

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 September 2003 (18.09.2003)

PCT

(10) International Publication Number  
**WO 03/076591 A2**

(51) International Patent Classification<sup>7</sup>: **C12N**

(21) International Application Number: PCT/US03/07177

(22) International Filing Date: 10 March 2003 (10.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10/093,953 8 March 2002 (08.03.2002) US  
10/336,566 3 January 2003 (03.01.2003) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:  
US 10/336,566 (CON)  
Filed on 3 January 2003 (03.01.2003)

(71) Applicant (for all designated States except US): **EMORY UNIVERSITY** [US/US]; 401 Administration Building, Atlanta, GA 30322 (US).

(71) Applicant (for US only): **UNITED STATES OF AMERICA, as represented by THE SECRETARY OF HEALTH AND HUMAN SERVICES** [US/US]; Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852-3804 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ROBINSON, Harriet, L.** [US/US]; 1051 Mason Woods Drive, Atlanta, GE

30329 (US). **SMITH, James** [US/US]; 6523 Sevenoaks Drive, Tucker, GA 30084 (US). **HUA, Jian** [US/CN]; 3251 Valleybrook Place, Decatur, GA 30333 (US). **MOSS, Bernard** [US/US]; 10301 Dickens Avenue, Bethesda, MD 20814 (US).

(74) Agent: **MEIKLEJOHN, Anita, L.**; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, 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, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

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

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: COMPOSITIONS AND METHODS FOR GENERATING AN IMMUNE RESPONSE

(57) Abstract: We have developed DNA and viral vectors that can be used, alone or in combination, as a vaccine against one HIV clade, subtype, or recombinant form of HIV or against multiple HIV clades, subtypes, or recombinant forms. Moreover, the vectors can encode a variety of antigens, which may be obtained from one clade or from two or more different clades, and the antigens selected and/or the manner in which the vectors are formulated (*e.g.*, mixed) can be manipulated to generate a protective immune response against a variety of clades (*e.g.*, the clades to which a patient is most likely to be exposed; with the proportions of the components of the vaccine tailored to the extent of the patient's risk to a particular clade or clades).



WO 03/076591 A2

**COMPOSITIONS AND METHODS FOR GENERATING AN IMMUNE RESPONSE****RELATED APPLICATIONS**

This application is a continuation-in-part of U.S.S.N. 10/093,953, which was filed on March 8, 2002, and which is a continuation-in-part of U.S.S.N. 09/798,675, which was filed  
5 on March 2, 2001, and which claims the benefit of the filing dates of four provisional applications (U.S.S.N. 60/251,083, filed December 1, 2000, U.S.S.N. 60/186,364, filed March 2, 2000, U.S.S.N. 60/324,845, filed September 25, 2001, and U.S.S.N. 60/325,004, filed September 26, 2001) and the benefit of the filing date of International Application  
10 No. PCT/US01/06795, which was filed on March 2, 2001, and which was published in English under International Publication No. WO 01/92470 on December 6, 2001. The contents of all earlier filed applications listed above are hereby incorporated by reference in their entirety.

**FIELD OF THE INVENTION**

15 The present invention is directed generally to the fields of molecular genetics and immunology. More particularly, the present invention features expression vectors and methods of administering those vectors to animals.

**GOVERNMENT SUPPORT**

20 The work described herein was supported, at least in part, by grants from the National Institutes of Health (P01 AI43045, P01 AI49364, and R21 AI44325). The United States Government may therefore have certain rights in this invention.

**BACKGROUND OF THE INVENTION**

25 Vaccines have had profound and long lasting effects on world health. Smallpox has been eradicated, polio is near elimination, and diseases such as diphtheria, measles, mumps, pertussis, and tetanus are contained. Nonetheless, current vaccines address only a handful of the infections suffered by people and domesticated animals. Common infectious diseases for which there are no vaccines cost the United States alone about \$120 billion dollars per year  
30 (Robinson *et al.*, American Academy of Microbiology, May 31 - June 2, 1996). In first world countries, emerging infections such as immunodeficiency viruses, as well as reemerging

diseases like drug resistant forms of tuberculosis, pose new threats and challenges for vaccine development. The need for both new and improved vaccines is even more pronounced in third world countries where effective vaccines are often unavailable or cost-prohibitive.

The prevalence of HIV-1 infection has made vaccine development for this recently  
5 emergent agent a high priority for world health. Pre-clinical trials on DNA vaccines have demonstrated that DNA alone can protect against highly attenuated HIV-1 challenges in chimpanzees (Boyer *et al.*, *Nature Med.* 3:526-532, 1997), but not against more virulent SIV challenges in macaques (Lu *et al.*, *Vaccine* 15:920-923, 1997). A combination of DNA priming plus an envelope glycoprotein boost has raised neutralizing antibody-associated protection  
10 against a homologous challenge with a non-pathogenic chimera between SIV and HIV (SHIV-IIIB) (Letvin *et al.*, *Proc. Natl. Acad. Sci. USA* 94:9378-9383, 1997). A comparative trial testing eight different protocols for the ability to protect against a series of challenges with SHIVs (chimeras between simian and human immunodeficiency viruses) revealed the best containment of challenge infections by an immunization protocol that included priming by  
15 intradermal inoculation of DNA and boosting with recombinant fowl pox virus vectors (Robinson *et al.*, *Nature Med.* 5:526, 1999). This containment of challenge infections was independent of the presence of neutralizing antibody to the challenge virus. Despite these and many other efforts, a vaccine for containing HIV infection is still not commercially available.

## 20 SUMMARY OF THE INVENTION

The continuing force of the AIDS epidemic illustrates the pressing need for effective vaccines against human immunodeficiency viruses (HIV), which frequently mutate and exist in several different clades (or subtypes) and recombinant forms. These subtypes and recombinant forms, which may arise either naturally or as the result of human intervention,  
25 can be distinguished by differences in the sequences of their nucleic acid. We have developed DNA vectors, viral vectors and recombinant viruses, *e.g.*, recombinant modified vaccinia Ankara (MVA) viruses (described at length below) that can be used, alone or in combination, as a vaccine against one HIV clade, subtype, or recombinant form of HIV or against multiple HIV clades, subtypes, or recombinant forms (unless otherwise specified, the  
30 term "clade(s)" is meant to encompass subtypes or recombinant forms of HIV). Moreover, the vectors can encode a variety of antigens, which may be obtained from one clade or from

two or more different clades, and the antigens selected and/or the manner in which the vectors are formulated (*e.g.*, mixed) can be manipulated to generate a protective immune response against a variety of clades (*e.g.*, the clades to which a patient is most likely to be exposed).

5           There is also a need for an effective vaccine against poxviruses, such as the variola virus that causes smallpox; the current smallpox vaccine carries a small risk of substantial adverse side effects. Although smallpox has been eradicated, the population is still threatened by smallpox as a biological weapon. The viral vectors described herein can be used to generate an immune response against poxviruses. Thus, methods in which such  
10       vectors are administered (regardless of the precise protocol followed) can also elicit an immune response that confers protective or therapeutic effects against conditions such as smallpox (*i.e.*, a pox viral vector can be administered before or after (*e.g.*, 1-4 or more days after) a subject has been exposed to an agent that causes a viral disease such as smallpox). These methods can be effective regardless of whether the vectors contain vaccine inserts or  
15       what that insert encodes (*e.g.*, proteins obtained from an HIV or proteins that elicit an immune response against one or more HIV clades).

          The present invention provides plasmid vectors as well as viral vectors that can be used to deliver nucleic acids to a cell; while the invention encompasses vectors that do not contain vaccine "inserts," when immunizing or treating a patient, the vectors will include  
20       nucleic acids that encode protein antigens that induce or enhance an immune response against a pathogen (*e.g.*, one or more HIV clades (or subtypes or recombinant forms)). The nucleic acids or polynucleotides described herein include those having linear arrays of naturally occurring and/or synthetic nucleotides (or nucleosides) derived from cDNA (or mRNA) or genomic DNA, or derivatives thereof (the pyrimidine or purine rings can be  
25       attached to a pentose sugar, such as a ribose or deoxyribose). The sequence of the nucleic acid may or may not be identical to a sequence occurring in nature (*e.g.*, the sequence can encode a mutant form of an HIV protein that may make the vaccine safer). Specific characteristics and specific sequences of the proteins that can be expressed by way of the vectors described herein are discussed below.

30           Plasmid or viral vectors can include nucleic acids representing one or more genes found in one or more HIV clades or any fragments or derivatives thereof that, when

expressed, elicit an immune response against the virus (or viral clade) from which the nucleic acid was derived or obtained. The nucleic acids may be purified from HIV or they may have been previously cloned, subcloned, or synthesized and, in any event, can be the same as or different from a naturally occurring nucleic acid sequence. The plasmid vectors of the present invention may be referred to herein as, *inter alia*, expression vectors, expression constructs, plasmid vectors or, simply, as plasmids, regardless of whether or not they include a vaccine insert (*i.e.*, a nucleic acid sequence that encodes an antigen or immunogen). Similar variations of the term "viral vector" may appear as well (*e.g.*, we may refer to the "viral vector" as a "poxvirus vector," a "vaccinia vector," a "modified vaccinia Ankara vector," or an "MVA vector"). The viral vector may or may not include a vaccine insert.

DNA molecules and viral vectors described herein can be used to generate recombinant viruses that can be used in combination with plasmid vectors to immunize patients prophylactically or therapeutically.

Accordingly, in one aspect, the invention features compositions (including pharmaceutically or physiologically acceptable compositions) that contain, but are not limited to, a vector, which may be a plasmid or viral vector, having a vaccine insert. The insert can include one or more of the sequences described herein (the features of the inserts and representative sequences are described at length below; any of these, or any combination of these, can be used as the insert). When the insert is expressed, the expressed protein(s) may generate an immune response against one or more HIV clades. One can increase the probability that the immune response will be effective against more than one clade by including sequences from more than one clade in the insert of a single vector (multi-vector vaccines are also useful and are described further below). For example, to increase the probability of generating an immune response against clade B and clade C, one can administer, to a subject, vectors that each include an insert that encodes proteins of clade B and clade C. The subject may be a person who lives in, or travels between, parts of the world where HIV clades B and C are prevalent. Of course, expressing one or more proteins of a single clade is also beneficial and vectors that do so are within the scope of the invention (again, any inserts having the features or sequences of the exemplary inserts described herein can be used, and the inserts *per se* are features of the invention).

In another aspect, the invention features compositions (including pharmaceutically or physiologically acceptable compositions) that contain, but are not limited to, two vectors: a first vector that encodes one or more antigens (*i.e.*, a vector that includes a vaccine insert) that elicit (*e.g.*, induces or enhances) an immune response against an HIV of a first clade and  
5 a second vector that encodes one or more antigens that elicit (*e.g.*, induces or enhances) an immune response against an HIV of a second clade. However, the compositions can contain more than first and second vectors; they can contain three, four, five, six, or more different vectors (by “different” vectors, we mean vectors that contain different regulatory elements (*e.g.* different promoters), or that encode different antigens or combinations of antigens, or  
10 that otherwise vary (*e.g.*, that vary in their “backbone” sequence)). In some embodiments, the compositions can contain as many vectors as are required to elicit an immune response against two, three, four, or a majority of, if not all, HIV clades. While one vector can encode one antigen (*e.g.*, Gag-Pol), one or more of the vectors (*i.e.*, the first vector, the second vector, or both; the first, second, third, or all three vectors; *etc.*) can include nucleic acids  
15 encoding at least three antigens (*e.g.* Gag-Pol and Env), each of which can elicit an immune response directed primarily against the same HIV clade (*i.e.*, the first vector can express three antigens, each of which generates a response against, primarily, clade A and a second vector can express three antigens, each of which generates a response against, primarily, clade B). In other embodiments, one or more of the vectors can elicit an immune response against  
20 more than one HIV clade (*i.e.*, the first vector can express a first antigen (*e.g.*, Gag-Pol) that generates a response against clade A and a second antigen (*e.g.*, Env) that generates a response against clade B). Thus, one or more of the vectors can elicit an immune response against more than one HIV clade. Any of the types of vectors described herein, whether they are plasmid or viral vectors, or whether they individually encode antigens that elicit immune  
25 responses against primarily one, or more than one, HIV clade, can be used alone or in combination with one another, depending on the particular HIV clades to which one wishes to generate immunity.

The vaccine inserts *per se* (*i.e.*, the sequences encoding HIV proteins that serve as antigens or immunogens) are also within the scope of the invention. While these inserts are  
30 described at length below, we note here that the invention features a variety of isolated nucleic acids that represent modified HIV genomes (*e.g.*, fragments or recombinant forms of

a genome or one or more HIV genes that are recombined or mutated in some way). For example, one or more nucleic acids can be deleted from one or more genes or replaced with other nucleic acids (*i.e.*, the sequences can be fragments of a gene or genes and can contain point mutations). More specifically, the invention features isolated nucleic acids that

5 represent HIV genomes having safety mutations (*e.g.*, deletion of the LTRs and of sequences encoding integrase (IN), Vif, Vpr and Nef). The nucleic acids can encode Gag, PR, RT, Env, Tat, Rev, and Vpu proteins, one or more of which may contain safety mutations (particular mutations are described at length below). Moreover, the isolated nucleic acids can be of any HIV clade and nucleic acids from different clades can be used in combination (as described

10 further below). In the work described herein, clade B inserts are designated JS (*e.g.*, JS2, JS7, and JS7.1), clade AG inserts are designated IC (*e.g.*, IC2, IC25, IC48, and IC90), and clade C inserts are designated IN (*e.g.*, IN2 and IN3). These inserts are within the scope of the present invention, as are vectors (whether plasmid or viral) containing them (particular vector/insert combinations are referred to below as, for example, pGA1/JS2, pGA2/JS2 *etc.*

15 Expression vectors that carry DNA are necessarily limited in that they can only be used to immunize patients with products (*i.e.*, proteins) encoded by DNA, and it is possible that bacterial and parasitic proteins may be atypically processed by eukaryotic cells. Another problem with existing DNA vaccines is that some vaccine insert sequences are unstable during the growth and amplification of DNA vaccine plasmids in bacteria. Instability can

20 arise during plasmid growth where the secondary structure of the vaccine insert or of the plasmid vector (the “backbone”) can be altered by bacterial endonucleases. The expression vectors of the present invention can include a termination sequence that improves stability. The termination sequence and other regulatory components (*e.g.*, promoters and polyadenylation sequences) are discussed at length below.

25 The compositions of the invention can be administered to humans, including children. Accordingly, the invention features methods of immunizing a patient (or of eliciting an immune response in a patient, which may include multi-epitope CD8<sup>+</sup> T cell responses) by administering one or more types of vectors (*e.g.*, one or more plasmids, which may or may not have identical sequences, components, or inserts (*i.e.*, sequences that can encode

30 antigens) and/or one or more viral vectors, which may or may not be identical or express identical antigens). As noted above, the vectors, whether plasmid or viral vectors, can

include one or more nucleic acids obtained from or derived from (*e.g.*, a mutant sequence is a derivative sequence) one or more HIV clades. When these sequences are expressed, they produce an antigen or antigens that elicit an immune response to one or more HIV clades. In particular embodiments, patients receive a first vector and a second vector. The first vector  
5 can encode one or more antigens of a first HIV clade (these antigens can elicit (*e.g.*, induce or enhance) an immune response against that HIV clade) and the second vector can encode one or more antigens of a second HIV clade (here again, these antigens can elicit (*e.g.*, induce or enhance) an immune response against the second HIV clade). In alternative  
10 embodiments, the subject can receive a third, fourth, fifth, *etc.* vector encoding one or more antigens from a third, fourth, fifth, *etc.* HIV clade (or mutants thereof). Moreover, and as in other embodiments, the antigen(s) can be from any clade (*e.g.*, from one or more of clades A-L) or any HIV isolate.

Where the compositions contain vectors that differ either in their backbone, regulatory elements, or insert(s), the ratio of the vectors in the compositions, and the routes  
15 by which they are administered, can vary. The ratio of one type of vector to another can be equal or roughly equal (*e.g.*, roughly 1:1 or 1:1:1, *etc.*). Alternatively, the ratio can be in any desired proportion (*e.g.*, 1:2, 1:3, 1:4 ... 1:10; 1:2:1, 1:3:1, 1:4:1 ... 1:10:1; *etc.*). Thus, the invention features compositions containing a variety of vectors, the relative amounts of antigen-expressing vectors being roughly equal or in a desired proportion. While preformed  
20 mixtures may be made (and may be more convenient), one can, of course, achieve the same objective by administering two or more vector-containing compositions (on, for example, the same occasion (*e.g.*, within minutes of one another) or nearly the same occasion (*e.g.*, on consecutive days)).

Plasmid vectors can be administered alone (*i.e.*, a plasmid can be administered on one  
25 or several occasions with or without an alternative type of vaccine formulation (*e.g.*, with or without administration of protein or another type of vector, such as a viral vector)) and, optionally, with an adjuvant or in conjunction with (*e.g.*, prior to) an alternative booster immunization (*e.g.*, a live-vectored vaccine such as a recombinant modified vaccinia Ankara (MVA) virus comprising a DNA insert that can encode antigens identical to or differing from  
30 those encoded by the "prime" portion of the immunization. For example, the viral vector or MVA virus can contain at least some of the sequence contained with the plasmid



administered as the “prime” portion of the inoculation protocol (*e.g.*, sequences encoding one or more, and possibly all, of the same antigens). The adjuvant can be a “genetic adjuvant” (*i.e.*, a protein delivered by way of a DNA sequence). Similarly, as described further below, one can immunize a patient (or elicit an immune response, which can include multi-epitope CD8<sup>+</sup> T cell responses) by administering a live-vectored vaccine (*e.g.*, an MVA vector or recombinant MVA virus) without administering a plasmid-based (or “DNA”) vaccine. Thus, in alternative embodiments, the invention features compositions having only recombinant virus or viral vectors (with, optionally, one or more of any of the inserts described here, or inserts having their features) and methods of administering them. The viral-based regimens (*e.g.*, “MVA only” or “MVA-MVA” vaccine regimens) are the same as those described herein for “DNA-MVA” regimens, and the MVAs in any vaccine can be in any proportion desired. For example, in any case (whether the immunization protocol employs only plasmid-based immunogens, only viral-carried immunogens, or a combination of both), one can include an adjuvant and administer a variety of antigens, including those obtained from any HIV clade, by way of the plurality of vectors administered.

As implied by the term “immunization” (and variants thereof), the compositions of the invention can be administered to a subject who has not yet become infected with a pathogen (thus, the terms “subject” or “patient,” as used herein encompasses apparently healthy or non-HIV-infected individuals), but the invention is not so limited; the compositions described herein can also be administered to treat a subject or patient who has already been exposed to, or who is known to be infected with, a pathogen (*e.g.*, an HIV of any clade, including those presently known as clades A-L or mutant or recombinant forms thereof).

An advantage of DNA and rMVA immunizations is that the immunogen may be presented by both MHC class I and class II molecules. Endogenously synthesized proteins readily enter processing pathways that load peptide epitopes onto MHC I as well as MHC II molecules. MHC I-presented epitopes raise CD8 cytotoxic T cell (Tc) responses, whereas MHC II-presented epitopes raise CD4 helper T cells (Th). By contrast, immunogens that are not synthesized in cells are largely restricted to the loading of MHC II epitopes and therefore raise CD4 Th but not CD8 Tc. In addition, DNA plasmids express only the immunizing antigens in transfected cells and can be used to focus the immune response on only those

antigens desired for immunization. In contrast, live virus vectors, *e.g.*, recombinant MVA virus, express many antigens (*e.g.*, those of the vector as well as the immunizing antigens) and prime immune responses against both the vector and the immunogen. Thus, we believed these live virus vectors could be highly effective at boosting a DNA-primed response by virtue of the large amounts of antigen that can be expressed by a live virus vector preferentially boosting the highly targeted DNA-primed immune response. The live virus vectors also stimulate the production of pro-inflammatory cytokines that augment immune responses. Thus, administering one or more of the DNA vectors described herein (as a “prime”) and subsequently administering one or more of the recombinant MVA viruses (as a “boost”), could be more effective than DNA-alone or live virus vectors-alone at raising both cellular and humoral immunity. Insofar as these vaccines may be administered by DNA expression vectors and/or recombinant viruses, there is a need for plasmids that are stable in bacterial hosts and safe in animals. Plasmid-based vaccines that may have this added stability are disclosed herein, together with methods for administering them to animals, including humans.

The antigens encoded by DNA or rMVA are necessarily proteinaceous. The terms “protein,” “polypeptide,” and “peptide” are generally interchangeable, although the term “peptide” is commonly used to refer to a short sequence of amino acid residues or a fragment of a larger protein. In any event, serial arrays of amino acid residues, linked through peptide bonds, can be obtained by using recombinant techniques to express DNA (*e.g.*, as was done for the vaccine inserts described and exemplified herein), purified from a natural source, or synthesized.

Other advantages of DNA-based vaccines (and of viral vectors, such as pox virus-based vectors) are described below. The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic illustration of the plasmid construct pGA1. The identities and positions of elements present in the vector (*e.g.*, the promoter (here, a CMV promoter including

intron A), the multiple-cloning site, a terminator sequence (here, the lambda T0 terminator), and a selection gene (here, the kanamycin resistance gene) are shown. Unique restriction endonuclease sites, which are useful for cloning vaccine inserts into the plasmid, are also shown.

5 Figs. 2a and 2b relate to pGA1. Fig. 2a is an illustration of the nucleotide sequence of pGA1 (SEQ ID NO:1), and Fig. 2b is a table listing the functional regions of pGA1, their positions within the SEQ ID NO:1, and the origins of the sequences.

Figs. 3a and 3b relate to pGA1.1. Fig. 3a is an illustration of the nucleotide sequence of pGA1.1 (SEQ ID NO:2), and Fig. 3b is a table listing the functional regions of pGA1.1, their positions within SEQ ID NO:2, and the origins of the sequences. pGA1.1 differs from  
10 pGA1 in that it includes an *EcoR* I restriction site in its multiple cloning site.

Figs. 4a and 4b relate to pGA1.2. Fig. 4a is an illustration of the nucleotide sequence of pGA1.2 (SEQ ID NO:3) and Fig. 4b is a table listing the functional regions of pGA1.2, their positions within SEQ ID NO:3, and the origins of the sequences. pGA1.2 differs from pGA1.1 in that it includes a *BamH* I site in its multiple cloning site.

15 Fig. 5 is a schematic illustration of the plasmid construct pGA2. The identities and positions of elements present in the vector (*e.g.*, a promoter (here the CMV promoter without intron A), the multi-cloning site, a terminator sequence (here, the lambda T0 terminator), and a selection gene (here, the kanamycin resistance gene) are shown. Unique restriction endonuclease sites, which are useful for cloning vaccine inserts into the plasmid, are also shown.

20 Figs. 6a and 6b relate to pGA2. Fig. 6a is an illustration of the nucleotide sequence of pGA2 (SEQ ID NO:4), and Fig 6b is a table listing the functional regions of pGA2, their positions within SEQ ID NO:4, and the origins of the sequences.

Figs. 7a and 7b relate to pGA2.1. Fig. 7a is an illustration of the nucleotide sequence of pGA2.1 (SEQ ID NO:5) and Fig. 7b is a table listing the functional regions of pGA2.1, the  
25 positions within SEQ ID NO:5, and the origins of the sequences. pGA2.1 differs from pGA2 in having an *EcoR* I site in its multiple cloning site.

Figs. 8a and 8b relate to pGA2.2. Fig. 8a is an illustration of the nucleotide sequence of pGA2.2 (SEQ ID NO:6), and Fig. 8b is a table listing the functional regions of pGA2.2, their positions with SEQ ID NO:6, and the origins of the sequences. pGA2.2 differs from  
30 pGA2.1 in having a *BamH* I site in its multiple cloning site.

Fig. 9 is a schematic representation of the proviral (integrated DNA) form of the HIV genome (HIV-1 wt) and a representative vaccine insert. This representative insert has safety mutations that include deletion of the LTRs, deletion of sequences encoding integrase (IN), Vif, Vpr and Nef. The insert encodes Gag, PR, RT, Env, Tat, Rev, and Vpu proteins. Clade B inserts are designated JS (clade B), IC (clade AG) and IN (clade C) with arabic numerals designating the specific vaccine constructs (*e.g.*, JS2, JS7 and JS7.1 are examples of specific clade B vaccine constructs; IC2, IC25, IC48 and IC90, examples of specific AG vaccine constructs; and IN2 and IN3 are examples of specific clade C vaccine constructs). When inserted into the pGA1 vector, the insert-bearing plasmids are referred to as pGA1/JS2 *etc.*; when inserted into the pGA2 vector, plasmids are referred to as pGA2/JS2 *etc.*

Figs. 10a-10c relate to pGA2/JS2. Fig. 10a is an illustration of the sequence of the pGA2/JS2 clade B vaccine vector (SEQ ID NO:7), and Fig 10b is a table listing the positions of seven functional regions of pGA2/JS2, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 10c is a table listing codons that were changed, the resulting amino acid change (*e.g.*, C392S indicates a substitution of serine for cysteine at amino acid residue 392), the region of the genome where the mutation resides, and the mutation's function.

Figs. 11a-11c relate to pGA2/JS7. Fig. 11a is an illustration of the sequence of the pGA2/JS7 clade B vaccine vector (SEQ ID NO:8), and Fig. 11b is a table listing the positions of seven functional regions of pGA2/JS7, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 11c is a table listing codons that were changed, the resulting amino acid change (*e.g.*, C395S indicates a substitution of serine for cysteine at amino acid residue 395), the region of the genome where the mutation resides, and the mutation's function.

Figs. 12a-12c relate to pGA2/JS7.1. Fig. 12a is an illustration of the sequence of the pGA2/JS7.1 clade B vaccine vector (SEQ ID NO:9), and Fig 12b is a table listing the positions of functional regions of pGA2/JS7.1, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 12c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 13a-13c relate to pGA1/IC25. Fig. 13a is an illustration of the sequence of the pGA1/IC25 clade AG vaccine vector (SEQ ID NO:10), and Fig. 13b is a table listing the positions of functional regions within pGA1/IC25, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 13c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 14a-14c relate to pGA1/IC2. Fig. 14a is an illustration of the sequence of the pGA1/IC2 clade AG vaccine vector (SEQ ID NO:11), and Fig. 14b is a table listing the positions of functional regions within pGA1/IC2, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 14c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 15a-15c relate to pGA1/IC48. Fig. 15a is an illustration of the sequence of the pGA1/IC48 clade AG vaccine vector (SEQ ID NO:12), and Fig. 15b is a table listing the positions of functional regions within pGA1/IC48, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 15c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 16a-16c relate to pGA1/IC90. Fig. 16a is an illustration of the sequence of the pGA1/IC90 clade AG vaccine vector (SEQ ID NO:13), and Fig. 16b is a table listing the positions of functional regions within pGA1/IC90, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 16c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 17a-17c relate to pGA1/IN3. Fig. 17a is an illustration of the sequence of the pGA1/IN3 clade C vaccine vector (SEQ ID NO:14), and Fig. 17b is a table listing the positions of functional regions within pGA1/IN3, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig. 17c is a table listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Figs. 18a-18c relate to pGA1/IN2. Fig. 18a is an illustration of the sequence of the pGA1/IN2 clade C vaccine vector (SEQ ID NO:15), and Fig. 18b is a table listing the positions of functional regions within pGA1/IN2, the gene sequences within those regions (describing mutations, where present), and the origins of their sequences. Fig 18c is a table  
 5 listing codons that were changed, the resulting amino acid change, the region of the genome where the mutation resides, and the mutation's function.

Fig. 19 is a schematic representation of an HIV-1 Env glycoprotein. The arrow indicates the site of gp160 cleavage to gp120 and gp41. In gp120, cross-hatched areas represent variable domains ( $V_1$  to  $V_2$ ) and open boxes depict conserved sequences ( $C_1$  to  $C_5$ ).  
 10 In the gp41 ectodomain, several domains are indicated: the N-terminal fusion peptide and the two ectodomain helices (N- and C-helices). The membrane-spanning domain is represented by a black box. In the gp41 cytoplasmic domain, the Tyr-X-X-Leu (YXXL) endocytosis motif and two predicted helical domains (helix-1 and helix-2) are shown. Amino acid residues are numbered at intervals of 100.

15 Figs. 20A and 20B relate to the plasmid transfer vector pLW-48. Fig. 20A is a map of pLW-48 and Fig. 20B is a representation of its sequence.

Fig. 21 is a representation of the sequences of the plasmid transfer vector pLW-48, the Psy II promoter (which controls ADA envelope expression), the ADA envelope (truncated), the PmH5 promoter (which controls HXB2 *gag* and *pol* expression), and HXB2  
 20 *gag-pol* (with safety mutations, inactivating point mutations in RT and the deletion of integrase).

Fig. 22 is a representation of the plasmid transfer vector pLW-48 and a scheme for making an MVA recombinant virus (MVA/HIV 48).

Fig. 23 is a representation of a clade B *gag pol*.

25 Fig. 24 is a representation of a Psyn II promoter.

### DETAILED DESCRIPTION

This invention encompasses a wide variety of vectors and types of vectors (e.g., plasmid and viral vectors), each of which can, but do not necessarily, include one or more  
 30 nucleic acid sequences that encode one or more antigens that elicit (e.g., that induce or enhance) an immune response against the pathogen from which the antigen was obtained or

derived (the sequences encoding proteins that elicit an immune response may be referred to herein as “vaccine inserts” or, simply, “inserts”; when a mutation is introduced into a naturally occurring sequence, the resulting mutant is “derived” from the naturally occurring sequence). We point out that the vectors do not necessarily encode antigens to make it clear  
5 that vectors without “inserts” are within the scope of the invention and that the inserts *per se* are also compositions of the invention.

Accordingly, the invention features the nucleic acid sequences disclosed herein, analogs thereof, and compositions containing those nucleic acids (whether vector plus insert or insert only; *e.g.*, physiologically acceptable solutions, which may include carriers such as  
10 liposomes, calcium, particles (*e.g.*, gold beads) or other reagents used to deliver DNA to cells). The analogs can be sequences that are not identical to those disclosed herein, but that include the same or similar mutations (*e.g.*, the same point mutation or a similar point mutation) at positions analogous to those included in the present sequences (*e.g.*, any of the JS, IC, or IN sequences disclosed herein). A given residue or domain can be identified in  
15 various HIV clades even though it does not appear at precisely the same numerical position. The analogs can also be sequences that include mutations that, while distinct from those described herein, similarly inactivate an HIV gene product. For example, a gene that is truncated to a greater or lesser extent than one of the genes described here, but that is similarly inactivated (*e.g.*, that loses a particular enzymatic activity) is within the scope of the  
20 present invention.

The pathogens and antigens, which are described in more detail below, include human immunodeficiency viruses of any clade (*e.g.* from any known clade or from any isolate (*e.g.*, clade A, AG, B, C, D, E, F, G, H, I, J, K, or L)). When the vectors include sequences from a pathogen, they can be administered to a patient to elicit an immune  
25 response. Thus, methods of administering antigen-encoding vectors, alone or in combination with one another, are also described herein. These methods can be carried out to either immunize patients (thereby reducing the patient’s risk of becoming infected) or to treat patients who have already become infected; when expressed, the antigens may elicit both cell-mediated and humoral immune responses that may substantially prevent the infection  
30 (*e.g.*, immunization can protect against subsequent challenge by the pathogen) or limit the extent of its impact on the patient’s health. While in many instances the patient will be a

human patient, the invention is not so limited. Other animals, including non-human primates, domesticated animals and livestock can also be treated.

The compositions described herein, regardless of the pathogen or pathogenic subtype (e.g., the HIV clade(s)) they are directed against, can include a nucleic acid vector (e.g., a plasmid). As noted herein, vectors having one or more of the features or characteristics (particularly the oriented termination sequence and a strong promoter) of the plasmids designated pGA1, pGA2 (including, of course, those vectors *per se*), can be used as the basis for a vaccine or therapy. Such vectors can be engineered using standard recombinant techniques (several of which are illustrated in the examples, below) to include sequences that encode antigens that, when administered to, and subsequently expressed in, a patient will elicit (e.g., induce or enhance) an immune response that provides the patient with some form of protection against the pathogen from which the antigens were obtained or derived (e.g., protection against infection, protection against disease, or amelioration of one or more of the signs or symptoms of a disease). The encoded antigens can be of any HIV clade or subtype or any recombinant form thereof. With respect to inserts from immunodeficiency viruses, different isolates exhibit clustal diversity, with each isolate having overall similar diversity from the consensus sequence for the clade (*see, e.g., Subbarao et al., AIDS 10(Suppl A):S13-23, 1996*). Thus, any isolate can be used as a reasonable representative of sequences for other isolates of the same clade. Accordingly, the compositions of the invention can be made with, and the methods described herein can be practiced with, natural variants of genes or nucleic acid molecules that result from recombination events, alternative splicing, or mutations (these variants may be referred to herein simply as “recombinant forms” of HIV).

Moreover, one or more of the inserts within any construct can be mutated to decrease their natural biological activity (and thereby increase their safety) in humans (these human-made variants may also be referred to herein as “recombinant forms” of HIV (there are naturally occurring recombinant forms as well)). As noted above in the description of JS2, JS7 and JS7.1 and as described below (*see, e.g., Examples 7-10*), mutations can be introduced into sequences that participate in encapsidation. For example, one can mutate (by, for example, deletion of all or a part of) a cis-acting RNA encapsidation sequence in the non-coding regulatory sequence of an HIV (e.g., HIV-1). Alternatively, or in addition, one



can mutate sequences that encode any antigenic proteins (*e.g.*, any HIV antigen, including those listed above (*e.g.*, the viral RT or protease).

For example, the compositions of the invention include those having two vectors: (a) a first vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against a human immunodeficiency virus (HIV) of a first subtype or recombinant form and (b) a second vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form. The compositions can be pharmaceutically acceptable and may include a carrier or adjuvant (discussed further below). Moreover, the insert of the first vector or the insert of the second vector can include the sequences of two or more of: (a) a *gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, or *vpu* gene or (b) mutants thereof and, optionally, (c) non-coding regulatory sequences (including the sequences of single promoters) of the HIV genome. At least one of the two or more sequences can be mutant or mutated so as to limit the encapsidation of viral RNA (preferably, the mutation(s) limit encapsidation appreciably).

One can introduce mutations and determine their effect (on, for example, expression or immunogenicity) using techniques known in the art; antigens that remain well expressed (*e.g.*, antigens that are expressed about as well as or better than their wild type counterparts), but are less biologically active than their wild type counterparts, are within the scope of the invention. Techniques are also available for assessing the immune response. One can, for example, detect anti-viral antibodies or virus-specific T cells.

The mutant constructs (*e.g.*, a vaccine insert) can include sequences encoding one or more of the substitution mutants described herein (*see, e.g.* the Examples) or an analogous mutation in another HIV clade. In addition to, or alternatively, HIV antigens can be rendered less active by deleting part of the gene sequences that encode them. Thus, the compositions of the invention can include constructs that encode antigens that, while capable of eliciting an immune response, are mutant (whether encoding a protein of a different length or content than a corresponding wild type sequence) and thereby less able to carry out their normal biological function when expressed in a patient. As noted above, expression, immunogenicity, and activity can be assessed using standard techniques for molecular biology and immunology.

Several plasmids have been constructed and used to express antigens (e.g., the pGA2/JS2 construct has gone through immunogenicity studies in macaques). The plasmids made and used include pGA1 and its derivatives pGA1.1 and pGA1.2; and pGA2, and its derivatives pGA2.1 and pGA2.2 (see Examples 1-8). The vaccine constructs we made are typically referred to with the "backbone" vector and the "insert" being separated by a backslash. These constructs express HIV-1 antigens, and those constructs can be administered to patients as described herein. While antigens (wild type and those containing mutations that render them safer for administration) are discussed at length below, we note here that, based upon our present evidence, plasmids containing JS7-like inserts appear to exhibit better immunogenicity and are more efficient in priming an immune response (as evidenced by anti-Env antibodies) than are plasmids containing JS2-like inserts. pGA2/JS7 and pGA2/JS7.1 differ from pGA2/JS2 in several ways, one of which is the source of their respective antigens. In pGA2/JS7 and pGA2/JS7.1, the Gag and Pol genes were obtained from HIV-1 HXB2, whereas in pGA2/JS2 those genes were obtained from a closely related isolate of HIV-1, HIV-1 BH10. Accordingly, the invention features inserts (as well as vectors and compositions containing them) that include *Gag* and *Pol* genes obtained from HIV-1 HXB2. Moreover, these inserts can contain mutations that inhibit one or more of the biological activities carried out by Gag-Pol. The vaccine inserts designated JS7 and JS2 also differ in that JS7 has an inactivating point mutation in its protease gene. This mutation facilitates the formation of viral like particles (VLPs) by, we believe, precluding premature intracellular cleavage of the pr55 Gag protein. pGA2/JS7 and pGA2/JS7.1 both contain this protease mutation and both constructs produce VLPs in abundance. Accordingly, the invention features inserts that include mutant *gag* and/or *pol* sequences (e.g., mutations (e.g., one or more deletions or point mutations) that inhibit the protease gene). Additional point mutations in the *vpu* gene in pGA2/JS7.1 resulted in a loss of Vpu expression and an increase in Env expression (in pGA2/JS7.1, the start site of Vpu is mutated along with a downstream ATG to eliminate translation of Vpu). The increase in Env expression does not compromise Gag expression.

Identical or analogous changes can be made in any vaccine insert that includes *gag*, *pol*; any vaccine insert that encodes a viral protease; or any vaccine insert that includes a *vpu* gene (regardless of the clade or isolate from which it was obtained). Moreover, these

changes can be made in vaccine inserts that are placed in any of the plasmid or live-vectored vaccines (e.g., MVA) described herein (i.e., in any plasmid having one or more of the features or characteristics of the pGA vectors, the pGA vectors themselves, or the vaccinia vectors that may be used alone or in conjunction with (e.g., to boost) a DNA-primed patient).

5           Any plasmid within the scope of the invention can be tested for expression by transfecting cells, such as 293T cells (a human embryonic kidney cell line) and assessing the level of antigen expression (by, for example, an antigen-capture ELISA or a Western blot). Plasmids that express immunogens at a level comparable to, or higher than, the plasmids tested herein are strong therapeutic candidates and are within the scope of the invention (of course, any construct that elicits an effective immune response (e.g., any desirable degree of protection from infection or other therapeutic benefit) is within the scope of the invention, regardless of the level of antigen expression it generates). One can similarly assess the ability of candidate vectors to produce VLPs; the more the vectors' products resemble VLPs, the more likely they are to elicit a strong antibody response (while this is a desirable feature, 15           vectors that fail to form VLPs are nevertheless useful and are within the scope of the present invention). In addition to assessing expression and VLP formation in cell culture, one can assess candidate vectors *in vivo*. For example, one can assess immunogenicity in animal models (and, eventually, in human patients). Plasmids that have substantially the same sequence as the pGA vectors described herein and that express one or more of the antigens described herein are within the scope of the invention so long as they are immunogenic 20           enough to induce or enhance a therapeutically beneficial response in a patient (a plasmid can have substantially the same sequence as a pGA vector even if one or more of the component parts of the plasmid, such as the marker gene or antibiotic-resistance gene, has been deleted). In tests in animals for immunogenicity, one can perform an intracellular cytokine assay or an ELISPOT assay for IFN- $\gamma$  production in response to stimulation with an antigenic peptide to 25           evaluate the frequency of responding T cells to that peptide. Proliferation assays can also be carried out. Antigens produced by transient transfection can be used for stimulation, and supernatants from mock-transfected cultures can serve as controls. If desired, the data can be presented as a stimulation index (the growth of cultures in the presence of pathogenic (e.g., viral) antigens divided by the growth of cultures in the presence of mock antigen). 30

The nucleic acid vectors of the invention, including pGA1 and pGA2 and their derivatives can encode at least one antigen (which may also be referred to as an immunogen) obtained from, or derived from, any HIV clade or isolate (*i.e.*, any subtype or recombinant form of HIV). The antigen (or immunogen) may be: a structural component of an HIV virus; glycosylated, myristoylated, or phosphorylated; one that is expressed intracellularly, on the cell surface, or secreted (antigens that are not normally secreted may be linked to a signal sequence that directs secretion). More specifically, the antigen can be all, or an antigenic portion of, Gag, gp120, Pol, Env (*e.g.*, a CCR5-using Env; see, for example, Fig. 19), Tat, Rev, Vpu, Nef, Vif, Vpr, or a VLP (*e.g.*, a polypeptide derived from a VLP that is capable of forming a VLP, including an Env-defective HIV VLP).

Particular inserts and insert-bearing compositions include the following. Where the composition includes either a vector with an insert or an insert alone, and that insert encodes a single antigen, the antigen can be a wild type or mutant *gag* sequence (*e.g.*, a *gag* sequence having a mutation in one or more of the sequences encoding a zinc finger (*e.g.*, a mutation at a nucleotide at any of positions 1279-1281, 1288-1290, 1342-1344, or 1351-1353 of SEQ ID NOs:7 or 8 or at an analogous position in an HIV *gag* sequence of another clade). As the mutation is intended to alter the encoded protein, it will not be a silent mutation (*e.g.*, one at the third-base wobble position of a codon (this is true in the context of *gag* or any other HIV sequence included in an insert of the invention). A mutation at one or more of the positions just listed would change one or more of the cysteine residues at positions 392, 395, 413, or 416 to another residue (*e.g.*, serine). Alternatively, the mutation can be at any of positions 1271-1273, 1280-1282, 1334-1336, or 1343-1345 of any of SEQ ID NOs:10-13) or at an analogous position in an HIV *gag* sequence of another clade. Such a mutation would change one or more of the cysteine residues at positions 390, 393, 411, or 414 to another residue (*e.g.*, serine). Alternatively, the mutation can be at any of positions 1260-1262, 1269-1271, 1323-1325, or 1332-1334 of SEQ ID NOs:14 or 15 or at an analogous position in an HIV *gag* sequence of another clade. Such a mutation would change one or more of the cysteine residues at positions 390, 393, 411, or 414 to another residue (*e.g.*, serine).

Where the composition includes either a vector with an insert or an insert alone, and that insert encodes multiple protein antigens, one of the antigens can be a wild type or mutant *gag* sequence, including those described above. Similarly, where a composition includes

more than one type of vector or more than one type of insert, at least one of the vectors or inserts (whether encoding a single antigen or multiple antigens) can include a wild type or mutant *gag* sequence, including those described above or analogous sequences from other HIV clades. For example, where the composition includes first and second vectors, the vaccine insert in either or both vectors (whether the insert encodes single or multiple antigens) can encode Gag; where both vectors encode Gag, the Gag sequence in the first vector can be from one HIV clade (*e.g.*, clade B) and that in the second vector can be from another HIV clade (*e.g.*, clade C).

Where the composition includes either a vector with an insert or an insert alone, and that insert encodes a single antigen, the antigen can be wild type or mutant Pol. The sequence can be mutated by deleting or replacing one or more nucleic acids, and those deletions or substitutions can result in a Pol gene product that has less enzymatic activity than its wild type counterpart (*e.g.*, less integrase activity, less reverse transcriptase (RT) activity, or less protease activity). For example, one can inhibit RT by introducing a mutation at one or more of positions 2454-2456 or 2697-2699 of SEQ ID NO:7 or at an analogous position in a sequence of another subtype or recombinant form. While the invention is not limited to mutations that have any particular effect on enzyme activity, we believe the mutation at position 2454-2456 inhibits RT by inactivating the polymerase's active site and that the mutation at position 2697-2699 inhibits RT by ablating strand transfer activity. Accordingly, these mutations and others that have similar effects on the activity of the gene product are within the scope of the invention. More specifically, the mutation can change the amino acid encoded by the nucleotides at 2454-2456 of SEQ ID NO:7 (aspartic acid (D)) to any another amino acid (*e.g.*, asparagine (N)). Alternatively, or in addition, one can inhibit the polymerase's RNase H activity by, for example, introducing a mutation at nucleotides 3333-3335 of SEQ ID NO:7 (*e.g.*, a mutation that changes the glutamic acid residue (E) to tryptophan (W)). Alternatively, the mutation can be at any of positions 2418-2420, 2661-2663, or 3297-3299 of SEQ ID NOs:8 or 9 (other clade B inserts). Alternatively, the mutation can be at any of positions 2410-2412, 2653-2655, or 3289-3291 of any of SEQ ID NOs:10-13 (for example, the aspartic acid (D), tryptophan (W) and glutamic acid (E) residues at those positions can be changed to asparagine (N), threonine (T), and/or glutamine (Q), respectively). Alternatively, the mutation can be at any of positions 2387-2389, 2630-

2632, or 3266-3268 of SEQ ID NOs:14 or 15. Nucleic acids encoding analogous residues in other clades can be identified by one of ordinary skill in the art, even if those residues are not found at precisely the same position as they were in the clades tested here.

Where the composition includes either a vector with an insert or an insert alone, and  
5 that insert encodes multiple protein antigens, one of the antigens can be a wild type or mutant *pol* sequence, including those described above (these multi-protein-encoding inserts can also encode the wild type or mutant *gag* sequences described above). Similarly, where a composition includes more than one type of vector or more than one type of insert, at least one of the vectors or inserts (whether encoding a single antigen or multiple antigens) can  
10 include a wild type or mutant *pol* sequence, including those described above (and, optionally, a wild type or mutant *gag* sequence, including those described above (*i.e.*, the inserts can encode Gag-Pol)). For example, where the composition includes first and second vectors, the vaccine insert in either or both vectors (whether the insert encodes single or multiple antigens) can encode Pol; where both vectors encode Pol, the Pol sequence in the first vector  
15 can be from one HIV clade (*e.g.*, clade B) and that in the second vector can be from another HIV clade (*e.g.*, clade AG).

Where an insert includes some or all of the *pol* sequence, another portion of the *pol* sequence that can be altered is the sequence encoding the protease activity (regardless of whether or not sequences affecting other enzymatic activities of Pol have been altered). For  
20 example, one can introduce a mutation at position 1641-1643 of SEQ ID NO:8 (*e.g.*, a mutation that changes the glutamic acid residue normally encoded by this codon to another amino acid residue, such as alanine (A)). As with the other mutants (*e.g.*, *gag* mutants) described herein, analogous mutations can be made in sequences obtained from other HIV clades. For example, one can introduce a mutation at position 1633-1635 of SEQ ID NO:10  
25 (changing arginine (R) to another amino acid, such as asparagine (N)), at position 1703-1705 of SEQ ID NO:12 (changing glycine (G) to another residue, such as valine (V)), or at position 1828-1830 of SEQ ID NO:13 (changing leucine (L) to another residue, such as methionine (M) (SEQ ID NOs:10, 12, and 13 all represent clade AG sequences). In an insert from clade C, one can introduce a mutation at position 1610-1612 of SEQ ID NO:14  
30 (changing aspartic acid (D) to another amino acid residue, such as asparagine (N)).

Where the composition includes either a vector with an insert or an insert alone, and that insert encodes a single antigen, the antigen can be a wild type or mutant Env, Tat, Rev, Nef, Vif, Vpr, or Vpu. Where the composition includes either a vector with an insert or an insert alone, and that insert encodes multiple protein antigens, one of the antigens can be a wild type or mutant Env. For example, multi-protein expressing inserts can encode wild type or mutant Gag-Pol and Env; they can also encode wild type or mutant Gag-Pol and Env and one or more of Tat, Rev, Nef, Vif, Vpr, or Vpu (each of which can be wild type or mutant). As with other antigens, Env, Tat, Rev, Nef, Vif, Vpr, or Vpu can be mutant by virtue of a deletion, addition, or substitution of one or more amino acid residues (*e.g.*, any of these antigens can include a point mutation). With respect to Env, one or more mutations can be in any of the domains shown in Fig. 19. For example, one or more amino acids can be deleted from the gp120 surface and/or gp41 transmembrane cleavage products. With respect to Gag, one or more amino acids can be deleted from one or more of: the matrix protein (p17), the capsid protein (p24), the nucleocapsid protein (p7) and the C-terminal peptide (p6). For example, amino acids in one or more of these regions can be deleted (this may be especially desired where the vector is a viral vector, such as MVA). With respect to Pol, one or more amino acids can be deleted from the protease protein (p10), the reverse transcriptase protein (p66/p51), or the integrase protein (p32).

More specifically, the compositions of the invention can include a vector (*e.g.*, a plasmid or viral vector) that encodes: (a) a Gag protein in which one or more of the zinc fingers has been inactivated to limit the packaging of viral RNA; (b) a Pol protein in which (i) the integrase activity has been inhibited by deletion of some or all of the *pol* sequence and (ii) the polymerase, strand transfer and/or RNase H activity of reverse transcriptase has been inhibited by one or more point mutations within the *pol* sequence; and (c) Env, Tat, Rev, and Vpu, with or without mutations. In this embodiment, as in others, the encoded proteins can be obtained or derived from a subtype A, B or C HIV (*e.g.*, HIV-1) or recombinant forms thereof. Where the compositions include non-identical vectors, the sequence in each type of vector can be from a different HIV clade (or subtype or recombinant form thereof). For example, the invention features compositions that include plasmid vectors encoding the antigens just described (Gag-Pol, Env *etc.*), where some of the plasmids include antigens that are obtained from, or derived from, one clade and other plasmids include antigens that are

obtained (or derived) from another clade. Mixtures representing two, three, four, five, six, or more clades (including all clades) are within the scope of the invention.

Where first and second vectors are included in a composition, either vector can be pGA1/JS2, pGA1/JS7, pGA1/JS7.1, pGA2/JS2, pGA2/JS7, pGA2/JS7.1 (pGA1.1 or pGA1.2  
5 can be used in place of pGA1 and pGA2.1 or pGA2.2 can be used in place of pGA2).

Similarly, either vector can be pGA1/IC25, pGA1/IC2, pGA1/IC48, pGA1/IC90, pGA2/IC25, pGA2/IC2, pGA2/IC48, or pGA2/IC90 (here again, pGA1.1 or pGA1.2 can be used in place of pGA1 and pGA2.1 or pGA2.2 can be used in place of pGA2). In alternative  
10 embodiments, the encoded proteins can be those of, or those derived from, a subtype C HIV (e.g., HIV-1) or a recombinant form thereof. For example, the vector can be pGA1/IN2, pGA1.1/IN2, pGA1.2/IN2, pGA1/IN3, pGA1.1/IN3, pGA1.2/IN3, pGA2/IN2, pGA2.1/IN2, pGA2.2/IN2, pGA2/IN3, pGA2.1/IN3, or pGA2.2/IN3.

The encoded proteins can also be those of, or those derived from, any of HIV clades (or subtypes) E, F, G, H, I, J, K or L or recombinant forms thereof. An HIV-1 classification  
15 system has been published by Los Alamos National Laboratory (HIV Sequence Compendium-2001, Kuiken et al, published by Theoretical Biology and Biophysics Group T-10, Los Alamos, NM, (2001); <http://hiv-web.lanl.gov>) .

The compositions of the invention can also include a vector (e.g., a plasmid vector) encoding: (a) a Gag protein in which one or both zinc fingers have been inactivated; (b) a  
20 Pol protein in which (i) the integrase activity has been inhibited by deletion of some or all of the *pol* sequence, (ii) the polymerase, strand transfer and/or RNase H activity of reverse transcriptase has been inhibited by one or more point mutations within the *pol* sequence and (iii) the proteolytic activity of the protease has been inhibited by one or more point mutations; and (c) Env, Tat, Rev, and Vpu, with or without mutations. As noted above,  
25 proteolytic activity can be inhibited by introducing a mutation at positions 1641-1643 of SEQ ID NO:8 or at an analogous position in the sequence of another HIV clade. For example, the plasmids can contain the inserts described herein as JS7, IC25, and IN3. As is true for plasmids encoding other antigens, plasmids encoding the antigens just described can be combined with (e.g., mixed with) other plasmids that encode antigens obtained from, or  
30 derived from, a different HIV clade (or subtype or recombinant form thereof). The inserts *per se* (*sans* vector) are also within the scope of the invention.



Other vectors of the invention include plasmids encoding a Gag protein (*e.g.*, a Gag protein in which one or both of the zinc fingers have been inactivated); a Pol protein (*e.g.*, a Pol protein in which integrase, RT, and/or protease activities have been inhibited; a Vpu protein (which may be encoded by a sequence having a mutant start codon); and Env, Tat, and/or Rev proteins (in a wild type or mutant form). As is true for plasmids encoding other antigens, plasmids encoding the antigens just described can be combined with (*e.g.*, mixed with) other plasmids that encode antigens obtained from, or derived from, a different HIV clade (or subtype or recombinant form thereof). The inserts *per se* (*sans* vector) are also within the scope of the invention.

The plasmids described above, including those that express the JS2 or JS7 series of clade B HIV-1 sequences, can be administered to any subject, but may be most beneficially administered to subjects who have been, or who are likely to be, exposed to an HIV of clade B (the same is true for vectors other than plasmid vectors). Similarly, plasmids or other vectors that express an IN series of clade C HIV-1 sequences can be administered to a subject who has been, or who may be, exposed to an HIV of clade C. As vectors expressing antigens of various clades can be combined to elicit an immune response against more than one clade (this can be achieved whether one vector expresses multiple antigens from different clades or multiple vectors express single antigens from different clades), one can tailor the vaccine formulation to best protect a given subject. For example, if a subject is likely to be exposed to regions of the world where clades other than clade B predominate, one can formulate and administer a vector or vectors that express an antigen (or antigens) that will optimize the elicitation of an immune response to the predominant clade or clades.

The antigens they express are not the only parts of the plasmid vectors that can vary. Useful plasmids may or may not contain a terminator sequence that substantially inhibits transcription (the process by which RNA molecules are formed upon DNA templates by complementary base pairing). Useful terminator sequences include the lambda T0 terminator and functional fragments or variants thereof. The terminator sequence is positioned within the vector in the same orientation and at the C terminus of any open reading frame that is expressed in prokaryotes (*i.e.*, the terminator sequence and the open reading frame are operably linked). By preventing read through from the selectable marker into the vaccine

insert as the plasmid replicates in prokaryotic cells, the terminator stabilizes the insert as the bacteria grow and the plasmid replicates.

Selectable marker genes are known in the art and include, for example, genes encoding proteins that confer antibiotic resistance on a cell in which the marker is expressed (e.g., resistance to kanamycin, ampicillin, or penicillin). The selectable marker is so-named because it allows one to select cells by virtue of their survival under conditions that, absent the marker, would destroy them. The selectable marker, the terminator sequence, or both (or parts of each or both) can be, but need not be, excised from the plasmid before it is administered to a patient. Similarly, plasmid vectors can be administered in a circular form, after being linearized by digestion with a restriction endonuclease, or after some of the vector “backbone” has been altered or deleted.

The nucleic acid vectors can also include an origin of replication (e.g., a prokaryotic *ori*) and a transcription cassette that, in addition to containing one or more restriction endonuclease sites, into which an antigen-encoding insert can be cloned, optionally includes a promoter sequence and a polyadenylation signal. Promoters known as strong promoters can be used and may be preferred. One such promoter is the cytomegalovirus (CMV) intermediate early promoter, although other (including weaker) promoters may be used without departing from the scope of the present invention. Similarly, strong polyadenylation signals may be selected (e.g., the signal derived from a bovine growth hormone (BGH) encoding gene, or a rabbit  $\beta$  globin polyadenylation signal (Bohm *et al.*, *J. Immunol. Methods* 193:29-40, 1996; Chapman *et al.*, *Nucl. Acids Res.* 19:3979-3986, 1991; Hartikka *et al.*, *Hum. Gene Therapy* 7:1205-1217, 1996; Manthorpe *et al.*, *Hum. Gene Therapy* 4:419-431, 1993; Montgomery *et al.*, *DNA Cell Biol.* 12:777-783, 1993)).

The vectors can further include a leader sequence (a leader sequence that is a synthetic homolog of the tissue plasminogen activator gene leader sequence (tPA) is optional in the transcription cassette) and/or an intron sequence such as a cytomegalovirus (CMV) intron A or an SV40 intron. The presence of intron A increases the expression of many antigens from RNA viruses, bacteria, and parasites, presumably by providing the expressed RNA with sequences that support processing and function as a eukaryotic mRNA. Expression can also be enhanced by other methods known in the art including, but not limited to, optimizing the codon usage of prokaryotic mRNAs for eukaryotic cells (Andre

*et al.*, *J. Virol.* 72:1497-1503, 1998; Uchijima *et al.*, *J. Immunol.* 161:5594-5599, 1998).

Multi-cistronic vectors may be used to express more than one immunogen or an immunogen and an immunostimulatory protein (Iwasaki *et al.*, *J. Immunol.* 158:4591-4601, 1997a; Wild *et al.*, *Vaccine* 16:353-360, 1998). Thus (and as is true with other optional components of the vector constructs), vectors encoding one or more antigens from one or more HIV clades or isolates may, but do not necessarily, include a leader sequence and an intron (*e.g.*, the CMV intron A).

The vectors of the present invention differ in the sites that can be used for accepting antigen-encoding sequences and in whether the transcription cassette includes intron A sequences in the CMVIE promoter. Accordingly, one of ordinary skill in the art may modify the insertion site(s) or cloning site(s) within the plasmid without departing from the scope of the invention. Both intron A and the tPA leader sequence have been shown in certain instances to enhance antigen expression (Chapman *et al.*, *Nucleic Acids Research* 19:3979-3986, 1991).

As described further below, the vectors of the present invention can be administered with an adjuvant, including a genetic adjuvant. Accordingly, the nucleic acid vectors, regardless of the antigen they express, can optionally include such genetic adjuvants as GM-CSF, IL-15, IL-2, interferon response factors, secreted forms of flt-3, and mutated caspase genes. Genetic adjuvants can also be supplied in the form of fusion proteins, for example by fusing one or more C3d gene sequences (*e.g.*, 1-3 (or more) C3d gene sequences) to an expressed antigen.

In the event the vector administered is a pGA vector, it can comprise the sequence of, for example, pGA1 (SEQ ID NO:1) or derivatives thereof (*e.g.*, SEQ ID NOs:2 and 3) or pGA2 (SEQ ID NO:4) or derivatives thereof (*e.g.*, SEQ ID NOs:5 and 6). The pGA vectors are described in more detail here (see also Examples 1-8). pGA1 is a 3897 bp plasmid that includes a promoter (bp 1-690), the CMV-intron A (bp 691-1638), a synthetic mimic of the tPA leader sequence (bp 1659 – 1721), the bovine growth hormone polyadenylation sequence (bp1761-1983), the lambda T0 terminator (bp 1984-2018), the kanamycin resistance gene (bp 2037-2830) and the ColEI replicator (bp 2831-3890). The DNA sequence of the pGA1 construct (SEQ ID NO:1) is shown in Fig. 2. In Fig. 1, the indicated restriction sites are useful for cloning antigen-encoding sequences. The *Cla* I or *Bsp*D I sites

are used when the 5' end of a vaccine insert is cloned upstream of the tPA leader. The *Nhe* I site is used for cloning a sequence in frame with the tPA leader sequence. The sites listed between *Sma* I and *Bln* I are used for cloning the 3' terminus of an antigen-encoding sequence.

5 pGA2 is a 2947 bp plasmid lacking the 947 bp of intron A sequences found in pGA1. pGA2 is the same as pGA1, except for the deletion of intron A sequences. pGA2 is valuable for cloning sequences which do not require an upstream intron for efficient expression, or for cloning sequences in which an upstream intron might interfere with the pattern of splicing needed for good expression. Fig. 5 presents a schematic map of pGA2 with useful restriction  
10 sites for cloning vaccine inserts. Fig. 6a shows the DNA sequence of pGA2 (SEQ ID NO:2). The use of restriction sites for cloning vaccine inserts into pGA2 is the same as that used for cloning fragments into pGA1. pGA2.1 and pGA2.2 are multiple cloning site derivatives of pGA2. Figs. 7a and 8a show the DNA sequence of pGA2.1 (SEQ ID NO:5) and pGA2.2 (SEQ ID NO:6) respectively.

15 pGA plasmids having "backbone" sequences that differ from those disclosed herein are also within the scope of the invention so long as the plasmids retain substantially all of the characteristics necessary to be therapeutically effective (e.g., one can substitute nucleotides, add nucleotides, or delete nucleotides so long as the plasmid, when administered to a patient, induces or enhances an immune response against a given or desired pathogen).  
20 For example, 1-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, 91-100, or more than 100 nucleotides can be deleted or replaced.

In one embodiment, the methods of the invention (e.g., methods of eliciting an immune response in a patient) can be carried out by administering to the patient a therapeutically effective amount of a physiologically acceptable composition that includes a  
25 vector, which can contain a vaccine insert that encodes one or more antigens that elicit an immune response against an HIV. The vector can be a plasmid vector having one or more of the characteristics of the pGA constructs described above (e.g., a selectable marker gene, a prokaryotic origin of replication, a termination sequence (e.g., the lambda T0 terminator) and operably linked to the selectable gene marker, and a eukaryotic transcription cassette  
30 comprising a promoter sequence, a nucleic acid insert encoding at least one antigen derived from an immunodeficiency virus, and a polyadenylation signal sequence). Of course, the

vaccine inserts of the invention may be delivered by plasmid vectors that do not have the characteristics of the pGA constructs (*e.g.*, vectors other than pGA1 or pGA2).

Alternatively, the composition can include any viral or bacterial vector that includes an insert described herein. The invention, therefore, encompasses administration of a single type of vector (*i.e.*, plasmid or viral vectors that contain the same vaccine insert (*i.e.*, an insert encoding the same antigens)). As is made clear elsewhere, the patient may receive two types of vectors, and each of those vectors can elicit an immune response against an HIV of a different clade. For example, the invention features methods in which a patient receives a composition that includes (a) a first vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against a human immunodeficiency virus (HIV) of a first subtype or recombinant form and (b) a second vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form. The first and second vectors can be any of those described herein. Similarly, the inserts in the first and second vectors can be any of those described herein.

A therapeutically effective amount of a vector (whether considered the first, second, third, *etc.* vector) can be administered by an intramuscular or an intradermal route, together with a physiologically acceptable carrier, diluent, or excipient, and, optionally, an adjuvant. A therapeutically effective amount of the same or a different vector can subsequently be administered by an intramuscular or an intradermal route, together with a physiologically acceptable carrier, diluent, or excipient, and, optionally, an adjuvant to boost an immune response. Such components can be readily selected by one of ordinary skill in the art, regardless of the precise nature of the antigens incorporated in the vaccine or the vector by which they are delivered.

The methods of eliciting an immune response can be carried out by administering only the plasmid vectors of the invention, by administering only the viral vectors or recombinant virus of the invention, or by administering both (*e.g.*, one can administer a plasmid vector (or a mixture or combination of plasmid vectors)) to “prime” the immune response and a viral vector or recombinant virus (or a mixture or combination of viral vectors)) to “boost” the immune response. Where plasmid and viral vectors or recombinant virus are administered, their inserts may be “matched.” To be “matched,” one or more of the

sequences of the inserts (*e.g.*, the sequences encoding Gag, or the sequences encoding Env, *etc.*) within the plasmid and viral vectors within the recombinant virus may be identical, but the term is not so limited. "Matched" sequences can also differ from one another. For example, inserts expressed by viral vectors are "matched" to those expressed by DNA  
5 vectors when the sequences used in the DNA vector are mutated or further mutated to allow (or optimize) replication of a viral vector that encodes those sequences and expression of the encoded antigens (*e.g.*, Gag, Gag-Pol, or Env) in cells infected with the viral vector.

At least some of the immunodeficiency virus vaccine inserts of the present invention were designed to generate non-infectious VLPs (a term that can encompass true VLPs as well  
10 as aggregates of viral proteins) from a single DNA. This was achieved using the subgenomic splicing elements normally used by immunodeficiency viruses to express multiple gene products from a single viral RNA. The subgenomic splicing patterns are influenced by (i) splice sites and acceptors present in full length viral RNA, (ii) the Rev responsive element (RRE) and (iii) the Rev protein. The splice sites in retroviral RNAs use the canonical  
15 sequences for splice sites in eukaryotic RNAs. The RRE is an approximately 200 bp RNA structure that interacts with the Rev protein to allow transport of viral RNAs from the nucleus to the cytoplasm. In the absence of Rev, the approximately 10 kb RNA of immunodeficiency virus mostly undergoes splicing to the mRNAs for the regulatory genes Tat, Rev, and Nef. These genes are encoded by exons present between RT and Env and at  
20 the 3' end of the genome. In the presence of Rev, the singly spliced mRNA for Env and the unspliced mRNA for Gag and Pol are expressed in addition to the multiply spliced mRNAs for Tat, Rev, and Nef.

The expression of non-infectious VLPs from a single DNA affords a number of advantages to an immunodeficiency virus vaccine. The expression of a number of proteins  
25 from a single DNA affords the vaccinated host the opportunity to respond to the breadth of T- and B-cell epitopes encompassed in these proteins. The expression of proteins containing multiple epitopes allows epitope presentation by diverse histocompatibility types. By using whole proteins, one offers hosts of different histocompatibility types the opportunity to raise broad-based T cell responses. This may be essential for the effective containment of  
30 immunodeficiency virus infections, whose high mutation rate supports ready escape from immune responses (Evans *et al.*, *Nat. Med.* 5:1270-1276, 1999; Poignard *et al.*, *Immunity*

10:431-438, 1999, Evans *et al.*, 1995). In the context of the present vaccination scheme, just as in drug therapy, multi-epitope T cell responses that require multiple mutations for escape will provide better protection than single epitope T cell responses (which require only a single mutation for escape).

5 Immunogens can also be engineered to be more or less effective for raising antibody or Tc by targeting the expressed antigen to specific cellular compartments. For example, antibody responses are raised more effectively by antigens that are displayed on the plasma membrane of cells, or secreted therefrom, than by antigens that are localized to the interior of cells (Boyle *et al.*, *Int. Immunol.* 9:1897-1906, 1997; Inchauspe *et al.*, *DNA Cell. Biol.* 16:185-195, 1997). Tc responses may be enhanced by using N-terminal ubiquitination  
10 signals which target the DNA-encoded protein to the proteosome causing rapid cytoplasmic degradation and more efficient peptide loading into the MHC I pathway (Rodriguez *et al.*, *J. Virol.* 71:8497-8503, 1997; Tobery *et al.*, *J. Exp. Med.* 185:909-920, 1997; Wu *et al.*, *J. Immunol.* 159:6037-6043, 1997). For a review on the mechanistic basis for DNA-raised  
15 immune responses, refer to Robinson and Pertmer, *Advances in Virus Research*, vol. 53, Academic Press (2000).

Another approach to manipulating immune responses is to fuse immunogens to immunotargeting or immunostimulatory molecules. To date, the most successful of these fusions have targeted secreted immunogens to antigen-presenting cells (APCs) or lymph  
20 nodes (Boyle *et al.*, *Nature* 392:408-411, 1998). Accordingly, the invention features the HIV antigens described herein fused to immunotargeting or immunostimulatory molecules such as CTLA-4, L-selectin, or a cytokine (*e.g.*, an interleukin such as IL-1, IL-2, IL-4, IL-7, IL-10, IL-15, or IL-21). Nucleic acids encoding such fusions and compositions containing them (*e.g.*, vectors and physiologically acceptable preparations) are also within the scope of the  
25 present invention.

DNA can be delivered in a variety of ways, any of which can be used to deliver the plasmids of the present invention to a subject. For example, DNA can be injected in, for example, saline (*e.g.*, using a hypodermic needle) or delivered biolistically (by, for example, a gene gun that accelerates DNA-coated beads). Saline injections deliver DNA into  
30 extracellular spaces, whereas gene gun deliveries bombard DNA directly into cells. The saline injections require much larger amounts of DNA (typically 100-1000 times more) than

the gene gun (Fynan *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11478-11482, 1993). These two types of delivery also differ in that saline injections bias responses towards type 1 T-cell help, whereas gene gun deliveries bias responses towards type 2 T-cell help (Feltquate *et al.*, *J. Immunol.* 158:2278-2284, 1997; Pertmer *et al.*, *J. Virol.* 70:6119-6125, 1996). DNAs injected in saline rapidly spread throughout the body. DNAs delivered by the gun are more localized at the target site. Following either method of inoculation, extracellular plasmid DNA has a short half life of about 10 minutes (Kawabata *et al.*, *Pharm. Res.* 12:825-830, 1995; Lew *et al.*, *Hum. Gene Ther.* 6:553, 1995). Vaccination by saline injections can be intramuscular (i.m.) or intradermal (i.d.); gene gun deliveries can be administered to the skin or to surgically exposed tissue such as muscle.

While other routes of delivery are generally less favored, they can nevertheless be used to administer the compositions of the invention. For example, the DNA can be applied to the mucosa or by a parenteral route of inoculation. Intranasal administration of DNA in saline has met with both good (Asakura *et al.*, *Scand. J. Immunol.* 46:326-330, 1997; Sasaki *et al.*, *Infect. Immun.* 66:823-826, 1998b) and limited (Fynan *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11478-82, 1993) success. The gene gun has successfully raised IgG following the delivery of DNA to the vaginal mucosa (Livingston *et al.*, *Ann. New York Acad. Sci.* 772:265-267, 1995). Some success at delivering DNA to mucosal surfaces has also been achieved using liposomes (McCluskie *et al.*, *Antisense Nucleic Acid Drug Dev.* 8:401-414, 1998), microspheres (Chen *et al.*, *J. Virol.* 72:5757-5761, 1998a; Jones *et al.*, *Vaccine* 15:814-817, 1997) and recombinant Shigella vectors (Sizemore *et al.*, *Science* 270:299-302, 1995; Sizemore *et al.*, *Vaccine* 15:804-807, 1997). Agents such as these (liposomes, microspheres and recombinant Shigella vectors) can be used to deliver the nucleic acids of the present invention.

The dose of DNA needed to raise a response depends upon the method of delivery, the host, the vector, and the encoded antigen. The method of delivery may be the most influential parameter. From 10 µg to 5mg of DNA is generally used for saline injections of DNA, whereas from 0.2 µg to 20 µg of DNA is used more typically for gene gun deliveries of DNA. In general, lower doses of DNA are used in mice (10-100 µg for saline injections and 0.2 µg to 2 µg for gene gun deliveries), and higher doses in primates (100 µg to 1 mg for saline injections and 2 µg to 20 µg for gene gun deliveries). The much lower amount of



DNA required for gene gun deliveries reflect the gold beads directly delivering DNA into cells.

In addition to the DNA vectors described above, a number of different poxviruses can be used either alone (*i.e.*, without a nucleic acid or DNA prime) or as the boost component of a vaccine regimen. MVA has been particularly effective in mouse models (Schneider *et al.*, *Nat. Med.* 4:397-402, 1998). MVA is a highly attenuated strain of vaccinia virus that was developed toward the end of the campaign for the eradication of smallpox, and it has been safety tested in more than 100,000 people (Mahnel *et al.*, *Berl. Munch Tierarztl Wochenschr* 107:253-256, 1994; Mayr *et al. Zentralbl. Bakteriologie* 167:375-390, 1978). During over 500 passages in chicken cells, MVA lost about 10% of its genome and the ability to replicate efficiently in primate cells. Despite its limited replication, MVA has proved to be a highly effective expression vector (Sutter *et al.*, *Proc. Natl. Acad. Sci. USA* 89:10847-10851, 1992), raising protective immune responses in primates for parainfluenza virus (Durbin *et al. J. Infect. Dis.* 179:1345-1351, 1999), measles (Stittelaar *et al. J. Virol.* 74:4236-4243, 2000), and immunodeficiency viruses (Barouch *et al.*, *J. Virol.* 75:5151-5158, 2001; Ourmanov *et al.*, *J. Virol.* 74:2740-2751, 2000; Amara *et al.*, *J. Virol.* 76:7625-7631, 2002). The relatively high immunogenicity of MVA has been attributed in part to the loss of several viral anti-immune defense genes (Blanchard *et al.*, *J. Gen. Virol.* 79:1159-1167, 1998).

Vaccinia viruses have been used to engineer viral vectors for recombinant gene expression and as recombinant live vaccines (Mackett *et al.*, *Proc. Natl. Acad. Sci. USA* 79:7415-7419; Smith *et al.*, *Biotech. Genet. Engin. Rev.* 2:383-407, 1984). DNA sequences, which may encode any of the HIV antigens described herein, can be introduced into the genomes of vaccinia viruses. If the gene is integrated at a site in the viral DNA that is non-essential for the life cycle of the virus, it is possible for the newly produced recombinant vaccinia virus to be infectious (*i.e.*, able to infect foreign cells) and to express the integrated DNA sequences. Preferably, the viral vectors featured in the compositions and methods of the present invention are highly attenuated. Several attenuated strains of vaccinia virus were developed to avoid undesired side effects of smallpox vaccination. The modified vaccinia Ankara (MVA) virus was generated by long-term serial passages of the Ankara strain of vaccinia virus on chicken embryo fibroblasts (CVA; *see* Mayr *et al.*, *Infection* 3:6-14, 1975). The MVA virus is publicly available from the American Type Culture Collection (ATCC;

No. VR-1508; Manassas, VA). The desirable properties of the MVA strain have been demonstrated in clinical trials (Mayr *et al.*, *Zentralbl. Bakteriolog.* 167:375-390, 1978; Stickl *et al.*, *Dtsch. Med. Wschr.* 99:2386-2392, 1974; *see also*, Sutter and Moss, *Proc. Natl. Acad. Sci. USA* 89:10847-10851, 1992). During these studies in over 120,000 humans, including  
5 high-risk patients, no side effects were associated with the use of MVA vaccine.

The MVA vectors can be prepared as follows. A DNA construct that contains a DNA sequence, *e.g.*, an insert sequence described herein, that encodes a foreign polypeptide (*e.g.*, any of the HIV antigens described herein) and that is flanked by MVA DNA sequences adjacent to a naturally occurring deletion with the MVA genome (*e.g.*, deletion III or other  
10 non-essential site(s); six major deletions of genomic DNA (designated deletions I, II, III, IV, V, and VI) totaling 31,000 base pairs have been identified (Meyer *et al.*, *J. Gen. Virol.* 72:1031-1038, 1991)) is introduced into cells infected with MVA under conditions that permit homologous recombination to occur. Once the DNA construct has been introduced into the eukaryotic cell and the foreign DNA has recombined with the viral DNA, the  
15 recombinant vaccinia virus can be isolated by methods known in the art (isolation can be facilitated by use of a detectable marker). The DNA constructed to be inserted can be linear or circular (*e.g.*, a plasmid, linearized plasmid, gene, gene fragment, or modified HIV genome). The foreign DNA sequence is inserted between the sequences flanking the naturally occurring deletion. For better expression of a DNA sequence, the sequence can  
20 include regulatory sequences (*e.g.*, a promoter, such as the promoter of the vaccinia 11 kDa gene or the 7.5 kDa gene). The DNA construct can be introduced into MVA-infected cells by a variety of methods, including calcium phosphate-assisted transfection (Graham *et al.*, *Virol.* 52:456-467, 1973 and Wigler *et al.*, *Cell* 16:777-785, 1979), electroporation (Neumann *et al.*, *EMBO J.* 1:841-845, 1982) microinjection (Graessmann *et al.*, *Meth.*  
25 *Enzymol.* 101:482-492, 1983), by means of liposomes (Straubinger *et al.*, *Meth. Enzymol.* 101:512-527, 1983), by means of spheroplasts (Schaffner, *Proc. Natl. Acad. Sci. USA* 77:2163-2167, 1980), or by other methods known in the art. The preparation of recombinant MVA virus is described in greater detail in Moss *et al.* PCT Publication WO 02/07275 A2 (International Application No. PCT/US02/06713)

30 One can arrive at an appropriate dosage when delivering DNA by way of a viral vector, just as one can when a plasmid vector is used. For example, one can deliver  $1 \times 10^8$

pfu of an MVA-based vaccine, and administration can be carried out intramuscularly, intradermally, intravenously, or mucosally.

Accordingly, the invention features a composition comprising: (a) a first viral vector comprising a vaccine insert encoding one or more antigens that elicit an immune response  
5 against a human immunodeficiency virus (HIV) of a first subtype or recombinant form and (b) a second viral vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form. The viral vector can be a recombinant poxvirus or a modified vaccinia Ankara (MVA) virus, and the insert can be any of the HIV antigens described herein from any clade (*e.g.*, one can  
10 administer a prophylactically or therapeutically effective amount of an MVA that encodes a clade A, B, or C HIV (*e.g.*, HIV-1 antigen). Moreover, when administered in conjunction with a plasmid vector (*e.g.*, when administered subsequent to a “DNA prime”), the MVA-borne sequence can be “matched” to the plasmid-borne sequence. For example, a vaccinia virus (*e.g.*, MVA) that expresses a recombinant clade B sequence can be matched to the JS  
15 series of plasmid inserts. Similarly, a vaccinia virus (*e.g.*, MVA) that expresses a recombinant clade A sequence can be matched to the IC series of plasmid inserts; a vaccinia virus (*e.g.*, MVA) that expresses a recombinant clade C sequence can be matched to the IN series of plasmid inserts. While particular clades are exemplified below, the invention is not so limited. The compositions that contain a viral vector, can include viral vectors that  
20 express an HIV antigen from any known clade (including clades A, B, C, D, E, F, G, H, I, J, K or L). Methods of eliciting an immune response can, of course, be carried out with compositions expressing antigens from any of these clades as well.

Either the plasmid or viral vectors described here can be administered with an adjuvant (*i.e.*, any substance that is added to a vaccine to increase the vaccine’s  
25 immunogenicity) and they can be administered by any conventional route of administration (*e.g.*, intramuscular, intradermal, intravenous or mucosally; *see below*). The adjuvant used in connection with the vectors described here (whether DNA or viral-based) can be one that slowly releases antigen (*e.g.*, the adjuvant can be a liposome), or it can be an adjuvant that is strongly immunogenic in its own right (these adjuvants are believed to function  
30 synergistically). Accordingly, the vaccine compositions described here can include known adjuvants or other substances that promote DNA uptake, recruit immune system cells to the

site of the inoculation, or facilitate the immune activation of responding lymphoid cells.

These adjuvants or substances include oil and water emulsions, *Corynebacterium parvum*, *Bacillus Calmette Guerin*, aluminum hydroxide, glucan, dextran sulfate, iron oxide, sodium alginate, Bacto-Adjuvant, certain synthetic polymers such as poly amino acids and co-

5 polymers of amino acids, saponin, REGRESSIN (Vetrepharm, Athens, GA.), AVRIDINE (N, N-dioctadecyl-N',N'-bis(2-hydroxyethyl)-propanediamine), paraffin oil, and muramyl dipeptide. Genetic adjuvants, which encode immunomodulatory molecules on the same or a co-inoculated vector, can also be used. For example, GM-CSF, IL-15, IL-2, interferon response factors, and mutated caspase genes can be included on a vector that encodes a  
10 pathogenic immunogen (such as an HIV antigen) or on a separate vector that is administered at or around the same time as the immunogen is administered. Expressed antigens can also be fused to an adjuvant sequence such as one, two, three or more copies of C3d.

The compositions described herein can be administered in a variety of ways including through any parenteral or topical route. For example, an individual can be inoculated by

15 intravenous, intraperitoneal, intradermal, subcutaneous or intramuscular methods.

Inoculation can be, for example, with a hypodermic needle, needleless delivery devices such as those that propel a stream of liquid into the target site, or with the use of a gene gun that bombards DNA on gold beads into the target site. The vector comprising the pathogen vaccine insert can be administered to a mucosal surface by a variety of methods including  
20 intranasal administration, i.e., nose drops or inhalants, or intrarectal or intravaginal administration by solutions, gels, foams, or suppositories. Alternatively, the vector comprising the vaccine insert can be orally administered in the form of a tablet, capsule, chewable tablet, syrup, emulsion, or the like. In an alternate embodiment, vectors can be administered transdermally, by passive skin patches, iontophoretic means, and the like.

25 Any physiologically acceptable medium can be used to introduce a vector (whether nucleic acid-based or live-vectored) comprising a vaccine insert into a patient. For example, suitable pharmaceutically acceptable carriers known in the art include, but are not limited to, sterile water, saline, glucose, dextrose, or buffered solutions. The media may include auxiliary agents such as diluents, stabilizers (*i.e.*, sugars (glucose and dextrose were noted  
30 previously) and amino acids), preservatives, wetting agents, emulsifying agents, pH buffering agents, additives that enhance viscosity or syringability, colors, and the like. Preferably, the

The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patent applications and patents cited throughout the present application  
5 are hereby incorporated by reference in their entirety. A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention.

## 10

25

30

The ColE1 replicator, the kanamycin resistance gene, and the transcriptional control elements for eukaryotic cells were combined in one plasmid using PCR fragments from the commercial vector pZErO-2.1 (Invitrogen, Carlsbad, CA) and a eukaryotic expression vector pJW4303 (Lu *et al.*, *Vaccine* 15:920-923, 1997).

5 An 1859 bp fragment from pZErO-2.1 (nucleotides 1319 to 3178) included the ColE1 origin of replication and the kanamycin resistance gene. A 2040 bp fragment from pJW4303 (nucleotides 376 to 2416) included the CMVIE promoter with intron A, a synthetic homolog of the tissue plasminogen activator leader (tPA), and the bovine growth hormone polyadenylation site (BGHpA). Fragments were amplified by polymerase chain reaction  
10 (PCR) with oligonucleotide primers containing *Sal* I sites. A ligation product with the transcription cassettes for kanamycin resistance from pZErO2 and the eukaryotic transcription cassette from pJW4303 in opposite transcriptional orientations, was identified for further development. Nucleotide numbering for this parent of the pGA vectors was started from the first bp of the 5' end of the CMV promoter.

15 The T0 terminator was introduced into this parent for the pGA vectors by PCR amplification of a 391 bp fragment with a *Bam*H I restriction endonuclease site at its 5' end and an *Xba* I restriction endonuclease site at its 3' end. The initial 355 bp of the fragment were sequences in the BGHpA sequence derived from the pJW4303 transcription cassette, the next 36 bases in a synthetic oligonucleotide introduced the T0 sequence and the *Xba* I site.

20 The introduced T0 terminator sequences comprised the sequence: 5'-ATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAA-3' (SEQ ID NO: 16).

The T0 terminator containing the *Bam*H I-*Xba* I fragment was substituted for the homologous fragment without the T0 terminator in the plasmid created from pZErO-2 and pJW4303. The product was sequenced to verify the T0 orientation (Fig. 1).

25 A region in the eukaryotic transcription cassette between nucleotides 1755-1845 contained the last 30 bp of the reading frame for SIV *nef*. This region was removed from pGA by mutating the sequence at nt 1858 and generating an *Avr* II restriction endonuclease site. A naturally occurring *Avr* II site is located at nt 1755. Digestion with *Avr* II enzyme and then religation with T4 DNA ligase allowed for removal of the SIV segment of DNA  
30 between nucleotides 1755-1845. To facilitate cloning of HIV-1 sequences into pGA vectors,

a *Cla* I site was introduced at bp1648 and an *Rsr* II site at bp 1747 using standard techniques for site directed mutagenesis. Constructions were verified by sequence analyses.

#### Example 2: pGA1.1

5 pGA1.1 (SEQ ID NO: 2) is identical to pGA1 except that the multiple cloning site has been altered to include an EcoRI site. This was accomplished by site directed mutagenesis using the following primers: 5'-GCTGCTGCTGTGTGGAGAATTCTTCGTTTCGGC-3' (forward) and 5'-GCCGAAACGAAGAATTCTCCACACAGCAGCAGC-3' (reverse) (SEQ ID  
10 NOs:17 and 18 respectively). Accordingly, the pGA1.1 vector is an embodiment of the invention; as are other vectors having one or more of the features or characteristics of a pGA plasmid (see the detailed description), but different restriction endonuclease sites in the multi-cloning site (e.g., the invention encompasses plasmids that are otherwise substantially similar to pGA1 but that have more, less, or different restriction endonuclease sites in their  
15 multi-cloning site).

#### Example 3: pGA1.2

pGA1.2 (SEQ ID NO: 3) is identical to pGA1.1 except that the multiple cloning site has been altered to include BamHI and XhoI sites 5' to the EcoRI site. This was  
20 accomplished by site directed mutagenesis using the primer 5'-CTGCAGTCACCATGGATCCTTGCACT-CGAGGATGCAATGAAGAG - 3' (SEQ ID NO:19) and the reverse primer 5'-CTCTTCATTGCATCCTCGAGTGCAAGGATCCATGGTGAAGTGCAG - 3' (SEQ ID NO:20).

25

#### Example 4: pGA2

pGA2 is schematically illustrated in Fig. 5, and its nucleotide sequence is shown in Fig. 6 (SEQ ID NO: 4). pGA2 is identical to pGA1 except that the intron A sequence has been deleted from the CMV promoter of pGA2. pGA2 was created from pGA1 by introducing a  
30 *Cla* I site 8 bp downstream from the mRNA cap site in the CMV promoter; the *Cla* I site was introduced using oligonucleotide-directed mutagenesis using complimentary primers having

the SEquences: 5'-CCGTCAGATCGCATCGATACGCCATCCACG-3' (SEQ ID NO: 19) and 5'-CGTGGATGGCGTATCGATGCGATCTGACGG-3' (SEQ ID NO: 20). After insertion of the new *Cla* I site, pGA1 was digested with *Cla* I to remove the 946 bp *Cla* I fragment from pGA1, and then religated to yield pGA2.

5

#### Example 5: pGA2.1

PGA2.1 (SEQ ID NO: 5) is identical to pGA2 except that the multiple cloning site has been altered to include an EcoRI sites. This was accomplished by site directed mutagenesis using using the following primers:

10 forward 5'-GCTGCTGCTGTGTGGAGAATTCTTCGTTTCGGC-3' (SEQ ID NO:17) and reverse 5'-GCCGAAACGAAGAATTCTCCACACAGCAGCAGC-3' (SEQ ID NO:18). Accordingly, the pGA2.1 vector is an embodiment of the invention; as are other vectors having one or more of the features or characteristics of a pGA plasmid (see the detailed description), but different restriction endonuclease sites in the multi-cloning site (e.g., the  
15 invention encompasses plasmids that are otherwise substantially similar to pGA1 but that have more, less, or different restriction endonuclease sites in their multi-cloning site).

#### Example 6: pGA2.2

PGA2.2 (SEQ ID NO: 6) is identical to pGA1.1 except that the multiple cloning site  
20 has been altered to include a BamHI and a XhoI site 5' to the EcoRI site. This was accomplished by site directed mutagenesis using the forward primer 5'-GAATCATTCTATGGATCCTTGC-TCGAGTGGATGCAATGAAGAG - 3' and the reverse primer 5'-CTCTTCATTGCATC-CACTCGAGCAAGGATCCATAGAATGAGTTC-3' (SEQ ID NOs:23 and 24 respectively)

25

#### Example 7: Immunodeficiency Virus Vaccine Inserts

HIV-1 vaccine inserts for the pGA1 and pGA2 series of vectors were constructed to express multiple HIV-1 proteins from a single RNA transcript using the same subgenomic splicing mechanisms used by immunodeficiency viruses. To ensure that these multiprotein-  
30 expressing vectors did not form infectious virus, deletions and point mutations were introduced to cripple essential steps in the retrovirus life cycle. Fig. 9 presents schematics of



the normal retroviral genome and a representative vaccine insert. Regions that have been deleted in the insert are stippled. X's indicate point mutations. The deletions included both of the long terminal repeat (LTR) sequences that encode cis-acting elements for reverse transcription, integration, and expression of proviral DNA. 5' sequences adjacent to the 5' LTR that promote encapsidation of viral RNA have been deleted. Coding sequences for the region of *pol* encoding integrase as well as the auxiliary genes *vif* and *vpr* have been deleted. And finally, *nef*, a gene encoding the Nef regulatory protein has been deleted. The seven point mutations that are common to all inserts described in the examples below are included in the schematic. These include four mutations in the zinc fingers in the nucleocapsid protein to limit zinc-finger-mediated packaging of viral RNA and three mutations in reverse transcriptase to prevent reverse transcription of viral RNA. Analogous changes can be made in any vaccine insert that includes *gag* and/or *pol*. Moreover, these changes (or analogous changes) can be made in vaccine inserts that are placed in any of the plasmid or live-vectored vaccines described herein (*i.e.*, in any plasmid having one or more of the features or characteristics of the pGA vectors, the pGA vectors themselves, or the vaccinia vectors that may be used alone or in conjunction with (*e.g.*, to boost) a DNA-primed patient).

The HIV-1 vaccine inserts described below can be expressed in any of the pGA vectors or further derivatives of these vectors. The examples for inserts that are given below are given with the example of the pGA vector that is planned for future use of that insert.

However, any of these inserts can be used in any of the pGA vectors as well as other eukaryotic expression vectors.

#### **Example 8: pGA2/JS2, multiprotein clade B HIV-1 insert**

The sequence of pGA2/JS2 is shown in Fig. 7a (SEQ ID NO:7), its functional regions and the origins of these regions in Fig. 7b and the positions of its point mutations in Fig. 7c. The JS2 insert described here was designed with clade B HIV-1 sequences so that it would elicit an immune response against HIV-1 sequences that are endemic in the United States, Europe, and Japan. As noted above, any clade B isolate can be used as a reasonable representative for other clade B isolates. Since HIV-1 isolates use different chemokine receptors as co-receptors, and the vast majority of viruses that are undergoing transmission use the CCR-5 co-receptor (Berger, *AIDS* 11(Suppl A):S3-16, 1997), the vaccine insert we

designed had a CCR-5-using Env. Of course, Envs that function through any other co-receptor or that have been constructed from naturally occurring or synthetic sequences so as to increase immunogenicity can be made and used as well.

To achieve a multiprotein-expressing clade B vaccine insert with high expression, candidate vaccines were constructed from seven different HIV-1 sequences, as shown in Table 1.

Table 1: Comparison of candidate vaccine inserts

| Plasmid designation | SEquences tested       | Ability to grow plasmid | Expression of Gag | Expression of Env | Comment                              |
|---------------------|------------------------|-------------------------|-------------------|-------------------|--------------------------------------|
| BH10-VLP            | BH10                   | Good                    | Good              | Good              | X4 Env                               |
| 6A-VLP              | 6A env in BH10-VLP     | Poor                    | Not tested        | not tested        |                                      |
| BAL-VLP             | BAL env in BH10-VLP    | Good                    | Poor              | Poor              |                                      |
| ADA-VLP             | ADA env in BH10-VLP    | Good                    | Good              | Good              | chosen for vaccine, renamed pGA1/JS1 |
| CDC-A-VLP           | CDC-A env in BH10-VLP  | Good                    | Good              | Poor              |                                      |
| CDC-B-VLP           | CDC-B-env in BH10-VLP  | Good                    | Good              | Good              | not as favorable expression as ADA   |
| CDC-C-VLP           | CDC -C env in BH10-VLP | Good                    | Good              | Good              | not as favorable expression as ADA   |

An initial construct, pBH10-VLP, was prepared from IIIB sequences that are stable in bacteria and have high expression in eukaryotic cells. The HIV-1-BH10 sequences were obtained from the NIH-sponsored AIDS Repository (catalog #90). The parental pHIV-1-BH10 was used as the template for PCR reactions to construct pBH10-VLP.

Primers were designed to yield a Gag-Rt PCR product (5' PCR product) encompassing (from 5' to 3') 105 bp of the 5' untranslated leader sequence and sequences from the start codon for Gag to the end of the RT coding sequence. The oligonucleotide primers introduced a *Cla* I site at the 5' end of the PCR product and *EcoR* I and *Nhe* I sites at the 3' end of the PCR product. Sense primer (5'-GAGCTCTATCGATGCAGGACTCGGCTTGC-3' (SEQ ID NO:25) and antisense primer (5'-GGCAGGTTTTTAATCGCTAGCCTATGCTCTCC-3' (SEQ ID NO:26) were used to amplify the 5' PCR product.

The PCR product for the env region of HIV-1 (3' PCR product) encompassed the *vpu*, *tat*, *rev*, and *env* sequences and the splice acceptor sites necessary for proper processing and expression of their respective mRNAs. An *EcoR* I site was introduced at the 5' end of this product and *Nhe* I and *Rsr* II sites were introduced into the 3' end. Sense primer  
 5 (5'-GGGCAGGAGTGCTAGCC-3' (SEQ ID NO:27) and antisense primer  
 5'-CCACACTACTTTCGGACCGCTAGCCACCC-3' (SEQ ID NO: 28)) were used to amplify the 3' PCR product. The 5' PCR product was cloned into pGA1 at the *Cla* I and *Nhe* I sites of pGA1 and the identity of the construct confirmed by sequencing. The 3' PCR product was then inserted into the 5' clone at the *EcoR* I and *Nhe* I sites to yield pBH10. The  
 10 construction of this plasmid resulted in proviral sequences that lacked LTRs, integrase, *vif*, *vpr* and *nef* sequences (see Fig. 9).

Because pBH10-VLP encoded a CXCR-4 using Env, rather than a CCR-5 using Env, sequences encoding six different R5 Envs were substituted for *env* sequence in the pBH10 intermediate (Table 1). *EcoR* I to *BamH* I fragments encompassing *tat*, *rev*, *vpu* and *env*  
 15 coding sequences from different viral genomes were substituted into pBH10. The resulting *env* and *rev* sequences were chimeras for the substituted sequences and HIV-1-BH10 sequences (see Fig. 9). In the case of the HIV-1-ADA envelope, a *BamH* I site was introduced into the HIV-1-ADA sequence to facilitate substituting an *EcoR* I to *BamH* I fragment for the *EcoR* I to *BamH* I region of pBH10. The results of these constructions are  
 20 summarized in Table 1. Of the six sequences tested, one, the 6A-VLP gave poor plasmid growth in transformed bacteria. The plasmid 6A-VLP was not developed further. Among the other constructs, the pBH10/ADA chimera produced the best expression of viral Gag and Env proteins (Table 1). In transient transfections in 293T cells, the expression from the pBH10/ADA chimera was higher than that of wt proviruses for HIV-1-ADA or HIV-1-IIIB.  
 25 Expression was also higher than for a previous multiprotein-expressing HIV-1 vaccine (dpol) (Richmond *et al.*, *J. Virol.* 72:9092-9100, 1998) that had successfully primed cytotoxic T cell responses in rhesus macaques (Kent *et al.*, *J. Virol.* 72:10180-10188, 1998). The pBH10/ADA chimera was now designated JS1. It should be recognized that plasmids having any given or desired HIV-1 inserts can be similarly assessed.

30 Next, inactivating point mutations were introduced into JS1 to further increase the safety of this construct for use in humans as a non-infectious vaccine agent (of course,

mutations can be made preemptively, before any testing at all) (see Fig 10c). Four codon mutations were introduced into the Zinc fingers in nucleocapsid to limit the encapsidation of viral RNA and three codon mutations were introduced into the reverse transcriptase region of *pol* to inactivate the viral reverse transcriptase. The JS1 insert with these mutations was designated JS2.

The mutations were made using a site directed mutagenesis kit (Stratagene) following the manufacturer's protocol. All mutations were confirmed by sequencing. Primer pairs used for the mutagenesis were:

(A) 5'-GGTTAAGAGCTTCAATAGCGGCAAAGAAGGGC-3' (C392S, C395S; SEQ ID NO:29) and 5'-GCCCTTCTTTGCCGCTATTGAAGCTCTTAACC-3' (C392S, C395S; SEQ ID NO:30);

(B) 5'-GGGCAGCTGGAAAAGCGGAAAGGAAGG-3' (C413S, C416S; SEQ ID NO:31) and 5'-CCTTCCTTTCCGCTTTTCCAGCTGCCC-3' (C413S, C416S; SEQ ID NO:32);

(C) 5'-CCAGACATAGTTATCTATCAATACATGAACGATTTGTATGTAGG-3' (D185N; SEQ ID NO:33) and 5'-CCTACATACAAATCGTTCATGTATTGATAGATAACTATGTCTGG-3' (D185N; SEQ ID NO:34);

(D) 5'-GGGGAAATTGAATACCGCAAGTCAGATTTACCC-3' (W266T; SEQ ID NO:35) and 5'-GGGTAAATCTGACTTGCGGTATTCAATTTCCCC-3' (W266T; SEQ ID NO:36);

(E) 5'-CCCTAACTAACACAACAAATCAGAAAACCTCAGTTACAAGC-3' (E478Q; SEQ ID NO:37) and 5'-GCTTGTAAGTGTGTTTCTGATTTGTTGTGTTAGTTAGGG-3' (E478Q; SEQ ID NO:38).

#### Example 9. pGA2/JS7 vaccine plasmid

The sequence of pGA2/JS7 is shown in Fig. 11a (SEQ ID NO:8), its functional regions and the origins of these regions in Fig. 11b and the positions of its codon mutations in Fig. 11c. In the JS7 insert, Gag sequences of HIV-1-HXB-2 are substituted for the Gag sequences of BH10. This was accomplished by PCR amplification of the HXB-2 sequence

(p5' plasmid, NIH AIDS Research and Reference Program, catalog # 3119) using the following primers:

forward 5'-GAGCTCTATCGATGCAGGACTCGGCTTGC-3' (SEQ ID NO:39)

and reverse 5'-CTCCAATTACTGTGAGAATTCTAATGTTTCATCTTGGG-3' (SEQ ID

5 NO:40). The forward primer introduced a *Cla* I site at the same position as that found in the JS2 insert and the reverse primer introduced a unique *EcoR* I site analogous to the same site in the JS2 insert. This PCR fragment was then inserted into pGA1.1 for mutagenesis. The safety mutations in the zinc finger regions and the RT mutations were then introduced as previously described for the JS2 insert. JS7 also differs from JS2 in having an inactivating  
 10 codon mutation at the active site of protease. This mutation was introduced using the primers: 5'-GGCAACTAAAGGAAGCTCTATTAGCCACAGGAGC-3' (D25A Prt1; forward; SEQ ID NO:41) and 5'-GCTCCTGTGGCTAATAGAGCTTCCTT-TAGTTGCC-3' (D25A Prt2; reverse; SEQ ID NO:42). Once the mutations were confirmed by sequencing, the HXB-2 Gag-Pol insert was introduced into pGA2/JS2 via the *Cla* I and *EcoR*  
 15 I sites. In contrast to the JS2 insert that expresses aggregates of HIV-1 proteins due to premature cleavage of the pr55Gag polyprotein by the viral protease, the JS7 insert forms immature virus like particles (VLPs) that bud from the plasma membrane of DNA-expressing cells.

#### 20 **Example 10 pGA2/JS7.1 vaccine plasmid**

The sequence of pGA2/JS7.1 is shown in Fig. 12a (SEQ ID NO:9), its functional regions and the origins of these regions in Fig. 12b and the positions of its codon mutations in Fig. 12c. pGA2/JS7.1 is a derivative of pGA2/JS7 in which the start codon as well as an immediately upstream ATG have been mutated in *vpu*. These mutations were introduced to  
 25 increase the level of the expression of Env. The mutations in the start codon for Vpu were accomplished using a site directed mutagenesis kit (Stratagene) and the oligonucleotides: forward 5'-GCAGTAAGTAGTAAATCTAATCCAACCTTTAC-3' (SEQ ID NO:43) and reverse 5'-GTAAAGGTTGGATTAGATTACTACTTACTGC-3' (SEQ ID NO:44).

### Example 11. pGA1/IC25 vaccine plasmid

The sequence of pGA1/IC25 is shown in Fig. 13a (SEQ ID NO:10), its functional regions and the origins of these regions in Fig. 13b and the positions of its point mutations in Fig. 13c. The IC25 insert described here was designed with a circulating recombinant form of clades A and G (designated AG) so that it would elicit an immune response against HIV-1 sequences that predominate in West Africa. As noted above, any clade AG isolate from West Africa could be used as a reasonable representative for other clade AG isolates. Since HIV-1 isolates use different chemokine receptors as co-receptors, and the vast majority of viruses that are undergoing transmission use the CCR-5 co-receptor (Berger, *AIDS* 11(Suppl A):S3-16, 1997), the AG vaccine insert we designed had a CCR-5-using Env. Of course, Envs that function through any other co-receptor or that have been constructed from naturally occurring or synthetic AG sequences so as to increase immunogenicity can be made and used as well.

To achieve a multiprotein-expressing clade AG vaccine insert with high expression, candidate vaccines were constructed from four different AG HIV-1 isolates, as shown in Table 2.

Table 2. Comparison of candidate AG vaccine inserts

| Plasmid designation | SEquences tested   | Ability to grow plasmid | Expression of Gag | Expression of Env | Comment                              |
|---------------------|--------------------|-------------------------|-------------------|-------------------|--------------------------------------|
| 418/928             | 418 gag in 928-VLP | Poor                    | Poor              | not tested        |                                      |
| 421/928             | 421 gag in 928-VLP | Good                    | Good              | Poor              |                                      |
| 896/928             | 896 gag in 928-VLP | Good                    | Good              | Poor              |                                      |
| 928/928             | 928                | Good                    | Good              | Good              | chosen for vaccine, renamed pGA1/IC1 |

For each isolate, the forward primer 5'-AAGATCTATCGATGCAAGGACTCGGCTTGC-3' (SEQ ID NO:45) and the reverse primer 5'-TTCCAATTGCTGTGAGAATTCTCA-TGCTCTTCTTGGG-3' (SEQ ID NO:46) were used to amplify the 5' Gag-RT PCR product. The 3' PCR product for the Env region encompassed the *vpu*, *tat*, *rev*, and *env* sequences and

the splice acceptor sites necessary for proper processing and expression of their respective mRNAs. An *EcoR* I site was introduced at the 5' end of this product and *Nhe* I and *Rsr* II sites were introduced into the 3' end. A forward primer

5'-AAGGGGTTAAAGCTATAATAAG-AATTCTGCA-3' (SEQ ID NO:47) and a reverse  
 5 primer 5'-CCTTTGCTGCCCTATCTGA-TTCTTCTAGG-3' (SEQ ID NO:48) were used to amplify the 3' PCR product. Of these, those from patient 928 proved particularly favorable for further development (Table 2. The 928 sequences with deletions but not codon mutations were designated IC1.

The strategy used to construct IC25, a more disabled virus than IC1, was similar to  
 10 that used to construct JS7 from JS1. Specifically four codon mutations were introduced into *gag* sequences to inactivate the zinc fingers that are involved in RNA packaging, three codon mutations were introduced into *pol* sequences to inactivate transcription, strand transfer and RNaseH activities of reverse transcriptase and the codon at the active site of the protease was mutated to limit proteolytic cleavage of viral Gag proteins and the maturation of  
 15 viral particles. The protease mutations also limited premature cleavage of the Gag polyprotein and allowed budding of immature VLPs.

The inactivating codon mutations were made using a site directed mutagenesis kit (Stratagene) following the manufacturer's protocol. All mutations were confirmed by sequencing. Primer pairs used for the mutagenesis were:

- 20 (A) 5'-GCCAGAGAATAATAAAGAGCTTCAACAGCGGCAAAGAAGG-3' (C390S, C393S; SEQ ID NO:49) and  
 5'-CCTTCTTTGCCGCTGTTGAAGCTCTTATTATTCTCTGGC-3' (C390S, C393S; SEQ ID NO:50);  
 (B) 5'-CCTAGAAAGAGAGGCAGCTGGAAAAGCGGAAAGGAAGG-3' (C411S, C414S; SEQ ID NO:51) and 5'-  
 25 CCTTCCTTCCGCTTTTCCAGCTGCCTCTCTTTCTAGG-3' (C414S 928 ZN4; SEQ ID NO:52);  
 (C) 5'-CCAATATATGAACGATTTATATGTAGGATCTGAC-3' (D185N; SEQ ID NO:53) and 5'-GTCAGATCCTACATATAAATCGTTCATATATTGG-3' (D185N;  
 30 SEQ ID NO:54);

(D) 5'-GGGAAACTAAATACCGCAAGTCAGATTTATGCAGG-3' (W266T; SEQ ID NO:55) and 5'-CCTGCATAAATCTGACTTGCGGTATTTAGTTTTCCC-3' (W266T; SEQ ID NO:56); and

(E) 5'-CCCTAATTGAGACAACAAATCAAAAGACTCAGTTACATGC-3' (E478Q; SEQ ID NO:57) and 5'-GCATGTAAGTCTTTTGATTTGTTGTCTCAATTAGGG-3' (E478Q; SEQ ID NO:58).

(F) 5'-GCCAATAGAAGCCCTATTAAACACAGGAGC-3' (D25A; SEQ ID NO:59) and 5'-GCTCCTGTGTTTAATAGGGCTTCTATTGGC-3' (D25A; SEQ ID NO:60).

#### Example 12. PGA1/IC2

The sequence of pGA1/IC2 is shown in Fig. 14a (SEQ ID NO:11), its functional regions and the origins of these regions in Fig. 14b and the positions of its point mutations in Fig. 14c. pGA1/IC2 is identical to pGA1/IC25 except for not containing the inactivating point mutation in protease.

#### Example 13. PGA1/IC48

The sequence of pGA1/IC48 is shown in Fig. 15a (SEQ ID NO:12), its functional regions and the origins of these regions in Fig. 15b and the positions of its point mutations in Fig. 15c. pGA1/IC48 is identical to pGA1/IC25 except that the codon mutation in protease is one that occurred in a drug resistant mutant (Jacobsen *et al.*, *Virology* 206:527-534, 1995). This mutation only partially inactivates the protease function. Mutagenesis was carried out using Stratagene kits and the following oligonucleotides:

5'-CCAAAAATGATAGtGGGAATTGGAGG -3' (G48V 928; SEQ ID NO:61) and 5'-CCTCCAATTCCCaCTATCATTTTTGG-3' (G48V 928; SEQ ID NO:62). This mutation only partially inactivates the protease function.

#### Example 14. PGA1/IC90

The sequence of pGA1/IC90 is shown in Fig. 16a (SEQ ID NO:13), its functional regions and the origins of these regions in Fig. 16b and the positions of its point mutations in Fig. 16c. pGA1/IC90 is identical to pGA1/IC25 except that the codon mutation in protease is one that occurred in a drug resistant mutant (Jacobsen *et al.*, *Virology* 206:527-534, 1995).



This mutation only partially inactivates the protease function. Mutagenesis was carried out using Stratagene kits and the following oligonucleotides,; 5'-

GGACGAAATATG<sub>a</sub>TGACTCAGATTGGT -3' (M90L; SEQ ID NO:63) and  
5'-ACCAATCTGAGTCA<sub>t</sub>CATATTTCTGCC-3' (M90L; SEQ ID NO:64).

5

#### Example 15. pGA1/IN3

The sequence of pGA1/IN3 is shown in Fig. 17a (SEQ ID NO:14), its functional regions and the origins of these regions in Fig. 17b and the positions of its point mutations in Fig. 17c. The IN3 insert described here was constructed from a clade C sequence recovered from a virus in India. As noted above, any clade C isolate could be used as a reasonable representative for other clade C isolates. Since HIV-1 isolates use different chemokine receptors as co-receptors, and the vast majority of viruses that are undergoing transmission use the CCR-5 co-receptor (Berger, *AIDS* 11(Suppl A):S3-16, 1997), the C vaccine insert we chose to construct had a CCR-5-using Env. Of course, Envs that function through any other co-receptor or that have been constructed from naturally occurring or synthetic C sequences so as to increase immunogenicity can be made and used as well.

To achieve a multiprotein-expressing clade C vaccine insert with high expression, candidate vaccines were constructed from four different clade C HIV-1 sequences that were obtained from the US NIAID AIDS repository, as shown in Table 3. Of these, those from the Indian clone proved particularly favorable for further development.

Table 3. Comparison of clade C candidate vaccine inserts

| Isolate and Genbank Accession # | Ability to grow plasmid | Expression of Gag | Expression of Env | Comment                              |
|---------------------------------|-------------------------|-------------------|-------------------|--------------------------------------|
| South Africa AF286227           | Good                    | Good              | Good              |                                      |
| Israel AF286233                 | Good                    | Good              | Good              |                                      |
| Tanzania AF286235               | Good                    | Good              | Good              |                                      |
| India AF286231                  | Good                    | Good              | Very good         | Chosen for vaccine, renamed pGA1/IN1 |

5' and 3' sequences from the Indian clone were cloned into pGA1.2 using oligonucleotides and PCR to generate 5' and 3' fragments. The 5' fragment encoding Gag and RT was generated using the forward primer 5'-CGCAGGATCCGGCTTGCTGAAG-3' (SEQ ID NO:65), which incorporated a *Bam*H I site at the 5' end of the fragment, and the reverse primer 5'-TCTACTCGAGCTTATTATAGCACTCTCCTG-3' (SEQ ID NO:66), which incorporated an *Xho* I site as well as two stop codons at the 3' end of the fragment. The 3' fragment encoding Tat, Rev, Vpu, and Env was generated using the forward primer 5'-CCTCTCGAGATACTTGGACAGGAG-3' (SEQ ID NO:67) and the reverse primer 5'-CACTTGCTAGCCATTTTACTGCAAAGC-3' (SEQ ID NO:68). These were designed such that *Xho* I and *Nhe* I restriction sites were incorporated at the 5' and 3' ends, respectively of the 3' fragment. These fragments were introduced into pGA1.2 using directed cloning to create pGA1.2/IN1.

The strategy used to construct IN3, a more disabled virus than IN1, was similar to that used to construct JS7 from JS1. Specifically four codon mutations were introduced into *gag* sequences to inactivate the zinc fingers that are involved in RNA packaging, three codon mutations were introduced into *pol* sequences to inactivate transcription, strand transfer and RNaseH activities of reverse transcriptase and the codon at the active site of the protease was mutated to limit proteolytic cleavage of viral Gag proteins and the maturation of viral particles. The protease mutations also limited premature cleavage of the Gag polyprotein and allowed budding of immature VLPs.

The inactivating codon mutations were made using a site directed mutagenesis kit (Stratagene) following the manufacturer's protocol. All mutations were confirmed by sequencing. Primer pairs used for the mutagenesis were:

(A) 5'-CTAAAAGAACTGTAAATCCTTCAACTCTGGCAAGGAAGGGCAC-3' (C390S, C393S; SEQ ID NO: 70) and 5'-GTGCCCTTCCTTGCCAGAGTTGAAGGATTAAACAGTTCTTTTAG-3' (C390S, C393S; SEQ ID NO: 71);  
(B) 5' -CTAGGAAAAAAGGCTCTTGGAATCTGGAAAGGAAGGACAC (C411S, C414S; SEQ ID NO:72) and

5'-GTGTCCTTCCTTTCCAGATTTCCAAGAGCCTTTTTTCCTAG-3' (C411S and C414S; SEQ ID NO:73):

(C) 5'-GTCATCTATCAATATATGAATGACTTGTATGTAG-3' (D185N, SEQ ID NO:74) and 5'-CTACATACAAGTCATTCATATATTGATAGATGAC-3' (D185N, SEQ ID NO:75);

(D) 5'-GTGGGAAAATTAAACACGGCAAGCCAGATTTAC-3' (W266T, SEQ ID NO:76) and 5'-GTAAATCTGGCTTGCCGTGTTAATTTTCCCAC-3' (W266T, SEQ ID NO:77);

(E) 5'-CAAATCAGAAGACTCAATTACAAGCAATTTATC-3' (E478Q, SEQ ID NO:78) and 5'-GATAAATTGCTTGTAATTGAGTCTTCTGATTTG-3' (E478Q, SEQ ID NO:79) and

(F) 5'-GGAGGCTCTCTTAGcCACAGGAGCAGATG-3' (D25N, SEQ ID NO: 80) and 5'-CATCTGCTCCTGTGgCTAAGAGAGCCTCC-3' (D25N, SEQ ID NO 81).

#### Example 16. pGA1/IN2

The sequence of pGA1/IN2 is shown in Fig. 18a (SEQ ID NO: 5), its functional regions and the origins of these regions in Fig. 18b and the positions of its point mutations in Fig. 18c. pGA1IN2 differs from pGA1/IN3 in not having the D25N Inactivating point mutation in protease.

#### Example 17: Sequences provided for matched rMVAs

Sequences for the JS, IC, and IN inserts were used to prepare matched recombinant modified vaccinia Ankara (rMVA) vectors. These matched vectors can be used as booster inoculations for the various DNAs. They can also be used for both priming and boosting an anti-HIV immune response. The sequences provided to generate the viral vector included the three inactivating point mutations in reverse transcriptase. A representative study, in which a recombinant MVA vector was constructed and characterized, follows.

MVA virus (which may be obtained from the American Type Culture Collection) 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). Secondary CEF cells were infected at an MOI of 0.05 of MVA and transfected with 2 pLW-48 (as described above; see Figs. 20A and 20B). Following a two-

day incubation at 37°C, the virus was harvested, frozen and thawed three times. It was then plated on CEF plates. At “four 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. 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 third replicate plates of those samples that had about equal numbers of mixed populations of GUS staining and non-staining foci as well as mostly envelope staining foci. These foci were again plated out in triplicate and analyzed the same way. After five passages, a virus was derived that 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.

Aliquots of MVA/HIV48 infected cell lysates were analyzed by radio-immunoprecipitation and immunostaining with monoclonal antibodies for expression of both the Env and Gag-Pol protein. In both of these tests, each of these proteins was detected. The recombinant virus was shown to produce gag particles in the supernatant of infected cells by pelleting the <sup>35</sup>S-labeled particles on a 20% sucrose cushion. By electron microscopy, gag particles were visualized both outside and budding from cells as well as within vacuoles of cells. The gag particles had envelope protein on their surface.

Thus, we made a recombinant MVA virus that expressed the ADA truncated envelope and the HXB2 Gag-Pol protein. 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; see, *e.g.*, Figs.21, 22, and 24). In addition, the envelope has been truncated, as this may enhance the amount of protein on the surface of the infected cells and hence enhance immunogenicity. Stability of the recombinant may also be enhanced.

**What is claimed is:**

1. A composition comprising:

(a) a first vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against a human immunodeficiency virus (HIV) of a first subtype or

5 recombinant form; and

(b) a second vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form.

2. The composition of claim 1, further comprising a pharmaceutically acceptable  
10 carrier or adjuvant.

3. The composition of claim 1, wherein the insert of the first vector or the insert of the second vector comprises the sequences of two or more of: (a) a *gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, or *vpu* gene or (b) mutants thereof and, optionally, (c) non-coding regulatory  
15 sequences of the HIV genome.

4. The composition of claim 3, wherein at least one of the two or more sequences comprises mutations that limit the encapsidation of viral RNA.

20 5. The composition of claim 3, wherein at least one of the two or more sequences comprises a *gag* sequence having a mutation in one or more of the sequences encoding a zinc finger.

6. The composition of claim 4, wherein all or part of cis-acting RNA encapsidation  
25 sequences have been deleted from the non-coding regulatory sequences of HIV-1.

7. The composition of claim 3, wherein the two or more sequences comprise a *pol* sequence having a mutation that inhibits one or more of the enzymatic activities of pol.

30 8. The composition of claim 7, wherein the enzymatic activity is integrase activity.

9. The composition of claim 7, wherein the enzymatic activity is reverse transcriptase activity.

10. The composition of claim 7, wherein the enzymatic activity is protease activity.

5

11. The composition of claim 8, 9, or 10, wherein the enzymatic activity is inhibited by deleting a portion of the *pol* sequence or introducing one or more point mutations into the *pol* sequence.

10 12. The composition of claim 3, wherein the insert of the first vector or the insert of the second vector comprises *tat*, *rev*, *nef*, *vif*, *vpr*, or *vpu* having one or more deletions or point mutations.

15 13. The composition of claim 12, wherein the deletion or point mutation inactivates one or more of the start codons of *tat*, *rev*, *nef*, *vif*, *vpr*, or *vpu*.

14. The composition of claim 3, wherein the first vector or the second vector is a plasmid encoding:

(a) a gag protein in which the zinc fingers have been inactivated;

20 (b) a pol protein in which (i) the integrase activity has been inhibited by deletion of some or all of the pol sequence and (ii) the polymerase, strand transfer and/or RNase H activity of reverse transcriptase has been inhibited by one or more point mutations within the pol sequence; and

(c) env, tat, rev, and vpu, without mutations.

25

15. The composition of claim 14, wherein the encoded proteins are derived from subtype B or D HIV-1 or a recombinant form of a subtype B or D HIV-1.

30 16. The composition of claim 14, wherein the first vector or the second vector comprises pGA2/JS2.

17. The composition of claim 14, wherein the encoded proteins are derived from a subtype A HIV-1 or a recombinant form of a subtype A HIV-1.

5 18. The composition of claim 17, wherein the first vector or the second vector comprises pGA1/IC2.

19. The composition of claim 14, wherein the encoded proteins are derived from a subtype C HIV-1 or a recombinant form of a subtype C HIV-1.

10 20. The composition of claim 19, wherein the first vector or the second vector comprises pGA1/IN2.

15 21. The composition of claim 14, wherein the encoded proteins are derived from a subtype E, F, G, H, I, J, K, or L subtype of HIV-1 or a recombinant form of a subtype E, F, G, H, I, J, K or L HIV-1.

22. The composition of claim 3, wherein the first vector or the second vector is a plasmid encoding:

- 20 (a) a gag protein in which both zinc fingers have been inactivated;
- (b) a pol protein in which (i) the integrase activity has been inhibited by deletion of some or all of the pol sequence, (ii) the polymerase, strand transfer and/or RNase H activity of reverse transcriptase has been inhibited by one or more point mutations within the pol sequence and (iii) the proteolytic activity of the protease has been inhibited by one or more point mutations; and
- 25 (c) env, tat, rev, and vpu, without mutations.

23. The composition of claim 22, wherein the encoded proteins are derived from subtype B or D HIV-1 or a recombinant form of a subtype B or D HIV-1.

30 24. The composition of claim 23, wherein the first vector or the second vector comprises pGA2/JS7.

25. The composition of claim 22, wherein the encoded proteins are derived from a subtype A HIV-1 or recombinant form of a subtype A HIV-1.

5           26. The composition of claim 25, wherein the first vector or the second vector comprises pGA1/IC25.

27. The composition of claim 22, wherein the encoded proteins are derived from a subtype C HIV-1 or recombinant form of a subtype C HIV-1.

10

28. The composition of claim 27, wherein the first vector or the second vector comprises pGA1/IN3.

29. The composition of claim 22, wherein the encoded proteins are derived from a subtype E, F, G, H, I, J, K or L subtype of HIV-1 or a recombinant form of a subtype E, F, G, H, I, J, K or L HIV-1.

15

30. The composition of claim 3, wherein the first vector or the second vector is a plasmid encoding:

20

(a) a gag protein in which the zinc fingers have been inactivated;

(b) a pol protein in which (i) the integrase activity has been inhibited, (ii) the polymerase, strand transfer and/or RNase H activities of reverse transcriptase have been inhibited by one or more point mutations within the pol sequence, and (iii) the proteolytic activity of the protease has been inhibited by one or more point mutations within the pol sequence;

25

(c) a vpu protein comprising a mutant start codon; and

(d) env, tat, and rev proteins in unmutated form.

31. The composition of claim 30, wherein the encoded proteins are derived from subtype B or D HIV-1 or a recombinant form of a subtype B or D HIV-1.

30



32. The composition of claim 31, wherein the first vector or the second vector comprises pGA2/JS7.1.

5 33. The composition of claim 30, wherein the encoded proteins are derived from a subtype A, C, E, F, G, H, I, J, K or L subtype of HIV-1 or a recombinant form of a subtype A, C, E, F, G, H, I, J, K, or L HIV-1.

34. The composition of claim 1, wherein the vector is a virus.

10 35. The composition of claim 34, wherein the virus is a recombinant poxvirus.

36. The composition of claim 34, wherein the virus is modified vaccinia Ankara virus.

15 37. The composition of claim 36, wherein the modified vaccinia Ankara virus harbors recombinant nucleic acid encoding clade B sequences matched to those in JS2 or JS7.

20 38. The composition of claim 36, wherein the modified vaccinia Ankara virus harbors recombinant nucleic acid encoding clade A sequences matched to those in IC2 or IC25.

25 39. The composition of claim 36, wherein the modified vaccinia Ankara virus harbors recombinant nucleic acid encoding clade C sequences matched to those in IN2 or IN3.

40. The composition of claim 36, wherein the modified vaccinia Ankara virus is a recombinant for subtype A, B, C, D, E, F, G, H, I, J, K or L HIV-1 or recombinant forms of subtype A, B, C, D, E, F, G, H, I, J, K or L HIV-1 matched to those in a DNA vaccine.

30

41. A method of treating a subject, the method comprising:

(a) identifying a subject in need of an HIV vaccine; and

(b) administering to the subject a therapeutically effective amount of a composition comprising a vector comprising a vaccine insert that induces or enhances an immune response against a human immunodeficiency virus.

42. A method of treating a subject, the method comprising:

(a) identifying a subject in need of an HIV vaccine; and

(b) administering to the subject a therapeutically effective amount of a composition comprising (i) a first vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against a human immunodeficiency virus (HIV) of a first subtype or recombinant form and (ii) a second vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form.

43. A method of treating a subject, the method comprising:

(a) identifying a subject in need of an HIV vaccine; and

(b) administering to the subject a therapeutically effective amount of a composition comprising (i) a first vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against a human immunodeficiency virus (HIV) of a first subtype or recombinant form, (ii) a second vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a second subtype or recombinant form, and (iii) a third vector comprising a vaccine insert encoding one or more antigens that elicit an immune response against an HIV of a third subtype or recombinant form.

44. The method of claim 40, 41, or 42, wherein administering the composition comprises administering a plasmid vector on more than one occasion for the purposes of priming and boosting a protective immune response.

45. The method of claim 40, 41, or 42, wherein administering the composition comprises administering a plasmid vector on one or more than one occasion for the purpose of priming the immune response and administering a modified vaccinia Ankara virus on one or more than one subsequent occasion for the purpose of boosting the immune response.

5

46. The method of claim 45, wherein the plasmid vector comprises the insert designated JS2, JS7, or JS7.1, and/or the insert designated IC2, IC25, IC48, or IC90, and/or the insert designated IN2 or IN3 and wherein the boosting comprises administering modified vaccinia Ankara virus containing HIV sequences matched to the plasmids used for priming.

10

47. The method of claim 45, wherein the plasmid vector comprises an insert obtained from an HIV clade A and/or clade B, and/or clade C, and/or clade D, and/or clade E and/or clade F, and/or clade G and/or clade H and/or clade J, and/or clade K, and/or clade L and/or a recombinant subtype thereof and wherein the boosting comprises administering modified vaccinia Ankara virus containing HIV sequences matched to the plasmids used for priming.

15

48. The method of claim 40, 41, or 42, wherein administering the composition comprises administering modified vaccinia Ankara virus on more than one occasion for the purposes of priming and boosting a protective immune response.

20

49. The method of claim 44, wherein the priming comprises administering clade B recombinant MVA virus and/or clade A recombinant MVA virus and/or clade C recombinant MVA virus and or any recombinant subtype thereof wherein the boosting comprises administering the same modified vaccinia Ankara virus used for priming.

25

50. The method of claim 44, wherein the priming comprises administering clade A and/or clade B, and/or clade C, and/or clade D, and or/clade E and/or clade F, and/or clade G and/or clade H and/or clade J, and/or clade K, and/or clade L and/or any recombinant subtype thereof in a modified vaccinia Ankara vector and wherein the boosting comprises
- 5 administering the same modified vaccinia Ankara vectors used for priming.

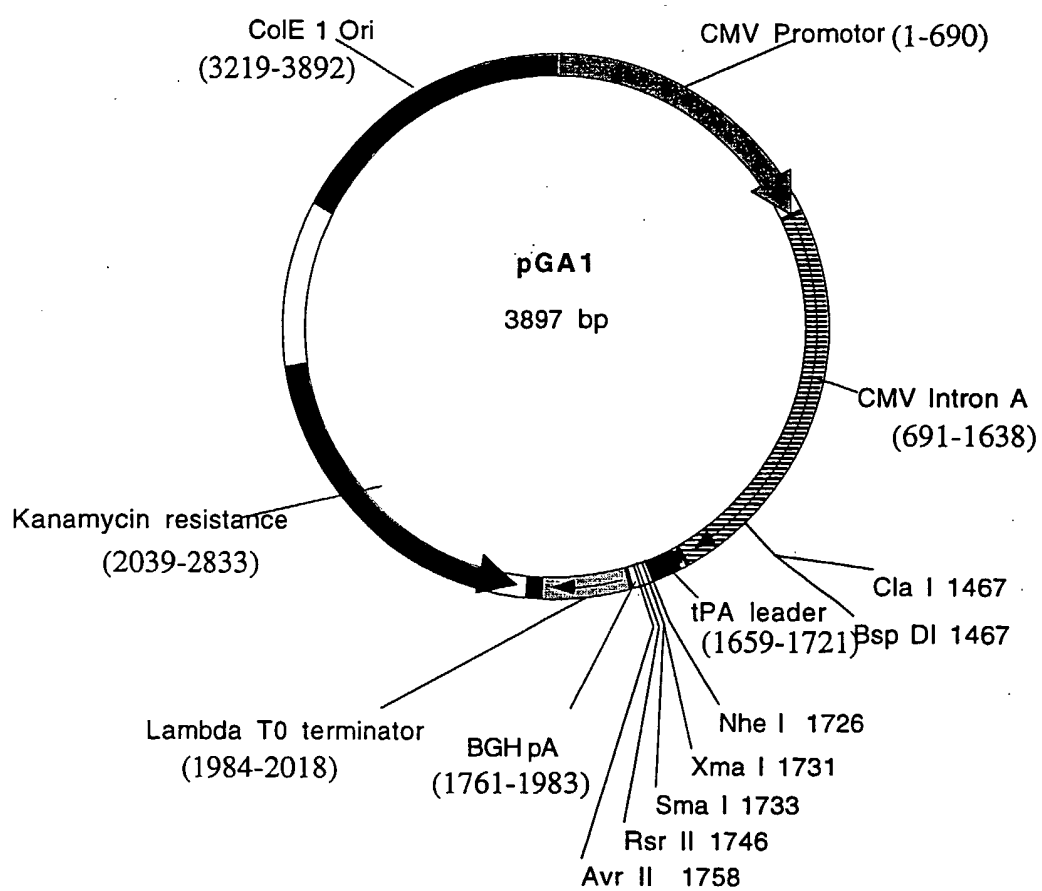


Fig. 1

Figure 2a. pGA1 Sequence

cgacaatattggctattggccattgcatacgttgtatctatatcataatatgtacattttatattggctcatgtccaatat  
gaccgccatgttgacattgattattgactagttattaatagtaataattacgggttcattagttcatagcccatatatg  
gagttccgcgttacataacttacggtaaatggcccgctggctgaccgcccacgacccccgcccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagattttacggtaaaactgccact  
tggcagtagacatcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggcccgctggcattat  
gcccagtagacacttacgggaactttcctacttggcagtagacatctacgggtattagtcacggctattaccatgggtgatg  
cggttttggcagtagaccaatgggcgtggatagcgggtttgactcacggggattttccaagtctccacccccattgacgtcaa  
tgggagtttgttttggcaccaaaatcaacgggaactttccaaaatgtcgttaataacccccgcccgttgacgcaaatgggcg  
gtaggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgctggagacgccatccacgc  
tgttttgacctccatagaagacacggggaccgatccagcctccgcccgggaacgggtgcattggaacgcggattccccg  
tgccaagagtgacgttaagtaccgcctatagactctataggcacaccccccttggctcttatgcatgctatactgtttttgg  
cttggggcctatacacccccgcttcccttatgctataggtgatgggtatagcttagcctataggtgtgggttattgaccatt  
attgaccactccccctattgggtgacgatactttccattactaatccataacatggctctttgccacaactatctctattgg  
ctatatgccaaactctgtccttcagagactgacacggactctgtatttttacaggatgggggtcccatttattatttaca  
aattcacatatacaacaacgcgctccccgctgcccgcagttttattaaacatagcgtgggatctccacgcgaatctcgg  
gtaccggtgttccggacatggggtcttctccggttagcggcgagcttccacatccgagccctgggtcccatgcctccagcgg  
ctcatggctcgctcggcagctccttgctcctaacagtggagccagacttaggcacagcacaatgccaccaccaccagt  
tgccgcacaaaggcgtggcgttaggggtatgtgtctgaaaaatgagctcggagattggggtcgacacgcgtgacgcagatgga  
agacttaaggcagcggcagaagaagatgcaggcagctgagttgtgtattctgataagagtcagaggtaactcccggttgc  
gggtgtgttaacgggtggagggcagtgtagctgagcagtagctcgttgctgcccgcgcgcaccagacataatagctgac  
agactaacagactgttcccttccatgggtctttctgacgtcaccatcgatgcttgcaatcatggatgcaatgaagagag  
ggctctgctgtgtgctgctgctggtggagcagctctcgtttcgggttagccccgggtgataaacggaccgcgcaatccct  
aggctgtgcttctagttgccagccatctgtgtttgccccctccccgtgcttcccttgacctggaagggtgccactccc  
actgtccttctcctaataaaaaatgaggaaattgcatcgcatgtgtgctgagtaggtgtcattctattctgggggttgggttggg  
gcaggacagcaagggggaggtatgggaagacaatagcaggcatgctggggatgagggtgggtctctataaaaaacgccc  
gcggcaaccgagcgttctgaaacgctagagtcgacaaattcagaagaactcgtcaagaaggcagatagaaggcagatgcgctg  
cgaatcgggagcggcgataccgtaaaagcacgaggaagcggctcagcccattcgccgccaagctcttcagcaatatcaagg  
tagccaacgctatgtcctgatagcggctctgccacaccagccggccacagtcgatgaatccagaaaagcggccattttcc  
accatgatattcggcagcagggcatcgccatgggtcacgacgagatcctcgccgtcgggcatgctcgcttgagcctggc  
gaacagttcggctggcgcgagccccctgatgctctcgtccagatcatcctgatcgacaagaccggttccatccgagtag  
gtgctcgctcgatgcgatgtttcgcttgggtggtcgaatgggcaggttagccggatcaagcgtatgcagccgcgcgcatgca  
tcagccatgatggatactttctcggcaggagcaaggtagatgacaggagatcctgccccggcacttcgcccataagcag  
ccagtccttccccgcttcagtgacaacgtcgagcacagctgcgcaaggaaacgcccgtcgtggccagccacgatagccgcg  
ctgctcgtcttgacgttcatcagggcacccggacaggtcggctcttgacaaaaagaacggggcgccccctgcgctgacagc  
cggaaacgcggcgccatcagagcagccgattgtctgttgcccagtcatagcccgaatagcctctccaccaagcggccgg  
agaacctgcgtgcaatccatctgttcaatcatgcgaaacgatcctcatcctgtctcttgatcagatcttgatccccctgc  
gccatcagatccttggcggcaagaaagccatccagtttactttgcagggttcccaaccttaccagagggcgccccagct  
ggcaattccgggttcgcttgctgtccataaaacgcccagcttagctatcgccatgtaagcccactgcaagctacctgctt  
tctctttgcgcttgctgtttcccttgtccagatagcccagtagctgacattcatccggggtcagcaccgtttctgcccac  
tggctttctacgtgaaaaggatctaggtgaagatccttttgataatctcatgacaaaaatcccttaacgtgagttttcg  
ttccactgagcgtcagaccctgtagaaaagatcaaaggatcttcttgagatcctttttctgcccgttaactgctgctt  
gcaaaacaaaaaaacccagcgtaccagcgggtgggtttggtttgcccggatcaagagctaccaactcttttccgaaggtaactg  
gcttcagcagagcgcagataccaaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagca  
ccgcctacatacctcgctctgctaactcgtgttaccagtggtgctgcccagtggaagtcgtgtcttaccgggttggga  
ctcaagacgatatgttaccggataaaggcgcagcggctcgggtgaaacggggggttcgtgcacacagcccagcttggagcga  
cgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccacgcttcccgaaggagaaaggcggacagg  
tatccggtaagcggcagggctcggaaacaggagagcgcaggggagcttccagggggaaacgcctgggtatctttatagtc  
tgtcgggtttcgccacctctgacttgagcgtcgatttttgtagtgcctcgtcagggggcgaggcctatggaaaacgcga  
gcaacgcggcccttttacgggttccctggccttttgctggccttttgctcacatgttgt

**Figure 2b. pGA1 Functional Regions**

| <b>Functional region</b>    | <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Origin of sequence</b>                                                                                                                 |
|-----------------------------|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| CMVIE promoter              | 1-690                                                          | pJW4303                                                                                                                                   |
| CMV intron A                | 691-1638                                                       | pJW4303                                                                                                                                   |
| tPA leader                  | 1659-1721                                                      | pJW4303                                                                                                                                   |
| Multiple cloning site       | 1648-1758<br>Cla I-Avr II                                      | pJW4303; mutation of Hind<br>III and Bam HI sites and<br>addition of Cla I and Rsr II<br>sites to the multiple cloning<br>site of pJW4303 |
| BGHpA                       | 1761-1983                                                      | pJW4303                                                                                                                                   |
| Lambda T0 terminator        | 1984-2018                                                      | Synthetic oligonucleotide                                                                                                                 |
| Kanamycin resistance        | 2039-2833                                                      | pZErO-2 (Invitrogen)                                                                                                                      |
| ColE1 origin of replication | 3219-3892                                                      | pZErO-2 (Invitrogen)                                                                                                                      |

Figure 3a. pGA1.1 Sequence

cgacaatattggctattggccattgcatacgttgatctatatcataatatgtacatttatattggctcatgtccaatat  
gaccgccatgttgacattgattattgactagttattaatagtaataacacgggttcattagttcatagcccatatatg  
gagttccgcgttacataaacttacggtaaatggccgcctggctgacggcccaacgacccccgcccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaacgcccact  
tggcagttacatcaagtgtatcatatgccaaagtcgccccctattgacgtcaatgacggtaaatggccgcctggcattat  
gcccagttacatgaccttacgggactttcctacttggcagttacatctacggtattagtcacggtcattaccatgggtgatg  
cggttttggcagttacaccaatgggctggatagcgggtttgactcacggggatttccaagtcctccacccccattgacgtcaa  
tgggagtttgggtttggcaccacaaatcaacgggactttccaaaatgtcgttaataacccccccccgttgacgcacaaatgggctg  
gtaggcgtgtacggtgggaggtctataaagcagagctcgttttagtgaaccgtcagatcgccctggagacgccatccacgc  
tgttttgacctccatagaagacacgggacggatccagcctccgcccggggaacggtgcattggaacgcggattccccg  
tgccaagagtgacgttaagtaccgctcatagactctataggcacaccccttggctcttatgcatgctatactgtttttgg  
cttggggcctatacacccccgcttccctatgctataggtgatgggtatagcttagcctataggtgtgggttattgaccatt  
attgaccactcccctattgggtgacgatactttccattactaataacataacatgggtcctttgccacaactatctctattgg  
ctatatgcccaatactctgtccttcagagactgacacggactctgtatttttacaggatggggtcccatttatatttaca  
aattcacatatacaaacacgcgtccccctgccccgcagtttttattaaacatagcgtggggtctccacgcgaatctcgg  
gtaccgtgttccggacatggggtcttctccggttagcggcggagcttccacatccgagccctgggtcccattgctccagcgg  
ctcatgggtcgtcggcagctccttgccttaacagtgaggccagacttaggcacagcacaaatgcccaccaccacagtg  
tgccgcacaaggccgtggcggtaggggtatgtgtctgaaaatgagctcggagattggggtcgcacccgtgacgcagatgga  
agacttaaggcagcggcagaagaagatgcaggcagctgagttgtgtattctgataagagtcagaggttaactcccgttgc  
gggtgctgttaacgggtggagggcagtgtagctgagcagctcgttgcctgcccgcgcgcacacagacataatagctgac  
agactaacagactgttcccttccatgggtctttctgcagtcacccatcgatgcttgcaatcatggatgcaatgaagagag  
ggctctgctgtgtgctgctgctgtgtggagaattcttcggttctgctgctgtgtggagaattcttcggttcggctagccc  
cgggtgataaacggacgcgcgaatccctaggtgtgccttctagttgcccagccatctgtgtttgcccctccccgtgcc  
ttccttgacctggaaggtgccactcccactgtccttcttaataaaatgaggaaattgcatcgcatgtctgagtaggt  
gtcattctattctggggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggat  
gcgggtgggtctctataataaaaaacggcccgccggaacgcagcgtctgaacgctagagtcgacaaattcagaagaactcgt  
caagaaggcgtatagaaggcgtgctgctggaatcgggagcggcgtatccgtaaaagcagaggaagcgggtcagcccattcg  
ccgccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcgggtctgccacacccagccggccacagtc  
gatgaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgc  
cgtcgggcatgctcgcccttgagcctggcgaaacagttcgggtggcgagcggccctgatgctcttcgtccagatcatcctga  
tcgacaagaccggcttccatccgagtagctgctcgtcgatgcatggttctcgttgggtggaatgggagggtagccgg  
atcaagcgtatgcaagcggcattgcatcagccatgatggatactttctcggcaggagcaaggtagatgacaggagat  
cctgccccggcaacttcgccaatagcagccagtccttccccgttcagtgacaacgtcgagcacagctgcgcaaggaaacg  
cccgtcgtggccagccagatagccgcgtgctcgtcttcgagttcattcagggcaccggacaggtcgggtcttgacaaa  
aagaaccggcgcccctgctgacagccggaacacggcgccatcagagcagccgattgtctgttgtgcccagtcacagc  
cgaatagcctctccacccaagcggccggagaacctgcgtgcaatccatctgttcaatcatgcgaaacgatcctcatcct  
gtctcttgatcagatcttgatcccctgcgcatcagatccttggcggaagaaagccatccagtttactttgaggggtt  
cccaaccttaccagagggcgcccagctggcaattccgggttcgcttgcgtgctccataaaacggccagctagctatcgcc  
atgtaagcccaactgcaagctacctgctttctcttggcgttgcgttttcccttgctccagatagcccagtagctgacattc  
atccgggggtcagcaccgtttctgcggactggctttctacgtgaaaaggatctagggtgaagatcctttttgataatctcat  
gacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatc  
ctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcgggtgggtttgtttgcccggatcaagag  
ctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagtt  
aggccaccacttcaagaactctgtgacccgctacatacctcgtctgcttaatcctgttaccagtggtgctgcccagtg  
gcgataagtcgtgtcttaccgggttggaactcaagacgatagttaccggataaggcgagcgggtcgggtgaaacggggggt  
tcgtgcacacagccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccac  
gcttcccgaaggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacagggagcttccag  
ggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtca  
ggggggcggagcctatggaaaaacgccagcaacgcggccctttacgggttcttgcccttttgcctggtccttttgcctcacat  
gttgt



Figure 3b. pGA1.1 Functional Regions

| Functional region           | Position within sequence<br>(starting-and ending pos.) | Origin of sequence                                                     |
|-----------------------------|--------------------------------------------------------|------------------------------------------------------------------------|
| CMVIE promoter              | 1-690                                                  | pGA1                                                                   |
| CMV intron A                | 691-1638                                               | pGA1                                                                   |
| tPA leader                  | none                                                   |                                                                        |
| Multiple cloning site       | 1648-1786<br>(Cla I-Avr II)                            | pGA1, addition of two EcoR I<br>sites to pGA1 multiple cloning<br>site |
| BGHpA                       | 1789-2011                                              | pGA1                                                                   |
| Lambda T0 terminator        | 2012-2046                                              | pGA1                                                                   |
| Kanamycin resistance        | 2067-2861                                              | pGA1                                                                   |
| ColE1 origin of replication | 3247-3920                                              | pGA1                                                                   |

cgaacaatttggctattggccattgcatcagttgtatctatatcāaāatgtacatttatattggctcatgtccaatat  
gaccgccatgttgacattgtatttgactagttattaatagtaatcaattacgggttcattagttcatagcccatatatg  
gagttccgcgttacataacttacggtaaatggccgcctggctgaccgcccaacgacccccgccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccact  
tggcagtagacatcaagtgtatcatatgccaaagtccgccccctattgacgtcaatgacggtaaatggcccgcctggcatat  
gcccagtagacatgacctacgggactttcctatttggcagtagacatcaccggtattagctcatcggtattaccattggtag  
cgggtttggcagtagaccaatggcggttagcgttagcgttagcgttagcgttagcgttagcgttagcgttagcgttagcgt  
tgggagtttgttttggcaccāaaatcaacgggactttccāaaatgtcgtaataaccccgccccgtgacgcaaatgggcg  
gtaggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgctggagacgccatccacgc  
tgttttgacctccatagaagacaccgggaccgatccagcctccgcggccgggaacgggtgcattggaacgcggattccccg  
tgccaagagtgacgtaaagtaccgcctatagactctataggcacaccccccttggctcttatgcatgctatactgttttgg  
cttggggcctatacacccccgccttcccttatgctataggtgatggtagttagcttagcctataggttgggttattgaccatt  
attgaccactccccatttggtagcagactttccatttactaatccataacatggctcttggccaactactctctattgg  
ctatattgccaaactctgtctctcagagactgacacggactctgtatttttacaggatgggggtcccatttattatttaca  
aattcacatataacaacacgcctgccccgcgtgcccgcagtttttattaacatagcgtgggatctccacgcgaatctcgg  
gtaccgtgttccggacatgggytcttctccggttagcggcggagcttcacatccgagccctgggtcccatgcctccagcgg  
ctcatgggtcgctcggcagctccttgctcctaacagtggaggccagacttaggcacagcacaatgccaccaccaccagtg  
tgccgcacaaggccgtggcggttagggtaggtatgtgtctgaaaatgagctcggagattgggctcgcaccgcgtcagcagatgga  
agacttaaggcagcggcaggaagaagtagcaggcagctgagtgtgtattctgtataagagtacagggtaacctccgttgc  
gggtcgtgttaacggtggagggaagtagtctgtagctgagcagtagctcgttgctgcccgcgcgcgccaccagacataatagctgac  
agactaacagactgttcccttccatgggtcttttctgcagtaccatggatccttgcaactcgaggatgcaatgaagagag  
ggctctgctgtgtgctgtgtgtggagaattcttcgtttctgctgctgtgtggagaattcttcgtttcggttagccc  
cgggtgataaaacggaccgcgcaatccctaggctgtgccttctagtgtgccagccatctgttgtttgccccctccccgcgtgcc  
ttccttgacccttggagggtgccactccccactgccttctcctaataaaatgaggaaatgcatcgcatgtctgtagttagt  
gtcattctattctgggggggtgggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgtctggggat  
gcggtgggctctatataaaaaacgcgccgcggcgaacaggagcgtttgaaacgttagagtgcacaaatcagaagaactcgt  
caagaaggcgatagaaggcgatgcgtgcgaatcgggagcggcgataaccgtaagcacgaggaagcggtcagcccatctcg  
ccgccaagctcttcagcaatatcacgggttagccaacgctatgtcctgatagcggcttgccacacccagcgggccacagt  
gatgaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatctcgc  
cgtcgggcatgctcgcttgagcctggcggaacagttcggctggcgcgagccccctgagtctcttcgtccagatcatctgta  
tcgaacaggccggcttccatccagtagctgtcgtcgtcgtcgttcttcgtttgttgggtcgaattgggcaggtagccgg  
atcaacgctatgcagcgcgcgcatgcatcagcatgatggatactttctcggcaggagcaaggtagatgacaggagat  
cctgccccggcacttcgccaatagcagccagctcccttccgcgttcagtgcacacgtcgagcacagctgcgcaagggaacg  
ccgctcgtggccagccaagatagccgcgtgcctcgtcttgcaattcattcagggcacgggacaggctcggcttgacaaa  
aagaaccggggcgccccctgcgtgcagcgcgggaacacggcggtcatcagagcggatgtctgttggccagctcatagc  
cgaatagcctctccaccaagcggcgagggaacctgcgtgcaattccatctgttctcaatcatcgcaaacgcatcctcatcct  
gtctcttgatcagatcttgatccccctggccatcagatccttggcggaagaaagccatccagtttactttgcagggtct  
cccaaccttaccagagggcgcgccagctggcaattccggttcgcttgctgtccataaaaacgcgccagcttagctatcgcc  
atgtaagcccaactgcaagctacctgctttctctttgcgttgcggtttcccttgctcagatagcccagtagctgacattc  
atccggggtcagcacctgttctgcggaactggctttctacgtgaaaaggatctaggtgaagatccttttgataatctcat  
gacaaaaatcccttaacgtgagtttctgttccactgagcgtcagacccccgtagaaaagatcaaaggatctcttgagatc  
cttttttctgcgcgttaattctgctgttgcaaacaaaaaacaccgcgtaccagcgggtgtgtttgttgcgggatcaagag  
ctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaatctgttctctagtgtagccgtagtt  
aggccaccacttcaagaactctgtagcacccgctacatacctcgctctgctaattctgttaccagtggctgctgccagt  
gcgataagtctgtcttaccgggttggaactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggggt  
tcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaagcgcac  
gcttcccgaaggggagaaaggcggacaggtatccggttaagcgcaggggtcggaacaggagagcgcacgaggaggtccac  
ggggaaaacgcctggtatctttatagtcctgtcgggtttcgccacctgacttgagcgtcgatttttgtgtgctcgtca  
gggggggcggagcctatggaaaaacgcgacgcaacgcggccctttacgggttctcgtggccttttgcgtcacat  
gttgt

Figure 4b. pGA1.2 Functional Regions

| Functional region           | Position within sequence<br>(starting-and ending pos.) | Origin of sequence                                                               |
|-----------------------------|--------------------------------------------------------|----------------------------------------------------------------------------------|
| CMVIE promoter              | 1-690                                                  | pGA1.1                                                                           |
| CMV intron A                | 691-1638                                               | pGA1.1                                                                           |
| tPA leader                  | none                                                   |                                                                                  |
| Multiple cloning site       | 1648-1786<br>(Bam HI-Avr II)                           | pGA1.1, addition of Bam HI<br>and Xho I sites to pGA1.1<br>multiple cloning site |
| BGHpA                       | 1789-2011                                              | pGA1.1                                                                           |
| Lambda T0 terminator        | 2012-2046                                              | pGA1.1                                                                           |
| Kanamycin resistance        | 2067-2861                                              | pGA1.1                                                                           |
| ColE1 origin of replication | 3247-3920                                              | pGA1.1                                                                           |

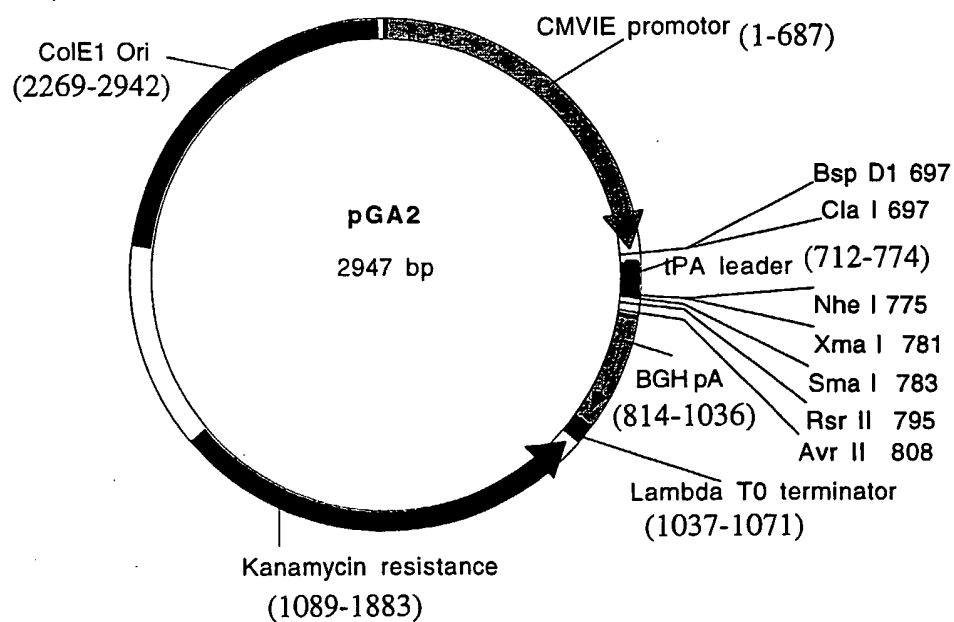


Fig. 5

Figure 6a. pGA2 Sequence

cgacaatattggctattggccattgcatacgttgatctatatcataatatgtacatttatattggctcatgtccaatat  
gaccgccatggttgacattgattattgactagttattaatagtaatacaattacggggtcatttagttcatagcccatatatg  
gagttccgcgttacataacttacggtaaatggcccgctggctgaccgccaacgaccccgccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaaactgccact  
tggcagtacatcaagtgtatcataatgccaaagtccgccccctattgacgtcaatgacggtaaatggcccgctggcattat  
gcccagtacatgaccttacgggaactttcctacttggcagtacatctacgtattagtcacgtctattaccatgggtgatgcg  
gttttggcagtacaccaatgggcgtggatagcgggttgactcacggggatttccaagtctccacccattgacgtcaatg  
ggagtttgttttggcaccaaaatcaacgggaactttccaaaatgtcgtaataaccccgccccgttgacgcaaatgggcggg  
aggcgtgtacgggtgggagggtctatataagcagagctcgtttagtgaactcattctatcgatgcttgcaatcatggatgca  
atgaagagaggggctctgctgtgtgctgctgctgtgctggtggagcagtcctcgtttcggctagccccgggtgataaacggaccg  
cgcaatccctaggctgtgcttctagttgccagccatctgtgtttgcccccccccgctgcttcccttgaccctgggaagg  
tgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctggggg  
gtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgagggtgggctctatataa  
aaaacgccccggcggaaccgagcgttctgaacgctagagtcgacaaattcagaagaactcgtcaagaaggcagatagaagg  
cgatcgctgcgaatcgggagcggcgataaccgtaaagcacagggaagcggtcagccattcgcgcgaagctcttcagca  
atatcacgggtagccaacgctatgtcctgatagcggctgtccacacccagccggccacagtcgatgaatccagaaaagcg  
gccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgccgtcgggcatgctcgct  
tgagcctggcgaaacagttcgggtggcgagccctgatgctcttcgtccagatcatcctgatcgacaagaccggttcc  
atccgagtagctgctcgctcgatgcatgctgttctcgttgggtggatggcaggttagccggatcaagcgtatgcagccg  
ccgcatgcatcagccatgatggatactttctcggcaggagcaaggtgagatgacaggagatcctgccccggcacttcgc  
ccaatagcagccagtccttcccgcttcagtgacaacgtcgagcacagctgcgcaaggaaacgcccgtcgtggccagccac  
gatagccgcgtgctcgtcttcaggttcatcagggcacgggacaggtcggctctgacaaaaagaacggggcgccccctg  
cgctgacagccggaacacggcgccatcagagcagccgattgtctgttgtgccagtcatacgccaatagccttccaccc  
aagcggccggagaacctgctgcaatccatcttgttcaatcatgcgaaacgatcctcatcctgtctcttgatcagatctt  
gatccccctgcccacatcagatccttggcggaagaaagccatccagtttactttgcagggttcccaaccttaccagaggg  
cgccccagctggcaattccgggtcgttgcgtgtccataaaacgcccagtcagctatcgccatgtaagccactgcaag  
ctacctgctttctcttgcgcttgcgttttcccttgcagatagcccagtagctgacattcatccgggggtcagcacccgt  
ttctgcggactggctttctacgtgaaaaggatctaggtgaagatcctttttgataatctcatgacaaaaatcccttaacg  
tgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgcgcgtaa  
tctgctgcttgcaaacaaaaaaaccacgctaccagcgggtggtttgtttgcccggatcaagagctaccaactcttttccg  
aaggtaaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaa  
ctctgtagcaccgcctacatacctcgtctgctaactcctgttaccagtggtgctgcccagtgccgataagtcgtgtctta  
ccgggttggactcaagacgatagttaccggataaggcgagcggctcgggctgaacgggggggttcgtgcacacagcccagc  
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggagagaaa  
ggcggaacaggtatccggtaagcggcaggggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggatc  
tttatagtcctgtcgggtttcggccacctctgacttgagcgtcgatttttgtgatgctcgtcagggggggcgagcctatgg  
aaaaacgccagcaacgcggcccttttacggttcctggccttttgcgtggccttttgcctcacatgttgt

**Figure 6b. PGA2 functional regions**

| <b>Functional region</b>    | <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Origin of sequence</b> |
|-----------------------------|----------------------------------------------------------------|---------------------------|
| CMVIE promoter              | 1-687                                                          | pGA1                      |
| tPA leader                  | 712-774                                                        | pGA1                      |
| Multiple cloning site       | 698-808<br>Cla I-Avr II                                        | pGA1                      |
| BGHpA                       | 814-1036                                                       | pGA1                      |
| Lambda T0 terminator        | 1037-1071                                                      | pGA1                      |
| Kanamycin resistance        | 1089-1883                                                      | pGA1                      |
| ColE1 origin of replication | 2269-2942                                                      | pGA1                      |

Figure 7a. pGA2.1 Sequence

cgacaatattggctattggccattgcatacgttgatctatatcataatatgtacatttatattggctcatgtccaatat  
gaccgccatgttgacattgattattgactagttattaatagtaatacaattacggggtcattagttcatagcccatatatg  
gagttccgcgttacataaacttacggtaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaaactgcccact  
tggcagtacatcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggcccgctggcattat  
gcccagtagcatgaccttacgggaactttcctacttggcagtacatctacgtatttagtcatcgctattaccatgggtgatgag  
gttttggcagtagaccaatgggcgtggatagcgggttgactcacggggatttccaagtctccacccattgacgtcaatg  
ggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgttaataacccccgccggttgacgcaaatggcggt  
aggcgtgtacgggtgggaggtctatatataagcagagctcgttttagtgaactcattctatcgatgcttgcaatcatggatgca  
atgaagagagggctctgctgtgtgctgtgctgtgtggagaattcttcggttcggctgctgtgtgtggagaattcttcg  
tttcggctagccccgggtgataaacggacgcgcaatccctaggtgtgccttctagttgcccagccatctgttgtttgcc  
cctccccgtgccttcttgacctggaaggtgccactcccactgtcctttcctaataaaaatgaggaaatgcatcgcat  
tgtctgagtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcag  
gcatgtgggggtgaggtgggtctatatataaaaaacgccccggcgcaaccgagcgttctgaacgctagagtcgacaaat  
cagaagaactcgtcaagaaggcgatagaaggcgatgcgtgcgaatcgggagcgcgataaccgtaaagcacgaggaagcg  
gtcagcccattcgccgccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcggctgtccacaccca  
gcccggccacagtcgatgaatccagaaaagcgccattttccaccatgatattcggcaagcaggcatcgccatgggtcacg  
acgagatcctcgccgtcgggcatgctcgcttgagcctggcgaaacgcttcggctggcgagccccctgatgctcttcgctc  
cagatcatcctgatcgacaagacgggttccatccgagtagctgctcgctcgatgcatggtttcgtttggtggtcgaatg  
ggcaggtagccggatcaagcgtatgcagccgcccattgcatcagccatgatggatactttctcggcaggagcaaggtga  
gatgacaggagatcctgccccggcacttcgccccaatagcagccagtccttcccgcttcagtgacaacgtcgagcacagc  
tgcgcaaggaaacgcccgtcgtggccagccacgatagccgctgctcgtcttgagttcattcagggcacccggacaggt  
cggctcttgacaaaaagaacccggcgccccctgcgctgacagccggaacacggcgcatcagagcagccgattgtctgtgt  
gcccagtcatagccgaatagcctctccacccaagcggcgaggagaacctgctgcaatccatcttgttcaatcatgcaaaa  
cgatcctcatcctgtctcttgatcagatcttgatccccctgcccattcagatccttggcggaagaaagccatccagttta  
ctttgcagggcttcccaaccttaccagagggcgccccagctggcaattccggttcgcttgctgtccataaaaaccgcccag  
tctagctatcgccatgtaagcccaactgcaagctacctgctttctcttgcgcttgctgtttcccttgctccagatagccca  
gtagctgacattcatccgggtcagcacccgtttctgcggactggctttctacgtgaaaaggatctaggtgaagatccttt  
ttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatcaaagga  
tcttcttgagatccttttttctgcgcgtaatctgctgcttgcaaaacaaaaaaccaccgctaccagcgggtggtttgttt  
gcccgatcaagagctaccaactctttttccgaaggttaactggcttcagcagagcgcagataccaaaatactgttcttctag  
tgtagccgtagtttaggccaccacttcaagaactctgtagcaccgcctacataacctcgctctgctaatcctgttaccagt  
gctgtgccagtgggcgataagtcgtgtcttaccgggttggaactcaagacgatagttaccggataaggcgagcgggtcggg  
ctgaacggggggttcgtgcacacagcccagcttgagcgaaacgacctacaccgaactgagatacctacagcgtgagctat  
gagaaagcgccacgcttcccgaaggagaaaggcgagaggtatccggttaagcggcagggtcggaacaggagagcgacg  
aggagcttccagggggaaacgcctggtatctttatagtcctgtcggtttcgccacctctgacttgagcgtcgattttt  
gtgatgctcgtcagggggcgaggcctatggaaaaacgccagcaacgcggcccttttacggttccctggccttttgctggg  
cttttgcctcacatgttgt

**Figure 7b. PGA2.1 functional regions**

| <b>Functional region</b>    | <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Origin of sequence</b>                                              |
|-----------------------------|----------------------------------------------------------------|------------------------------------------------------------------------|
| CMVIE promoter              | 1-687                                                          | PGA2                                                                   |
| tPA leader                  | None                                                           |                                                                        |
| Multiple cloning site       | 698-839<br>Cla I-Avr II                                        | PGA2, addition of two EcoR I<br>sites to pGA2 multiple cloning<br>site |
| BGHpA                       | 842-1064                                                       | PGA2                                                                   |
| Lambda T0 terminator        | 1065-1099                                                      | PGA2                                                                   |
| Kanamycin resistance        | 1914-1120                                                      | PGA2                                                                   |
| ColE1 origin of replication | 2300-2973                                                      | PGA2                                                                   |



Figure 8a. pGA2.2 Sequence

cgacaatatattggctatttgccattgcatacgttgatctatatcataatatgtacatttatattggctcatgtccaatat  
gaccgccatgttgacattgattattgactagttattaatagtaataacacggggtcattagttcatagcccatatatg  
gagttccgctgttacataacttacggtaaatggccgcctggctgaccgccaacgacccccgccattgacgtcaataat  
gacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaaactgccact  
tggcagtacatcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggccgcctggcattat  
gccagtacatgaccttacgggaactttcctacttggcagtacatctacgtatttagtcatcgctattaccatgggtgatgcg  
gttttggcagtacaccaatgggcgtggatagcgggttgactcacggggatttccaagtctccaccccatgacgtcaatg  
ggagtgtgttttggcaccaaaatcaacgggaactttccaaaatgtcgttaataacccccgcccgttgacgcaaatgggcggg  
aggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaactcattctatggatccttgctcgagtggaatgca  
atgaagagaggggtctgctgtgtgctgctgctgtgtggagaattcttcgcttcggctgctgctgtgtggagaattcttcg  
tttcggctagccccgggtgataaacggaccgcgaatccctaggctgtgccttctagtgtgcagccatctgttggttgcc  
cctccccgctgccttccctgaccctggaagggtgccactcccactgtccttccctaataaaatgaggaaatgcatcgcat  
tgtctgagtaggtgtcattctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcag  
gcatgctggggatgcgggtgggtctatatataaaaaacgcccggcgccgaaccgagcgttctgaacgctagagtcgacaaat  
cagaagaactcgtcaagaaggcgatagaaggcgatgcgtgcggaatcgggagcggcgataccgtaaaagcacgaggaagcg  
gtcagcccatctcgccgccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcgggtctgccacacca  
gccggccacagtcgatgaatccagaaaagcgccattttccaccatgatattcggcaagcaggcatcgccatgggtcacg  
acgagatcctcgccgtcgggcatgctcgcccttgagcctggcgaaacagttcggtggcgcgagccccctgatgctcttcgct  
cagatcatcctgatcgacaagaccggcttccatccgagtagctgctcgctcgatgcgatgtttcgcttgggtgggtcgaaatg  
ggcaggtagccggatcaagcgtatgcagccgcccagcttgcatcagccatgatggatactttctcgccaggagcaagggtga  
gatgacaggagatcctgccccggcacttcgcccataagcagccagtccttccccgcttcagtgacaacgtcgagcacagc  
tgcgcaaggaaacgcccgtcgtggccagccacgatagccgcgtgcctcgtcttcgagttcattcagggcaccggacaggt  
cggtcttgacaaaaagaaccggcgcccctgcgtgacagccggaacacggcgccatcagagcagccgattgtctgttgt  
gccagtcatagcgaatagcctctccacccaagcgccgggagaacctgcgtgcaatccatcttgttcaatcatgcgaaa  
cgatcctcatcctgtctcttgatcagatcttgatccccctgcgcatcagatccttggcggaagaaagccatccagttta  
ctttgcagggttcccaacctaccagagggcgcccagctggcaattccggttcgcttgctgtccataaaaaccgcccag  
tctagctatcgccatgtaagcccaactgcaagctacctgctttctctttgctgcttgctgttcccttgctccagatagccca  
gtagctgacattcatccgggggtcagcacggtttctgcggactggctttctacgtgaaaaggatctaggtgaagatccttt  
ttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaagga  
tcttcttgagatccttttttctgcgcgtaatctgctgcttgcaacaaaaaaaccaccgctaccagcgggtgggttggtt  
gccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaataactgttctctag  
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaactcctgttaccagtg  
gctgctgccagtggcgataagtcgtgtcttaccgggttggaactcaagacgatagttaccggataaaggcgagcggtcggg  
ctgaacgggggggttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctat  
gagaaagcgccacgcttcccgaaggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg  
agggagcttccagggggaaacgcctgggtatctttatagtcctgtcggggttcgccacctctgacttgagcgtcgattttt  
gtgatgctcgtcagggggggcggagcctatggaaaaacgcccagcaacgcggcccttttacgggttcctggccttttgctggc  
cttttgctcacatgttgt

Figure 8b. PGA2.2 functional regions

| Functional region           | Position within sequence<br>(starting-and ending pos.) | Origin of sequence                                                               |
|-----------------------------|--------------------------------------------------------|----------------------------------------------------------------------------------|
| CMVIE promoter              | 1-687                                                  | PGA2.1                                                                           |
| tPA leader                  | None                                                   |                                                                                  |
| Multiple cloning site       | 698-839<br>Bam HI-Avr II                               | PGA2.1, addition of Bam HI<br>and Xho I sites to pGA2.1<br>multiple cloning site |
| BGHpA                       | 842-1064                                               | PGA2.1                                                                           |
| Lambda T0 terminator        | 1065-1099                                              | PGA2.1                                                                           |
| Kanamycin resistance        | 1914-1120                                              | PGA2.1                                                                           |
| ColE1 origin of replication | 2300-2973                                              | PGA2.1                                                                           |

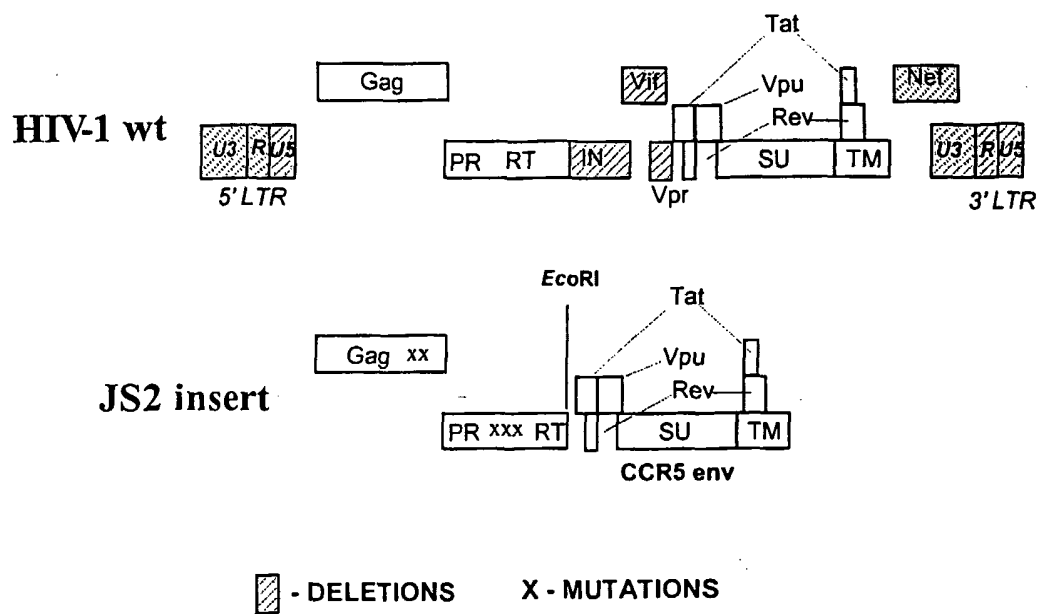


Fig. 9

Figure 10a. pGA2/JS2 sequence

atcgatgcaggactcggcttgctgaagcgcgcacggcaagaggcagggggcgagctgggtgggtacgccaaaaattttg  
actagcggaggctagaaggagagagatgggtgcgagagcgctcagttatgaagcgggggagaattagatcgatgggaaaaa  
ttcgggttaaggccagggggaaagaaaaatatataatataaacatatagtagggcaagcaggagctagaacgattcgc  
gttaatcctggcctgttagaaacatcagaaggctgtagacaaatactgggacagctacaacatcccttcagacaggatc  
agaagaacttagatcattatataatacagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaagg  
aagctttagacaagatagaggaagagcaaaacaaaagtaagaaaaagcacagcaagcagcagctgacacaggacacagc  
agttaggtcagccaaaattaccctatagtcgagaacatccaggggcaaatggtacatcaggccatatacctagaacttt  
aatgcatgggtaaaagttagtagaagagaaggctttcagccagaagtaatacccatgttttcagcattatcagaaggag  
ccacccacaagatttaaacacacatgctaaacacagtggggggacatcaagcagccatgcaaatgttaaaagagaccatc  
aatgaggaagctgcagaatgggatagagtacatccagtcgcatgcagggcctattgcaccaggccagatgagagaaccaag  
gggaagtgcacatagcaggaactactagtagcccttcaggaacaaataggatggatgacaaataatccacatcccagtag  
gagaaatttataaaagatggataatcctgggattataaaaatagtgaagatgtatagccctaccagcatcttgacata  
agacaaggacaaaaagaaccttttagagactatgtagaccggttctataaaactctaagagccgagcaagcttcacagga  
ggtaaaaaattggatgcagagaaccttgggtgggtccaaaatgcgaaccagattgttaagactattttaaaagcattgggac  
cagcggctacactagaagaatgatgcagcatgtcaggggagtaggaggaccggccataaggcaagagttttgggtgaa  
gcaatgagccaagtaacaaatacagctaccataatgatgcagagaggcaatttttaggaaccaaagaagatgggttaagag  
cttcaatagcggcaagaaggccacacagccagaaatgcagggcccttaggaaaaagggcagctggaaaagcggaaagg  
aaggacaccaaatagaagattgtactgagagacaggctaattttttagggaagatctggccttctacaagggaaggcca  
gggaattttcttcagagcagaccagagccaacagccccaccatttcttcagagcagaccagagccaacagccccaccaga  
agagagcttcaggctcgggttagagacaacaactccccctcagaagcaggagccgatagacaaggaactgtatctcttaa  
ctccctcagatcactctttggcaacgaccctcgtcaataaaagataggggggcaactaaaggaagctctattagata  
caggagcagatgatacagtagtagaagaatgagtttgccaggaagatggaaacaaaaatgatagggggaattggaggt  
tttatcaaagtaagacagtagatgcagatactcatagaaatctgtggacataaaagctataggtagcagtagtaggacc  
tacacctgtcaacataattggaagaaatctgttgactcagattgggtgcaactttaaattttccattagccctattgaga  
ctgtaccagtaaaattaaagccaggaatggatggcccaaaaggttaacaatggccattgacagaagaaaaataaaagca  
ttagtagaatttgtacagaatggaaaaggaagggaataatcaaaaattgggctgagaatccatacaataactccagt  
atttggcataaagaaaaaagacagtagtaaatggagaaaattagtagatttcagagaacttaataagagaactcaagact  
tctgggaagttcaattaggaataccacatcccgcagggttaaaaaagaaaaaatcagtaacagtagtggtgtgggtgat  
gcatatttttcagttcccttagatgaagacttcaggaagtatactgcatttaccatacctagtagataaaacaatgagacacc  
agggattagatatcagtagcaatgtgcttccacagggatggaaaagatcaccagcaatattccaaagtagcatgacaaaa  
tcttagagccttttaaaaaacaaaatccagacatagttatctatcaatacatgaacgatttgtatgtaggatctgactta  
gaaatagggcagcatagaacaaaaatagaggagctgagacaacatctgttgaggtggggacttaccacaccagacacaaa  
acatcagaagaagacccctcattcctttggatgggttatgaaactccatccctgataaatggacagtagcagcctatagtgctgc  
cagaaaaagacagctggactgtcaatgacatacagaagttagtggggaaattgaataccgcaagtcagatttaccaggg  
attaaagtaaggcaattatgtaaactccttagaggaaacaaagcactaacagaagtaataccactaacagaagaagcaga  
gctagaactggcagaaaacagagagattctaaaagaaccagtagcatggagtgattatgacccatcaaaagacttaatag  
cagaaatacagaagcaggggcaaggccaatggacatacaaatttatcaagagccatttaaaaaatctgaaaacaggaaaa  
tatgcaagaatgaggggtgcccacactaatgatgtaaaacataaacagagggcagtgcaaaaaataaccacagaaagcat  
agtaatatggggaagactcctaataatttaactacccatacaaaaaggaaacatgggaaacatgggtggacagagtagttggc  
aagccacctggattcctgagtgagggtttgttaataccctccttttagtgaaattatggtagccagttagagaagaaccc  
atagtaggagcagaaaccttctatgtagatggggcagctaacagggagactaaattaggaaaagcaggatattgtactaa  
caaagggaagacaaaaggttgtccccctaactaacacacaaatcagaaaactcagttacaagcaatttatctagctttgc  
aggattcaggattagaagtaaacatagtagaacagactcacaatatgcattaggaatcattcaagcacaccagataaaagt  
gaatcagagttagtagcaataatagagcagtttaaaaaaggaaaaggtctatctggcatgggtaccacaccagacacaaa  
aggaattggaggaatgaacaagtagataaattagtcagtgctggaaatcaggaaaatactatttttagatggaaatagata  
aggcccaagtagaacattagaattctgcaacaactgctgtttatccattttcagaattgggtgtcgacatagcagaatag  
gcgttactcgacagaggagagcaagaaatggagccagtagatccttagactagagccctggaaagcatccagggaagtcagcc  
taaaactgcttgtagcaatgtctattgtaaaaagtggtgctttcattggccaagtttggttcataacaaaagccttaggca  
tctcctatggcaggaagaagcggagacagcgacgaagacctcctcaagacagtcagactcatcaagtttctctatcaaag  
cagtaagtagtaaatgtaatgcaacctttacaaatattagcaatagtagcatttagtagtagcagcaataatagcaatag  
tgtgtggaccatagtagtaataataggaataatgaagcaaaagaaaatagacaggttaattgataggataaacag  
aaagagcagaagacagtggaatgaaagtgaaggggatcaggaagaattatcagcactgtggaaatggggcatcatgct  
ccttgggatgttgatgatctgttagtgctgtagaaaatttgggggtcacagtttattatggggtagctgtgtgaaagaag  
caaccaccactctattttgtgcatcagatgctaaagcatatgatacagaggtacataatgtttggggcacacatgcctgt  
gtaccacagaccccaaccacaagaagtagtattggaaaatgtgacagaaaattttaacatgtggaaaaataacatgggt  
agaacagatgcattagagatataatcagtttatgggatcaaagcctaaagccatgtgtaaaaataacccactctgtgtta  
cttttaattgcactgatttgaggaatgttactaatcaataatagtagtgagggaatgagaggagaataaaaaactgc  
tcttcaatatcaccacaagcagaagataaagtgaaagactatgcactttttatagacttgatgtagtaccaat  
agataatgataatactagctataggttgataaattgtaatacctcaaccattacacaggcctgtccaaaggtatcctttg  
agccaattcccatacattattgtaccccggtggttttgcgattctaaagtgtaaagacaagaagttcaatggaaacggg  
ccatgtaaaaatgtcagcagtagtaaatgtacacatggaaataggccagtagtgtcaactcaactgctgttaaatggcag

tctagcagaagaagaggttagtaattagatctagtaatttcacagacaatgcaaaaaacataatagtagcagttgaaagaat  
ctgtagaaatataattgtacaagacccaacaacaatacaaggaaaagtatacatataggaccaggaagagcattttatataca  
acaggagaaataataggagatataagacaagcacattgcaacattagtagaacaataatggaataaacacttttaatacaaat  
agctacaaaattaaaagaacaatttgggaataataaaaacaatagtcctttaatcaatcctcaggaggggacccagaaatttg  
taatgacacagttttaaattgtggagggggaatttttctactgtaattcaacacaactgtttaatagtagtacttggaaattttaa  
ggtagtctggaaatttaacacaatcgaaatggtagtgaaggaaatgacactatcacactcccatgtagaataaaacaaattat  
aaatatgtggcaggaagtaggaaaagcaatgtatgcccctcccatcagaggacaaattagatgctcatcaaatattacag  
ggctaataattaacaagagatgggtggaactaacagtagtgggtccgagatcttcagacctgggggaggagatagagggac  
aattggagaagtgaattatataaataaaagtagtaaaaattgaaccattaggagtagcaccaccaaggcaaaaagaag  
agtgggtgcagagagaaaaagagcagtggaacgataggagctatgttcttgggttcttgggagcagcaggaagcacta  
tgggagcagcgtcaataacgctgacggtagcggccagactattatgtctggtatagtgcaacagcagaacaatttgcctg  
agggctattgaggcgcaacagcatctgttgcaactcacagctctggggcatcaagcagctccaggcaagagctcctggctct  
ggaaagatcacctaagggtatcaacagctcctagggatttgggggtgctctggaaaactcatctgcaccactgctgtgctct  
ggaatgctagttggagtaataaaactctggatattgatttgggataacatgacctggatggagtggaagagaaatcgaa  
aattacacaggcttaataatacaccttaattgaagaatcgacagaaccaacaagaaaagaatgaacaagacttattagcatt  
agataagtgggcaagtgtgtggaattgggttgacatatcaaatgggtgtggtgtataaaaaatcttcataatgtagtag  
gaggcttgtagaggtttaagaatagttttactgtactttctatagtaaatagagttaggcaggatactcaccatttgc  
tttcagaccacactcccagccccgaggggacccgacaggccgaaggaaatcgaaagaagggtggagacagagacagaga  
cagatccgtgcatagtagtggtggtccttagcacttatctgggacgatctgcggagcctgtgctctctcagctaccacc  
gcttgagagacttactcttgattgtaacgaggattgtggaactctgggacgcaggggggtgggaagccctcaaatattgg  
tggaatctcctacagtagttggagtcaggagctaaagaatagtgctgttagcttgctcaatgccacagctatagcagtagc  
tgaggggacagatagggttatagaagtagtagacaaggagcttatagagctattcgccacatacctagaagaataagacagg  
gcttggaaggatttgcataaagatgggtgggttagcctcggggtgataaaacgacccgcgaatcctcaggctgtgctctc  
tagttgcccagccatctgtgtgttgccctccccgtgcttctctgacctggaaagggtgacctcccactgtcctttcct  
aataaaatgaggaaattgcatcgcatgtctgtagtaggtgtcattctattctgggggggtgggggtggggcaggacagcaag  
ggggaggattgggaagacaatagcaggcatgtctggggtgctggtggtctctataaaaaacgcccggcggaacccgagc  
gttctgaacgctagagtcgacaaattcagaagaactcgtcaagaaggcgatagaaggcgatgctgctgcaatcgggagcg  
gcgataaccgtaaacgacgaggaagcgggtcagcccatcgccgccaagctcttcagcaatatcacgggttagccaacgctat  
gtcctgtagcgggtctgccacaccagccggccacagctcgatgaatccagaaaagcggccattttccaccatgatattcg  
gcaagcaggcatcgccatcgccatcacgacagatcctcgccgtcgggcatgctcgcttgagcctggcgcaacagttcggt  
ggcgcgagccctgatgctctctcgccagatcatcctgatcgacaagaccggcttccatccgagtagctgctcgctcgat  
gcgtagtttgccttggtggtcgaaatgggcaggtagccggatcaagcgtagtcagccgcatgcatcagccatgatgg  
atactttctcggcaggagcaagggtgagatgacaggagatcctgccccggcacttcgcccataagcagccagctccctccc  
gcttcagtgacaacgctcgagcacagctgcgcaaggaaacgcccgtcggtggccagccagatagcccgctgctcctgcttg  
cagttcattcagggcaccggacaggtcggtcttgacaaaagaacccggcgccctgctgctgacagccggaacacggcg  
catcagagcagccgattgtctgtgtgcccagctcatagccgaatagccttccacccaagcggcgccggaacactgcgtgc  
aatccatctgttcaatcatgcaaacgatcctcatcctgtctcttgatcagatcttgatccctgcgccatcagatcct  
tgggcgcgagaaagccatccagtttactttgcagggttcccaaccttaccagagggcgccccagctggcaattccgggt  
cgcttgctgtccataaaacggccagcttagctatcgccatgtaagcccaactgcaagctacctgctttctctttgcgctt  
gcgttttcccttgctccagatagccagtagctgacattcatccgggtcagcaccgtttctgcggactggctttctacgt  
gaaaaggatctaggtgaagatccttttgataatctcatgacaaaatcccttaacgtgagttttcgttccactgagcgt  
cagaccccgtagaaaagatcaaaggatccttcttgatccctttttctgcgctgaatctgctgcttgcaaaacaaaaaa  
ccaccgctaccagcggtgggtttgtttgcccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagc  
gcagataccaaatactgtccttctagtgtagccgtagtttaggccaccacttcaagaactctgtagcaccgctacatacc  
tcgctctgctaatacctgttaccagtggtgctgctgcaagtggtgataagtcgtgtcttaccgggttggaactcaagacgatag  
ttaccggataaggcgagcggtcggttgaaacgggggggttcgtgcacacagccagcttgagcgcaacgacctacaccga  
actgagatacctacagcgtgagctatgagaaagcgccacgcttccgaaggagaaaggcggaacaggtatccggtaagcg  
gcaggggtcggaacaggagagcgacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcggggttcgc  
cactctgacttgagcgtcgatttttgtgatgctcgtagggggggcgagcctatggaaaacgcccagcaacgccccttt  
ttacgggttccctggggttttgcctggccttttgcctcacatgttgcgacggacaatattggctattggccattgcatacgtt  
gtatctatatcataaatatgtacatttatattgggtcatgtccaatatgaccgcatgttgacattgattattgactagtt  
attaatagtaatacaattacggggtcatttagttcatagcccatatagaggttccgcttacataacttacggtaaatggc  
ccgctcgtgaccgcccacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggact  
ttccattgacgtcaatgggtggagtaatttacgggtaaactggccacttggcagtagacatcaagtgtatcatatgccaagtcc  
gcccctattgacgtcaatgacggttaaatggcccgcctggcattatgccagtagacatgaccttacgggactttcctacttg  
gcagtagacatctacgtattagtcagctattaccatgggtgatgcgggtttggcagtagaccaatggggcggttagcggt  
ttgactcacggggatttccaagtctccacccattgacgtcaatgggagttgttttggcaccaaaatcaacgggacttt  
ccaaaatgtcgtataaaccgcccccggttgacgcaaatgggcggtaggcggtgacgggtgggaggtctatataagcagagc  
tcgttttagtgaaccgtcagatcgc

Fig. 10a

**Figure 10b. pGA2/JS2 Functional Regions**

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 106-1644                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | HIV-1 IIIb (BH10)                        |
| 1401-3620                                                      | Pol, inactivating point mutations in reverse transcriptase, deletion of integrase               | HIV-1 IIIb (BH10)                        |
| 3708-3922 and 6247-6337                                        | Tat exons #1 and #2                                                                             |                                          |
| 3847-3922 and 6247-6521                                        | Rev exons #1 and #2                                                                             |                                          |
| 3939-4184                                                      | Vpu                                                                                             |                                          |
| 4102-6663                                                      | Env                                                                                             | HIV-1 ADA                                |
| 6664-9544                                                      | Plasmid vector                                                                                  | PGA2, Gen Bank accession # AF425298      |

**Figure 10c. JS2 Mutation chart**

| <b>Codon change</b>          | <b>Amino acid change</b> | <b>Region</b>            | <b>Function</b>                                 |
|------------------------------|--------------------------|--------------------------|-------------------------------------------------|
| to AGC Position<br>1279-1281 | C392S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1288-1290 | C395S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1342-1344 | C413S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1351-1353 | C416S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AAC Position<br>2454-2456 | D185N                    | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC Position<br>2697-2699 | W266T                    | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG Position<br>3333-3335 | E478Q                    | Rnase H                  | Inactivation of<br>RNase H activity             |

Figure 11a. pGA2/JS7 Sequence

ATCGATGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTG  
ACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGAGAAATTAGATCGATGGGAAAAA  
TTCGGTTAAGGCCAGGGGAAAGAAAAATATAAATTAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTTCGCA  
GTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATC  
AGAAGAACTTAGATCATTATATAATACAGTAGCAACCTCTATTGTGTGCATCAAAGGATAGAGATAAAGACACCAAGG  
AAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAGTAAGAAAAAGCACAGCAAGCAGCAGCTGACACAGGACACAGC  
AATCAGGTCAGCCAAAATTACCTTATAGTGCAGAACATCCAGGGGCAATGGTACATCAGGCCATATCACCTAGAACTTT  
AAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTAGCCCAAGTGTATCCCATGTTTTTTCAGCATTATCAGAAGGAG  
CCACCCCAACAGATTTAAACACCATGCTAAACACAGTGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATC  
AATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAG  
GGGAAGTGACATAGCAGGAACCTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAG  
GAGAAATTTATAAAGATGGATAATCCTGGGATTAATAAAAATAGTAAGAATGTATAGCCCTACCAGCATCTGGACATA  
AGACAAGGACCAAAAGAACCTTTAGAGACTATGTAGACCGGTTCTATAAACTCTAAGAGCCGAGCAAGCTTCACAGGA  
GGTAAAAAATTTGGATGACAGAAACCTTGTGTGTCAAAAATGCGAACCCAGATTGTAAGACTATTTTAAAGCATTGGGAC  
CAGCGGCTACACTAGAAGAAATGATGACAGCATGTACAGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTGGCTGAA  
GCAATGAGCCAAGTAACAAATTCAGCTACCATAATGATGCAGAGAGGCAATTTTAGGAACCAAGAAAGATTGTTAAGAG  
CTTCAATAGCGGCAAAAGAAAGGGCACACAGCCAGAAATTGCAGGGCCCTAGGAAAAAGGGCAGCTGGAAAAGCGGAAAGG  
AAGGACACCAATGAAAGATTGTACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCTTACAAGGGAAGGCCA  
GGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGTAGAGACAACAACCTCC  
CCCTCAGAAGCAGGAGCCGATAGACAAGGAAGCTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGT  
CACAATAAAGATAGGGGGGCACTAAAGGAAGCTCTATTAGCCACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTT  
TGCCAGGAAGATGGAACCAAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATA  
GAAATCTGTGGACATAAAGCTATAGGTACAGTATAGGTACAGCTACACCTGTCAACATAATTGGAAGAAATCTGTTGAC  
TCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAATTAAGCCAGGAATGGATGGCC  
CAAAAGTTAAACAATGGCCATTGACAGAAGAAAAGATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAGGAAGGG  
AAAATTTCAAATTTGGGCTGAAAATCCATACAATCTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAG  
AAAATTTAGTAGATTTAGAGAACTTAATAAGAGAAGCTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAG  
GGTTAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTCAGTTCCTTAGATGAAGACTTCAGG  
AATATACTGCATTTACCATACCTAGTATAAACAATGAGACACCCAGGATTAGATATCAGTACAATGTGCTTCCACAGGG  
ATGGAAGGATCACCAGCAATATTCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAAACAAATCCAGACATAG  
TTATCTATCAATACATGAACGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAATAATAGAGGAGCTG  
AGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGGTTA  
TGAATCCATCTGATAAATGGACAGTACAGCCTATAGTGTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGA  
AGTTAGTGGGAAATTAATACCGCAAGTCAGATTTACCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGA  
aCCAAAGCACTAACAGAAGTAATACCCTAACAGAAGAACGAGCTAGAAGTGGCAGAAAAACAGAGAGATTCTAAAGAG  
ACCAGTACATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACAT  
ATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCACACTAATGATGTA  
AAACAATTAACAGAGGCAGTGCAAAAAATAACCAAGAGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTGCC  
CATACAAAAGGAAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGAGTTTGTAAAT  
CCCCTCCTTTAGTGAATTTAGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCA  
GCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTACCCTAACTAACAC  
AACAAATCAGAAAACTCAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACT  
CACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTA  
ATAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCACAAAGGAATGGAGGAATGAACAAGTAGATAAATTAGT  
CAGTGTCTGGAATCAGGAAAGTACTATTTTTAGATGGAATAGATAAGGCCCAAGATGAACATTAGAATCTGCAACAACATG  
CTGTTTATCCATTTAGAAATTTGGGTGTGCATAGCAGAAATAGGCGTTACTCGACAGAGGAGAGCAAGAAATGGAGCCAG  
TAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGT  
TGCTTTTCAATGCCAAGTTTGTTCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAG  
AGCTCCTCAAGACAGTCAGACTCATCAAGTTTCTCTATCAAAGCAGTAAGTAGTAAATGTAATGCAACCTTTACAAATAT  
TAGCAATAGTAGCATTAGTAGTAGCAGCAATAATAGCAATAGTTGTGTGGACCATAGTATTCATAGAATATAGGAAAAATA  
TTAAGACAAAAGAAAAATAGACAGGTTAATTGATAGGATAACAGAAAGAGCAGAAGACAGTGGCAATGAAAGTGAAGGGGA  
TCAGGAAGAATTATCAGCACTTGTGGAATGGGGCATCATGCTCCTTGGGATGTTGATGATCTGTAGTGTCTGTAGAAAAT  
TTGTGGGTACAGTTTATTATGGGGTACCTGTGTGGAAGAAAGCAACCCACTCTATTTTGTGCATCAGATGCTAAAGC  
ATATGATACAGAGGTACATAATGTTTGGGGCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGG  
AAAATGTGACAGAAAATTTTAACATGTGGAATAAATCATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGAT  
CAAAGCCTAAAGCCATGTGTAAAATTAACCCACTCTGTGTACTTTAAATTGCACTGATTTGAGGAATGTTACTAATAT  
CAATAATAGTAGTGAGGGAATGAGAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGCATAAGAGATAAGGTGA  
AGAAAGACTATGCACTTTTTATAGACTTGATGTAGTACCAATAGATAATGATAACTAGCTATAGGTTGATAAATGT  
AATACCTCAACCATTACACAGGCCTGTCCAAGGTATCCTTTGAGCCAATCCCATACATTATTGTACCCCGGCTGTTT  
TGCGATTCTAAAGTGTAAGACAAAGAGTTCAATgAACAGGGCCATGTAAAAATGTGAGCACAGTACAATGTACACATG  
GAATTAGGCCAGTAGTGTCAACTCACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTAGTAAT



TTACAGACAATGCAAAAAACATAATAGTACAGTTGAAAGAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATAC  
AAGGAAAAGTATACATATAGGACCAGGAAGAGCATTTTATACAACAGGAGAAATAATAGGAGATATAAGACAAGCACATT  
GCAACATTAGTAGAACAATAATGGAATAACACTTTAAATCAAATAGCTACAAAATTAAAAGAACAATTTGGGAATAATAAA  
ACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTA  
CTGTAAATCAACACAACCTGTTTAAATAGTACTTGGAAATTTAATGGTACTTGGAAATTTAACACAATCGAATGGTACTGAAG  
GAAATGACACTATCACACTCCCATGTAGAATAAAACAAATTATAAATATGTGGCAGGAAGTAGGAAAAGCAATGTATGCC  
CCTCCCATCAGAGGACAAATTAGATGCTCATCAAATATTACAGGGCTAATATTAACAAGAGATGGTGGAACCTAACAGTAG  
TGGGTCCGAGATCTTCAAGCCTGGGGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAA  
AAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAAAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAACGATA  
GGAGCTATGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATAACGCTGACGGTACAGGCCAG  
ACTATTATTGTCTGGTATAGTGAACAGCAGACAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCA  
CAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCTGGCTGTGGAAAGATACCTAAGGGATCAACAGCTCCTAGGGATT  
TGGGGTTGCTCTGGAAAACCTCATCTGCACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAACTCTGGATATGAT  
TTGGGATAACATGACCTGGATGGAGTGGGAAAGAGAAATCGAAAATTACACAGGCTTAATATACACCTTAATTGAAGAAT  
CGCAGAACCAACAGAAAAGAATGAACAAGACTTATTAGCATTAGATAAGTGGGCAAGTTTGTGGAATTGGTTTGACATA  
TCAAATTGGCTGTGGTATGTAAAAATCTTCATAATGATAGTAGGAGGCTTGATAGGTTTAAAGATAGTTTACTGTACT  
TTCTATAGTAAATAGAGTTAGGCAGGGATACTCACCATTGTCTATTTAGACCACCTCCAGCCCGAGGGGACCCGACA  
GGCCCGAAGGAATCGAAGAAGAAGGTGGAGACAGACAGACAGATCCGTGCGATTAGTGGATGGATCCTTAGCACTT  
ATCTGGGACGATCTGCGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGT  
GGAACCTCTGGGACGACGGGGTGGGAAGCCCTCAAATATTGGTGAATCTCCTACAGTATTGGAGTCAGGAGCTAAAGA  
ATAGTGCTGTTAGCTTGTCTCAATGCCACAGCTATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGA  
GCTTATAGAGCTATTCCGCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGATTTTGTCTATAAGATGGGTGGCTAGC  
CCCGGGTGATAAACGGACCGCGCAATCCCTAGGCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTG  
CCTTCCCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAG  
GTGTCACTTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG  
ATGCGGTGGGCTCTATATAAAAAACGCCCGCGGCAACCGAGCGTTCTGAACGCTAGAGTCGACAAATTAGAGAAGCTC  
GGCAAGAAGGCATAGAAGGCGATGCGCTGCGAATCGGGAGCGCGATACCGTAAAGCACGAGGAAGCGGTAGCCCAT  
CGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCTGATAGCGGTCTGCCACACCCAGCCGCCACAG  
TCGATGAATCCAGAAAAGCGGCCATTTCCACCATGATATTCCGCAAGCAGGCATCGCCATGGGTACGACAGATCCTC  
GCGCTCGGGCATGCTCGCCTTGACCTGGCGAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCGTCCAGATCATCCT  
GATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATGGGCAGGTAGCC  
GGATCAAGCGTATGCAGCCGCCGCTTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAG  
ATCCTGCCCCGCACTTCCGCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAA  
CGCCCGTCTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTTATTAGGGCACCGGACAGGTCCGTCTTGACA  
AAAAGAACCAGGGCGCCCTGCGCTGACAGCCGGAACAGCGCGGCATCAGAGCAGCGATTGTCTGTTGCCCCAGTCATA  
GCGGAATAGCCTCTCCACCCAAGCGGCCGGAAGCACTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATC  
CTGTCTCTTGATCAGATCTTGATCCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAAGGC  
TTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCGGTTTCGCTTGTCTCCATAAAACCGCCAGTCTAGCTATCG  
CCATGTAAGCCCACTGCAAGTACCTGCTTTCTTTGCGCTTTCGCTTTTCCCTTGTCCAGATAGCCAGTAGCTGACAT  
TCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGAAAAGGATCTAGGTGAAGATCCTTTTGATAATCTC  
ATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGA  
TCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACACCGCTACCAGCGGTGGTTTGTGTTGCGGATCAAG  
AGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAG  
TTAGGCCACCACTTCAAGAACTCTGTAGACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAG  
TGGCGATAAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGG  
GTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCC  
ACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC  
AGGGGAAACGCTGCTGCTTATAGTCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGT  
CAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTACGGTTCTTGGCCTTTTGTGCTGCTTTTGTCTCAC  
ATGTTGTGACAAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGT  
CCAATATGACCGCCATGTTGACATTGATTATGACTAGTTATTAATAGTAATCAATTACGGGTTCAATTAGTTTATAGCCC  
ATATATGGAGTTCCGCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCCATGACGT  
CAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT  
GCCCACTTGGCAGTACATCAAGTGATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGTAATAGGCCCGCCTG  
GCATTATGCCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG  
TGATGCGGTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAATCACGGGGATTTCAGTCTCCACCCCATGAC  
GTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTGTAATAACCCCGCCCCGTTGACGCAAT  
GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGC

Fig. 11a

Figure 11b. pGA2/JS7 functional regions

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 106-1608                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | HIV-1 HXB2                               |
| 1401-3584                                                      | Pol, inactivating point mutations in protease and reverse transcriptase, deletion of integrase  | HIV-1 HXB2                               |
| 3671-3885 and 6210-6300                                        | Tat exons #1 and #2                                                                             |                                          |
| 3810-3885 and 6210-6884                                        | Rev exons #1 and #2                                                                             |                                          |
| 3902-4147                                                      | Vpu                                                                                             |                                          |
| 4165-6626                                                      | Env                                                                                             | HIV-1 ADA                                |
| 6627-9506                                                      | Plasmid vector                                                                                  | PGA2, Gen Bank accession # AF425298      |

**Figure 11c. JS7 Mutation Chart**

| <b>Codon change</b>          | <b>Amino acid change</b> | <b>Region</b>            | <b>Function</b>                                 |
|------------------------------|--------------------------|--------------------------|-------------------------------------------------|
| to AGC Position<br>1279-1281 | C392S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1288-1290 | C395S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1342-1344 | C413S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1351-1353 | C416S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to GCC position<br>1641-1643 | D25A                     | Protease                 | Inactivation of<br>Protease active site         |
| to AAC Position<br>2418-2420 | D185N                    | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC Position<br>2661-2663 | W266T                    | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG Position<br>3297-3299 | E478Q                    | RNase H                  | Inactivation of<br>RNase H activity             |

Figure 12a. pGA2/JS7.1 Sequence

atcgatgcaggactcggcttgctgaagcgcgcacggcaagaggcgagggcgggcgactgggtgagtacgccccaaatt  
ttgactagcggaggctagaaggagagagatgggtgagagagcgtcagattaaagcgggggagaattagatcgatggg  
aaaaaattcgggttaaggccagggggaaagaaaaatataaattaaaacatatagatgggcaagcaggagctagaa  
cgattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaatactgggacagctacaaccatccct  
tcagacaggatcagaagaacttagatcattatataatcacagtagcaaccctctattgtgtgcatcaaaggatagaga  
taaaagacaccaaggaagcttagacaagatagaggaagagcaaaacaaaagtaagaaaaagcacagcaagcagca  
gctgacacaggacacagcaatcaggtcagccaaaattaccctatagtgagcaacatccaggggcaaatggtacatca  
ggccatatcacctagaactttaaatgcatgggtaaaagtagtagaagagaaggctttcagcccagaagtgtaccca  
tgttttcagcattatcagaaggagccacccccacaagatttaaacaccatgctaaacacagtggggggacatcaagca  
gccatgcaaatgttaaaagagaccatcaatgaggaagctgcagaatgggatagagtgcacagtgcatgcagggcc  
tattgcaccaggccagatgagagaaccaaggggaagtgcacatagcaggaactactagtacccttcaggaacaaatag  
gatggatgacaaataatccacctatcccagtagggagaaatttataaaagatggataatcctgggattaaataaaata  
gtaagaatgtatagccctaccagcattctggacataagacaaggacaaaagaaccctttagagactatgtagaccg  
gttctataaaactctaagagccgagcaagcttcacaggaggtaaaaaattggatgacagaaaccttgttgggtccaaa  
atgcgaaccagattgttaagactattttaaaagcattgggaccagcggctacactagaagaaatgatgacagcatgt  
cagggagtaggaggacccggccataaggcaagagttttggctgaagcaatgagccaagtaacaaattcagctacat  
aatgatgcagagaggcaatttttaggaaccaaagaaagattgttaagagcttcaatagcggcgaaggaaggcacacag  
ccagaaattgcagggcccttaggaaaaagggcagctggaaaagcggaaaggaaggacaccaaatgaaagattgtact  
gagagacaggctaatttttagggaagatctggccttctacaaggggaaggccagggaaatttcttcagagcagacc  
agagccaacagccccaccagaagagagcttcaggtctggggtagagacaacaactccccctcagaagcaggagccga  
tagacaaggaactgtatcctttaacttccctcagatcactctttggcaacgacccctcgtcacaataaagatagggg  
ggcaactaaaggaagctctattagccacaggagcagatgatcacgtattagaagaaatgagtttggcaggaagatgg  
aaacaaaaaatgatagggggaattggaggttttatcaaagtaagacagatgatcagatactcatagaaatctgtgg  
acataaagctataggtacagttatagtaggacctacacctgtcaacataattggaagaaatctgttgactcagattg  
gttgacttttaattttcccattagccctattgagactgtaccagtaaaattaaagccaggaatggatggcccaaaa  
gttaaacaaatggccattgacagaagaaaagataaaagcatttagtagaaatttgtacagagatggaaaaggaaggaa  
aatttcaaaaattgggctgaaaatccatacaatactccagttatttgccataaagaaaaagacagtaactaaatgga  
gaaaattagtagatttcagagaacttaataagagaactcaagacttctgggaagttcaattaggaataccacatccc  
gcagggttaaaaaagaaaaaatcagtaacagtaactggatgtgggtgatgcataatttttcagttcccttagatgaaga  
cttcaggaaatatactgcatttaccatacctagtataaacaatgagacaccagggttagatatcagtaaatgtgc  
ttccacagggttggaaggatcaccagcaatatccaaagtagcatgacaaaaatcttagagccttttagaaaacaa  
aatccagacatagttatctatcaatcacatgaacgatttgtatgtaggatctgacttagaaatagggcagcatagaac  
aaaaatagaggagctgagacaacatctgttgaggtggggacttaccacaccagacaaaaaacatcagaaagaacctc  
cattccttggatgggttatgaactccatcctgataaattggacagtagcagcctatagtgctgccagaaaaagacagc  
tggactgtcaatgacatacagaagttagtggggaaattgaataccgcaagtcagatttaccagggttaaaagtaag  
gcaattatgtaaaactccttagaggaaccaaagcactaacagaagtaataccactaacagaagaagcagagctagaac  
tggcagaaaaacagagagattctaaaagaaccagtagcatggagtgattatgacccatcaaagacttaatagcagaa  
atacagaagcaggggcaaggccaatggacatatcaaatttatcaagagccatttaaaaaatctgaaaacaggaaaata  
tgcaagaatgaggggtgccacactaatgatgtaaaacaattaacagaggcagtgcaaaaaataaccacagaaagca  
tagtaatatggggaagactcctaaatttaaaactgcccatacaaaaggaaacatgggaaacatggtggacagagtat  
tggcaagccacctggattcctgagtgagggttggtaataccctcctttagtgaatttatggtaccagtttagagaa  
agaacctatagtaggagcagaaaccttctatgtagatggggcagctaacaggagagactaaattagggaaagcaggat  
atgttactaatagaggaagacaaaaagttgtcaccctaactaacacacaaatcagaaaactcagttacaagcaatt  
tatctagctttgcaggattcgggattagaagtaaacatagtaacagactcacaaatagcattaggaatcattcaagc  
acaaccagatcaaagtgaatcagagttagtcaatcaataatagagcagtttaataaaaaaggaaaaggtctatctgg  
catgggtaccagcacacaaaggaattggaggaaatgaacaagtagataaattagttagtgctggaatcaggaaagta  
ctatttttagatggaatagataaggcccaagatgaacattagaattctgcaacaactgctgtttatccatttcagaa  
tgggtgtcgacatagcagaataggcgttactcgacagaggagagcaagaaatggagccagtagatccttagactaga  
gccctggaagcatccaggaagtcagcctaaaactcgttgtagcaaatgtctattgtaaaaagtggttcttctattgcc  
aagttgtttcatacaaaaagccttaggcattccttatcagaaggaagcggagacagcgaagagctcctcaa  
gacgtcagactcatcaagtttctctatcaaagcagtagagtagtaaatctaatccaaccttacaatatattagcaat  
agtgcattagtagtagcagcaataatagcaatagttgtgtggaccatagttatcatagaatataggaaaatattaa  
gacaaaagaaaaatagacaggttaattgataggataacagaaagagcagaagacagtggaatgaaagtgaaggggat  
caggaaagaaattatcagcacttgtggaaatggggcatcatgctccttgggatgttgatgatctgtagtgttagaaa  
atgtgtgggtcacagtttattatggggtaacctgtgtggaaagaagcaaccaccactctattttgtgcatcagatgct

aaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtacccacagaccccaaccacaagaagt  
agtattggaaaatgtgacagaaaattttaacatgtggaaaaataacatggtagaacagatgcatgaggatataatca  
gtttatgggatcaaagcctaaagccatgtgtaaaat taacccactctgtgttactttaaat tgcactgatttgagg  
aatgttactaatatcaataatagtagtgagggaaatgagaggagaaataaaaaactgctctttcaatatcaccacaag  
cataagagataaaggtgaagaaagactatgcactttttatagacttgatgtagtaccaatagataatgataacta  
gctataggttgataaattgtaatacctcaaccattacacaggcctgtccaaaggtatcctttgagccaattcccata  
cattattgtaccccgctgggttttgcgattctaaagtgtaaagacaagaagttcaatggaacaggggccatgtaaaaa  
tgtcagcacagtacaatgtacacatggaattaggccagtagtgtcaactcaactgctgttaaattggcagctctagcag  
aagaagaggtagtaattagatctagtaatttcacagacaatgcaaaaaacataatagtacagttgaaagaatctgta  
gaaattaattgtacaagacccaacaacaatacaaggaaaagtatacatataggaccaggaagagcattttatacaac  
aggagaaataataggagatataagacaagcacattgcaacattagtagaacaataatggaataacactttaaatcaaa  
tagctacaaaattaaaagaacaatttggaataataaaaacaatagtcctttaatcaatcctcaggaggggacccagaa  
attgtaatgcacagttttaattgtggaggggaatttttctactgttaattcaacacaactgtttaatagtacttgaa  
ttttaatggtacttggaatttaacacaatcgaatggtactgaaggaaatgacactatcacactcccattgtagaataa  
aacaattataaatatgtggcaggaagtaggaaaagcaatgtatgcccctcccattcagaggacaaattagatgctca  
tcaaatattacagggtcaatattaacaagagatggtggaactaacagtagtggtccgagatcttcagacctggggg  
aggagatatgagggacaatttggaagtggaattatataaatataaagttagtaaaaattgaaccattaggagtagcac  
ccaccaaggcaaaaagaagagtggtgcagagagaaaaagagcagtggaacgataggagctatgttccctggggttc  
ttgggagcagcaggaagcactatgggcgcagcgtcaataacgctgacggtagcaggccagactattattgtctggtat  
agtgaacagcagaacaatttgctgagggctattgagggcgaacagcactctgttgcaactcacagctctggggcatca  
agcagctccaggcaagagctctggctgtggaaagatacctaagggtacacagctcctagggtttgggggtgctct  
ggaaactccatctgcaccactgctgtgccttgggaatgctagtggagtaataaaaactctggatatgatttgggataa  
catgacctggatggagtggaagagaaatcgaaaattacacaggcttaatatcaccttaattgaagaatcgcaga  
accaacaagaaaagaatgaacaagacttattagcattagataagtgggcaagtttgtggaattgggttgacatatca  
aattggctgtggtatgtaaaaatcttcataatgatagtaggggcttgatagggttaagaatagtttttactgtact  
ttctatagtaaatagagttaggcagggatactcaccattgtcatttcagacccacctcccagccccgaggggacccg  
acagggccgaaggaatcgaagaagaaggtggagacagagacagagacagatccgtgcgattagtggtggatcctta  
gcacttatctgggacgatctgcggagcctgtgcctcttcagctaccaccgcttgagagacttactcttgattgtaac  
gaggatttgtggaacttctgggacgcaggggtgggaagccctcaaatattggtggaatctcctacagtagttggagtc  
aggagctaaagaatagtgctgttagcttgctcaatgccacagctatagcagtagctgaggggacagatagggttata  
gaagtagtacaaggagcttatagagctattcgccacatacctagaagaataagacagggttggaaggattttgct  
ataagatgggtggctagccccgggtgataaacggaccgcgcaatcctagggtgtgccttctagtggccagccatct  
gttgtttgccccctccccgtgccttccctgaccctggaaaggtgccactcccactgtccttccctaataaaaatgagga  
aattgcatcgcattgtctgagtaggtgtcattctattctggggggtgggggtggggcaggacagcaaggggaggatt  
gggaagacaatagcaggcatgctgggagtcggtgggtctctataaaaaacgcccggcggaacccgagcgttctga  
acgtagagtgcagaaattcagaagaactcgtcaagaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgga  
taccgtaaagcacgaggaagcgggtcagccattcgccgccaagctcttcagcaatatcacgggtagccaacgctatg  
tctgatagcggctctgccacacccagccggccacagtcgatgaatccagaaaagcggccattttccaccatgatatt  
cggcaagcaggcatcgccatgggtcacgacgagatcctcgccgtcgggcatgctgccttgagcctggcgaacagtt  
cggctggcgcgagccccctgatgctcttcgtccagatcatcctgatcgacaagaccggcttccatccgagtacgtgct  
cgctcgatgcgatgtttcgttgggtggtcgaatgggcaggtagccggatcaagcgtatgcagccgcgcgcattgcatc  
agccatgatggatactttctcggcaggagcaaggtgagatgacaggagatcctgccccggcacttcgccccaatagca  
gccagtccttcccgttccagtgcacaacgtcgagcacagctgcgcaaggaaacgcccgtcgtggccagccacgatagc  
cgcgctgcctcgtcttgcaattcaggggcacccggacaggtcggtcttgacaaaaagaaccggggcgccccctgcgc  
tgacagccggaacacggcgccatcagagcagccgattgtctgttgtgccagtcataagccgaatagcctctccaccc  
aagcggcgaggagaacctgctgcaatccatcttgttcaatcatgcgaaacgatcctcatcctgtctcttgatcagat  
cttgatccccctgcgccatcagatccttggcggcragaaagccatccagtttactttgcagggcttcccaaccttacc  
agagggcgccccagctggcaattccgggttcgcttgctgtccataaaaaccgcccagcttagctatcgccatgtaagcc  
cactgcaagctacctgcttctcttgcgcttgctgttcccttgctccagatagccagtagctgacattcatccgg  
gctcagcaccgtttctgcggactggctttctacgtgaaaaggatctaggtgaagatcctttttgataatctcatgac  
caaaatcccttaacgtgagtttctgcttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatc  
ctttttttctgcgcgtaattctgctgcttgcaacaaaaaaaccaccgctaccagcgggtgggtttgtttgcccgatcaa  
gagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagcc  
gtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggtg  
ctgccagtgggcataagtcgtgtcttaccgggttggtactcaagacgatagttaccggataaaggcgcagcgggtcgggc  
tgaaacgggggggttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagct  
atgagaaagcggccagcttcccgaaggagaaaggcggacaggtatccggttaagcggcaggggtcggaacaggagagc

Fig. 12a

gcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgt  
cgatTTTTgtgatgctcgtcaggggggcgagcctatggaaaaacgccagcaacgcggcccttttacggttcctggc  
cttttgcTggccttttgctcacatgttgTcgacaatatTggctattggccattgcatacgttgTatctatatcataa  
tatgtacatttatattggctcatgtccaatatgaccgccatgttgacattgattattgactagttattaatagtaat  
caattacgggktcattagttcatagcccatatatggagttccgcgttacataaacttacggtaaatggcccgctggc  
tgaccgccccaacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttcca  
ttgacgtcaatgggtggagtatTTacggtaaaactgcccacttggcagTacatcaagtgtatcatatgccaagtccgc  
ccctattgacgtcaatgacggtaaatggcccgctggcattatgccagTacatgaccttacgggactttcctactt  
ggcagTacatctacgtattagTcatcgctattaccatggTgatgcggTTTTggcagTacaccaatgggcgtggatag  
cggtttgactcacggggatttccaagtctccacccattgacgtcaatgggagtttgTTtkgscacccaaatcaacg  
ggactttccaaaatgtcgtaataaccccgcccgTtgacgcaaatgggcggtaggcgtgtacggTgggaggtctata  
taagcagagctcgTTtagTgaaccgtcagatcgc

Fig. 12a

**Figure 12 b. pGA2/JS7.1 functional regions**

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 106-1608                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | HIV-1 HXB2                               |
| 1401-3584                                                      | Pol, inactivating point mutations in protease and reverse transcriptase, deletion of integrase  | HIV-1 HXB2                               |
| 3671-3885 and 6210-6300                                        | Tat exons #1 and #2                                                                             | HIV-1 ADA                                |
| 3810-3885 and 6210-6884                                        | Rev exons #1 and #2                                                                             | HIV-1 ADA                                |
| 3902-4147                                                      | Vpu start site mutated (G3904C) and upstream ATG mutated (G3899C)                               | HIV-1 ADA                                |
| 4165-6626                                                      | Env                                                                                             | HIV-1 ADA                                |
| 6627-9506                                                      | Plasmid vector                                                                                  | PGA2, Gen Bank accession # AF425298      |

**Figure 12c. JS7.1 Mutation Chart**

| <b>Codon change</b>          | <b>Amino acid change</b> | <b>Region</b>            | <b>Function</b>                                 |
|------------------------------|--------------------------|--------------------------|-------------------------------------------------|
| to AGC Position<br>1279-1281 | C392S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1288-1290 | C395S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1342-1344 | C413S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC Position<br>1351-1353 | C416S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to GCC position<br>1641-1643 | D25A                     | Protease                 | Inactivation of<br>Protease active site         |
| to AAC Position<br>2418-2420 | D185N                    | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC Position<br>2661-2663 | W266T                    | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG Position<br>3297-3299 | E478Q                    | RNase H                  | Inactivation of<br>RNase H activity             |
| to ATC Position<br>3902-3904 | NA                       | Vpu                      | Vpu start site<br>mutated                       |
| to ATC Position<br>3897-3899 | NA                       | Non-coding region        | ATG upstream of<br>Vpu mutated                  |



atcgatgcaaggactcggcttgctgaggtgcacacagcaaggcgcagagcgacgactggtgagtagcccaattttgtac  
tagcggaggcttagaaggagagagatgggtgcgagagcgtcagtggttaacggggggaaaattagattcatgggagaaaatt  
aggtttaaggccagggggaaaagaaaagatatagactaaaacacctagttatgggcaagcaggagctggagagattcgact  
taaccctggcctatttagaaaacagcagaaggatgtcaacaactaatgggacagttacaaccagctctcaggacaggatcag  
aagagtttaaatcattatataatatagtagcaaccctttggtgcgtacatcaaagaatagacataaaaagacaccaggag  
gccttagataaagtagaggaaaaaataaagagcaagcaaaaggcacagcagcagctgcaacagccgcgcaagg  
aagcagagccaaaattaccctatagtgcaaaatgcacaagggcaaatggtacatcagtcctgctcacttaggactttaa  
atgcatgggtgaaggttaatagaagaaaaggcttttagccagcagaggttaatacccatgttttcagcattatcagaggagcc  
acccacaagatttaaatatgatgctaaacatagtggggggacaccaggcagcaatgcagatgttaaagataccatcaa  
tgatgaagctgcagaatgggacagagtacatccagtacatgcagggcctattccaccaggccaaatgagggaaaccaaggg  
gaagtgcatagcaggaactactagtacccttcaagaacaaataggatggatgacaagtaatccacctatccagtgagg  
gaaatctataaaaagatggatagtcctgggattaaataaaatagtaagaatgtatagccctaccagcattttggacataag  
acaagggccaaaagaacccttttagagattatgtagacaggttctttaaaactttgagagctgcaacagctacgcaggagg  
taaaaaactggatgacagaacccctgttgggtccaaaatgcgaatccagactgcaagtccttttaagagcataggacca  
ggggctacatagaagaaatgatgacatcagtcaggaggtgggaggacctggccataaagcaagggttttggctgaggc  
aatgagtcaggtacaacagaccaatgtaatgatgcagagaggcaatttttagaggccagagaataataaagagcttcaaca  
gcggcaaaagaaggacacctagccagaaaatgcaaggctcctagaagagagggcagctggaaaagcggaaagggaaggacac  
caaatgaaagactgtactgaaagacaggctaattttttagggaaaatttggccttcccacaaggggaggccaggaaattt  
tcctcagagcagaccagaaccaacagccccgcagcagagagctttggagtgggggaagagataccctcctctccgaagc  
aggagccgagggcaaggagctatatcctccttaacttccctcaaatcactctttggcaacgaccagtagtcacagtaag  
gaatagggggacagccaatagaagccctataaaccacaggagcagatgatcagattatagaagaaataagtttaccagga  
aaatggaaaacaaaaatgatagggggaattggagggttttatcaaagtaagacagtatgatcagatatctatagaaaattg  
tggaaaaaggggccataggtacagtatttagtaggacctacacctgtcaacataattggacgaaatatgttgactcagattg  
gttgacttttaattttccaattagtctctattgaaactgtgccagtaaaattaaagtcaggaatggatggcccaaagggt  
aaacaattggccattgacagaagaaaaataaaaagcattaaaagaaatttgtgcagagatggaaaaggaaggaataattc  
aaaaattgggcctgaaaacccatacaatactccaattttggccataaagaaaaagatagtaataaaggagaaaatttag  
tagatttcagagaactcaataagagaactcaagactctctggagggtccaattaggaatacctcactcctgcgggattaaaa  
aagaaaaaatcagtaacagctactagttgtgggggatgcataatttttcagttcccttagatgaagactttagaaaaatac  
tgcattcaccatacctagtttaataatgagacaccagggttagatatcagtacaatgtactcccacaggggatggaaag  
gatcaccagcaatatttcaggcaagcatgacaaaaatcttagagccctttagagcaaaaaatccagagatagtgatctac  
caatatatgaacgatttatatgtaggatctgacttagaaaatagggcagcatagagcaaaaaatagaggagttgagagaaca  
tctattgaaattggggattttaccacaccagacaaaaaacatcagaaagaacctccatttctttggatgggataatgaactcc  
atcttgacaaatggacagctcagccttatcagctgcgcagaaaaagacagctggactgtcaatgatatacaaaaaattagtg  
ggaaaaactaaataccgcaagtcagatttatgcaggaaataaagtaaaagcaatttggtgtagactcctcaggggagccaaagc  
gctaacagatgtagtaacactgactgaggaagcagaattagaattggcagagaacaggggaaattctaaaagaacctgtac  
atggagtatatattagcccaacaaaagacttagtggcagaaatacagaaaaaaggggcaagatcaatggacatatcaatt  
tatcaagagccatttaaaaatctaaagacaggaaaatatgcaaaaaagagggtcggcccaactaatgatgtaaaaacaatt  
aacagagggtagtgcaaaaaatagccatagaaaagcatagtaatatggggaaagaccctcaatttagactaccatacaaa  
gagaaaaactgggaagcatggtggatgggattttggcaggctacctggattcctgtaattgggagtttgtcaataccctcct  
ctagtaaaattatgtaccagttagagaaggaccccaataatgggagcagaaaactttctatgtagatggggcagcctaagtg  
ggagactaagctaggaaaagcagggtatgtcactgacagaggaagacaaaagggttgtttccctaattcagacaacaaatc  
aaaagactcagttacatgcaattcatctagccttgaggattcaggatcagaagtaatatagtaacagactcacagtat  
gcattaggaatcattcaggcacaaccagacaggagtgaatcagagtttagtcaatcaaatatagagaaactaatagaaaa  
ggacaaagtctacctgtcatgggtaccagcacacaaagggtatggaggaaatgaacaagtagataaaattagtcagtagtg  
gaatcagaaagggtactatttttagatggaatagataaaagcccaagatgaacatagaattctgcaacagctactgtttgt  
tccattcagaattgggtgtcaacatagcagaataggcatattccaggagaagaggcaggaatggagctggttagatcct  
agccttagagccctggaaaccacccgggaagctcagcctcaaacctgtgtgcaagtgtagtgaataaaggctgtgctggca  
ttgccaaatttgtctttctgaacaagggttaggcattcctatggcaggaagaagcggagagcgcgacgaggaactcctc  
aggaccgtcagggttcatcaaaatcctgtacaaaacagtaagtagtagtaatttagtatatgtgatgcaatctttacaaat  
agctgcaatagtaggactagtagtagcatccatagtagccatagtttgtgtggtccatagtatattatagaatatagaaaaa  
taaggaaaacagaagaaaatagacaggttacttgagagaataagagaaaagagcagaagatagtggaatgagagtgatggg  
gatacagaagaattaccactcttatggagagggggatgacaatattttggttaatgatgatttgtaatgtgaaaagt  
tgtgggtgcagctctactatgggttacctgtgtggagagcgcagagaccacctatctgtgcatcagatgctcaagca  
tatgacaagaagcacaacaatgtctgggtcagcgtgcctgcgtaccacagacacctgaccacaagaattacctttggt  
aaatgtaacagaagagtttaacatgttgaaaaataatatggtagaacagatgcatgaagatataaattagtcctatgggacc  
aaagcttaaagccatgtgtacagtaacccctctctgcgttacttttaggggtgtgctgacgctcaaaacgtcaccgcaccc  
aacaccaccatatctaatgaaatgcaaggggaaataaaaaaactgctcttcaatatgaccacagaataaagagataagaa  
gcagaaagtgtatgcacttttttatagacctgatgtaatagaaatataaaaactaagataaacaatagtaatagtagtgc  
agtatatgttaataaattgtaatacctcaacctatcacagactgtccaaaggattcctttgagccaattcccatacat  
tatttgtccccagctgggttttgcaattctaaagtgtaatgatcaggagttcagtggaaggaagacatgcaagagtgctcag  
cacagtacaattgcacacatggaattcaaqccagtatgatacaactcaactgctgtttaaattggcagcttagcagaaggaaag

tagcgattagatctgagaatatctcaacaatgccaaaactataatagtagacaattgactgagcctgtagaaattaattgt  
atcagacctggcaacaatacaagaaaaagtgtacgcataaggaccaggacaaacattctatgcaacaggtgacataatagg  
agatataagacaagcacactgttaattgttagtaaaatagcatgggaagaaactttacaaaaggtagctgcaacaattaaggga  
agcactttcagaatgccacaataaaatttactaaacactcaggaggggatttagaaattacaacaaatagtttttaattgt  
ggaggagaatttttctattgcaatacaacaagctgtttaatagcacttggataatgataactcaaacctcacagagga  
aaagagaaaggaaaacataactctccactgcagaataaagcaaattgttaatatgtggccaagagtaggacaagcaatat  
atgccccctccatcccaggaaacataaacttgtggatcaaacattactgggctactattaacaagagatggagggaataat  
ggtacaaatgatactgagaccttcaggcctggaggaggagatagagggaacaattggagaagtgaattatataaaatataa  
agtagtaaaaaattgaaccactaggtgttagcaaacccctgcataaagaagagtggtggaaagagaaaaagagcagttg  
gaattgggagctttgatctttgagttcttaggagcagcaggaagcactatgggcgcggcgtcaatggcgtcagcggtagag  
gccagacaattattgtctggtatagtgcacacgagagcaatctgctgaaggctatagaggctcaacaacatctgttgag  
actcacggctctggggcattaaacagctccaggcaagagtcctggctctggaaagatacctaaaggatcaacagctcctag  
gaatttggggctgctctggaaaactcatttgcaccactgctgtaccttggaaactctagctggagtaataaaagttataat  
gacatatgggataacatgacctggctgcaatgggataaagaaattaacaattacacatacataatataatctacttga  
aaaaatcgcagaaccagcaggaataatgaacaagacttatggcattagacaagtgggcaagtcctgtggaattgggttg  
acataacaagctgggttatggtatataagattaggtataatgatatagtaggagcgtaataaggcttaagaaataattttgt  
gtgcttactatagtgaatagagttaggcagggatactcacctttgtcattccagaccttgcccaccaccagagggaacc  
cgacaggccccgaagaatcgaagaaggaggtggcgagcaagacagagagagatccgtgcgcttagtgagcggattcttag  
cacttgccctgggaagatctgcggagcctgtgcctctcagctaccgcccgttagagagacttagtcttgattgcagcaagg  
actgtggaactcctgggacacagcagctcgaagggtgagactggggtgggaagccctcaaatatctgtggaaccttct  
atcatactggggtcaggaaactaaagaatagtgtctatatttggcttgatacaatagcaatagcagtagcttaactggcag  
atagagttataaaaaatagtaacaaagactggtagagctattcttaacatacctagaaggatcagatagggctagccccgg  
gtgataaacggacgcgcaatccctaggctgtgccttctagttgccagccatctgttgtttgccccctcccccgctgccttc  
cttgacctgggaagggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcttctgtgagtaggtgtc  
attctattctgggggggtgggggtgggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgctg  
gtgggctctataaaaaacgcccggcggaaccgagcggttctgaacgctagagtcgacaaaattcagaagaactcgtcaa  
gaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaaaagcacgaggaagcggctcagccccattcgccg  
ccaagctcttcagcaatatcacgggttagccaacgctatgctctgtagcggctcgtccacacccagcgcgctctgacaaaaag  
gaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgccgt  
cgggcatgctcgcttgagcctggcgaacagttcggctggcgagcggcctgatgctctcgctccagatcatcctgatcg  
acaagaccggcttccatccgagtagctgctcgctcgatgcatggtttcgcttgggtggtcgaatgggcaggtagccggatc  
aagcgtatgcagccgcccgttgcatcagccatgatggatactttctcgccaggagcaaggtagatgacaggagatcct  
gccccgctctcgcccaatagcagccagtccttcccgtctcagtgacaacgctcgagcacagctgcgcaaggaaacgccc  
gtcgtggccagccagcttagcgcgcgtgcctgcttgcagttcattcagggcaccggacaggtcgctctgacaaaaag  
aacggggcgccccctgcgctgacagccggaacacggcgccatcagagcagccgattgtctgttgtgccagtcatagcccga  
atagcctctccacccaagcggccggagaaactgcgtgcaatccatctgttcaatcatgcgaaacgatcctcatcctgtc  
tcttgatcagatcttgatccccctgcgccatcagatccttggcggaagaaagccatccagtttactttgcagggttccc  
aacettaccagagggcgccccagctggcaattccgggtcgcttgcgtgctccataaaacgcccagcttagctatcgccatg  
taagcccatgcaagctacctgctttctctttgcgttgcggtttcccttgcagatagcccagtagctgacattcatc  
cggggtcagcacccgtttctgcggactggctttctacgtgaaaaggatctagggtgaagatcctttttgataatctcatgac  
caaatcccttaacgtgagtttctgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctt  
ttttctgcgctaatctgctgcttgcaacaaaaaaaccacggctaccagcgggtggtttgtttgcccggatcaagagcta  
ccaactcttttccgaaggtaactggcttcagcagagcgcagataccaataactgttcttctagtgtagccgtagttagg  
ccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaattcctgttaccagtggtgctgcccagtggtg  
ataagtgcgtgcttaccgggttggaactcaagacgtagttaccggataaggcgcagcggctcgggctgaacgggggggtcgt  
tgcacacagcccagcttgagcgcgaacgacctacaccgactgagatacctacagcgtgagctatgagaaagcgccacgct  
tcccgaaggggagaaaggcggacaggtatccggtaagcggcagggctcggaacaggagagcgcacgaggggagcttccagggg  
gaaacgcctggtatctttagtctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgtagctcgtcaggg  
gggcgagcctatggaaaaacgcccagcaacgcggcccttttacggttccctggccttttgctggccttttgctcacatgtt  
gtcgacaatatgggtattggccattgcatacgttgatctatatcataatatgtacatttatattggctcatgtccaat  
tgaccgcagctgttgacattgatttagctagtttatgaatagtaatcaattacgggttcatttagttacgcccata  
tggagtcccggttacataacttacggtaaatgcccctggcagcccaacgacccccgcccattgacggtcaata  
atgacgtatgttcccatagtaacgcaatagggaactttccattgacgtcaatgggtggagtagtttacgggtaaactgccc  
cttggcagtagatcaagtgtatcatatgccaagtcgccccctattgacgtcaatgacggtaaatggcccgcctggcatt  
atgcccagtagatgaccttacgggactttcctacttggcagtagatctacggtattagtcacggctattaccatgggtga  
tgcggttttggcagtagaccaatggcggtgagcgggttgactcacggggatttccaagctcaccacccattgacgtc  
aatggggagtttggtttggcaccaaaaacacgggactttccaaaatgtcgtaataacccccgcccgttgacgcaaatggg  
cggtagggcgtgacgggtgggaggtctatataagcagagctcgttttagtgaaccgtagatcgctggagacggccatccac  
cggttttgacctcatagaaagacacgggacggatccagcctccgcccgggaacgggtgcattggaacgggtattccc  
cgtgccaagagtgacgtaagtaccgctatagactctataggcacacccctttggctcttatgcatgctatactgtttt  
ggcttggggcctatacacccccgcttccctatgctataggtgatggtagcttagcctataggtgtgggttattgacca  
ttattgaccactccccatttgggtgacgatacttccattactaattccataacatggctctttgcccacactatctctatt  
ggctatatgccaatactctgtccttcagagactgacacggactctgtatttttacaggatgggggtcccatattattatta

Fig. 13a

caaattcacatatacaacaacgccgtcccccgtagccgcagttttattaaacatagcgtgggatctccacgcgaatctc  
gggtaccgtgttccggacatgggytcttctccggtagcggcgagcttccacatccgagccctggccccatgcctccagc  
gggtcatggtcgctcggcagctccttgctcctaacagtggaggccagaettaggcacagcacaatgccaccaccaccag  
tgtgccgcacaaggccgtggcggtagggtagtggtctgaaaatgagctcggagattgggctcgcaccgctgacgcagatg  
gaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgattctgataagagtcagaggtaactcccggt  
gcggtgctgttaacggtggagggcagtgtagtctgagcagtagctggtgctgcccgcgcgcgccaccagacataatagctg  
acagactaacagactgttcctttccatgggtcttttctgcagtcacc

Fig. 13a

**Figure 13b. pGA1/IC25 Functional Regions**

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 104-1591                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | AG HIV-1 isolate 928                     |
| 1393-3576                                                      | Pol, inactivating point mutations in protease and reverse transcriptase, deletion of integrase  | AG HIV-1 isolate 928                     |
| 3663-3877 and 6215-6305                                        | Tat exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3802-3877 and 6215-6510                                        | Rev exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3904-4149                                                      | Vpu                                                                                             | AG HIV-1 isolate 928                     |
| 4067-6628                                                      | Env                                                                                             | AG HIV-1 isolate 928                     |
| 6629-10447                                                     | Plasmid vector                                                                                  | PGA1, Gen Bank accession # AF425297      |

Figure 13c. IC25 Mutation Chart

| Codon change                 | Amino acid change | Region                   | Function                                        |
|------------------------------|-------------------|--------------------------|-------------------------------------------------|
| to AGC position<br>1271-1273 | C390S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1280-1282 | C393S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1334-1336 | C411S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1343-1345 | C414S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AAC position<br>1633-1635 | R25N              | Protease                 | Inactivation of<br>Protease active site         |
| to AAC position<br>2410-2412 | D185N             | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC position<br>2653-2655 | W266T             | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG position<br>3289-3291 | E478Q             | Rnase H                  | Inactivation of<br>RNase H activity             |

Figure 14a. pGA1/IC2 Sequence

atcgatgcaaggactcggcttgctgaggtgcacacagcaaggaggagagagagactgggtgagtagcgaatttttgac  
tagcggaggctagaaggagagagatgggtgcgagagcgtcagtggttaacgggggaaaattagattcatgggagaaaatt  
aggttaaggccaggggaaaagaaaagatatagactaaaacacctagtagggcaagcagggagctggagagattcgact  
taaccctggcctattagaaacagcagaaggatgtcaacaactaatgggacagttacaaccagctctcaggacaggatcag  
aagagtttaaatcattatataatatagtagcaaccctttgggtgcgtacatcaaagaatagacataaaaagacaccaggag  
gccttagataaagtagaggaaaaaataaagagcaagcaaaaggcacagcaggcagcagctgcaacagccggccacagg  
aagcagcagccaaaattaccctatagtgcaaaatgcacaagggcaaatggtacatcagtcctatgctacctaggactttaa  
atgcatgggtgaaggtaatagaagaaaaggcttttagccagaggtaatacccatgttttcagcattatcagagggagcc  
acccacaaagatttaaatatgatgctaaacatagtggggggacaccaggcagcaatgcagatgttaaaagataccatcaa  
tgatgaagctgcagaatgggacagagtacatccagtacatgcagggcctattccaccaggccaaatgagggaaaccaaggg  
gaagtgcatagcaggaactactagtagccttcaagaacaaataggatggatgacaagtaatccacctatcccagtgagg  
gaaatctataaaagatggatagtcctgggattaaataaaatagtaagaatgtatagccctaccagcattttggacataag  
acaagggccaaaagaaccctttagagattatgtagacaggttctttaaaactttgagagctgaacaagctacgcagagg  
taaaaactggatgacagaaaaccttgggtccaaaatgccaatccagactgcaagtcattttaagagcatttaggacca  
ggggctacattagaagaaatgatgacatcatgtcaggagtgaggagcctggccataaaagcaagggttttggtgagggc  
aatgagtagcagtaacagaccaatgtaatgatgcagagaggcaattttagaggccagagaataataaagagcttcaaca  
gcggcaaaagaaggacacctagccagaaatgcaaggctcctagaaaagagaggcagctggaaaagcggaagggaagcac  
caaatgaaagactgtactgaaagacaggctaatttttagggaaaaattggccttcccacaaggggagggcaggaaattt  
tctcagagcagacagaaaccaacagcccgccagcagagagctttggagtgggggaagagataccctcctctccgagc  
aggagccgaggagacaggactatcctccttaacttccctcaaatcactctttggcaacgaccagtagtcacagttaa  
gaatagggggcagccaatagaagccctattagacacaggagcagatgatacagtagtagaagaaataagtttaccagga  
aaatggaaacccaaaatgatagggggaattggaggttttatcaaagtaagacagtagatcagatctctatagaaatttg  
tggaagaaaggccataggtacagtagtagtaggacctacacctgtcaacataattggacgaaatattgtgactcagattg  
gtgtacttttaattttccaattagtcctattgaaactgtgccagtaaaatataagtcaggaatggatggcccaaagggt  
aaacaattggccattgacagagaagaaaaataaaagcattaaaagaaatttggtgcagagatggaaaagggaagaaatttc  
aaaaattgggctgaaacccatacaatactccaattttgcccataaagaaaaagatagtagtaaaatggagaaaaattag  
tagatttcagagaactcaataagagaactcaagacttctgggaggtccaattaggaatacctcatcctgcgggattaaaa  
aagaaaaaatcagtaacagtagtagtagtgaggatgcatatttttcagttcccttagatgaagactttagaaaaatatac  
tgcattcaccatacctagtttaataatgagacaccagggttagatatacagtagaactgtagtcccacagggatggaaag  
gatcaccagcaatatttcaggcaagcatgacaaaaatccttagagccctttagagcaaaaaatccagagatagtgatctac  
caatatagaacgatttatgttagatctgacttagaaaatagggcagcatagagcaaaaaatagaggagtgagagaaca  
tctattgaaatgggctgaaacccatacaatactccaattttgcccataaagaaaaagatagtagtaaaatggagaaaaatc  
atcctgacaaatggacagtcagcctatacagctgccagaaaaagacagctggactgtcaatgatatacaaaaaatagtg  
ggaaaactaaataccgcaagtcagatttatgcaggaataaagtaaaagcaattgtgtagactcctcaggggagccaaagc  
gctaacagatgtagtaacactgactgaggaagcagaattagaattggcagagaaacagggaaattctaaagaacctgtac  
atggagtagattatgacccaacaaaagacttagtggcagaaatacagaaacaagggcaagatcaatggacatacaatt  
tatcaagagccatttaaaaatctaaagacaggaataatgcaaaaaagaggctcgccacactaatgtatgtaaaacaatt  
aacagaggtagtcagaaaaatagccatagaaagcatagtaatatgggggaaagacccctaaatttagactaccatacaaaa  
gagaaacatgggaagcatgggtggatggagtagtggcaggctacctggattcctgaatgggaggtttgtcaataccctcct  
ctagtaaaattatggtagcagtagagaaggacccataatgggagcagaaactttctatgtagatggggcagctaatag  
ggagactaagctaggaagagcagggtatgtcactgacagaggaagacaaaagggtgtttccctaattcagacaacaaatc  
aaaagactcagttacatgcaattcatctagccttgaggattcaggatcagaagtaaatatagtaacagactcacagtag  
gcattaggaatcattcaggcacaaaccagacaggagtgaaatcagagtagtcaatcaaataatagagaaactaatagaaaa  
ggacaaaagtcacctgtcattgggtaccagcacacaaaagggtatggaggaaatgaacaagtagataaaatagtcagtagtg  
gaatcagaaaggtagtatttttagatggaatagataaaagcccaagatgaacattagaattctgcaacagctactgtttgt  
tcatttcagaattgggtgtcaacatagcagaataggcattattccaggggagaagaggcaggaatggagctggtagatcct  
agcctagagccctggaccacccgggaagtcagcctacaactgctgttagcaagtgtagtgaataatgtgtagtaattttaca  
ttgccaatgtgctttctgaacaagggttaggcattcctatggcaggaagaagcggagacgcccagcaggaactcctc  
aggaccgtcaggttcacataaactcctgtaccaaaacagtaagtagtaatttagtatatgtgtagcaacttttaca  
agctgcaattagtaggactagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag  
taaggaaaacagaagaaaatagacaggttacttgagagaataaagagaagagcagaagatagtggaatgagagtgatggg  
gatacagaagaattatccactcttatggagagggggtatgacaatatttggttaatgatgatttgtaatgctgaaaagt  
tgtgggtcacagctctactatggggtacctgtgtggagagacgcagagaccaccctattctgtgcatcagatgctaaagca  
tatgacaaaagaagcaccaatgtctgggctacgcatgcctgcgtacccacagacctgaccacaagaattacctttgggt  
aaatgtaacagaagagtttaacatgtggaaaaataataggtagaacagatgcatgaagatataattagtcctatgggacc  
aaagcttaaagccatgtgtacagctaacccctctctgcgttactttagggtgtgtgtagcgtcaaaacgtcaccgacacc  
aacaccacatatctaatgaaatgcaagggaataaaaaactgcctcttcaatatgaccacagaattaaagagataagaa  
gcagaaagtgtatgcacttttttatagacctgatgtaatagaatataaaaaactaagattaacaatagtaatagtagtc  
agtatatgttaataaattgtaatacctcaaccattacacagactgtccaaagggtatcctttgagccaattcccatacat  
tattgtgccccagctgggtttgcaattctaaagtgtaatgtacggagttcagtggaagggacatgcaagagtgtag  
cacagtagaatgcacacatggaatcaagccagtagtatcaactcaactgctgttaaatggcagtagtagcagaaggaaaga

tagcgattagatctgagaatatctcaacaatgccaaaactataatagtacaattgactgagcctgtagaatttaattgt  
atcacacctggcaacaatacagaagaaagtgtacgcataaggaccaggacaaacattctatgcaacaggtgacataatagg  
agatataagacaagcacactgtaattgttagtaaaatagcatgggaagaaactttacaaaaggtagctgcacaattaagga  
agcactttcagaatgccacaataaaattactaaacactcaggaggggatttagaaattacaacaaatagttttaattgt  
ggaggagaatttttctattgcaatacaacaaagctgtttaatagcacttggaaataatgataactcaaacctcacagagga  
aaagagaaaggaaaacataactctccactgcagaataaagcaaatgttaaatatgtggccaagagtaggacaagcaatat  
atgccccctcccatcccaggaaacataacttgtggatcaaacattactgggctactattaacaagagatggagggaataat  
ggtacaaatgatactgagaccttcaggcctggaggaggagatatgagggacaattggagaagtgaattatataatataa  
agtagtaaaaaattgaaccactagggtgtagcaccaaccctgcaaaaagaagagtgggtggaaagagaaaaagagcagttg  
gaatgggagctttgatctttgagttcttaggagcagcaggaagcactatggcgcgcgctcaatggcgctgacggtacag  
gccagacaattattgtctggtatagtgaacagcagagcaatctgctgaaggctatagaggctcaacaacatctgttgag  
actcacggtctggggcattaaacagctccaggcaagagctcctggctctggaaagatacctaaaggatcaacagctcctag  
gaatttggggctgctctggaaaactcatttgcaccactgctgtaccttggaaactctagctggagtaataaaagtataat  
gacatatgggataacatgacctggctgcaatgggataaagaaattacaattacacatacataatataatctacttga  
aaaaatcgagaaccagcaggaaattaatgaacaagacttatggcattagacaagtgggcaagctctgtggaatttgggttg  
acataacaagctggctatgggtatataagattagggtataatgatagtaggagcgtaataggcttaagaataattttgtct  
gtgcttactatagtgaatagagtttaggcagggatactcacctttgtcattccagacccttggccaccaccagagggaaacc  
cgacaggcccgaaagaatcgaagaaggaggtggcgagcaagacagagagagatccgtgctcttagtgagcggatctcttag  
cacttgctgggaagatctgcggagcctgtgctctctcagctaccgcgatggagagacttagtcttgattgcagcaagg  
actgtggaactcctgggacacagcagctctcaaggagctgagactggggtgggaagccctcaaatactgttggaacctctt  
atcatactggggtcaggaaactaaagaatagtgtctattaatttgccttgatacaatagcaatagcagtagcactggacag  
atagagtataaaaaatagtacaaagaactggtagagctattcttaacatacctagaaggatcagatagggttagccccgg  
gtgataaaacggaccgcgcaatccctaggctgtgctctctagtggccagccatctgttgtttgccccctccccgtgctctc  
cttgaccttggaaggtgccactcccactgtccttcttaataaaatagggaattgcatcgcatgtctgagtaggtgtc  
attctattctggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgctg  
gtgggctctatataaaaaacgcccggcggaaccgagcgttctgaacgctagagtcgacaaaattcagaagaactcgtcaa  
gaaggcgatagaaggcgatgctgctgcaatcgggagcggcgataccgtaaaagcagcaggaagcggtcagcccattcgcgg  
ccaagctcttcagcaatatcacgggtagccaacgctatgtcctgtagcggctgcccacaccagccggccacagtcgat  
gaaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgccgt  
cgggcatgctcgcttgagcctggcgaaacagttcggctggcgagcggcctgtatgctcttcgtccagatcatcctgatcg  
acaagaccggttccatccgagtagctgctcgctcgatgctgatttgcgttggtggtcgaatgggaggttagccggatc  
aagcgtatgcagccgcccagttgcatcagccatgatggatactttctcggcaggagcaaggtgagatgacaggagatcct  
gccccggcacttcgccaatagcagccagtccttcccgtctcagtgacaacgctcgagcacagctgcgaaggaacgccc  
gtcgtggccagccagcttgagcgcagcgtgctctcttcagttcattcagggcaccggacaggtcgggtcttgacaaaaag  
aacggggcgcccctgctgacagccgggaacacggcgccatcagagcagccgattgtctgttgtgcccagtcatagcccga  
atagcctctccacccaagcggcgaggagaacctgctgcaatccatctgttcaatcatgcgaaacgatcctcatcctgtc  
tcttgatcagatcttgatccccctgcgccatcagatccttggcggaagaaagccatccagtttactttgcagggttccc  
aaccttaccagagggcgccccagctggcaattccgggtcgcttgctgtccataaaacgcccagctctagctatcgccatg  
taagcccactgcaagctacctgttctctttgcgttgctgttccccttgccagatagccagtagctgacattcctc  
cggggtcagcaccgttctgcggactggcttctcagtgaaaaggatctagggtgaagatccttttgataatctcatgac  
caaaatcccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaagatcaaaggatctcttgagatcctt  
ttttctgcgcgtaatctgctgcttgcaacaaaaaaaccacgctaccagcgggtggtttgtttgcccggatcaagagcta  
ccaaactcttttccgaaggtaactggcttcagcagagcgagataccaaatactgttcttctagtgtagccgtagttagg  
ccaccacttcaagaactctgtagaccgcctacatacctcgctctgctaactctgttaccagtggtgctgcccagtgggc  
ataagtctgtcttaccgggttgagctcaagacgatagttaccggataaggcgagcggctcgggtgaaacgggggttcg  
tgcaacagccagcttggaagcagacctaaccgaactgagatacctacagcgtgagctatgagaaagcggccacgct  
tcccgaaggagaaaggcgagcaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggg  
gaaacgcctggtatctttatagtcctgtcgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcaggg  
ggcgagcctatggaaaacgcccagcaacgcggcccttttacgggttccctggccttttgcgtggccttttgcctcacatgtt  
gtcgacaatatggctattggccattgcatacgttgtatctatatacataatgtacatttatattggctcatgtccaat  
atgaccgccaatgttgacattgatttagactagttatgaatgaataatcaggggttcattagttcatagcccata  
tggagtccgcgttacataacttacggtaaatggccccctggtgaccgcccacgacccccgcccattgacgtcaata  
atgacgtatgttcccataagtaacgccaatagggaacttccattgacgtcaatgggtggagattttacggtaaaactgccc  
cttggcagtagacatcaagtgtatcatatgccaaagtcgccccctattgacgtcaatgacggtaaatggccgcctggcatt  
atgcccagtagacatgaccttacgggacttctacttggcagtagacatctacgggtatagtcacggtattacatgggtga  
tgcggttttggcagtagaccaatggcggtggatagcgggttgactcacggggatttccaagcttccacccattgacgtc  
aatgggagttgttttgccacaaaatcaacgggacttccaaaatgtcgtataataaccccgcccgttgacgcaaatggg  
cggtaggcgtgtacggtgggaggtctatataagcagagctcgttttagtgaaccgtcagatcgctggagacggccatccac  
gctgtttgacctccatagaagacacgggacgactcagcctcccgcggggaacgggtgcatggaaacgggattccc  
cgtgccaagagtgacgtaagtaccgctatagactctataggcacacccctttggctcttatgcatgctatactgtttt  
ggcttggggcctatacacccccgcttcttatgctataggtgatgggtatagcttagcctataggtggtgggttatgacca  
ttattgaccactcccctattgggtgacgatacttccattactaatccataacatggctctttgcccacactatctctatt  
ggctatatgccaatactctgtccttcagagactgacacggactctgtatttttacaggatgggggtcccatttattattta

Fig. 14a

caaattcacatatacaacaacgccgtcccccgtagcccgagttttattaaacatagcgtgggatctccacgcgaatctc  
gggtaccgtgttccggacatgggytcttctccggtagcggcgagcttccacatccgagccctgggtcccatgcctccagc  
ggctcatggtcgctcggcagctccttgctcctaacagtggaggccagacttaggcacagcacaatgcccaccaccaccag  
tgtgccgcacaaggccgtggcggtagggtaggtgtgtctgaaaatgagctcggagattgggctcgccaccgctgacgcagatg  
gaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgattctgataagagtcagaggtaactcccgtt  
gcgggtgctgttaacggtggagggcagtgtagtctgagcagtactcgttgetgccgcgcgcgccaccagacataatagctg  
acagactaacagactgttcctttccatgggtcttttctgcagtcacc

Fig. 13a



Figure 14b. pGA1/IC2 Functional Regions

| <b>Position within sequence<br/>(starting and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 104-1591                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | AG HIV-1 isolate 928                     |
| 1393-3576                                                      | Pol, inactivating point mutations in reverse transcriptase, deletion of integrase               | AG HIV-1 isolate 928                     |
| 3663-3877 and 6215-6305                                        | Tat exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3802-3877 and 6215-6510                                        | Rev exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3904-4149                                                      | Vpu                                                                                             | AG HIV-1 isolate 928                     |
| 4067-6628                                                      | Env                                                                                             | AG HIV-1 isolate 928                     |
| 6629-10447                                                     | Plasmid vector                                                                                  | PGA1, Gen Bank accession # AF425297      |

Figure 14c. IC2 Mutation Chart

| Codon change                 | Amino acid change | Region                   | Function                                        |
|------------------------------|-------------------|--------------------------|-------------------------------------------------|
| to AGC position<br>1271-1273 | C390S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1280-1282 | C393S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1334-1336 | C411S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1343-1345 | C414S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AAC position<br>2410-2412 | D185N             | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC position<br>2653-2655 | W266T             | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG position<br>3289-3291 | E478Q             | RNase H                  | Inactivation of<br>RNase H activity             |

Figure 15a. pGA1/IC48 sequence

atcgatgcaaggactcggcttgctgaggtgcacacagcaagaggcgagagcgacgactggtagtacgccaatttttgac  
tagcggaggctagaaggagagagatgggtgcgagagcgtcagtggttaacgggggaaaatttagattcatgggagaaaatt  
agggttaaggccagggggaaaagaaagatatagactaaaaacactagtagtggaagcagggagctggagagattcgcact  
taacctggcctattagaacacgacagaggatgtcaacaactaatgggacagttacaaccagctctcaggacaggatcag  
aagagtttaaatcattatataatatagtagcaaccctttgggtgcgtacatcaaagaatagacataaaagaccccaggag  
gccttagataaaagtagaggaaaaaataaagagcaagcaaaaggcacagcagcagcagctgcaacagccgacagg  
aagcagcagccaaaattaccctatagtgcaaaatgcacaagggcaaatggtacatcagtcctatgtcacctaggactttaa  
atgcatgggtgaaggtaataagaagaaaaggcttttagccagaggtataacccatgttttcagcattatcagagggagcc  
accccaacaagatttaaatatgatgctaaacatagtggggggacaccaggcagcaatgcagatgttaaaagataccatcaa  
tgatgaagctgcagaatgggacagagtacatccagtacatgcagggcctattccaccaggccaaatgagggaaaccaagg  
gaagtgcacatagcaggaactactagtagccttcaagaacaaataggatggatgacaagtaatccacctatcccagtgga  
gaaatctataaaagatggatagtcctgggattaaataaaatagtaagaatgtatagccctaccagcatttttggacataag  
acaagggccaaaagaaccccttagagattatgtagacaggttctttaaactttgagagctgaacaagctacgcaggagg  
taaaaaactggatgacagaaccccttggtgggtccaaaatgcgaatccagactgcaagtccattttaagagcattaggacca  
ggggctacattagaagaaatgatgacatcatgtcagggagtgaggagcctggccataaagcaagggttttggctgaggc  
aatgagtagcagtaacagaccaatgtaatgatgcagagaggcaatttagaggccagagaataataaagagcttcaaca  
gcggcaaaagagacacccctagccagaatgcgaaggctcctagaagagagggcagctggaaaagcggaaaggaaggacac  
caaatgaaagactgtactgaaagacaggctaattttttagggaaaaatttggccttcccacaaggggaggccgaaattt  
tctcagagcagaccagaaccaacagccccgcagcagagagctttggagtgggggaagagataccctcctcctccgaagc  
aggagccgaggggacaagggactatatcctcccttaacttccctcaaatcactctttggcaacgaccagtagtcacagtaa  
gaatagggggacagccaatagaagccctattagacacaggagcagatgatagctattagaagaaataagtttaccagga  
aaatggaaacaaaaatgataggtggaattggagggttttatcaaagtaagacagtagatgatcagatatctatagaatttg  
tgaaaaaaggccatagggtacagtagtaggacctaactgtcacaataattggacgaaatattgtgactcagattg  
gttgtacttttaattttccaattagtcctatttgaacctgtgagcagtaaaattaaagttaggaatggatggcccaagggt  
aaacaatggccattgacagaagaaaaataaaagcattaaaagaaatttgtgcagagatggaaaaggaagaaatttc  
aaaaattgggcctgaaaacccatacaatactccaatatttgccataaagaaaaagatagtagtaaatggagaaaattag  
tagatttcagagaactcaataagagaactcaagacttctgggagggtccaattaggaaatccctcatcctgagggttaaaa  
aagaaaaaatcagtaacagtagtagatgtgggggatgcatttttcagttcccttagatgaagacttttagaaaataac  
tgacttaccataccttagtttaataatagacacacagggttagatatcagtagcaatgtactcccacagggttggaag  
gatcaccagcaatatttcaggcaagcatgacaaaaatcttagagccctttagagcaaaaaatccagagatagtgactac  
caatatatgaacgatttatatgtaggactgtgacttagaaatagggcagcatagagcaaaaaatagaggagttgagagaaca  
tctattgaaatggggatttaccacaccagacaaaaacatcagaagaacccctccatttctttggatgggatatgaactcc  
atcctgacaaatggacagctccagcctatacagctgccagaaaaagacagctggactgtcaatgatatacaaaaaatagtg  
ggaaaactaaatccgcaggtcagatttatgcaggaattaaagtaaaagcaatttgtgtagactcctcaggggagccaaagc  
gctaacagatgtagtaacactgactgaggaagcagaattagaattggcagagaaacagggaaattctaaaagaacctgtac  
atggagtagatattatgacccaacaaaagacttagtggcagaaatcacagaacaaagggcaagatcaatggacatatcaaat  
tatcaagagccatttaaaaaatctaagacagggaaaatagcaaaaaagagggtcgccccacactaatgatgtaaaaaatt  
aacagaggttagtcagaaaaatagccatagaaagcatagtaatatggggaaagacccctaaatttagactaccatacaaaa  
gagaaacatgggaagcatgggtggatggagtagtggcagggtcactggattcctgaatgggagtttgtcaataccctcct  
ctagtaaaattatggtaccagtagagaaggacccataatgggagcagaaacttctatgtagatggggcagctaatag  
ggagactaagctaggaaaaagcagggtatgtcactgacagaggaagacaaaagggtgtttccctaattcagacaacaaatc  
aaaagactcagttacatgcaattcatctagccttgaggattcaggatcagaagtaaatatagtaaacagactcacagat  
gcattaggaatcattcaggcacacaccagacaggagtgatcagagtttagtcaataataatagagaaactaatagaaaa  
ggacaaagtctacctgtcatgggtaccagcacacaaagggttgaggagaaatgaacaagtagataaatttagtcagtagtg  
gaatcagaaggtactatttttagatggaatagataaaagcccaagatgaacattagaattctgcaacagctactgtttgt  
tcatttcagaattgggtgtcaacatagcagaataggcattattccaggggagaagaggcaggaatggagctggtagatcct  
agcctagagccctggaaaccccggaagtcagcctacaactgctgttagcaagtgttactgtaaaaaatgctgctggca  
ttgccaatgtgctttctgaacaagggttaggcctccttatggcaggaagaagcggagacgcccagcagggaactcctc  
aggaccgtcagggttcatcaaaatcctgtacaaaacagtaagtagtagtaatttagtatatgtgatgcaatctttacaaat  
agctgcaatagtaggacttagtagtagcatccatagtagccatagttgtgtgggtccatagtagttatagaaatagaaaaa  
taaggaaacagaagaaatagacaggttacttgagagaataagagaaagagcagaagatagtggaatgagagtgatggg  
gatacagaagaattatccactcttatggagagggggtatgacaatattttggttaattgatgatttgtaattgctgaaaagt  
tgtgggtcacagctactactatgggttacctgtgtggagagcgcagagaccacccctattctgtgcactcagatgctaaagca  
tatgacaaaagcagcaaatgtctgggtacgcagctgtcggtacccacagaccctgaccacaagaattacctttgggt  
aaatgtaacagaagagtttaacatgtggaaaaataataggtagaacagatgcatgaagataaatttagtctatgggacc  
aaagcttaagccatgtgtacagctaaccctctctgctgtacttttaggggtgtgctgacgctcaaaacgtcaccgacacc  
aacaccacatattctaataagaaatgcaaggggaaataaaaaactgctcttcaatatgaccacagaattaagagataagaa  
gcagaaagtgtatgcactttttatagacctgatgtaataagaaatataaaaactaagatttaacaatagtaataagtagtc  
agtagtagtataataaattgtaataacctcaaccattacacagactgtccaaagggtatcctttgagccaattcccatacat  
tattgtgccccagctgggttttgcaattctaaagtgaatgatacggagttcagtggaaggggacatgcaagagtggtcag  
cacagtaaatgcacacatggaatcaagccagtagtatcaactcaactgctgtttaaattggcagcttagcagaaggaaaga

tagcgattagatctgagaatatctcaaacatgcccactataatagtacaattgactgagcctgtagaattaattgt  
atcagacctggcaacaatacaagaaaaagtgtacgcataaggaccaggacaaacattctatgcaacaggtgacataatagg  
agatataagacaagcacactgtaattgttagtaaaatagcatgggaagaaactttacaaaaggtagctgcacataaagga  
agcactttcagaatgcccaataaaatttactaaacactcaggaggggatttagaaattacaacaaatagttttaattgt  
ggaggagaattttctattgcaatacaacaaagctgtttaatagcacttggataatgataactcaaactcacagagga  
aaagagaaaggaaaacataactctccactgcagaataaaagcaaattgtaaatatgtggccaagagtaggacaagcaatat  
atgcccctcccatcccaggaaacataacttgtggatcaaacattactgggctactattaacaagagatggagggaataat  
ggtacaaatgatactgagaccttcaggcctggaggaggagataatgagggacaattggagaagtgaattatataataataa  
agtagtaaaaaattgaaccactaggtgtgacaccaacccctgcaaaaagaagagtgggtggaaagagaaaaaagagcagttg  
gaatgggagctttgactcttgagttcttaggagcagcaggaagcactatgggcgcggtcaatggcgctgacggtacag  
gcagacaattattgtctggtatagtgcaacagcagagcaatctgctgaaggctatagaggctcaacaacatctgttgag  
actcacggtctggggcattaaacagctccaggcaagagtcctggctctggaaagatacctaaggatcaacagctcctag  
gaatttggggctgctctggaaaactcatttgcaccactgctgtaccttggaaactctagctggagtaataaaagtataat  
gacatatgggataacatgacctggctgcaatgggataaagaatttaacaattacacatacataatatataatctacttga  
aaaatcgcagaaccagcaggaaattaatgaacaagacttattggcattagacaagtgggcaagtctgtggaattgggttg  
acataacaagctggctatggatatagaattaggtataatgatagtaggagcgtaataaggcttaagaataatttttgc  
gtgcttactatagtgaatagagttaggcagggaactcaccttgtcattccagacccttggccaccaccagaggggaacc  
cgacaggcccgaaagaatcgaagaaggaggtggcgagcaagacagagagagatccgtgcgcttagtgagcggattcttag  
cacttgcctgggaagatctgcggagcctgtgctcttcagctaccgagcttgagagacttagtcttgattgagcagcaag  
actgtggaactcctgggacacagcagcttcaagggactgagactggggtgggaagccctcaaatactgtggaacctct  
atcatactggggctcaggaaactaaagaatagtgtctattaatttgccttgatacaatagcaatagcagtagctaactggacag  
atagagttataaaaaatagtaacaaagaactggtagagctattcttaacatacctagaaggatcagatagggctagccccg  
gtgataaacggacggcgcaatccctaggctgtgcttctagttgccagccatctgttgtttgcccctccccctgctctc  
cttgacctgggaaggtggcactccccactgtcctttcctaataaaatgaggaattgcatcgcatgtctgtagtaggtgtc  
attctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatctggggatg  
gtgggctctataaaaaaacgcccggcggaaccgagcgttctgaacgctagagtcgacaaattcagaagaactcgtcaa  
gaaggcgatagaaggcgatgctgctgcgaatcgggagcggcgataccgtaaaagcacgaggaagcggctcagccattcgc  
ccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcggcttgcacacccagccggccacagtcgat  
gaatccagaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgcctg  
cgggcatgctgccttgagcctggcgaacagttcggctggcgagccctgatgctcttcgtccagatcatcctgatcg  
acaagaccggttccatccgagtagctgtcgtcgtatgctggttctcgcttgggtggatgggaggtgagtagcaggagatc  
aagcgtatgcagccgcccgttgcacagccatgatggatacttctcggcaggagcaaggtgagtagcaggagatcct  
ggcccgccacttcgcccataagcagccagtccttcccgttccagtgcacacgtcgagcacagctgcgaaggaacgccc  
gtcgtggccagccagatagccgctgctcgtcttcagttcattcagggcaccggacaggtcggtcttgacaaaaag  
aacggggcgcccctgcgctgacagccgggaacacggcgccatcagagcagccgatgtctgttggccagtcagagccga  
atagcctctccaccgaagcggcggaacctgcgtgcgaatccatctgttcaatcatgcgaacgatcctcatcctgtc  
cttgatcagatcttgatccccctgcgcatcagatccttggcggaagaaagccatccagtttactttgacagcttccc  
aaccttaccagagggcgcccagctggcaattccggttcgcttgcgtgctccataaaacggccagcttagctatcgccatg  
:aagcccactgcaagctacctgcttctctttgcgcttgcggtttcccttgcagatagccagtagctgacattcatc  
:ggggtcagcaccgtttctgcggaactggcttctacgtgaaaaggatctaggtgaagatccttttgataatctcatgac  
:aaaatcccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctt  
:tttctgcgcttaactgctgcttgcaaaacaaaaaacccagcgtaccagcgggtgggttgggttgcgggataagagcta  
:caactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagcgtagttagg  
:caccacttcaagaactctgtagcaccgctacatacctcgctctgctaactctgttaccagtgggtgctgcccagtgggc  
ataagtcgtgtcttaccgggttggactcaagacgatatgttaccggataaaggcgcagcggctcgggtgaacggggggttcg  
:gcacacagcccagcttggagcgaacgacctacacgaactgagatacctacagcgtgagctatgagaaagcggccacgct  
:cccgaagggaagaaaggcggacaggtatccggtgaagcgaagggtcggaacaggagagcgcacgagggagcttccagggg  
:aaacgcctgggtatctttatagtcgtgcttgcgacactctgacttgagcgtcgatttttgtagtgcctcgtcaggg  
:ggcgagcctatggaaaaacggcagcaacggcgccctttacgggttctcggccttttgctggccttttgctcacatggt  
:tcgacaatatggctattggccattgcatacgttgcatactatataatataatgtacatttatattggctcatgtccaat  
:tgaccgcatgttgacattgattattgactagttatataatagtaataacacgggttcatagttcatagcccatata  
:ggagttccgcttacataacttacggtaaatggcccgctgggtgacggcccaacgacccccgcccattgacgtcaata  
:tgacgtatgttcccatagtaacgccaatagggaacttccattgacgtcaatgggtggagttttacggtaaaactgccc  
:ttggcagtagcatcaagtgtagcatatgccaagtcggcccttatgacgtcaatgacggtaaatggcccgctggcatt  
:tgccagtagcatgaccttacgggaacttccctacttggcagtagcatatgacgtatgacgttaaatggcccgctggcatt  
:gcgggtttggcagtagacaaatgggctggatagcgggttgactcacggggatttccaagcttccacccattgacgtc  
atgggagttgttttggcaccaaaatcaacgggacttccaaaatgtcgtaataaccccgcccgttgacgcaaatggg  
ggtaggcgtgtacgggtgggaggtctatataagcagagctcggttagtgaaacgctcagatcgctggagacgcatccac  
ctgttttgacctccatagaagacacgggacgcatccagcctccgcgccgggaacgggtgcattggaacggcggtatccc  
gtgccaagtagtagtaagtagcggctatagactctataggcaccctttggctcttgatgctatagctgtttt  
gcttggggcctatacaccggcttctcttatgctataggtggtatagcttagcctataggtggttattgacca  
tattgaccactccctattggtagcagatacttccattactaatccataacatggctcttgccacaactatctctatt  
gctatatgccaatactctgctcctcagagactgacacggactctgtattttacaggatgggggtccattttattatta

Fig. 15a

caaattcacatatacaacaacgccgtcccccgtagcccgagttttattaaacatagcgtgggatctccacgcgaatctc  
gggtaccgtgttccggacatgggytcttctccggtagcggcggagcttccacatccgagccctgggtcccatgcctccagc  
ggctcatggtegctcggcagctccttgctcctaacagtggaggccagacttaggcacagcacaatgccaccaccaccag  
tgtgccgcacaaggccgtggcggtagggatatgtgtctgaaaatgagctcggagattgggctcgcaccgctgacgcagatg  
gaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgattctgataagagtcagaggtaactcccggtt  
gcggtgctgttaacggtggaggggcagtgtagtctgagcagtagctcggtgctgccgcgcgcgccaccagacataatagctg  
acagactaacagactgttcctttccatgggtcttttctgcagtcacc

Fig. 15a

Figure 15b. pGA1/IC48 Functional Regions

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 104-1591                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | AG HIV-1 isolate 928                     |
| 1393-3576                                                      | Pol, inactivating point mutations in protease and reverse transcriptase, deletion of integrase  | AG HIV-1 isolate 928                     |
| 3663-3877 and 6215-6305                                        | Tat exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3802-3877 and 6215-6510                                        | Rev exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3904-4149                                                      | Vpu                                                                                             | AG HIV-1 isolate 928                     |
| 4067-6628                                                      | Env                                                                                             | AG HIV-1 isolate 928                     |
| 6629-10447                                                     | Plasmid vector                                                                                  | PGA1, Gen Bank accession # AF425297      |

Figure 15c. IC48 Mutation Chart

| Codon change                 | Amino acid change | Region                   | Function                                        |
|------------------------------|-------------------|--------------------------|-------------------------------------------------|
| to AGC position<br>1271-1273 | C390S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1280-1282 | C393S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1334-1336 | C411S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1343-1345 | C414S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to GTG position<br>1703-1705 | G48V              | Protease                 | Partial inactivation<br>of Protease             |
| to AAC position<br>2410-2412 | D185N             | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC position<br>2653-2655 | W266T             | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG position<br>3289-3291 | E478Q             | Rnase H                  | Inactivation of<br>RNase H activity             |

Figure 16a. pGA1/IC90 sequence

atcgatgcaaggactcggcttgcgtgaggtgcacacagcaagaggcgagagcgacgactggtgagtagcgaatttttgc  
tagcggaggcttagaaggagagagatgggtgcgagagcgtcagtggttaacggggggaaaatttagattcatgggagaaaatt  
agggttaaggccagggggaaagaaaagatatagactaaaacactagtagtggtgcaagcaggagctggagagattcgcact  
taacctggcctatttagaaacagcagaaggatgtcaacaactaatgggacagttacaaccagctctcaggacaggatcag  
aagagtttaaatcattatataatatagtagcaaccctttggtgcgtacatcaaagaatagacataaaagacaccaggag  
gccttagataaagtagagggaaaaacaaaataagagcaagcaaaaggcacagcaggcagcagctgcaacagccgccacagg  
aagcagcagccaaaattaccctatagtgcaaaatgcacaaggcgaatggtacatcagtcctatgtcacctaggactttaa  
atgcatgggtgaaaggtaatagaagaaaaggcttttagccagaggtaatacccatggttttcagcattatcagagggagcc  
acccacaagatttaaatatgatgctaacaatagtggggggacaccaggcagcaatgcagatggttaaaagataccatcaa  
tgatgaagctgcagaattgggacagagtagacatccagtgaggcctattccaccaggccaaatgaggggaaccaaagg  
gaagtgcatagcaggaactactagtacccttcaagaacaaataggatggatgacaagttaaccacctatcccagtgga  
gaaatctataaaagatggatagtcctgggattaaataaaatagtaagaatgtatagccctaccagcattttggacataag  
acaaggcccaaaagaacccttttagagattatgtagacaggttctttaaaactttgagagctgaacaagctacgcaggagg  
taaaaaactggatgacagaaaccttgttgggtccaaatgcgaatccagactgcaagtcctatttaagagcatttaggacca  
ggggctacattagaagaaatgatgacatcatgtcagggagtgaggaggacctggccataaaagcaagggttttggctgaggc  
aatgagtcaggtacaacagaccaatgtaatgatgcagagaggcaatttttagaggccagagaataataaagagcttcaaca  
gcggcaaaaggaagacaccttagccagaaaattgcaaggctcctagaaagagagggcagctggaaaagcggaagggacac  
caaatgaaagactgtactgaaagacaggctaattttttagggaaaatttggccttcccacaaggggaggccaggaaattt  
tcctcagagcagaccagaaccaacagcccgccagcagagagctttggagtgggggaagagataccctcctctccgaagc  
aggagccgaggggacaagggaactatctcctccttaacttccctcaaatcactcttgggcaacgaccagtagtcacagtaa  
gaatagggggacagccaatagaagccctattagacacaggagcagatgatacagattagaagaaataagtttaccagga  
aaatggaaaccaaataatgatagggggaattggaggttttatcaagtaagacagtagtatcagatatctatagaataattg  
tgaaaaagaggaactaggtacagttatagtaggacctacactgtcaacataattggacgaaataggtgactcagattg  
gtgtacttttaattttccaatttagtcctattgaaactgtccagtaaaattaaagttaggaatggatggcccaaagggt  
aaacaatggccattgacagaagaaaaataaaagcattaaaagaaatttgtgcagagatggaaaaggaaggaataattc  
aaaaattgggcctgaaaaccatacaatactccaatatttggccataaaagaaaaagatagtactaaatggagaaaaattag  
tagatttcagagaactcaataagagaactcaagacttctgggaggtccaattaggaatacctcatcctgcgggattaaaa  
aagaaaaaacagtaaacagtagtagatgtgggggtagcatattttcagttcccttagatgaagacttagaaaaatatac  
tgcatccacatacctagtttaataaatgagacaccagggttagatatcagtaacaatgtactcccacagggttaggaag  
gatcaccagcaatatttcaggcaagcatgacaaaaatcttagagcccttagagcaaaaaatccagagatagtgatctac  
caatatatgaacgattttatgttaggtctgacttagaaaatagggcagcatagagcaaaaaatagaggagttgagagaaca  
tctattgaaatggggatttaccacaccagacaaaaaacatcagaaagaacctccatttcttggatgggatatgaactcc  
atcctgacaaaatggacagtcacgcctatacagctgccagaaaaagacagctggactgtcaatgatatacaaaaaattagt  
ggaaaactaaataccgcaagtgcagatttatgcaggaattaaagtaaaagcaattgtgtagactcctcaggggagccaaagc  
gctaacagatgtagtaacacagtagtggaagcagaattagaattggcagagaacagggaattctaaaagaacctgtac  
atggagtagatattatgacccaacaaaagacttagtggcagaaaatacagaacaaagggaagatcaatggacatatcaaat  
tatcaagagccatttaaaaaatctaaagacaggaaaatatgcaaaaaagaggtcgggccacactaatgatgtaaaacaatt  
aacagaggttagtgcaaaaatagccatagaagcagatagtaatatggggaagagaccctaaatttagactaccatacaaaa  
gagaaacatgggaagcatggtggatggagtatggcagggtacctggattcctgaatgggagtttgtcaataccccctcct  
ctagtaaaaattatggtaccagtagtagaaggaagcccaataatgggagcagaactttctatgtagatggggcagctaatag  
ggagactaaagtaggaaaagcagggtatgtcactgacagagggaagacaaaagggtgtttccctaatcagacacaaaatc  
aaaagactcagttacatgcaattcatctagccttgaggattcaggatcagaagtaaatatagtaaacagactcacagtat  
gcattaggaatcattcaggcacaaccagacaggagtgaaatcagagtttagtcaatcaataatagagaaactaatagaaaa  
ggacaaagtctacctgtcatgggtaccagcacacaaagggtatggaggaatgaacaagtagataaaattagtcagtagtg  
gaatcagaaggtactatttttagatggaatagataaaagcccaagatgaacattagaattctgcaacagctactgtttgt  
tcatttcagaattgggtgtcaacatagcagaataggcatattccaggggagaagaggcaggaaatggagctggtagatcct  
agcctagagccctggaaccaccgggaagtcagcctacaactgctttagcaaggttactgtaaaaaatgctgctggca  
ttgccaatttgtgtttctgaacaagggttaggcatctcctatggcaggaagaagcggagacgcccagcagggaactcctc  
aggaccgtcaggttcatcaaaatcctgtacaaaacagtaagtagtagtaattagtagtatgtgatgcaatctttacaaat  
agctgcaatagtaggacttagtagtagcatccatagtagccatagttgtgtggtccatagtagtttatagaatatagaaaaa  
taaggaaacagaagaaaaatagacaggttacttgagagaataagagaaagagcagaagatagtggaatgagagtgatggg  
gatacagaagaattatccactcttagggaggggggtatgacaatattttgggtaatgatgatttgtaatgctgaaaagt  
tgtgggtcacagcttactatggggtaacctgtgtggagagagcagagaccacctattctgtgcatcagatgctaaagca  
tatggcaaaagaagcacacaatgtctgggtacgcatgcctgcgtacccacagaccctgaccacagaattacctttggt  
aatgttaacagaagagtttaacatgtggaaaaataatatggtagaacagatgcatgaagatataatttagtctatgggacc  
aaagcttaagccatgtgtacagctaaccctctctgcgttactttagggtgtgctgacgctcaaaacgtcaccgacacc  
aacaccacatattctaatgaaatgcaaggggaaataaaaaactgctcttcaatatgaccacagaattaagagataagaa  
gcagaaagtgtatgacttttttatagacctgatgtaatagaaatttaataaaaactaagatttaacaatagtaatagtagtc  
agtatatgttaataaaattgtaataacctcaaccttacacagactgtccaaagggtatcctttgagccaattcccatacat  
tattgtgccccagctgggttttgcaattctaaagtgtaatgatacggaggttcagtggaaggggacatgcaagaggtgtcag  
cacagtacaatgcacacatggaatcaagccagtagtatcaactcaactgctgttaaatggcagctctagcagaaggaaaga



tagcgattagatctgagaatatctcaaacaatgccaaaactataatagtacaattgactgagcctgtagaaattaattgt  
atcagacctggcaacaatacaagaaaaagtgtacgcataaggaccaggacaaacattctatgcaacagggtgacataatagg  
agatataagacaagcacactgttaattgttagtaaaatagcatgggaagaaactttacaaaaggtagctgcacaattaagga  
agcactttcagaatgccacaataaaatttactaaacactcaggaggggatttagaaattacaacaaatagttttaattgt  
ggaggagaatttttctattgcaatacaacaaagctgtttaatagcacttggaaataatgataactcaaacctcacagagga  
aaagagaaaggaaaaacataactctccactgcagaataaagcaaatgttaaatatgtggccaagagtaggacaagcaatat  
atgccccctcccatcccaggaaacataacttgtggatcaaacattactgggctactattaacaagagatggagggaataat  
ggtacaaatgatactgagaccttcaggcctggaggaggagatatgagggacaattggagaagtgaattatataataataa  
agtagtaaaaaattgaaccactagggtgtagcaccacccctgcaaaaagaagagtgggtggaaagagaaaaagagcagttg  
gaatgggagcctttgatctttgagttcttaggagcagcaggaagcactatgggcgcggtcaatggcgctgacggtacag  
gccagacaattattgtctgggtatagtgtcaacagcagagcaatctgctgaaggctatagaggctcaacaacatctgttgag  
actcacggtctggggcattaaacagctccaggcaagagtcctggctctggaaagataacctaaaggatcaacagctccttag  
gaatttggggctgctctggaaaactcatttgcaccactgctgtaccttggaaactctagctggagtaataaaagtataat  
gacatatgggataacatgacctggctgcaatgggataaagaaatatacaattacacatacataatataatctacttga  
aaaaatcgcagaaccagcaggaaattaatgaacaagacttattggcattagacaagtgggcaagtctgtggaattgggttg  
acataacaagctggctatgggtatataagattaggtataatgatagtagggcgtaataaggcttaagaataatttttgc  
gtgcttactatagtgaatagagttaggcagggtatactcaccttgtcattccagaccttgcccaccaccagagggaacc  
cgacaggcccgaaagaatcgaagaaggaggtggcgagcaagacagagagatccgtgcgcttagtgagcggatcttag  
cacttgcctgggaagatctgcggagcctgtgctcttcagctaccgctgattgagagacttagtcttgattgagcgaagg  
actgtggaactcctgggacacagcagctctcaagggactgagactggggtgggaagccctcaaatatctgtggaaccttct  
atcatactggggtcaggaactaaagaatagtgtatttaatttgccttgatacaatagcaatagcagtagctaaactggacag  
atagagttataaaaaatagtaaaagaactggtagagctattcttaacataacctagaaggatcagatagggttagccccgg  
gtgataaacggaccgcaatccctaggctgtgcttctagtggcagccatctgttgttggccctccccctgcttct  
cttgacctggaaggtgccactcccactgtccttccataaaaatgaggaaatgcatcgcatgtctgtgagtggtgtc  
attctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcg  
gtgggctctatataaaaaacgcccggcgcaaccgagcgttctgaacgctagagtgcacaaattcagaagaactcgtcaa  
gaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaaagcacgaggaagcgggtcagccattcgcg  
ccaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcgtctgccacaccagccggccacagtcgat  
gaatccagaaaagcggcatttccaccatgatattcggcaagcaggcatcgccatgggtcacgacgagatcctcgccgt  
cgggcatgctcgcttgagcctggcgaaacagttcggctggcgcgagccctgatgctcttcgctccagatcatcctgag  
acaagaccggcttccatccgagtagctgtcgtcgtatgcgagtgttctgcttgggtggtcgaatgggcaggtagccggatc  
aagcgtatgcagccgcccgtatgcacagccatgatggatactttctcggcaggagcaaggtgagatgacaggagatcct  
jccccggcacttcgcccataagcagccagtccttcccgttctagtgacaacgtcgagcacagctgcgcaaggaacgccc  
gtcgtggccagccagatagccgctgctcgtcttcaggttcattcagggcaccggacaggtcggctcttgacaaaaag  
aacggggcgccccctgcgtgacagccggaacacggcgagcagcagcagcagctgtctgttggccagtcagatccgga  
atagcctctccaccaagcggccggagaacctgcgtgcaatccatctgttcaatcatgcgaaacgatcctcatcctgtc  
:cttgatcagatcttgatccccctgcgccatcagatccttggcggaagaaagccatccagtttactttgcagggttccc  
aaccttaccagagggtcgccccagctggcaattccggttcgcttgcgttccataaaaaccgcccagcttagctatcgccatg  
:aagcccactgcaagctacttgccttctctttgcgcttgcgctttcccttgcagatagcccagtagctgacattcatc  
:ggggtcagcacgcttctgcgagctggcttctacgtgaaaaggatctaggtgaagatccttttgataatctcatgac  
:aaaaatcccttaacgtgagtttctgcttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctt  
:tttctgcgctaatctgctgcttgcaaaacaaaaaaccaccgctaccagcgggtggtttggttgcgggataagagcta  
:caactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttagg  
:caccacttcaagaactctgtagcaccgctacatacctcgctctgctaactcctgttaccagtggctgctgccagtggcg  
:taagtctgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggctcgggctgaacgggggggttcg  
:gcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccagct  
:cccgaaggggagaaaggcggacaggtatccggttaagcggcagggctcggaacaggagagcgcacgaggagcttccagggg  
:jaacgcctgggtatctttatagtcctgtcgggttgcggcactctgagcttgagcgtcgatttttgtgatgctcgtcaggg  
:ggcgagcctatggaaaaacgccagcaacgcggccctttacgggtcctggccttttgcgtggttcttgcacatgtt  
:tcgacaatatgtggctattggccatgcatagcttgatctatatacataatatgtacatttatattggctcatgtccaat  
:tgaccgcatgttgacattgattattgactagttattaatagtaatacaattacgggttcatttagtctatagcccatata  
:ggagttccgcttgataacttaacggttaaatggccgctggctgaccgccaacgacccccgcccattgacgtcaata  
:tgacgtatgttcccatagtaacgccaatagggaacttccattgacgtcaatgggtggagtatttacggtaaaactggccat  
:ttggcagtagacatcaagtgtatcatatgccaagtcgccccctattgacgtcaatgacggtaaatggccgctggcatt  
:tgcccagtagacattacgggacttccctacttggcagtagacatctacgggtatttagtcatcggtattaccatggtga  
:gcgggttttggcagtagaccaatgggctggatagcgggttgactcacggggatttccaagtctccacccattgacgtc  
:atgggagtttgttttggcaccaaaatcaacgggacttccaaaatgtcgtaataaccccggccggtgacgcaaatggg  
:ggtaggcgtgacgggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgctggagacgccatccac  
:ctgttttgacctccatagaagacacgggacggatccagcctccgcgccgggaacgggtgcatggaacgcggattccc  
:gtgccaagagtgacgttaagtaccgctatagactctataggcacccttggctcttatgcatgctatgctattgttt  
:gcttggggcctatacaccctcgcttccctatgctataggtgaggtatagcttagcctataggtggttatttgacca  
:tattgaccactccccattgggtgacgatacttccattactaatccataacatggctcttggccacaactatctctatt  
:gctatatgccaatactctgtccttcagagactgacacggactctgtattttacaggatgggggtccattttatttta

Fig. 16a

caaattcacatatacaacaacgccgtcccccgtagcccgagtttttattaaacatagcgtgggatctccacgcgaatctc  
gggtaccgtgttccggacatgggytcttctccggtagcggcgagcttccacatccgagccctgggtcccatgcctccagc  
ggctcatggtcgctcggcagctccttgctcctaacagtggaggccagacttaggcacagcacaatgccaccaccaccag  
tgtgccgcacaaggccgtggcggtagggtatgtgtctgaaaatgagctcggagattgggctcgcaccgctgacgcagatg  
gaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgattctgataagagtcagaggtaactcccggt  
gcggtgctgttaacggtggagggcagtgtagtctgagcagtagctggtgctgccgcgcgcgccaccagacataatagctg  
acagactaacagactgttcctttccatgggtcttttctgcagtcacc

Fig. 16a

**Figure 16b. pGA1/IC90 Functional Regions**

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b> | <b>Origin of sequence<br/>(organism)</b> |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|
| 104-1591                                                       | Gag, inactivating point mutations in zinc finger packaging sequences for viral RNA              | AG HIV-1 isolate 928                     |
| 1393-3576                                                      | Pol, inactivating point mutations in protease and reverse transcriptase, deletion of integrase  | AG HIV-1 isolate 928                     |
| 3663-3877 and 6215-6305                                        | Tat exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3802-3877 and 6215-6510                                        | Rev exons #1 and #2                                                                             | AG HIV-1 isolate 928                     |
| 3904-4149                                                      | Vpu                                                                                             | AG HIV-1 isolate 928                     |
| 4067-6628                                                      | Env                                                                                             | AG HIV-1 isolate 928                     |
| 6629-10447                                                     | Plasmid vector                                                                                  | PGA1, Gen Bank accession # AF425297      |

**Figure 16c. IC90 Mutation Chart**

| <b>Codon change</b>          | <b>Amino acid change</b> | <b>Region</b>            | <b>Function</b>                                 |
|------------------------------|--------------------------|--------------------------|-------------------------------------------------|
| to AGC position<br>1271-1273 | C390S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1280-1282 | C393S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1334-1336 | C411S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to AGC position<br>1343-1345 | C414S                    | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to ATG position<br>1828-1830 | L90M                     | Protease                 | Partial inactivation<br>of Protease             |
| to AAC position<br>2410-2412 | D185N                    | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACC position<br>2653-2655 | W266T                    | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAG position<br>3289-3291 | E478Q                    | Rnase H                  | Inactivation of<br>RNase H activity             |

Figure 17a. PGA1/IN3 Sequence

ggatccggcttgctgaagtgcaactcggcaagagggcgaggggtggcggtggtgagtagcgcacaaattttatttgactagcg  
gaggttagaaggagagagatgggtgagagagcgtaataatagagggggaaaattagataaatgggaaaagattaggtt  
aaggccagggggaaagaaacacataatgctaaaacacctagtagtggaagcagggagctggaaagatttgcaactaacc  
ctggccttttagagacatcagaaggctgtaaacaaatataaaacagctacaaccagctcttcagacaggaacagaggaa  
cttaggtcattattcaatgcagtagcaactctctattgtgtacatgcagacatagaggtacgagacaccaaagaagcatt  
agacaagatagaggaagaacaaaacaaaagtcagcaaaaaacgcagcaggcaggaagggctgacaaaaaggtcgtagtc  
aaaattatcctatagtgagaatcttcaagggcaaatggtacaccaggcactatcacctagaactttgaatgcatgggtg  
aaagtaatagaagaaaaagccttttagcccgagggttaatacccatgttcacagcattatcagaaggagccacccccacaaga  
tttaaacaccatgttaaataccgtggggggacatcaagcagccatgcacaaatgttaaaagataccatcaatgaggaggtg  
cagaatgggtagattacatccagtacatgcagggcctgttgaccaggccaaatgagagaaccaaggggaagtgcata  
gcaggaactactagtaaccttcaggaacaaatagcatggatgacaaagtaaccacctattccagtgggagatatctataa  
aagatggataattctggggttaataaaatagtaagaatgtatagccctgtcagcatttttagacataagacaagggccaa  
aggaaaccttttagagattatgtagaccggtctcttaaaactttaagagctgaacaagcttcacaagatgtaaaaaattgg  
atggcagacaccttgttggtccaaaaatgcgaacccagattgtaagaccattttaagagcatttaggaccaggagctacatt  
agaagaaatgatgacagcatgtcaaggagtgaggagcctagccacaaagcaagagtggtgggtgaggcaatgagccaaa  
caggcagtagccataatgtatgcagagaagcaatttttaaggctctaaagaactgttaaatccttcaactctggcaaggaa  
gggcacatagctagaaattgcagggcccttaggaaaaaggctcttggaatctggaaagggaacaccaaattgaaaga  
ctgtgctgagaggcaggtcaatttttttagggaaaaattggccttcccacaagggaggccagggaatttccttcagaaca  
ggccagagccaaacagccccaccagcagagagcttcagggttcgaggagacaacccctgctccgaagcaggagctgaaagac  
agggaaccttaacctccctcaaatcactctttggcagcgacccttgtctcaataaaaaatagggggccagataaaaggag  
gctctcttagccacagggcagatgatacagtagtagaagaatgaatttgccaggaaaaatggaaacaaaaaatgatagg  
aggaattggaggttttatcaaaagtaagacagtagtacaacttatagaaatttggtggaaaaaggctaataggtacag  
tattagtaggaccacacctgtcaacataattggaagaaatgtctgactcagattggatgcacgctaaattttccaatt  
agtcctatgaaactgtaccagtaaaattaaagccaggaatggatggcccaaagggttaaaacatggccattgacagagga  
gaaaataaaagcatttaacagcaatttgatgataatggagaagggaaggaataacaaaaattgggctgaaaaatccat  
ataacactccaatatcgccataaaaaagaaggacagtagtaagtgaggaaaaattagtagatttcagagaacttaataaa  
agaactcaagacttctgggaagttcaattaggaataccacacccagcagggttaaaaaagaaaaaatcagtgacagtagt  
agatgtgggggtagcatatttttcagttcctttagatgaaagcttttaggaggtatactgcattcaccaatccttagtaga  
acaatgaaacaccagggttagatatcaatataatgtgcttcacaaggatggaaaggatcaccagcaaatattccagagt  
agcatgacaaaaatcttagagcccttttagagcacaaaaatccagaaatagtcattctatcaatataatgaatgacttgtatgt  
aggatctgacttagaaatagggcaacatagagcaaatagaggaatgaagagaacatctattaagggtggggatttacca  
caccagacaagaacatcagaagaacccccatttctttggatgggggtatgaactccatcctgcacaaatggacagtacag  
cctatacagctgccagaaaaaggagagctggactgtcaatgatatacagaagtagtgggaaaaattaaacacggcaagcca  
gatttaccaggggttaaagtaagacaactttgtagactccttagaggggccaagcactaacagacatagtagcacttaa  
ctgaagaagcagaattagaattggcagagaacagggaatttcaaaagaaccagtagatggagtatattagacccttca  
aaagacttgatagctgaaatacagaacacaggacatgaccaatggacatatcaaatttaccaagaaccattcaaaaaatct  
gaaaacaggggaagtatgcaaaaaatgaggactgccacactaatgatgtaaaacgggttaacagaggcagtgcaaaaaatag  
ccttagaaagcatagtaatatggggaaagatttctaaacttaggttacccatccaaaaagaaacatgggagacatggtgg  
actgactattggcaagccacctggattcctgagtggggaatttggttaatactcctcccttagtataaattatggtaccagct  
agagaaggaaaccataataggagtagaaacttctatgtagatggagcagctaaatgggaaaccaaataaggaaagcag  
ggtagttactgacagaggaaggcagaaaaattgtttctctaactgaaacaacaaatcagaagactcaattacaagcaatt  
tatctagctttgcaagattcaggatcagaagtaaacatagtaacagactcacagtagcattaggaattattcaagcaca  
accagataagagtgaaatcagggttagtcaaccaaataatagaacaatttaataaaaaaggaaagggtctacctgtcatggg  
taccagcacataaaggatttgagggaatgaacaagtagacaaattagtaagtagtggaatcaggagagtgctataataa  
gctcgagatacttggaaggaggttgaaactatcataagaatgctgcaacaactactgtttattcatttcagaattgggtg  
ccagcatagcagaataggcatatgagacagagaagagcaagaaatggagccagtagatcctaacctagagccctggaa  
catccaggaagtcagcctgaaactgcttgcaataactgttatgtaaacgctatagctaccatttgtctagttgtcttca  
gagaaaaggcttaggcatttctatggcaggaagaagcggagacagcgagcgaagcgctcctcagagcagtgaggatcatc  
agaattttgtatcaaagcagtaagtagtctgtaattgttagattataaattagcagtaggagcatttatagtagca  
ctactcatagcaatagttgtgtggaccatagtagttatagaatataggaaattgttaagacaaagaaaaatagactgggt  
aattaaaaagaattagggaagagcagaagacagtggaatgagagtgaaaggggatactgaggaaattatcgacaatgggtg  
atatggggcatcttaggcttttggatgttaatgattgtaattggaacttggtgggtcacagctctattatggggtagctgt  
gtggaaagaagcaaaaactactctattctgtgcatcaaatgctaaagcatatgagaaagaagtacataatgtctgggcta  
cacatgctgtgtacccacagacccccacacagaatgggttttggaacacgtaacagaaaaatttaacatgtggaaa  
aatgacatgggtgaatcagatgcatgaggatgtaatcagcttatgggatcaaagcctaaagccatgtgtaaagttgacccc  
actctgtgtcactttagaatgtagaaagggttaatgctaccataatgctaccaataatggggatgctaccataatgtta  
ccaataatgggcaagaaatacaaaatgtgctctttcaatgcaaccacagaaataagagataggaaagcagagagtgtagca  
cttttttagacttgatagtagtaccacttgataagaacaactctagtaagaacaactctagtgagtagttatagtagta  
aaattgttaactcctcagccataacacaagcatgtccaaaggctcagtttgatccaattcctatacactattgtgtctcag  
ctgggtatgcgattctaaagtgtacaataagacattcaatgggacaggacatgcaataatgtcagcacagacaatgt  
acacatgggaattaagccagtggtatcaactcagctattgttaacggtagcctagcagaaggagagataataattagatc

tgaaaatctgacagacaatgtcaaaacaataatagtacatcttgatcaatctgtagaaattgtgtgtacaagacccaaca  
ataatacaagaaaaagtataaggatagggccaggacaaacattctatgcaacaggagcataataggggaacatacgacaa  
gcacattgtaacattagtgaagacaaatggaatgaaactttacaaagggtgggtaaaaaatttagtagaacacttccctaa  
taagacaataaaaatttgcacccatcctcaggaggggacctagaaattacaacacatagctttaaattgtagaggagaatttt  
tctattgcagcacatcaagactgtttaatagtacatacatgccaatgatatacaaaaagtaagtcaaaacaaacccatcaca  
atcccatgcagcataaaacaaattgtaaacatgtggcaggaggtaggacgagcaatgtatgccctcccatgaaggaaa  
cataacctgtagatcaaatatcacaggaatactatttggtacgtgtatggaggagtagattcagaagatccagaaaaataa  
agacagagacattccgacctggaggaggagatatgaggaacaattggagaagtgaattatataaatataaagcggcagaa  
attaagccattgggagtagcaccactccagcaaaaaggagagtgggtggagagagaaaaagagcagtaggattaggagc  
tgtgttccctggattcttgggagcaggaagcactatggcgagcgtcaataacgctgacgggtacaggccagacaat  
tgttgtctgggtatagtgcacagcaaaagcaatttgctgagggctatcgaggcgcaacagcatctgttgcactcacggct  
tggggcattaaagcagctccagacaagagtccctggctatcgaaagatacctaaggaatcaacagctcctagggtttgggg  
ctgctctggaaaactcatctgcaccactaatgtaccttggaaactccagtggagtaacaaatctcaaacagatatttggg  
aaaacatgacctggatgcagtgggataaagaagtttagtaattacacagacacaatataggttgcttgaagactcgcaa  
accagcaggaaagaaatgaaaaggatttattagcattggacaattggaaaaatctgtggaattgggtttagtataacaaa  
ctggctgtggtatataaaaaatttcataatgatagtagagggttgataggcttaagaataatttttgcctgtgctttcta  
tagtgaatagagttaggcagggtatactcacctttgtcgtttcagaccttaccceaaacccaaggggacccgacaggctc  
ggaagaatcgagaagaagggtggaggggcaagacagagacagatcgattcgattagtgaacggattcttagcacttgctc  
ggacgacctgtggagcctgtgacctctcagctaccaccgattgagagacttaatttgggtgacagcgagagcgggtggaac  
ttctgggacacagcagctcctcaggggactacagaggggggtgggaagcccttaagtatctgggaggtattgtgcagattgg  
ggtctggaactaaaaaagagggtctattagtctgcttgatactgtagcaatagcagtagctgaaggcacagataggattat  
agaattcctccaaagaattttagtagctatccgcaacatacctagaaggataagacaggggtttgaagcagctttgcagt  
aaaatggctgacccccgggtgataaaacggaccgagccatctcctaggctgtgcttctagtgtgccagccatctgtgttgc  
ccctccccctgtccttcccttgaccttgaaggtgccactcccactgtccttctcctaataaaaaatgaggaaattgcacgc  
ttgtctgagtaggtgtcattctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagca  
ggcatgctggggatgcgggtgggctctatataaaaaacgcccggcggaaccgagcgttctgaacgctagagtcgacaaat  
tcagaagaactcgtcaagaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaaagcacgaggaagc  
ggtcagcccatctgcgcgaagctcttcagcaatatcacgggtagccaacgctatgtcctgatagcgggtctgccacaccc  
agccggccacagtcgatgaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcac  
gacgagatcctcgccgtcgggcatgctcgccttgagcctggcgagacagttcggctggcgagccctgtatgctctcgt  
ccagatcatcctgatcgacaagaccggcttccatccaggtacgtgctcgcctcgatgcgatgtttcgcttgggtggtcgaat  
ggcaggttagccggtcaagcgtatgcagccgcccgtattgcatcagccatgatggatactttctcggcaggagcaagggtg  
agatgacaggagatcctgccccggcacttcgccaatagcagccagtccttcccgttccagtgaacacgtcgagcacag  
ctgcgaaggaaacgcccgtcgtggccagccagatagccgctgctcctcgtcttcaggttccattcagggcaccggacagg  
tcggctcttgacaaaaaagaacccggcgccctcgcgtgacagcgggaacacggcgagcagacagcagctgtctgtgtg  
tgcccagtcatagccaatagcctctccaccaagcggcgaggaacctgcgtgcaatccatcttgttcaatcatgcgaa  
acgatcctcatcctgtctcttgatcagatcttgatccctgcgccatcagatccttggcggaagaaagccatccagttt  
actttgcaggggttcccaaccttaccagagggcgccccagctggcaattccggttcgcttgctgtccataaaaccgccc  
gtctagctatcgccatgtaagcccaactgcaagctacctgcttctctttgcgcttgctgtttcccttgtccagatagccc  
agtagctgacattcatccggggtcagcacctgttctgcggactggctttctacgtgaaaaggatctagggtgaagatcctt  
tttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaagg  
atcttcttgagatccttttttctgcgcgtaattctgctgcttgcaaaacaaaaaaccacggctaccagcgggtgtgtt  
tgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttcta  
gtgtagccgtagtttaggccaccacttcaagaactctgtagaccgcctacatacctcgctctgctaactcctgttaccagt  
ggctgctgccagtgggcagataagtcgtgtcttaccgggttggaactcaagacgatatgttaccggataaggcgcagcggctcgg  
gctgaacgggggggttcgtgcacacagcccagcttgagcgaaacgacctacaccgaactgagatacctacagcgtgagcta  
tgagaaagcgccacgcttcccgaagggaagaaaggcggaagggtatccggtaagcggcagggtcggaacaggagagcgcac  
gaggagcgttccaggggggaaacgcctgggtatctttatagtcctgtcgggttccgccaccttgacttgagcgtcgatattt  
tgtgatgctcgtcagggggcgagcctatggaaaaacgccgaacgcggccctttacgggttccctggccttttgcctgg  
ccttttgcctcacatgttgtcgacaatattggctattggccattgcatacgttgatctatatcataatatgtacatttat  
attggctcatgtccaatatgaccgccatgttgacattgattattgactagttattaatagtaatacaattacgggttcatt  
agttcatagcccatataatggagttccgcgttacataacttacggtaaatggccgcctggctgacggcccaacgaccccc  
gcccattgacgtcaataatgagctatgttcccatagtaaacgccaatagggactttccattgacgtcaatgggtggaggtat  
ttacggtaaaactgcccacttggcagtagatcaagtgatcatatgccaagtccgccccctattgacgtcaatgacgggtaa  
atggcccgctggcatatgcccagtagacacgttacgggaactttccctacttggcagtagatctacgggtattagtcac  
ggctattaccatgggtgatgcgggttttggcagtagaccaatggcggtggatagcgggttgactcacggggatttccaagtc  
tccaccccatgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaataacccccgcc  
ccgttgacgcaaatggcggttaggcgtgtacgggtgggagggtctatataagcagagctcgtttagtgaaccgtcagatcgc  
ctggagacgcctccacgctgttttgacctccatagaagacacggggaccgatccagcctccggggcgggaacgggtgca  
ttggaacgggattccccgtgccaaagagtgacgttaacgccttagactctataggcacacccctttggctcttatg  
catgctatactgttttggcttggggcctatacaccgcccttccctatgctataggtgaggtatagcttagcctatag  
gtgtgggttattgaccattattgaccactcccctattgggtgacgatactttccattactaatccataacatggctctttg  
ccacaactatctctattggctatattgccaatactctgtccttcagagactgacacggactctgtatttttacaggatggg

Fig. 17a

gtccattttattttacaaattcacatatacaacaacgccccgtgccccgagtttttattaaacatagcgtggg  
atctccacgcgaatctcgggtaccgtgttccggacatgggytcttctccggtagcggcggagcttccacatccgagccct  
ggccccatgcctccagcggctcatggctcgctcggcagctccttgctcctaacagtggaggccagacttaggcacagcaca  
atgcccaccaccaccagtgtgccgcacaaggccgtggcggtaggggtatgtgtctgaaaatgagctcggagattgggctcg  
caccgtgacgcagatggaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgtattctgataagagt  
cagaggtaactcccgttgcggtgctgttaacgggtggagggcagtgtagctgagcagtactcgttgctgccgcgcgcgcc  
accagacataatagctgacagactaacagactgttcctttccatgggtcttttctgcagtcaccat

Fig. 17a

**Figure 17b. pGA1 IN52/73 Inunctional Regions**

| <b>Position within sequence<br/>(starting-and ending pos.)</b> | <b>Gene/ORF (indicate<br/>complete name of the gene<br/>and the expressed gene<br/>product)</b>                   | <b>Origin of sequence<br/>(organism)</b>       |
|----------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|------------------------------------------------|
| <b>99-1577</b>                                                 | <b>Gag, inactivating point<br/>mutations in zinc finger<br/>packaging sequences for<br/>viral RNA</b>             | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>1382-3520</b>                                               | <b>Pol, inactivating point<br/>mutations in protease and<br/>reverse transcriptase,<br/>deletion of integrase</b> | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>3645-3859 and 6205-6295</b>                                 | <b>Tat exons #1 and #2</b>                                                                                        | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>3784-3859 and 6205-6452</b>                                 | <b>Rev exons #1 and #2</b>                                                                                        | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>3873-4121</b>                                               | <b>Vpu</b>                                                                                                        | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>4039-6642</b>                                               | <b>Env</b>                                                                                                        | <b>C HIV-1 isolate 98IN012</b>                 |
| <b>6643-10466</b>                                              | <b>Plasmid vector</b>                                                                                             | <b>pGA1, Gen Bank accession #<br/>AF425297</b> |



Figure 17c. IN3 Mutation Chart

| Codon change                 | Amino acid change | Region                   | Function                                        |
|------------------------------|-------------------|--------------------------|-------------------------------------------------|
| to TCC position<br>1260-1262 | C390S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1269-1271 | C393S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1323-1325 | C411S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1332-1334 | C414S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to GCC position<br>1610-1612 | D25A              | Protease                 | Inactivation of<br>Protease active site         |
| to AAT position<br>2387-2389 | D185N             | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| to ACG position<br>2630-2632 | W266T             | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAA position<br>3266-3268 | E478Q             | Rnase H                  | Inactivation of<br>RNase H activity             |

Figure 18a. pGA1/IN2 Sequence

ggatccggccttgctgaagtgcactcggcaagaggcgaggggtggcggtggtgagtagcgcacaaattttatttgactagcg  
gaggctagaaggagagagatgggtgcgagagcgtcaatatataagaggggaaaattagataaatgggaaaagattaggtt  
aaggccagggggaagaaacactatatgctaaaacacctagtagtggaagcagggagctggaaagatttgcacttaacc  
ctggccttttagagacatcagaaggctgtaaacaaataataaaacagctacaaccagctcttcagacaggaacagaggaa  
cttaggtcattattcaatgcagtagcaactctctattgtgtacatgcagacatagaggtacgagacaccaaagagcatt  
agacaagatagaggaagaacaaaacaaagtcagcaaaaaacgcagcaggcaaaagaggctgacaaaaaggtcgtcagtc  
aaaattatcctatagtcagaatcttcaagggcaaatggtacaccaggcactatcacctagaactttgaatgcatgggtta  
aaagtaatagaagaaaaagcctttagcccggaggttaatacccatgttcacagcattatcagaaggagccaccccacaaga  
tttaaacaccatgtttaaataccgtggggggacatcaagcagccatgcaaatgttaaaagataccatcaatgaggaggctg  
cagaatgggtagattacatccagtacatgcagggcctgttgaccaggccaaatgagagaaccaaggggaagtgcata  
gcaggaactactagtaaaccttcagggaacaaatagcatggatgacaagtaaccacctattccagtgggagatatctataa  
aagatggataattctggggttaataaaaatagtaagaatgtatagccctgtcagcatttttagacataagacaagggccaa  
aggaacccttttagagattatgtagaccggttctttaaaactttaagagctgaacaagcttcacaagatgtaaaaaatggg  
atggcagacaccttgttgggtccaaaatgcgaacccagattgttaagaccattttaagagcattaggaccaggagctacatt  
agaagaaatgatgacagcatgtcaaggagtgggaggacctagccacaaagcaagagtggtgggtgaggcaatgagccaaa  
caggcagtaccataatgatgcagagaagcaattttaagggtctaaaagaactgttaaatgcttcaactgtggcaaggaa  
gggcacatagctagaaattgcagggcccttaggaaaaaggctgttggaattgtggaaggagacaccaaataagaa  
ctgtgctgagaggcaggctaattttttagggaaaatttggccttcccaaggggaggccagggaatttccctcagaaca  
ggccagagccaaacagccccacgcagcagagagcttcagggttcgaggagacaacccctgctccgaagcaggagctgaaagac  
agggaaacccttaacctccctcaaatcactcttggcagcagcccttgtctcaataaaaaatagggggccagataaaggag  
gctctcttagacacaggagcagatgatagctattagaagaaatgaatttgccaggaaaaatggaaacaaaaatgatagg  
aggaattggagggttttatcaagtaagacagtagatgataaatactttatagaatttgtggaaaaaggctataggtacag  
tatttagtaggacccacacctgtcaacataattggaagaaatattgctgactcagattggatgcagctaaatttccaatt  
agtcccatgaaactgtaccagtaaaaattaaagccaggaattgggtggcccaaagggttaaaccaatggccattgcagagga  
gaaaaataaaagcatttaacagcaatttgtgatgaaatggagaagggaagaaaaattacaaaaattgggctgaaaaatccat  
ataacactccaatattcgccataaaaaagaggacagtagtaagtggaagaaatagtagatttcagagaacttaataaa  
agaactcaagacttctgggaagttcaattaggaataccacacccagcaggggttaaaaaagaaaaaatcagtgcagtagt  
agatgtgggggatgcataattttcagttccttttagatgaaagcttttagagggtatactgcattcaccatacctagtagaa  
acaatgaaacacccagggattagatatcaatataatgtgcttccacaaggatggaaaggatcaccagcaatattccagagt  
agcatgacaaaaatcttagagcccttttagagcacaaaaatccagaaatagtcattctatcaatatatgaatgacttgtatgt  
aggatctgacttagaaatagggaacatagagcaaaagatagaggaattaaagagaacatctattaaggtggggattacca  
caccagacaagaaacatcagaaagaacccccatttcttggatgggggtatgaactccatcctgacaaatggacagtagacg  
cctatacagctgccagaaaaggagagctggactgtcaatgatatacagaagttagtgggaaaaataaacacgggcaagcca  
gatttaccaggggattaaagtaagacaactttgtgactccttagaggggccaagcactaacagacatagtaccactaa  
ctgaagaagcagaattagaattggcagagaacagggaatttctaaaagaaacagtagatggagtagattatgacccctca  
aaagacttgatagctgaaatacagaaacagggacatgaccaatggacatatcaaatttaccaagaaccattcaaaaatct  
gaaaacaggggaagtatgcaaaaatgaggactgccacactaatgatgtaaaacgggttaacagaggcagtgcaaaaaatag  
ccttagaaaagcatagtaatatggggaaagatttcttaaaacttaggttaccatccaaaaagaaacatgggagacatgggtg  
actgactattggcaagccacctggatttctgagtggggaatttgttaatactcctcccttagtaaaattatgggtaccagct  
agagaaggaaacccataataggagttagaaacttctatgtagatggagcagctaatagggaacaaaaataggaaagcag  
ggtaggttactgacagaggaaggcagaaaaattgttctcctaactgaaacaacaaatcagaagactcaattacaagcaatt  
tatctagcttgcagattcaggatcagaagtaaacatagtaacagactcacagtagcatttaggaattattcaagcaca  
accagataagagtgaaatcagggttagtcaaccaaataatagaacaattataaaaaaggaaaggggtctacctgtcatggg  
taccagcacataaagggtattggaggaaatgaacaagtagacaaattagtaagtagtggaatcaggagagtgctataataa  
gctcgagatacttgacagaggagtgaaactatcataagaatgctgcaacaactactgtttattcatttcagaattgggtg  
ccagcatagcagaataggcattatgagacagagaagagcaagaaatggagccagtagatcctaacctagagccctggaac  
catccaggaagtcagcctgaaactgttgcaataactgttattgttaaacgctatagctaccattgtctagtttgccttca  
gagaaaaaggcttaggcatttctatggcaggaagaagcggagacagcgacgaagcgtcctcagagcagtgaggatcatc  
agaattttgtatcaaagcagtaagtagtgaatgttagatttagattataaattagcagtaggagcatttatagtagca  
ctactcatagcaatagttgtgtggaccatagtagttatagaatataggaaattgttaagacaaagaaaaatagactgggt  
aattaaaagaattagggaagagcagaagacagtggaatgagagtgaaggggatactgaggaattatcgacaatgggtg  
atatggggcatcttaggcttttggatgttaatgatttgaatggaaactgtgggtcacagctctattatggggtacctgt  
gtggaaagaagcaaaaactactctattctgtgcataaattgttctaaagcatatgagaagaagtacataatgtctgggcta  
cacatgcctgtgtaccacagaccccaaccacagaagaatgggtttggaaaaacgtaacagaaaaatttacaatgtggaaa  
aatgacatggtgaatcagatgcagtaggagtgtaatcagcttatgggatcaaagcctaaagccatgtgtaaagttagcccc  
actctgtgtcactttagaatgtagaagggttaagtgtaccataatgctaccaataatggggatgctaccataatgtta  
ccaataatgggcaagaaatacaaaattgctcttcaatgcaaccacagaaataagagataggaagcagagagtgtagca  
cttttttatagacttgatatagtaccacttgataagaacaactctagtaagaacaactctagtgagtagtatatagattaat  
aatgttaataacctcagccataacacaagcatgtccaaagggtcagttttgatccaattcctatacactattgtgtccag  
ctgggttagcgattctaaagtgtacaataagacattcaatgggacaggaccatgcaataatgtcagcagtagcaatgt  
acacatgggaattaaagcagtggtatcaactcagctattgtttaaaccggtagcctagcagaaggagagataataattagatc

tgaaaatctgacagacaatgtcaaaacaataatagtacatcttgatcaatctgtagaaattgtgtgtacaagacccaaca  
ataatacaagaaaaagtataaggatagggccaggacaaacattctatgcaacaggaggcataatagggacatacgacaa  
gcacattgtaacattagtgaagacaaatggaaatgaaactttacaaaggggtgggtaaaaaattagtagaacacttccctaa  
taagacaataaaaatttgcaccatcctcaggaggggacctagaaattacaacacatagctttaattgtagaggagaatttt  
tctattgcagcacatcaagactgtttaatagtacatacatgcctaattgatacaaaaaagtaagtcaaaaaaacatcaca  
atcccatgcagcataaaaaaaatgttaaacatgttggcaggaggtaggacgagcaatgtatgcccctccattgaaggaaa  
cataacctgttagatcaaatatcacaggaatactatttggtagctgtagggagtagattcagaagatccagaaaaataa  
agacagagacattccgacctggaggaggagatatgaggaaacaattggagaagtgaattatataaaataaaagcggcagaa  
attaagccattgggagtagcaccactccagcaaaaaggagagtggtggagagagaaaaagagcagtaggattaggagc  
tgtgttccttggattcttgggagcagcaggaagcactatgggagcagcgtcaataacgctgacggtacaggccagacaa  
tgttgtctggtatagtgaacagcaagcaatttgcctgagggtatcgaggcgcaacagcatctgttgcaactcacggtc  
tggggcattaagcagctccagacaagagtctggtctatcgaaagatacctaaaggatcaacagctccttagggctttggg  
ctgctctggaaaaactcatctgcaccactaatgtaccttggaaactccagtggagtaacaaatctcaaacagatatttggg  
aaaacatgacctggatgcagtgggataaagaagtttagtaattcacacagacacaatatagaggttgcttgaagactcgcaa  
acccagcaggaagaaatgaaaaggatttattagcattggacaattggaaaaatctgtggaattgggttagtataacaaa  
ctggctgtggtatataaaaaatattcataatgatagtaggaggttgtaggcttaagaataatttttgcctgtgctttcta  
tagtgaatagagttaggcagggaatactcacctttgtcgtttcagacccttaccctaaacccaaggggacccgacaggctc  
ggaagaatcgaagaagaagggtggagggcaagacagagacagatcgattcgatttagtaacggattcttagcacttgctg  
ggacgacctgtggagcctgtgcctctcagctaccaccgatttgagagacttaataattggtgacagcagagcgggtggaac  
ttctgggacacagcagctctcaggggactacagaggggtgggaagcccttaagtatctgggaggtattgtgcagtttgg  
ggtctggaactaaaaaagagggtctattagctgcttgatactgtagcaatagcagtagctgaaggcacagataggattat  
agaattcctccaaagaattttagtagctatccgcaacatacctagaaggataagacagggctttgaagcagctttgcagt  
aaaatggctagccccgggtgataaacggaccgagcaatccctaggctgtgccttctagtgtccagccatctgtgtttgc  
ccctccccgtgcttcttgaccttgaagggtgcccactcccactgtcctttcctaaataaatgaggaattgcactcgca  
ttgtctgagtaggtgtcatttcttctgggggttgggggtggggcaggacagcaagggggaggattgggaagacaaatgca  
ggcatgctggggatgagggtgggtctatataaaaaacgcccggcggaaccgagcgttctgaacgctagagtcgacaaat  
tcagaagaactcgtcaagaaggcgatagaaggcgatgcgctgcgaatcgggagcggcgataccgtaagcacgaggaagc  
ggtagcccatctgcgcgcaagctcttcagcaatatcagggtagccaacgctatgtcctgatagcggctctgccacaccc  
agccggccacagtcgatgaatccagaaaagcggccattttccaccatgatattcggcaagcaggcatcgccatgggtcac  
gacgagatcctgcgctcgggcatgctgccttgagcctggcggaacagttcggctggcgagcctctgatgctcttcgt  
ccagatcatcctgatcgacaagacgggttccatccgagtagctgctcgtcgtatgcttgccttgggtgggtcgaaat  
gggaggttagccgatcaagcgtatgcagccgcccatttgcacagccatgatggatactttctcggcaggagcaagggtg  
agatgacaggagatcctgccccggcacttcgcccataagcagccagtccttcccgttcagtgacaacgtcgagcacag  
ctgcgcaaggaaacgcccgtcgtggccagccacgatagccgcgctgcctcgtcttgagttcattcagggcaccggacagg  
tcggctcttgacaaaaaagaacggggcgccccctgcgctgacagccggaacacggcgccatcagagcagccgatgtctgtg  
tgcccagtcataagccgaatagcctctccacccaagcggcggaagacctgcgtgcaatccatctgttcaatcatcgcaa  
acgatcctcatcctgtctcttgatcagatcttgatccctgcgcctcagatccttgccggcaagaaagccatccagttt  
actttgcagggttcccaaccttaccagagggcgccccagctggcaattccgggttcgcttgctgtccataaaacgccc  
gtctagctatcgccatgtaagccactgcaagctacctgctttctctttgcgcttgcttttcccttgctccagatagccc  
agtagctgacattcatccggggtcagcacccgtttctgcggactggctttctacgtgaaaaggatctaggtgaagatcctt  
tttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaagg  
atcctcttgagatcctttttctgcgctaatctgctgcttgcaaaaaaaaccaccgctaccagcgggtgggtttgtt  
tgccggatcaagagctaccaactctttccgaaggtaactggcttcagcagagcgagataccaaatctgttcttcta  
gtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactcgtgtaccagt  
ggctgctgccagtgccgataagtcgtgtcttaccgggttggactcaagacgatatgtaccggataaggcgagcgggtcgg  
gctgaacgggggggttcgtgcacacagcccagcttgagagcgaacgacctacaccgaactgagatacctacagcgtgagcta  
tgagaaagcgccacgcttcccgaagggaagggcgagcaggtatccggtaagcggcagggtcggaacaggagagcgcac  
gaggagcttccagggggaaacgcctggatctttatagtcctgtcgggtttcgccaccttgacttgagcgtcgatctt  
tgtgtgctcgtcagggggggcggagcctatggaaaaaacgcagcaacgcggccctttacgggttccctggccttttgcgtg  
ccttttgcctcacatgttgtcgacaatatggctattggccattgcatacgttgtatctatatcataatatgtacatttat  
attggctcatgtccaatatgaccgcatgttgacattgattattgactagttattaatagtaataattacgggttcatt  
agttcatagcccatatatggagttccgcgttacataacttacggtaaatggccgcctggctgaccgcccacgaccccc  
gcccattgacgtcaataatgacgtatgttcccatagtaaacgccaatagggactttccattgacgtcaatgggtggagtat  
ttacggtaaaactgcccacttggcagtagatcaagtgatcatatgccaaagtccgccccctattgacgtcaatgacggtaa  
atggcccgccctggcatatgcccagtagacaccttacgggactttcctacttggcagtagacatctacgggtattagtcac  
ggctattaccattgggtgatgcgggtttggcagtagaccaaaggcggtgagcgggttgactcacggggatttccaagtc  
tccacccattgacgtcaatgggagttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaataaccccgcc  
ccgttgacgcaaatggcggttaggcgtgtacgggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgc  
ctggagacgcatccagcgtgtttgacctccatagaagacacgggacggatccagcctccgcgccgggaacgggtgca  
ttggaacgggattccccgtgccaagagtgcgtaagtacggcctatagactctataggcacacccctttggctcttatg  
catgctatactgttttggcttggggcctataccccccgcttcttatgctatagggtgatgggtatagcttagcctatag  
gtgtgggttattgaccattattgaccactccccctattggtagcagtaacttccattactaataacataacatggctcttg  
ccacaactatctctattggctatatgccaaactctgtccttcagagactgacacggactctgtatttttacaggatggg

Fig. 18a

gtcccatTTATTATTtACAAATtCACATATACAACAACgCGTCCCCCGTgCCCGcAGTttttATTAAACATAgCGTggg  
atctccacgcgaatctcgggtaccgtgttccggacatgggytcttctccggtagcggcggagcttccacatccgagccct  
ggTcccatgcctccagcggctcatggtcgctcggcagctccttgctcctaacagtggaggccagacttaggcacagcaca  
atgccccaceaceagtgTgcgcacaaggccgtggcggtagggtagtggtctgaaaatgagctcggagattgggctcg  
caccgctgacgcagatggaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgattctgataagagt  
cagaggtaactcccgTTcggtgctgttaacggTggagggcagtgtagtctgagcagTactcgttgctgccgcgcgcgcc  
accagacataatagctgacagactaacagactgttcctttccatgggtcttttctgcagtcaccat

Fig. 18a

Figure 18b pGA1/IN2 Functional Regions

| Position within sequence<br>(starting-and ending pos.) | Gene/ORF (indicate<br>complete name of the gene<br>and the expressed gene<br>product)      | Origin of sequence<br>(organism)       |
|--------------------------------------------------------|--------------------------------------------------------------------------------------------|----------------------------------------|
| 99-1577                                                | Gag                                                                                        | C HIV-1 isolate 98IN012                |
| 1382-3520                                              | Pol, inactivating point<br>mutations in reverse<br>transcriptase, deletion of<br>integrase | C HIV-1 isolate 98IN012                |
| 3645-3859 and 6205-6295                                | Tat exons #1 and #2                                                                        | C HIV-1 isolate 98IN012                |
| 3784-3859 and 6205-6452                                | Rev exons #1 and #2                                                                        | C HIV-1 isolate 98IN012                |
| 3873-4121                                              | Vpu                                                                                        | C HIV-1 isolate 98IN012                |
| 4039-6642                                              | Env                                                                                        | C HIV-1 isolate 98IN012                |
| 6643-10466                                             | Plasmid vector                                                                             | pGA1, Gen Bank accession #<br>AF425297 |

Figure 18c. IN2 Mutation Chart

| Codon change                 | Amino acid change | Region                   | Function                                        |
|------------------------------|-------------------|--------------------------|-------------------------------------------------|
| to TCC position<br>1260-1262 | C390S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1269-1271 | C393S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1323-1325 | C411S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| to TCT position<br>1332-1334 | C414S             | Gag                      | Ablation of zinc<br>finger used in<br>packaging |
| To AAT position<br>2387-2389 | D185N             | Reverse<br>Transcriptase | Inactivation of<br>Polymerase active<br>site    |
| To ACG position<br>2630-2632 | W266T             | Reverse<br>Transcriptase | Ablation of strand<br>transfer activity         |
| to CAA position<br>3266-3268 | E478Q             | Rnase H                  | Inactivation of<br>RNase H activity             |

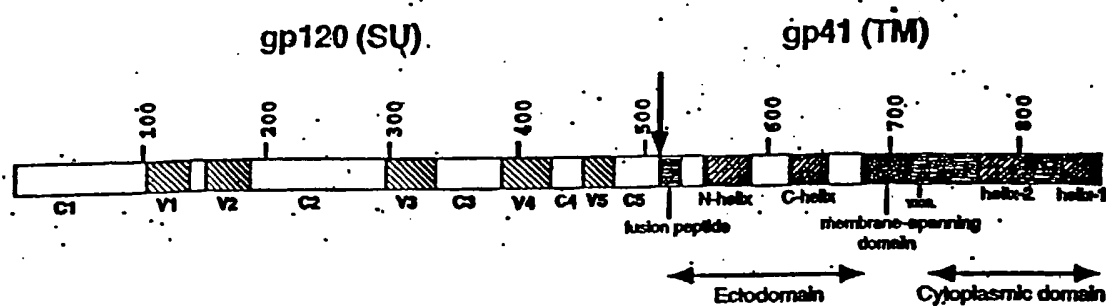


Fig. 19

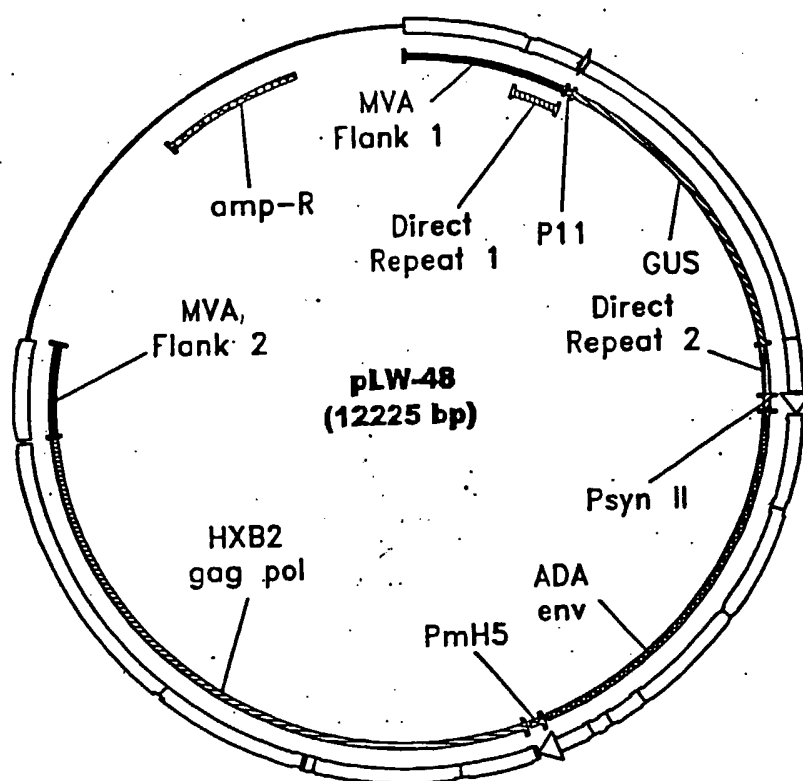


Fig. 20A



1 GAATTCGTTG GTGGTCGCCA TGGATGGTGT TATTGTATAC TGTCTAAACG CGTTAGTAAA ACATGGCGAG  
CTTAAGCAAC CAGCAGCGGT ACCTACCCACA ATAACATATG ACAGATTTCG GCAATCATTT TGTACCGGCTC

71 GAAATAAATC ATATAAAAA TGATTTCATG ATTAAACCAT GTTGTGAAAA AGTCAAGAAC GTTCACATTG  
CTTTATTAG TATATTTTTT ACTAAAGTAC TAATTGGTA CAACACTTTT TCAGTTCTTG CAAGTGTAAC

141 GCGGACAATC TAAACAAT ACAGTGATTG CAGATTGGC ATATATGGAT AATGCGGTAT CCGATGTATG  
CGCCTGTTAG ATTTTGTGA TGCTACTAAC GTCTAAACGG TATATACCTA TTACGCCATA GGCTACATAC

211 CAATTCAC TG TATAAAGA ATGTATCAAG AATATCCAGA TTTGCTAATT TGATAAAGAT AGATGACGAT  
GTTAAGTGAC ATATTTTCT TACATAGTTC TTATAGGTCT AAACGATTAA ACTATTTCTA TCTACTGCTA

281 GACAAGACTC CTA CTGGTGT ATATAATTAT TTTAAACCTA AAGATGCCAT TCCTGTTAAT ATATCCATAG  
CTGTTCTGAG GATGACCACA TATATTAATA AAATTGGAT TTCTACGGTA AGGACAATAA TATAGGTATC

351 GAAAGGATAG AGATGTTTGT GAACTATTAA TCTCATCTGA TAAAGCGTGT CCGTGTATAG AGTTAAATTC  
CTTTCCTATC TCTACAAACA CTTGATAATT AGAGTAGACT ATTTCGCACA CGCACATATC TCAATTTAAG

Fig. 20B

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

**Entire pLW-48 plasmid sequence:**

GAATTCGTTGGTGGTCCCATGGATGGTGTTATTGTATACTGTCTAAACGCG  
T TAGTAAACATGGCGAGGAAATAAATCATATAAAAAATGATTTTCATGATTAA  
ACCATGTTGTGAAAAAGTCAAGAACGTTACATTGGCGGACAATCTAAAAAC  
AATACAGTGATTGCAGATTTGCCATATATGGATAATGCGGTATCCGATGTAT  
GCAATTCAGTGTATAAAAAGAATGTATCAAGAATATCCAGATTTGCTAATTTG  
ATAAAGATAGATGACGATGACAAGACTCCTACTGGTGTATATAATTATTTTAA  
ACCTAAAGATGCCATTCTGTTATTATATCCATAGGAAAGGATAGAGATGTTT  
GTGAACTATTAATCTCATCTGATAAAGCGTGTGCGTGTATAGAGTTAAATTCA  
TATAAAGTAGCCATTCTTCCCATGGATGTTTCCTTTTTTACCAAAGGAAATGC  
ATCATTGATTATTCTCCTGTTTGATTTCTCTATCGATGCGGCACCTCTCTTAA  
GAAGTGTAACCGATAATAATGTTATTATATCTAGACACCAGCGTCTACATGA  
CGAGCTTCCGAGTTCCAATTGGTTCAAGTTTTACATAAGTATAAAGTCCGAC  
TATTGTTCTATATTATATATGGTTGTTGATGGATCTGTGATGCATGCAATAGC  
TGATAATAGAACTTACGCAAATATTAGCAAAAATATATTAGACAATACTACAA  
TTAACGATGAGTGTAGATGCTGTTATTTTGAACCACAGATTAGGATTCTTGAT  
AGAGATGAGATGCTCAATGGATCATCGTGTGATATGAACAGACATTGTATTA  
TGATGAATTTACCTGATGTAGGCGAATTTGGATCTAGTATGTTGGGGAAATA  
TGAACCTGACATGATTAAGATTGCTCTTTCCGGTGGCTGGGTACCAGGCGCG  
CCTTTTCATTTTGTTTTTTCTATGCTATAAATGGTACGTCCTGTAGAAACCCC  
AACCCGTGAAATCAAAAACTCGACGGCCTGTGGGCATTCAAGTCTGGATCG  
CGAAAACGTGGAATTGATCAGCGTTGGTGGGAAAGCGCGTTACAAGAAAG  
CCGGGCAATTGCTGTGCCAGGCAGTTTAAACGATCAGTTCGCCGATGCAGA  
TATTCGTAATTATGCGGGCAACGTCTGGTATCAGCGCGAAGTCTTTATACCG  
AAAGGTTGGGCAGGCCAGCGTATCGTGCTGCGTTTCGATGCGGTCATCAT  
TACGGCAAAGTGTGGGTCAATAATCAGGAAGTGATGGAGCATCAGGGCGG  
CTATACGCCATTGGAAGCCGATGTCACGCCGTATGTTATTGCCGGGAAAG  
TGTACGTATCACCGTTTGTGTGAACAACGAACCTGAACTGGCAGACTATCCC  
GCCGGGAATGGTGATTACCGACGAAACGGCAAGAAAAAGCAGTCTTACTT  
CCATGATTTCTTTAACTATGCCGGAATCCATCGCAGCGTAATGCTCTACACC  
ACGCCGAACACCTGGGTGGACGATATCACCGTGGTGACGCATGTCGCGCA  
AGACTGTAACCACGCGTCTGTTGACTGGCAGGTGGTGGCCAATGGTGATGT  
CAGCGTTGAACTGCGTGATGCGGATCAACAGGTGGTTGCAACTGGACAAG  
GCACTAGCGGGACTTTGCAAGTGGTGAATCCGCACCTCTGGCAACCGGGT  
GAAGGTTATCTCTATGAACTGTGCGTCACAGCCAAAAGCCAGACAGAGTGT  
GATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGTGAAGGGCGA  
ACAGTTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCAT  
GAAGATGCGGACTTGCGTGGCAAAGGATTGATAACGTGCTGATGGTGCAC  
GACCACGCATTAATGGACTGGATTGGGGCCAACTCCTACCGTACCTCGCAT  
TACCCTTACGCTGAAGAGATGCTCGACTGGGCAGATGAACATGGCATCGTG

GTGATTGATGAACTGCTGCTGTCTGGCTTTAACCTCTCTTTAGGCATTGGTT  
TCGAAGCGGGCAACAAGCCGAAAGAACTGTACAGCGAAGAGGCAGTCAAC  
GGGGAACTCAGCAAGCGCACTTACAGGCGATTAAAGAGCTGATAGCGCGT  
GACAAAAACCAACCAAGCGTGGTGATGTGGAGTATTGCCAACGAACCGGAT  
ACCCGTCCGCAAGGTGCACGGGAATATTTCCGCGCCACTGGCGGAAGCAAC  
GCGTAACTCGACCCGACGCGTCCGATCACCTGCGTCAATGTAATGTTCTG  
CGACGCTCACACCGATACCATCAGCGATCTCTTTGATGTGCTGTGCCTGAA  
CCGTTATTACGGATGGTATGTCCAAAGCGGCGATTTGGAAACGGGCAGAGAA  
GGTACTGGAAAAAGAACTTCTGGCCTGGCAGGAGAACTGCATCAGCCGAT  
TATCATCACCGAATACGGCGTGGATACGTTAGCCGGGCTGCACTCAATGTA  
CACCGACATGTGGAGTGAAGAGTATCAGTGTGCATGGCTGGATATGTATCA  
CCGCGTCTTTGATCGCGTCAGCGCCGTCGTCGGTGAACAGGTATGGAATTT  
CGCCGATTTTGCACCTCGCAAGGCATATTGCGCGTTGGCGGTAACAAGAA  
AGGGATCTTCACTCGCGACCGCAAACCGAAGTCGGCGGGCTTTTCTGCTGCA  
AAAACGCTGGACTGGCATGAACTTCGGTGAAAAACCGCAGCAGGGAGGCA  
AACAAATGAGAGCTCGGTTGTTGATGGATCTGTGATGCATGCAATAGCTGATA  
ATAGAAGTTACGCAAATATTAGCAAAAATATATTAGACAATACTACAATTAAC  
GATGAGTGTAGATGCTGTTATTTGAACCACAGATTAGGATTCTTGATAGAG  
ATGAGATGCTCAATGGATCATCGTGTGATATGAACAGACATTGTATTATGAT  
GAATTTACCTGATGTAGGCGAATTTGGATCTAGTATGTTGGGGAAATATGAA  
CCTGACATGATTAAGATTGCTCTTTCCGTGGCTGGCGGCCCGCTCGAGTAA  
AAAATGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTCTAT  
GCTATAAATAATAAATAGCGGCCGCAACCATGAAAGTGAAGGGGATCAGGAA  
GAATTATCAGCACTTGTGGAAATGGGGCATCATGCTCCTTGGGATGTTGATG  
ATCTGTAGTGCTGTAGAAAATTTGTGGGTCACAGTTTATTATGGGGTACCTG  
TGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATA  
TGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGA  
CCCCAACCCACAAGAAGTAGTATTGGAAAATGTGACAGAAAATTTTAACATG  
TGGAAAATAACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGG  
ATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTACTTTAAAT  
TGCCTGATTTGAGGAATGTTACTAATATCAATAATAGTAGTGAGGGAATGA  
GAGGAGAAATAAAAACTGCTCTTTCAATATCACCACAAGCATAAGAGATAA  
GGTGAAGAAAGACTATGCACTTTTCTATAGACTTGATGTAGTACCAATAGATA  
ATGATAATACTAGCTATAGGTTGATAAATTGTAATACCTCAACCATTACACAG  
GCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTACCCCGG  
CTGGTTTTGCGATTCTAAAGTGTAAGACAAGAAGTTCAATGGAACAGGGCC  
ATGTAAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTG  
TCAACTCACTGCTGTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTA  
GATCTAGTAATTTACAGACAATGCAAAAAACATAATAGTACAGTTGAAAGAA  
TCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATAC  
ATATAGGACCAGGAAGAGCATTATATACAACAGGAGAAATAATAGGAGATAT  
AAGACAAGCACATTGCAACATTAGTAGAACAAAATGGAATAACACTTTAAAT  
CAAATAGCTACAAAATTAAGAACAATTTGGGAATAATAAACAAATAGTCTT  
TAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGT  
GGAGGGGAATTTCTTCTACTGTAATTCAACACAACCTGTTTAATAGTACTTGA  
ATTTAATGGTACTTGAATTTAACACAATCGAATGGTACTGAAGGAAATGA

Fig. 21

CACTATCACACTCCCATGTAGAATAAAACAAATTATAAATATGTGGCAGGAA  
GTAGGAAAAGCAATGTATGCCCCTCCCATCAGAGGACAAATTAGATGCTCAT  
CAAATATTACAGGGCTAATATTAACAAGAGATGGTGGAACATAACAGTAGTGG  
GTCCGAGATCTTCAGACCTGGGGGAGGAGATATGAGGGACAATTGGAGAA  
GTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCC  
ACCAAGGCCAAAAAGAAGAGTGGTGACAGAGAGAAAAAAGAGCAGTGGGAAC  
GATAGGAGCTATGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGG  
CGCAGCGTCAATAACGCTGACGGTACAGGCCAGACTATTATTGTCTGGTAT  
AGTGCAACAGCAGACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCT  
GTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGG  
CTGTGGAAAGATACCTAAGGGATCAACAGCTCCTAGGGATTGGGGTTGCT  
CTGGAACACTCATCTGCACCACTGCTGTGCCCTTGAATGCTAGTTGGAGTA  
ATAAACTCTGGATATGATTTGGGATAACATGACCTGGATGGAGTGGGAAA  
GAGAAATCGAAAATTACACAGGCTTAATATACACCTTAATTGAGGAATCGCA  
GAACCAACAAGAAAAGAATGAACAAGACTTATTAGCATTAGATAAGTGGGCA  
AGTTTGTGGAATTGGTTTGACATATCAAATTGGCTGTGGTATGTAAAAATCTT  
CATAATGATAGTAGGAGGCTTGATAGGTTTAAGAATAGTTTTACTGTACTTT  
CTATAGTAAATAGAGTTAGGCAGGGATACTCACCATTGTCATTTTCAGACCCA  
CCTCCAGCCCCGAGGGGACCCGACAGGCCCGAAGGAATCGAAGAAGAAG  
GTGGAGACAGAGACTAATTTTTATGCGGCCGCTGGTACCCAACCTAAAAATT  
GAAAATAAATACAAAGGTTCTTGAGGGTTGTGTTAAATTGAAAGCGAGAAAT  
AATCATAAATAAGCCCGGGGATCCTCTAGAGTCGACACCATGGGTGCGAGA  
GCGTCAGTATTAAGCGGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTA  
AGGCCAGGGGGGAAAGAAAAAATATAAATTAACATATAGTATGGGCAAGCA  
GGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAG  
GCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAG  
AAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAA  
AGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAG  
CAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGACAC  
AGCAATCAGGTGAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGGCAA  
ATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAG  
TAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTTCAGCATTATC  
AGAAGGAGCCACCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGG  
ACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGC  
AGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCA  
GATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACCTACTAGTACCCTTCA  
GGAACAAATAGGATGGATGACAAATAATCCACCTATCCAGTAGGAGAAATT  
TATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCC  
TACCAGCATTCTGGACATAAGACAAGGACCAAAAGAACCCTTTAGAGACTAT  
GTAGACCGGTTCTATAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTA  
AAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCAGATTGTA  
AGACTATTTTAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGAC  
AGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTG  
AAGCAATGAGCCAAGTAACAAATTCAGCTACCATAATGATGCAGAGAGGCA  
ATTTTAGGAACCAAGAAAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGG  
GCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAT

Fig. 21  
4

GTGGAAAGGAAGGACACCAAATGAAAGATTGTACTGAGAGACAGGCTAATT  
TTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCA  
GAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGG  
TAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTGT  
ATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCAACAATA  
AAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGAT  
ACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAG  
GGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGA  
AATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTC  
AACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCC  
CATTAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGC  
CCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAG  
AAATTTGTACAGAAATGGAAAAGGAAGGGGAAAATTTCAAAAATTGGGCCTGA  
GAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAAT  
GGAGGAAATTAGTAGATTTTCAAGAACTTAATAAGAGAACTCAAGACTTCTG  
GGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCA  
GTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAG  
ACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAACAATGAGACACC  
AGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACC  
AGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAAAAAACAAA  
ATCCAGACATAGTTATCTATCAATACATGAACGATTTGTATGTAGGATCTGAC  
TTAGAAATAGGGCAGCATAGAACAATAATAGAGGAGCTGAGACAACATCTG  
TTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCA  
TTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTA  
TAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAG  
TGGGGAAATTGAATACCGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGC  
AATTATGTAACTCCTTAGAGGAACCAAGCACTAACAGAAGTAATACCACT  
AACAGAAGAAGCAGAGCTAGAAGTGGCAGAAAACAGAGAGATTCTAAAAGA  
ACCAGTACATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATA  
CAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTT  
AAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCACACTAAT  
GATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAG  
TAATATGGGGAAAGACTCCTAAATTTAACTACCCATACAAAAGGAAACATG  
GGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGA  
GTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTACCAGTTAGAGAAAGAA  
CCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAG  
ACTAAATTAGGAAAAGCAGGATATGTTACTAACAAAGGAAGACAAAAGGTTG  
TCCCCCTAACTAACACAACAAATCAGAAAACCTCAGTTACAAGCAATTTATCTA  
GCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACAATATG  
CATTAGGAATCATTCAAGCACAACCAGATAAAAGTGAATCAGAGTTAGTCAA  
TCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTA  
CCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGT  
GCTGGAATCAGGAAAATACTATTTTTAGATGGAATAGATAAGGCCCAAGATG  
AACATTAGTTTTTATGTGACCTGCAGGGAAAGTTTTATAGGTAGTTGATAG  
AACAAAATACATAATTTTGTAAAAATAAATCACTTTTTATACTAATATGACACG  
ATTACCAATACTTTTGTACTAATATCATTAGTATACGCTACACCTTTTCCTCA

Fig. 21

GACATCTAAAAAATAGGTGATGATGCAACTTTATCATGTAATCGAAATAATA  
CAAATGACTACGTTGTTATGAGTGCTTGGTATAAGGAGCCCAATTCCATTAT  
TCTTTTAGCTGCTAAAAGCGACGTCTTGTATTTTGATAATTATACCAAGGATA  
AAATATCTTACGACTCTCCATACGATGATCTAGTTACAACCTATCACAATTAAA  
TCATTGACTGCTAGAGATGCCGGTACTTATGTATGTGCATTCTTTATGACATC  
GCCTACAAATGACACTGATAAAGTAGATTATGAAGAATACTCCACAGAGTTG  
ATTGTAAATACAGATAGTGAATCGACTATAGACATAATACTATCTGGATCTAC  
ACATTCACCAGAACTAGTTAAGCTTGTCTCCCTATAGTGAGTCGTATTAGA  
GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCT  
CACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG  
TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCT  
TTCGAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC  
GCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCGCTTCCTCGCTCAC  
TGAATCGCTGCGCTCGGTCTGCTCGGCTGCGGCGAGCGGTATCAGCTCACT  
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA  
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG  
TTGCTGGCGTTTTTCGATAGGCTCCGCCCCCTGACGAGCATCAGAAAAAT  
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG  
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG  
CTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCTTTCT  
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG  
CTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGCTGCGCCTTATC  
CGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT  
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG  
CTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAG  
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG  
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT  
TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG  
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGG  
ATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA  
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACA  
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTG  
TTCATCCATAGTTGCCTGACTCCCCGTCTGTGTAGATAACTACGATACGGGAG  
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA  
CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG  
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGC  
CGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCACAACGTTGTT  
GGCATTGCTACAGGCATCGTGGTGTCAAGGTTGGTGGTTTGGTATGGCTTCA  
TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT  
GCAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGT  
TGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC  
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG  
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCA  
ATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTG  
GAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGAT  
CCAGTTCGATGTAACCCACTCGTGCAACCAACTGATCTTCAGCATCTTTTAC

Fig. 21

TTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA  
AAAGGGAATAAGGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT  
CAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATT  
TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGCGGCACATTTCCCCGA  
AAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA  
AAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGG  
TGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA  
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTT  
GGCGGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACT  
GAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA  
ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGG  
GCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGAT  
GTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGAC  
GTTGTAAAACGACGGCCAGTGAATTGGATTAGGTGACACTATA

**New Psyn II Promoter which controls ADA envelope expression:**

TAAAAAATGAAAAAATATTCTAATTTATAGGACGGTTTTGATTTTCTTTTTTC  
TATGCTATAAATAATAATA

**ADA envelope truncated:**

ATGAAAGTGAAGGGGATCAGGAAGAATTATCAGCACTTGTGGAAATGGGGC  
ATCATGCTCCTTGGGATGTTGATGATCTGTAGTGCTGTAGAAAATTTGTGGG  
TCACAGTTTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTATT  
TTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCC  
ACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGAA  
AATGTGACAGAAAATTTAACATGTGGAATAAACATGGTAGAACAGATGC  
ATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAATTT  
AACCCCACTCTGTGTTACTTTAAATTGCACTGATTTGAGGAATGTTACTAATA  
TCAATAATAGTAGTGAGGGAATGAGAGGAGAAATAAAAACTGCTCTTTCAA  
TATCACCACAAGCATAAGAGATAAGGTGAAGAAAGACTATGCACTTTTCTAT  
AGACTTGATGTAGTACCAATAGATAATGATAATACTAGCTATAGGTTGATAAA  
TTGTAATACCTCAACCATTACACAGGCCCTGTCCAAAGGTATCCTTTGAGCCA  
ATTCCCATACATTATTGTACCCCGGCTGGTTTTGCGATTCTAAAGTGTAAG  
ACAAGAAGTTCATGGAACAGGGCCATGTAAAAATGTCAGCACAGTACAAT  
GTACACATGGAATTAGGCCAGTAGTGTCAACTCAACTGCTGTTAAATGGCAG  
TCTAGCAGAAGAAGAGGTAGTAATTAGATCTAGTAATTTACAGACAATGCA  
AAAAACATAATAGTACAGTTGAAAGAATCTGTAGAAATTAATTGTACAAGACC  
CAACAACAATACAAGGAAAAGTATACATATAGGACCAGGAAGAGCATTTTAT  
ACAACAGGAGAAATAATAGGAGATATAAGACAAGCACATTGCAACATTAGTA  
GAACAAAATGGAATAACACTTTAAATCAAATAGCTACAAAATTAAGAACA  
TTTGGGAATAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAG  
AAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTCTTCTACTGTAATTCA  
ACACAACCTGTTTAATAGTACTTGGAATTTAATGGTACTTGGAATTTAACACA

Fig. 21

ATCGAATGGTACTGAAGGAAATGACACTATCACACTCCCATGTAGAATAAAA  
CAAATTATAAATATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCA  
TCAGAGGACAAATTAGATGCTCATCAAATATTACAGGGCTAATATTAACAAG  
AGATGGTGGAACTAACAGTAGTGGGTCCGAGATCTTCAGACCTGGGGGAG  
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAA  
AATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAAAGAAGAGTGGTGCA  
GAGAGAAAAAAGAGCAGTGGGAACGATAGGAGCTATGTTCTTGGGTTCTT  
GGGAGCAGCAGGAAGCAGTATGGGCGCAGCGTCAATAACGCTGACGGTAC  
AGGCCAGACTATTATTGTCTGGTATAGTGCAACAGCAGAACAAATTTGCTGAG  
GGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAA  
GCAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGATACCTAAGGGATCAACA  
GCTCCTAGGGATTGTTGGGGTTGCTCTGGAAAACCTCATCTGCACCACTGCTGT  
GCCTTGAATGCTAGTTGGAGTAATAAACTCTGGATATGATTTGGGATAAC  
ATGACCTGGATGGAGTGGGAAAGAGAAATCGAAAATTACACAGGCTTAATAT  
ACACCTTAATTGAGGAATCGCAGAACCAACAAGAAAAGAATGAACAAGACTT  
ATTAGCATTAGATAAGTGGGCAAGTTTGTGGAATTGGTTTGACATATCAAATT  
GGCTGTGGTATGTAAAAATCTTCATAATGATAGTAGGAGGCTTGATAGGTTT  
AAGAATAGTTTTTACTGTACTTTCTATAGTAAATAGAGTTAGGCAGGGATACT  
CACCATTGTCATTTACAGACCCACCTCCCAGCCCCGAGGGGACCCGACAGG  
CCCGAAGGAATCGAAGAAGAAGGTGGAGACAGAGAC

PmH5 promoter (which controls HXB2 gag pol expression):

AAAAATTGAAAATAAATACAAAGGTTCTTGAGGGTTGTGTTAAATTGAAAGC  
GAGAAATAATCATAAATA

HXB2 gag pol (with safety mutations,  $\Delta$  integrase):

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGA  
AAAAATTCGGTTAAGGCCAGGGGGGAAAGAAAAAATATAAATTAAACATATA  
GTATGGGCAAGCAGGGAGCTAGAACGATTGCGAGTTAATCCTGGCCTGTTA  
GAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTT  
CAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCT  
ATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAA  
GATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGC  
TGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAA  
CATCCAGGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCA  
TGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGGCCCAGAAGTGATACCCATG  
TTTTAGCATTATCAGAAGGAGCCACCCACAAAGATTTAAACACCATGCTAA  
ACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCA  
ATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCT  
ATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAAC  
TACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATC  
CCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAAATAG  
TAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAAGA  
ACCCTTTAGAGACTATGTAGACCGGTTCTATAAACTCTAAGAGCCGAGCAA

Fig. 21



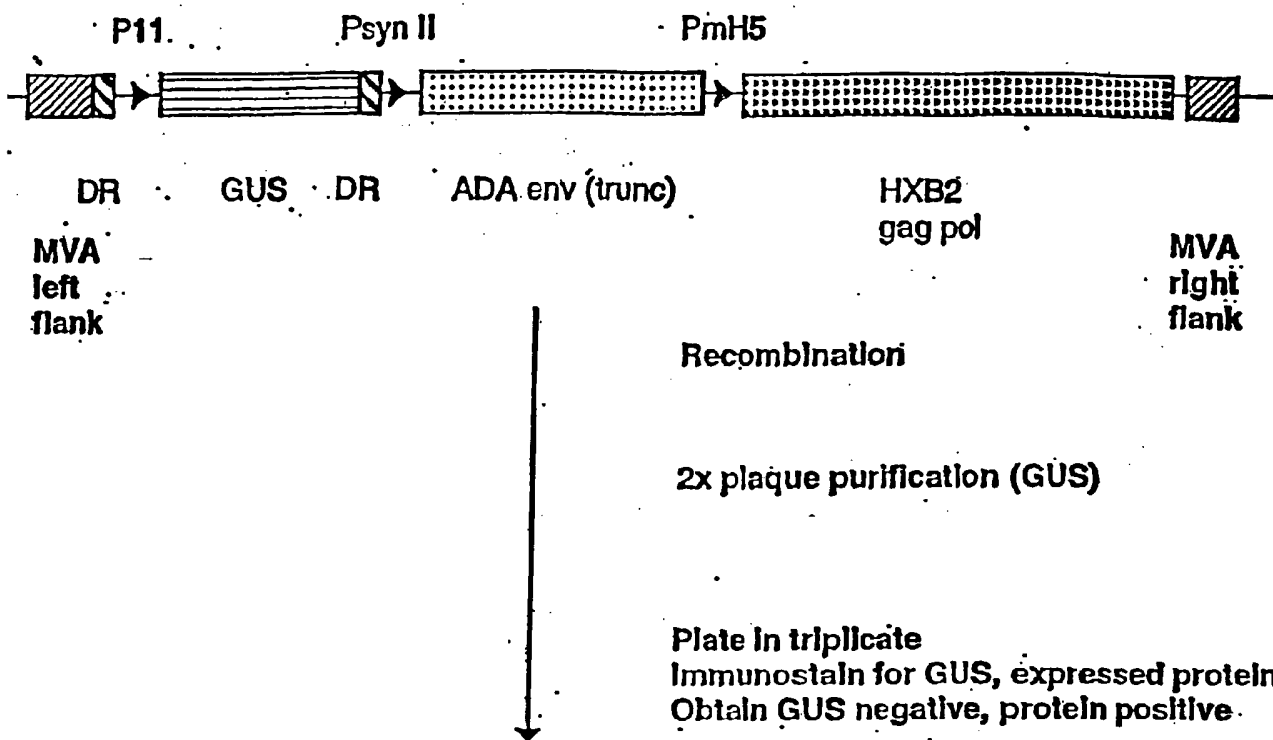
GCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATG  
CGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACT  
AGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGG  
CAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACCATAAT  
GATGCAGAGAGGCAATTTTAGGAACCAAAGAAAGATTGTTAAGTGTTCAT  
TGTGGCAAAGAAGGGGCACACAGCCAGAAATTGCAGGGCCCCCTAGGAAAAA  
GGGCTGTTGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACTGA  
GAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCC  
AGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAG  
CTTCAGGTCTGGGGTAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGAT  
AGACAAGGAACTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGAC  
CCCTCGTCACAATAAAGATAGGGGGGGCAACTAAAGGAAGCTCTATTAGATA  
CAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGA  
AACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGA  
TCAGATACTCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTA  
GGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT  
GCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAATTAAG  
CCAGGAATGGATGGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAA  
ATAAAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGGAAAATTT  
CAAAAATTGGGCCTGAGAATCCATACAATACTCCAGTATTTGCCATAAAGAA  
AAAAGACAGTACTAAATGGAGGAAATTAGTAGATTTTCAGAGAACTTAATAAG  
AGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGG  
TTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTC  
AGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGT  
ATAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGG  
GATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGA  
GCCTTTTAAAAAACAAAATCCAGACATAGTTATCTATCAATACATGAACGATT  
TGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGA  
GCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACA  
TCAGAAAGAACCTCCATTCTTTGGATGGGTTATGAACTCCATCCTGATAAA  
TGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAAT  
GACATACAGAAGTTAGTGGGGAAATTGAATACCGCAAGTCAGATTTACCCA  
GGGATTAAAGTAAGGCAATTATGTAACTCCTTAGAGGAACCAAAGCACTAA  
CAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAAACA  
GAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAGA  
CTTAATAGCAGAAATACAGAAGCAGGGGGCAAGGCCAATGGACATATCAAT  
TTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGG  
GGTGCCCACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAA  
CCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAACTACCCAT  
ACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTG  
GATTCCTGAGTGGGAGTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTAC  
CAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGG  
GCAGCTAACAGGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAACAAA  
GGAAGACAAAAGGTTGTCCCCCTAACTAACACAACAAATCAGAAAATCAGT  
TACAAGCAATTTATCTAGCTTTGCAGGATTGAGGATTAGAAGTAAACATAGTA  
ACAGACTACAATATGCATTAGGAATCATTCAAGCACAACCAGATAAAAGTG

Fig. 21

AATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGT  
CTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGT.  
AGATAAATTAGTCAGTGCTGGAATCAGGAAAATACTATTTTATAGATGGAATA  
GATAAGGCCCAAGATGAACATTAG

Fig. 21

## Plasmid pLW-48



## Virus MVA/HIV 48

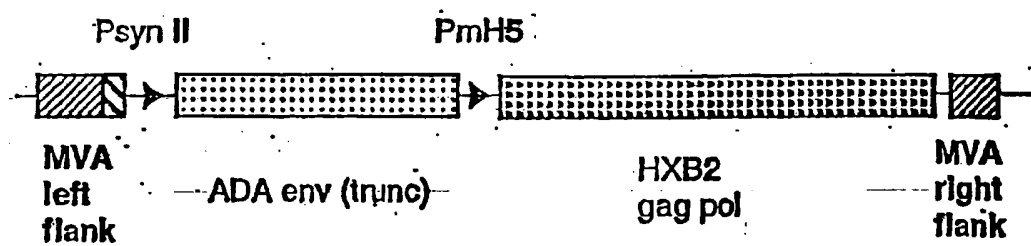


Fig. 22

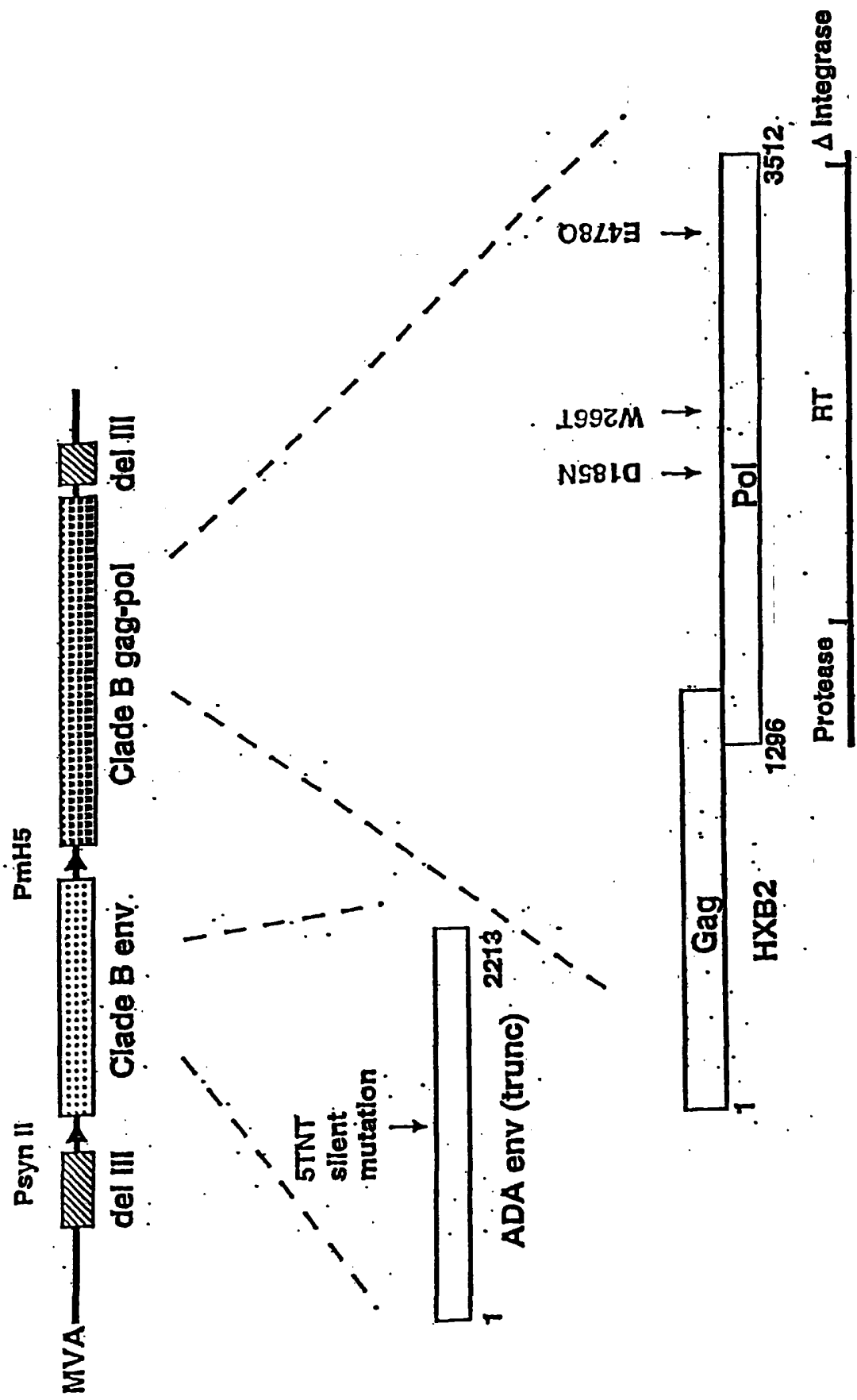


Fig. 28

## Sequence of new Psyn II promoter:

### Early part of promoter

Critical region

Early start site

TAAAAAATGAAAAAATATTCTAATTTATAGGACGGT

### Late part of promoter

TTTGATTTTCTTTTTTCTATGCTATAAATAATAAATA

Fig. 24