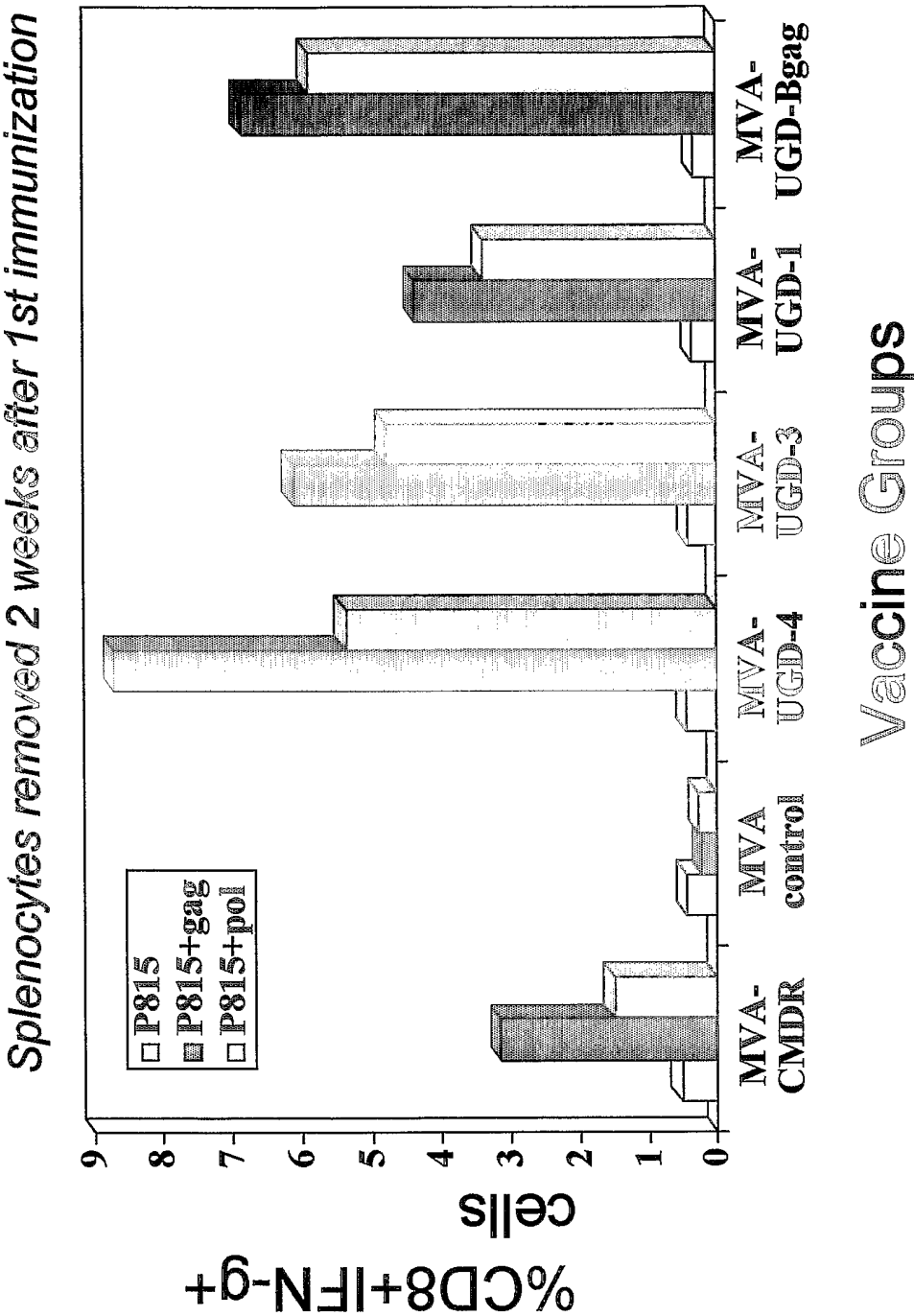


Figure 31B.

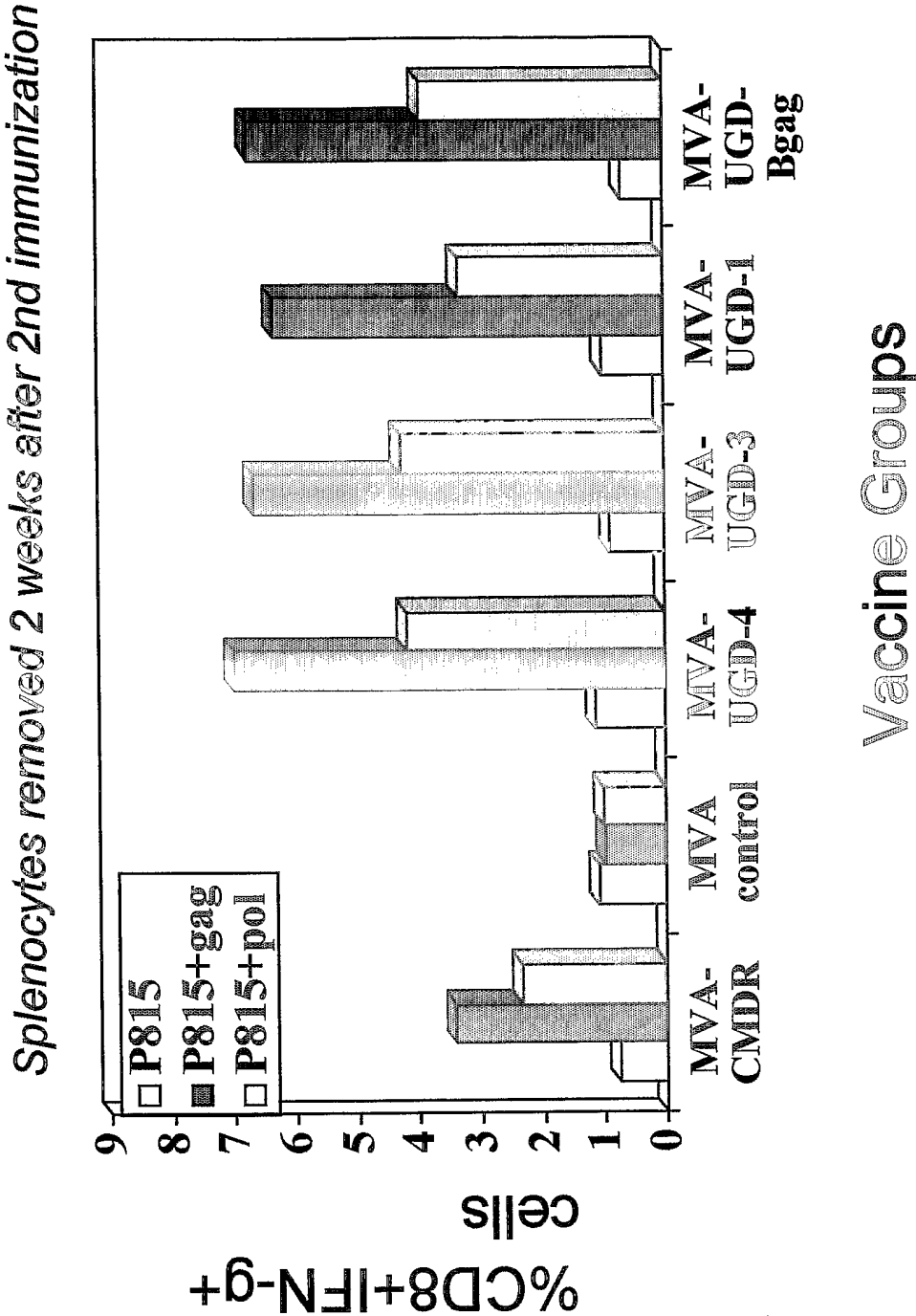


Figure 32A.

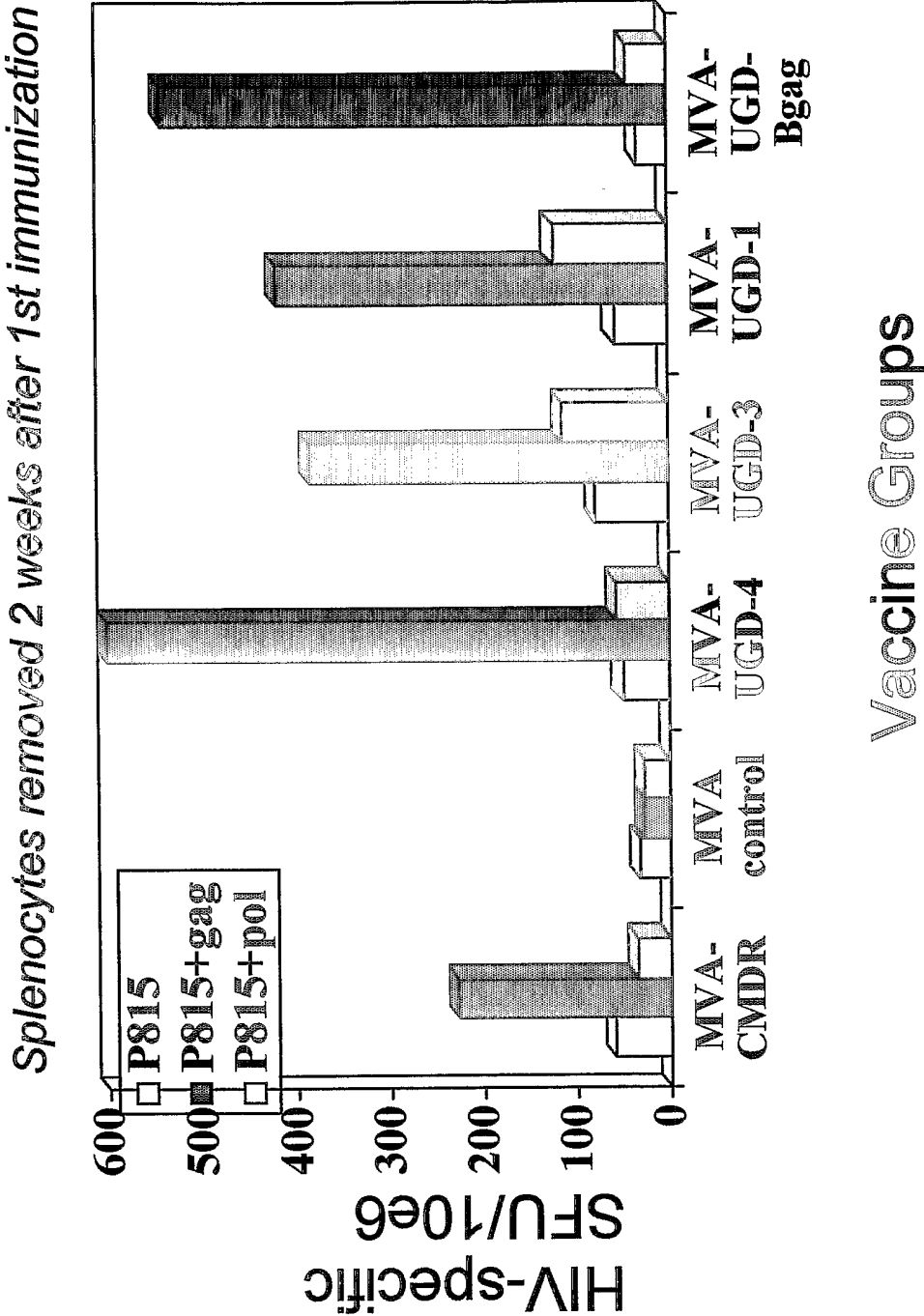


Figure 32B.

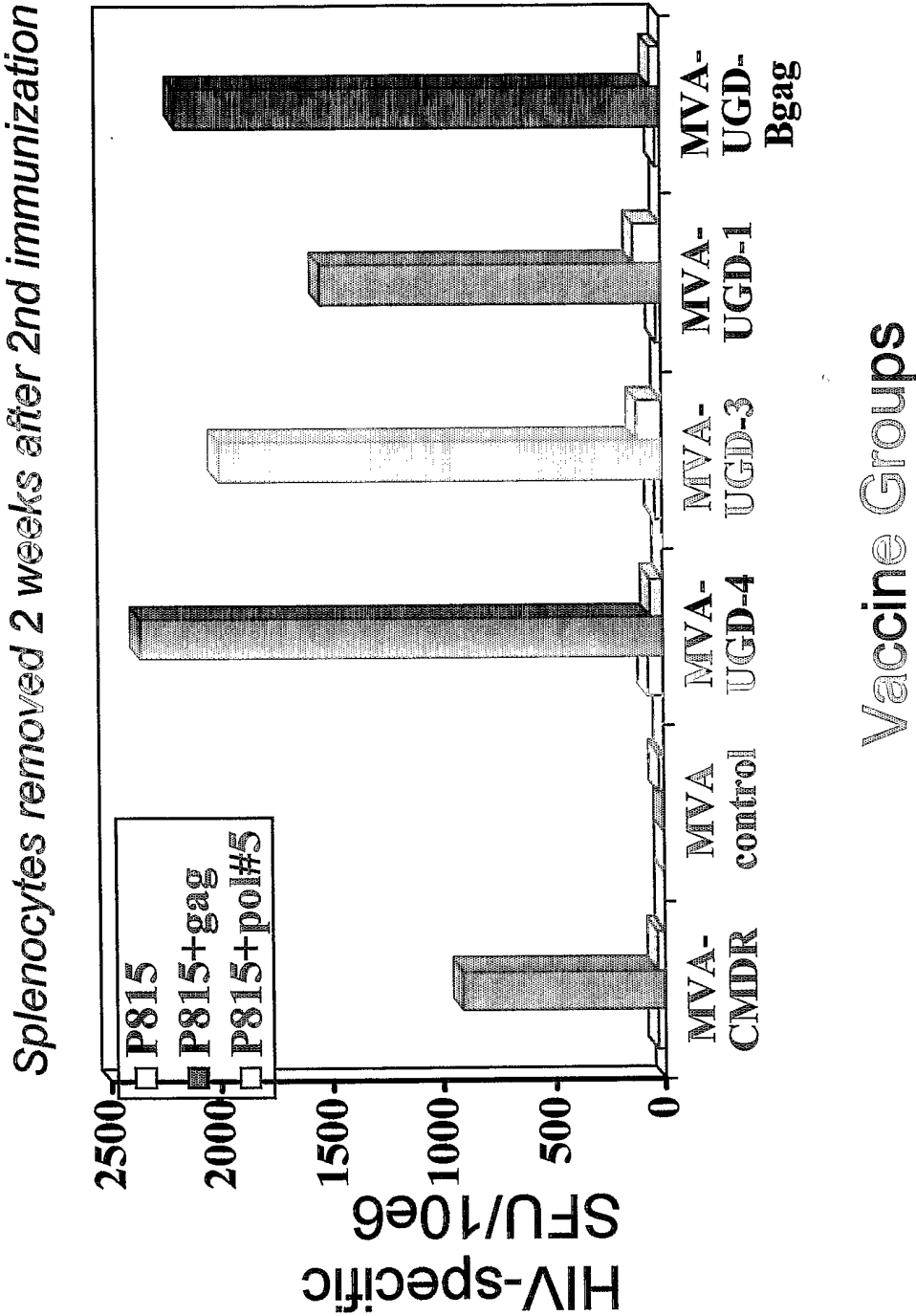


Figure 32C.

Splenocytes removed 2 weeks after 2nd immunization

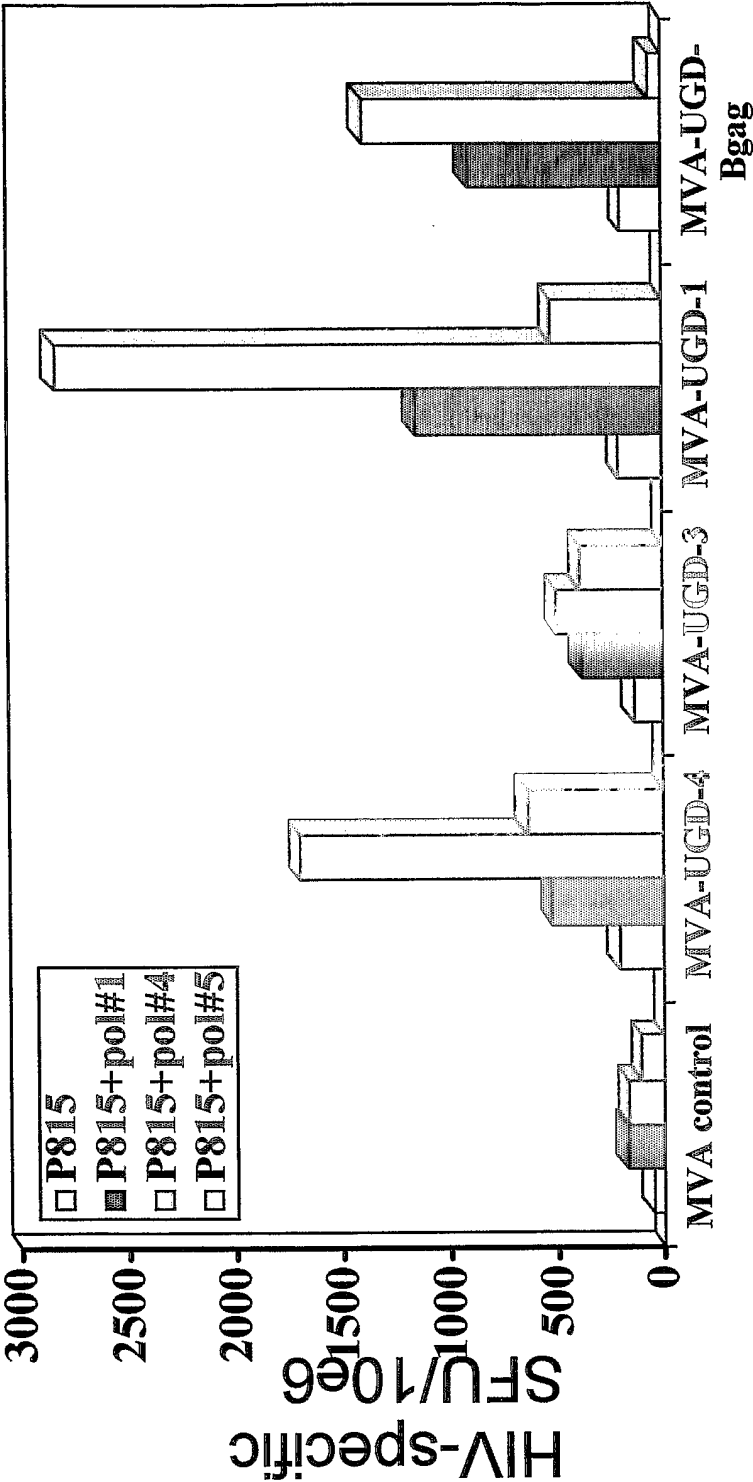


Figure 33.

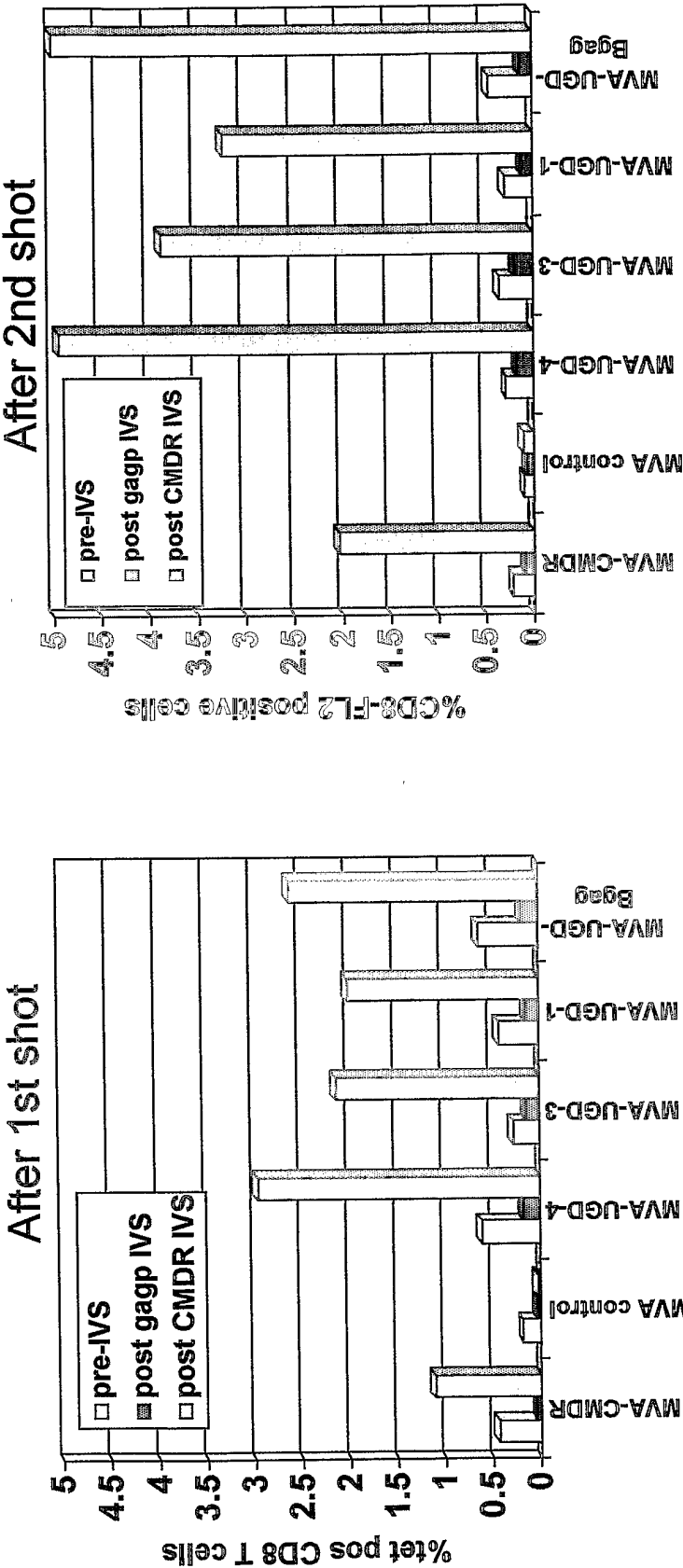
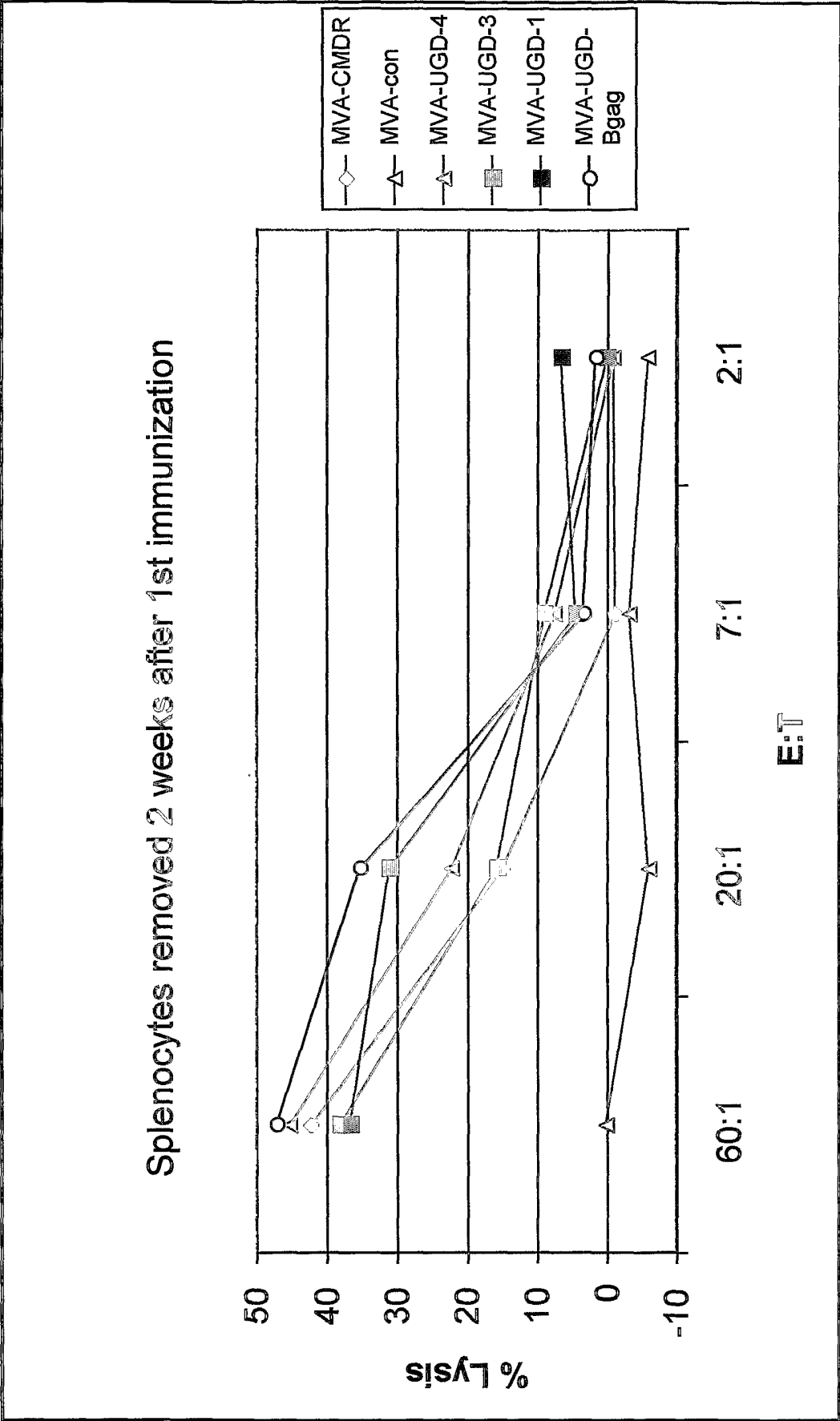
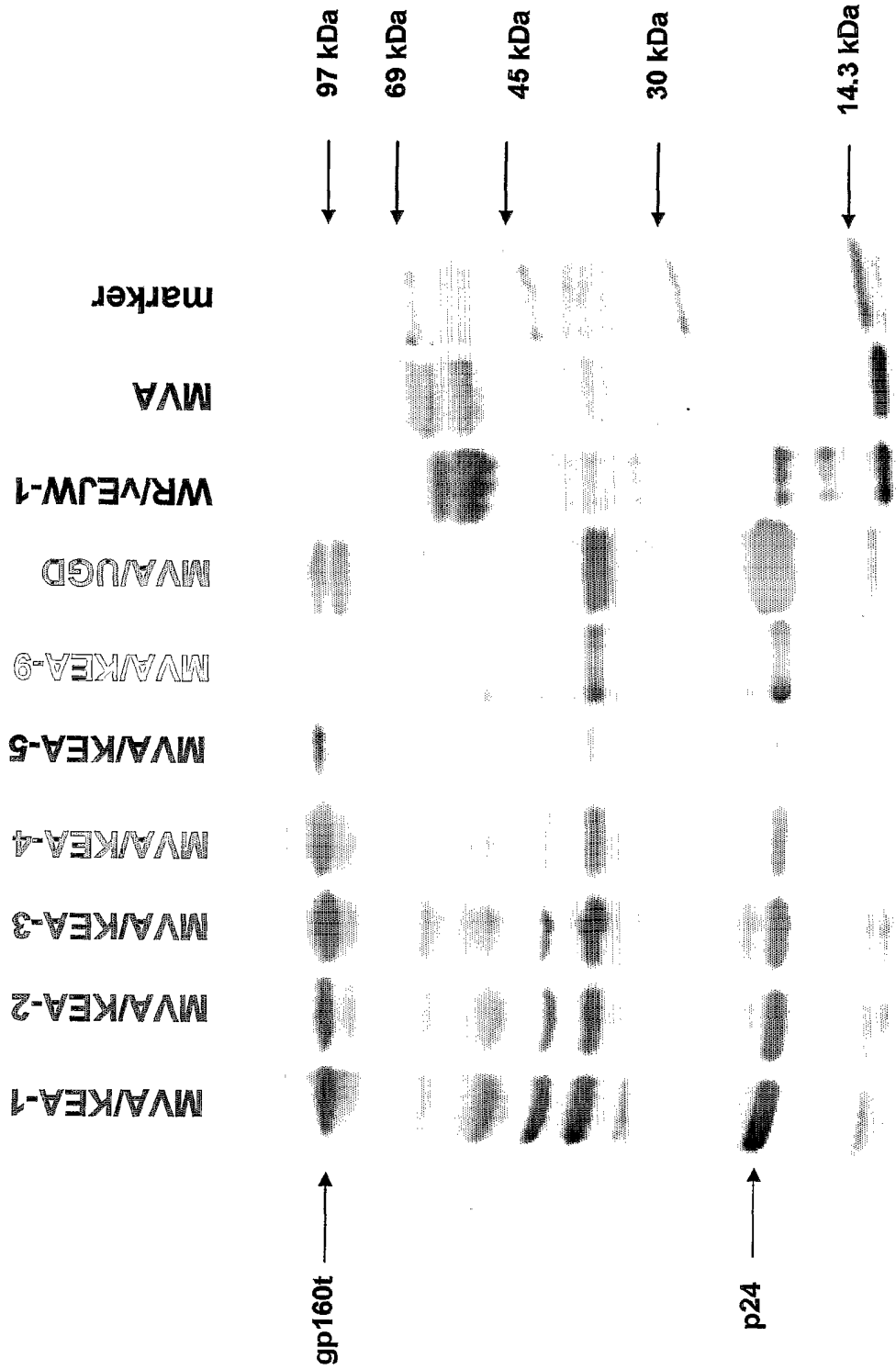


Figure 34.



56/64

Figure 35.



Antibodies used: T24 (env); 183-H12-5C (gag); pooled HIV+sera (clade A)

Figure 36.

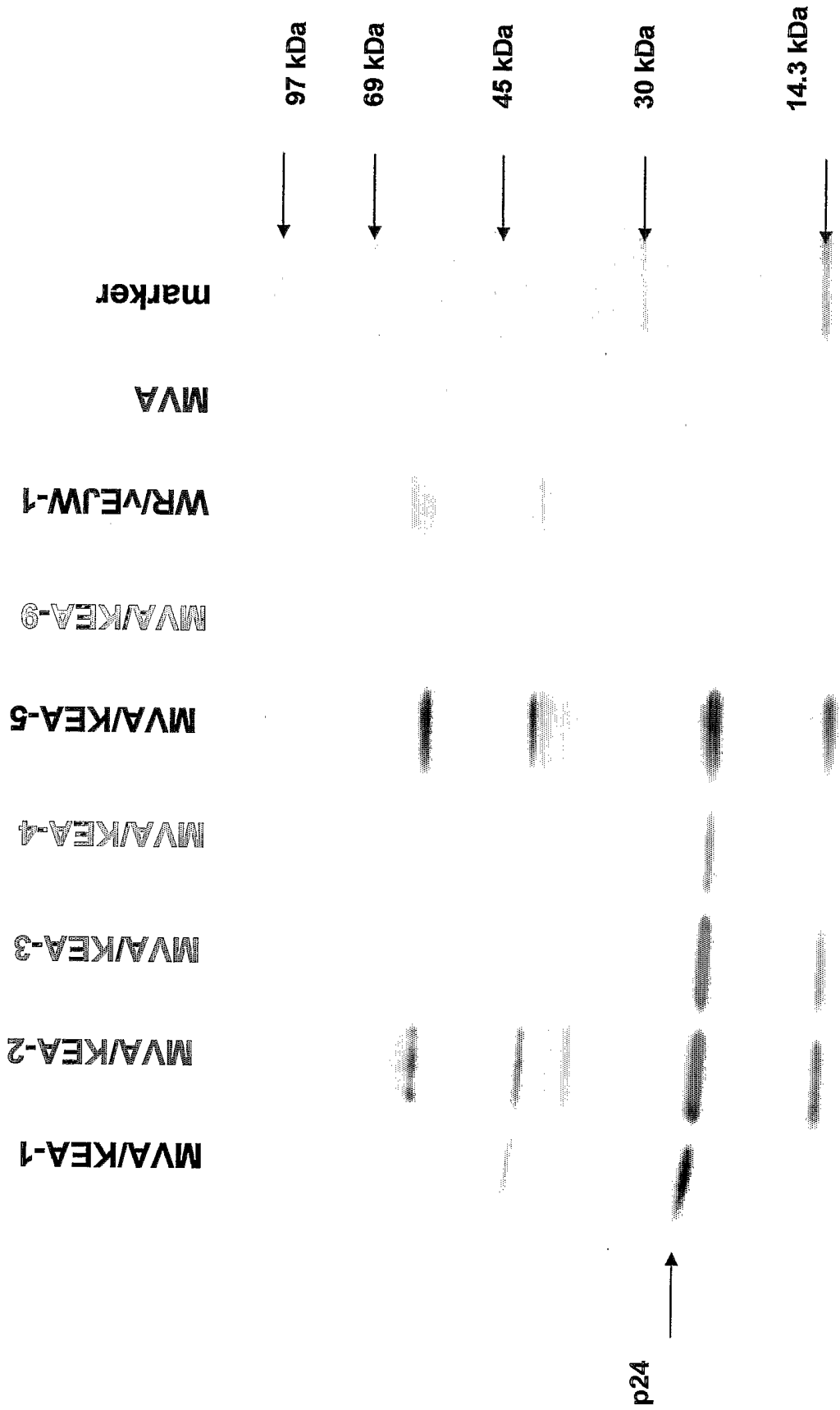


Figure 37

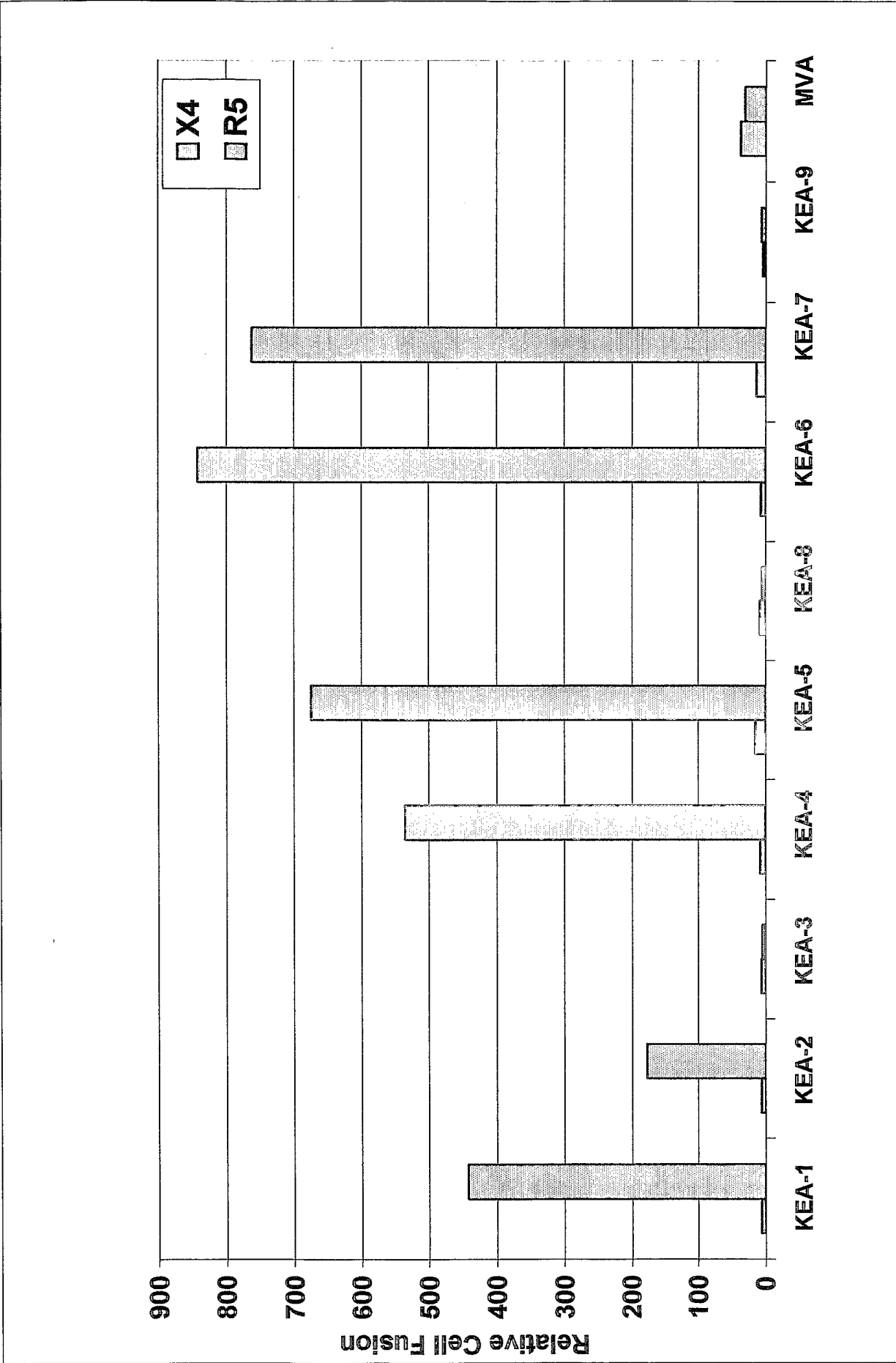


Figure 38.

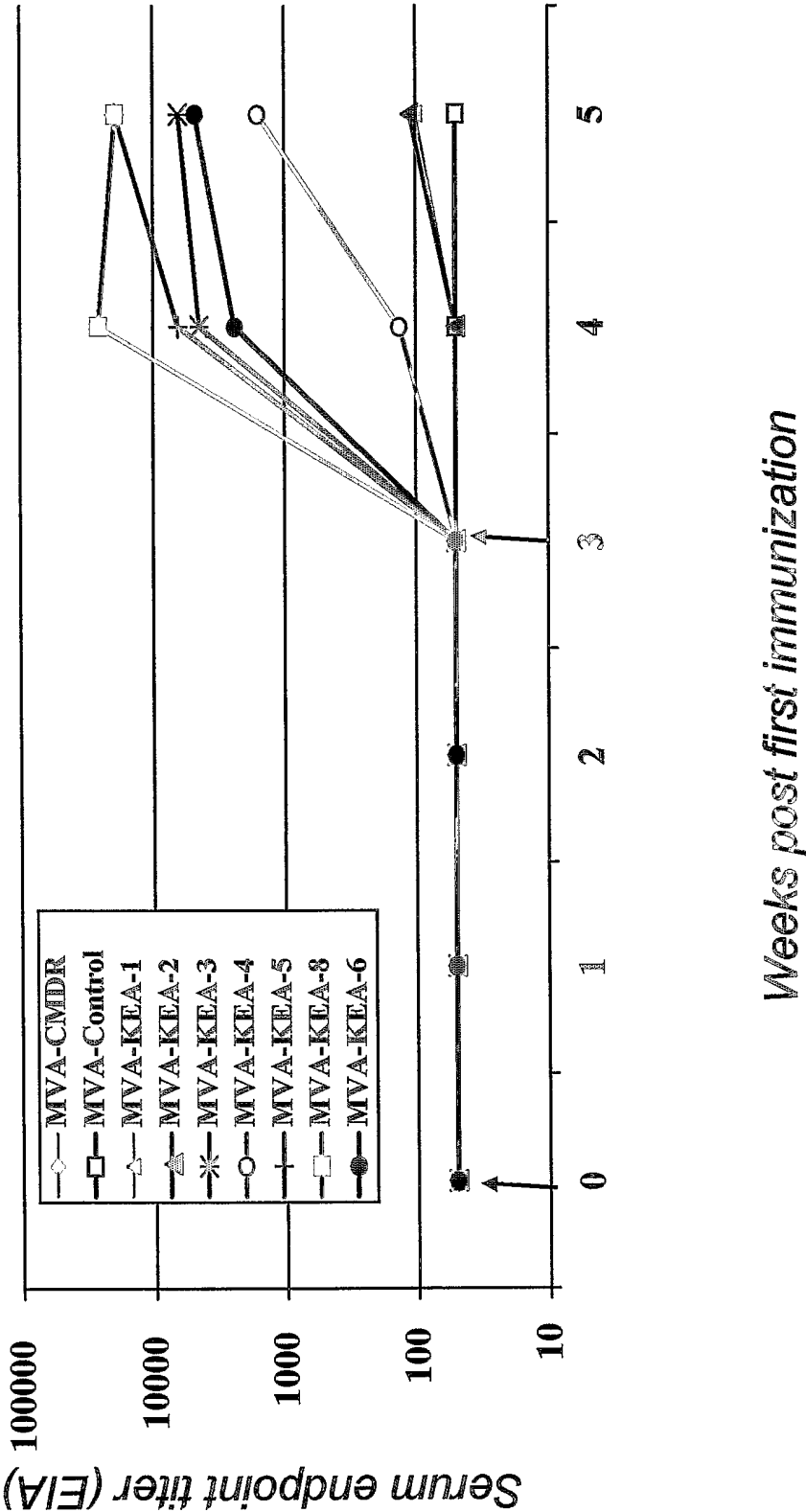


Figure 39.

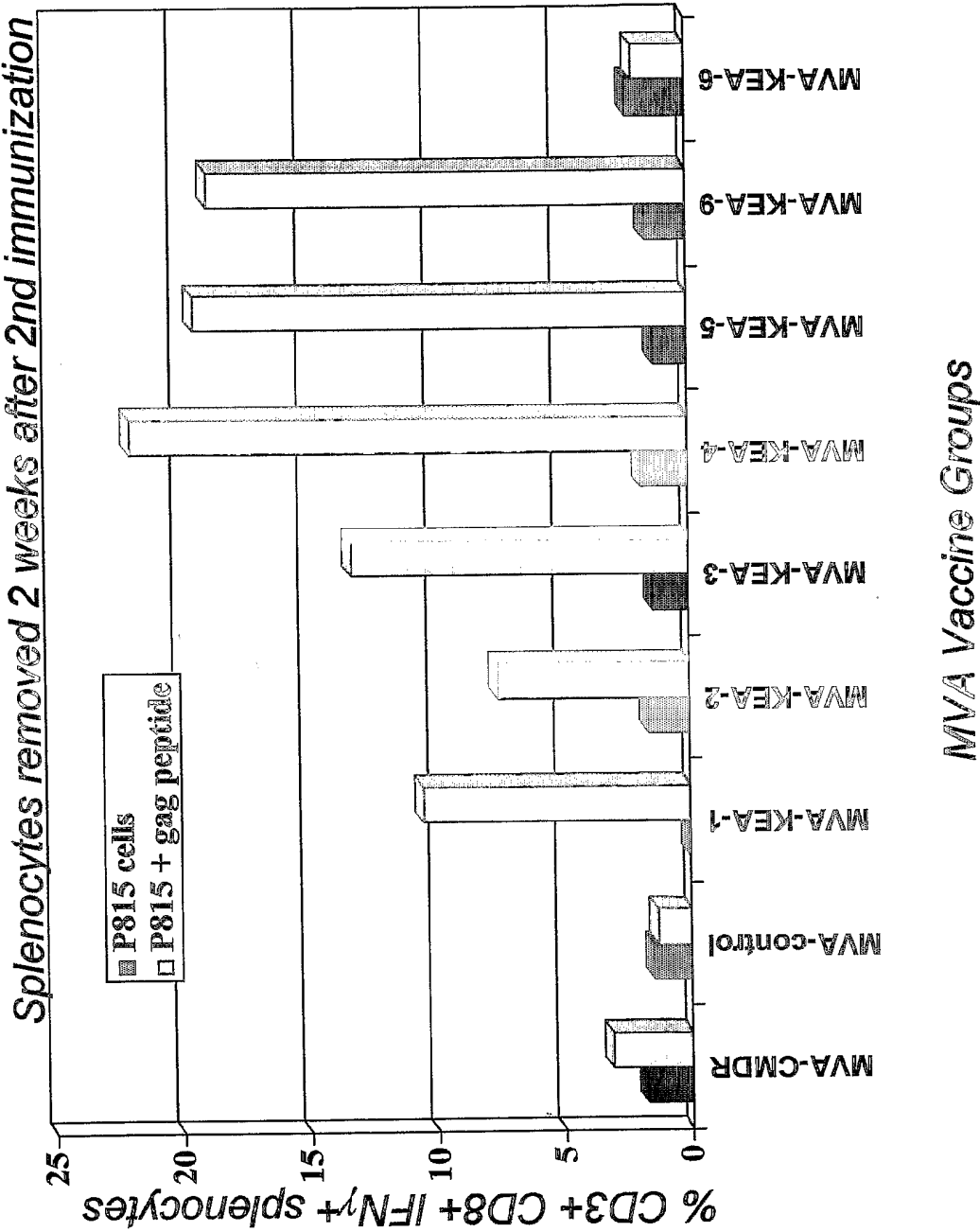


Figure 40.

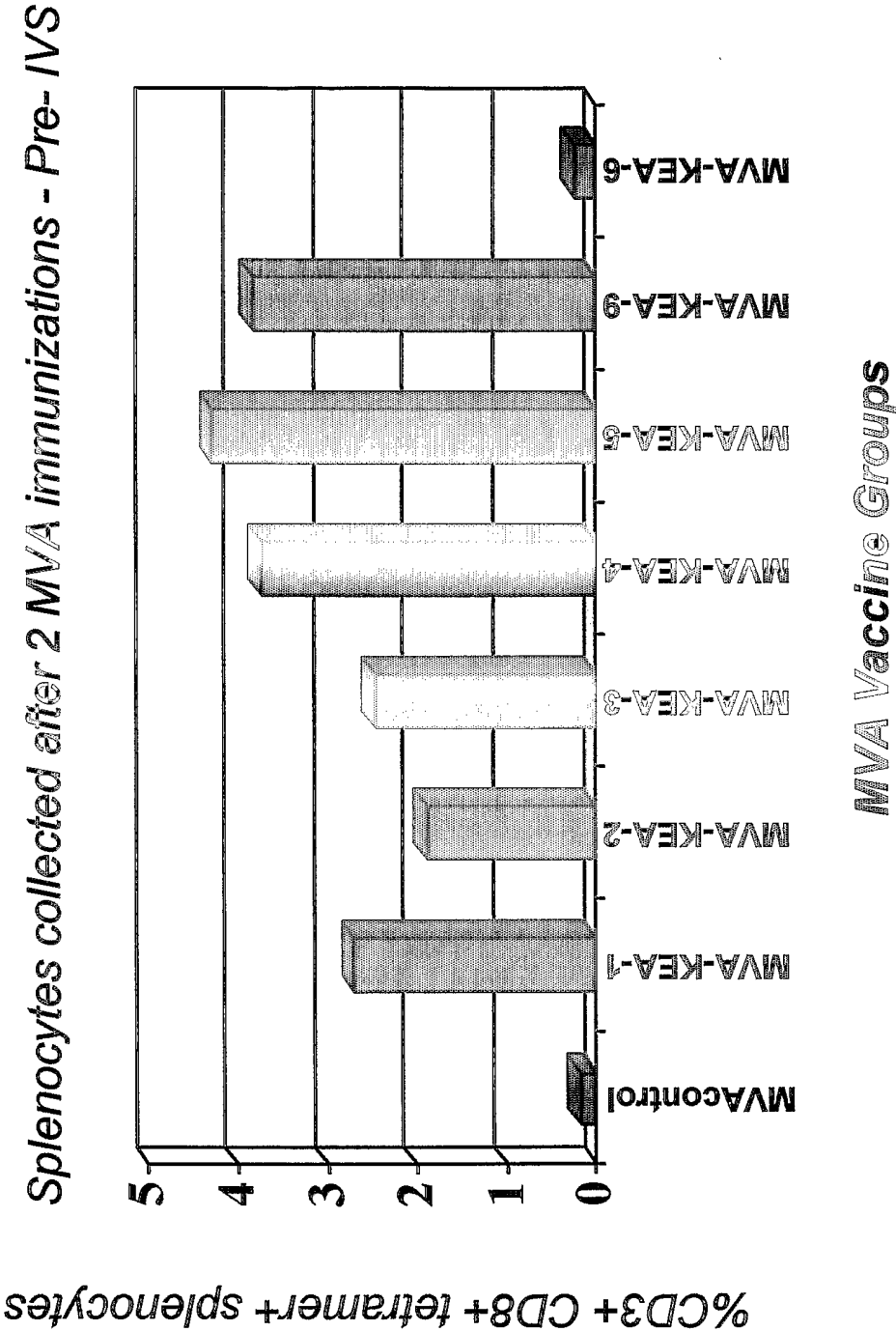


Figure 41.

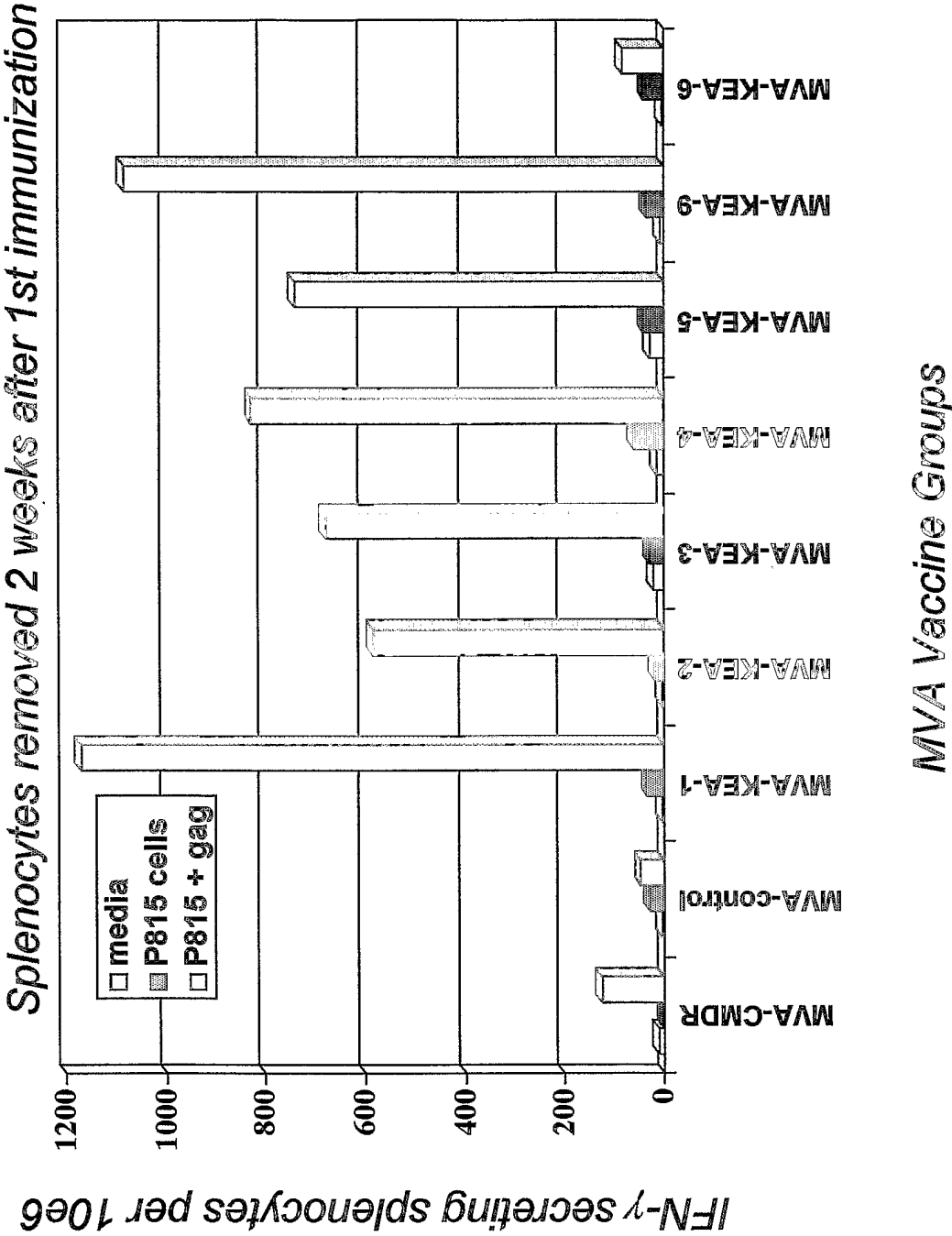


Figure 42:

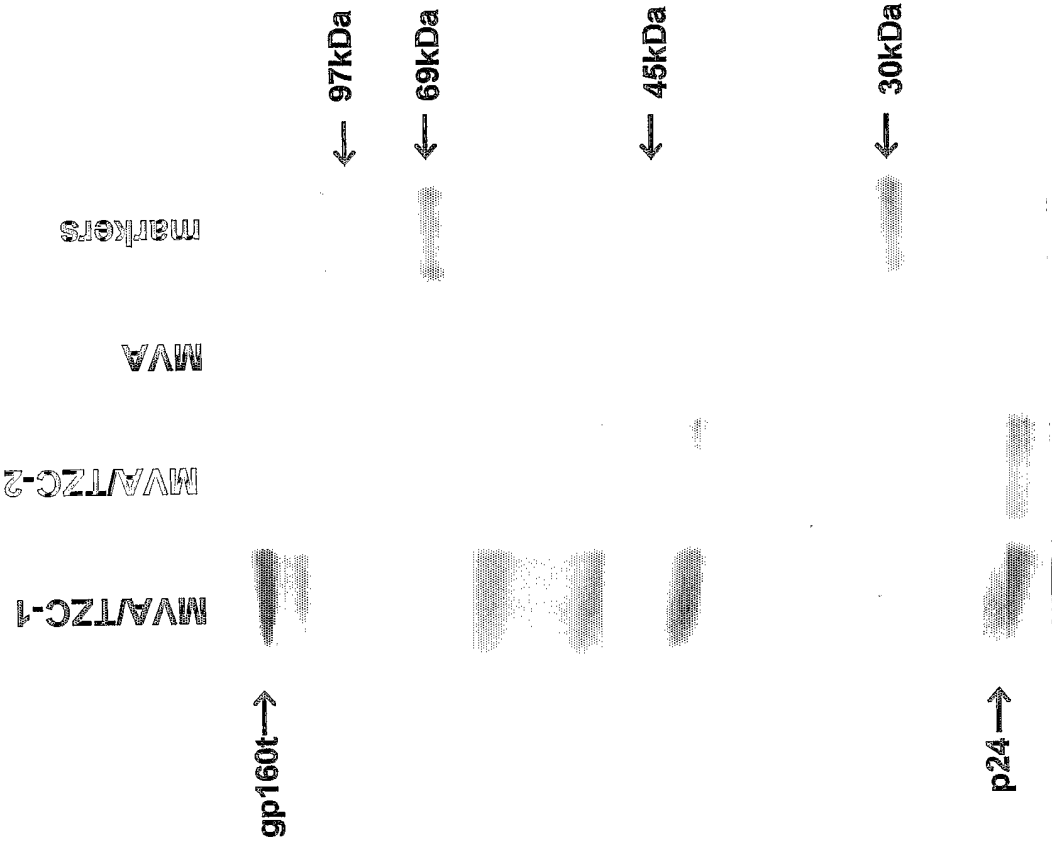
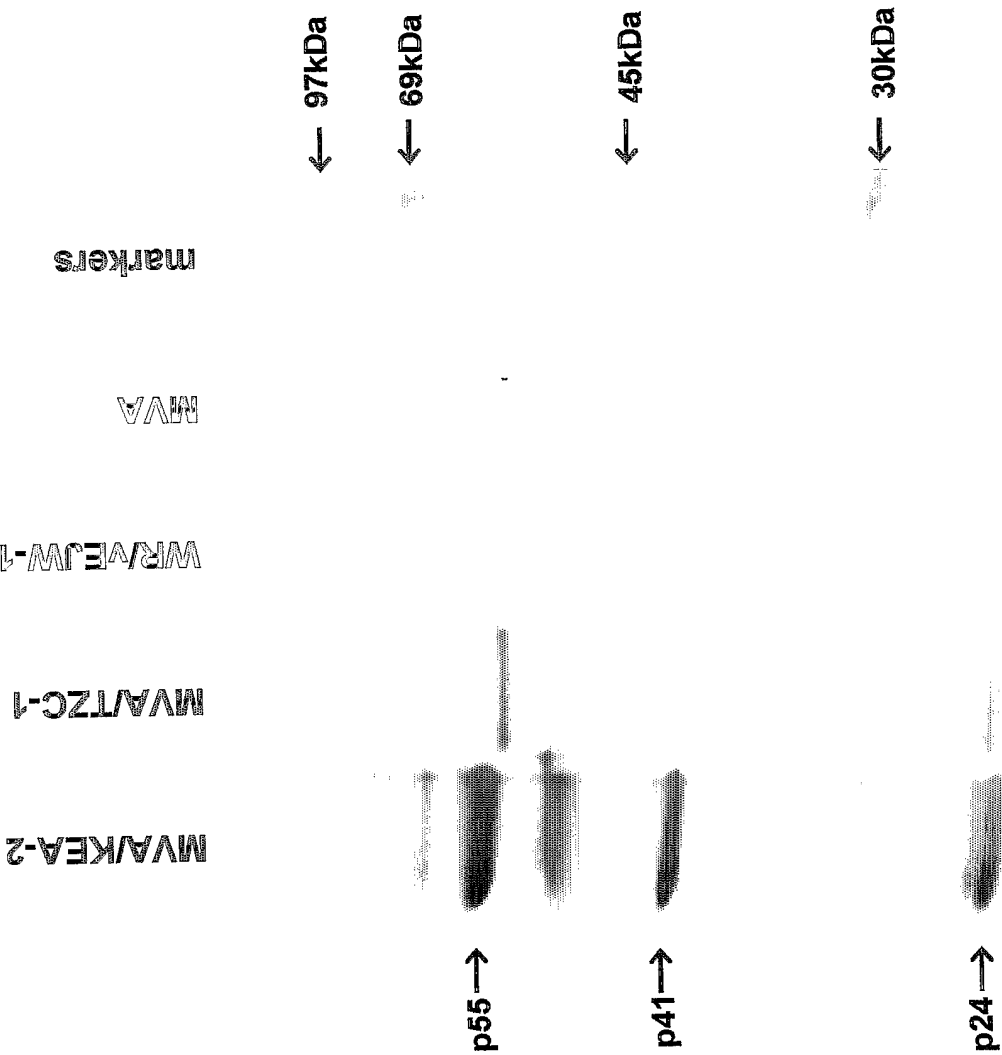


Figure 43.



NIH265.001VPC SEQLIST.TXT

SEQUENCE LISTING

<110> The Government of the United States of America, as represented by
Secretary, Health and Human Services
Moss, Bernard
Earl, Patricia L.
Wyatt, Linda
Steinmeyer, Leigh Anne
VanCott, Thomas
Harris, Matthew Edward

<120> MVA EXPRESSING MODIFIED HIV ENVELOPE,
GAG, AND POL GENES

<130> NIH265.001VPC

<150> US 60/459,175

<151> 2003-03-28

<160> 63

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 12225

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pLW-48

<400> 1

gaattcggttg	gtggctcgcca	tggatgggtgt	tattgtatac	tgtctaaacg	cgttagtaaaa	60
acatggcgag	gaaataaatc	atataaaaaa	tgatttcatg	attaaaccat	gttgtgaaaa	120
agtcaagaac	gttcacattg	gcggacaatc	taaaaacaat	acagtgattg	cagatttgcc	180
atatatggat	aatgcggtat	ccgatgtatg	caattcactg	tataaaaaga	atgtatcaag	240
aatatccaga	tttgctaatt	tgataaagat	agatgacgat	gacaagactc	ctactgggtg	300
atataattat	tttaaacctt	aagatgccat	tcctgttatt	atatccatag	gaaaggatag	360
agatgtttgt	gaactattaa	tctcatctga	taaagcgtgt	gcgtgtatag	agttaaattc	420
atataaagta	gccattcttc	ccatggatgt	ttcctttttt	accaaaggaa	atgcatcatt	480
gattattctc	ctgtttgatt	tctctatcga	tgcggcacct	ctcttaagaa	gtgtaaccga	540
taataatggt	attatatcta	gacaccagcg	tctacatgac	gagcttccga	gttccaattg	600
gttcaagttt	tacataagta	taaagtccga	ctattgttct	atattatata	tggttggtga	660
tggtactgtg	atgcatgcaa	tagctgataa	tagaacttac	gcaaataata	gcaaaaatat	720
attagacaat	actacaatta	acgatgagtg	tagatgctgt	tattttgaac	cacagattag	780
gattcttgat	agagatgaga	tgctcaatgg	atcatcgtgt	gatatgaaca	gacattgtat	840
tatgatgaat	ttacctgatg	taggcgaatt	tggtactagt	atgttgggga	aatatgaacc	900
tgacatgatt	aagattgctc	tttcgggtggc	tggtgaccag	gcgcgccttt	cattttgttt	960
ttttctatgc	tataaatggt	acgtcctgta	gaaaccccaa	cccgtgaaat	caaaaaactc	1020
gacggcctgt	gggcattcag	tctggatcgc	gaaactgtgt	gaattgtatc	gcgttggtgg	1080
gaaagcgcgt	tacaagaaag	ccgggcaatt	gctgtgccag	gcagttttta	cgatcagttc	1140
gccgatgcag	atattcgtaa	ttatgcgggc	aacgtctggt	atcagcgcca	agtctttata	1200
ccgaaagggt	gggcaggcca	gcgtatcgtg	ctgcgtttcg	atgcggtcac	tcattacggc	1260
aaagtgtggg	tcaataatca	ggaagtgatg	gagcatcagg	gcggctatac	gccatttgaa	1320
gccgatgtca	cgccgtatgt	tattgccggg	aaaagtgtac	gtatcaccgt	ttgtgtgaac	1380
aacgaactga	actggcagac	tatcccgccg	ggaatgggtga	ttaccgacga	aaacggcaag	1440
aaaaagcagt	cttactttcca	tgatttcttt	aactatgccg	gaatccatcg	cagcgtaatg	1500
ctctacacca	cgccgaacac	ctgggtggac	gatatcaccg	tggtgacgca	tgctcgcgcaa	1560
gactgttaacc	acgcgtctgt	tgactggcag	gtgggtggcca	atggtgatgt	cagcgttgaa	1620
ctgcgtgatg	cggatcaaca	ggtggttgca	actggacaag	gcactagcgg	gactttgcaa	1680
gtggtgaatc	agcacctctg	gcaaccgggt	tctatgaact	tctatgaact	gtgcgtcaca	1740
gccaaaagcc	agacagagtg	tgatatctac	ccgcttcgcg	tcggcatccg	gtcagtggca	1800
gtgaaggggc	aacagttcct	gattaaccac	aaaccgttct	actttactgg	ctttggctgt	1860
catgaagatg	cggacttgcg	tggcaaagga	ttcgataacg	tgctgatggg	gcacgaccac	1920
gcattaatgg	actggattgg	ggccaactcc	taccgtacct	cgattaccc	ttacgctgaa	1980
gagatgctcg	actgggcaga	tgaacatggc	atcgtgggtg	ttgatgaaac	tgctgctgtc	2040
ggctttaacc	tctcttttag	cattggtttc	gaagccggca	acaagccgaa	agaactgtac	2100

NIH265.001VPC SEQLIST.TXT

agcgaagagg	cagtcaacgg	ggaaactcag	caagcgcact	tacaggcgat	taaagagctg	2160
atagcgcgtg	acaaaaacca	cccaagcgtg	gtgatgtgga	gtattgccaa	cgaaccggat	2220
acccgtccgc	aaggtgcacg	ggaatatttc	gcgccactgg	cggaagcaac	gcgtaaactc	2280
gacccgacgc	gtccgatcac	ctgcgtcaat	gtaatgttct	gcgacgctca	caccgatacc	2340
atcagcgatc	tctttgatgt	gctgtgcctg	aaccgttatt	acggatggta	tgtccaaagc	2400
ggcgatttgg	aaacggcaga	gaaggctactg	gaaaaagaac	ttctggcctg	gcaggagaaa	2460
ctgcatcagc	cgattatcat	caccgaatac	ggcgtggata	cgttagccgg	gctgcactca	2520
atgtacaccg	acatgtggag	tgaagagtat	cagtgtgcat	ggctggatat	gtatcaccgc	2580
gtctttgatc	gcgtcagcgc	cgctcgtcgt	gaacagggtat	ggaatttcgc	cgattttgcg	2640
acctcgcaag	gcatattgcg	cgttggcggg	aacaagaaag	ggatcttcac	tcgcgaccgc	2700
aaaccgaagt	cggcggcctt	tctgctgcaa	aaacgcctgga	ctggcatgaa	cttcgggtgaa	2760
aaaccgcagc	agggaggcaa	acaatgagag	ctcggttgtt	gatggatctg	tgatgcatgc	2820
aatagctgat	aatagaactt	acgcaaatat	tagcaaaaat	atattagaca	atactacaat	2880
taacgatgag	tgtagatgct	gttattttga	accacagatt	aggattcctg	atagagatga	2940
gatgctcaat	ggatcatcgt	gtgatatgaa	cagacattgt	attatgatga	atttacctga	3000
tgtaggcgaa	tttgatctta	gtatgttggg	gaaatatgaa	cctgacatga	ttaagattgc	3060
tctttcgggtg	gctggcggcc	cgctcgagta	aaaaatgaaa	aaatattcta	attttatagga	3120
cggttttgg	tttctttttt	tctatgctat	aaataataaa	tagcggccgc	accatgaaag	3180
tgaaggggat	caggaagaat	tatcagcact	tgtggaaatg	gggcatcatg	ctccttggga	3240
tgttgatgat	ctgtagtgtc	gtagaaaatt	tgtgggtcac	agtttattat	ggggtacctg	3300
tgtggaaaga	agcaaccacc	actctatttt	gtgcatcaga	tgctaaagca	tatgatacag	3360
aggtacataa	tgtttgggccc	acacatgcct	gtgtacccac	agaccccaac	ccacaagaag	3420
tagtatttga	aaatgtgaca	gaaaaattta	acatgtggaa	aaataacatg	gtagaacaga	3480
tgcatgagga	tataatcagt	ttatgggata	aaagcctaaa	gccatgtgta	aaattaaccc	3540
cactctgtgt	tactttaaat	tgacttgatt	tgaggaatgt	tactaatatc	aataatagta	3600
gtgaggggat	gagaggagaa	ataaaaaact	gctctttcaa	tatcaccaca	agcataagag	3660
ataaggtgaa	gaaagactat	gcacttttct	atagacttga	tgtagtacca	atagataatg	3720
ataatactag	ctatagggtg	ataaattgta	atacctcaac	cattacacag	gcctgtccaa	3780
aggtatcctt	tgagccaatt	cccatacatt	attgtacccc	ggctgggtttt	gcgattctaa	3840
agtgtaaaaga	caagaagttc	aatggaacag	ggccatgttaa	aaatgtcagc	acagtacaat	3900
gtacacatgg	aattaggcca	gtagtgtcaa	ctcaactgct	gtttaaattggc	agtctagcag	3960
aagaagaggt	agtaattaga	tctagtaatt	tcacagacaa	tgcaaaaaaac	ataatagtac	4020
agttgaaaga	atctgtagaa	attaattgta	caagacccaa	caacaataca	aggaaaagta	4080
tacatatagg	accaggaaga	gcattttata	caacaggaga	aataatagga	gatataagac	4140
aagcacattg	caacattagt	agaacaaaat	ggaataaacac	tttaaatcaa	atagctacaa	4200
aattaaaaga	acaatttggg	aataataaaa	caatagtctt	taatcaatcc	tcaggagggg	4260
acccagaaat	tgtaatgcac	agtttttaatt	gtggagggga	attctttctac	tgtaattcaa	4320
cacaactgtt	taatagtact	tgggaatttta	atgggtacttg	gaattttaaca	caatcgaatg	4380
gtactgaagg	aaatgacact	atcacactcc	catgtagaat	aaaacaaatt	ataaatatgt	4440
ggcaggaagt	aggaaaagca	atgtatgccc	ctcccacag	aggacaaatt	agatgtctat	4500
caaataattac	agggtctaata	ttacaagag	atgggtggaac	taacagtagt	gggtccgaga	4560
tcttcagacc	tgggggagga	gatatgaggg	acaattggag	aagtgaatta	tataaatata	4620
aagtagtaaa	aattgaaacca	ttaggagtag	caccacacaa	ggcaaaaaga	agagtgggtg	4680
agagagaaaa	aagagcagtg	ggaacgatag	gagctatgtt	ccttgggttc	ttgggagcag	4740
caggaagcac	tatgggcgca	gcgtcaataa	cgctgacggt	acaggccaga	ctattattgt	4800
ctggtatagt	gcaacagcag	aacaatttgc	tgagggctat	tgaggcgcaa	cagcatctgt	4860
tgcaactcac	agtctggggc	atcaagcagc	tccaggcaag	agtcctggct	gtggaaagat	4920
acctaaggga	tcaacagctc	ctagggattt	gggggttgctc	tggaaaactc	atctgcacca	4980
ctgctgtgcc	ttggaatgct	agttggagta	ataaaaactct	ggatatgatt	tggtataaca	5040
tgacctggat	ggagtgggaa	agagaaatcg	aaaattacac	aggcttaata	tacaccttaa	5100
ttgaggaatc	gcagaaccaa	caagaaaaga	atgaacaaga	cttattagca	ttagataagt	5160
gggcaagttt	gtggaatttg	tttgacatat	caaattggct	gtggtatgta	aaaatcttca	5220
taatgatagt	aggaggcttg	ataggtttaa	gaatagtttt	tactgtactt	tctatagtaa	5280
atagagttag	gcagggtatac	tcaccattgt	catttcagac	ccacctcca	gccccgaggg	5340
gacccgacag	gcccgaagga	atcgaagaag	aaggtggaga	cagagactaa	tttttatgcg	5400
gccgctggta	cccaacctaa	aaattgaaaa	taaaatacaa	ggttcttgag	ggttgtgtta	5460
aattgaaagc	gagaaataat	cataaataag	cccggggatc	ctctagagtc	gacaccatgg	5520
gtgcgagagc	gtcagtatta	agcgggggag	aattagatcg	atgggaaaaa	attcgggttaa	5580
ggccaggggg	aaagaaaaaa	tataaattaa	aacatatagt	atgggcaagc	agggagctag	5640
aacgattccg	agttaatcct	ggcctgttag	aaacatcaga	aggctgtaga	caaatactgg	5700
gacagctaca	accatccctt	cagacaggat	cagaagaact	tagatcatta	tataatacag	5760
tagcaaccct	ctattgtgtg	catcaaaagg	tagagataaa	agacaccaag	gaagctttag	5820
acaagataga	ggaagagcaa	aacaaaagta	agaaaaaagc	acagcaagca	gcagctgaca	5880
caggacacag	caatcagggtc	agccaaaatt	accctatagt	gcagaacatc	caggggcaaa	5940
tggtacatca	ggccatatca	cctagaactt	taaatgcatg	ggtaaaaagta	gtagaagaga	6000
aggctttcag	cccagaagtg	ataccatgt	tttcagcatt	atcagaagga	gccacccac	6060
aagattttaa	caccatgcta	aacacagtgg	ggggatagag	agcagccatg	caaattgttaa	6120
aagagaccat	caatgaggaa	gctgcagaat	gggatagagt	gcatccagtg	catgcagggc	6180

NIH265.001VPC SEQLIST.TXT

ctattgcacc	aggccagatg	agagaaccaa	ggggaagtga	catagcagga	actactagta	6240
cccttcagga	acaaatagga	tggatgacaa	ataatccacc	tatcccagta	ggagaaatth	6300
ataaaagatg	gataatcctg	ggattaaata	aaatagtaag	aatgtatagc	cctaccagca	6360
ttctggacat	aagacaagga	ccaaaagaac	ccttttagaga	ctatgtagac	cggttctata	6420
aaactctaag	agccgagcaa	gcttcacagg	aggtaaaaaa	ttggatgaca	gaaaccttgt	6480
tgggtccaaa	tgcgaaccca	gattgtgaaga	ctatttttaa	agcattggga	ccagcggcta	6540
cactagaaga	aatgatgaca	gcatgtcagg	gagtaggagg	acccggccat	aaggcaagag	6600
ttttggctga	agcaatgagc	caagtaacaa	attcagctac	cataatgatg	cagagaggca	6660
atttttaggaa	ccaaagaaaag	attgttaagt	gtttcaattg	tggcaaaagaa	gggcacacag	6720
ccagaaattg	cagggccctt	aggaaaaagg	gctgttggaa	atgtggaaag	gaaggacacc	6780
aaatgaaaga	ttgtactgag	agacaggcta	atttttttagg	gaagatctgg	ccttcctaca	6840
aggggaaggcc	aggggaatth	cttcagagca	gaccagagcc	aacagcccca	ccagaagaga	6900
gcttcagggtc	tggggtagag	acaacaactc	cccctcagaa	gcaggagccg	atagacaagg	6960
aactgtatcc	tttaacttcc	ctcagatcac	tctttggcaa	cgacccctcg	tcacaataaa	7020
gatagggggg	caactaaagg	aagctctatt	agatacagga	gcagatgata	cagtattaga	7080
agaaatgagt	ttgccaggaa	gatggaaacc	aaaaatgata	gggggaattg	gaggttttat	7140
caaagtaaga	cagtatgatc	agatactcat	agaaatctgt	ggacataaag	ctataggtac	7200
agtattagta	ggacctacac	ctgtcaacat	aatttgaaga	aatctgttga	ctcagattgg	7260
ttgcacttta	aatttttcca	ttagccctat	tgagactgtg	ccagtaaaat	taaagccagg	7320
aatggatggc	ccaaaagtta	aacaatggcc	attgacagaa	gaaaaaataa	aagcattagt	7380
agaaatttgt	acagaaatgg	aaaaggaagg	gaaaatttca	aaaattgggc	ctgagaatcc	7440
atacaatact	ccagtatttg	ccataaagaa	aaaagacagt	actaaatgga	ggaaattagt	7500
agatttcaga	gaacttaata	agagaactca	agacttctgg	gaagttcaat	taggaatacc	7560
acatcccgcg	gggttaaaaa	agaaaaaatc	agtaacagta	ctggatgtgg	gtgatgcata	7620
tttttcagtt	cccttagatg	aagacttcag	gaagtatact	gcattttacca	tacctagtat	7680
aaacaatgag	acaccagggg	ttagatatca	gtacaatgtg	cttccacagg	gatggaaagg	7740
atcaccagca	atattccaaa	gtagcatgac	aaaaatctta	gagcctttta	aaaaacaaaa	7800
tccagacata	gttatctatc	aatacatgaa	cgatttgtat	gtaggatctg	acttagaaat	7860
agggcagcat	agaacaaaaa	tagaggagct	gagacaacat	ctgttgaggt	ggggacttac	7920
cacaccagac	aaaaaacatc	agaaagaacc	tccattcctt	tggatgggtt	atgaactcca	7980
tcctgataaa	tggacagtac	agcctatagt	gctgccagaa	aaagacagct	ggactgtcaa	8040
tgacatacag	aagtttagtg	ggaaattgaa	taccgcaagt	cagattttacc	cagggattaa	8100
agtaaggcaa	ttatgtaaac	tccttagagg	aaccaaagca	ctaacagaag	taataccact	8160
aacagaagaa	gcagagctag	aactggcaga	aaacagagag	attctaaaag	aaccagttaca	8220
tggagtgtat	tatgacctat	caaaagactt	aatagcagaa	atacagaagc	aggggcaagg	8280
ccaatggaca	tatcaaatth	atcaagagcc	atttataaat	ctgaaaacag	gaaaatatgc	8340
agaatgagg	ggtgcccaca	ctaattgatg	aaaacaatta	acagaggcag	tgcaaaaaat	8400
aaccacagaa	agcatagtaa	tatggggaaa	gactcctaaa	tttaaaactac	ccatacaaaa	8460
ggaaacatgg	gaaacatggt	ggacagagta	ttggcaagcc	acctggattc	ctgagtggga	8520
gtttgttaat	acccctcctt	tagtgaaatt	atggtaccag	ttagagaaag	aacccatagt	8580
aggagcagaa	actttctatg	tagatggggc	agctaaccag	gagactaaat	taggaaaagc	8640
aggatatgtt	actaacaag	gaagacaaaa	ggttgtcccc	ctaactaaca	caacaaatca	8700
gaaaactcag	ttacaagcaa	tttatctagc	tttgaggat	tcaggattag	aagtaaacat	8760
agtaacagac	tcacaatatg	cattaggaat	cattcaagca	caaccagata	aaagtgaatc	8820
agagttagtc	aatcaaataa	tagagcagtt	aataaaaaag	gaaaaggctc	atctggcatg	8880
ggtaccagca	cacaaaggaa	ttggaggaaa	tgaacaagta	gataaattag	tcagtgtctg	8940
aatcaggaaa	atactattth	tagatgggaa	caagatgaac	attagtthtt	taattttttt	9000
atgtcgacct	gcagggaaaag	ttttataggt	agttgataga	acaaaataca	taatttttga	9060
aaaataaatc	actttttata	ctaatatgac	acgattacca	atacttttgt	tactaatatc	9120
attagtatac	gctacacctt	ttcctcagac	atctaaaaaa	ataggtgatg	atgcaacttt	9180
caatgttaat	cgaaataata	caaatgacta	agtgcttatg	agtgtcttgt	ataaggagcc	9240
atattccatt	attctttttag	ctgctaaaag	cggtgtcttg	tatttttgata	attataccaa	9300
ggataaaaata	tcttacgact	ctccatacga	tgatctagtt	acaactatca	caattaaatc	9360
attgactgct	agagatgccg	gtactttatgt	atgtgcattc	tttatgacat	cgcctacaaa	9420
tgacactgat	aaagtagatt	atgaagaata	ctccacagag	ttgattgtaa	atacagatag	9480
tgaatcgact	atagacataa	tactatctgg	atctacacat	tcaccagaaa	ctagttaagc	9540
ttgtctccct	atagttagtc	gtatttaggc	ttggcgtaat	catggtcata	gctgtttctt	9600
gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccgggaag	cataaagtgt	9660
aaagcctggg	gtgcctaatt	agttagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	9720
gctttcgagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggcca	acgcgcgggg	9780
agaggcggtt	tgcgtatttg	gcgctcttcc	gcttctctgc	tcactgactc	gctgcgtctg	9840
gtcgttccgc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatagc	gttatccaca	9900
gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	ggcagggaac	9960
cgtaaaaagg	ccgcgtttgt	ggcgtttttc	gataggctcc	gccccctga	cgagcatcac	10020
aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	10080
tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatag	10140
ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	ctgtaggtat	10200
ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgacagcaac	ccccgttcag	10260

NIH265.001VPC SEQLIST.TXT

```

ccccgaccgct ggcgccttatt cggtaactat cgtcttgagt ccaacccgggt aagacacgac 10320
ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 10380
gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggag agtatttggt 10440
atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 10500
aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 10560
aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaaac 10620
gaaaactcac gttaagggat tttggtcatt agattatcaa aaaggatctt cacctagatc 10680
cttttaaatt aaaaatgaag ttttaaatac atctaaagta tataatgagta aacttggtct 10740
gacagttacc aatgcttaat cagtgaaggc cctatctcag cgatctgtct atttcgttca 10800
tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 10860
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 10920
ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 10980
atccagctta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 11040
cgcaacgttg ttggcattgc tacaggcatc gtggtgtcac gctcgtcgtt ttgttggttg 11100
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccatt gttgtgcaaa 11160
aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagtggc cgcagtgtta 11220
tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 11280
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 11340
agttgctctt gcccgcgctc aatacgggat aataccgcg cacaatagcag aactttaaaa 11400
gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 11460
agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 11520
accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa ggaataagg 11580
gcgacacgga aatgttgaaat actcatactc ttcctttttc aatattattg aagcatttat 11640
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 11700
ggggttccgc gcacatttcc cgaaaaagt ccacctgacg tctaagaaac cattattatc 11760
atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtctcgc gcgtttcggg 11820
gatgacggtg aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa 11880
gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag cgggtgttg cggtgtcgg 11940
ggctggctta actatgcggc atcagagcag attgtactga gagtgcacca tatgcggtgt 12000
gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggcgccattc gccattcagg 12060
ctgcgcaact gttgggaagg gcgatcgggt cgggcctctt cgctattacg ccagctggcg 12120
aaagggggat gtgctgcaag gcgattaagt tgggtaacgc caggggtttt ccagtcacga 12180
cgttgtaaaa cgacggccag tgaattggat ttaggtgaca ctata 12225

```

<210> 2
 <211> 74
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Psyn II promoter

<400> 2
 taaaaaatga aaaaatattc taatttatag gacgggtttg attttctttt tttctatgct 60
 ataaataata aata 74

<210> 3
 <211> 2214
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> ADA envelope truncated

<400> 3
 atgaaagtga aggggatcag gaagaattat cagcacttgt ggaaatgggg catcatgctc 60
 cttgggatgt tgatgatctg tagtgctgta gaaaatttgt gggtcacagt ttattatggg 120
 gtacctgtgt ggaaagaagc aaccaccact ctattttgtg catcagatgc taaagcatat 180
 gatacagagg tacataatgt ttgggccaca catgcctgtg taccacacaga cccaaccca 240
 caagaagtga tattggaaaa tgtgacagaa aattttaaca tgtggaaaaa taacatggta 300
 gaacagatgc atcagatat aatcagttta tgggatcaaa gcctaaagcc atgtgtaaaa 360
 ttaacccac tctgtgttac tttaaattgc actgatttga ggaatgttac taatatcaat 420
 aatagtagtg agggaatgag aggagaaata aaaaactgct ctttcaatat caccacaagc 480
 ataagagata aggtgaagaa agactatgca cttttctata gacttgatgt agtaccaata 540
 gataatgata atactagcta taggttgata aattgtaata cctcaaccat tacacaggcc 600
 tgtccaaagg tatcctttga gccaatccc atacattatt gtaccccgcc tggttttgcy 660

NIH265.001VPC SEQLIST.TXT

```

attctaaagt gtaaagacaa gaagttcaat ggaacagggc catgtaaaaa tgtcagcaca 720
gtacaatgta cacatggaat taggccagta gtgtcaactc aactgctgtt aaatggcagt 780
ctagcagaag aagaggtagt aattagatct agtaatttca cagacaatgc aaaaaacata 840
atagtacagt tgaaagaatc tgtagaaatt aattgtacaa gacccaacaa caatacaagg 900
aaaagtatac atataggacc aggaagagca ttttatacaa caggagaaat aataggagat 960
ataagacaag cacattgcaa cattagtaga acaaaatgga ataacacttt aaatcaaata 1020
gctacaaaat taaaagaaca atttgggaat aataaaacaa tagtctttta tcaatcctca 1080
ggagggggacc cagaaattgt aatgcacagt ttttaattgt gaggggaatt ctctactgt 1140
aattcaaacac aactgtttta tagtacttgg aatttttaatg gtacttggaa ttttaacacaa 1200
tcgaatggta ctgaaggaaa tgacactatc acactcccat gtagaataaa acaaattata 1260
aatatgtggc aggaagtagg aaaagcaatg tatgcccctc ccatcagagg acaaattaga 1320
tgctcatcaa atattacagg gctaataatta acaagagatg gtggaactaa cagtagtggg 1380
tccgagatct tcagacctgg gggaggagat atgagggaca attggagaag tgaattatat 1440
aaatataaag tagtaaaaat tgaaccatta ggagtagcac ccaccaaggc aaaaagaaga 1500
gtggtgcaga gagaaaaaag agcagtggga acgataggag ctatgttcct tgggttcttg 1560
ggagcagcag gaagcactat gggcgagcg tcaataacgc tgacggtaca ggccagacta 1620
ttattgtctg gtatagtgcac acagcagaac aatttgctga gggctattga ggcgcaacag 1680
catctgttgc aactcacagt ctggggcatc aagcagctcc aggcaagagt cctggctgtg 1740
gaaagatacc taagggatca acagctccta gggatttggg gttgctctgg aaactctatc 1800
tgcaccactg ctgtgccttg gaatgctagt tggagtaata aaactctgga tatgatttgg 1860
gataacatga cctggatgga gtgggaaaga gaaatcgaaa attacacagg cttaatatatac 1920
accttaattg aggaatcgca gaaccaacaa gaaaagaatg aacaagactt attagcatta 1980
gataagtggg caagtttgtg gaattggttt gacatatcaa attggctgtg gtatgtaaaa 2040
atcttcataa tgatagtagg aggcttgata ggtttaagaa tagtttttac tgtactttct 2100
atagtaataa gagttaggca gggatactca ccattgtcat ttcagaccca cctcccagcc 2160
ccgaggggac ccgacaggcc cgaaggaatc gaagaagaag gtggagacag agac 2214

```

<210> 4
<211> 70
<212> DNA
<213> Artificial sequence

<220>
<223> PmH5 promoter

<400> 4
aaaaattgaa aataaatata aaggttcttg agggttgtgt taaattgaaa gcgagaaata 60
atcataaata 70

<210> 5
<211> 3479
<212> DNA
<213> Artificial sequence

<220>
<223> HXB2 gag pol

<400> 5
atgggtgcga gagcgctcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 120
ctagaacgat tcgcagttta tcctggcctg ttagaacat cagaaggctg tagacaaata 180
ctgggacagc tacaaccatc ccttcagaca ggaatcagaag aacttagatc attatataat 240
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caaggaagct 300
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 360
gacacaggac acagcaatca ggtcagccaa aattacccta tagtgcagaa catccagggg 420
caaatggtag atcaggccat atcacctaga actttaaatg catgggtaaa agtagtagaa 480
gagaaggctt tcagcccaga agtgataccc atgttttcag cattatcaga aggagccacc 540
ccacaagatt taaacaccat gctaaacaca gtggggggac atcaagcagc catgcaaatg 600
ttaaaagaga ccatcaatga ggaagctgca gaattgggata gagtgcattc agtgcattga 660
gggcctattg caccaggcca gatgagagaa ccaaggggaa gtgacatagc aggaactact 720
agtacccttc aggaacaaat aggatggatg acaataatc cacctatccc agtaggagaa 780
atttataaaa gatgataat cctgggatta aataaaaatag taagaatgta tagccctacc 840
agcattctgg atataagaca aggaccctta gaacctatgt agaccgggtt 900
tataaaactc taagagccga gcaagcttca caggaggtaa aaaattggat gacagaaacc 960
ttgttggtcc aaaatgagaa cccagattgt aagactattt taaaagcatt gggaccagcg 1020
gctacactag aagaaatgat gacagcatgt cagggagtag gaggaccccg ccataaggca 1080
agagtttttg ctgaagcaat gagccaagta acaaatcag ctaccataat gatgcagaga 1140
ggcaatttta ggaaccaaaag aaagattgtt aagtgtttca attgtggcaa agaagggcac 1200

NIH265.001VPC SEQLIST.TXT

acagccagaa	attgcagggc	ccctaggaaa	aagggctgtt	ggaaatgtgg	aaaggaagga	1260
caccaaata	aagattgtac	tgagagacag	gctaattttt	tagggaagat	ctggccttcc	1320
tacaagggaa	ggccagggaa	ttttcttcag	agcagaccag	agccaacagc	cccaccagaa	1380
gagagcttca	gggtctgggt	agagacaaca	actccccctc	agaagcagga	gccgatagac	1440
aaggaactgt	atccttttaac	ttccctcaga	tcactctttg	gcaacgaccc	ctcgtcacaa	1500
taaagatagg	ggggcaacta	aaggaagctc	tattagatac	aggagcagat	gatacagtat	1560
tagaagaaat	gagtttgcca	ggaagatgga	aacccaaaat	gataggggga	attggagggtt	1620
ttatcaaagt	aagacagtat	gatcagatac	tcatagaaat	ctgtggacat	aaagctatag	1680
gtacagtatt	agtaggacct	acacctgtca	acataattgg	aagaaatctg	ttgactcaga	1740
ttggttgcac	tttaaatttt	cccattagcc	ctattgagac	tgtaccagta	aaattaaagc	1800
caggaatgga	tggcccaaaa	gttaaacaaat	ggccattgac	agaagaaaaa	ataaaagcat	1860
tagtagaaat	ttgtacagaa	atggaaaagg	aagggaaaat	ttcaaaaatt	gggcctgaga	1920
atccatacaa	tactccagta	tttgccataa	agaaaaaaga	cagtactaaa	tggaggaaat	1980
tagtagattt	cagagaactt	aataagagaa	ctcaagactt	ctgggaagt	caattaggaa	2040
taccacatcc	cgcaggggta	aaaaagaaaa	aatcagtaac	agtactggat	gtgggtgatg	2100
catatttttc	agttccctta	gatgaagact	tcaggaagta	tactgcattt	accataccta	2160
gtataaacaa	tgagacacca	gggattagat	atcagtacaa	tgtgcttcca	cagggatgga	2220
aaggatcacc	agcaatatct	caaagtagca	tgacaaaaat	cttagagcct	tttaaaaaac	2280
aaaatccaga	catagtattc	tatcaataca	tgaacgattt	gtatgtagga	tctgacttag	2340
aaatagggca	gcatagaaca	aaaatagagg	agctgagaca	acatctgttg	aggtggggag	2400
ttaccacacc	agacaaaaaa	catcagaaa	aacctccatt	cctttggatg	ggttatgaac	2460
tccatcctga	taaatggaca	gtacagccta	tagtgctgcc	agaaaaagac	agctggactg	2520
tcaatgacat	acagaagtta	gtggggaaat	tgaataccgc	aagtcagatt	taccagggga	2580
ttaaagtaag	gcaattatgt	aaactcctta	gaggaaccaa	agcactaaca	gaagtaatac	2640
cctaacaga	agaagcagag	ctagaactgg	cagaaaaacag	agagattcta	aaagaaccag	2700
tacatggagt	gtattatgac	ccatcaaaag	acttaatagc	agaaatacag	aagcaggggc	2760
aaggccaatg	gacatatcaa	atztatcaag	agccatttaa	aaatctgaaa	acaggaaaat	2820
atgcaagaat	gaggggtgcc	cacactaatg	atgtaaaaca	attaacagag	gcagtgcata	2880
aaataaccac	agaaagcata	gtaatatggg	gaaagactcc	taaattttaa	ctaccatac	2940
aaaaggaaac	atgggaaaca	tggtggacag	agtattggca	agccacctgg	attcctgagt	3000
gggagtttgt	taatacccct	cctttagtga	aattatggta	ccagtttagag	aaagaaccca	3060
tagtaggagc	agaaaccttc	tatgtagatg	gggcagctaa	cagggagact	aaattaggaa	3120
aagcaggata	tgttactaac	aaaggaagac	aaaagggtgt	ccccctaact	aacacaacaa	3180
atcagaaaa	tcagttacaa	gcaatttatc	tagctttgca	ggattcagga	ttagaagtaa	3240
acatagtaac	agactcaca	tatgcattag	gaatcattca	agcacaacca	gataaaagt	3300
aatcagagtt	agtcaatcaa	ataatagagc	agtttaataa	aaaggaaaag	gtctatctgg	3360
catgggtacc	agcacacaaa	ggaattggag	gaaatgaaca	agtagataaa	ttagtcaagt	3420
ctggaatcag	gaaaatacta	tttttagatg	gaatagataa	ggccaagat	gaacattag	3479

<210> 6

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Gag-CM9 peptide

<400> 6

Cys Thr Pro Tyr Asp Ile Asn Glu Met

1

5

<210> 7

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> ovalbumin peptide

<400> 7

Ser Ile Ile Asn Phe Glu Lys Leu

1

5

<210> 8

<211> 20

<212> DNA

NIH265.001VPC SEQLIST.TXT

<213> Artificial Sequence

<220>

<223> synthetic probe

<400> 8

ctgtctgcgt catttggtgc

20

<210> 9

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> endocytosis motif

<221> VARIANT

<222> (1)...(4)

<223> Xaa = Any Amino Acid

<221> VARIANT

<222> (1)...(4)

<223> Xaa = Any Amino Acid

<400> 9

Tyr Xaa Xaa Leu

1

<210> 10

<211> 93

<212> DNA

<213> Artificial Sequence

<220>

<223> m7.5 promoter

<400> 10

cgcttttttat agtaagtttt tcaccataa ataataaata caataattaa tttctcgtaa 60
aaattgaaaa actattctaa tttattgcac ggt 93

<210> 11

<211> 74

<212> DNA

<213> Artificial Sequence

<220>

<223> Psyn III promoter

<400> 11

taaaaattga aaaaatattc taatttatag gacggttttg attttctttt tttctatact 60
ataaataata aata 74

<210> 12

<211> 74

<212> DNA

<213> Artificial Sequence

<220>

<223> Psyn IV promoter

<400> 12

taaaaattga aaaactattc taatttatag gacggttttg attttctttt tttctatact 60
ataaataata aata 74

<210> 13

<211> 75

<212> DNA

NIH265.001VPC SEQLIST.TXT

<213> Artificial Sequence

<220>

<223> Psyn V promoter

<400> 13

aaaaaatgat aaagtaggtt cagttttatt gctggtttaa aatcacgctt tcgagtaaaa 60
 actacgaata taaat 75

<210> 14

<211> 7

<212> DNA

<213> Artificial Sequence

<220>

<223> early transcription termination signal

<221> misc_feature

<222> (1)...(7)

<223> n = A,T,C or G

<400> 14

ttttnt

7

<210> 15

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> 5' primer

<400> 15

gcgccccggg tcgacgcggc cgcgccatga gagtgagggg gatacagagg aac 53

<210> 16

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> 3' primer

<400> 16

gcgccccggg cggccgcaga aaaattagcc ttgctctcca ccttcttctt ctattcc 57

<210> 17

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> 5' primer

<400> 17

gcgccccggg tcgacgcggc cgcgccatga gagtgagggg gacagtgagg aattat 56

<210> 18

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> 3' primer

<400> 18

gcgccccggg cggccgcaga aaaattagcc ttgctctcca ccttcttctt ctattcc 57

NIH265.001VPC SEQLIST.TXT

<210> 19
 <211> 56
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 5' primer

 <400> 19
 gcgccccggg tcgacgcggc cgcgccatga gagtgagggg gatagagagg aattat 56

 <210> 20
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 3' primer

 <400> 20
 gcgccccggg cggccgcaga aaaattagcc ttgctctcca cttctcctt c 51

 <210> 21
 <211> 40
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 5' primer

 <400> 21
 gcgccccggg gccatgggtg cgagagcgtc agtattaagc 40

 <210> 22
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 3' primer

 <400> 22
 gcgccccggg agaaaaatta gaaggtttct gtcctacta tgggttcct 49

 <210> 23
 <211> 40
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 5' primer

 <400> 23
 gcgccccggg gccatgggtg cgagagcgtc agtgtaagt 40

 <210> 24
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 3' primer

 <400> 24
 gcgccccggg agaaaaatta gaaagtttct gtcctacta tgggttcct 49

 <210> 25
 <211> 30

NIH265.001VPC SEQLIST.TXT

```

<212> DNA
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 25
ggaatagaag aagaaggtgg agagcaaggc
30

<210> 26
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 26
Gly Ile Glu Glu Glu Gly Gly Glu Gln Gly
1          5          10

<210> 27
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 27
ggaatagaag aagaaggtgg agagcaaggc
30

<210> 28
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 28
Gly Ile Glu Glu Glu Gly Gly Glu Gln Gly
1          5          10

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 29
ggaacagaag gagaaggtgg agagcaaggc
30

<210> 30
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV env C-terminal truncation sequence

<400> 30
Gly Thr Glu Gly Glu Gly Gly Glu Gln Gly
1          5          10

```

NIH265.001VPC SEQLIST.TXT

<210> 31
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV gagpol C-terminal truncation sequence

<400> 31
aaggaaccca tagtaggagc agaaaccttc 30

<210> 32
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV gagpol C-terminal truncation sequence

<400> 32
Lys Glu Pro Ile Val Gly Ala Glu Thr Phe
1 5 10

<210> 33
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV gagpol C-terminal truncation sequence

<400> 33
aaggaaccca tagtaggagc agaaactttc 30

<210> 34
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV gagpol C-terminal truncation sequence

<400> 34
Lys Glu Pro Ile Val Gly Ala Glu Thr Phe
1 5 10

<210> 35
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> HIV envelope C-terminal truncation sequence

<400> 35
agaatcgaag gagaaggtgg agagcaagac 30

<210> 36
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> HIV envelope C-terminal truncation sequence

NIH265.001VPC SEQLIST.TXT

<400> 36
 Gly Arg Ile Glu Gly Glu Gly Gly Glu Gln Asp
 1 5 10

<210> 37
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 37
 agaatcgaag gagaagggtgg agagcaagac 30

<210> 38
 <211> 11
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 38
 Gly Arg Ile Glu Gly Glu Gly Gly Glu Gln Asp
 1 5 10

<210> 39
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 39
 agaatcgaag gagaagggtgg agagcaagac 30

<210> 40
 <211> 11
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 40
 Gly Arg Ile Glu Gly Glu Gly Gly Glu Gln Asp
 1 5 10

<210> 41
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 41
 agaatcgaag gagaagggtgg agagcaagac 30

<210> 42
 <211> 11
 <212> PRT

NIH265.001VPC SEQLIST.TXT

<213> Artificial Sequence

<220>

<223> HIV envelope C-terminal truncation sequence

<400> 42

Gly Arg Ile Glu Gly Glu Gly Glu Gln Asp
1 5 10

<210> 43

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV envelope C-terminal truncation sequence

<400> 43

gacatatcaa attggctgtg gtatataaga

30

<210> 44

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV envelope C-terminal truncation sequence

<400> 44

Asp Ile Ser Asn Trp Leu Trp Tyr Ile Arg
1 5 10

<210> 45

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV gagpol C-terminal truncation sequence

<400> 45

gaccccatag caggagcaga gactttc

27

<210> 46

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> HIV gagpol C-terminal truncation sequence

<400> 46

Lys Asp Pro Ile Ala Gly Ala Glu Thr Phe
1 5 10

<210> 47

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> HIV envelope C-terminal truncation sequence

<400> 47

ggaatcgaag aagaaggtgg agagcaagac

30

NIH265.001VPC SEQLIST.TXT

<210> 48
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV envelope C-terminal truncation sequence

<400> 48
 Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp
 1 5 10

<210> 49
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> HIV gagpol C-terminal truncation sequence

<400> 49
 aaagaaccca tagtaggagc agaaactttc t 31

<210> 50
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> HIV gagpol C-terminal truncation sequence

<400> 50
 Lys Glu Pro Ile Val Gly Ala Glu Thr Phe
 1 5 10

<210> 51
 <211> 3068
 <212> DNA
 <213> HIV-1

<400> 51
 atgggtgcga gagcgtcagt attaagcggg ggaaaattag atgaatggga aaaaattcgg 60
 ttacggccag ggggaaacaa aaaatataga ttaaaacatt tagtatgggc aagcagggag 120
 ctagaacgat ttgcacttaa tcctgggtctt ttagaaacat cagaaggctg tagacaaata 180
 atagaacagc tacaaccatc tattcagaca ggatcagagg aacttaaattc attacataat 240
 acagtagtaa ccctctattg tgtacatgaa aggataaagg tagcagatac caaggaagct 300
 ttagataaga taaaggaaga acaaaccaaa agtaagaaaa aagcacagca agcaacagct 360
 gacagcagcc aggtcagcca aaattatcct atagtacaaa acctacaggg gcaaattggt 420
 caccagtcct tatcacctag gactttgaat gcatgggtta aagtaataga agagaaggct 480
 ttcagcccag aagtaatacc catgttttca gcattatcag aaggagccac cccaacagat 540
 ttaaacacca tgctaaacac agtgggggga catcaagcag ccatgcaaat gttaaaagag 600
 actatcaatg aggaagctgc agaatgggat aggctacatc cagtgcctgc agggcctgtt 660
 gcaccaggcc aaatgagaga accaagggga agtgatatag caggaactac cagtaccctt 720
 caggaacaaa taggatggat gacaagcaat ccacctatcc cagtaggaga aatctataaa 780
 agatggataa tcctaggatt aaataaaata gtaagaatgt atagccctgt cagcattttg 840
 gacataagac aaggaccaa ggaacccttt agagactatg tagatcgggt ctataaaact 900
 ctacgagccg agcaagcttc acaggatgta aaaaatttga tgactgaaac cttgttagtc 960
 caaaatgcga atccagattg taaaactatc ttaaaagcat tgggaccagc ggctacatta 1020
 gaagaaatga tgacagcatg tcagggagtg gggggaccga gtcataaagc aagagttttg 1080
 gctgaggcaa tgagccaagc atcaaacaca aatgctgtta taatgatgca gaggggcaat 1140
 ttcaagggca agaaaatcat taagtgtttc aactgtggca aagaaggaca cctagcaaaa 1200
 aattgtaggg ctctaggaa aagaggctgt tggaaatgtg gaaaggaagg gcaccaaattg 1260
 aaagattgta atgaaagaca ggctaatttt ttagggagaa tttggccttc ccacaagggg 1320
 aggccaggga atttccttca gagcagacca gagccaacag cccaccagc agagagcttc 1380
 gggtttgggg aagagataac accctcccag aaacaggagg ggaaagagga gctgtatcct 1440

NIH265.001VPC SEQLIST.TXT

tcagcctccc	tcaaatacact	ctttggcaac	gacccctagt	cacaataaaaa	ataggggggac	1500
agctaaagga	agctctatta	gatacaggag	cagatgatac	agtagtagaa	gaaatgaatt	1560
tgccaggaaa	atggaaacca	aaaatgatag	ggggaattgg	gggctttatc	aaagtaagac	1620
agtatgatca	aatactcgta	gaaatctatg	gatataaggc	tacagggtaca	gtatttagtag	1680
gacctacacc	tgtcaacata	attggaagaa	atttgttgac	tcagatttgt	tgcactttta	1740
atcccccaat	tagtcctatt	gaaactgtac	cagtaaaatt	aaagtacagg	atggatgggtc	1800
caagagttaa	acaatggcca	ttgacagaag	agaaaaataa	agcactaata	gaaatttgta	1860
cagaaattga	aaaggaagga	aaactttcaa	gaattggacc	tgaaaaacca	tacaatactc	1920
caatatattgc	cataaagaaa	aaagacagta	ctaagtggag	aaaattagta	gatttcagag	1980
aacttaataa	gagaactcaa	gatttctggg	aagttcaact	aggaatacca	catcctgcag	2040
ggctaaaaaa	gaaaaaatca	gtaacagtac	tggaggtggg	tgatgcatat	ttttcagttc	2100
ccttatatga	agacttttaga	aaatacactg	cattcaccat	acctagtata	aacaatgaga	2160
caccaggaat	tagatatcag	tacaatgtgc	ttccacaagg	atggaaagga	tcaccggcaa	2220
tattccaaag	tagcatgaca	aaaatttttag	aaccttttag	aaaacaaaat	ccagaagtgg	2280
ttatctacca	atacatgcac	gatttgtatg	taggattctga	cttagaaaata	gggcagcata	2340
gaataaaaaat	agaggaatta	aggggacacc	tattgaagtg	gggatttacc	acaccagaca	2400
aaaatcatca	gaaggaacct	ccatttcttt	ggatgggtta	tgaactccat	cctgataaat	2460
ggacagtaca	gcctataaaa	ctgccagaaa	aagaaagctg	gactgtcaat	gatctgcaga	2520
agtttagtggg	gaaattaaat	tgggcaagtc	aaattttatt	aggaattaaa	gtaagacaat	2580
tagtcaaatg	ccttagggga	accaaagcac	tgacagaagt	agtaccactg	acagaagaag	2640
cagaattaga	actggcagaa	aacagggaac	ttctaaaaga	aacagtacat	ggagtgtatt	2700
atgacccatc	aaaagactta	atagcagaaa	tacagaaaca	agggcaagac	caatggacat	2760
atcaaattta	tcaagaacaa	tataaaaatt	tgaaaacagg	aaagtatgca	aagaggagga	2820
gtacccacac	taatgatgta	aaacaattaa	cagaggcagt	gcaaaaaata	gcccaagaat	2880
gtatagtgtg	atggggaaaag	actcctaatt	tcgactacc	catacaaaag	gaaacatggg	2940
aaacatgggtg	gacagagtat	tggcaggcca	cctggattcc	tgagtggggag	tttgtcaata	3000
cccctccctt	ggtttaaatta	tggtagcagt	tagagaagga	acccatagta	ggagcagaaa	3060
ccttctaa						3068

<210> 52
 <211> 3080
 <212> DNA
 <213> HIV-1

<400> 52						
atgggtgcga	gagcgtcagt	gttaagtggg	ggaaaattag	atgaatggga	aagaattcgg	60
ttacgggccag	ggggaaacaa	aagatataaa	ctaaaacata	tagtatgggc	aagcaggggag	120
ctagagcgat	ttgcacttaa	tcctggcctt	ttagaaacat	cagaaggctg	taaacaaata	180
ttgggacagc	tacaaccagc	tattcagaca	ggatcagaag	aacttaaattc	attatataat	240
acagtagcaa	ccctctattg	tgtacatgag	aggctaaagg	taacagacac	caaggaagct	300
ttagacaaaa	tagaggaaga	acaaacccaa	agtaagaaaa	aagcacagca	agcaacagct	360
gacacaaaaa	acagcagcca	ggtcagccaa	aattatccta	tagtacaaaa	cctacagggg	420
caaatgggtac	accaggctat	atcacctaga	acgttgaacg	catgggtaaa	agtaatagag	480
gagaaggctt	tcagcccaga	agtaataccc	atgttttcag	cattatcaga	aggagccacc	540
ccacaagatt	taaacaccat	gctaaacaca	gtggggggac	atcaggcagc	catgcagatg	600
ttaaaagaga	ccatcaatga	ggaagctgca	gaatgggata	ggttacatcc	agtagatgca	660
gggcctattg	caccaggaca	aatgagagaa	ccaacaggaa	gtgatataag	aggaactact	720
agtacccttc	aggaacaaat	aggatggatg	accagcaatc	cacctatccc	agtaggagaa	780
atctataaaa	gatggataat	cctaggatta	aataaaatag	taaggatgta	tagccctgtc	840
agtatttttg	acataaaaca	agggccaaag	gaacccttta	gagactatgt	agatcggttc	900
tataaaactc	taagggccga	gcaagcttca	caggaggtaa	aaggttggat	gaccgaaacc	960
ttgttggtcc	aaaatgcaaa	cccagattgt	aaaaccatct	taaaagcatt	gggacacagc	1020
gctacattag	aagaaatgat	gacagcatgt	cagggaagtgg	ggggaccggg	tcataaagca	1080
agagttttgg	ctgaggcaat	gagtcaagtc	tcaacaaata	ctgctataat	gatgcagaga	1140
ggcaatttta	agggcccaaa	gaaaagcatt	aagtgtttta	actgtggcaa	agaaggtcac	1200
acagcaaaaa	actgtagagc	tcctaggaaa	aggggctgtt	ggaaatgttg	aaggggaagga	1260
catcaaatga	aagattgcac	tgaaagacag	gctaattttt	tagggaaaaat	ttggcctttc	1320
cacaagggaa	ggccagggaa	tttccttcag	aacagaccag	agccaacagc	cccaccagaa	1380
gaaagcttcg	ggtttgggga	agagataaca	ccctctcaga	aacaggagaa	gaaggacaag	1440
gagctgtatc	ctgtagcttc	cctcaaatca	ctctttggca	acgacccctt	gtcacataaa	1500
agataggggg	acagctaaag	gaagctctac	tagatacagg	agcagatgat	acagtattag	1560
aagaaataaa	tttgccagga	aaatggaaac	caaaaatgat	aggggggaatt	ggaggcttta	1620
tcaaagtaag	acagtatgag	caataatctg	tggaacagaa	gctataggta	gctataggta	1680
cagtattagt	agggcctaca	cctgtcaaca	taatttggaa	aaatttgttg	actcagattg	1740
gttgcacttt	aaattttcca	attagcccta	ttgaaactgt	accagtaaaa	ttaaagccag	1800
ggatggacgg	tccaaaagtt	aaacaatggc	cattgacaga	agaaaaggta	aaagcactaa	1860
tagaaatttg	tacagaaatg	gaaaagggaag	gaaaaatttc	agaatttgga	cctgaaaatc	1920
catacaatac	tccaatattt	gccataaaga	aaaaggacag	tactaagtgg	agaaaaattag	1980

NIH265.001VPC SEQLIST.TXT

tagattttcag	ggaacttaat	aagagaactc	aagactttctg	ggaagttcaa	ctaggaatac	2040
cacatcctgc	ggggctaaaa	aagaaaaaat	cagtaacagt	actggagggtg	ggtgatgcat	2100
atttttcagt	tcccttatat	gaagatttta	gaaaatatac	tgcatccacc	atacctagta	2160
taaacaatga	aacaccagga	attagatatc	agtacaatgt	gcttccacaa	gggtggaaag	2220
gatcaccagc	aatattccaa	agtagcatga	caaaaatctt	agaacctttt	agaaaacaaa	2280
atccagaaat	ggttatctat	caatacatgc	acgatttgta	tgtaggatct	gacttagaaa	2340
tagggcagca	tagaataaaa	atagaagaat	taaggggaca	cctggtgaag	tggggattta	2400
ccacaccaga	caaaaagcat	cagaaagaac	ctccatttct	ttggatgggt	tatgaactcc	2460
atcctgataa	atggacagta	cagtctataa	aactgccaga	acaagaaagc	tggactgtca	2520
atgatataca	gaagttagtg	ggaaaattaa	attgggcaag	ccagatttat	ccaggaatta	2580
aggtaagaca	attatgcaaa	tgcattagggt	gtaccaaagc	actgacagaa	gtagtaccac	2640
tgacagaaga	agcagaatta	gaactggcag	aaaacaggga	aattctaaga	gaaccagtac	2700
atggagtgtg	ttatgaccca	tcaaaagact	taatagcaga	gatacagaaa	caagggcaag	2760
accagtggac	ataccaaatt	tatcaagaac	aatataaaaa	tctgaaaaca	ggaaagtatg	2820
caaaagttag	gggtaccac	actaatgatg	taaaacaatt	aacagaggca	gtacaaaaaa	2880
taaccaaga	atgtatagtg	atatggggaa	agcctcctaa	atttagacta	cccatacaaa	2940
aagaaacatg	ggaaatatgg	tggacagagt	attggcaggc	cacctggatt	cctgagtggg	3000
agtttgtcaa	taccctctct	ttagttaaat	tatggtacca	attagagaag	gaaccatag	3060
taggagcaga	aactttctaa					3080

<210> 53
 <211> 2181
 <212> DNA
 <213> HIV-1

<400> 53						
atgagagtga	gggggataca	gaggaactat	caaaacttgt	ggagatgggg	caccttgctc	60
cttgggatgt	tgatgatatg	taaggctaca	gaacagttgt	gggtcacagt	ttactatggg	120
gtacctgtgt	ggaaagaagc	aaccactact	ctattttgtg	catcagatgc	taaatcatat	180
aaagaagaag	cacataatat	ctgggctaca	catgcctgtg	taccaacaga	ccccaaccca	240
cgagaattaa	taatagaaaa	tgtcacagaa	aactttaaca	tgtggaaaaa	taacatgggtg	300
gagcagatgc	atgaggatat	aatcagttta	tgggatcaaa	gcctaaaacc	atgtgtaaaa	360
ttaacccac	tctgtgtcac	tttaaaactgc	actgaatgga	ggaagaataa	cactatcaat	420
gccaccagaa	tagaaatgaa	aaactgtctct	ttcaatctaa	ccacagaaat	aagagatagg	480
aaaaagcaag	tcagatgcact	tttctataaa	cttgatgtgg	taccaataga	tgataataat	540
agtactaata	ccagcttatag	gttaataaat	tgtaatacct	cagccattac	acaggcgtgt	600
ccaaaggtaa	cctttgagcc	aattcccata	cattattgtg	ccccagctgg	atatgcgatt	660
ctaaaatgta	acaataagaa	gttcaatggg	acaggtccat	gcgataatgt	cagtacagta	720
cagtgtacac	atggaattag	gccagtagta	tccactcaat	tgttgttgaa	tggcagctca	780
gcagaagaag	acataataat	tagatctgag	aatctcacaa	ataatgctaa	aatcataata	840
gtacagctta	atgagctctgt	aacaattaat	tgacaaaggc	cctacaacaa	tacaagaaga	900
ggtgtacata	taggaccagg	gagagcatac	tatacaacag	acataatagg	agatataaga	960
caagcacatt	gtaacattag	tggagcagaa	tggaaataaga	ctttacatcg	ggtagctaaa	1020
aaattaagag	acctatttaa	aaagacaaca	ataattttta	aaccgtcctc	cggagggggac	1080
ccagaaatta	caacacacag	ctttaattgt	agaggggaat	tcttctactg	caatacaaca	1140
agactgttta	atagcatatg	gggaaataat	agtacaggag	ttgatgagag	tataacactc	1200
ccatgcagaa	taaaacaaat	tataaacatg	tggcaggagg	taggaaaagc	aatgtatgcc	1260
cctcccatgt	aaggactaat	cagctgctca	tcaaataatta	caggattact	gttgacaaga	1320
gatggtggtg	gaagtaacag	tagtcagaat	gagaccttca	gacctggagg	gggagatatg	1380
agagacaatt	ggagaagtga	attatataaa	tataaagtag	taagaattga	accattaggt	1440
ctagcacctc	ccaaggcaaa	aagaagagta	gtagaaagag	agaaaagagc	aataggacta	1500
ggagctatgt	tcccttgggtt	cttggggagca	gcagaaagca	cgatgggagc	agcgctactg	1560
acgctgacgg	tacaggccag	acagctattg	tctggtatag	tgcaacagca	aaacaatttg	1620
ctgaaggcta	tagaggcgca	acagcacctg	ttgcaactca	cagtctgggg	cgttaaacag	1680
ctccaggcaa	gagtcctggc	tgtggaaagc	tacctaaagg	atcaacagct	cctaggaatt	1740
tggggttgct	ctggaaaaca	catttgcacc	accaatgtgc	cctggaactc	tagctggagt	1800
aataaaaactc	taaaatcaat	ttgggataac	atgacctgga	tggagtggga	aagagaaatt	1860
gacaattaca	cagggataat	atacaattta	cttgaagaat	cgcaaaccga	gcaagaaaga	1920
aatgaacaag	acctattgaa	attggaccaa	tgggcaagtt	tgtggaattg	gttttagcata	1980
acaaaatggc	tgtggtatat	aaaaatatat	ataatgatag	taggaggctt	gataggctta	2040
aggatagttt	ttgtctgtgt	ttctatagta	aatagagtta	ggcagggata	ttcacctctg	2100
tcgtttcaga	ccctcctccc	agccccgcgg	ggacccgaca	ggcccgaagg	aatagaagaa	2160
gaagggtggag	agcaaggcta	a				2181

<210> 54
 <211> 2214
 <212> DNA
 <213> HIV-1

NIH265.001VPC SEQLIST.TXT

<400> 54
 atgagagtga gggagacagt gaggaattat cagcacttgt ggagatgggg catcatgctc 60
 cttgggatgt taatgatatg tagtgctgca gaccagctgt gggcacagt gtattatggg 120
 gtacctgtgt ggaagaagc aaccactact ctattttgtg catcagatgc taaagcacat 180
 aaagcagagg cacataatat ctgggctaca catgcctgtg taccaacaga cccaatcca 240
 cgagaaataa tactaggaaa tgtcacagaa aactttaaca tgtggaagaa taacatggta 300
 gagcagatgc atgaggatat aatcagttta tgggatcaaa gtctaaaacc atgtgtaaaa 360
 ttaacccac tctgtgttac tttaaactgc actacatat ggaatggaac ttacagggg 420
 aatgaaacta aagggaagaa tagaagtgc ataatgacat gctctttcaa tataaccaca 480
 gaaataagag gtagaaagaa gcaagaaact gcacttttct ataaacttga tgtggtacca 540
 ctagaggata aggatagtaa taagactacc aactatagca gctatagatt aataaattgc 600
 aatacctcag tcgtgacaca ggcgtgtcca aaagtaacct ttgagccaat tcccatatc 660
 tattgtgccc cagctggatt tgcgattctg aaatgtaata ataagacgtt caatggaacg 720
 ggtccatgca aaaatgtcag cacagtacag tgtacacatg gaattaggcc agtagtgtca 780
 actcaactgt tgttgaatgg cagtctagca gaagaagaga taataattag atctgaaaat 840
 atcacaaata atgcaaaaac cataatagta cagcttaatg agtctgtaac aattgattgc 900
 ataaggccca acaacaatac aagaaaaagt atacgcatag gaccagggca agcactctat 960
 acaacagaca taataggaa tataagacaa gcacattgta atgttagtaa agtaaaatgg 1020
 ggaagaatgt taataagggt agctgaaaaa ttcttaacca gacaaagaac gacaaagaac 1080
 ataacttttg aaccatcctc agggggggac ccagaaatta caacacacag ctttaattgt 1140
 ggaggggaat tcttctactg caatacatca ggactattta atgggagtct gcttaattgag 1200
 cagtttaatg agacatcaaa tgatactctc acactccaat gcagaataaa acaaattata 1260
 aacatgtggc aaggagtagg aaaagcaatg tatgcccctc ccattgcagg accaatcagc 1320
 tgttcatcaa atattacagg actattgttg acaagagatg gtggtaatac tggtaattgat 1380
 tcagagatct tcagacctgg agggggagat atgagagaca attggagaag tgaattatac 1440
 aaatataaag tagtaagaat tgaaccaatg ggtctagcac ccaccagggc aaaaagaaga 1500
 gtggtggaaa gagaaaaaag agcaatagga ctgggagcta tgttccttgg gttcttggga 1560
 gcggcaggaa gcacgatggg cgacgctga ctgacgctga cggtagaggc cagacagtta 1620
 ttgtctggta tagtgcaaca gcaaaaacta ttgctgagag ctatagaggc gcaacagcat 1680
 ctgttgcaac tcacagtctg gggcattaaa cagctccagg caagagtcct ggctatggaa 1740
 agctacctaa aggatcaaca gctcctagga atttgggggt gctctggaaa acacatttgc 1800
 accactactg tgccctggaa ctctacctgg agtaatagat ctgtagagga gatttggaa 1860
 aatattgacct ggatgcagtg ggaagagaa attgagaatt acacaggttt aatatacacc 1920
 ttaattgaag aatcgcaaac ccagcaagaa aagaatgaac agaactatt gcaattggat 1980
 aaatgggcaa gtttgggaa ttggtttagt ataacaaaat ggctgtggta tataaaaaata 2040
 ttcataatga tagtaggagg cttaataggt ttaagaatag ttttggctgt gctttcttta 2100
 gtaaatagag ttaggcaggg atattcacct ctgtcttttc agaccctcct cccagccccg 2160
 aggggacccg acaggcccga aggaatagaa gaagaagggt gagagcaagg ctaa 2214

<210> 55
 <211> 2265
 <212> DNA
 <213> HIV-1

<400> 55
 atgagagtga gggggataga gaggaattat cagcacttat ggtggagatg gggcaccatg 60
 ctccttggga tattgatgat atgtagtgtc gcagaacaat tgtgggtcac agttttattat 120
 ggggtacctg tgtggaaaga agcaaccact actctatttt gtgcatcaga tgctaaagca 180
 tataaagcag aggcacacaa tatctgggct acacatgcct gtgtaccaac agacccaac 240
 ccacaagaaa tagtactaga aaatgtcaga gaaaacttta acatgtggaa aatagcatg 300
 gtggagcaga tgcattgagg ttatgggatc aaagcctaaa accatgtgtg 360
 aaattaaccc cactctgtgt cacttttaaac tgcactaatg ccactgccac taatgccact 420
 gccactagtc aaaatagcac tgatggtagt aataaaaactg ttaacacaga cacaggaatg 480
 aaaaactgct ctttcaatgt aaccacagat ctaaaagata agaagaggca agactatgca 540
 cttttctata aacttgatgt ggtacgaata gatgataaga ataccaatgg tactaatacc 600
 aactatagat taataaattg taataacctc gccattacac aagcgtgtcc aaagataacc 660
 tttgagccaa tttccataca ttattgtgcc ccagctggat atgcgattct aaaatgtaat 720
 aataagacat tcaatgggac ggggtccatgc aaaaacgtca gcacagtaca gtgtacacat 780
 gggattaggc cagtagtgtc aactcaactg ttgttgaatg gcagtctagc agaggaagag 840
 atagtaatta gatctgaaaa cctcacaaat aatgctaaaa ttataatagt acagcttaat 900
 gaagctgtaa caattaattg cacaagacc tccaacaata caagacgaag tgtacatata 960
 ggaccagggc aagcaatcta ttcaacagga taacaatag gagatataag aaaagcacat 1020
 tgtaatatata gtagaaaaga atggaatagc accttacaac aggttaactaa aaaattagga 1080
 agcctgttta acacaacaaa aataattttt aatgcatcct cgggagggga cccagaaatt 1140
 acaacacaca gctttaattg taacggggaa ttcttctact gcaatacagc aggactgttt 1200
 aatagtagat ggaacaggac aaatagttaa tggataaata gtaaatggac aaataagaca 1260
 gaagatgtaa atatcacact tcaatgcaga ataaaacaaa ttataaacat gtggcagggg 1320

NIH265.001VPC SEQLIST.TXT

gtaggaaaag	caatgtatgc	ccctcccgtt	agtggaaata	tccgatgttc	atcaaatatt	1380
acaggactgt	tgctgacaag	agatgggtgt	gggtgcagata	ataataggca	gaatgagacc	1440
ttcagacctg	ggggaggaga	tatgagagac	aattggagaa	gtgaattata	caaatataaa	1500
gtagtaagaa	ttgaaccact	aggtatagca	cccaccaagg	caaggagaag	agtgggtgaa	1560
agagaaaaaa	gagcaatagg	actgggagcc	ttgttccttg	ggttcttggg	aacagcagga	1620
agcacgatgg	gcgcagtgct	aatgacgctg	acggtacagg	ccagacaagt	attgtctggg	1680
atagtgaac	agcaaaacaa	tctgctgagg	gctatagagg	cgcaacagca	tctgttgcaa	1740
ctcacagtct	ggggcattaa	acagctccag	gcaagaatcc	tggctgtgga	aagctaccta	1800
aaggatcaac	agctcctagg	aatttggggg	tgctctggaa	aacacatttg	caccactaat	1860
gtgccctgga	actctagctg	gagtaataaa	tctctaaatt	atatttggaa	taacatgacc	1920
tggatggagt	gggaaaagga	aattgacaat	tacacagaat	taatatacag	cttaattgaa	1980
gtatcgcaaa	tccagcaaga	aaagaatgaa	caagaactat	tgaaattgga	cagttgggca	2040
agtttgtgga	attggtttag	cataacaaaa	tggctgtggg	atataaaaa	attcataatg	2100
atagtaggag	gcttgatagg	cttaagaata	gtttttgctg	tgctttcttt	agtaaataga	2160
gttaggcagg	gatactcacc	tctgtcgttt	cgagaccctt	tcccagcctc	gaggggaccc	2220
gacaggcccc	aaggaacaga	aggagaaggt	ggagagcaag	gctaa		2265

<210> 56

<211> 5326

<212> DNA

<213> Artificial Sequence

<220>

<223> MVA shuttle plasmid pLAS-1

<400> 56

gaattcggtg	gtggctcgcca	tggatgggtg	tattgtatac	tgtctaaacg	cgttagtaaa	60
acatggcgag	gaaataaatc	atataaaaaa	tgatttcattg	attaaacat	gttgtgaaaa	120
agtcaagaac	gttcacattg	gcggacaatc	taaaaacaat	acagtgattg	cagatttgcc	180
atatatggat	aatgcgggat	ccgatgtatg	caattcactg	tataaaaaga	atgtatcaag	240
aatatccaga	tttgctaatt	tgataaagat	agatgacgat	gacaagactc	ctactgggtg	300
atataattat	tttaaaccct	aagatgccat	tctgtttatt	atatccatag	gaaaggatag	360
agatgtttgt	gaactattaa	tctcatctga	taaagcgtgt	gcgtgtatag	agttaaattc	420
atataaagta	gccattcttc	ccatggatgt	ttcctttttt	accaaaggaa	atgcatcatt	480
gattattctc	ctgtttgatt	tctctatcga	tgccggcacct	ctcttaagaa	gtgtaaccga	540
taataatggt	attatatcta	gacaccagcg	tctacatgac	gagcttccga	gttccaattg	600
gttcaagttt	tacataagta	taaagtccga	ctattgttct	atattatata	tgggtgttga	660
tggatctgtg	atgcatgcaa	tagctgataa	tagaacttac	gcaaataatta	gcaaaaatat	720
attagacaat	actacaatta	acgatgagtg	tagatgctgt	tattttgaac	cacagattag	780
gattcttgat	agagatgaga	tgctcaatgg	atcatcgtgt	gatatgaaca	gacattgtat	840
tatgatgaat	ttacctgatg	taggcgaatt	tggatctagt	atgttggggg	aatatgaacc	900
tgacatgatt	aagattgctc	tttcgggtggc	tgggtaccag	gcgcgccttt	cattttgttt	960
ttttctatgc	tataaatggt	gagcaagggc	gaggagctgt	tcaccggggg	ggtgcccatc	1020
ctggctcgagc	tggacggcga	cgtaaacggc	cacaagttca	gcgtgtccgg	cgagggcgag	1080
ggcgatgcca	cctacggcaa	gctgaccctg	aagttcatct	gcaccaccgg	caagctgccc	1140
gtgccctggc	ccaccctcgt	gaccaccctg	acctacggcg	tgcatgtcct	cagccgctac	1200
cccgcaccaca	tcgaagcaga	cgacttcttc	aagtccgcca	tgcccgaagg	ctactgccag	1260
gagcgcacca	tcttcttcaa	ggacgacggc	aactacaaga	cccgcgcgca	ggtgaagttc	1320
gagggcgaca	ccctggtgaa	ccgcatcgag	ctgaagggca	tcgacttcaa	ggaggacggc	1380
aacatcctgg	ggcacaagct	ggagtacaac	tacaacagcc	acaacgtcta	tatcatggcc	1440
gacaagcaga	agaacggcat	caaggtgaa	ttcaagatcc	gccacaacat	cgaggacggc	1500
agcgtgcagc	tcgccgacca	ctaccgagac	aacaccccca	tcggcgacag	ccccgtgctg	1560
ctgcccgcaca	accactacct	gagcaccag	tccgccctga	gcaaagaccc	caacgagaag	1620
cgcgatcaca	tggctcctgt	ggagtctgtg	accgccgcgg	ggatcactct	cggcatgcac	1680
gagctgtaca	agtaagagct	cggttggtga	tggatctgtg	atgcatgcaa	tagctgataa	1740
tagaacttac	gcaaatatta	gcaaaaatat	attagacaat	actacaatta	acgatgagtg	1800
tagatgctgt	tattttgaac	cacagattag	gattcttgat	agagatgaga	tgctcaatgg	1860
atcatcgtgt	gatattgaaca	gacattgtat	tatgatgaat	ttacctgatg	taggcgaatt	1920
tggatctagt	atgttggggg	aatatgaacc	tgacatgatt	aagattgctc	tttcgggtggc	1980
tggcggcccc	ctcgaggccg	ctgggtaccca	acctaaaaat	tgaaaataaa	tacaaagggt	2040
cttgagggtt	gtgttaaat	gaaagcgaga	aataatcata	aataagcccc	gggatcctct	2100
agagtcgacc	tgcagggaaa	gttttatagg	tagttgatag	aacaaaatac	ataattttgt	2160
aaaaataaat	cactttttat	actaatatga	cacgattacc	aatacttttg	ttactaatat	2220
cattagtata	cgctacacct	tttcctcaga	catctaaaaa	aataggtgat	gatgcaactt	2280
tatcatgtaa	tcgaaataat	acaaatgact	acgttggtat	gagtgcttgg	tataaggagc	2340
ccaattccat	tattctttta	gctgtcaaaa	gcgacgtctt	gtattttgat	aattatacca	2400
aggataaaat	atcttacgac	tctccatacg	atgatctagt	tacaactatc	acaattaaat	2460
cattgactgc	tagagatgcc	ggtacttatg	tatgtgcatt	ctttatgaca	tcgcctacaa	2520

NIH265.001VPC SEQLIST.TXT

atgacactga	taaagtagat	tatgaagaat	actccacaga	gttgattgta	aatacacagata	2580
gtgaactcgac	tatagacata	atactatctg	gatctacaca	ttcaccagaa	actagttaag	2640
cttgtctccc	tatagttagt	cgtatttagag	cttggcgtaa	tcattggtcat	agctgtttcc	2700
tgtgtgaaat	tgttatccgc	tcacaattcc	acacaacata	cgagccggaa	gcataaagt	2760
taaagcctgg	ggtgccta	gagttagcta	actcacatta	attgctgtgc	gctcactgcc	2820
cgctttcgag	tcgggaaacc	tgctgtgcca	gctgcattaa	tgaatcggcc	aacgcgcggg	2880
gagaggcgg	ttgcgtattg	ggcgctcttc	cgcttctctg	ctcactgact	cgctgcgctc	2940
ggctcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaa	gcggtaatac	gggtatccac	3000
agaatcagg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	3060
ccgtaaaaag	gccgcgttgc	tggcgTTTT	cgataggctc	cgccccctg	acgagcatca	3120
caaaaatcga	cgctcaagtc	agagggtggc	aaaccgcgac	ggactataaa	gataccaggc	3180
gtttccccct	ggaagctccc	tcgtgcgctc	tcctgtttcc	accctgccc	ttaccggata	3240
cctgtccgc	tttctccctt	cggaagcgt	ggcgctttct	catagctcac	gctgtaggta	3300
tctcagttcg	gtgtaggctg	ttcgctccaa	gctgggctgt	gtgcacgaac	ccccggttca	3360
gcccgaaccg	tgcgcttat	ccggtacta	tcgtcttgag	tccaaccgg	taagacacga	3420
cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	agagcgagg	atgtaggcgg	3480
tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	3540
tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	gttggttagct	cttgatccgg	3600
caacaaaacc	accgctggta	gcgggtggtt	ttttgtttgc	aagcagcaga	ttacgcgcag	3660
aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	3720
cgaaaactca	cgtaaggga	ttttggtcat	gagattatca	aaaaggatct	tcacctagat	3780
ctttttaaat	taaaaatgaa	gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	3840
tgacagttac	caatgcttaa	tcagttaggc	acctatctca	gcgatctgtc	tatttcgttc	3900
atccatagtt	gcctgactcc	ccgtcgtgta	gataactacg	atacgggagg	gcttaccatc	3960
tggtcccgag	gtgcaatga	taccgcgaga	cccacgtca	ccggctccag	atttatcagc	4020
aataaaccag	ccagccggaa	gggcccagcg	cagaagtgg	cctgcaactt	tatccgcctc	4080
catccagttc	attaattgtt	gccgggaagc	tagagtaagt	agttcgccag	ttaatagttt	4140
gcgcaacggt	gttggcattg	ctacaggcat	cgtggtgtca	cgctcgtcgt	ttggtatggc	4200
ttcattcagc	tccggttccc	aacgatcaag	gcgagttaca	tgatccccc	tggtgtgcaa	4260
aaaagcgggt	agctccttcg	gtcctccgat	cgttgtcaga	agtaagttag	ccgcagtgtt	4320
atcactcatg	gttatggcag	cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	4380
cttttctgtg	actggtgagt	actcaaccaa	gtcattctga	gaatagtgtg	tgccggcgacc	4440
gagttgctct	tgcccggcgt	caatacggga	taataccgcg	ccacatagca	gaactttaaa	4500
agtgtcatc	attggaaaac	gttcttcggg	gcgaaaaactc	tcaaggatct	taccgctgtt	4560
gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	4620
caccagcggt	tctgggtgag	caaaaacagg	aaggcaaaat	gccgcaaaaa	agggaataag	4680
ggcgacacgg	aaatgttgaa	tactcatact	cttctttttt	caatattatt	gaagcattta	4740
tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	atttagaaaa	ataaacaat	4800
aggggttccg	cgcacatttc	cccgaagaat	gccacctgac	gtctaagaaa	ccattattat	4860
catgacatta	acctataaaa	ataggcgtat	cacgaggccc	tttctgtctg	cgcgtttcgg	4920
tgatgacggt	gaaaacctct	gacacatgca	gctcccggag	acggtcacag	cttctctgta	4980
agcggtatgcc	gggagcagac	aagcccgtca	gggcccgtca	gcgggtgttg	gcgggtgtcg	5040
gggctggctt	aactatgcgg	catcagagca	gattgtactg	agagtgcacc	atatgcgggtg	5100
tgaaataaccg	cacagatgcg	taaggagaaa	ataccgcatac	aggcgccatt	cgccattcag	5160
gctgcgcaac	tgttggaag	ggcgatcgg	gcgggcctct	tcgctattac	gccagctggc	5220
gaaaggggga	tgtgctgcaa	ggcgattaag	tgggtaacg	ccagggtttt	cccagtcacg	5280
acgttgttaaa	acgacggcca	gtgaattgga	tttaggtgac	actata		5320

<210> 57

<211> 5143

<212> DNA

<213> MVA shuttle plasmid pLAS-2

<400> 57

cctcctgaaa	aactggaatt	taatacacca	tttgtgttca	tcatacagaca	tgatattact	60
ggattttatat	tgtttatggg	taaggtagaa	tctccttaat	atgggtacgg	tgtaagggaat	120
cattattttta	tttatattga	tggttacgtg	aaatctgaat	tttcttaata	aatattattt	180
ttattaaatg	tgtatatgtt	gttttgcgat	agccatgtat	ctactaatca	gatctattag	240
agatattatt	aattctggtg	caatatgaca	aaaattatac	actaattagc	gtctcgtttc	300
agacatggat	ctgtcacgaa	ttaatacttg	gaagtctaag	cagctgaaaa	gctttctctc	360
tagcaaaagt	gcatttaagg	cggatgtcca	tggaatagat	gccttgattt	atgcaatagc	420
tgataataac	gtgcgtctag	tatgtacgtt	gttgaacgct	ggagcattga	aaaatcttct	480
agagaatgaa	tttccattac	atcaggcagc	cacattggaa	gataccaaaa	tagtaaagat	540
tttgctattc	agtggactgg	atgattcgag	gtaccaggcg	cgccctttca	ttttgttttt	600
ttctatgcta	taaatgggtg	gcaagggcga	ggagctgttc	accgggggtg	tgcccatcct	660
ggctgagctg	gacggcgacg	taaacggcca	caagttcagc	gtgtccggcg	agggcgaggg	720

NIH265.001VPC SEQLIST.TXT

cgatgccacc	tacggcaagc	tgaccctgaa	gttcatctgc	accaccggca	agctgcccgt	780
gccctggccc	accctcgtga	ccaccctgac	ctacggcgctg	cagtgcctca	gccgctaccc	840
cgaccacatg	aagcagcacg	acttcttcaa	gtccgccatg	cccgaaggct	acgtccagga	900
gcgcaccatc	ttcttcaagg	acgacggcaa	ctacaagacc	cgcgccgagg	tgaagttcga	960
gggcgacacc	ctggtgaacc	gcatcgagct	gaaggcgatc	gacttcaagg	aggacggcaa	1020
catcctgggg	cacaagctgg	agtacaacta	caacagccac	aacgtctata	tcatggccga	1080
caagcagaag	aacggcatca	aggtgaactt	caagatccgc	cacaacatcg	aggacggcag	1140
cgtgcagctc	gccgaccact	accagcagaa	cacccccatc	ggcgacggcc	ccgtgctgct	1200
gcccgaacaac	cactacctga	gcacccagtc	cgccctgagc	aaagacccca	acgagaagcg	1260
cgatcacatg	gtcctgctgg	agttcgtgac	cgccgcccgg	atcactctcg	gcatgcacga	1320
gctgtacaag	taagagctcg	ctttctctct	agcaaagatg	catttaaggc	ggatgtccat	1380
ggacatagtg	ccttgattta	tgcaatagct	gataataacg	tgctgttagt	atgtacgttg	1440
ttgaacgctg	gagcattgaa	aaatcttcta	gagaatgaat	ttccattaca	tcaggcagcc	1500
acatttgaag	ataccaaaat	agtaaagatt	ttgctattta	gtggactgga	tgattctccg	1560
gatgtgtacc	aacctaaaaa	ttgaaaataa	atacaaaggt	tcttgagggt	tgtgttaaat	1620
tgaagcgag	aaataatcat	aaataagccc	ggggatcctc	tagagtgcac	ctgcaggcat	1680
gctcgagcgg	ccgccagtgt	gatggatatc	tgcagaattc	ggcttggggg	gctgcagggtg	1740
gatgcgatca	tgacgtcctc	tgcaatggat	aacaatgaac	ctaaagtact	agaaatggta	1800
tatgatgcta	caatctttacc	cgaaggtagt	agcatggatt	gtataaacag	acacatcaat	1860
atgtgtatac	aacgcaccta	tagttctagt	ataattgccca	tattggatag	attcctaattg	1920
atgaacaagg	atgaactaaa	taatacacag	tgtcatataa	ttaaagaatt	tatgacatac	1980
gaacaaatgg	cgattgacca	ttatggagaa	tatgtaaacg	ctattctata	tcaaattcgt	2040
aaaagaccta	atcaacatca	caccattaat	ctgttttaaaa	aaataaaaaag	aacccggtat	2100
gacactttta	aagtggatcc	cgtagaattc	gtaaaaaaag	ttatcggatt	tgtatctatc	2160
ttgaacaaat	ataaacgggt	ttatagttac	gtcctgtacg	agaacgtcct	gtacgatgag	2220
ttcaaattgtt	tcattgacta	cgtggaaact	agatatttct	aaaattaatg	atgcattaat	2280
ttttgtattg	attctcaatc	ctaaaaacta	aaatatgaat	aagtattaaa	catagcgggtg	2340
tactaattga	tttaacataa	aaaatagttg	ttaactaatc	atgaggactc	tacttattag	2400
atatatttctt	tggagaaatg	acaacgatca	aaccgggcat	gcaagcttgt	ctccctatag	2460
ttagctcgat	tagagcttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaaattgtta	2520
tcgcgtcacca	attccacaca	acatacagagc	cggaagcata	aagtgtaaag	cttgggggtgc	2580
ctaattgagtg	agctaactca	cattaattgc	gttgcgctca	ctgcccgtt	tcgagtcggg	2640
aaacctgtcg	tgccagctgc	attaatgaat	cggccaacgc	gcggggagag	gcggtttgcg	2700
tattggggcgc	tcttcgcgtt	cctcgctcac	tgactcgctg	cgctcggctg	ttcgggtcgcg	2760
gcgagcggta	tcagctcact	caaaggcggg	aatacggtta	tccacagaat	caggggataa	2820
gcgaggaag	aacatgtgag	caaaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	2880
ggttctggcg	tttttcgata	ggctcgcgcc	ccctgacgag	catcacaaaa	atcgacgctc	2940
aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	3000
ctccctcgctg	cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	3060
cccttcgggga	agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	gttcgggtgta	3120
ggctcgttcgc	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccg	accgtctgcg	3180
cttatccggg	aactatcgtc	ttgagtccta	cccggtaaga	cacgacttat	cgccactggc	3240
agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	ggcgggtgta	cagagttctt	3300
gaagtgggtg	cctaactacg	gctacactag	aaggacagta	tttggtatct	gcgctctgct	3360
gaagccagtt	accttcggaa	aaagagtttg	tagctcttga	tccggcaaac	aaaccaccgc	3420
tggtagcggg	ggtttttttg	tttgcaagca	gcagattacg	cgcaaaaaa	aaggatctca	3480
agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	tggaaacgaa	actcacgtta	3540
agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	ttaaattaaaa	3600
atgaagtttt	aaatcaatct	aaagtatata	tgagtaaact	tggtctgaca	gttaccatg	3660
cttaatcagt	gaggcaccta	tctcagcgat	ctgtctatct	cgttcatcca	tagttgcctg	3720
actccccgtc	gtgtagatac	ctacgatacg	ggagggctta	ccatctggcc	ccagtctgc	3780
aatgatactc	cgagacccac	gctcaccggc	tccagattta	tcagcaataa	accagccagc	3840
cgggaagggcc	gagcgcagaa	gtggtcctgc	aactttatcc	gcctccatcc	agtctattaa	3900
ttgttgccgg	gaagctagag	taagtagttc	gccagttaat	agtttgcgca	acgttggttg	3960
cattgctaca	ggcatcgctg	gtcacgctc	gtcgtttggg	atggcttcat	tcagctccgg	4020
ttcccaacga	tcaaggcgag	ttacatgatc	ccccatgttg	tgcaaaaaag	cgggttagctc	4080
cttcggctct	ccgatcgttg	tcagaagtaa	gttgcccgca	gtgttatcac	tcatggttat	4140
ggcagcatcg	cataattctc	ttactgtcat	ggcatccgta	agatgctttt	ctgtgactgg	4200
tgagtactca	accaagtcac	tctgagaata	gtgtatgcgg	cgaccgagtt	gctcttgccc	4260
ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	ttaaaagtgc	tcattcattgg	4320
aaaacggtct	tccggggcgaa	aactctcaag	gatcttaccg	ctgttgagat	ccagttcgat	4380
gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	actttcacca	gcgtttctgg	4440
gtgagcgaata	acaggaaggc	aaaatgccc	ataaagggga	ataagggcga	cacgggaatg	4500
ttgaatactc	atactcttcc	tttttcaata	ttattgaagc	atttatcagg	gttattgtct	4560
catgagcgga	tacatatattg	aatgtattta	gaaaaataaa	caaatagggg	ttccgcgcac	4620
atttccccga	aaagtgccac	ctgacgtcta	agaaaccatt	attatcatga	cattaacctta	4680
taaaaatagg	cgtatcacga	ggccctttcg	tctcgcgcgt	ttcgggtgatg	acgggtgaaa	4740
cctctgacac	atgcagctcc	cggagacggg	cacagcttgt	ctgtaagcgg	atgccgggag	4800

NIH265.001VPC SEQLIST.TXT

cagacaagcc	cgtcagggcg	cgtcagcggg	tggttgccggg	tgctcggggct	ggcttaacta	4860
tgccgcatca	gagcagattg	tactgagagt	gcaccatattg	cggtgtgaaa	taccgcacag	4920
atgcgtaagg	agaaaaatacc	gcatcaggcg	ccattcgcca	ttcaggctgc	gcaactgttg	4980
ggaagggcga	tcggtgcggg	ctcttcgct	attacgccag	ctggcgaaaag	ggggatgtgc	5040
tgcaaggcga	ttaagttggg	taacgccagg	gttttcccag	tcacgacgtt	gtaaacgac	5100
ggccagtga	ttggatttag	gtgacactat	agaatacgaa	ttc		5143

<210> 58
 <211> 3077
 <212> DNA
 <213> HIV-1 clade A

<400> 58

atgggtgcga	gagcgtcagt	attaagtggg	ggaaaattag	atgcatggga	gaaaattcgg	60
ttaaggccag	ggggaaagaa	aaaatataga	ctgaaacact	tagtatgggc	aagcaggag	120
ctggaaaaat	tcgtacttaa	ccctagcctt	ttagaaactt	cagaaggatg	tcagcaata	180
atgaaccaa	tacaaccagc	tcttcagaca	ggaacagaag	aacttagatc	attattta	240
gcagtagcaa	ccctctattg	tgtacatcaa	cggatagagg	taaaagacac	caagggaagct	300
ttagataaag	tagaggaaat	acaaaacaag	agcaagcaaa	agacacaaca	ggcagcagct	360
gatacaggaa	acaacagcaa	ggtcagccat	aattacccta	tagtgcaaaa	tgcaaaagg	420
caaatgatac	atcagtcctt	atcaccaagg	actttgaatg	catgggtaaa	ggtaatagaa	480
gaaaggggtt	tcagcccaga	agtaataccc	atgttctcag	cattatcaga	aggagccatc	540
ccacaagatt	taaatatgat	gctgaacata	gtggggggac	accaggcagc	tatgcaaatg	600
ttaaaagaaa	ctatcaatga	ggaagctgca	gaatgggaca	ggttacatcc	agcacaggca	660
gggcctattc	caccaggcca	gataagagac	ccaaggggaa	gtgacatagc	aggaactact	720
agtacccttc	aggaacaaat	aacatggatg	acaaacaacc	cacctatccc	agtgggagac	780
atctataaaa	gatggataat	cctaggatta	aataaaaatag	taagaatgta	tagccctgtt	840
agcatttttag	atataaaaaca	ggggccaaaa	gaacccttca	gagactatgt	agatagggtc	900
tttaaagttc	tcagagccga	acaagctaca	cagggaagtaa	aaggctggat	gacagagacc	960
ctgctgggtc	aaaatgcaaa	tccagattgt	aagtccattt	taagagcatt	aggaacaggg	1020
gctacattag	aagaaatgat	gacagcatgt	caggagagtgg	gaggaccggg	ccataaagca	1080
aggggttttag	ctgaggcaat	gagtcaagca	caacaggcaa	atgtaatgat	gcagaggggc	1140
agctttaagg	ggcagaaaaag	aattaagtgc	ttcaactgtg	gcaaagaggg	acacctagcc	1200
agaaattgca	gagcccctag	gaaaaaaggc	tgttggaagt	gtgggaaaga	aggacaccaa	1260
atgaaagatt	gcaatgagag	acaggcta	tttttaggga	aaatttgcc	ttccagcaag	1320
gggaggccag	gaaattttcc	ccagagcaga	ccggagccaa	cagccccacc	agcagagatc	1380
tttgggatgg	gggaagagat	aacctcccct	ccgaagcagg	agcagaaaga	gagggaacaa	1440
acccccacct	ttgtttccct	caaatcactc	tttggcaacg	acccgttgtc	acagtaaaag	1500
taggaggaga	aatgagagaa	gctctattag	atacaggagc	agatgatata	gtattagaag	1560
atataaat	cccaggaaaa	tggaaaccaa	aaatgatagg	gggaattgga	ggttttatca	1620
aggtaaaaca	atatgatcag	gtatctatag	aaatttgtgg	aaaaaaggct	ataggtacgg	1680
tattagtagg	acctacacct	gtcaacataa	ttggaagaaa	tatgttgact	cagattgggt	1740
gtaccttaaa	ttttccaatt	agtcctattg	agactgtacc	agtaacatta	aagccaggaa	1800
tggtatggccc	aagggttaaa	caatggccat	tgacagaaga	gaaaataaaa	gcattgacag	1860
aaatttgtaa	agagatggaa	aaggaaggaa	aaatttcaaa	aattgggcct	gaaatccat	1920
acaatactcc	aatatttgca	ataaagaaaa	aagatagcac	taaatggagg	aaattagtag	1980
atttcagaga	gctcaataaa	agaacacaag	acttttgga	agttcaatta	gggataccgc	2040
atccagcggg	cctaaaaaag	aaaaaatcag	taacagtact	agaggtgggg	gatgcatatt	2100
tttcagttcc	cctagataaa	aacttttagaa	agtatactgc	atttaccata	cctagtttaa	2160
ataatgaaac	accaggaatc	aggtatcagt	acaatgtgct	tccacaagga	tgaaaaggat	2220
caccagcaat	attccagtg	agtatgacaa	aaatcttaga	gccctttaga	tcaaaaaatc	2280
cagaaataat	tatctatcaa	tacatgcacg	acttgtatgt	aggatcagat	ttagaaatag	2340
ggcagcatag	agcaaaaata	gaagaattaa	gagctcatct	actgagctgg	ggatttacta	2400
caccagacaa	aaagcatcag	aaagaacctc	cattcctttg	gatgggatat	gagctccatc	2460
ctgacaagtg	gacagtccag	cctatagagc	tgccagaaaa	agaaagctgg	actgtcaatg	2520
atatacagaa	attagtggga	aaactaaatt	gggcagctca	aatttatcca	ggaattaaag	2580
taaagcaatt	gtgtaaactt	ctcaggggag	ccaaagccct	aacagatata	gtaacactga	2640
ctgaggaagc	agaattagaa	ttagcagaga	acagggagat	tctaaaagac	cctgtgcatg	2700
gggtatatta	tgacccatca	aaagacttaa	tagcagaaat	acagaaacaa	gggcaagacc	2760
aatggacata	ccaaatttat	caggagccat	ttaaaaatct	aaaaacagga	aaatatgcaa	2820
gaaaagggtc	tgctcacact	aatgatgtaa	gacaatttagc	agaagtgtg	cagaagtgg	2880
tcatggaaag	cataagtaata	tggggaaaga	ctcctaaatt	taaactaccc	atacaaaaag	2940
agacatggga	gacatggtgg	atggactatt	ggcaagctac	ctggattcct	gagtgaggag	3000
ttgtcaatac	ccctccccta	gtaaaattat	ggtaccagtt	agagaaagac	cccatagcag	3060
gagcagagac	tttctaa					3077

NIH265.001VPC SEQLIST.TXT

<210> 59
 <211> 2241
 <212> DNA
 <213> HIV-1 Clade A

<400> 59
 atgagagtga tggggatata gatgaattgt cagcacttat tgagatgggg aactatgatc 60
 ttgggattga taataatctg taatgctgta aacagcaact tgtgggttac tgtctattat 120
 ggggtacctg tgtggaaaga tgcagagacc accttatttt gtgcatcaga tgctaaagca 180
 tataaaacag aaaagcataa tgtctgggct acacatgcct gtgtgcccac agacccaac 240
 ccacaagaaa taccttttga aaatgtgaca gaagagttaa acatgtggaa aaataaaatg 300
 gtagaacaata tgcatacaga tataatcagt ctatgggacc aaagcctaca gccatgtgta 360
 aagttaaccc ctctctgcat tacttttaac tgtacagatg ttactaatgt tacagatgtt 420
 agtggtagca ggggcaacat caccatcatg aaagagatgg agggagaaat aaaaaactgt 480
 tctttcaata tgaccacaga aataagggat aagaaacaga aagtatatc actcttttat 540
 agacttgatg tagtaccaat aaatcagggt aatagtagta gtaaaaacag tagtgagtat 600
 agattaataa gttgtaatac ctacagccatt acacaagctt gcccaaagggt aagctttgag 660
 ccaattccca tacattattg tgccccagct ggttttgcca tcctgaagtg tagggataag 720
 gagtccaatg gaacagggga atgcaagaat gtcagcacag tccaatgcac acatggaatc 780
 aagccagtag tatcaactca actactgtta aatggcagtc tagcagaaga aaaggtaaaa 840
 atcagaactg aaaatatcac aaacaatgcc aaaactatag tagtacaact tgtcagacct 900
 gtgagaatta attgtactag acctaataac aatacaagag agagtgtgca tataggcca 960
 ggacaagcat tctttgcaac aggtgacata ataggggata taagacaagc acattgtaat 1020
 gtcagtagat cacaatggaa taagacttta caacaggtag ctgaacaatt aagagaacac 1080
 tttaaaaaca aaacaataat atttaacagt tcctcaggag gggatctaga aatcacaca 1140
 catagtttca attgtggagg agaattcttc tattgtaata catcaggctc gttcaatagc 1200
 acctggaata ccagcatgtc aggggtcaagt aacacggaga caaatgacac tataactctc 1260
 caatgcagaa taaagcaaat tataaatatg tggcagagaa caggacaagc aatataatgcc 1320
 cctcccatcc agggagtgat aagggtgtgaa tcaaacaatca caggactact gttacaaga 1380
 gagggtgggg aggagaagaa cagtacaaat gaaatcttca gacctggagg aggagatatg 1440
 agggacaact ggagaagtga attatataag tataaagtag taaaaattga accactagga 1500
 gtagcaccca ccagggaag gagaagagtgt gtgggaagag aaaaaagagc agttggaata 1560
 ggagctgttt tccttgggtt cttaggagca gcagggaagca ctatgggcgc ggcgtcaata 1620
 acgctgacgg tacaggccag gcaattattg tctggcatag tgcagcagca gagcaatttg 1680
 ctgagggcta tagaggctca acaacatatg ttgaaactca cggctctgggg cattaaacag 1740
 ctccaggcaa ggtccttgc tgtggaaaga tacctaaggg atcaacagct cctagggaatt 1800
 tgggctgtct ctggaaaact catctgcacc actaatgtgc cctggaactc tagttggagt 1860
 aataaatctc aggatgaaat atggaacaac atgacctggc tgcaatggga taaagaaatt 1920
 agcaattaca taaacctaat atatagtcta attgaagaat cgcaaaacca gcaggaaaag 1980
 aatgaacaag acttattggc attgggcaag tgggcaaatc tgtggacttg gtttgacata 2040
 tcaaatggc tgtggtatat aagaatattt ataattgatg taggaggctt aataggatta 2100
 agaattagtt ttgctgtgct tgctgtaata aagagagtta ggcagggata ctacactgtg 2160
 tcatttcaga tccatgcccc aaaccagggt ggtctcgaca ggcccgaag aatcgaagga 2220
 gaagggtggag agcaagacta a 2241

<210> 60
 <211> 2211
 <212> DNA
 <213> HIV-1 Clade A

<400> 60
 atgagagtga tggggatata gatgaattgt caaagcttgt ggagatgggg aactatgatc 60
 ttgggaattgt taatgatttg tagtgttgca ggaacttgt gggttactgt ctactatggg 120
 gtacctgtgt ggaaagaggc agacaccacc ttattttgtg tatcaaatgc tagagcatat 180
 gatacagaag tgcataatgt ctgggctaca catgcctgtg tacctacgga cccaaccca 240
 caagaaatag atttgagaga tgtgacagaa gagtttaaca tgtggaaaaa taacatggtg 300
 gagcaaatgc atacagatat aattagtcta tgggaccaa gcctaaaacc atgtgtaaag 360
 ttaacccctc tctgcgttac tttagattgt ggcataatg taaccaactt gaatttcacc 420
 agtaacatga aaggagacat aacaaactgc tcttacaata tgaccacaga aataagggat 480
 aggaacaga aagtgtattc acttttctat aggttgata tagtaccaat taatgaagaa 540
 aagaataata gcaggagagac tagtccgtat agattaataa attgtaatac ctacagccatt 600
 acacaagctt gtcctaaggt atcttttgaa ccaattccca tacattattg tgccccagcc 660
 ggttttgcca ttctaataatg taaggatgca aggttcaatg gaacagggcc atgcaagaat 720
 gtcagcacag tacaatgtac acatggaaatc aggccagtaa tatcaactca actgctgtta 780
 aatggcagtt tagcagagaa tgggacaaag attagatctg aaaatatcac aaacaatgcc 840
 aaaaccataa tagtacaact taacgagacc gtacaaatta attgtaccag acctagcaac 900
 aatacaagaa aaagtgtacg tataggacca ggacaagcat tctatacaac aggtgatata 960
 acaggggata taagacaagc atattgtaat gtcagttagc aagaatggga acaagcatta 1020

NIH265.001VPC SEQLIST.TXT

aaaggggtag	ttatacaatt	aagaaaacac	tttaacaaaa	caataatctt	taacagttcc	1080
tcaggagggg	atttagaaat	tacaacacat	agttttaatt	gtggaggaga	attcttctat	1140
tgtgatacat	caggcctgtt	taatagcacc	tggaacacga	acaccaccga	gccaacaac	1200
acaacgtcaa	atggcactat	cattctccaa	tgcaagaata	agcaaattat	aaatctgttg	1260
cagagaaccg	gacaagcaat	gtatgccctt	cccatccaag	gggtaataag	gtgtgattcc	1320
aacattacag	gactactatt	aacaagagat	ggtggagtag	ttgatagtag	aaatgaaacc	1380
gaaatcttca	gacctggagg	aggagatatg	agggacaatt	ggagaagtga	attatataag	1440
tataaagtag	taaaaattga	accactagga	gtagcaccca	ccggggcaaa	gagaagagt	1500
gtggagagag	aaaaaagagc	agttggcata	ggagctgtat	tcattggggt	cttagggagc	1560
gcaggaagca	ctatgggctg	ggcgtcaata	acgctgacgg	tacaggccag	acaattattg	1620
tctggcatag	tgcaacagca	aagcaatttg	ctgagggtca	tagagggtca	acagcatatg	1680
ttgagactca	cggctctggg	cattaagcag	ctccaggcaa	gagtcctggc	tgtggaaaga	1740
tacctaaggg	atcaacagct	cctaggaatt	tggggctgct	ctggaaaact	catctgcacc	1800
actaatgtgc	cctggaactc	tagttggagt	aataaatctc	aggaggaaat	atggggtaac	1860
atgacctggc	tgcaatggga	taaagaaatt	agcaattaca	cacaaacaat	atataacctc	1920
cttgaagaat	cgcagaacca	gcaggaaaag	aatgaacaag	acttattggc	attggacaag	1980
tgggcaaat	tgcgacttg	gtttgacata	acaaattggc	tgtggtatat	aaaaatgttt	2040
ataatgatag	taggaggctt	aataggatta	agaatagttt	ttgctgtgct	ttctgtaata	2100
aatagagtta	ggcagggata	ctcacctctg	tcgtttcaga	cccataatcc	gagcccaagg	2160
ggtctcgata	ggcccgggaag	aatcgaagga	gaagggtggag	agcaagacta	a	2211

<210> 61

<211> 3083

<212> DNA

<213> HIV-1 Clade C

<400> 61

atgggtgcga	gagcgtcaat	attaagaggg	ggaaaattag	atcgatggga	aaaaattagg	60
ttaaggccag	ggggaaagaa	aagctatatg	ataaaacact	tagtatgggc	aagcagggag	120
ctggaaagat	ttgcacttaa	ccctagcctt	ttagagacat	cagaaggctg	taaacaaata	180
atgaaacagc	tacaaccagc	tcttcagaca	tgacagagaag	aacttaaatc	attattcaat	240
gcaatagcag	ttctctattg	tgtacatgaa	gggatagatg	taaaagacac	caaggaagcc	300
ttagacaaga	tagaggaaga	acagaacaaa	agtcagcaaa	aaacacagca	ggcagaagca	360
gctggcgga	aagtcagtag	aaattatcct	atagtgcaga	atctccaagg	acaaatggta	420
caccagtcca	tatcacctag	aactttgaat	gcatgggtaa	aagtaataga	ggaaaaggct	480
tttagccag	aggtaatacc	catgtttaca	gcattatcag	aaggagccac	cccacaagat	540
ttaaacacca	tgctaaatac	agtgggggga	catcaagcag	ccatgcaaat	gttaaaagat	600
accatcaatg	aggaggctgc	agaatgggat	aggatacatc	cagtacatgc	agggcctact	660
gcaccaggcc	aaatgagaga	accaagggga	agtgacatag	caggaactac	tagtaccctt	720
caggaacaaa	tagcatggat	gacagctaac	ccacctgttc	cagtgggaga	aatctacaaa	780
agatggataa	tactgggttt	aaataaaata	gtaagaatgt	atagccctgt	cagcattttg	840
gacataaaac	aagggccaaa	ggaacccttt	agagactatg	tagatcggtt	ctttaaaact	900
ttaagagctg	aacaggctac	acaagatgta	aaaaattgga	tgacagacac	cttgttggtc	960
caaaatgcga	accagattg	taagaccatt	ttaagagcat	taggaccagg	ggctacatta	1020
gaagaaatga	tgacagcatg	tcaaggagtg	ggaggacctg	gccacaaagc	cagagttttg	1080
gctgaggcaa	tgagccaagc	aaacacacac	ataatgatgc	agagaagcaa	ttttaaaggc	1140
tctaaaagaa	tgtttaaagt	tttcaactgt	tgcaaggaa	ggcacatagc	cagaaattgc	1200
agggccctta	ggaaaaaagg	ctgttgga	ttgtggaaag	aaggacacca	aatgaaagac	1260
tgtactgaga	ggcaggctaa	tttttttagg	aaaatttggc	cttccacaa	ggggaggcca	1320
gggaatttcc	ttcagaacag	gtcagagcca	acagcccac	caacgaacag	gccagagcca	1380
acagctccac	cagcagagag	cttcagggtc	gaggaagcaa	cccctgctcc	gaagcaggag	1440
ctgaaagaca	gggaaccttt	aatttccctc	aatcactctt	ttggcagcga	cccctcgtct	1500
caataaaagt	aggggggtcaa	acaaaggagg	ctctttttaga	cacaggagca	gatgatacag	1560
tattagaaga	aataaatttg	ccaggaaaat	ggaaacccaa	aatgatagga	ggaattggag	1620
gttttatcaa	agtaagacag	tatgatcaga	tagttataga	aatttgtgga	aaaaaggcta	1680
taggtacagt	attagtagga	cccacccttg	tcaacataat	tggaagaaat	atgttgactc	1740
agcttgtagt	cacactaaat	tttccaatta	gtcctattga	aactgtacca	gtaaagttaa	1800
agccaggaat	ggatggccca	aagggttaaac	aatggccatt	gacagaagaa	aaaataaagg	1860
cattaacagc	aatgtgtgaa	gaaatggaga	aggaaggaaa	aattacaaag	attgggcctg	1920
aaaatccata	taacactcca	gtatttgcca	taaaaaagaa	ggacagtact	aagtggagaa	1980
aattagtaga	tttcagggaa	cgcaataaaa	gaactcaaga	tttttgggaa	gttcaattag	2040
gcataccaca	cccagcaggg	ttaaaaaaga	aaaaatcagt	gacagtactg	gagggtgggg	2100
atgcatactt	ctcagttcct	ttagatgaag	gcttcaggaa	atatactgca	ttcaccatac	2160
ctagtataaa	caatgaaaca	ccaggaaatta	gtatcataata	caatgtgctt	ccacagggat	2220
ggaaaggatc	accagcaata	ttccagagta	gcatgacaaa	aatcttagag	cccttttagag	2280
cacaaaatcc	agaaatagtc	atctatcaat	atatgcacga	cttatatgta	ggatctgact	2340
tagaaatagg	gcaacataga	gcaaaaatag	aggaattaag	agaacatcta	ttaaagtggg	2400
gatttaccac	accagacaag	aaacatcaga	aagaaccccc	atttcttttg	atgggggtatg	2460

NIH265.001VPC SEQLIST.TXT

aactccatcc	tgacaaatgg	acagtacagc	ctataacgct	gccagaaaag	gaaagctgga	2520
ctgtcaatga	tatacagaag	ttagtgggaa	aactaaactg	ggcaagtcag	atztatgcag	2580
ggattaaagt	aaggcaactg	tataaactcc	ttaggggagc	caaagcacta	acagacatag	2640
taccactaac	tgaagaggca	gaattagaat	tggcagagaa	cagggaaatt	ctaaaagaac	2700
cagtacatgg	ggtatattat	gacccatcaa	aagacttgat	agctgaaata	cagaaacaag	2760
ggcatgacca	atggacatat	caaattttacc	aagaaccatt	caaaaatctg	aaaacaggga	2820
agtatgcaaa	aatgaggagt	gcccacacta	atgatgtaaa	acaattaaca	gaggcagtg	2880
aaaaaatagc	catggaaggc	atagtaatat	ggggaaagac	tcctaaattt	agactgcccc	2940
ttcaaaaagga	aacatgggaa	acatgggtgga	cagactattg	gcaagccacc	tggatttcctg	3000
agtgggagtt	tgtaataacc	cctcccctag	taaaattatg	gtaccagctg	gagaaagaac	3060
ccatagtagg	agcagaaact	ttc				3083

<210> 62

<211> 2226

<212> DNA

<213> HIV-1 Clade C

<400> 62

atgagagtga	aggggatatt	gaggaattgg	caacacaggt	ggatatggat	ctggatcatc	60
ttaggctttt	ggatgctaata	gatttgtaata	gggaacttgt	gggtcactgt	ctactatggg	120
gtacctgtgt	ggaaagaagc	aaatgtctct	ctatttttgg	catcagatgc	taaagcatat	180
gagaaagaag	tgcataatgt	ctgggctaca	catgcctgtg	taccacacaga	ccccaaccca	240
caagaactag	acttgggtaaa	tgtaacagaa	aatttttaaca	tgtggaaaaa	tgacatggta	300
gatcagatgc	atgaggatat	aatcagttta	tgggatgaaa	gcctaaagcc	atgtgtaaaag	360
ttgacccccc	tctgtgtcac	tctaaactgt	actaatgcta	atattaataa	tgatactgtt	420
gctaatagtg	gtacttttta	ggttgataat	agtagtaaat	tagtaaaaaa	ttgtcttttc	480
aatataacca	cagaaataag	agataagaag	aaaaaagaat	attcattgtt	ttatagactt	540
gatataattac	cacttgataa	ctctagttag	tctaagaact	atagttagta	tgtattaata	600
aatttgtaatg	cctcaaccgt	aacacaagcc	tgtccaaagg	tctcttttga	cccaatttct	660
atacattatt	gtgctccagc	tggttatgcg	attctaaagt	gtaaagataa	gacatttcaat	720
ggaacaggac	catgcagtaa	tgtcagcaca	gtactatgta	cacatgggaat	taagccagtg	780
gtatcaactc	aattactgtt	aaatggtagc	ctagcagaag	aagggatagt	aattagatct	840
gaaaatctga	caaacaatgc	caaaacaaca	atagtacagc	ttaatgaacc	tgtagaaatt	900
atgtgtgttaa	gacccggcaa	taatacaaga	aaaagtgtga	ggataggacc	aggacaaaca	960
ttctatgcaa	caggaggcat	aataggagat	ataagacaag	cacattgtaa	cattagttaga	1020
agtgtattgga	ataaaaacttt	acaagaggta	ggtaaaaaat	tacgagaata	cttccacaat	1080
aaaacaataa	gattttaaacc	ggcggtcgta	ggaggggacc	tggaaattac	aacacatagc	1140
tttaattgta	gaggagaatt	cttctattgc	aatacatcag	aactgtttac	aggtgaatat	1200
aatgggtactg	agtataagaa	tacttcaaat	tcaaactcta	acatcacact	cccatgtaga	1260
ataaaaacaat	ttgtaaacat	gtggcagagg	gtaggacgag	caatgtatgc	ccctcctatt	1320
gaaggaaaca	taacatgtaa	ctcaagtatc	acaggactac	tattgacatg	ggatggaggga	1380
aacaatacta	atggcacaga	gacatttaga	cctggaggag	gagatatgag	ggataatttg	1440
agaagtgaat	tatataaata	taaagtggta	gaaatttaaac	cattaggaat	agcaccact	1500
agtgcaaaaa	ggagagtgg	ggagagagag	aaaagagcag	tgggaatagg	agctttgttc	1560
cttgggttct	taggagcagc	aggaagcact	atgggcgcag	catcaataac	gctgacggta	1620
caggccagac	aattattgtc	tggatatgtg	caacagcaaa	gcaatttgc	gagggccata	1680
gaggcgcaac	agcatatgtt	gcaactcaca	gtctggggca	ttaaacagct	ccagacaaga	1740
gtcctggcta	tagaaagata	cctaaaggat	caacagctcc	tagggatttg	gggctgctct	1800
ggaaaactca	tctgcaccac	tgtgtgacct	tggaaacta	gttggagtaa	taaaactgaa	1860
caggacattt	ggaatctaac	ctggatgcag	tgggatagag	aagttagtaa	ttacacagac	1920
ataatataca	agttgcttga	agactcacia	atccagcagg	aaaacaatga	aaaggattta	1980
ctagcattgg	acagttggaa	aaatctgttg	aatttggttg	acataacaaa	ttggttggtg	2040
tatataagaa	cattcataat	gatagtagga	ggcttgatag	gcttaaggat	aatttttgct	2100
gtaatttcta	tagtgaatag	agttaggcag	ggatactcac	ctttgtcatt	tcagaccctt	2160
accccaaccc	cgaggggacc	agaaaggctc	ggaggaatcg	aagaagaagg	tggagagcaa	2220
gactaa						2226

<210> 63

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic peptide

<400> 63

Ala Met Gln Met Leu Lys Glu Thr Ile

1

5