



(86) **Date de dépôt PCT/PCT Filing Date:** 2006/02/21
(87) **Date publication PCT/PCT Publication Date:** 2006/08/31
(45) **Date de délivrance/Issue Date:** 2016/01/19
(85) **Entrée phase nationale/National Entry:** 2007/08/20
(86) **N° demande PCT/PCT Application No.:** US 2006/005869
(87) **N° publication PCT/PCT Publication No.:** 2006/091511
(30) **Priorité/Priority:** 2005/02/21 (US60/654,484)

(51) **Cl.Int./Int.Cl. C12N 15/09** (2006.01),
C12N 7/00 (2006.01)

(72) **Inventeurs/Inventors:**
BURKE, DONALD, US;
SWITZER, WILLIAM, US;
HENEINE, WALID, US;
FOLKS, THOMAS, US;
WOLFE, NATHAN, US;
EITEL, MPOUDI, CM

(73) **Propriétaires/Owners:**

(54) **Titre : VIRUS T-LYMPHOTROPIQUES DE PRIMATE**
(54) **Title: PRIMATE T-LYMPHOTROPIC VIRUSES**

(57) **Abrégé/Abstract:**

Disclosed are compositions and methods related to the isolation and identification of the novel primate T-lymphotropic viruses, HTLV-3 and HTLV-4. The diversity of HTLVs was investigated among central Africans reporting contact with NHP blood and body fluids through hunting, butchering, and keeping primate pets. Herein it is shown that this population is infected with a variety of HTLVs, including two novel retroviruses; HTLV-4 is the first member of a novel phylogenetic lineage that is distinct from all known HTLVs and STLVs; HTLV-3 falls within the genetic diversity of STLV-3, a group that has not previously been seen in humans. The present invention also relates to vectors and vaccines for use in humans against infection and disease. The invention further relates to a variety of bioassays and kits for the detection and diagnosis of infection with and diseases caused by HTLV-3 and HTLV-4 and related viruses.



(73) **Propriétaires(suite)/Owners(continued):**

THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF THE
DEPARTMENT OF HEALTH AND HUMAN SERVICES, CENTERS FOR DISEASE CONTROL AND PREVENTION, US;
JOHNS HOPKINS UNIVERSITY, US

(74) **Agent:** SMART & BIGGAR

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 August 2006 (31.08.2006)

PCT

(10) International Publication Number
WO 2006/091511 A3

(51) International Patent Classification:

C12N 15/09 (2006.01) C12N 7/00 (2006.01)

WOLFE, Nathan [US/US]; 1101 St. Paul, #604, Baltimore, Maryland 21202 (US). **EITEL, Mpoudi** [CM/CM]; P.o. Box 906, Yaounde (CM).

(21) International Application Number:

PCT/US2006/005869

(74) **Agents:** **SPRATT, Gwendolyn** et al.; NEEDLE & ROSENBERG, P.C., Suite 1000, 999 Peachtree Street, Atlanta, Georgia 30309-3915 (US).

(22) International Filing Date:

21 February 2006 (21.02.2006)

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

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/654,484 21 February 2005 (21.02.2005) US

(71) **Applicant** (*for all designated States except US*): **THE GOVERNMENT OF THE USA, AS REPRESENTED BY THE SECRETARY, DEPT OF HEALTH AND HUMAN SERVICES** [US/US]; THE CENTERS FOR DISEASE CONTROL, AND PREVENTION, TECHNOLOGY TRANSFER OFFICE, 4770 Buford Highway (k-79), Atlanta, Georgia 30341 (US).

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

(71) **Applicant** (*for US only*): **JOHNS HOPKINS UNIVERSITY** [US/US]; JH TECHNOLOGY TRANSFER, 100 North Charles Street, 5th Floor, Baltimore, Maryland 21201 (US).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(72) Inventors; and

(75) **Inventors/Applicants** (*for US only*): **BURKE, Donald** [US/US]; 3241 Livingston St., Washington, District Of Columbia 20015 (US). **SWITZER, William** [US/US]; 5745 Redcoatrun, Stone Mountain, Georgia 30087 (US). **HENEINE, Walid** [US/US]; 2830 Woodland Park Drive, Atlanta, Georgia 30345 (US). **FOLKS, Thomas** [US/US]; 3815 Belle Glade Trail, Snellville, Georgia 30039 (US).

(88) **Date of publication of the international search report:**
28 December 2006

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:** PRIMATE T-LYMPHOTROPIC VIRUSES

(57) **Abstract:** Disclosed are compositions and methods related to the isolation and identification of the novel primate T-lymphotropic viruses, HTLV-3 and HTLV-4. The diversity of HTLVs was investigated among central Africans reporting contact with NHP blood and body fluids through hunting, butchering, and keeping primate pets. Herein it is shown that this population is infected with a variety of HTLVs, including two novel retroviruses; HTLV-4 is the first member of a novel phylogenetic lineage that is distinct from all known HTLVs and STLVs; HTLV-3 falls within the genetic diversity of STLV-3, a group that has not previously been seen in humans. The present invention also relates to vectors and vaccines for use in humans against infection and disease. The invention further relates to a variety of bioassays and kits for the detection and diagnosis of infection with and diseases caused by HTLV-3 and HTLV-4 and related viruses.

WO 2006/091511 A3

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 83

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 83

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

63198-1563

PRIMATE T-LYMPHOTROPIC VIRUSES

1. This invention was made with United States government support. The United States government has certain rights in the invention.

I. BACKGROUND

2. Primate T-lymphotropic viruses (PTLVs) are diverse deltaretroviruses, composed of 3 distinct species (PTLV-1, -2, -3) which by conventional nomenclature are named 'STLV' (simian T-lymphotropic virus) when found in non-human primates (NHPs) and 'HTLV' (human T-lymphotropic virus) when found in humans, regardless of suspected zoonotic origin (Mahieux et al. 1998; Salemi et al. 1999; Slattery et al. 1999; Courgnaud et al. 2004). Like HIV, HTLV has the potential to cause disease and circulate globally in humans sexually, from mother-to-child, and by exposure to contaminated blood from transfusions and intravenous drug use. HTLV-1 causes adult T-cell leukaemia and HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) and other inflammatory diseases (Gessain & Mahieux 2000) and HTLV-2 has been associated with a neurologic disease similar to HAM/TSP (Araujo & Hall 2004). There has been no evidence to date of STLVs crossing into people occupationally exposed to NHPs in laboratories and primate centers, as has been documented with other primate retroviruses, including simian immunodeficiency virus (SIV) (Khabbaz et al. 1994), simian foamy virus (SFV) (Switzer et al. 2004, Heneine et al. 1998), and simian type D retrovirus (Lerche et al. 2001). Nevertheless, ongoing zoonotic transmission of STLV to widespread human populations naturally exposed to NHPs through hunting or butchering, similar to that recently reported for SFV in African hunters (Wolfe et al. 2004b), would be of particular public health significance due to the transmissible and pathogenic nature of this group of viruses among humans. Previous studies have not documented evidence of HTLV outside of the PTLV-1 and PTLV-2 groups (Busch et al. 2000; VanDamme et al. 1997; Salemi et al. 1999; Slattery et al. 1999).

II. SUMMARY

3. Disclosed are methods and compositions related to the isolation of the novel primate T-lymphotropic viruses, HTLV-3 and HTLV-4.

4. The present invention is directed to compositions and methods comprising a novel primate T-lymphotropic virus or primate T-lymphotropic virus, known as HTLV-3 or HTLV-4. The novel primate T-lymphotropic virus of the present invention has utility as a reagent for the immunological screening of the human population for the prevalence of such viruses in the population. The novel primate T-lymphotropic virus of the present invention can also serve as a vector in gene therapy because the virus appears to cause no disease in humans and may not be transmitted to other humans. Additionally, the novel primate T-lymphotropic virus of the present invention can be used as a reagent in pathogenicity studies of these and related viruses. Moreover, the sequences of the novel primate T-lymphotropic virus of the present invention can be used as probes to detect virus in biological samples. Vectors include, but are not limited to, prokaryotic, eukaryotic and viral vectors. The primate T-lymphotropic virus of the present invention can also be used as a live recombinant virus vaccine. Additionally, the primate T-lymphotropic virus of the present invention can be used as a replicating viral system to kill live dividing cells, either *in vitro* or *in vivo*.

5. The present invention includes the isolation and characterization of primate T-lymphotropic viruses, HTLV-3 or HTLV-4, that are believed to have been transmitted from non-human primates to humans at some point in the past. The primate T-lymphotropic virus of the present invention does not appear to be readily transmitted from human to human. The primate T-lymphotropic virus of the present invention can be used in constructing protocols for diagnosing primate T-lymphotropic virus infections and may be used as a vector in gene therapy procedures.

6. The present invention provides compositions comprising live replicating retroviral vectors, wherein the vector is derived from a primate T-lymphotropic virus, and wherein the vector further comprises a nucleic acid that encodes a non-primate T-lymphotropic virus peptide, polypeptide, or protein. For example, the vector can be derived from an HTLV-3 or HTLV-4 virus. Thus, in one aspect, the present invention provides compositions comprising live replicating primate T-lymphotropic virus vectors.

7. Also provided by the present invention is a method of treating a subject with a condition, wherein the condition can be a viral infection, bacterial infection, parasitic infection, proliferative disorder (e.g., cancer), or a condition associated with a genetic or autoimmune disorder; comprising administering to the subject a live replicating viral vector, wherein the immunizing construct is specific for the condition.

8. Also provided by the present invention is a method of preventing a condition in a subject, wherein the condition can be a viral infection, bacterial infection, parasitic infection, proliferative disorder, or a condition associated with a genetic or autoimmune disorder; comprising administering to the subject a live replicating viral vector, wherein the antigen-encoding nucleic acid is specific for the condition.

9. Also provided are methods of using the present vector, isolated viruses, or infectious clones for making viral infection models and using models to study diseases and potential treatments, as well as the models themselves.

10. The present invention also includes methods and compositions for detecting primate T-lymphotropic virus in biological fluids. The methods and compositions, including kits, can be in any configuration well known to those of ordinary skill in the art. The present invention also includes antibodies specific for the primate T-lymphotropic virus and antibodies that inhibit the binding of antibodies specific for the primate T-lymphotropic virus. These antibodies can be polyclonal antibodies or monoclonal antibodies, which also includes fragments of any type of antibody. Thus, specifically disclosed and herein contemplated are antibodies to HTLV-3 and HTLV-4. The antibodies specific for the primate T-lymphotropic virus can be used in diagnostic kits to detect the presence and quantity of primate T-lymphotropic virus in biological fluids or in organs from nonhuman primates for xenotransplantation. For example, the HTLV-3 antibody can be used in a diagnostic kit to detect HTLV-3. Antibodies specific for primate T-lymphotropic virus may also be administered to a human or animal to passively immunize the human or animal against primate T-lymphotropic virus, thereby reducing infection after accidental exposure to nonhuman primate bodily fluids.

11. The present invention also includes compositions and methods, including kits, for detecting the presence and quantity of antibodies that bind primate T-lymphotropic virus in body fluids. The methods, including kits, can be in any configuration well known to those of ordinary skill in the art. Such kits for detection of primate T-lymphotropic virus itself or detection of antibodies to the primate T-lymphotropic virus can be used to monitor the blood supply for the presence of primate T-lymphotropic virus in the blood supply. It is understood and herein contemplated that the disclosed kits can include, for example, a kit for the detection of antibodies to HTLV-3 or HTLV-4.

12. The present invention also includes methods and compositions comprising recombinant live virus vaccines. The virus of the present invention has areas of its genome

63198-1563

that make it ideal for the insertion of exogenous genes. The genes can code for any protein for which vaccination or gene therapy is desired. A novel aspect of such recombinant live viruses is that the recombinant HTLV-3 or HTLV-4 does not cause disease in the host organism. The recombinant live virus vaccines of the present invention are a safe way to
5 provide antigen in a most optimum method to the immune system.

13. Accordingly, one aspect of the present invention provides a composition comprising a novel primate T-lymphotropic virus. It is understood and herein contemplated that an example of the disclosed primate T-lymphotropic virus includes, but is not limited to HTLV-3 and HTLV-4.

10 14. Another aspect of the present invention provides a method of detecting a primate T-lymphotropic virus, such as HTLV-3 or HTLV-4.

15. Another aspect of the present invention provides methods and compositions for detecting the presence and amount of primate T-lymphotropic virus in a body fluid or organ.

15 16. A further aspect of the present invention provides compositions and methods for treating genetic and physiologic disorders using gene therapy techniques comprising the novel primate T-lymphotropic virus of the present invention as a vector for nucleic acid sequences and antisense sequences.

20 17. Another aspect of the present invention provides compositions and methods useful for manipulating the expression of genes.

18. Yet another aspect of the invention provides vaccines.

19. Yet another aspect of the present invention provides compositions and methods for treating viral infections in humans or animals.

25 20. Another aspect of the present invention provides compositions and methods that are effective in treating genetic diseases.

63198-1563

21. Yet another aspect of the present invention provides a method of treating microbial infections in humans or animals.

22. Another aspect of the present invention provides treatments of conditions that are caused in part by rapidly dividing cellular growth.

5 23. Another aspect of the present invention provides live recombinant virus vaccines.

24. Another aspect of the present invention provides diagnostic tools such as antibodies or antigens for the monitoring of the blood supply or organ and tissue donation for the presence of primate T-lymphotropic virus.

10 24a. The present invention as claimed relates to:

- an isolated primate T-lymphotropic virus comprising SEQ ID NOs: 1, 3 and 5;

- an isolated primate T-lymphotropic virus comprising SEQ ID NOs: 1, 3, 35, and 45;

15 - an isolated primate T-lymphotropic virus comprising SEQ ID NOs: 2, 4, and 6;

- an isolated primate T-lymphotropic virus comprising a *pol* gene at least 92.8% identical to SEQ ID NO: 1, an *env* gene at least 92.5% identical to SEQ ID NO: 3, and a *tax* gene at least 94.2% identical to SEQ ID NO: 5;

20 - an isolated primate T-lymphotropic virus comprising a *pol* gene at least 71.5% identical to SEQ ID NO: 2, an *env* gene at least 73.5% identical to SEQ ID NO: 4, and a *tax* gene at least 82% identical to SEQ ID NO: 6;

- a vector comprising the primate T-lymphotropic virus as described herein;

63198-1563

- a use of the vector as described herein for vaccinating a subject against an illness associated with a pathogen;

- a method of identifying new primate T-lymphotropic viruses comprising: a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is SEQ ID NOs: 19 and 20, wherein the second set of primers is SEQ ID NOs: 21 and 22; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus;

- a method of identifying new primate T-lymphotropic viruses comprising:

10 a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is selected from the group of primers pairs consisting of SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, and SEQ ID NOs: 77 and 78, wherein the second set of

15 primers is selected from the group consisting of SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new

20 virus; and

- a method of identifying new primate T-lymphotropic viruses comprising:

a) contacting a nucleic acid using a set of primers, wherein the set of primers is selected from the set of primers consisting of SEQ ID NOs: 19 and 20, SEQ ID NOs: 21 and 22, SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24,

25 SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, SEQ ID NOs: 77 and 78, SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-

63198-1563

lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus.

25. These and other features and advantages of the present invention will become apparent after a review of the following detailed description of the disclosed
5 embodiments and the appended claims.

III. BRIEF DESCRIPTION OF THE DRAWINGS

26. The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments and together with the description illustrate the disclosed compositions and methods.

10 27. Figure 1 shows the Western blot serological pattern of Human T-cell lymphotropic virus (HTLV) infected African hunters. HTLV classification based on phylogenetic analyses is provided above specimen names. Reactivity to HTLV-specific proteins is indicated on left.

15 28. Figure 2 shows the phylogenetic relationships of (a) primate T-cell lymphotropic virus polymerase (a) PTLV pol (662-bp), (b) PTLV env (297-bp), (c) PTLV tax (730-bp), (d) PTLV-3 long terminal repeat (LTR) (398-bp), and (e) PTLV pol-env-tax region (5258-bp) sequences by neighbour joining analysis. Sequences generated in the current study are noted with boxes. Nonhuman primate taxon codes are provided in Methods. Support for the branching order was determined by 1,000 bootstrap replicates, and only values 60% or
20 greater are shown. Branch lengths are proportional to the evolutionary distance (scale bar) between the taxa.

29. Figure 3 shows the Phylogenetic relationships of PTLV type 1 LTR (377-bp) sequences by neighbour joining analysis. Sequences generated in the current study are noted with boxes. Nonhuman primate taxon codes are provided in Methods. Support for
25 the branching order was determined by 1,000 bootstrap replicates and only values 60% or greater are shown. Branch lengths are proportional to the evolutionary distance (scale bar) between the taxa.

63198-1563

30. Figure 4 shows the strategy for PCR-amplifying the entire HTLV-3(2026ND) genome. Small proviral sequences were first amplified in each major gene region and the long terminal repeat (stippled bars) using generic primers as described in methods. The complete proviral sequences was then obtained by using PCR primers located within each

major gene region by genome walking as indicated with arrows and orange bars. The typical HTLV-1 genomic organization is provided for reference.

31. Figure 5(a) shows the nucleotide sequence of the HTLV-3(2026ND) LTR and pre-*gag* region. The U3-R-U5 locations (vertical lines), the approximate cap site (cap), the polyadenylation signal, TATA box, the predicted splice donor site (sd-LTR), and two 21-bp repeats are indicated. In the R and U5 regions, the predicted Rex core elements and nuclear riboprotein A1 binding sites are underlined. The pre-*gag* region and primer binding site (PBS, underlined) are in italics. Figure 5(b) shows the plot of predicted RNA stem loop secondary structure of HTLV-3(2026ND) LTR region. Position of the Rex responsive element (RexRE) core is indicated.

32. Figure 6 shows a comparison of predicted Tax amino acid sequences of primate T-cell lymphotropic viruses (PTLV). Shown in boxes are known functional motifs: NLS, nuclear localization signal; (CBP)/P300, cAMP response element (CREB) binding protein; NES, nuclear export signal; CR2, C-terminal transcriptional activating domain binding; PDZ.

33. Figure 7 shows a comparison of predicted amino acid sequences of primate T-cell lymphotropic viruses (PTLV) and cellular basic leucine zipper (bZIP) transcription factors. Conserved arginine rich and leucine zipper regions of the bZIP proteins are boxed. Frameshift mutation of HTLV-2(MoT) leucine zipper region is depicted with italics.

34. Figure 8 shows the phylogenetic relationship of HTLV-3(2026ND) to other PTLVs (a) entire genome sans long terminal repeat (LTR), (b) *gag*, (c), polymerase (*pol*), and (d) envelope (*env*). Sequences generated in the current study are shown in boxes. Support for the branching order was determined by 1,000 bootstrap replicates; only values of 60% or more are shown. Branch lengths are proportional to the evolutionary distance (scale bar) between the taxa.

35. Figure 9 shows the estimated divergence dates for the most recent common ancestor of HTLV-3(2026ND) and other PTLVs. Divergence dates are provided for each major node of a neighbor-joining tree rooted with PTLV-1 as the outgroup; estimates are provided as ranges using as calibration points 40,000 and 60,000 years ago (YA) as the separation of the Melanesian HTLV-1 (MEL5) sequence from other PTLV-1 strains. Bootstrap analysis of 1000 replicates is shown on the tree branches; only values > 60% are shown.

IV. DETAILED DESCRIPTION

36. Before the present compounds, compositions, articles, devices, and/or methods are disclosed and described, it is to be understood that they are not limited to specific synthetic methods or specific recombinant biotechnology methods unless otherwise specified, or to particular reagents unless otherwise specified, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

A. Definitions

37. As used in the specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “a pharmaceutical carrier” includes mixtures of two or more such carriers, and the like.

38. Ranges can be expressed herein as from “about” one particular value, and/or to “about” another particular value. When such a range is expressed, another embodiment includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent “about,” it will be understood that the particular value forms another embodiment. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as “about” that particular value in addition to the value itself. For example, if the value “10” is disclosed, then “about 10” is also disclosed. It is also understood that when a value is disclosed that “less than or equal to” the value, “greater than or equal to the value” and possible ranges between values are also disclosed, as appropriately understood by the skilled artisan. For example, if the value “10” is disclosed the “less than or equal to 10” as well as “greater than or equal to 10” is also disclosed. It is also understood that the throughout the application, data is provided in a number of different formats, and that this data, represents endpoints and starting points, and ranges for any combination of the data points. For example, if a particular data point “10” and a particular data point 15 are disclosed, it is understood that greater than, greater than or equal to, less than, less than or equal to, and equal to 10 and 15 are considered disclosed as well as between 10 and 15.

39. In this specification and in the claims which follow, reference will be made to a number of terms which shall be defined to have the following meanings:

63198-1563

40. "Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

41. "Primers" are a subset of probes which are capable of supporting some type of enzymatic manipulation and which can hybridize with a target nucleic acid such that the enzymatic manipulation can occur. A primer can be made from any combination of nucleotides or nucleotide derivatives or analogs available in the art which do not interfere with the enzymatic manipulation.

42. "Probes" are molecules capable of interacting with a target nucleic acid, typically in a sequence specific manner, for example through hybridization. The hybridization of nucleic acids is well understood in the art and discussed herein. Typically a probe can be made from any combination of nucleotides or nucleotide derivatives or analogs available in the art.

43.

B. Compositions

44. Disclosed are the components to be used to prepare the disclosed compositions as well as the compositions themselves to be used within the methods disclosed herein. These and other materials are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc. of these materials are disclosed that while specific reference of each various individual and collective combinations and permutation of these compounds may not be explicitly disclosed, each is specifically contemplated and described herein. For example, if a particular HTLV-3 or HTLV-4 or gene of the HTLV-3 or HTLV-4 such as *gag*, *pol*, *env*, LTR, *rex*, and *tax* is disclosed and discussed and a number of modifications that can be made are discussed, specifically contemplated is each and every combination and permutation of HTLV-3 or HTLV-4 or genes of the HTLV-3 or HTLV-4 such as *gag*, *pol*, *env*, LTR, *rex*, and *tax* and the modifications that are possible unless specifically indicated to the contrary. Thus, if a class of molecules A, B, and C are disclosed as well as a class of molecules D, E, and F and an example of a combination

molecule, A-D is disclosed, then even if each is not individually recited each is individually and collectively contemplated meaning combinations, A-E, A-F, B-D, B-E, B-F, C-D, C-E, and C-F are considered disclosed. Likewise, any subset or combination of these is also disclosed. Thus, for example, the sub-group of A-E, B-F, and C-E would be considered disclosed. This concept applies to all aspects of this application including, but not limited to, steps in methods of making and using the disclosed compositions. Thus, if there are a variety of additional steps that can be performed it is understood that each of these additional steps can be performed with any specific embodiment or combination of embodiments of the disclosed methods.

45. Disclosed herein are compositions relating to novel primate T-lymphotropic viruses HTLV-3 (SEQ ID NO: 36) and HTLV-4 (SEQ ID NO: 53). It is understood and herein contemplated that the compositions of the invention can comprise the entire HTLV-3 or HTLV-4 virus. It is also understood that the disclosed compositions of the invention can comprise proteins of the disclosed primate T-lymphocyte viruses or fragments of the disclosed proteins. For example, specifically disclosed and herein contemplated are compositions comprising SEQ ID NOs: 1, 3, 5, 35, 45, 47, 49, 51, and 52, or any combination thereof. Also disclosed are compositions comprising SEQ ID NOs: 2, 4, 6, 59, 61, and 63 or any combination thereof. Also disclosed are compositions comprising SEQ ID NOs: 37, 40, 44, 46, 48, and 50 or any combination thereof. Also disclosed are compositions comprising SEQ ID NOs: 54, 57, 58, 60, and 62 or any combination thereof. Also disclosed are compositions comprising fragments of the disclosed proteins. Thus, for example are compositions comprising SEQ ID NOs: 38, 39, 41, 42, and 43 or any combination thereof. Also disclosed are compositions comprising SEQ ID NOs: 55 and 56. It is understood and herein contemplated that any of the disclosed proteins can be used in combination with any of the protein fragments in the compositions disclosed herein. Thus, for example, disclosed herein are compositions comprising SEQ ID NOs: 37, 38, 39, 40, 41, 42, 43, 44, 46, 48, and 50 or any combination thereof. Also disclosed are SEQ ID NOs: 54, 55, 56, 57, 58, 60, and 62 or any combination thereof. SEQ ID NOs 1-6, 35, and 45 can be used for all the molecular biological techniques known to those skilled in the art. Such uses include, but are not limited to, generation of probes and vectors containing the sequences, antisense sequences derived from such sequences, and proteins synthesized using the sequences. RNA and other nucleic acid derivatives are contemplated by the present invention.

46. It is understood that there are known viruses in the art that based on certain genomic or sequence similarity or taxonomically related to the novel viruses disclosed herein. It is also understood that the known viruses in the art thought related taxonomically do not encode the specific viruses disclosed herein. Thus specifically disclosed and herein contemplated are isolated primate T-lymphotropic viruses having a *pol* gene that has less than 63.5% identity to the *pol* gene of HTLV-1, HTLV-2, STLV-2, and STLV-3, for example, HTLV-4. Also disclosed are isolated primate T-lymphotropic viruses having a *gag* gene that has less than 69% identity to the *gag* gene of HTLV-1, HTLV-2, STLV-2, and STLV-3, for example, HTLV-3. Also disclosed are isolated primate T-lymphotropic viruses having a *pol* gene that has less than 62% identity to the *pol* gene of HTLV-1, HTLV-2, STLV-2, and less than 86% identity to the *pol* gene of STLV-3, for example, HTLV-3. Similarly, the disclosed viruses can be distinguished based on the genes encoded by the viruses of the invention and specifically the identity of said genes to the corresponding genes of known viruses. Thus, specifically disclosed are isolated primate T-lymphotropic viruses having a LTR that has less than 41% identity to the LTR of HTLV-1, HTLV-2 and STLV-3. Also disclosed are isolated primate T-lymphotropic viruses having at least 92.8% identity to the nucleic acid SEQ ID NO: 1.

47. Also disclosed are isolated primate T-lymphotropic virus having at least 92.5% identity to the nucleic acid SEQ ID NO: 3. Also disclosed are primate T-lymphotropic viruses having at least 94.2% identity to the nucleic acid SEQ ID NO: 5. Also disclosed are primate T-lymphotropic viruses having at least 91.5% identity to the nucleic acid SEQ ID NO: 35. Also disclosed are isolated primate T-lymphotropic viruses having at least 92.8% identity to the nucleic acid SEQ ID NO: 1, at least 92.5% identity to the nucleic acid SEQ ID NO: 3, and at least 94.2% identity to the nucleic acid SEQ ID NO: 5.

48. HTLV-4 represents a unique delta primate T-lymphotropic virus. The most notable finding was the discovery of a novel human virus that is distinct from all known PTLV lineages with 29–34.4% and 18.3–25% nucleotide divergence in the conserved *pol* and *tax* genes, respectively, a range of divergence similar to that between PTLV-1, PTLV-2, and PTLV-3. This virus formed a separate phylogenetic lineage with a long branch length and significant bootstrap support in both the *pol* (Figure 2a) and *tax* (Figure 2c) trees. Identical topologies were obtained by using maximum likelihood analysis. Phylogenetic analyses combined with GenBank blast searches show that this is the only known virus in this novel group. For these reasons, this virus, which was designated HTLV-4, qualifies as

the first member of a novel group in the deltaretrovirus genus. Following the guidelines of the International Committee on Taxonomy of Viruses and pending formal classification, primate T-lymphotropic virus 4 (PTLV-4) was proposed as the name for this species, and PTLV-4(1863LE) as the prototype strain. Due to the classification of the virus within the family retroviridae, certain sequence similarity is expected to exist with known retroviruses. It is understood that the known viruses in the art thought to be related taxonomically do not encode the specific viruses disclosed herein. Thus specifically disclosed and herein contemplated are isolated primate T-lymphotropic viruses having at least 71.5% identity to the nucleic acid SEQ ID NO: 2. Also disclosed are isolated primate T-lymphotropic viruses having at least 73.5% identity to the nucleic acid SEQ ID NO: 4. Also disclosed are isolated primate T-lymphotropic viruses having at least 82% identity to the nucleic acid SEQ ID NO: 6. Also disclosed are isolated primate T-lymphotropic viruses having at least 71.5% identity to the nucleic acid SEQ ID NO: 2, at least 73.5% identity to the nucleic acid SEQ ID NO: 4, and at least 82% identity to the nucleic acid SEQ ID NO: 6.

49. Knowing the sequence for HTLV-3 or HTLV-4, allows for various uses of the virus and viral sequences. The *env* gene of HTLV-3 or HTLV-4 is necessary for primate T-lymphotropic virus entry into animal cells. The gene of the present invention is effective in permitting infection of cells in a human host. Thus, for example, the *env* gene is used for uptake of foreign DNA by a wide range of human cells. There has long been a need for vectors for getting foreign nucleic acids into cells, both *in vivo* and *in vitro*. The introduction of foreign or exogenous nucleic acids into cells has been a technological hurdle for many gene therapy applications and has now been solved by the virus and sequences of the present invention. The *env* sequences can be used with any vector known to those skilled in the art, and with any other genetic sequences of choice, to allow for entry of the nucleic acids into the cells.

50. The recent advent of technology, and advances in the understanding of the structure and function of many genes makes it possible to selectively turn off or modify the activity of a given gene. Alteration of gene activity can be accomplished many ways. For example, oligonucleotides that are complementary to certain gene messages or viral sequences, known as "antisense" compounds, have been shown to have an inhibitory effect against viruses. By creating an antisense compound that hybridizes with the targeted RNA message of cells or viruses the translation of the message into protein can be interrupted or prevented. In this fashion gene activity can be modulated.

51. The ability to deactivate specific genes provides great therapeutic benefits. For example, it is theoretically possible to fight viral diseases with antisense molecules that seek out and destroy viral gene products. In tissue culture, antisense oligonucleotides have inhibited infections by herpes-viruses, influenza viruses and the human immunodeficiency virus that causes AIDS. It may also be possible to target antisense oligonucleotides against mutated oncogenes. Antisense technology also holds the potential for regulating growth and development. However, in order for the gene therapy to work, antisense sequences must be delivered across cellular plasma membranes to the cytosol.

52. Gene activity is also modified using sense DNA in a technique known as gene therapy. Defective genes are replaced or supplemented by the administration of “good” or normal genes that are not subject to the defect. Instead of being defective, the gene may have been deleted, thus replacement therapy would provide a copy of the gene for use by the cell. The administered normal genes can either insert into a chromosome or may be present as extracellular DNA and can be used to produce normal RNA, leading to production of the normal gene product. In this fashion gene defects and deficiencies in the production of a gene product may be corrected.

53. Still further gene therapy has the potential to augment the normal genetic complement of a cell. For example, it has been proposed that one way to combat HIV is to introduce into an infected person’s T cells a gene that makes the cells resistant to HIV infection. This form of gene therapy is sometimes called “intracellular immunization.” Genetic material such as a polynucleotide sequence may be administered to a mammal in a viral vector to elicit an immune response against the gene product of the administered nucleic acid sequence. Such gene vaccines elicit an immune response in the following manner. First, the viral vector containing the nucleic acid sequence is administered to a human or animal. Next, the administered sequence is expressed to form a gene product within the human or animal. The gene product inside the human or animal is recognized as foreign material and the immune system of the human or animal mounts an immunological response against the gene product. The virus of the present invention may be used as a viral vector to provide the foreign nucleic acid sequences to the intracellular metabolic processes.

54. Additionally, gene therapy may be used as a method of delivering drugs *in vivo*. For example, if genes that code for therapeutic compounds can be delivered to endothelial cells, the gene products would have facilitated access to the blood stream. Additionally, cells could be infected with a retroviral vector such as the present invention carrying nucleic

acid sequences coding for pharmaceutical agents that prevent infection from occurring in the retrovirally infected cells.

55. The novel primate T-lymphotropic viruses of the present invention can also be used as a safe and effective vaccine agent. Genetic sequences for immunogenic proteins from a variety of infectious agents can be incorporated into the primate T-lymphotropic virus RNA. Once inside a cell, the gene product is expressed and releases the immunizing peptide to the body's immune system. In another method, the virus of the present invention can be used to immunize the body against cell markers found on cancer or tumor cells. The genetic sequence of the cancer cell marker is incorporated into the primate T-lymphotropic virus RNA and after infection with the virus, the expressed gene product stimulates the immune system. The patient's immune system is used to remove the cancerous cells, obviating the need for chemotherapeutic methods.

56. The antibodies of the present invention can be used to detect the presence of the virus or viral particles of the present invention. These antibodies can be used in diagnostic or screening kits to assess the presence of the virus. Additionally, the antibodies can be used to screen organs from nonhuman primates that may be used in humans. Detection of the presence of a virus that is transmitted from nonhuman primates to humans would be crucial in providing virus-free organs for transplantation.

57. Such treatment with HTLV-3 or HTLV-4 can be used for any condition in which rapidly dividing cells provide an aspect of the pathology of the condition. One such condition is the presence of uncontrolled angiogenesis within the body. Angiogenesis dependent diseases are well known in the art and are caused in part by the rapid growth of blood vessels. Another such condition is cancer or tumor growth. Cancer or tumors include both solid tumors and other types. Infection with the virus of the present invention, which can cause no disease and does not effect the host systemically, is an improvement over currently known treatments that involved systemically administered agents. Such chemotherapeutic agents kill rapidly dividing cells but also cause trauma to the entire person. The dosages of such chemotherapeutic agents must be titrated between killing the cancer and killing the patient.

58. In contrast, treatments of cancer with the present invention are not as harmful to the patient. The virus can either be administered systemically or injected in situ into the tumor. The infected cells are killed and tumor growth is stopped. The virus may be administered in one treatment or in a series of treatments.

59. The HTLV-3 or HTLV-4 of the present invention can be recombinantly modified to be selective for cellular receptors on the tumor to make the virus even more specifically targeted to just those cells. Additionally, the virus may have altered promoter regions that can be selectively activated to cause a productive infection. The combination of different levels of control of the virus, both natural and recombinantly-produced, are contemplated in the present invention. A virus could be made specific for attachment to only certain types of cellular receptors, for those cells that are dividing, and will only undergo replication if another exogenous promoter factor is present. Viral infection by two or more individually defective viruses, that require factors or promoters supplied by other primate T-lymphotropic viruses or any type of virus, could provide for many levels of control of infection or treatment of specific conditions.

60. The virus may be administered to the host, for cancer treatment, gene therapy or vaccination by any methods known to those skilled in the art. Such methods include but are not limited to injection, inhalation, ingestion, topical administration and implantation. The virus may be killed or live, depending on the treatment considered.

61. The inventors of the present invention believe that the virus of the present invention, comprising the isolates from HTLV-3, is the first definitive isolation of an STL-3-like primate T-lymphotropic virus from persons exposed to nonhuman primates. This belief is supported by the epidemiology data, the PCR and sequencing data and the serology data and the absence of such reports in the literature. It is understood that HIV-1 and HIV-2 used to be called HTLV-III and HTLV-IV before it was known they were different types of viruses. Additionally, the virus of the present invention comprising the isolates from HTLV-4, are a new species in the delta primate T-lymphotropic viruses.

1. Vectors

62. Disclosed are live replicating human primate T-lymphotropic virus vectors suitable for human use comprising an immunizing construct, wherein the immunizing construct is inserted in nontranslated region between env and tax/rex. The disclosed immunizing construct can be an antigen-encoding nucleic acid.

63. Where reference is made to "antigen"-encoding nucleic acid, it is understood that in the context of the invention antigens encoded by the antigen-encoding nucleic acid can include but are not limited to immunogenic or non-immunogenic peptides, polypeptides, proteins, enzymes, cytokines. These antigens can be non-human exogenous antigenic sequences from viruses, bacteria, or parasites. The antigens can also be antigenic

endogenous human or human derived sequences from a condition such as a cancer. Also, peptides encoded by the antigen-encoding nucleic acid can include non-antigenic sequences for the purposes of gene therapy.

64. In another embodiment of the present invention, sequences of the primate T-lymphotropic viruses of the present invention can be used for other molecular biological applications. Regions of the *gag* gene are important in packaging genetic material. For example, the *gag* sequence or regions of the sequence are incorporated into other vectors and direct the packaging of the resultant genetic material for the particular application desired, such as packaging recombinant sequences to make altered infectious virions. Regions of the *pol* gene are known to be critical for the stable integration of foreign/viral DNA into the host genome. Vectors comprising the *pol* gene sequences can be used to integrate any DNA into a genome. The primate T-lymphotropic virus and sequences of the present invention infect human cells, and thus, these sequences are used with other foreign or exogenous sequences in humans in methods, including, but not limited to, entry into cells, packaging, and insertion into the genome. Additionally, methods of using the primate T-lymphotropic virus and sequences of the present invention are not limited to human cells, but all cells that allow for infection or entry of the nucleic acids.

65. The present invention is directed to compositions and methods comprising a new primate T-lymphotropic virus, HTLV-3 or HTLV-4, particularly compositions and methods for the sequences of the viral genome. The virus was obtained from humans. The new virus of the present invention can be an excellent vector for gene therapy and for vaccination purposes. Additionally, the antibodies or other detection methods for detecting the new virus may be important in detecting the presence of this and related viruses for xenotransplantation. In addition, the novel primate T-lymphotropic virus of the present invention can be used as a reagent in pathogenicity studies of these and related viruses. Moreover, the sequences of the novel primate T-lymphotropic virus of the present invention can be used as probes to detect virus in biological samples. Vectors include but are not limited to prokaryotic, eukaryotic and viral vectors.

66. Many new and potentially useful technologies are being developed which use viral vectors and may form the basis of future medical cures and therapies. Examples of such technologies include, but are not limited to, gene replacement, antisense gene therapy, in situ drug delivery, treatment of cancer or infectious agents, and vaccine therapy.

However, to be successful, these technologies require an effective means for the delivery of the genetic information across cellular membranes.

67. It is well-known in the art that vaccinations can be used prophylactically for the prevention of infections as well as therapeutically for the treatment of ongoing conditions. Such infections or conditions can be but are not limited to viral infections. Thus, also disclosed are vectors of the invention, wherein the antigen-encoding nucleic acid is an antigen from a virus. The viral antigen can be selected from the group of viruses consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-lymphotropic virus, Hantavirus, Rubella virus, Simian Immunodeficiency virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2, and Simian Immunodeficiency virus (SIV). Also disclosed are vectors of the invention, wherein the antigen-encoding nucleic acid is SIV-*GAG*. The art is replete with examples of viral antigens whose sequences and methods of obtaining them are well known.

68. Vaccinations are also known for the prevention of bacterial infections. Additionally, antibiotics are well-known in the art for the treatment of various bacterial infections. Herein contemplated and disclosed are vectors of the invention, wherein the antigen-encoding nucleic acid is an antigen from a bacterium. The bacterial antigen can be selected from the group consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other *Nocardia* species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria*

monocytogenes, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowaria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas aeruginosa*, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Hemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species. The art is replete with examples of bacterial antigens whose sequences and methods of obtaining them are well known.

69. Vaccinations are also known for the prevention of fungal infections. Additionally, antibiotics are well-known in the art for the treatment of various fungal infections. Herein contemplated and disclosed are vectors of the invention, wherein the antigen-encoding nucleic acid is an antigen from a fungus. The fungal antigen can be selected from the group consisting of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatitidis*, *Pneumocystis carinii*, *Penicillium marneffi*, and *Alternaria alternata*.

70. The vectors of the invention are not limited to fungi, bacteria, and viruses. Also disclosed are vectors of the invention, wherein the antigen-encoding nucleic acid is an antigen from a parasite. The parasitic antigen can be selected from the group consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, other *Plasmodium* species., *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, other *Leishmania* species., *Schistosoma mansoni*, other *Schistosoma* species., and *Entamoeba histolytica*. The art is replete with examples of parasitic antigens whose sequences and methods of obtaining them are well known.

71. There are instances wherein it is advantageous to administer the vector of the invention in a pharmaceutical composition that comprises other vaccines. Pharmaceutical compositions comprising multiple vaccines can be for therapeutic or prophylactic purposes. An example Examples of such a composition is other vaccines include the mumps, measles, rubella (MMR) vaccine, and vaccines against *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other

63198-1563

Nocardia species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas aeruginosa*, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Haemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species .

Specifically contemplated and disclosed are pharmaceutical compositions comprising the vector of the invention and one or more additional vaccines. Also disclosed and herein contemplated are instances in which the vector of the invention comprises more than one antigen-encoding nucleic acid. In such a situation the vector of the invention will produce each antigen encoded in the vector as a separate antigen.

72. There are instances in which a vector of the invention alone may not be suitable for a given purpose (e.g., A kit designed to screen potential drugs for the treatment of a condition such kit, intended for use in laboratories without the capabilities to transfect a cell-line with the vector). In such cases, cells previously transfected with the vector of the invention are needed. Thus, also disclosed are cells comprising the vector of the invention.

73. In an embodiment of the invention the antigen-encoding nucleic acid can encode a non-antigenic sequence of DNA. This sequence can provide a functional copy of a disrupted, mutated, disregulated or deleted gene. Examples of nucleic acids encoding proteins that play a role in genetic disorders are known in the literature relating to genetic disorders. Methods of making these cells are described and exemplified herein and in the art.

74. The ability to detect the presence of a construct can be a desirable feature of any vector. As such, vectors are often contain a marker to show that the construct of interest has been delivered to the subject (e.g., a cell), and once delivered, is being expressed. A marker can take the form of a gene that is detectable when expressed. Thus, also disclosed are vectors of the invention further comprising a reporter gene. One example of a reporter gene is green fluorescence protein (GFP).

2. Delivery of the compositions to cells

75. There are a number of compositions and methods which can be used to deliver nucleic acids to cells, either in vitro or in vivo. These methods and compositions can largely be broken down into two classes: viral based delivery systems and non-viral based delivery systems. For example, the nucleic acids can be delivered through a number of direct delivery systems such as, electroporation, lipofection, calcium phosphate precipitation, plasmids, viral vectors, viral nucleic acids, phage nucleic acids, phages, cosmids, or via transfer of genetic material in cells or carriers such as cationic liposomes. Appropriate means for transfection, including viral vectors, chemical transfectants, or physico-mechanical methods such as electroporation and direct diffusion of DNA, are described by, for example, Wolff, J. A., et al., Science, 247, 1465-1468, (1990); and Wolff, J. A. Nature, 352, 815-818, (1991) Such methods are well known in the art and readily adaptable for use with the compositions and methods described herein. In certain cases, the methods will be modified to specifically function with large DNA molecules. Further, these methods can be used to target certain diseases and cell populations by using the targeting characteristics of the carrier.

a) Nucleic acid based delivery systems

76. Transfer vectors can be any nucleotide construction used to deliver genes into cells (e.g., a plasmid), or as part of a general strategy to deliver genes, e.g., as part of recombinant retrovirus or adenovirus (Ram et al. Cancer Res. 53:83-88, (1993)).

77. As used herein, plasmid or viral vectors are agents that transport nucleic acids into the cell without degradation and include a promoter yielding expression of the gene in the cells into which it is delivered. In some embodiments the vectors are derived from either a virus or specifically a retrovirus. Viral vectors can include for example, for example, HTLV-1, HTLV-2, HTLV-3, HTLV-4, Adenovirus, Adeno-associated virus, Herpes virus, Vaccinia virus, Polio virus, AIDS virus, neuronal trophic virus, Sindbis and other RNA viruses, including these viruses with the HIV backbone. Also preferred are any viral families which share the properties of these viruses which make them suitable for use as vectors. Retroviruses include Murine Maloney Leukemia virus, MMLV, and retroviruses that express the desirable properties of MMLV as a vector. Retroviral vectors are able to carry a larger genetic payload, i.e., a transgene or marker gene, than other viral vectors, and for this reason are a commonly used vector. However, they are not as useful in non-proliferating cells. Adenovirus vectors are relatively stable and easy to work with, have

63198-1563

high titers, and can be delivered in aerosol formulation, and can transfect non-dividing cells. Pox viral vectors are large and have several sites for inserting genes, they are thermostable and can be stored at room temperature. A preferred embodiment is a viral vector which has been engineered so as to suppress the immune response of the host organism, elicited by the viral antigens. Preferred vectors of this type will carry coding regions for Interleukin 8 or 10.

78. Viral vectors can have higher transaction (ability to introduce genes) abilities than chemical or physical methods to introduce genes into cells. Typically, viral vectors contain, nonstructural early genes, structural late genes, an RNA polymerase III transcript, inverted terminal repeats necessary for replication and encapsidation, and promoters to control the transcription and replication of the viral genome. When engineered as vectors, viruses typically have one or more of the early genes removed and a gene or gene/promoter cassette is inserted into the viral genome in place of the removed viral DNA. Constructs of this type can carry up to about 8 kb of foreign genetic material. The necessary functions of the removed early genes are typically supplied by cell lines which have been engineered to express the gene products of the early genes in trans.

3. Retroviral Vectors

79. Primate T-lymphotropic viruses are retroviruses. A retrovirus is an animal virus belonging to the virus family of Retroviridae, including any types, subfamilies, genus, or tropisms. Retroviral vectors, in general, are described by Verma, I.M., Retroviral vectors for gene transfer. In Microbiology-1985, American Society for Microbiology, pp. 229-232, Washington, (1985). Examples of methods for using retroviral vectors for gene therapy are described in U.S. Patent Nos. 4,868,116 and 4,980,286; PCT applications WO 90/02806 and WO 89/07136; and Mulligan, (Science 260:926-932 (1993)). Although the present primate T-lymphotropic virus vector is unique, the methods described for using other types of viral vectors can be useful in certain contexts. See for example U.S. Patent No. 5,646,032.

80. A retrovirus is essentially a package which has packed into it nucleic acid cargo. The nucleic acid cargo carries with it a packaging signal, which ensures that the replicated daughter molecules will be efficiently packaged within the package coat. In addition to the package signal, there are a number of molecules which are needed in cis, for the replication, and packaging of the replicated virus. Typically a retroviral genome, contains the *gag*, *pol*,

and *env* genes which are involved in the making of the protein coat. It is the *gag*, *pol*, and *env* genes which are typically replaced by the foreign DNA that is to be transferred to the target cell. Retrovirus vectors typically contain a packaging signal for incorporation into the package coat, a sequence which signals the start of the *gag* transcription unit, elements necessary for reverse transcription, including a primer binding site to bind the tRNA primer of reverse transcription, terminal repeat sequences that guide the switch of RNA strands during DNA synthesis, a purine rich sequence 5' to the 3' LTR that serves as the priming site for the synthesis of the second strand of DNA synthesis, and specific sequences near the ends of the LTRs that enable the insertion of the DNA state of the retrovirus to insert into the host genome. The removal of the *gag*, *pol*, and *env* genes allows for large fragments of foreign sequence to be inserted into the viral genome, become reverse transcribed, and upon replication be packaged into a new retroviral particle. This amount of nucleic acid is sufficient for the delivery of a one to many genes depending on the size of each transcript. It is preferable to include either positive or negative selectable markers along with other genes in the insert.

81. Since the replication machinery and packaging proteins in most retroviral vectors have been removed (*gag*, *pol*, and *env*), the vectors are typically generated by placing them into a packaging cell line. A packaging cell line is a cell line which has been transfected or transformed with a retrovirus that contains the replication and packaging machinery, but lacks any packaging signal. When the vector carrying the DNA of choice is transfected into these cell lines, the vector containing the gene of interest is replicated and packaged into new retroviral particles, by the machinery provided in cis by the helper cell. The genomes for the machinery are not packaged because they lack the necessary signals.

82. A packaging cell line is a cell line which has been transfected or transformed with a retrovirus that contains the replication and packaging machinery, but lacks any packaging signal. When the vector carrying the DNA of choice is transfected into these cell lines, the vector containing the gene of interest is replicated and packaged into new retroviral particles, by the machinery provided in cis by the helper cell. The genomes for the machinery are not packaged because they lack the necessary signals.

83. It is also understood and herein contemplated that the pX region can be used to construct a vector. The pX region is located between the end of *env* and the beginning of *Tax* and contains small ORFs hence this can be another good region for insertion of foreign DNA in an HTLV genome based vector.

84. Disclosed are methods of detecting the expression of the vector of the invention comprising using a first antibody to the antigen to measure protein expression in a quantitative or qualitative way, further comprising detecting the first antibody directly via a colorimetric measurement produced through the use of a substrate and a conjugated antibody or indirectly via a first antibody to the antigen which in turn is bound by a second antibody which is conjugated and will result in a colorimetric measurement when combined with a substrate.

85. Also disclosed are methods of the invention, wherein the antigen is detected by placing an aliquot of the vector of the invention in a lane on a gel and probing the gel for the antigen.

86. Disclosed are methods of detecting the expression of the vector of the invention comprising using a fluorescently labeled first antibody specific for the antigen and visualizing the antigen using a flow cytometer, fluorescence microscope, or chemiluminescence.

87. Also disclosed are methods of the invention, wherein the first antibody is not fluorescently labeled, but a target for a second antibody with a fluorescent label.

88. Disclosed are methods of detecting the expression of a disclosed vector comprising using cytolytic killing assay to assess activity.

89. Disclosed are methods of detecting the vector of the invention further comprising obtaining a sample from a subject comprising a tissue biopsy or removal of blood or bone marrow.

a) Non-nucleic acid based systems

90. The disclosed compositions can be delivered to the target cells in a variety of ways. For example, the compositions can be delivered through electroporation, or through lipofection, or through calcium phosphate precipitation. The delivery mechanism chosen will depend in part on the type of cell targeted and whether the delivery is occurring for example in vivo or in vitro.

91. In the methods described above which include the administration and uptake of exogenous DNA into the cells of a subject (i.e., gene transduction or transfection), delivery of the compositions to cells can be via a variety of mechanisms. As one example, delivery can be via a liposome, using commercially available liposome preparations such as LIPOFECTIN, LIPOFECTAMINE (GIBCO-BRL, Inc., Gaithersburg, MD), SUPERFECT (Qiagen, Inc. Hilden, Germany) and TRANSFECTAM (Promega Biotec, Inc., Madison,

WI), as well as other liposomes developed according to procedures standard in the art. In addition, the disclosed nucleic acid or vector can be delivered *in vivo* by electroporation, the technology for which is available from Genetronics, Inc. (San Diego, CA) as well as by means of a SONOPORATION machine (ImaRx Pharmaceutical Corp., Tucson, AZ).

92. The materials may be in solution, suspension (for example, incorporated into microparticles, liposomes, or cells). These may be targeted to a particular cell type via antibodies, receptors, or receptor ligands. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Senter, et al., *Bioconjugate Chem.*, 2:447-451, (1991); Bagshawe, K.D., *Br. J. Cancer*, 60:275-281, (1989); Bagshawe, et al., *Br. J. Cancer*, 58:700-703, (1988); Senter, et al., *Bioconjugate Chem.*, 4:3-9, (1993); Battelli, et al., *Cancer Immunol. Immunother.*, 35:421-425, (1992); Pietersz and McKenzie, *Immunolog. Reviews*, 129:57-80, (1992); and Roffler, et al., *Biochem. Pharmacol.*, 42:2062-2065, (1991)). These techniques can be used for a variety of other specific cell types. Vehicles such as "stealth" and other antibody conjugated liposomes (including lipid mediated drug targeting to colonic carcinoma), receptor mediated targeting of DNA through cell specific ligands, lymphocyte directed tumor targeting, and highly specific therapeutic retroviral targeting of murine glioma cells *in vivo*. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Hughes et al., *Cancer Research*, 49:6214-6220, (1989); and Litzinger and Huang, *Biochimica et Biophysica Acta*, 1104:179-187, (1992)). In general, receptors are involved in pathways of endocytosis, either constitutive or ligand induced. These receptors cluster in clathrin-coated pits, enter the cell via clathrin-coated vesicles, pass through an acidified endosome in which the receptors are sorted, and then either recycle to the cell surface, become stored intracellularly, or are degraded in lysosomes. The internalization pathways serve a variety of functions, such as nutrient uptake, removal of activated proteins, clearance of macromolecules, opportunistic entry of viruses and toxins, dissociation and degradation of ligand, and receptor-level regulation. Many receptors follow more than one intracellular pathway, depending on the cell type, receptor concentration, type of ligand, ligand valency, and ligand concentration. Molecular and cellular mechanisms of receptor-mediated endocytosis has been reviewed (Brown and Greene, *DNA and Cell Biology* 10:6, 399-409 (1991)).

93. Nucleic acids that are delivered to cells which are to be integrated into the host cell genome, typically contain integration sequences. These sequences are often viral

related sequences, particularly when viral based systems are used. These viral integration systems can also be incorporated into nucleic acids which are to be delivered using a non-nucleic acid based system of deliver, such as a liposome, so that the nucleic acid contained in the delivery system can be come integrated into the host genome.

94. Other general techniques for integration into the host genome include, for example, systems designed to promote homologous recombination with the host genome. These systems typically rely on sequence flanking the nucleic acid to be expressed that has enough homology with a target sequence within the host cell genome that recombination between the vector nucleic acid and the target nucleic acid takes place, causing the delivered nucleic acid to be integrated into the host genome. These systems and the methods necessary to promote homologous recombination are known to those of skill in the art.

b) In vivo/ex vivo

95. As described above, the compositions can be administered in a pharmaceutically acceptable carrier and can be delivered to the subject's cells *in vivo* and/or *ex vivo* by a variety of mechanisms well known in the art (e.g., uptake of naked DNA, liposome fusion, intramuscular injection of DNA via a gene gun, endocytosis and the like).

96. If *ex vivo* methods are employed, cells or tissues can be removed and maintained outside the body according to standard protocols well known in the art. The compositions can be introduced into the cells via any gene transfer mechanism, such as, for example, calcium phosphate mediated gene delivery, electroporation, microinjection or proteoliposomes. The transduced cells can then be infused (e.g., in a pharmaceutically acceptable carrier) or homotopically transplanted back into the subject per standard methods for the cell or tissue type. Standard methods are known for transplantation or infusion of various cells into a subject.

4. Expression systems

97. The nucleic acids that are delivered to cells typically contain expression controlling systems. For example, the inserted genes in viral and retroviral systems usually contain promoters, and/or enhancers to help control the expression of the desired gene product. A promoter is generally a sequence or sequences of DNA that function when in a relatively fixed location in regard to the transcription start site. A promoter contains core elements required for basic interaction of RNA polymerase and transcription factors, and may contain upstream elements and response elements.

a) Viral Promoters and Enhancers

98. Preferred promoters controlling transcription from vectors in mammalian host cells may be obtained from various sources, for example, the genomes of viruses such as: polyoma, Simian Virus 40 (SV40), adenovirus, retroviruses, hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian promoters, e.g. beta actin promoter. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication (Fiers et al., *Nature*, 273: 113 (1978)). The immediate early promoter of the human cytomegalovirus is conveniently obtained as a *HindIII* E restriction fragment (Greenway, P.J. et al., *Gene* 18: 355-360 (1982)). Of course, promoters from the host cell or related species also are useful herein. Such preferred promoters are in the LTRs of HTLV.

99. Enhancer generally refers to a sequence of DNA that functions at no fixed distance from the transcription start site and can be either 5' (Laimins, L. et al., *Proc. Natl. Acad. Sci.* 78: 993 (1981)) or 3' (Lusky, M.L., et al., *Mol. Cell Bio.* 3: 1108 (1983)) to the transcription unit. Furthermore, enhancers can be within an intron (Banerji, J.L. et al., *Cell* 33: 729 (1983)) as well as within the coding sequence itself (Osborne, T.F., et al., *Mol. Cell Bio.* 4: 1293 (1984)). They are usually between 10 and 300 bp in length, and they function in cis. Enhancers function to increase transcription from nearby promoters. Enhancers also often contain response elements that mediate the regulation of transcription. Promoters can also contain response elements that mediate the regulation of transcription. Enhancers often determine the regulation of expression of a gene. While many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, fetoprotein and insulin), typically one will use an enhancer from a eukaryotic cell virus for general expression. Preferred examples are the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

100. The promotor and/or enhancer may be specifically activated either by light or specific chemical events which trigger their function. Systems can be regulated by reagents such as tetracycline and dexamethasone. There are also ways to enhance viral vector gene expression by exposure to irradiation, such as gamma irradiation, or alkylating chemotherapy drugs.

101. In certain embodiments the promoter and/or enhancer region can act as a constitutive promoter and/or enhancer to maximize expression of the region of the transcription unit to be transcribed. In certain constructs the promoter and/or enhancer

region be active in all eukaryotic cell types, even if it is only expressed in a particular type of cell at a particular time. A preferred promoter of this type is the CMV promoter (650 bases). Other preferred promoters are SV40 promoters, cytomegalovirus (full length promoter), and retroviral vector LTR.

102. It has been shown that all specific regulatory elements can be cloned and used to construct expression vectors that are selectively expressed in specific cell types such as melanoma cells. The glial fibrillary acetic protein (GFAP) promoter has been used to selectively express genes in cells of glial origin.

103. Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or nucleated cells) may also contain sequences necessary for the termination of transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding tissue factor protein. The 3' untranslated regions also include transcription termination sites. It is preferred that the transcription unit also contain a polyadenylation region. One benefit of this region is that it increases the likelihood that the transcribed unit will be processed and transported like mRNA. The identification and use of polyadenylation signals in expression constructs is well established. It is preferred that homologous polyadenylation signals be used in the transgene constructs. In certain transcription units, the polyadenylation region is derived from the SV40 early polyadenylation signal and consists of about 400 bases. It is also preferred that the transcribed units contain other standard sequences alone or in combination with the above sequences improve expression from, or stability of, the construct.

b) Markers

104. The viral vectors can include nucleic acid sequence encoding a marker product. This marker product is used to determine if the gene has been delivered to the cell and once delivered is being expressed. Preferred marker genes are the *E. Coli* lacZ gene, which encodes β -galactosidase, and green fluorescent protein.

105. In some embodiments the marker may be a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase, neomycin, neomycin analog G418, hydromycin, and puromycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure. There are two widely used distinct categories of selective regimes. The first category is based on a

63198-1563

cell's metabolism and the use of a mutant cell line which lacks the ability to grow independent of a supplemented media. Two examples are: CHO DHFR- cells and mouse LTK- cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented media. An alternative to supplementing the media is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells which were not transformed with the DHFR or TK gene will not be capable of survival in non-supplemented media.

106. The second category is dominant selection which refers to a selection scheme used in any cell type and does not require the use of a mutant cell line. These schemes typically use a drug to arrest growth of a host cell. Those cells which have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of such dominant selection use the drugs neomycin, (Southern P. and Berg, P., J. Molec. Appl. Genet. 1: 327 (1982)), mycophenolic acid, (Mulligan, R.C. and Berg, P. Science 209: 1422 (1980)) or hygromycin, (Sugden, B. et al., Mol. Cell. Biol. 5: 410-413 (1985)). The three examples employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid) or hygromycin, respectively. Others include the neomycin analog G418 and puramycin.

5. Sequence similarities

107. It is understood that as discussed herein the use of the terms homology and identity mean the same thing as similarity. Thus, for example, if the use of the word homology is used between two non-natural sequences it is understood that this is not necessarily indicating an evolutionary relationship between these two sequences, but rather is looking at the similarity or relatedness between their nucleic acid sequences. Many of the methods for determining homology between two evolutionarily related molecules are routinely applied to any two or more nucleic acids or proteins for the purpose of measuring sequence similarity regardless of whether they are evolutionarily related or not.

108. In general, it is understood that one way to define any known variants and derivatives or those that might arise, of the disclosed genes and proteins herein, is through defining the variants and derivatives in terms of homology to specific known sequences. This identity of particular sequences disclosed herein is also discussed elsewhere herein. In general, variants of genes and proteins herein disclosed typically have at least, about 70, 71,

63198-1563

72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99 percent homology to the stated sequence or the native sequence. Those of skill in the art readily understand how to determine the homology of two proteins or nucleic acids, such as genes. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

109. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Adv. Appl. Math.* 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970), by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

110. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 for at least material related to nucleic acid alignment. It is understood that any of the methods typically can be used and that in certain instances the results of these various methods may differ, but the skilled artisan understands if identity is found with at least one of these methods, the sequences would be said to have the stated identity, and be disclosed herein.

111. For example, as used herein, a sequence recited as having a particular percent homology to another sequence refers to sequences that have the recited homology as calculated by any one or more of the calculation methods described above. For example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using the Zuker calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by any of the other calculation methods. As another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using both the Zuker calculation method and the Pearson and Lipman calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by the Smith and Waterman calculation method, the Needleman and Wunsch calculation

method, the Jaeger calculation methods, or any of the other calculation methods. As yet another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using each of calculation methods (although, in practice, the different calculation methods will often result in different calculated homology percentages).

6. Nucleic acids

112. There are a variety of molecules disclosed herein that are nucleic acid based, including for example the nucleic acids that encode, for example HTLV-3 or HTLV-4. The disclosed nucleic acids are made up of, for example, nucleotides, nucleotide analogs, or nucleotide substitutes. Non-limiting examples of these and other molecules are discussed herein. It is understood that for example, when a vector is expressed in a cell, that the expressed mRNA will typically be made up of A, C, G, and U. Likewise, it is understood that if, for example, an antisense molecule is introduced into a cell or cell environment through for example exogenous delivery, it is advantageous that the antisense molecule be made up of nucleotide analogs that reduce the degradation of the antisense molecule in the cellular environment.

a) Nucleotides and related molecules

113. A nucleotide is a molecule that contains a base moiety, a sugar moiety and a phosphate moiety. Nucleotides can be linked together through their phosphate moieties and sugar moieties creating an internucleoside linkage. The base moiety of a nucleotide can be adenin-9-yl (A), cytosin-1-yl (C), guanin-9-yl (G), uracil-1-yl (U), and thymin-1-yl (T). The sugar moiety of a nucleotide is a ribose or a deoxyribose. The phosphate moiety of a nucleotide is pentavalent phosphate. A non-limiting example of a nucleotide would be 3'-AMP (3'-adenosine monophosphate) or 5'-GMP (5'-guanosine monophosphate). There are many varieties of these types of molecules available in the art and available herein.

114. A nucleotide analog is a nucleotide which contains some type of modification to either the base, sugar, or phosphate moieties. Modifications to nucleotides are well known in the art and would include for example, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, and 2-aminoadenine as well as modifications at the sugar or phosphate moieties. There are many varieties of these types of molecules available in the art and available herein.

115. Nucleotide substitutes are molecules having similar functional properties to nucleotides, but which do not contain a phosphate moiety, such as peptide nucleic acid

63198-1563

(PNA). Nucleotide substitutes are molecules that will recognize nucleic acids in a Watson-Crick or Hoogsteen manner, but which are linked together through a moiety other than a phosphate moiety. Nucleotide substitutes are able to conform to a double helix type structure when interacting with the appropriate target nucleic acid. There are many varieties of these types of molecules available in the art and available herein.

116. It is also possible to link other types of molecules (conjugates) to nucleotides or nucleotide analogs to enhance for example, cellular uptake. Conjugates can be chemically linked to the nucleotide or nucleotide analogs. Such conjugates include but are not limited to lipid moieties such as a cholesterol moiety. (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989,86, 6553-6556). There are many varieties of these types of molecules available in the art and available herein.

117. A Watson-Crick interaction is at least one interaction with the Watson-Crick face of a nucleotide, nucleotide analog, or nucleotide substitute. The Watson-Crick face of a nucleotide, nucleotide analog, or nucleotide substitute includes the C2, N1, and C6 positions of a purine based nucleotide, nucleotide analog, or nucleotide substitute and the C2, N3, C4 positions of a pyrimidine based nucleotide, nucleotide analog, or nucleotide substitute.

118. A Hoogsteen interaction is the interaction that takes place on the Hoogsteen face of a nucleotide or nucleotide analog, which is exposed in the major groove of duplex DNA. The Hoogsteen face includes the N7 position and reactive groups (NH₂ or O) at the C6 position of purine nucleotides.

b) Sequences

119. There are a variety of sequences related to the protein molecules, for example the protein coding regions *gag*, *pol*, *env*, *tax*, *rex*, and protease (*pro*) genes and noncoding regions such as the LTR of HTLV-3 and HTLV-4, or any of the nucleic acids disclosed herein for making HTLV-3 or HTLV-4, all of which are encoded by nucleic acids or are nucleic acids. The sequences for the human analogs of these genes, as well as other analogs, and alleles of these genes, and splice variants and other types of variants, are available in a variety of protein and gene databases, including GenBank. GenBank (National Center for Biotechnology Information National Library of Medicine, Building 38A, Bethesda, MD, USA, 20894). Those of skill in the art understand how to resolve sequence discrepancies and differences and to adjust the

compositions and methods relating to a particular sequence to other related sequences. Primers and/or probes can be designed for any given sequence given the information disclosed herein and known in the art.

c) Primers and probes

120. Disclosed are compositions including primers and probes, which are capable of interacting with the disclosed nucleic acids, such as the HTLV-3 or HTLV-4 as disclosed herein. In certain embodiments the primers are used to support nucleic acid (DNA, RNA, etc.) amplification reactions. Thus, for example, disclosed herein are primers wherein the primer comprises SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, SEQ ID NOs: 77 and 78, SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 64 and 65, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80. Typically the primers will be capable of being extended in a sequence specific manner. Extension of a primer in a sequence specific manner includes any methods wherein the sequence and/or composition of the nucleic acid molecule to which the primer is hybridized or otherwise associated directs or influences the composition or sequence of the product produced by the extension of the primer. Extension of the primer in a sequence specific manner therefore includes, but is not limited to, PCR, DNA sequencing, DNA extension, DNA polymerization, RNA transcription, or reverse transcription. Techniques and conditions that amplify the primer in a sequence specific manner are preferred. In certain embodiments the primers are used for the DNA amplification reactions, such as PCR or direct sequencing. Thus, herein are disclosed primer pairs used in conjunction with a second nested set of primers pairs. For example, disclosed herein are PCR amplification methods comprising a first primer pair and a second primer pair, wherein the second primer pair is internal to the first primer pair and wherein the first primer pair is selected from the group consisting of SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, and SEQ ID NOs: 77 and 78, wherein the second set of primers is selected from the group consisting of SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80. It is understood that in certain

embodiments the primers can also be extended using non-enzymatic techniques, where for example, the nucleotides or oligonucleotides used to extend the primer are modified such that they will chemically react to extend the primer in a sequence specific manner.

Typically the disclosed primers hybridize with the disclosed nucleic acids or region of the nucleic acids or they hybridize with the complement of the nucleic acids or complement of a region of the nucleic acids.

d) Functional Nucleic Acids

121. Functional nucleic acids are nucleic acid molecules that have a specific function, such as binding a target molecule or catalyzing a specific reaction. Functional nucleic acid molecules can be divided into the following categories, which are not meant to be limiting. For example, functional nucleic acids include antisense molecules, aptamers, ribozymes, triplex forming molecules, and external guide sequences. The functional nucleic acid molecules can act as effectors, inhibitors, modulators, and stimulators of a specific activity possessed by a target molecule, or the functional nucleic acid molecules can possess a de novo activity independent of any other molecules.

122. Functional nucleic acid molecules can interact with any macromolecule, such as DNA, RNA, polypeptides, or carbohydrate chains. Thus, functional nucleic acids can interact with the mRNA of any of the disclosed nucleic acids, such as the *pol*, *tax*, *env*, *gag*, *rex* and *pro* genes and noncoding regions such as the LTR of HTLV-3 and HTLV-4, or the nucleic acids used for the generation of HTLV-3 and HTLV-4, or the genomic DNA of any of the disclosed viruses, such as HTLV-3 and HTLV-4, or they can interact with the polypeptide encoded by any of the disclosed nucleic acids, such as *pol*, *tax*, *rex*, *env*, *gag*, or *pro* genes of HTLV-3 and HTLV-4, or the nucleic acids used for the generation of *pol*, *tax*, *rex*, *env*, *gag*, or LTR proteins of HTLV-3 and HTLV-4. Often functional nucleic acids are designed to interact with other nucleic acids based on sequence homology between the target molecule and the functional nucleic acid molecule. In other situations, the specific recognition between the functional nucleic acid molecule and the target molecule is not based on sequence homology between the functional nucleic acid molecule and the target molecule, but rather is based on the formation of tertiary structure that allows specific recognition to take place.

7. Peptides

a) Protein variants

123. As discussed herein there are numerous variants of the HTLV-3 genes encoded herein such as *gag* (SEQ ID NO: 40), *pol* (SEQ ID NO: 44), *env* (SEQ ID NO: 37), *tax* (SEQ ID NO: 50), *rex* (SEQ ID NO: 48), protease (SEQ ID NO: 46) and noncoding regions such as the LTR and HTLV-4 genes encoded herein such as *gag*, *pol* (SEQ ID NO: 57), *env* (SEQ ID NO: 54), *tax* (SEQ ID NO: 62), *rex* (SEQ ID NO: 60), protease (SEQ ID NO: 58) and noncoding regions such as the LTR that are known and herein contemplated. In addition, to the known functional HTLV-3 and HTLV-4 strain variants there are derivatives of the HTLV-3 and HTLV-4 *gag*, *pol*, *tax*, *rex*, and *env*, LTR proteins which also function in the disclosed methods and compositions. Protein variants and derivatives are well understood to those of skill in the art and in can involve amino acid sequence modifications. For example, amino acid sequence modifications typically fall into one or more of three classes: substitutional, insertional or deletional variants. Insertions include amino and/or carboxyl terminal fusions as well as intrasequence insertions of single or multiple amino acid residues. Insertions ordinarily will be smaller insertions than those of amino or carboxyl terminal fusions, for example, on the order of one to four residues. Immunogenic fusion protein derivatives, such as those described in the examples, are made by fusing a polypeptide sufficiently large to confer immunogenicity to the target sequence by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding the fusion. Deletions are characterized by the removal of one or more amino acid residues from the protein sequence. Typically, no more than about from 2 to 6 residues are deleted at any one site within the protein molecule. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the protein, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis and PCR mutagenesis. Amino acid substitutions are typically of single residues, but can occur at a number of different locations at once; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. The mutations must not place the sequence out of

reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. Substitutional variants are those in which at least one residue has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Tables 1 and 2 and are referred to as conservative substitutions.

124. TABLE 1:Amino Acid Abbreviations

Amino Acid	Abbreviations
alanine	AlaA
allosoleucine	Alle
arginine	ArgR
asparagine	AsnN
aspartic acid	AspD
cysteine	CysC
glutamic acid	GluE
glutamine	GlnK
glycine	GlyG
histidine	HisH
isoleucine	IleI
leucine	LeuL
lysine	LysK
phenylalanine	PheF
proline	ProP
pyroglutamic acidp	Glu
serine	SerS
threonine	ThrT
tyrosine	TyrY
tryptophan	TrpW
valine	ValV

TABLE 2:Amino Acid Substitutions

Original Residue Exemplary Conservative Substitutions, others are known in the art.
Alaser
Arglys, gln
Asngln; his
Aspglu
Cysser
Glnasn, lys
Gluasp
Glypro
Hisasn;gln
Ileleu; val
Leuile; val
Lysarg; gln;

MetLeu; ile
Phemet; leu; tyr
Serthr
Thrser
Trptyr
Tyrtrp; phe
Valile; leu

125. Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the protein properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine, in this case, (e) by increasing the number of sites for sulfation and/or glycosylation.

126. For example, the replacement of one amino acid residue with another that is biologically and/or chemically similar is known to those skilled in the art as a conservative substitution. For example, a conservative substitution would be replacing one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as, for example, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. Such conservatively substituted variations of each explicitly disclosed sequence are included within the mosaic polypeptides provided herein.

127. Substitutional or deletional mutagenesis can be employed to insert sites for N-glycosylation (Asn-X-Thr/Ser) or O-glycosylation (Ser or Thr). Deletions of cysteine or other labile residues also may be desirable. Deletions or substitutions of potential proteolysis sites, e.g. Arg, is accomplished for example by deleting one of the basic residues or substituting one by glutamyl or histidyl residues.

63198-1563

128. Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and asparyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the o-amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, *Proteins: Structure and Molecular Properties*, W. H. Freeman & Co., San Francisco pp 79-86 [1983]), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

129. It is understood that one way to define the variants and derivatives of the disclosed proteins herein is through defining the variants and derivatives in terms of homology/identity to specific known sequences. For example, SEQ ID NO: 1 sets forth a particular sequence of HTLV-3 *pol* protein and SEQ ID NO: 2 sets forth a particular sequence of a HTLV-4 *pol* protein. Specifically disclosed are variants of these and other proteins herein disclosed which have at least, 70% or 75% or 80% or 85% or 90% or 95% homology or any amount of homology in between to the stated sequence. Those of skill in the art readily understand how to determine the homology of two proteins. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

130. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Adv. Appl. Math.* 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970), by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci. U.S.A.* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

131. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 for at least material related to nucleic acid alignment.

63198-1563

132. It is understood that the description of conservative mutations and homology can be combined together in any combination, such as embodiments that have at least 70% homology to a particular sequence wherein the variants are conservative mutations.

133. As this specification discusses various proteins and protein sequences it is understood that the nucleic acids that can encode those protein sequences are also disclosed. This would include all degenerate sequences related to a specific protein sequence, i.e. all nucleic acids having a sequence that encodes one particular protein sequence as well as all nucleic acids, including degenerate nucleic acids, encoding the disclosed variants and derivatives of the protein sequences. Thus, while each particular nucleic acid sequence may not be written out herein, it is understood that each and every sequence is in fact disclosed and described herein through the disclosed protein sequence. For example, one of the many nucleic acid sequences that can encode the protein sequence set forth in SEQ ID NO: 44 is set forth in SEQ ID NO: 1. In addition, for example, disclosed are conservative derivatives of SEQ ID NO: 44.

134. It is understood that there are numerous amino acid and peptide analogs which can be incorporated into the disclosed compositions. For example, there are numerous D amino acids or amino acids which have a different functional substituent than the amino acids shown in Table 1 and Table 2. The opposite stereo isomers of naturally occurring peptides are disclosed, as well as the stereo isomers of peptide analogs. These amino acids can readily be incorporated into polypeptide chains by charging tRNA molecules with the amino acid of choice and engineering genetic constructs that utilize, for example, amber codons, to insert the analog amino acid into a peptide chain in a site specific way (Thorson et al., *Methods in Molec. Biol.* 77:43-73 (1991), Zoller, *Current Opinion in Biotechnology*, 3:348-354 (1992); Ibba, *Biotechnology & Genetic Engineering Reviews* 13:197-216 (1995), Cahill et al., *TIBS*, 14(10):400-403 (1989); Benner, *TIB Tech*, 12:158-163 (1994); Ibba and Hennecke, *Bio/technology*, 12:678-682 (1994) at least for material related to amino acid analogs).

135. Molecules can be produced that resemble peptides, but which are not connected via a natural peptide linkage. For example, linkages for amino acids or amino acid analogs can include $\text{CH}_2\text{NH--}$, $\text{--CH}_2\text{S--}$, $\text{--CH}_2\text{--CH}_2\text{--}$, --CH=CH-- (cis and trans), $\text{--COCH}_2\text{--}$, $\text{--CH(OH)CH}_2\text{--}$, and $\text{--CHH}_2\text{SO--}$ (These and others can be found in Spatola, A. F. in *Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins*, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., *Vega Data* (March 1983),

63198-1563

Vol. 1, Issue 3, Peptide Backbone Modifications (general review); Morley, Trends Pharm Sci (1980) pp. 463-468; Hudson, D. et al., Int J Pept Prot Res 14:177-185 (1979) (--CH₂NH--, CH₂CH₂--); Spatola et al. Life Sci 38:1243-1249 (1986) (--CH H₂--S); Hann J. Chem. Soc Perkin Trans. I 307-314 (1982) (--CH--CH--, cis and trans); Almquist et al. J. Med. Chem. 23:1392-1398 (1980) (--COCH₂--); Jennings-White et al. Tetrahedron Lett 23:2533 (1982) (--COCH₂--); Szelke et al. European Appln, EP 45665 CA (1982): 97:39405 (1982) (--CH(OH)CH₂--); Holladay et al. Tetrahedron. Lett 24:4401-4404 (1983) (--C(OH)CH₂--); and Hruby Life Sci 31:189-199 (1982) (--CH₂--S--).

A particularly preferred non-peptide linkage is --CH₂NH--.

It is understood that peptide analogs can have more than one atom between the bond atoms, such as β -alanine, γ -aminobutyric acid, and the like.

136. Amino acid analogs and analogs and peptide analogs often have enhanced or desirable properties, such as, more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity, and others.

137. D-amino acids can be used to generate more stable peptides, because D amino acids are not recognized by peptidases and such. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (e.g., D-lysine in place of L-lysine) can be used to generate more stable peptides. Cysteine residues can be used to cyclize or attach two or more peptides together. This can be beneficial to constrain peptides into particular conformations. (Rizo and Gierasch Ann. Rev. Biochem. 61:387 (1992)).

8. Pharmaceutical carriers/Delivery of pharmaceutical products

138. As described above, the compositions can also be administered *in vivo* in a pharmaceutically acceptable carrier. By "pharmaceutically acceptable" is meant a material that is not biologically or otherwise undesirable, i.e., the material may be administered to a subject, along with the nucleic acid or vector, without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained. The carrier would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

139. The compositions may be administered orally, parenterally (e.g., intravenously), by intramuscular injection, by intraperitoneal injection, transdermally,

63198-1563

extracorporeally, topically or the like, including topical intranasal administration or administration by inhalant. As used herein, "topical intranasal administration" means delivery of the compositions into the nose and nasal passages through one or both of the nares and can comprise delivery by a spraying mechanism or droplet mechanism, or through aerosolization of the nucleic acid or vector. Administration of the compositions by inhalant can be through the nose or mouth via delivery by a spraying or droplet mechanism. Delivery can also be directly to any area of the respiratory system (e.g., lungs) via intubation. The exact amount of the compositions required will vary from subject to subject, depending on the species, age, weight and general condition of the subject, the severity of the allergic disorder being treated, the particular nucleic acid or vector used, its mode of administration and the like. Thus, it is not possible to specify an exact amount for every composition. However, an appropriate amount can be determined by one of ordinary skill in the art using only routine experimentation given the teachings herein.

140. Parenteral administration of the composition, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as emulsions. A more recently revised approach for parenteral administration involves use of a slow release or sustained release system such that a constant dosage is maintained. See, e.g., U.S. Patent No. 3,610,795.

141. The materials may be in solution, suspension (for example, incorporated into microparticles, liposomes, or cells). These may be targeted to a particular cell type via antibodies, receptors, or receptor ligands. The following references are examples of the use of this technology to target specific proteins to tumor tissue (Senter, et al., *Bioconjugate Chem.*, 2:447-451, (1991); Bagshawe, K.D., *Br. J. Cancer*, 60:275-281, (1989); Bagshawe, et al., *Br. J. Cancer*, 58:700-703, (1988); Senter, et al., *Bioconjugate Chem.*, 4:3-9, (1993); Battelli, et al., *Cancer Immunol. Immunother.*, 35:421-425, (1992); Pietersz and McKenzie, *Immunolog. Reviews*, 129:57-80, (1992); and Roffler, et al., *Biochem. Pharmacol.*, 42:2062-2065, (1991)). Vehicles such as "stealth" and other antibody conjugated liposomes (including lipid mediated drug targeting to colonic carcinoma), receptor mediated targeting of DNA through cell specific ligands, lymphocyte directed tumor targeting, and highly specific therapeutic retroviral targeting of murine glioma cells *in vivo*. The following references are examples of the use of this technology to target specific proteins to tumor

tissue (Hughes et al., *Cancer Research*, 49:6214-6220, (1989); and Litzinger and Huang, *Biochimica et Biophysica Acta*, 1104:179-187, (1992)). In general, receptors are involved in pathways of endocytosis, either constitutive or ligand induced. These receptors cluster in clathrin-coated pits, enter the cell via clathrin-coated vesicles, pass through an acidified endosome in which the receptors are sorted, and then either recycle to the cell surface, become stored intracellularly, or are degraded in lysosomes. The internalization pathways serve a variety of functions, such as nutrient uptake, removal of activated proteins, clearance of macromolecules, opportunistic entry of viruses and toxins, dissociation and degradation of ligand, and receptor-level regulation. Many receptors follow more than one intracellular pathway, depending on the cell type, receptor concentration, type of ligand, ligand valency, and ligand concentration. Molecular and cellular mechanisms of receptor-mediated endocytosis has been reviewed (Brown and Greene, *DNA and Cell Biology* 10:6, 399-409 (1991)).

a) Pharmaceutically Acceptable Carriers

142. The compositions, including antibodies, can be used therapeutically in combination with a pharmaceutically acceptable carrier.

143. Suitable carriers and their formulations are described in *Remington: The Science and Practice of Pharmacy* (19th ed.) ed. A.R. Gennaro, Mack Publishing Company, Easton, PA 1995. Typically, an appropriate amount of a pharmaceutically-acceptable salt is used in the formulation to render the formulation isotonic. Examples of the pharmaceutically-acceptable carrier include, but are not limited to, saline, Ringer's solution and dextrose solution. The pH of the solution is preferably from about 5 to about 8, and more preferably from about 7 to about 7.5. Further carriers include sustained release preparations such as semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, liposomes or microparticles. It will be apparent to those persons skilled in the art that certain carriers may be more preferable depending upon, for instance, the route of administration and concentration of composition being administered.

144. Pharmaceutical carriers are known to those skilled in the art. These most typically would be standard carriers for administration of drugs to humans, including solutions such as sterile water, saline, and buffered solutions at physiological pH. The compositions can be administered intramuscularly or subcutaneously. Other compounds will be administered according to standard procedures used by those skilled in the art.

145. Pharmaceutical compositions may include carriers, thickeners, diluents, buffers, preservatives, surface active agents and the like in addition to the molecule of choice. Pharmaceutical compositions may also include one or more active ingredients such as antimicrobial agents, antiinflammatory agents, anesthetics, and the like.

146. The pharmaceutical composition may be administered in a number of ways depending on whether local or systemic treatment is desired, and on the area to be treated. Administration may be topically (including ophthalmically, vaginally, rectally, intranasally), orally, by inhalation, or parenterally, for example by intravenous drip, subcutaneous, intraperitoneal or intramuscular injection. The disclosed antibodies can be administered intravenously, intraperitoneally, intramuscularly, subcutaneously, intracavity, or transdermally.

147. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

148. Formulations for topical administration may include ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.

149. Compositions for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, capsules, sachets, or tablets. Thickeners, flavorings, diluents, emulsifiers, dispersing aids or binders may be desirable..

150. Some of the compositions may potentially be administered as a pharmaceutically acceptable acid- or base- addition salt, formed by reaction with inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, and phosphoric acid, and organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, and fumaric acid, or by reaction with an inorganic base such as sodium

hydroxide, ammonium hydroxide, potassium hydroxide, and organic bases such as mono-, di-, trialkyl and aryl amines and substituted ethanolamines.

b) Therapeutic Uses

151. Effective dosages and schedules for administering the compositions may be determined empirically, and making such determinations is within the skill in the art. The dosage ranges for the administration of the compositions are those large enough to produce the desired effect in which the symptoms disorder are effected. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of the disease in the patient, route of administration, or whether other drugs are included in the regimen, and can be determined by one of skill in the art. The dosage can be adjusted by the individual physician in the event of any counterindications. Dosage can vary, and can be administered in one or more dose administrations daily, for one or several days. Guidance can be found in the literature for appropriate dosages for given classes of pharmaceutical products. For example, guidance in selecting appropriate doses for antibodies can be found in the literature on therapeutic uses of antibodies, e.g., Handbook of Monoclonal Antibodies, Ferrone et al., eds., Noyes Publications, Park Ridge, N.J., (1985) ch. 22 and pp. 303-357; Smith et al., Antibodies in Human Diagnosis and Therapy, Haber et al., eds., Raven Press, New York (1977) pp. 365-389. A typical daily dosage of the antibody used alone might range from about 1 $\mu\text{g/kg}$ to up to 100 mg/kg of body weight or more per day, depending on the factors mentioned above.

152. Following administration of a disclosed composition, such as an antibody, for treating, inhibiting, or preventing an HTLV-3 or HTLV-4 infection, the efficacy of the therapeutic antibody can be assessed in various ways well known to the skilled practitioner. For instance, one of ordinary skill in the art will understand that a composition, such as an antibody, disclosed herein is efficacious in treating or inhibiting an HTLV-3 or HTLV-4 infection in a subject by observing that the composition reduces viral load or prevents a further increase in HTLV-3 or HTLV-4 viral load. Techniques used to measure viral HTLV-3 or HTLV-4 treatment with an antibody partially or completely inhibits the appearance of the virus in the blood or other body fluid.

153. Other molecules that interact with HTLV-3 or HTLV-4 or the genes encoded by those viruses to inhibit HTLV-3 or HTLV-4 and cell interactions which do not have a specific pharmaceutical function, but which may be used for tracking changes within

cellular chromosomes or for the delivery of diagnostic tools for example can be delivered in ways similar to those described for the pharmaceutical products.

154. The disclosed compositions and methods can also be used for example as tools to isolate and test new drug candidates for a variety of primate T-lymphotropic virus related diseases.

C. Treatment and Prevention Methods

155. By “treating” is meant an improvement in the disease state (i.e., viral infection, bacterial infection, parasitic infection, cancer, genetic disorder, or autoimmune disease) is observed and/or detected upon administration of a substance of the present invention to a subject. Treatment can range from a positive change in a symptom or symptoms of the disease to complete amelioration of the viral infection, bacterial infection, parasitic infection, or cancer (e.g., reduction in severity or intensity of disease, alteration of clinical parameters indicative of the subject’s condition, relief of discomfort or increased or enhanced function), as detected by art-known techniques. The methods of the present invention can be utilized to treat an established viral infection, bacterial infection, parasitic infection, or cancer. One of skill in the art would recognize that viral infection, bacterial infection, parasitic infection, or cancer refer to conditions characterized by the presence of a foreign pathogen or abnormal cell growth. Clinical symptoms will depend on the particular condition and are easily recognizable by those skilled in the art of treating the specific condition. Treatment methods can include, but are not limited to therapeutic vaccinations. Thus, disclosed are methods of treating a subject with a condition comprising administering to the vector of the invention.

156. Also disclosed are methods of the invention, wherein the condition being treated is a viral infection. The viral infection can be selected from the list of viruses consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever

virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-lymphotropic virus, Hantavirus, Rubella virus, Simian Immunodeficiency virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2.

157. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a virus. The viral antigen can be selected from the group of viruses consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-lymphotropic virus, Hantavirus, Rubella virus, Simian Immunodeficiency virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2.

158. Also disclosed are methods of the invention, wherein the condition being treated is a bacterial infection. The bacterial infection can be selected from the list of bacterium consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other *Nocardia* species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas*

aeruginosa, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Hemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species.

159. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a bacterium. The bacterial antigen can be selected from the group consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other *Nocardia* species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas aeruginosa*, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Hemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species.

160. Also disclosed are methods of the invention, wherein the condition being treated is a fungal infection. The fungal infection can be selected from the list of fungus consisting of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatidis*, *Pneumocystis carinii*, *Penicillium marneffi*, and *Alternaria alternata*.

161. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a fungus. The fungal antigen can be selected from the group consisting of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatidis*, *Pneumocystis carinii*, *Penicillium marneffi*, and *Alternaria alternata*.

162. Also disclosed are methods of the invention, wherein the condition being treated is a parasitic infection. The parasitic infection can be selected from the list of parasites consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, other *Plasmodium* species., *Trypanosoma brucei*, *Trypanosoma*

cruzi, *Leishmania major*, other *Leishmania* species., *Schistosoma mansoni*, other *Schistosoma* species., and *Entamoeba histolytica*.

163. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a parasite. The parasitic antigen can be selected from the group consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, other *Plasmodium* species., *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, other *Leishmania* species., *Schistosoma mansoni*, other *Schistosoma* species., and *Entamoeba histolytica*.

164. Also disclosed are methods of the invention, wherein the condition being treated is cancer.

165. The disclosed vectors and vector containing compositions can be used to treat any disease where uncontrolled cellular proliferation occurs such as cancers. A non-limiting list of different types of cancers is as follows: lymphomas (Hodgkins and non-Hodgkins), leukemias, carcinomas, carcinomas of solid tissues, squamous cell carcinomas, adenocarcinomas, sarcomas, gliomas, high grade gliomas, blastomas, neuroblastomas, plasmacytomas, histiocytomas, melanomas, adenomas, hypoxic tumors, myelomas, AIDS-related lymphomas or sarcomas, metastatic cancers, or cancers in general.

166. A representative but non-limiting list of cancers that the disclosed compositions can be used to treat is the following: lymphoma, B cell lymphoma, T cell lymphoma, mycosis fungoides, Hodgkin's Disease, myeloid leukemia, bladder cancer, brain cancer, nervous system cancer, head and neck cancer, squamous cell carcinoma of head and neck, kidney cancer, lung cancers such as small cell lung cancer and non-small cell lung cancer, neuroblastoma/glioblastoma, ovarian cancer, pancreatic cancer, prostate cancer, skin cancer, liver cancer, melanoma, squamous cell carcinomas of the mouth, throat, larynx, and lung, colon cancer, cervical cancer, cervical carcinoma, breast cancer, and epithelial cancer, renal cancer, genitourinary cancer, pulmonary cancer, esophageal carcinoma, head and neck carcinoma, large bowel cancer, hematopoietic cancers; testicular cancer; colon and rectal cancers, prostatic cancer, or pancreatic cancer.

167. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is a tumor antigen. The tumor antigen can be selected from the list consisting of human epithelial cell mucin (Muc-1; a 20 amino acid core repeat for Muc-1 glycoprotein, present on breast cancer cells and pancreatic cancer cells), the Ha-ras oncogene product, p53, carcino-embryonic antigen (CEA), the raf oncogene product, gp100/pmell17, GD2,

GD3, GM2, TF, sTn, MAGE-1, MAGE-3, BAGE, GAGE, tyrosinase, gp75, Melan-A/Mart-1, gp100, HER2/neu, EBV-LMP 1 & 2, HPV-F4, 6, 7, prostate-specific antigen (PSA), HPV-16, MUM, alpha-fetoprotein (AFP), CO17-1A, GA733, gp72, p53, the ras oncogene product, HPV E7, Wilm's tumor antigen-1, telomerase, and melanoma gangliosides.

168. Disclosed are methods of treating a condition in a subject comprising administering to the subject the vector of the invention, wherein the condition is due to a mutated, dysregulated, disrupted, or deleted gene; autoimmunity; or inflammatory diseases.

169. Disclosed are methods of treating a condition in a subject, wherein the condition can be selected from list consisting of cystic fibrosis, asthma, multiple sclerosis, muscular dystrophy, diabetes, tay-sachs, spinobifida, cerebral palsy, parkinson's disease, lou gehrigg disease, alzheimer's, systemic lupus erythamatosi, hemophilia, Addison's disease, Cushing's disease.

170. By "preventing" is meant that after administration of a substance of the present invention to a subject, the subject does not develop the symptoms of the viral, bacterial, or parasitic infection, and/or does not develop the viral, bacterial, or parasitic infection. "Preventing" or "prevention" can also refer to the ultimate reduction of an infection, condition, or symptoms of an infection, or condition relative to infections or conditions in subjects that do not receive the substance. Methods of prevention can include, but are not limited to prophylactic vaccination. As such, disclosed are methods of preventing an infection in a subject comprising administering to the subject the vector of the invention.

171. Also disclosed are methods of the invention, wherein the infection prevented is a viral infection. The viral infection can be selected from the list of viruses consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-

lymphotropic virus, Hantavirus, Rubella virus, Simian Immunodeficiency virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2.

172. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a virus. The viral antigen can be selected from the group consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-lymphotropic virus, Hantavirus, Rubella virus, Simian Immunodeficiency virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2.

173. Also disclosed are methods of the invention, wherein the infection prevented is a bacterial infection. The bacterial infection can be selected from the list of bacterium consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other *Nocardia* species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhoea*, *Pseudomonas aeruginosa*, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Haemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species.

174. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a bacterium. The bacterial antigen can be selected from the group consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, other *Nocardia* species, *Legionella pneumophila*, other *Legionella* species, *Salmonella typhi*, other *Salmonella* species, *Shigella* species, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, other *Pasteurella* species, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, other *Brucella* species, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, other *Rickettsial* species, *Ehrlichia* species, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter* species, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas aeruginosa*, other *Pseudomonas* species, *Haemophilus influenzae*, *Haemophilus ducreyi*, other *Hemophilus* species, *Clostridium tetani*, other *Clostridium* species, *Yersinia enterocolitica*, and other *Yersinia* species.

175. Also disclosed are methods of the invention, wherein the infection prevented is a fungal infection. The fungal infection can be selected from the list of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatitidis*, *Pneumocystis carinii*, *Penicillium marneffii*, and *Alternaria alternata*.

176. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a fungus. The fungal antigen can be selected from the group consisting of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatitidis*, *Pneumocystis carinii*, *Penicillium marneffii*, and *Alternaria alternata*.

177. Also disclosed are methods of the invention, wherein the infection prevented is a parasitic infection. The parasitic infection can be selected from the list of parasites consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, other *Plasmodium* species., *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, other *Leishmania* species., *Schistosoma mansoni*, other *Schistosoma* species., and *Entamoeba histolytica*.

178. Also disclosed are methods of the invention, wherein the antigen-encoding nucleic acid is an antigen from a parasite. The parasitic antigen can be selected from the group consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, other *Plasmodium* species., *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, other *Leishmania* species., *Schistosoma mansoni*, other *Schistosoma* species., and *Entamoeba histolytica*.

179. Also disclosed are methods of the invention, wherein the subject is a horse, cow, pig, dog, cat, mouse, monkey, human, or a cell isolated from such an animal.

D. Screening Methods

180. Disclosed herein are methods of identifying new primate T-lymphotropic viruses comprising: a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is SEQ ID NOs: 19 and 20, and wherein the second set of primers is SEQ ID NOs: 21 and 22 under conditions that permit primer extension; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein a sequence divergence greater than 5% indicates a new virus.

181. Also disclosed are methods of identifying new primate T-lymphotropic viruses comprising: a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is selected from the group of primers pairs consisting of SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, and SEQ ID NOs: 77 and 78, wherein the second set of primers is selected from the group consisting of SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus.

182. It is also understood that the disclosed methods of identifying a new primate T-lymphotropic virus can be achieved using non-nested PCR techniques such as real-time PCR. Thus, for example, specifically disclosed are methods of identifying new primate T-lymphotropic viruses comprising a) contacting a nucleic acid using a set of primers, wherein the set of primers is selected from the set of primers consisting of SEQ ID NOs: 19 and 20,

SEQ ID NOs: 21 and 22, SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, SEQ ID NOs: 77 and 78, SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, SEQ ID NOs: 79 and 80, and SEQ ID NOs: 64 and 65; b) identifying any amplified nucleic acid; and c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus. Also disclosed are identification methods wherein the method is a real-time PCR method.

183. It is also contemplated that the disclosed methods can be used in conjunction with probes to detect the presence of amplification product. Specifically disclosed are fluorescently labeled probes that can be used to detect the amplification product of the disclosed methods. For example, a fluorescent probe, can comprise TTCCCAAGGCTTCAAAAACAGCCCCACGC (SEQ ID NO: 66).

184. It is understood and herein contemplated that the surface antigen (SU) and transmembrane are regions of *env* can be used serologically for the identification and differentiation of PTLVs (the type specific peptides MTA-1 and K55 are in SU; likewise the p24 region of gag can be used for the serological identification of PTLV). Thus, disclosed herein are methods of identifying a PTLV comprising contacting a nucleic acid with a set of primers specific for the surface antigen or transmembrane regions of *env* and indentifying any amplified nucleic acid.

185. It is also understood that the disclosed peptides, polypeptides, proteins, and protein fragments can be used to generate antibodies that can be used to identify new and known primate T-lymphotropic viruses. Specifically disclosed are methods of identifying the presence of a primate T-lymphotropic virus in a subject comprising taking a tissue sample from the subject and contacting the sample with an antibody directed to an HTLV-3 or HTLV-4 peptide, polypeptide, protein, or protein fragment, wherein the peptide, polypeptide, protein, or protein fragment can be selected from the group consisting of SEQ ID NOs. 1, 2, 3, 4, 5, 6, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 67, and 68, and wherein biding of the antibody to the sample indicates the presence of a new or known primate T-lymphotropic virus. It is understood and herein contemplated that the disclosed methods can be used to identify new primate T-lymphotropic viruses as well as detect all primate T-lymphotropic viruses or a

group of particular primate T-lymphotropic viruses. Those of skill in the art will know which antibodies to use to accomplish their detection goal. For example, to detect more than one of the known HTLV viruses (HTLV-1, 2, and 3, or HTLV-1, 2, and 4) one can use type specific peptide of HTLV-1 and HTLV-2 such as SEQ ID NO: 67 and 68.

186. Also provided by the present invention is a method of screening a substance for effectiveness in treating or reducing the severity of the condition (e.g., HTLV-3 or HTLV-4 infection) comprising: a) obtaining an animal having the condition or characteristic (e.g., symptom) of the condition; b) administering the substance to an animal having one or more characteristics of the condition; and assaying the animal for an effect on the condition, thereby identifying a substance effective in reducing the condition.

187. The ability of a substance to reduce the severity of a condition can be determined by evaluating the histological and clinical manifestations, of the animal with condition before and after administration of the substance of interest and quantitating the amount of reduction of the condition.

188. The animal in which the condition or characteristic (e.g., symptom) of the condition is produced can be any mammal and can include but is not limited to mouse, rat, guinea pig, hamster, rabbit, cat, dog, goat, monkey, and chimpanzee. The condition or characteristic (e.g., symptom) of the condition can be produced in the animal by any method known in the art. For example, HTLV-3 or HTLV-4 can be produced by introducing into the animal (eg., chimpanzee infected with HTLV-3 or HTLV-4 or rhesus macaques or nemestrina macaques infected with an HTLV-3 or HTLV-4 *env* on an SIV backbone. Pullium, JK, et. al., J. Infectious Dis. 183:1023, 2001) an infectious amount of HTLV-3 or HTLV-4.

189. The present invention also provides a method of screening for a substance effective in preventing the condition (e.g., HTLV-3 or HTLV-4 infection) comprising: a) administering the substance to an animal susceptible to the condition; b) subjecting the animal to treatment that will induce the condition or characteristic (e.g., symptom) of the condition; and c) assaying cells from the animal for an change in immune responses as compared to an the immune responses in a control animal having the condition in the absence of the substance identifies a substance that is effective in preventing the condition.

190. A model for use in screening for substances effective in treating or preventing a disease comprising an animal capable of manifesting a characteristic of the disease is provided, wherein the animal has been administered the vector of the invention.

191. A method of making the model of the invention comprising obtaining administering to an animal capable of manifesting a characteristic of the disease and administering to said animal the vector of the invention which encodes an antigen associate with the disease.

192. A method of screening for a substance effective in treating a disease associated with an immunizing construct comprising: a) administering the substance to the model of the invention; and b) assaying for an change in the course of the disease as compared to an the course of the disease in a control subject; an improvement in the course of the disease in the presence of the substance identifies a substance that is effective in treating the disease.

193. A method of screening for a substance effective in preventing a disease associated with an immunizing construct comprising: a) administering a the vector of the invention to a subject; b) subjecting the subject to treatment that will induce the disease or characteristic (e.g., symptom) of the disease; and c) assaying for an change in the course of the disease as compared to an the course of the disease in a control subject; an improvement in the course of the disease in the presence of the substance identifies a substance that is effective in preventing the disease.

194. A method of screening for a substance effective in treating a disease associated with an immunizing construct comprising: a) subjecting the subject to treatment that will induce the disease or characteristic (e.g., symptom) of the disease; b) administering a the vector of the invention to a subject; and c) assaying for an change in the course of the disease as compared to an the course of the disease in a control subject; an improvement in the course of the disease in the presence of the substance identifies a substance that is effective in treating the disease.

1. Methods of using the compositions as research tools

195. The disclosed compositions can also be used diagnostic tools related to primate T-lymphotropic diseases such as HTLV-3 and HTLV-4.

E. Methods of making the compositions

196. The compositions disclosed herein and the compositions necessary to perform the disclosed methods can be made using any method known to those of skill in the art for that particular reagent or compound unless otherwise specifically noted.

1. Process claims for making the compositions

197. Disclosed are processes for making the compositions as well as making the intermediates leading to the compositions. For example, disclosed are nucleic acids in SEQ ID NOs: 1-6, 35, and 45. There are a variety of methods that can be used for making these compositions, such as synthetic chemical methods and standard molecular biology methods. It is understood that the methods of making these and the other disclosed compositions are specifically disclosed.

198. Disclosed are nucleic acid molecules produced by the process comprising linking in an operative way a nucleic acid comprising the sequence set forth in SEQ ID NOs: 1-6, 35, and 45 and a sequence controlling the expression of the nucleic acid.

199. Also disclosed are nucleic acid molecules produced by the process comprising linking in an operative way a nucleic acid molecule comprising a sequence having 80% identity to a sequence set forth in SEQ ID NOs: 1-6, 35, and 45, and a sequence controlling the expression of the nucleic acid.

200. Disclosed are nucleic acid molecules produced by the process comprising linking in an operative way a nucleic acid molecule comprising a sequence that hybridizes under stringent hybridization conditions to a sequence set forth SEQ ID NOs: 1-6, 35, and 45 and a sequence controlling the expression of the nucleic acid.

201. Disclosed are nucleic acid molecules produced by the process comprising linking in an operative way a nucleic acid molecule comprising a sequence encoding a peptide set forth in SEQ ID NOs: 1-6, 35, and 45 and a sequence controlling an expression of the nucleic acid molecule.

202. Disclosed are nucleic acid molecules produced by the process comprising linking in an operative way a nucleic acid molecule comprising a sequence encoding a peptide having 80% identity to a peptide set forth in SEQ ID NOs: 1-6, 35, and 45 and a sequence controlling an expression of the nucleic acid molecule.

203. Disclosed are cells produced by the process of transforming the cell with any of the disclosed nucleic acids. Disclosed are cells produced by the process of transforming the cell with any of the non-naturally occurring disclosed nucleic acids.

204. Disclosed are any of the disclosed peptides produced by the process of expressing any of the disclosed nucleic acids. Disclosed are any of the non-naturally occurring disclosed peptides produced by the process of expressing any of the disclosed

nucleic acids. Disclosed are any of the disclosed peptides produced by the process of expressing any of the non-naturally disclosed nucleic acids.

F. Examples

205. The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the compounds, compositions, articles, devices and/or methods claimed herein are made and evaluated, and are intended to be purely exemplary and are not intended to limit the disclosure. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts, temperature, etc.), but some errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C or is at ambient temperature, and pressure is at or near atmospheric.

1. Example 1

206. To determine whether novel HTLVs are present among individuals exposed to the blood and body fluids of wild primate populations (Wolfe et al. 2004a) known to be infected with STL (Courgnaud et al. 2004), individuals were examined from twelve villages in southern Cameroon proximal to both forested and nonforested NHP habitats. Individuals were asked to identify and quantify their exposure to NHPs, which were organized according to three categories reliably distinguished by this population: chimpanzee, gorilla and monkey (Wolfe et al. 2004a). A total of 930 who reported exposure to NHP blood and body fluids, mainly through hunting and butchering were selected for further analysis. Plasma specimens from exposed people were screened for HTLVs using an HTLV-1/2 EIA capable of detecting antibodies to a broad range of HTLVs, followed by confirmation with an HTLV Western blot (WB) assay that can distinguish HTLV-1 and HTLV-2 (van Dooren et al. 2004). A total of 97 (10.4%) persons were EIA reactive of which 90 (9.7%) were also reactive in the WB assay. A broad range of WB profiles were seen, including HTLV-1-like (1.1 %), HTLV-2-like (0.5 %), HTLV-positive but untypeable (1.4%), and HTLV indeterminate (6.7%).

207. DNA from peripheral blood mononuclear cells (PBMCs) available from 86 of the 90 WB reactive samples were then subjected to PCR amplification of several viral regions. Viral sequences from 13 persons were obtained using this strategy. The WB reactivities of these 13 persons is shown in Fig. 1 and included HTLV-1-like (n=9), HTLV-2-like (n=1), and HTLV indeterminate (n=3) profiles. All 13 HTLV-infected persons were exclusively from lowland forest sites, including both men and women who often reported

multiple opportunities for contact with the blood and body fluids of NHPs (Table 3). Since PTLV diversity is influenced more by geography than by primate species (Salemi et al 1999, Slattery et al. 1999, Gessain & Mahieux 2000), viral sequences were analyzed phylogenetically along with African and global representatives of HTLV and STLV.

Table 3. Nonhuman primate exposures for human T-lymphotropic virus (HTLV)-infected central African hunters

Central African Hunters														Reported Injuries			
D	Site	HTLV	Nearest PTLV	Sex	Age	Hunting Technique	NHP Exposure										
							Hunt			Butcher			Pet				
							m	c	g	m	c	g	m	c	g		
42	LE	HTLV-1	Group D - Mandrill clade	m	32					x							bitten/scatched by wild animal
63	LE	HTLV-4	Distinct from all major PTLV groups	m	48	s	x	x	x								
72	LE	HTLV-1	Group A Cosmopolitan	f	27					x							
79	MO	HTLV-1	Group B- Central African	m	30	g	x			x							monkey bite
127	MO	HTLV-1	Group D - Mandrill clade	m	44	g, s	x			x	x	x					
180	MV	HTLV-1	Group B- Central African	f	55					x	x	x					
143	MV	HTLV-1	Group B- Central African	f	71					x	x	x					
103	MV	HTLV-1	Group B- Central African	f	75					x	x	x					
137	MV	HTLV-1	Group B- Central African	m	39	g, s	x			x							wild animal injured finger
126	ND	HTLV-3	STLV-3	m	63	s	x			x							
156	ND	HTLV-1	Group G - Central West Africa	m	65	g	x			x					x		
159	NG	HTLV-1	Group D - Mandrill clade	m	71	g, s	x	x									bitten/scatched by wild animal
110	YI	HTLV-1	Group G - Central West Africa	m	55	s	x			x							

*,m = monkey, c = chimpanzee, g = gorilla.

†, PTLV, primate T lymphotropic virus; STLV, simian T-lymphotropic virus

Table 4: Nucleotide and Amino Acid Percent Identities¹

	HTLV-1 (ATK)	HTLV-2 (MoT)	STLV-2 (PP1664)	STLV-3 (TGE2117)	HTLV-3 (2026ND)
HTLV-3 _{2026ND}					
Genome (8917-bp)	61.6	62.9	62.6	87.0	-
LTR(697-bp)	48.7	43.7	41.4	86.7	-
<i>gag</i> (1268-bp)	69.3 (83.2)	69.4 (80.5)	70.6 (80.7)	87.5 (96.0)	-
<i>pro</i> (534-bp)	59.7 (62.6)	59.2 (66.7)	59.4 (59.3)	84.3 (88.1)	-
<i>pol</i> (2670-bp)	62.2 (66.2)	63.9 (71.2)	63.5 (69.9)	86.2 (93.1)	-
<i>env</i> (1476-bp)	65.9 (73.8)	69.0 (78.2)	67.1 (77.4)	87.8 (95.7)	-
<i>tax</i> (1053-bp)	76.3 (81.4)	75.1 (83.4)	74.4 (80.4)	91.2 (97.4)	-
<i>rex</i> (549-bp)	76.9 (61.9)	76.3 (60.6)	75.8 (63.5)	87.6 (89.6)	-
<i>pX</i> (699-bp)	43.3	50.5	49.8	85.6	-
HTLV-4 _{1863LE}					
Genome (5320-bp)	64.0	72.2	71.4	66.2	66.1
<i>pro</i> (273-bp) ²	71.4(55.6)	79.5(28.1)	79.5(36.0)	71.8(29.2)	73.3(31.7)
<i>pol</i> (2549-bp) ²	63.6(68.7)	71.4(80.1)	71.0(79.7)	65.2(71.7)	64.8(71.6)
<i>env</i> (1458-bp)	65.8(75.9)	73.1(85.3)	72.0(85.5)	67.2(78.8)	68.5(79.4)
<i>tax</i> (765-bp) ²	77.4(85.1)	81.7(92.6)	79.4(92.9)	75.2(86.7)	75.0(86.3)
<i>rex</i> (512-bp)	76.0(63.9)	79.5(74.1)	80.7(68.8)	72.5(57.7)	72.7(59.4)
<i>pX</i> (559-bp)	46.1	60.8	59.9	53.6	51.3

1 Amino acid identities are in parentheses.

2 Only partial sequences are available

208. Most notable of the findings was the discovery of a novel human virus that is distinct from all known PTLV lineages with 26-34% and 18-25% nucleotide divergence in the conserved *pol* and *tax* genes, respectively, a range of nucleotide divergence similar to that seen between HTLV-1, HTLV-2, and STLV-3 (Meertens *et al.* 2002)(Table 4). This virus formed a separate phylogenetic lineage with a long branch length and significant bootstrap support in both the *pol* (Figure 2a: *pol* tree) and *tax* trees. Phylogenetic analyses combined with GenBank blast searches show that this is the only known virus in this novel group. For these reasons this virus, which is designate HTLV-4, qualifies as the first member of a novel species in the deltaretrovirus genus. Following ICTV guidelines (van Regenmortel 2000) and pending formal classification, it is proposed that primate T-lymphotropic virus 4 (PTLV-4) be the name for this species, with PTLV-4(1863LE) as the prototype strain. HTLV-4 was found in a 48 year old male hunter (1863LE) from the southern forests of Cameroon who had an HTLV-2-like WB result and reported hunting monkeys, chimpanzees, and gorillas, and also being bitten and scratched by a wild animal, although the animal causing the injury was not specified.

209. Also documented, with significant phylogenetic bootstrap support, is the first evidence of human infection within the PTLV-3 group (Figure 2a: *pol* tree). This virus, which is designated HTLV-3, clusters with STLV-3 viruses present in west African NHPs as expected (Figure 2d: LTR tree). HTLV-3 was found in a 63 year old male (2026ND) from the southern forests of Cameroon who had an HTLV-1-like WB result and who reported hunting and butchering of monkeys. The fact that this virus falls within the diversity of a group of STLVs first identified in 1994 (Goubau et al 1994) without evidence of a human counterpart to date, indicates that this infection was most likely acquired zoonotically through exposure to the blood or body fluids of a hunted NHP from this region (Courgnaud et al. 2004).

210. In addition, broad diversity of HTLV-1 viruses was also found in this collection. Of the 11 HTLV-1 sequences, two did not fall within any of the known HTLV-1 subtypes but clustered clearly within a clade that included only STLV-1 from central and west Africa (Fig 2d: LTR tree). One of these viruses clustered with STLV-1 from monkeys in Cameroon and was from a 65 year-old male (2656ND) from the southern forest zone of Cameroon. He reported hunting and butchering of monkeys and kept a gorilla as a pet (Table 3). The second virus clustered with STLV-1 recently identified in chimpanzees and red colobus monkeys (Leendertz et al. 2004) and was from a 55 year old male (2810YT) who reported hunting and butchering of monkeys (Table 3). The presence of these viruses in hunters, seen previously only among NHPs, indicates that these persons were infected zoonotically. This distinct clade is referred to as HTLV-1 subtype G. Three subjects (1259NG, 1127MO, 1842LE) from different villages were found to have HTLV-1 subtype D, viruses known to infect geographically overlapping populations of humans and mandrills in central Africa (Fig 3 LTR tree (Mahieux et al. 1998). Two of the three viruses were found in hunters (Table 3), providing indirect evidence of cross-species transmission between humans and mandrills within subtype D and supporting further the claims of cross-species transmission of this subtype (Mahieux et al. 1998). These results are consistent with SFV infection from mandrills that was documented previously in this population (Wolfe et al 2004) and indicate that the frequent hunting of mandrills may explain the widespread transmission of mandrill retroviruses. Five persons (979MO, 1380MV, 1443MV, 1503MV, 1537MV) were infected with HTLV-1 subtype B viruses, which are known to be endemic among humans in central Africa and which are believed to have originated from STLV-1 in this region (Mahieux, R. et al 1997, Gessain, A. & Mahieux, R 2000) (Fig 3: LTR tree).

Thus, these five new subtype B viruses may have been acquired either zoonotically from STLV-1-infected primates or from human-to-human transmission, or both. Notably, a 71 year old female (1443MV) who reported butchering gorillas was found to be infected with a virus most closely related to STLV-1 found in two gorillas from Cameroon (Nerrienet 2004, Courgnaud et al 2004), although without significant bootstrap support (Fig. 3 LTR). Interestingly, person 1503MV is also WB positive for SFV (Wolfe et al 2004), indicating that zoonotic transmission in an individual is not limited to a single retrovirus and providing a biological setting for viral recombination and altered pathogenicity and transmissibility of these viruses. One person (2472LE) was infected with the HTLV-1 subtype A virus, a clade consisting of sequences from only globally disseminated HTLV-1 and thus this infection was most likely acquired through human-to-human transmission. DNA samples from the remaining 73 persons with reactive WB results were all negative by the generic PCR assay for *tax* sequences and four other sequences specific for each PTLV clade, including HTLV-4. The results presented in this study demonstrate that HTLV diversity is far greater than previously understood. The data indicates that contact with the blood and body fluids of NHPs is a major factor in the emergence of novel HTLVs, which are known to be transmissible among humans and have the potential to cause disease. Because the hunting and butchering of wild NHPs is widespread throughout central Africa (Bowen-Jones & Pendry 1999) and STLVs are known to be highly prevalent among hunted NHPs (Courgnaud et al. 2004), it is suspected that zoonotic transmission of STLV is not a restricted risk. Since blood banks in central Africa do not generally screen for HTLV, further spread of these viruses among central Africans may be facilitated by blood donations from infected persons. While a recent simian origin could be recognized for at least 6 of the 13 viruses identified, the origin of HTLV-4 is presently unclear since this virus does not have a known primate counterpart. The possibility that HTLV-4 represents a previously unrecognized virus being transmitted between humans indicates that more substantial screening for this virus in central African populations is needed. The finding that both HTLV-4 and HTLV-3 are serologically indistinguishable from HTLV-1 and HTLV-2 in current assays can explain why these viruses have not been previously identified, and highlight the importance of improved diagnostic assays. The increasing evidence that primate hunting is associated with the emergence of a range of simian retroviruses (Wolfe et al. 2004b) calls for increased surveillance and follow-up of individuals exposed to the blood and body fluids of wild NHPs, and for effective strategies to control the hunting of NHPs.

a) Methods

(1) Ethical approvals.

211. Studies were conducted in the context of a community-based HIV prevention campaign designed to provide information using Cameroonian educators and counselors and therefore to decrease transmission. Participation in the study was completely voluntary. The study protocol was approved by the Johns Hopkins Committee for Human Research, the Cameroon National Ethical Review Board, and the HIV Tri-Services Secondary Review Board. Questionnaires and matching samples were anonymized by removing all personal identifiers to provide an unlinked study population.

(2) Sample preparation and serology.

212. Blood was collected from participants, transported to a central laboratory, processed into plasma and PBMC aliquots and stored at -80°C. Initial screening for HTLV antibodies in serum and plasma samples was performed by using the Vironostika HTLV-1/2 microelisa system (Organon-Teknika, Durham, N.C.) following the manufacturer's instructions. Reactive samples were then tested in a WB test (HTLV Blot 2.4, Genelabs Diagnostics, Singapore) that contains disrupted HTLV-1 virions, a gp21 recombinant protein (GD21) common to both HTLV-1 and HTLV-2, and two HTLV-type specific recombinant envelope (*Env*) peptides, MTA-1 and K55, which allow serological differentiation of HTLV-1 and HTLV-2, respectively. Samples with reactivity to the *Gag* (p24) and *Env* (GD21) proteins were considered seropositive. Seropositive samples with reactivity to MTA-1 or K55 were considered HTLV-1-like or HTLV-2-like, respectively. Samples with reactivity to either p24 or GD21 alone or in combination with other HTLV proteins (Fig. 1) were considered indeterminate.

(3) PCR and sequence analysis.

213. DNA was prepared from uncultured PBMCs and its integrity was confirmed by β -actin PCR as previously described. All DNA preparation and PCR assays were performed in a laboratory where only human samples are processed and tested following recommended precautions to prevent contamination. DNA samples were first screened with a generic PTLV *tax* PCR assay capable of detecting 222-bp sequences from each of the three major PTLV groups (Busch *et al.* 2000, van Dooren *et al.* 2004). Sequence analysis of this *tax* sequence provided broad genetic classification into each PTLV group. Phylogenetic resolution within the PTLV-1 and PTLV-3 groups was done using LTR sequences as described previously (van Dooren *et al.* 2004, Meertens *et al.* 2001). A portion of the 3'

HTLV-1 LTR from selected samples (1259NG, 1127MO, 1842LE, and 2810YT) was amplified by nested PCR using external primers 5VLTRext 5' AACCAACCCATTTCCTCCCCATG 3' (SEQ ID NO: 19)(Meertens *et al.* 2001) and 1MNDR1 5'GTCGTGAATGAAAGGGAAAGGGGT 3' (SEQ ID NO: 20) (Meertens *et al.* 2001) and the internal primers Enh280 5' TGACGACAACCCCTCACCTCAA 3' (SEQ ID NO: 21) (Meertens *et al.* 2001) and 1MNDR2 5' AGGGGTGGAACTTTCGATCTGTAA 3'(SEQ ID NO: 22) (Meertens *et al.* 2001). The *tax* (577-bp) and polymerase (*pol*) (709-bp) sequences of HTLV-3 and HTLV-4 were amplified by nested PCR using primers designed from conserved PTLV regions. The external and internal *tax* primers are PTLVTPG 5'T(C/T)ACCT(G/A)GGACCCCATCGATGGACG 3' (SEQ ID NO: 7) and PGTAXR1 5' GAIGA(T/C)TGI A(C/G)TAC(T/C)AAAGATGGCTG 3' (SEQ ID NO: 8) and PH2Rrev 5' CCTTATCCCTCGICTCCCCTC CTT 3' (SEQ ID NO: 9) and PGTAXR2 5' TTIGGG(T/C)AIGGICCGG AAATCAT 3'(SEQ ID NO: 10), respectively. The external and internal *pol* primers are PGPOLF1 5' C(T/G)TTAAACCIGA(A/G)CGCCT CCAGGC 3' (SEQ ID NO: 11) and PGPOLR1 GG(T/C)(A/G)TGIA (A/G)CCA(A/G)(A/G)CIAG(T/G)GG CCA 3' (SEQ ID NO: 12) and PGPOLF2 5' AC(T/C)TGGT(C/T)(C/T) (G/C)(G/C)A(A/G)GGCCCTGGAGG 3' (SEQ ID NO: 13) and PGPOLR2 5' G(A/G)(T/C)(A/G)GGIGTIC CTTTIGAGACCCA 3'(SEQ ID NO: 14), respectively. Inosines (I) and wobble bases (N/N) were used to accommodate areas of heterogeneity (Table 5).

TABLE 5: Sequences of primers used for amplifying partial *tax*, envelope (*env*), polymerase (*pol*) and LTR regions of primate T-cell lymphotropic viruses

Name	Primer sequence ^a (5' to 3')	SEQ ID NO.	Location ^b	Expected PCR product size (bp)	Annealing temp (°C), No. of cycles
PH1F	TTGTCATCAGCCCACTTCCCAGG	(SEQ ID NO: 23)	<i>tax</i> , 7243-7262, outer		
PH2R	AAGGAGGGAGTCGAGGGGATAAGG	(SEQ ID NO: 24)	<i>tax</i> , 7478-7455, outer	236	50, 40
PH2F	CCCAGGTTTCGGGCAAAAGCCTTCT	(SEQ ID NO: 25)	<i>tax</i> , 7257-7280, inner		
PH2R ^c	AAGGAGGGAGTCGAGGGGATAAGG	(SEQ ID NO: 26)	<i>tax</i> , 7478-7455, inner	222	50, 40
PTLVTPG	T(C/T)ACCT(G/A)GGACCCCATCGATGGACG	(SEQ ID NO: 7)	<i>tax</i> , 7480-7504, outer		
PGTAXR1	GAIGA(T/C)TGIA(C/G)TAC(T/C)AAAGATGGCTG	(SEQ ID NO: 8)	<i>tax</i> , 8140-8115, outer	660	45, 40
PH2R ^{rev}	CCTTATCCCTCGICTCCCCTCCTT	(SEQ ID NO: 9)	<i>tax</i> , 7529-7552, inner		
PGTAXR2	TTIGGG(T/C)AIGGICCGGAAATCAT	(SEQ ID NO: 10)	<i>tax</i> , 8106-8085, inner	577	45, 40
PGPOLF1	C(T/G)TTAAACCIGA(A/G)CGCCTCCAGGC	(SEQ ID NO: 11)	<i>pol</i> , 2611-2634, outer		
PGPOLR1	GG(T/C)(A/G)TGIA(A/G)CCA(A/G)(A/G)CIAG(T/G)GGCCA	(SEQ ID NO: 12)	<i>pol</i> , 3598-3575, outer	987	45, 40
PGPOLF2	AC(T/C)TGGT(C/T)(C/T)(G/C)(A/G)GGCCCTGGAGG	(SEQ ID NO: 13)	<i>pol</i> , 2643-2666, inner		
PGPOLR2	G(A/G)(T/C)(A/G)GGIGTICCTTTIGAGACCCA	(SEQ ID NO: 14)	<i>pol</i> , 3352-3329, inner	709	45, 40
PGENVF1	TGGATCCCGTGG(A/C)GI(C/T)TCCTTAA	(SEQ ID NO: 27)	<i>env</i> , 5114-5136, outer		
PGENVR1	GT(A/G)TAIG(C/G)(A/G)(C/G)AIGTCCAIG(A/C)(T/C)TGG	(SEQ ID NO: 28)	<i>env</i> , 5576-5552, outer	462	45, 40
PGENVF2	AIAGACC(T/A)(C/T)CAAC(A/T)CCATGGGTAA	(SEQ ID NO: 29)	<i>env</i> , 5186-5209, inner		
PGENVR2	G(A/C)(T/C)TGGCAICCA(A/G)GTAIGGGCA	(SEQ ID NO: 30)	<i>env</i> , 5557-5535, inner	371	45, 40
GPLTRF1	(G/A)CCACCAICTIGGGACAAATAGCTGA	(SEQ ID NO: 31)	LTR, 8256-8282, outer		
GPLTRR2	C(C/T)GGGCCAAGCCTCGCTGCAGGCA	(SEQ ID NO: 32)	LTR, 8830-8807, outer	575	45, 40
GPLTRF2	ACCHGGCTCTGACGTCTCTCCCT	(SEQ ID NO: 33)	LTR, 8333-8356, inner		
GPLTRR2	GGCAGIAGAACTGCTACTTTCGAT	(SEQ ID NO: 34)	LTR, 8810-8787, inner	478	45, 40

^a Inosines and wobble nucleotides were included in the primers to accommodate sequence heterogeneity.

^b The positions of the *pol*, *env*, and *tax* primers are given according to human T-cell lymphotropic virus type 1 (strain ATK); the LTR primer positions are given according to the simian T-cell lymphotropic virus type 3 (strain PH969) genome.

^c The primer PH2R is used with PH2F in a semi-nested PCR.

Additional diagnostic PCR with PTLV-specific primers was done on samples with negative results for the generic 222-bp *tax* fragments. Assays described previously were used for PTLV-1 *env* and STLTV-3 LTR (van Dooren et al. 2004) and HTLV-2 *env* (Switzer et al. 1995). For HTLV-4, a new nested PCR assay was developed based on the HTLV-4 *tax* sequence using the external primers 1863TF1 5' CTCCTTCTTTCAGTCCGTGCGGAG 3' (SEQ ID NO: 15) and 1863TR1 5' GGGGTAGTCAGGTTTGGCTGGTAT 3' (SEQ ID NO: 16) and the internal primers 1863TF2 5' CCTACCGCAACGGATGTCTTGAAA 3' (SEQ ID NO: 17) and 1863TR2 5' TATGGCGCC GGTGTGATGATAAAG 3' (SEQ ID NO: 18) and standard conditions to generate a 275-bp fragment. Percent nucleotide divergence was calculated using the Gap program in the Genetic Computer Group's Wisconsin package. Sequences were aligned using the Clustal W program, gaps were removed, and distance-based trees were generated by using the Kimura two-parameter model in conjunction with the NJ method in the MEGA program (version 2.1) as described elsewhere (van Dooren *et al* 2004). 1000 bootstrap replicates were used to test the reliability of the final topology of the trees.

(4) Primate taxonomic nomenclature.

214. Nomenclature used herein was as described. NHPs were coded using the first letter of the genus and the first two letters of the species names with their house names or codes within parentheses. Cmo=*Cercopithecus mona* (Mona monkey), Cne=*C. neglectus* (De Brazza's guenon), Cmi=*C. mitis* (Sykes's monkey), Cni=*C. nictitans* (greater spot-nosed guenon), Cae=*Chlorocebus species* (African green monkey), Cpo=*C. pogonias* (crowned monkey), Cto=*Cercocebus torquatus* (red-capped mangabey), Cag=*Cercocebus agilis* (agile mangabey), Mog=*Miopithecus ogouensis* (talapoin monkey), Ani=*Allenopithecus nigroviridis* (Allen's swamp monkey), Msp=*Mandrillus sphinx* (mandrill (mnd)), Pan=*Papio anubis* (olive baboon (bab)), Pcy=*P. cynocephalus* (yellow baboon), Pha=*P. hamadryas* (sacred baboon), Ppu=*P. ursinus* (chacma baboon), Ppa=*P. papio* (Guinea baboon), Pba=*Piliocolobus badius* (red colobus monkey), Mto=*Macaca tonkeana* (Celebes macaque), Ptr=*Pan troglodytes* (chimpanzee), Ppn=*Pan paniscus* (bonobo), Ggo=*Gorilla gorilla* (western lowland gorilla).

(5) Nucleotide sequence accession numbers.

215. The GenBank accession numbers for the 28 new HTLV sequences include AY818406 and AY818433.

2. Example 2: Ancient Origin and Molecular Features of the Novel Human T-lymphotropic virus type 3 Revealed by Complete Genome Analysis

a) Results

(1) Comparison of the HTLV-3(2026ND) proviral genome with prototypical PTLVs.

216. Using a combination of primers designed from small sequences obtained in each of the three major genes of PTLV and the LTR region the complete genome of HTLV-3(2026ND) was successfully generated as depicted in Fig. 4. Sequence analysis of the overlapping regions, followed by comparison with the genetic structure of other PTLVs, demonstrated that the complete proviral genome of HTLV-3(2026ND) is 8917-bp. Despite being genetically equidistant from HTLV-1 and HTLV-2, the genomic structure of HTLV-3(2026ND) was similar to that of other PTLVs and included the structural, enzymatic, and regulatory proteins all flanked by long terminal repeats (LTRs). Comparison of HTLV-3(2026ND) with prototypical PTLV genomes demonstrates that this new human virus is equidistant from the PTLV-1 (62% identity) and PTLV-2 (63% identity) groups across the genome. The results also confirm that HTLV-3 has the closet nucleotide and protein sequence identity to STL-3 (87-92% identity) (Table 6).

Table 6. Percent Nucleotide and Amino Acid Identity of HTLV-3(2026ND) with other PTLV Prototypes¹

	HTLV-1 (ATK)	HTLV-2 (MoT)	STLV-2 (PP1664)	STLV-3 (PH969)	STLV-3 (PPAF3)	STLV-3 (CTO604)	STLV-3 (NG409)
Genome	61.6	62.9	62.6	86.7	92.0	88.4	90.6
LTR	48.7	43.7	41.4	86.2	91.1	86.9	86.9
<i>gag</i>	69.3 (83.2)	69.4 (80.5)	70.6 (80.7)	86.4 (95.5)	91.3 (97.6)	89.4 (96.2)	90.6 (96.7)
p19	(74.4)	(68.3)	(67.2)	(95.9)	(95.9)	(95.9)	(94.3)
p24	(90.1)	(90.1)	(90.6)	(98.1)	(99.1)	(98.6)	(99.1)
p15	(78.0)	(73.8)	(72.6)	(88.4)	(96.5)	(90.7)	(94.2)
<i>pro</i>	59.7 (62.6)	59.2 (66.7)	59.4 (59.3)	83.3 (87.0)	88.8 (91.5)	85.0 (89.3)	88.0 (90.4)
<i>pol</i>	62.2 (66.2)	63.9 (71.2)	63.5 (69.9)	86.1 (92.7)	92.6 (94.9)	88.4 (92.9)	92.0 (92.9)
<i>env</i>	65.9 (73.8)	69.0 (78.2)	67.1 (77.4)	88.1 (95.1)	92.3 (95.1)	88.4 (94.3)	91.2 (95.3)
SU ²	(68.4)	(70.7)	(69.7)	(92.7)	(97.1)	(92.4)	(94.0)
TM ²	(83.5)	(91.6)	(91.0)	(99.4)	(98.9)	(97.8)	(97.8)
<i>rex</i>	76.9 (61.9)	76.3 (60.6)	75.8 (63.5)	87.1 (88.5)	90.9 (94.5)	88.5 (94.0)	88.3 (92.3)
<i>tax</i>	75.4 (81.4)	73.1 (83.4)	72.3 (80.4)	90.2 (97.4)	94.0 (98.3)	91.4 (96.6)	92.8 (96.9)

¹ amino acid identity in parentheses; strain names given in parentheses below PTLV designation

² SU, surface protein; TM, transmembrane protein

The most genetic divergence between the PTLV groups was seen in the LTR region (52-59%) while the highest inter-group identity was observed in the highly conserved regulatory genes, *tax* and *rex* (72-77%). Interestingly, within the PTLV-3 group, HTLV-3(2026ND), which was identified in a hunter from Cameroon, was unique but shared the most overall sequence identity to STLTV-3(PPAF3) (92%) from a Senegalese baboon instead of STLTV-3(CTO604) (88.4%) identified in red-capped mangabeys, also from Cameroon. This relationship is highlighted further by comparison of HTLV-3(2026ND) with all available full-length STLTV-3 genomes in similarity plot analysis where the highest identity was seen in the highly conserved *tax* gene. As seen within other PTLV groups, there was no clear evidence of genetic recombination of HTLV-3(2026ND) with STLTV-3 or PTLV-1 and PTLV-2 proviral sequences by using bootscanning analysis. HTLV-3(2026ND) was not compared to the recently reported second strain of HTLV-3 because only two short sequences were available at GenBank and in these region this virus has been shown to be nearly identical to STLTV-3(CTO604) (Callatini, S., et al. (2005) *Retrovirology*. 2:30).

**(2) Organization of the LTR and pre-gag region. **

217. As with STLTV-3, the HTLV-3(2026ND) LTR (697-bp) was smaller than that of HTLV-1 (756-bp) and HTLV-2 (764-bp), by having two and not three of the 21-bp transcription regulatory repeat sequences in the U3 region (Fig. 5a) (Meertens, L., and A. Gessain. (2003) *J. Virol.* 77:782-789; Meertens, L., et al. (2002) *J. Virol.* 76:259-268; Van Brussel, M., et al. (1997) *J. Virol.* 71:5464-5472; Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519). Other regulatory motifs such as the polyadenylation signal, TATA box, and cap site were all conserved in the HTLV-3(2026ND) LTR (Fig. 5a). By secondary structure analysis of the LTR RNA sequence, a stable stem loop structure from nucleotides 421- 464 (Fig. 5b) was also observed similar to that shown to be essential for Rex-responsiveness control of viral expression in both HTLV-1 and HTLV-2.

(3) Analysis of the genomic structure of HTLV-3(2026ND).

218. Translation of predicted protein open reading frames (ORFs) across the viral genome identified all major Gag, Pol, Pro (protease), and Env proteins, as well as the regulatory proteins, Tax and Rex. Translation of the overlapping *gag* and *pro* and *pro* and *pol* ORFs occurs by one or more successive -1 ribosomal frameshifts that align the different ORFs. The conserved slippage nucleotide sequence 6(A)-8nt-6(G)-11nt-6(C) is present in the Gag-Pro overlap, while a point mutation in the Pro-Pol overlap slippage sequence

(GTTAAAC compared to TTTAAAC in HTLV-1 and HTLV-2) was observed in HTLV-3(2026ND) but the asparagine codon (AAC) crucial for the slippage mechanism was unaffected.

219. The structural and group specific precursor Gag protein consisted of 422 amino acids (aa) that is predicted to be cleaved into the three core proteins p19 (matrix), p24 (capsid), and p15 (nucleocapsid) similar to HTLV-1, HTLV-2, and STL-3. Across PTLVs, Gag was one of the most conserved proteins with identities ranging from 81% and 83% for HTLV-1 and PTLV-2, to 95% for STL-3 supporting the observed cross-reactivity seen with PTLV-3 antisera in Western blot assays using HTLV-1 antigens. Within Gag, the capsid protein showed greater than 90% identity to HTLV-1, while the matrix and nucleocapsid proteins were more divergent sharing less than 78% identity to PTLV-1 and PTLV-2 indicating their potential use in serologic assays for discriminating the three major PTLV groups.

220. The predicted size of the Env polyprotein is 491 aa, which is slightly shorter than that found in STL-3s (313 aa versus 314 and 315 for STL-3(PH969) and STL-3(CTO-604) due to sequence variation at the carboxy terminus of the surface (SU) protein. In contrast, the transmembrane (TM) protein (178 aa) was highly conserved across all PTLVs supporting further the use of the recombinant HTLV-1 GD21 protein spiked onto WB strips for the identification of divergent PTLVs. Despite the weak reactivity of anti-HTLV-3(2026ND) antibodies to the HTLV-1 type specific SU peptide (MTA-1) Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A.* 102:7994-7999 spiked onto WB strips, there was only 70.8% identity of MTA-1 to HTLV-3(2026ND), which is similar to the 68.8% identity of MTA-1 to HTLV-2, demonstrating no clear correlation of WB profile and predicted SU sequence.

221. The HTLV-1 and HTLV-2 Tax proteins (Tax1 and Tax2, respectively) transactivate initiation of viral replication from the promoter in the 5' LTR and are thus essential for viral expression (Feuer, G., and P. L. Green. (2005) *Oncogene.* 24:5996-6004). Tax1 and Tax2 have also been shown to be important for T-cell immortalization, while the HTLV-3 Tax (Tax3) has not yet been characterized (Feuer, G., and P. L. Green. (2005) *Oncogene.* 24:5996-6004). Hence, the Tax3 sequences were compared with those of prototypic HTLV-1, PTLV-2, and STL-3s to determine if motifs associated with these functional characteristics are preserved. Alignment of predicted Tax3 sequences shows excellent conservation of the critical functional regions, including the nuclear localization

signal (NLS), cAMP response element (CREB) binding protein (CBP)/p300 binding motifs, and nuclear export signal (NES) (Fig. 6). The C-terminal transcriptional activating domain (CR2), essential for CBP/p300 binding, was also very conserved except for a single I/V to F mutation at position five of the motif compared to HTLV-1 and PTLV-2, respectively (Fig. 6). However, this single amino acid change in the STL-3 Tax has recently been shown in transient transfection assays to have no deleterious effect on viral transactivation (Chevalier, S., et al. (2005) *AIDS Res. Hum. Retrovir.* 21:513 (Abs. P174)). Since the predicted CR2 domain is conserved in Tax3, similar transactivation activity can be seen with HTLV-3.

222. Interestingly, although these important functional motifs are highly conserved in PTLV, phenotypic differences of HTLV-1 and HTLV-2 Tax proteins have been observed leading to speculation that these differences account for the different pathologies associated with both HTLVs (Feuer, G., and P. L. Green. (2005) *Oncogene*. 24:5996-6004). Recently, the C-terminus of Tax1, and not Tax2, has been shown to contain a conserved PDZ domain present in cellular proteins involved in signal transduction and induction of the IL-2-independent growth required for T-cell transformation (Rousset, R., et al. (1998) *Oncogene*. 6:643-654; Tsubata, C., et al. (2005) *Retrovirol.* 2:46). The presence of a PDZ domain in PTLV-1 and its absence in PTLV-2 indicates a potential role of this motif in the phenotypic differences of the two viral groups. The consensus PDZ domain has been defined as S/TXV-COOH, where the first amino acid is serine or threonine, X is any amino acid, followed by valine and the carboxy terminus. Examination of the PTLV-3 Tax sequences showed that both HTLV-3 and STL-3 have predicted PDZ domains with the consensus sequence S(P/S)V compared to T(E/D)V in PTLV-1 (Fig. 6).

223. Besides Tax and Rex, two additional ORFs coding for four proteins (p27^I, p12^I, p30^{II}, and p13^{II} where I and II denote ORFI and ORFII, respectively) have been identified in the pX region of HTLV-1 (Fig. 4) and are important in viral infectivity and replication, T-cell activation, and cellular gene expression (Bindhu, M., et al. (2004) *Front. Biosci.* 9:2556-2576). Analysis of the pX region of HTLV-3(2026ND) revealed a total of four putative ORFs (named I-IV, respectively) coding for 96, 122, 72, and 118 aa in length. While both ORFIII (72 aa) and ORFIV (118 aa) shared identity to the ORFII of STL-3 and HTLV-1 and STL-2/HTLV-2, respectively, and each contained two PXXP motifs, only ORF III was leucine rich like that seen in the leucine zipper motifs of ORFI p12^I (Bindhu, M., et al. (2004) *Front. Biosci.* 9:2556-2576). However, ORFIII did not share any sequence homology with p12^I and both ORFI and ORFII shared only weak sequence identity to

miscellaneous cellular proteins available at GenBank. Interestingly, 22 of 28 (79%) amino acids in ORFIV (pos 64-91) were identical among the ORFIIIs of all PTLVs indicating a conserved functionality of this motif. Analysis of alternatively spliced messenger RNA expression in viable cells or tissue culture, and/or in vitro characterization, will be required to investigate the functionality of these putative accessory proteins.

224. A novel protein termed the HTLV-1 basic leucine zipper ZIP (bZIP) factor (HBZ) was recently identified in translation of the complementary strand of the viral RNA genome between the *env* and *tax/rex* genes (Gaudray, G., et al. (2002) *J Virol.* 76:12813-12822). Although originally reported to be exclusive to PTLV-1 (Gaudray, G., et al. (2002) *J Virol.* 76:12813-12822), HBZ is conserved among PTLVs (Fig. 7), including HTLV-3(2026ND), demonstrating further the potential importance of this protein in viral replication and oncogenesis. The carboxy terminus of the HBZ ORF contains a 21 aa arginine rich region that is relatively conserved in PTLV and known cellular bZIP transcription factors, followed by a leucine zipper region possessing five or four conserved leucine heptads in HTLV-1 and all other PTLVs, respectively (Fig. 7). PTLV-1 has 5 leucine heptads similar to that found in mammalian bZIP proteins, while PTLV-1 and PTLV-2 have four leucine heptads followed by leucine octet (Fig. 7). Of all PTLVs with full length genomes available at GenBank, only HTLV-2(MoT) did not have the full complement of leucine heptads but was limited to the initial three leucine motifs due to a one nucleotide deletion at position 6823 causing a frameshift in the predicted HBZ sequence.

(4) Phylogenetic analysis.

225. The genetic relationship of HTLV-3(2026ND) to PTLV-3 was confirmed by using aligned full-length prototype sequences excluding the LTR region (Fig. 8a). Phylogenetic analysis inferred three major PTLV groups with very high bootstrap support (100%) with HTLV-1, HTLV-2 and HTLV-3 each clustering in separate clades (Fig. 8a). Within the PTLV-3 phylogroup, HTLV-3(2026ND) formed a separate lineage but clustered with high bootstrap support with STL-3s from west central Africa (strains CTO604, CTO-NG409, and PPA-F3) indicating a possible primate origin for this human infection in this geographic region. The relationship of HTLV-3 to STL-3 was supported further by phylogenetic inference of identical tree topologies using an alignment of each major gene region (Fig. 8b-8d). The phylogenetic stability seen across the PTLV genome also demonstrates further the absence of major recombination events occurring in PTLV despite

evidence of dual infections in humans and primates (Courgnaud, V., S. Et al. (2004) *J. Virol.* 78 :4700-4709), compared to other retroviruses such as HIV which undergo frequent recombination.

(5) Dating the origin of HTLV-3(2026ND) and other PTLVs.

226. The finding of HTLVs in three distinct clades indicates an ancient, independent evolution of these viruses. Hence, additional molecular analyses was undertaken in order estimate the divergence times of the PTLV lineages. Although others have reported finding a clock-like behavior of STL-3 sequences (Meertens, L., and A. Gessain. (2003) *J. Virol.* 77:782-789; Meertens, L., et al. (2002) *J. Virol.* 76:259-268; Meertens, L., et al. (2003) *J. Gen. Virol.* 84:2723-2727; 22)), these results were not confirmed and instead found that PTLVs evolved at different rates by using an alignment of full-length PTLV genomes sans LTR sequences. However, reliable retrovirus divergence times can be obtained by using nonparametric rate smoothing of the sequences to relax the stringency of a clock assumption followed by time calibration of the tree using a value of 40,000 – 60,000 YA for the origin of the Melanesian HTLV-1 (Sanderson, M. J. (2003) *Bioinformatics.* 19:301-2; Switzer, W. M., et al. (2005) *Nature.* 434:376-380; Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519). By using these dates and methods, the mean evolutionary rate for PTLV was estimated to be 1.12×10^{-6} (confidence interval 6.82×10^{-7} to 1.56×10^{-6}) substitutions/site/ year, respectively, which is consistent with rates determined previously both with and without enforcing a molecular clock (Lemey, P., et al. (2005) *Infect. Gen. Evol.* 5:291-298; Meertens, L., and A. Gessain. (2003) *J. Virol.* 77:782-789; Meertens, L., et al. (2002) *J. Virol.* 76:259-268; Meertens, L., et al. (2003) *J. Gen. Virol.* 84:2723-2727; Salemi, M., et al. (2000) *Mol. Biol. Evol.* 17:374-386; Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519). The mean evolutionary rate for HTLV-3(2026ND) is estimated to be 9.94×10^{-7} (confidence interval 6.04×10^{-7} to 1.38×10^{-6}). The PTLV ancestor was estimated to have originated about 630,000 – 947,000 YA confirming an archaic evolution of the primate deltaretroviruses (Fig. 9) (Salemi, M., et al. (2000) *Mol. Biol. Evol.* 17:374-386). The separation of PTLV-1 and PTLV-2 occurred about 579,077 – 867,458 YA, while HTLV-2 and STL-2 diverged around 191,621 – 286,730 YA (Fig. 9). The origin of all PTLV-3s was estimated to be between 63,294 - 94,700 YA with the ancestor of HTLV-3(2026ND) occurring about 36,087 – 54,067 YA (Fig. 9) indicating an ancient origin of this virus in humans. Alternatively, HTLV-3 may represent a

recent zoonoses from a primate infected with a very old, divergent STL-3. However, if HTLV-3 is an old human infection, then it appeared during the same period as the ancestor of both HTLV-1 and HTLV-2 (40,000 – 60,000 and 28,800 – 43,392 YA, respectively) and may have also spread to become endemic in specific populations yet to be identified.

b) Discussion

227. The complete nucleotide sequence and genomic characterization of the first HTLV-3 that is clearly distinct from all STL-3s and is genetically equidistant to HTLV-1 and HTLV-2 is described herein. HTLV-3(2026ND) is also unique from the second HTLV-3(Pyl43) reported recently in a Bakola pygmy from Cameroon since the latter strain is nearly identical to STL-3 found in a red-capped mangabey, based on the limited sequence data available for this virus (Callatini, S., et al. (2005) *Retrovirology*. 2:30). Although HTLV-1 and HTLV-2 are pathogenic and have spread globally to become endemic in different human populations, little is known about the epidemiology of HTLV-3 infection. However, detailed, comparative sequence analyses of viral genomes can help provide important molecular clues to the origin, evolution, and public health importance of novel human infections.

228. Like other PTLVs, HTLV-3(2026ND) is genetically stable and its slow evolutionary rate, combined with estimates of known human migrations, can then be used to infer divergence times for HTLV. The finding that the predecessor of HTLV-3(2026ND) originated over 30 millenia ago, an age which is estimated that the ancestors of both HTLV-1 and -2 to have appeared, combined with the wide geographic distribution of STLVs and the recent finding of another HTLV-3 in an African pygmy (Callatini, S., et al. (2005) *Retrovirology*. 2:30; Gessain, A., and R. Mahieux. (2000) *Bull. Soc. Pathol. Exot.* 93:163-171; Meertens, L., and A. Gessain. (2003) *J. Virol.* 77:782-789; Meertens, L., et al. (2002) *J. Virol.* 76:259-268; Meertens, L., et al. (2003) *J. Gen. Virol.* 84:2723-2727; Takemura, T., et al. (2002) *J. Virol.* 76:1642-1648; Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519), collectively indicate that HTLV-3 infection be more frequent than previously understood. In addition, the archaic age of the ancestral HTLVs and the recent finding of STL-like infections in African hunters collectively imply that cross-species transmission of STLVs to humans is both an ancient and contemporary phenomenon coupled to behavior that exposes humans to nonhuman primates. The ancient origin of HTLV contrasts with that reported for HIV, which is believed to have only crossed over into humans from SIV-infected NHPs within the last century, and indicates a long period of viral evolution and

adaptation in humans possibly resulting in the observed lower pathogenicity for HTLV compared to HIV (Hahn, B. H., et al. (2000) *Science* 287:607-614; Sharp, P. M., et al. (2000) *Biochem Soc Trans.* 28:275-282).

229. Screening for HTLV-3 can be facilitated by the application of diagnostic serologic and molecular assays based on the sequences reported here. For example, the data show that the Gag matrix and nucleocapsid regions and the envelope surface protein are relatively conserved within PTLV-3 but are divergent from PTLV-1 and PTLV-2 and can thus be used to differentiate the three PTLV groups with serological methods.

230. At the molecular level, examination of the genomic structure showed that the enzymatic, regulatory, and structural proteins are well preserved in HTLV-3(2026ND), including conserved functional motifs in Tax important for viral expression and T-cell proliferation. The finding of a PDZ domain in the Tax protein of HTLV-3(2026ND), like that seen in HTLV-1 but not HTLV-2 (Feuer, G., and P. L. Green. (2005) *Oncogene*. 24:5996-6004), which has been shown to be important in cellular signal transduction and T-cell transformation (Rousset, R., et al. (1998) *Oncogene*. 6:643-654; Tsubata, C., et al. (2005) *Retrovirol.* 2:46), indicates that the HTLV-3 Tax is more phenotypically similar to HTLV-1 than HTLV-2. The high amino acid identity of the PTLV-3 Tax proteins combined with the ability of STLV-3 to transform human cells in vitro indicates that the HTLV-3 Tax functions similarly (Goubau, P., et al. (1994) *Proc. Natl. Acad. Sci. U S A* 91:2848-2852).

231. In contrast to the *tax* gene, the HTLV-3(2026ND) LTR has only two of the three conserved promoters identified in HTLV-1 and HTLV-2 that are responsible for basal viral transcription levels and like STLV-3 is missing the TATA-distal 21-bp repeat element (Meertens, L., and A. Gessain. (2003) *J. Virol.* 77:782-789; Meertens, L., et al. (2002) *J. Virol.* 76:259-268; Meertens, L., et al. (2003) *J. Gen. Virol.* 84:2723-2727; Van Brussel, M., et al. (1997) *J. Virol.* 7:5464-5472; Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519). All of the remaining functional elements in the LTR were conserved, including the stem loop structure necessary for Rex responsive control of viral expression in HTLV-1 and -2.

232. Recently, a novel HBZ protein was identified in translation of the complementary strand of the viral RNA genome between the *env* and *tax/rex* genes (Gaudray, G., et al. (2002) *J Virol.* 76:12813-12822). Protein translation on the minus-strand RNA is a unique feature of HTLV-1 not previously seen in retroviruses. HBZ was shown to be involved in the negative regulation of viral replication (Gaudray, G., et al.

(2002) *J Virol.* 76:12813-12822). The more recent finding of HBZ mRNA expression in ATL patients indicates a role of HBZ mRNA in the survival of leukemic cells *in vivo* and in HTLV-1-associated oncogenesis (Satou, Y., et al. (2006) *Proc. Natl. Acad. Sci. USA.* 103:720-725). Although originally reported to be exclusive to PTLV-1 (Gaudray, G., et al. (2002) *J Virol.* 76:12813-12822), HBZ is conserved among PTLVs, including HTLV-3(2026ND), demonstrating further the potential importance of this protein in viral replication and oncogenesis. Of all PTLVs with full length genomes available at GenBank, only HTLV-2(MoT) did not have the full complement of leucine heptads in the leucine zipper due to a frameshift mutation in the predicted HBZ sequence.

233. In summary, disclosed herein, the novel and unique HTLV-3(2026ND) is genetically stable and has an ancient origin. HTLV-3(2026ND) genomic structure is relatively conserved and contains many of the functional motifs important for the viral expression and pathology associated with HTLVs.

c) Materials and Methods

(1) DNA Preparation and PCR-based Genome Walking.

234. DNA was prepared from uncultured PBMCs available from person 2026ND identified in the original PTLV surveillance study in Cameroon reported in detail elsewhere (Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A.* 102:7994-7999). DNA integrity was confirmed by β -actin polymerase chain reaction (PCR) as previously described (Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A.* 102:7994-7999). All DNA preparation and PCR assays were performed in a laboratory where only human specimens are processed and tested according to recommended precautions to prevent contamination. To obtain the full-length genomic sequence of HTLV-3 small regions of each major coding region were PCR-amplified by using nested PCR and degenerate PTLV primers. The *tax* (577-bp) and polymerase (*pol*) (709-bp) sequences were amplified by using primers and conditions provided elsewhere (Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A.* 102:7994-7999). Envelope (*env*) (371-bp) sequences were amplified by using standard PCR conditions with a 45°C annealing temperature and the external and internal primers PGENVF1 5' TGGATCCCGTGG(A/C)GI(C/T)TCCTIAA 3' (SEQ ID NO: 27) and PGENVR1 5' GT(A/G)TAIG(C/G)(A/G)(C/G)AIGTCCAIG(A/C)(T/C)TGG 3' (SEQ ID NO: 28) and PFENVF2 5' AIAAGACC(T/A)(C/T)CAAC(A/T)CCATGGGTAA 3' (SEQ ID NO: 29) and PGENVR2 5' G(A/C)(T/C)TGGCAICCA(A/G)GTAIGGGCA 3' (SEQ ID NO: 30), respectively. A 398-bp fragment of the long terminal repeat (LTR) was obtained

by using conserved STL-3 primers as previously reported (Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A.* 102:7994-7999).

235. HTLV-3(2026ND)-specific primers were then designed from sequences obtained in each of the four viral regions described above and were used in nested, long-template PCRs to fill in the gaps in the genome as depicted in Fig. 4 by using an expand high fidelity kit containing both Taq and Tgo DNA polymerases (Roche). The external and internal primer sequences for the LTR-*pol* and *pol-env* fragments are 2026LF1 5' GGTAAGATCCCCTGGGTCGAGC 3' (SEQ ID NO: 69) and 2026PR1 5' GAAGCCAGGTCTCGGGTGACG 3' (SEQ ID NO: 70) and 2026LF2 5' CGCTCCCCTGGAGCTCTCTCG 3' (SEQ ID NO: 71) and 2026PR2 5' GCCACTTCCCATTGGGCTTTTGTGACGG 3' (SEQ ID NO: 72) and 2026PF3 5' GCTCTCACCGATAAAGTAACAAACG 3' (SEQ ID NO: 73) and 2026ER1 5' GGTAAGGAAGAGGCTCCTATGAACAG 3' (SEQ ID NO: 74) and 2026PF2 5' CAGGACTGCATAACATACGAGACCCTCC 3' (SEQ ID NO: 75) and 2026ER3 5' CCTATGAACAGGGTGCATCGACTGGG 3' (SEQ ID NO: 76), respectively. The external and internal primer sequences used to obtain about 3 kb of the 3' end of the genome (*env-tax*-LTR) are 2026EF1 5' CCTAAGCCCCCATGTCCAGAC 3' (SEQ ID NO: 77) and 2026LR1 5' CGAGAGAGCTCCAGGGGAGCG 3' (SEQ ID NO: 78) and 2026EF3 5' CCTACTCCCTGTATGTATTCCCCCATTTGG 3' (SEQ ID NO: 79) and 2026LR2 5' GCTCGACCCAGTGGGATCTTACCGAGTGG 3' (SEQ ID NO: 80), respectively.

236. PCR products were revealed on 1.5% agarose gels stained with ethidium bromide, purified with a Qiaquick PCR purification kit (Qiagen) and sequenced in both directions with a BigDye terminator cycle kit and automated sequencers (Applied Biosystems). Selected PCR products were also cloned into the pCR4-TOPO vector using the TOPO TA Cloning kit (Invitrogen) and recombinant plasmid DNA was prepared using the Qiagen plasmid purification kit prior to automated sequencing.

(2) Sequence and phylogenetic analysis.

237. Percent nucleotide divergence was calculated by using the GAP program in the Genetic Computer Group's (GCG) Wisconsin package (Thompson, J. D., et al. (1994) *Nucleic Acids Res.* 22:4673-4680). LTR RNA secondary structure was determined using the program RNAstructure v4.2 (Mathews, D. et al. (1999) *J. Mol. Biol.* 288:911-940). Sequences were aligned by using the Clustal W program (Womble, D. D. (2000) *Methods*

Mol. Biol. 132:3-22), gaps were removed, and distance-based trees were generated by using the Kimura two-parameter model in conjunction with the neighbor-joining (NJ) method in the MEGA program (version 2.1) and maximum likelihood (ML) analysis in the PAUP* program as described in detail elsewhere (Switzer, W. M., et al. (2005) *Nature*. 434:376-380; Wolfe, N. D., et al. (2005) *Proc. Natl. Acad. Sci. U S A*. 102:7994-7999). The reliability of the final topology of the trees was tested with 1,000 bootstrap replicates. Comparison of full-length PTLV genomes available at GenBank was done using HTLV-3(2026ND) as the query sequence and the F84 (ML) model and a transition/transversion ratio of 2.0 implemented in the program SimPlot (Lole, K. S., et al. (1999) *J. Virol.* 73:152-160).

238. For dating of HTLV-3(2026ND), full-length genomes from prototypical PTLVs available at GenBank were aligned with HTLV-3(2026ND) by using Clustal W, gaps were removed, and minor adjustments in the alignment were made manually. LTR sequences were excluded from the analysis since this region does not align accurately in PTLVs. The best fitting evolutionary model for the aligned sequences was determined with Modeltest v3.6 (Posada, D., and K. A. Crandall. (1998) *Bioinformatics*. 14:817-818). The general time-reversible model, allowing six different substitution rate categories, with gamma-distributed rate heterogeneity (1.9724) and an estimated proportion of invariable sites (0.3687), was determined to be the best fit to the data. Little substitution saturation was observed in the 7213-bp alignment ($P < 0.0001$) as determined with the DAMBE program, and was therefore satisfactory for use in phylogenetic analyses. Likewise, using the best-fitting evolutionary model defined above, good phylogenetic signal in the alignment was also found with likelihood mapping analysis using the program Tree-Puzzle v5.2.

239. The molecular clock hypothesis, or constant rate of evolution, was tested by using the likelihood ratio test with the likelihoods for the ML and clock-like ML trees obtained in PAUP*. The clock was tested with the best-fitting evolutionary model estimated in Modeltest, and ML trees were constructed in PAUP* starting from the NJ tree that is iteratively optimized using two consecutive heuristic searches with nearest neighbor interchange followed by a final heuristic search with the tree-bisection-reconnection algorithm. To adjust for rate heterogeneity among different PTLV taxa, clock-like ML trees were then transformed into ultrametric trees using the nonparametric rate smoothing (NPRS) algorithm in the program TreeEdit (v1.0a10 carbon) (Sanderson, M. J. (2003) *Bioinformatics*. 19:301-2). The branches of the NPRS tree were then scaled by using a

divergence time of 40,000 – 60,000 years ago (ya) for the Melanesian HTLV-1mel lineage based on genetic and archaeological evidence of when the ancestors of indigenous Melanesians and Australians migrated from Southeast Asia (Lemey, P., et al. (2005) *Infect. Gen. Evol.* 5:291-298; Salemi, M., et al. (2000) *Mol. Biol. Evol.* 17:374-386; Salemi, M., et al. (1999) *AIDS Rev.* 1:131-139). Variance in age estimates (branch lengths) was determined in PAUP* with 100 bootstrap repetitions by enforcing topological constraints and using a heuristic search without branch swapping on the clock-like ML tree. Branch lengths in all 100 trees were calibrated as before and average divergence times and confidence intervals ($\alpha = 0.05$) were calculated in Excel. The evolutionary rate was estimated based on a known divergence time point of 40,000 – 60,000 ya and on the branch length of the ML clock-like tree according to the formula: evolutionary rate (r) = branch length (bl)/divergence time (t) (Van Dooren, S., et al. (2004) *J. Gen. Virol.* 85:507-519).

(3) Nucleotide sequence accession number.

240. The HTLV-3(2026ND) proviral sequence has the GenBank accession number DQ093792.

G. References

- Araujo, A. & Hall, W.W. Human T-lymphotropic virus type II and neurological disease. *Ann. Neurol.* 56, 10-19 (2004).
- Barnhart, M. K., L. M. Connor, S. J. Marriott. 1997. Function of the human T-cell leukemia virus type 1 21-base-pair repeats in basal transcription. *J. Virol.* 71:337-344.
- Bindhu, M., A. Nair, and M. D. Lairmore. 2004. Role of accessory proteins of HTLV-1 in viral replication, T cell activation, and cellular gene expression. *Front. Biosci.* 9:2556-2576.
- Bowen-Jones, E. & Pendry, S. The threat to primates and other mammals from the bushmeat trade in Africa, and how this threat could be diminished. *Oryx* 33, 233-246 (1999).
- Busch, M.P., Switzer, W.M., Murphy E., L. & Heneine, W. and the Retrovirus Epidemiology donor study (REDS). Absence of evidence of infections with divergent simian T-lymphotropic viruses in US blood donors with seroindeterminate human T-lymphotropic results. *Transfusion* 40, 443-449 (2000).
- Callatini, S., S. A. Chevalier, R. Duprez, S. Bassot, A. Froment, R. Mahieux, and A. Gessain. 2005. Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. *Retrovirology*. 2:30.
- Chevalier, S., L. Meertens, C. Pise-Masison, J. Brady, A. Gesain, and R. Mahieux. 2005. Molecular characterization of the Tax protein from the highly divergent simian T-cell lymphotropic virus type 3 strain. *AIDS Res. Hum. Retrovir.* 21:513 (Abs. P174).

- Courgnaud, V., Van Dooren, S., Liegeois, F., Pourrut, X., Abela, B., Loul, S., Mpoudi-Ngole, E., Vandamme, A., Delaporte, E. & Peeters, M. Simian T-cell leukemia virus (STLV) infection in wild primate populations in Cameroon: evidence for dual STLV type 1 and type 3 infection in agile mangabeys (*Cercocebus agilis*). *J. Virol.* 78, 4700-4709 (2004).
- Digilio, L., A. Giri, N. Cho, J. Slattery, P. Markham, G. Franchini. 1997. The simian T-lymphotropic/leukemia virus from *Pan paniscus* belongs to the type 2 family and infects Asian macaques. *J Virol.* 71:3684-3692.
- Feuer, G., and P. L. Green. 2005. Comparative biology of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2. *Oncogene.* 24:5996-6004.
- Gaudray, G., F. Gachon, J. Basbous, M. Biard-Piechaczyk, C. Devaux, and J. M. Mesnard. 2002. The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. *J Virol.* 76:12813-12822.
- Gessain, A. & Mahieux, R. Epidemiology, origin and genetic diversity of HTLV-1 retrovirus and STLV-1 simian affiliated retrovirus. *Bull. Soc. Pathol. Exot.* 93, 163-171 (2000).
- Goubau, P., Van Brussel, M., Vandamme, A.M., Liu, H.F. & Desmyter, J. A primate T-lymphotropic virus, PTLV-L, different from human T-lymphotropic viruses types I and II, in a wild-caught baboon (*Papio hamadryas*). *Proc. Natl. Acad. Sci. U S A* 91, 2848-2852 (1994).
- Groves, C. *Primate Taxonomy* (Smithsonian Institution Press, Washington, D.C., 2001).
- Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp. 2000. AIDS as a zoonosis: scientific and public health implications. *Science* 287:607-614.
- Heneine, W., Switzer, W.M., Sandstrom, P., Brown, J., Vedapuri, S., Schable, C.A., Khan, A.S., Lerche, N.W., Schweizer, M., Neumann-Haefelin, D., Chapman, L.E. & Folks, T.M. Identification of a human population infected with simian foamy viruses. *Nat. Med.* 4, 403-407 (1998).
- Khabbaz, R.F., Heneine, W., George, J.R., Parekh, B., Rowe, T., Woods, T., Switzer, W.M., McClure, H.M., Murphey-Corb, M., & Folks, T.M. Brief report: infection of a laboratory worker with simian immunodeficiency virus. *N. Engl. J. Med.* 330, 172-177 (1994).
- Leendertz, F.H., Junglen, S., Boesch, C., Formenty, P., Couacy-Hymann, E., Courgnaud, V., Pauli, G. & Ellerbrok, H. High variety of different simian T-cell leukemia virus type 1 strains in chimpanzees (*Pan troglodytes verus*) of Tai National Park, Cote d'Ivoire. *J. Virol.* 78, 4352-4356 (2004).
- Lemey, P., O. G. Pybus, S. Van Dooren, and A. M. Vandamme. 2005. A Bayesian statistical analysis of human T-cell lymphotropic virus evolutionary rates. *Infect. Gen. Evol.* 5:291-298.
- Lerche, N.W., Switzer, W.M., Yee, J.L., Shanmugam, V., Rosenthal, A.N., Chapman, L.E.,

- Folks, T.M. & Heneine W. Evidence of infection with simian type D retrovirus in persons occupationally exposed to nonhuman primates. *J. Virol.* 75, 1783-1789 (2001).
- Lole, K. S., R. C. Bollinger R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73:152-160.
- Mahieux, R., Chappey, C., Georges-Courbot, M.C., Dubreuil, G., Mauclore, P., Georges, A. & Gessain, A. Simian T-cell lymphotropic virus type 1 from *Mandrillus sphinx* as a simian counterpart of human T-cell lymphotropic virus type 1 subtype D. *J. Virol.* 72, 10316-10322 (1998).
- Mahieux, R., Ibrahim, F., Mauclore, P., Herve, V., Michel, P., Tekiaia, F., Chappey, C., Garin, B., Van Der Ryst, E., Guillemain, B., Ledru, E., Delaporte, E., de The, G. & Gessain, A. Molecular epidemiology of 58 new African human T-cell leukemia virus type 1 (HTLV-1) strains: identification of a new and distinct molecular subtypes in central Africa and in pygmies. *J. Virol.* 71, 1317-1333 (1997).
- Mahieux, R., P. Horal, P. Mauclore, O. Mercereau-Puijalon, M. Guillotte, L. Meertens, E. Murphy, and A. Gessain. 2000. Human T-cell lymphotropic virus type 1 gag indeterminate western blot patterns in Central Africa: relationship to *Plasmodium falciparum* infection. *J. Clin. Microbiol.* 38 :4049-4057.
- Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 288:911-940.
- Meertens, L., and A. Gessain. 2003. Divergent simian T-cell lymphotropic virus type 3 (STLV-3) in wild-caught *Papio hamadryas papio* from Senegal: widespread distribution of STLV-3 in Africa. *J. Virol.* 77:782-789.
- Meertens, L., Mahieux, R., Mauclore, P., Lewis, J. & Gessain, A. Complete sequence of a novel highly divergent simian T-cell lymphotropic virus from wild-caught red-capped mangabeys (*Cercocebus torquatus*) from Cameroon: a new primate T-lymphotropic virus type 3 subtype. *J. Virol.* 76, 259-268 (2002).
- Meertens, L., Rigoulet, J., Mauclore, P., Van Beveren, M., Chen, G.M., Diop, O., Dubreuil, G., Georges-Goubot, M.C., Berthier, J.L., Lewis, J. & Gessain, A. Molecular and Phylogenetic Analyses of 16 Novel Simian T Cell Leukemia Virus Type 1 from Africa: Close Relationship of STLV-1 from *Allenopithecus nigroviridis* to HTLV-1 Subtype B Strains. *Virology* 287, 275-285 (2001).
- Meertens, L., V. Shanmugam, A. Gessain, B. E. Beer, Z. Tooze, W. Heneine, and W. M. Switzer. 2003. A novel, divergent simian T-cell lymphotropic virus type 3 in a wild-caught red-capped mangabey (*Cercocebus torquatus torquatus*) from Nigeria. *J. Gen. Virol.* 84:2723-2727.
- Milner-Gulland, E.J., Bennett, E.L. & the SCB 2002 Annual Meeting Wild Meat Group Wild meat: the bigger picture. *TREE* 18, 351-357 (2003).

Nerrienet E, Meertens L, Kfutwah A, Foupouapouognigni Y, Ayouba A, Gessain A. Simian T cell leukaemia virus type I subtype B in a wild-caught gorilla (*Gorilla gorilla gorilla*) and chimpanzee (*Pan troglodytes vellerosus*) from Cameroon. *J Gen Virol.* 85: 25-9 (2004).

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 14:817-818.

Rousset, R., S. Fabre, C. Desbios, F. Bantignies, and P. Jalinot. 1998. The C-terminus of the HTLV-1 Tax oncoprotein mediates interaction with the PDZ domain of cellular proteins. *Oncogene.* 6:643-654.

Salemi, M., Desmyter, J. & Vandamme, A.M. Tempo and mode of human and simian T-lyphotropic virus (HTLV/STLV) evolution revealed by analyses of full-genome sequences. *Mol. Biol. Evol.* 17, 374-386 (2000).

Salemi, M., Van Dooren, S. & Vandamme, A.M. Origin and evolution of human and simian T-cell lymphotropic viruses. *AIDS Rev.* 1, 131-139 (1999).

Sanderson, Mj J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics.* 19:301-2.

Satou, Y., J.-I. Yasunaga, M. Yoshida, and M. Matsuoka. 2006 HTLV-1 basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc. Natl. Acad. Sci. USA.* 103:720-725..

Sharp, P. M., E. Bailes, F. Gao, B. E. Beer, V. M. Hirsch, B. H. Hahn. 2000. Origins and evolution of AIDS viruses: estimating the time-scale. *Biochem Soc Trans.* 28:275-282.

Slattery, J.P., Franchini, G. & Gessain, A. Genomic evolution, patterns of global dissemination, and interspecies transmission of human and simian T-cell leukemia/lymphotropic viruses. *Genome Res.* 9, 525-549 (1999).

Switzer, W. M., M. Salemi, V. Shanmugam, F. Gao, M. E. Cong, C. Kuiken, V. Bhullar, B. E. Beer, D. Vallet, A. Gautier-Hion, Z. Tooze, F. Villinger, E. C. Holmes, and W. Heneine. 2005. Ancient co-speciation of simian foamy viruses and primates. *Nature.* 434:376-380.

Switzer, W.M., Pieniazek, D., Swanson, P., Samdal, H.H., Soriano, V., Khabbaz, R.F., Kaplan, J.E., Lal, R.B. & Heneine, W. Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. *J. Virol.* 69, 621-632 (1995).

Takemura, T., M. Yamashita, M. K. Shimada, S. Ohkura, T. Shotake, M. Ikeda, T. Miura and M. Hayami. 2002. High prevalence of simian T-lymphotropic virus type L in wild ethiopian baboons. *J. Virol.* 76:1642-1648.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.

Tsubata, C., M. Higuchi, M. Takahashi, M. Oie, Y. Tanaka, F. Geyjo, and M. Fujii. 2005. PDZ domain-binding motif of human T-cell leukemia virus type 1 Tax oncoprotein is essential for the interleukin 2 independent growth induction of a T-cell line. *Retrovirol.*

2:46.

Van Brussel, M., M. Salemi, H. F. Liu, J. Gabriels, P. Goubau, J. Desmyter, and A. M. Vandamme. 1998. The simian T-lymphotropic virus STLV-PP1664 from *Pan paniscus* is distinctly related to HTLV-2 but differs in genomic organization. *Virology*. 243:366-379.

Van Brussel, M., P. Goubau, R. Rousseau, J. Desmyter, and A. M. Vandamme. 1997. Complete nucleotide sequence of the new simian T-lymphotropic virus, STLV-PH969 from a *Hamadryas baboon*, and unusual features of its long terminal repeat. *J. Virol.* 7:5464-5472.

Van Dooren S, Salemi M, Pourrut X, Peeters M, Delaporte E, Van Ranst M, Vandamme AM. Evidence for a second simian T-cell lymphotropic virus type 3 in *Cercopithecus nictitans* from Cameroon. *J Virol.* 2001 Dec;75(23):11939-41.

Van Dooren, S., Shanmugam, V., Bhullar, V., Parekh, B., Vandamme, A.M., Heneine, W. & Switzer, W.M. Identification in gelada baboons (*Theropithecus gelada*) of a distinct simian T-cell lymphotropic virus type 3 with a broad range of Western blot reactivity. *J. Gen. Virol.* 85, 507-519 (2004).

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R. & Wickner, R.B. Seventh Report of the International Committee on Taxonomy of Viruses. (Academic Press, San Diego, Wien, New York, 2000). (online version <http://www.virustaxonomyonline.com/>)

Vandamme, A.M., Van Laethem, K., Liu, H.F., Van Brussel, M., Delaporte, E., de Castro Costa, C.M., Fleischer, C., Taylor, G., Bertazzoni, U., Desmyter, J. & Goubau, P. Use of a generic polymerase chain reaction assay detecting human T-lymphotropic virus (HTLV) types I, II and divergent simian strains in the evaluation of individuals with indeterminate HTLV serology. *J. Med. Virol.* 52, 1-7 (1997).

Wolfe, N. D., W. Heneine, J. K. Carr, A. D. Garcia, V. Shanmugam, U. Tamoufe, J. N. Torimiro, A. T. Prosser, M. Lebreton, E. Mpoudi-Ngole, F. E. McCutchan, D. L. Birx, T. M. Folks, D. S. Burke, and W. M. Switzer. 2005. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc. Natl. Acad. Sci. U S A*. 102:7994-7999.

Wolfe, N.D., Prosser, A.T., Carr, J.K., Tamoufe, U., Mpoudi-Ngole, E., Torimiro, J.N., LeBreton, M., McCutchan, F.E., Birx, D.L. & Burke D.S. 2004. Exposure to nonhuman primates in rural Cameroon. *Emerg. Infect. Dis.* 10, 2094 (2004).

Wolfe, N.D., Switzer, W.M., Carr, J.K., Bhullar, V.B., Shanmugam, V., Tamoufe, U., Prosser, A.T., Torimiro, J.N., Wright, A., Mpoudi-Ngole, E., McCutchan, F.E., Birx, D.L., Folks, T.M., Burke, D.S. & Heneine, W. Naturally acquired simian retrovirus infections in central African hunters. *Lancet* 363, 932-937 (2004b).

Wolfe, N.D., Switzer, W.M., Folks, T.M., Burke, D.S. & Heneine, W. Simian retroviral infections in human beings. *Lancet* 364, 139-140 (2004c).

Womble, D. D. 2000. GCG: The Wisconsin Package of sequence analysis programs.

Methods Mol. Biol. 132:3-22.

Yamashita, M., Ido, E., Miura, T., Hayami, M. Molecular Epidemiology of HTLV-I in the world. *J Acquir Immune Defic Syndr Hum Retrovirol.* 13: S124-31 (1996).

H. Sequences

SEQ ID NO: 1 (HTLV-3 *pol*) (pos 2407-5076)

SEQ ID NO: 2 (HTLV-4 *pol*) (pos 3-2549)

SEQ ID NO: 3 (HTLV-3 *env*) (pos 5069-6544)

SEQ ID NO: 4 (HTLV-4 *env*) (pos 2542-3999)

SEQ ID NO: 5 (HTLV-3 *tax*)

SEQ ID NO: 6 (HTLV-4 *tax*)

SEQ ID NO: 7 (PTLVTPG)

SEQ ID NO: 8 (PGTAXR1)

SEQ ID NO: 9 (PH2Rrev)

SEQ ID NO: 10 (PGTAXR2)

SEQ ID NO: 11 (PGPOLF1)

SEQ ID NO: 12 (PGPOLR1)

SEQ ID NO: 13 (PGPOLF2)

SEQ ID NO: 14 (PGPOLR2)

SEQ ID NO: 15 (1863TF1)

SEQ ID NO: 16 (1863TR1)

SEQ ID NO: 17 (1863TF2)

SEQ ID NO: 18 (1863TR2)

SEQ ID NO: 19 (5VLTRext)

SEQ ID NO: 20 (1MNDR1)

SEQ ID NO: 21 (Enh280)

SEQ ID NO: 22 (1MNDR2)

SEQ ID NO: 23 (PH1F)

SEQ ID NO: 24 (PH2R)

SEQ ID NO: 25 (PH2F)

SEQ ID NO: 26 (PH2R)

SEQ ID NO: 27 (PGENVF1)

SEQ ID NO: 28 (PGENVR1)

SEQ ID NO: 29 (PGENVF2)

SEQ ID NO: 30 (PGENVR2)

SEQ ID NO: 31 (GPLTRF1)

SEQ ID NO: 32 (GPLTRR1)

SEQ ID NO: 33 (GPLTRF2)
SEQ ID NO: 34 (GPLTRR2)
SEQ ID NO: 35 (HTLV-3 *gag*) (pos 756-2023)
SEQ ID NO: 36 (HTLV-3 Complete genome: 2026ND.seq (8917 bp)
SEQ ID NO: 37 (HTLV-3 *env* amino acid)
SEQ ID NO: 38 (HTLV-3 *env* surface antigen (SU)=aa 1-315)
SEQ ID NO: 39 (HTLV-3 *env* transmembrane=aa 316-491)
SEQ ID NO: 40 (HTLV-3 *gag* amino acid)
SEQ ID NO: 41 (HTLV-3 *gag* p15=aa 337-422)
SEQ ID NO: 42 (HTLV-3 *gag* p19=aa 1-123)
SEQ ID NO: 43 (HTLV-3 *gag* p24=aa 124-336)
SEQ ID NO: 44 (HTLV-3 *pol*) amino acid
SEQ ID NO: 45 (HTLV-3 LTR) (pos 1-697 & 8221-8917)
SEQ ID NO: 46 (HTLV-3 *pro*) amino acid
SEQ ID NO: 47 (HTLV-3 *pro*) (pos 1976-2509)
SEQ ID NO: 48 (HTLV-3 *rex*) amino acid
SEQ ID NO: 49 (HTLV-3 *rex*)(pos 5010-5071..7245-7730)
SEQ ID NO: 50 (HTLV-3 *tax*) amino acid
SEQ ID NO: 51 (HTLV-3 *tax*) (pos 5069-5071..7244-8293)
SEQ ID NO: 52 (HTLV-3 pX (pos 6545-7243)
SEQ ID NO: 53 (HTLV-4 (pol-env-tax region)
SEQ ID NO: 54 (HTLV-4 *env*) amino acid
SEQ ID NO: 55 (HTLV-4 *env* surface antigen (SU)=aa 1-307)
SEQ ID NO: 56 (HTLV-4 *env* transmembrane=aa 308-485)
SEQ ID NO: 57 (HTLV-4 *pol*) amino acid
SEQ ID NO: 58 (HTLV-4 *pro*) amino acid
SEQ ID NO: 59 (HTLV-4 *pro*) (pos 1-273)
SEQ ID NO: 60 (HTLV-4 *rex*) amino acid
SEQ ID NO: 61 (HTLV-4 *rex*) (pos 2483-2545..4560-5009)
SEQ ID NO: 62 (HTLV-4 *tax*) amino acid
SEQ ID NO: 63 (HTLV-4 pX) (pos.4000-4558)
SEQ ID NO: 64 1863PF1
SEQ ID NO: 65 1863PR2

SEQ ID NO: 66 1863PP2 FAM (flourescent labeled probe)

SEQ ID NO: 67 region of HTLV-3 where type specific peptides of HTLV-1 and HTLV-2 are located

SEQ ID NO: 68 region of HTLV-4 where type specific peptides of HTLV-1 and HTLV-2 are located

SEQ ID NO: 69 2026LF1

SEQ ID NO: 70 2026PR1

SEQ ID NO: 71 2026LF2

SEQ ID NO: 72 2026PR2

SEQ ID NO: 73 2026PF3

SEQ ID NO: 74 2026ER1

SEQ ID NO: 75 2026PF2

SEQ ID NO: 76 2026ER3

SEQ ID NO: 77 2026EF1

SEQ ID NO: 78 2026LR1

SEQ ID NO: 79 2026EF3

SEQ ID NO: 80 2026LR2

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 83

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 83

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

63198-1563

CLAIMS:

1. An isolated primate T-lymphotropic virus comprising SEQ ID NOs: 1, 3 and 5.
2. The primate T-lymphotropic virus of claim 1, wherein the isolated primate T-lymphotropic virus comprises SEQ ID NOs: 1, 3, 5 and 35.
- 5 3. The primate T-lymphotropic virus of claim 1, wherein the primate T-lymphotropic virus comprises SEQ ID NOs: 1, 3, 5 and 45.
4. The primate T-lymphotropic virus of claim 1, wherein the primate T-lymphotropic virus comprises SEQ ID NOs: 1, 3, 5, 35 and 45.
5. An isolated primate T-lymphotropic virus comprising SEQ ID NOs: 1, 3, 35,
10 and 45.
6. An isolated primate T-lymphotropic virus comprising SEQ ID NOs: 2, 4, and 6.
7. An isolated primate T-lymphotropic virus comprising a *pol* gene at least 92.8% identical to SEQ ID NO: 1, an *env* gene at least 92.5% identical to SEQ ID NO: 3, and
15 a *tax* gene at least 94.2% identical to SEQ ID NO: 5.
8. An isolated primate T-lymphotropic virus comprising a *pol* gene at least 71.5% identical to SEQ ID NO: 2, an *env* gene at least 73.5% identical to SEQ ID NO: 4, and a *tax* gene at least 82% identical to SEQ ID NO: 6.
9. A vector comprising the primate T-lymphotropic virus of any one of
20 claims 1-8.
10. The vector of claim 9, further comprising a heterologous nucleic acid.
11. The vector of claim 10, wherein the heterologous nucleic acid is a nucleic acid encoding an antigen from a known pathogen under the control of a promoter element.

63198-1563

12. The vector of claim 11, wherein the pathogen is a virus selected from the group of viruses consisting of Herpes simplex virus type-1, Herpes simplex virus type-2, Cytomegalovirus, Epstein-Barr virus, Varicella-zoster virus, Human herpesvirus 6, Human herpesvirus 7, Human herpesvirus 8, Variola virus, Vesicular stomatitis virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rhinovirus, Coronavirus, Influenza virus A, Influenza virus B, Measles virus, Polyomavirus, Human Papillomavirus, Respiratory syncytial virus, Adenovirus, Coxsackie virus, Dengue virus, Mumps virus, Poliovirus, Rabies virus, Rous sarcoma virus, Yellow fever virus, Ebola virus, Marburg virus, Lassa fever virus, Eastern Equine Encephalitis virus, Japanese Encephalitis virus, St. Louis Encephalitis virus, Murray Valley fever virus, West Nile virus, Rift Valley fever virus, Rotavirus A, Rotavirus B, Rotavirus C, Sindbis virus, Simian Immunodeficiency virus, Human T-lymphotropic virus type-1, Human T-lymphotropic virus type-2, Primate T-lymphotropic virus, Hantavirus, Rubella virus, Human Immunodeficiency virus type-1, and Human Immunodeficiency virus type-2, and Simian Immunodeficiency virus (SIV).

13. The vector of claim 11, wherein the pathogen is a bacterium selected from the group of bacteria consisting of *M. tuberculosis*, *M. bovis*, *M. bovis* strain BCG, BCG substrains, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subspecies *paratuberculosis*, *Nocardia asteroides*, *Nocardia species*, *Legionella pneumophila*, *Legionella species*, *Salmonella typhi*, *Salmonella species*, *Shigella species*, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, *Pasteurella species*, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, *Brucella species*, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, *Rickettsial species*, *Ehrlichia species*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, *Campylobacter species*, *Neisseria meningitidis*, *Neisseria gonorrhea*, *Pseudomonas aeruginosa*, *Pseudomonas species*, *Haemophilus influenzae*, *Haemophilus ducreyi*, *Haemophilus species*, *Clostridium tetani*, *Clostridium species*, *Yersinia enterocolitica*, and *Yersinia species*.

14. The vector of claim 11, wherein the pathogen is a parasite selected from the group of parasites consisting of *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium*

63198-1563

vivax, *Plasmodium malariae*, *Plasmodium species*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, *Leishmania species*, *Schistosoma mansoni*, *Schistosoma species*, and *Entamoeba histolytica*.

15. The vector of claim 11, wherein the pathogen is a fungi selected from the
5 group of fungi consisting of *Candida albicans*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Aspergillus fumigatus*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Blastomyces dermatidis*, *Pneumocystis carinii*, *Penicillium marneffi*, and *Alternaria alternata*.

16. The vector of claim 10, wherein the heterologous nucleic acid encodes a
cancer related antigen selected from the group of cancers consisting of lymphoma, B cell
10 lymphoma, T cell lymphoma, mycosis fungoides, Hodgkin's Disease, myeloid leukemia, bladder cancer, brain cancer, nervous system cancer, head and neck cancer, squamous cell carcinoma of head and neck, kidney cancer, lung cancers such as small cell lung cancer and non-small cell lung cancer, neuroblastoma/glioblastoma, ovarian cancer, pancreatic cancer, prostate cancer, skin cancer, liver cancer, melanoma, squamous cell carcinomas of the mouth,
15 throat, larynx, and lung, colon cancer, cervical cancer, cervical carcinoma, breast cancer, and epithelial cancer, renal cancer, genitourinary cancer, pulmonary cancer, esophageal carcinoma, head and neck carcinoma, large bowel cancer; hematopoietic cancers; testicular cancer; colon and rectal cancers, prostatic cancer, and pancreatic cancer.

17. The vector of claim 16, wherein the cancer antigen is selected from the group
20 of cancer antigens consisting of human epithelial cell mucin (Muc-1), the Ha-ras oncogene product, p53, carcino-embryonic antigen (CEA), the raf oncogene product, gp100/pm117, GD2, GD3, GM2, TF, sTn, MAGE-1, MAGE-3, BAGE, GAGE, tyrosinase, gp75, Melan-A/Mart-1, gp100, HER2/neu, EBV-LMP 1 & 2, HPV-F4, 6, 7, prostate-specific antigen (PSA), HPV-16, MUM, alpha-fetoprotein (AFP), CO17-1A, GA733, gp72, p53, the ras
25 oncogene product, HPV E7, Wilms tumor antigen-1, telomerase, and melanoma gangliosides.

18. A use of the vector of claim 9 for vaccinating a subject against an illness associated with a pathogen.

19. A method of identifying new primate T-lymphotropic viruses comprising:

63198-1563

a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is SEQ ID NOs: 19 and 20, wherein the second set of primers is SEQ ID NOs: 21 and 22;

b) identifying any amplified nucleic acid; and

5 c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus.

20. A method of identifying new primate T-lymphotropic viruses comprising:

a) contacting a nucleic acid using a first set of primers and a second set of primers internal to the first set of primers, wherein the first set of primers is selected from the
10 group of primers pairs consisting of SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70, SEQ ID NOs: 73 and 74, and SEQ ID NOs: 77 and 78, wherein the second set of primers is selected from the group consisting of SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30,
15 SEQ ID NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80;

b) identifying any amplified nucleic acid; and

c) comparing the sequence to known primate T-lymphotropic viral sequences, wherein sequence divergence greater than 5% indicates a new virus.

20 21. A method of identifying new primate T-lymphotropic viruses comprising:

a) contacting a nucleic acid using a set of primers, wherein the set of primers is selected from the set of primers consisting of SEQ ID NOs: 19 and 20, SEQ ID NOs: 21 and 22, SEQ ID NOs: 7 and 8, SEQ ID NOs: 11 and 12, SEQ ID NOs: 15 and 16, SEQ ID NOs: 23 and 24, SEQ ID NOs: 27 and 28, SEQ ID NOs: 31 and 32, SEQ ID NOs: 69 and 70,
25 SEQ ID NOs: 73 and 74, SEQ ID NOs: 77 and 78, SEQ ID NOs: 9 and 10, SEQ ID NOs: 13 and 14, SEQ ID NOs: 17 and 18, SEQ ID NOs: 25 and 26, SEQ ID NOs: 29 and 30, SEQ ID

63198-1563

NOs: 33 and 34, SEQ ID NOs: 71 and 72, SEQ ID NOs: 75 and 76, and SEQ ID NOs: 79 and 80;

b) identifying any amplified nucleic acid; and

c) comparing the sequence to known primate T-lymphotropic viral sequences,

5 wherein sequence divergence greater than 5% indicates a new virus.

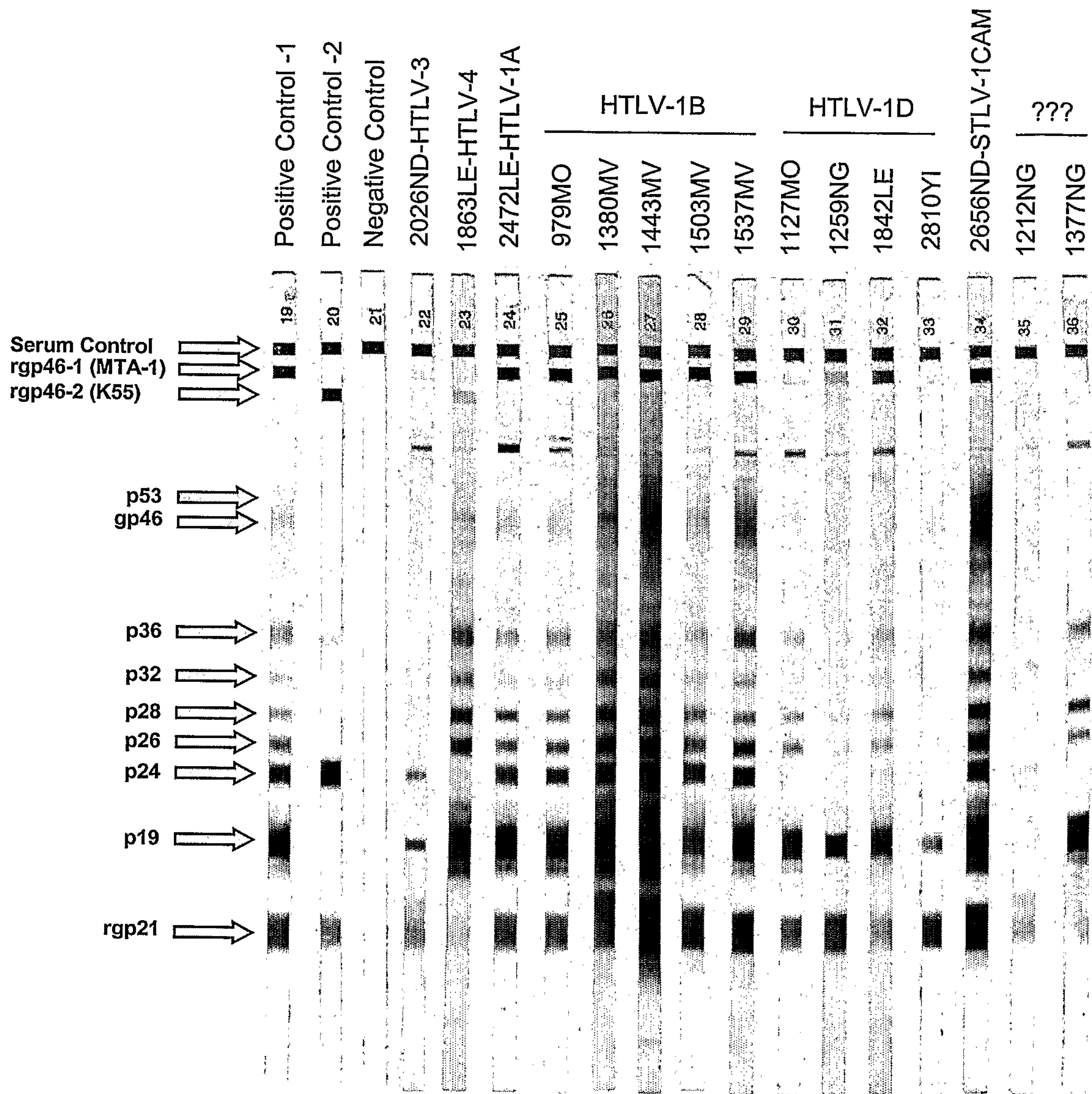
Fig. 1

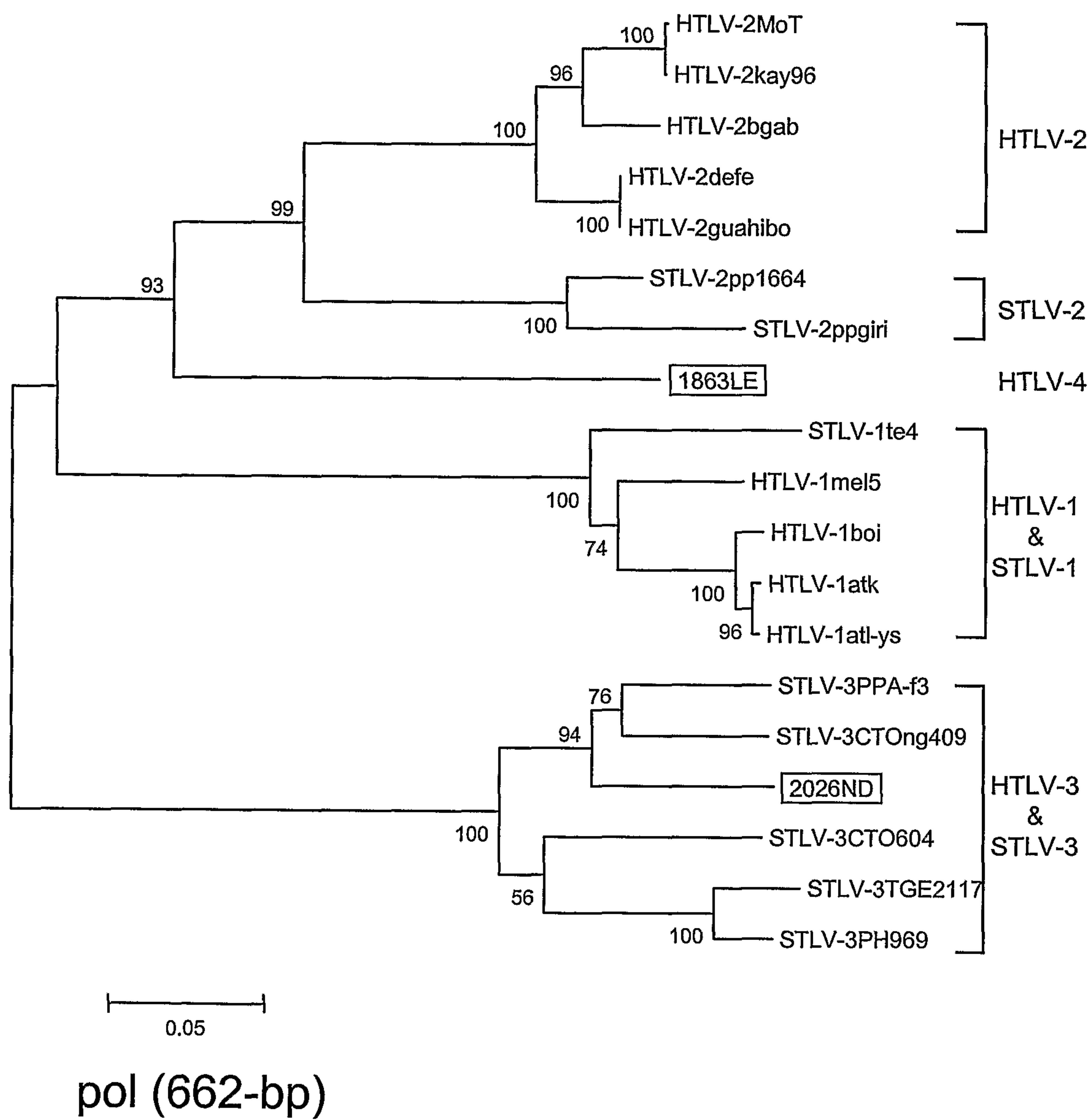
Fig. 2a

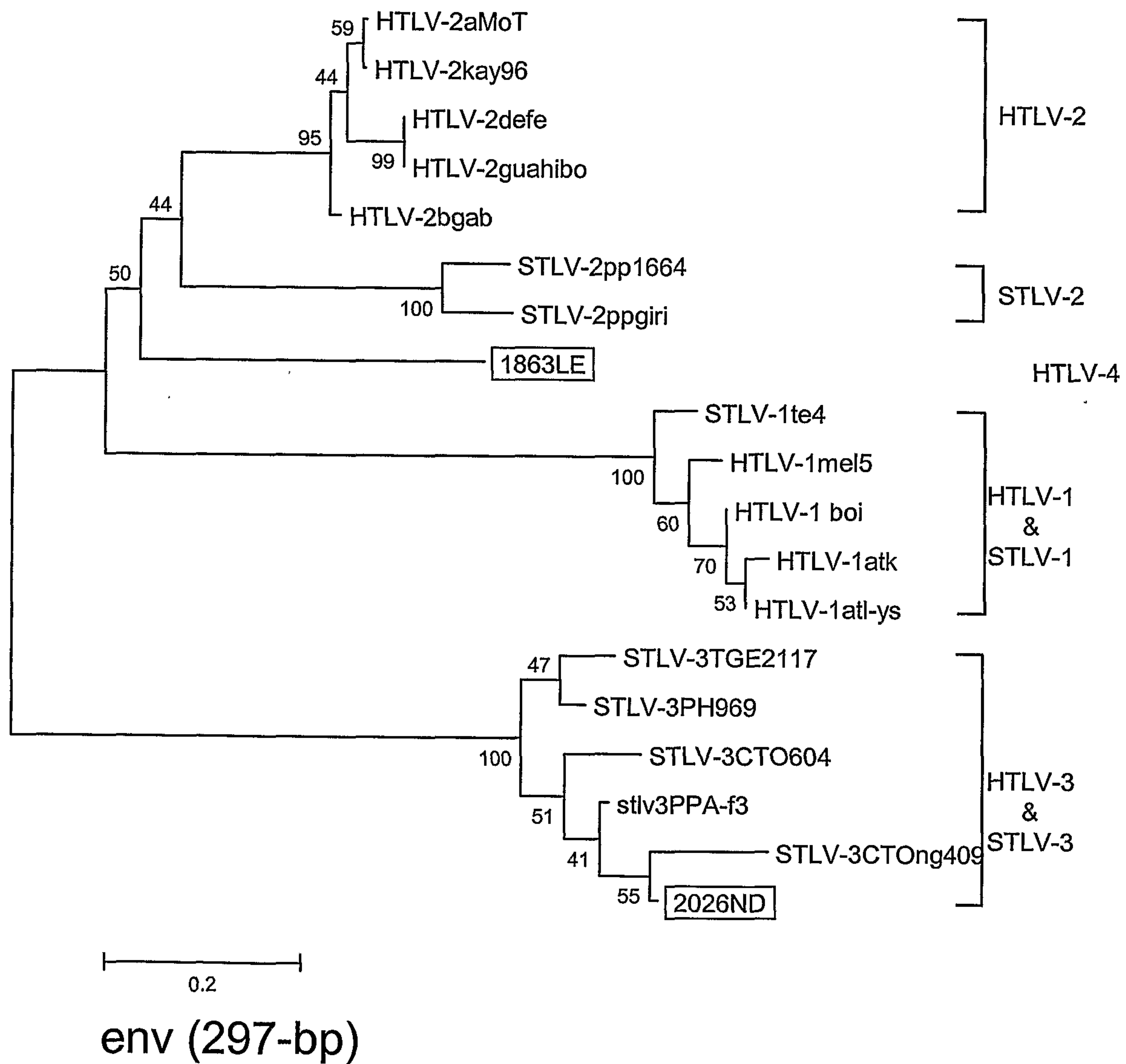
Fig. 2b

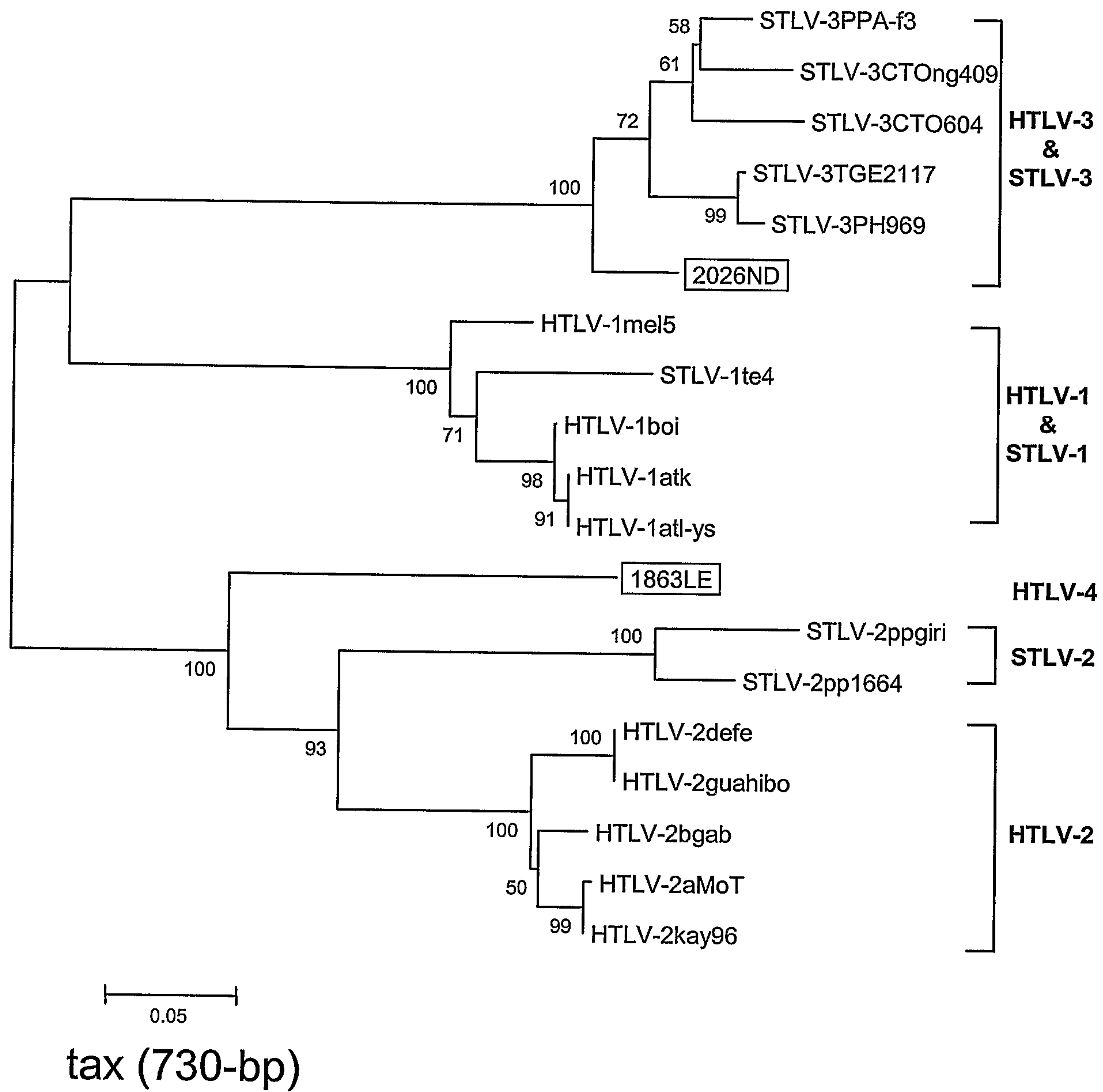
Fig. 2c

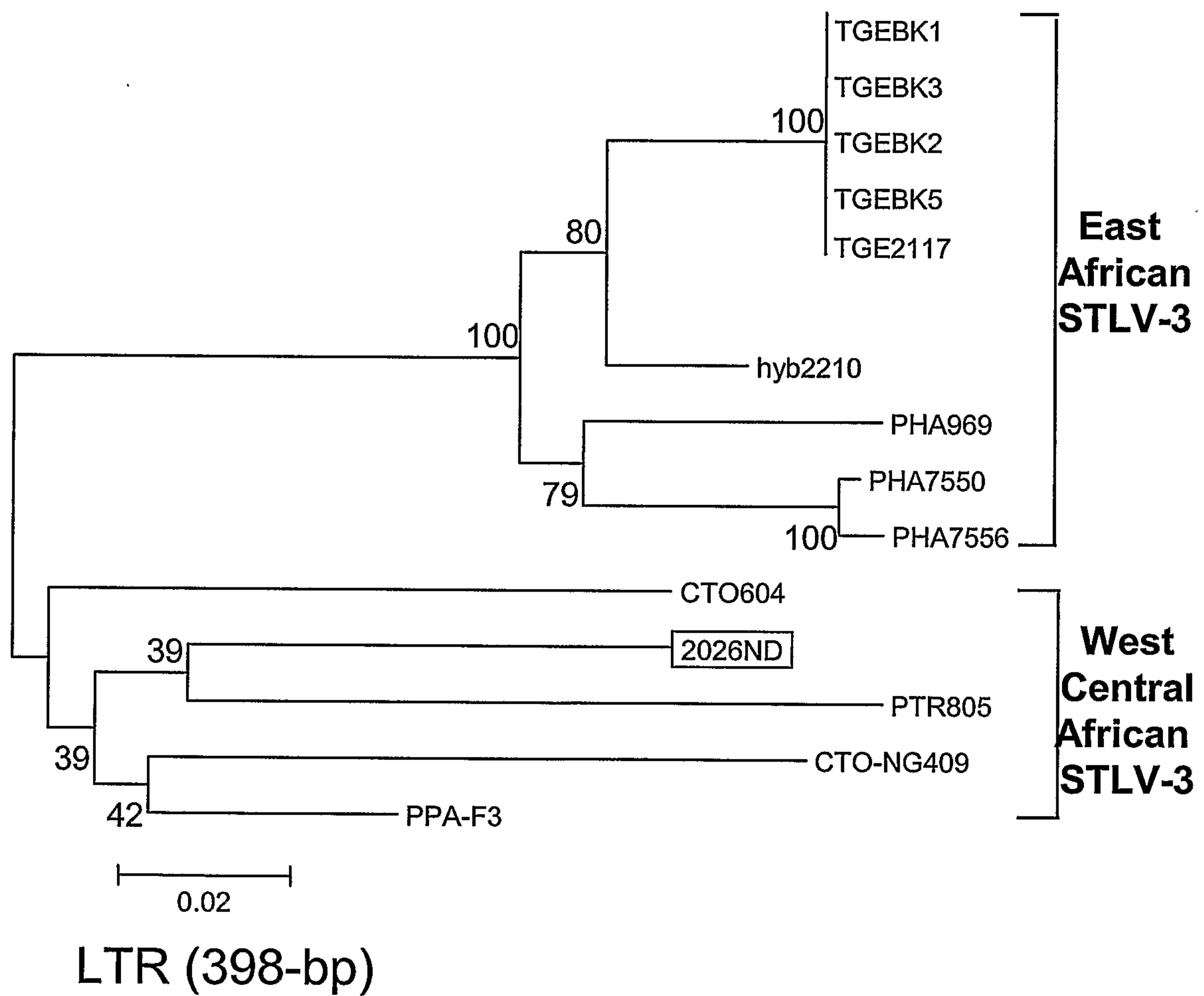
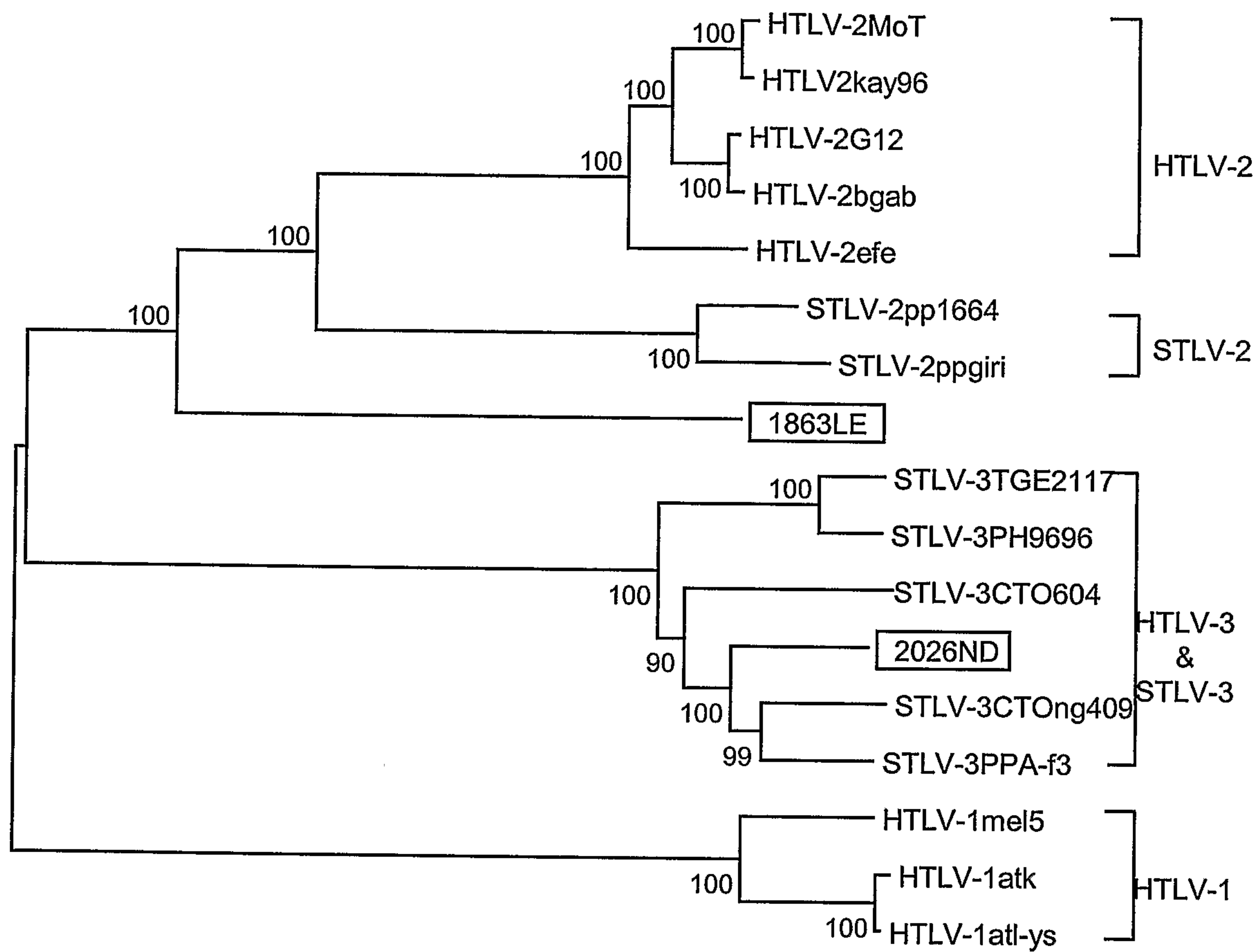
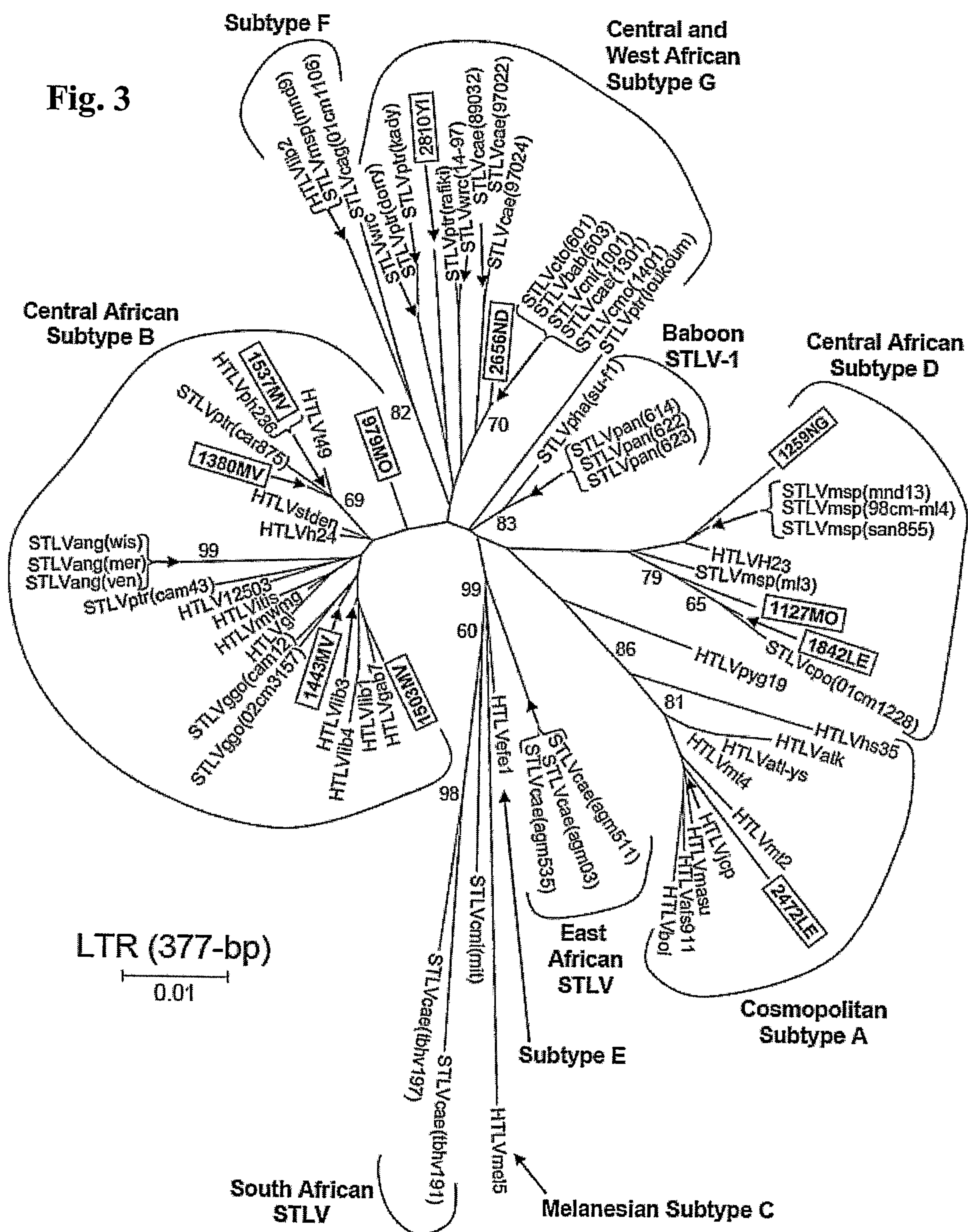
Fig. 2d

Fig. 2e

pol-env-tax (5258-bp)



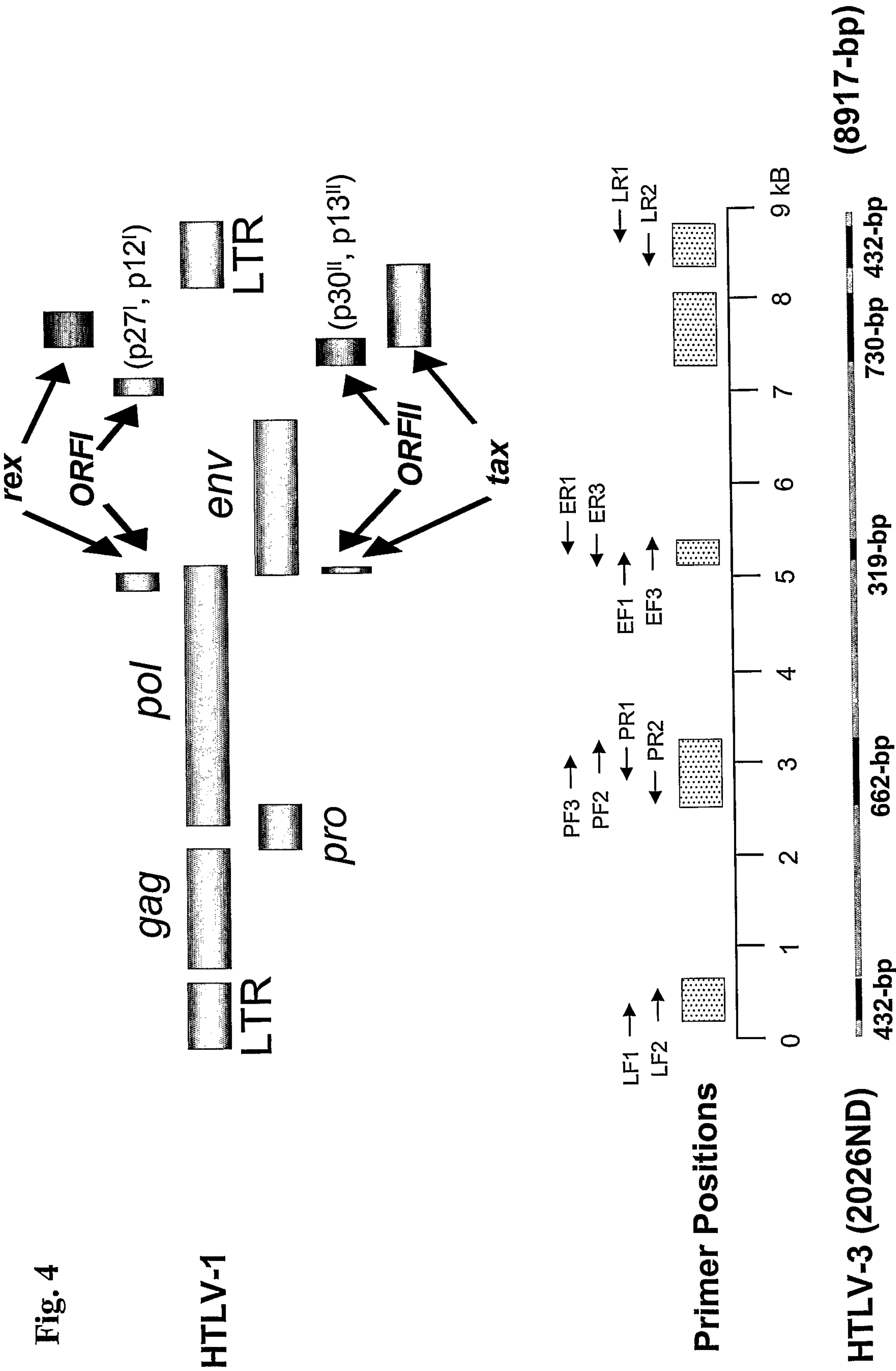


Fig 5a.

TGTCGATGATGAGCCCCGAGACGGGTCACACCACAGCTAGAGGACAAATA
GCTGAGTACCCGCTGAGAACCGTCTCACACCGGGAATTGTGCCAAAAAGAACAA
CCGGGCTCTGACGTCTCTCCCTAC **21 R** CCTGGCTCCCGGAAAAAACCAACCACCC
ATTTCCTCATGTTTGCCTAAAGCTCTGACGATAACCCCTAAAAAATTGACTAGCAA **U3 R**
ATAAAGAAACCCTGGGCCCTATAAAAGGGGAGAGCAACCTAAAAATGGGATCCCTT **U3 R**
poly (A) signal **TATA box** **cap** *******
TTCTGCACTCGCCAACCCCTCCTCTGGCCACGGTCCGACTTTGGTCATTCCCTGCCT
ACCTGAATCGCCGCTTCGGGATCGAGCCATCCCTCTTCTATTGGTGGCACTTCGC
GCACTCCGCCGCTTCCACTCGGCTAAGATCCCACTGGGTCGAGCTAGGCCATCACC
CCTGGGCCGCTCCCTGGAGCTCTCTCGCGGCTCTTAAGTTGCTCCCCCTCAG
CAAAGGCCCAGGGCTTCTCTACTTCCCTTGTTCAAAGTCTCTTCTTGGCGGTCG **R U5**
ACCTAAATCGAAAGTAGCACTTCTGCTGTCTCAGCAGCAGGCTTGGCCAGGGCCA
GCGCCTGTAAAGTTACCCAGCTCGGAGTTGGGTCTCTAGAGAATCAGGGCTAAAG
CTGCTAGCCCTAGGAAAGAAAGGCAACAGGTGGGGGCTCGTCCGGGATTGATCACC **PBS**
TCTCTGTATTGCCCCTTCCCTGTCGAAGCC

Fig. 5b

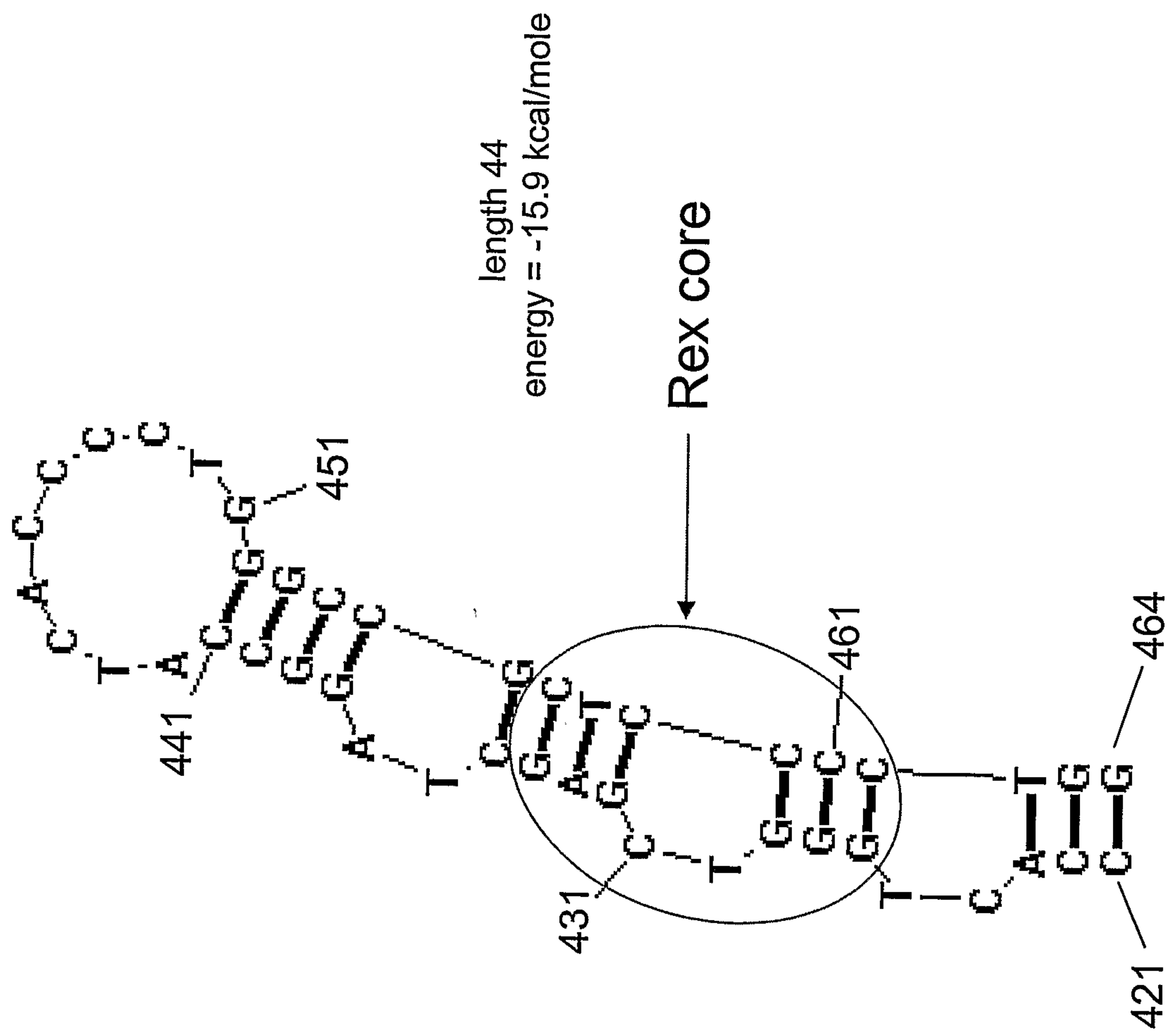


Fig. 6

		NLS	
HTLV-3 (2026ND)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-3 (CTO604)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-3 (PH969)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-3 (PPAF3)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-3 (NG409)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-3 (TGE2117)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
HTLV-2 (MoT)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
HTLV-2 (G12)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-2 (PanP)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
STLV-2 (PP1664)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
HTLV-1 (ATK)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	70
HTLV-1 (Me1)	MAHPEPCGCSILHVMCPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	CPERQITWDPEDERVVSSALQY	69
		CBP/P300 binding	
HTLV-3 (2026ND)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-3 (CTO604)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-3 (PH969)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-3 (PPAF3)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-3 (NG409)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-3 (TGE2117)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
HTLV-2 (MoT)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
HTLV-2 (G12)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-2 (PanP)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
STLV-2 (PP1664)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
HTLV-1 (ATK)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	140	
HTLV-1 (Me1)	HTPRPSPSPRRSRKSKVAPPVAVPKPPSPPTAVKKKPPRRNNCLRLHGEQPPNCPSPCLRRP	139	
		NES	
HTLV-3 (2026ND)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-3 (CTO604)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-3 (PH969)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-3 (PPAF3)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-3 (NG409)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-3 (TGE2117)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
HTLV-2 (MoT)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
HTLV-2 (G12)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-2 (PanP)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
STLV-2 (PP1664)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
HTLV-1 (ATK)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	210
HTLV-1 (Me1)	QNYVITWCSVWCNLYOCPPEWNPVAVREDCVQADWCPVSGGLCSARLIRHALLAA	KRIEELLYKILTSALHTEPE	209
HTLV-3 (2026ND)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-3 (CTO604)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-3 (PH969)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-3 (PPAF3)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-3 (NG409)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-3 (TGE2117)	NCEPRLHLPTRAPALQAPWAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
HTLV-2 (MoT)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
HTLV-2 (G12)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-2 (PanP)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
STLV-2 (PP1664)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
HTLV-1 (ATK)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	280	
HTLV-1 (Me1)	EDPPTVPCVVRPQVQVANDAGLEPCCKEHVPPHAWHNDGCPMISCPCKKGGPRLAVGSSHEIFQQ	279	
		CR2 binding	
HTLV-3 (2026ND)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-3 (CTO604)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-3 (PH969)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-3 (PPAF3)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-3 (NG409)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-3 (TGE2117)	FOYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
HTLV-2 (MoT)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
HTLV-2 (G12)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-2 (PanP)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
STLV-2 (PP1664)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
HTLV-1 (ATK)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
HTLV-1 (Me1)	FEYKALPAPATLSPKKEHYS-----SFSSTHILFREYINVPGLLEFNEKCANVDDDEERDGS---CPPAR	342	
		PDZ	
HTLV-3 (2026ND)	GOYALSP-----	350	
STLV-3 (CTO604)	GOYALSP-----	350	
STLV-3 (PH969)	GOYALSP-----	350	
STLV-3 (PPAF3)	GOYALSP-----	350	
STLV-3 (NG409)	GOYALSP-----	350	
STLV-3 (TGE2117)	GOYALSP-----	350	
HTLV-2 (MoT)	GOYALSP-----	350	
HTLV-2 (G12)	KVRPSHTNNPK-----	356	
STLV-2 (PanP)	QLKGHVVSHKQSREQALTFFPADHLRNQEFVSRKITSPLPTSPCLEKQRL	400	
STLV-2 (PP1664)	SA-----	347	
HTLV-1 (ATK)	KHFRFMEV-----	353	
HTLV-1 (Me1)	KHFRFMEV-----	352	

Fig. 7

