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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 April 2010 (15.04.2010)

PCT

(10) International Publication Number
WO 2010/042817 A1

(51) International Patent Classification:
A61K 39/21 (2006.01) *A61P 31/18* (2006.01)
A61K 39/295 (2006.01)

(21) International Application Number:
PCT/US2009/060167

(22) International Filing Date:
9 October 2009 (09.10.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/195,845 10 October 2008 (10.10.2008) US

(71) Applicants *(for all designated States except US)*: **THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; Office of Technology Transfer, National Institutes of Health, 6011 Executive Boulevard, Suite 325, MSC 7660, Bethesda, Maryland 20892-7660 (US). **Los Alamos National Security, LLC** [US/US]; Los Alamos National Laboratory, LC/IP, Los Alamos, New Mexico 87545 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **NABEL, Gary J.** [US/US]; 2520 30th Street N.W., Washington, District of Columbia 20008 (US). **KONG, Wing-pui** [GB/US]; 13308 Copper Ridge Road, Germantown, Maryland 20874 (US). **WU, Lan** [US/US]; 13322 Copper Ridge Road, Germantown, Maryland 20874 (US). **KORBER, Bette** [US/US]; 1290 Big Rock Loop, Los Alamos, New Mexico 87544 (US). **FISCHER, William M.** [US/US]; 619 Galisteo Street, Santa Fe, New Mexico 87505 (US).

(74) Agents: **KOLOM, Melissa E.** et al.; Leydig, Voit & Mayer, Ltd., Two Prudential Plaza, Suite 4900, 180 North Stetson Avenue, Chicago, Illinois 60601 (US).

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

(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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2010/042817 A1

(54) **Title**: USE OF HIV MOSAIC ENVELOPE PROTEINS TO OPTIMIZE T-CELL RESPONSES AGAINST HIV ENVELOPE

(57) **Abstract**: The invention is directed to a nucleic acid sequence encoding a HIV-1 Env polypeptide comprising an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide, as well as a method of making the nucleic acid sequence. The invention also provides a method of inducing an immune response against HIV-1 in a mammal.

USE OF HIV MOSAIC ENVELOPE PROTEINS TO OPTIMIZE T-CELL RESPONSES AGAINST HIV ENVELOPE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This patent application claims the benefit of U.S. Provisional Patent Application No. 61/195,845, filed October 10, 2008, which is incorporated by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

[0002] This invention was made with Government support under Contract No. DE-AC52-06NA25396 awarded by the U.S. Department of Energy. The Government has certain rights in this invention.

INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ELECTRONICALLY

[0003] Incorporated by reference in its entirety herein is a computer-readable nucleotide/amino acid sequence listing submitted concurrently herewith and identified as follows: One 288,764 Byte ASCII (Text) file named "705228_ST25.TXT," created on October 8, 2009.

BACKGROUND OF THE INVENTION

[0004] The development of an AIDS vaccine has been advanced recently by demonstrations of increased survival and decreased viral load following vaccination with T-cell vaccines in non-human primate models (see, e.g., Kawada et al., *J. Virol.*, *82*: 10199-101206 (2008); Letvin et al., *Science*, *312*: 1530-1533 (2006); Matano et al., *J. Exp. Med.*, *199*: 1709-1718 (2004); Santra, *Proc. Natl. Acad. Sci. USA*, *105*: 10489-10494 (2008); Wilson et al., *J. Virol.*, *80*: 5875-5885 (2006)). Although such vaccines have suggested that T-cells may contribute to the control of HIV viremia in the highly lethal SIVmac251 challenge model, how these results apply to human studies remains uncertain. The major concern regarding the efficacy of HIV vaccines in humans is the extraordinary genetic diversity of the virus. The sequence similarity of HIV-1 Envelope

protein (Env) from diverse isolates within a clade can diverge as much as 15%, and between alternative clades can diverge as much as 30% (see, e.g., Gaschen et al., *Science*, *296*: 2354-2360 (2002)). In addition, the diversity of the HIV-1 Gag protein can approach similar levels, particularly in the p17 and p15 regions which are much more diverse than the p24 region (see, e.g., Fischer et al., *Nat. Med.*, *13*: 100-106 (2007)), although Gag does not have the extreme localized diversity observed in the highly variable regions of Env (see, e.g., Fischer et al., *supra*, and Gaschen et al., *supra*). While viral diversity has been addressed in existing vaccines through the use of envelopes derived from representative viruses in the major clades, increasing knowledge about the genetic diversity of naturally occurring isolates has enabled alternative approaches that enhance population coverage of vaccine-elicited T-cell responses.

[0005] Approaches under consideration include the use of ancestral, central or consensus, and "center of the tree" gene sequences (see, e.g., Doria-Rose et al, *J. Virol.*, *79*: 11214-11224 (2005); Gaschen et al., *supra;* Kothe et al., *Virology*, *352*: 438-449 (2006); Santra et al., *supra;* and Weaver et al., *J. Virol.*, *80*: 6745-6756 (2006)). These can be derived using a number of alternative approaches, including the alignment of HIV gene sequences with selection of the most common amino acids at each residue (see, e.g., Gaschen et al., *supra;* Korber et al., *Br. Med. Bull.*, *58*: 19-42 (2001); Kothe et al., *Virology*, *360*: 218-234 (2007); Liao et al., *Virology*, *353*: 268-282 (2006); Novitsky et al., *J. Virol.*, *76*: 5435-5451 (2002); Weaver et al., *supra*), modeling the most recent common ancestor of diverging viruses in a vaccine target population (see, e.g., Doria-Rose et al., *supra;* Gaschen et al., *supra;* Kothe et al., *Virology*, *352*: 438-449 (2006); Weaver et al., *supra*), or modeling the sequence at the center of the phylogenetic tree (see, e.g., Rolland et al., *J. Virol.*, *81*: 8507-8514 (2007)). Peptides based on any of these three centralized protein strategies enhance the detection of T-cell responses in a natural HIV-1 infection relative to the use of peptides based on natural strains; however, all three strategies produce equivalent results (see, e.g., Frahm et al., *AIDS*, *22*: 447-456 (2008)).

[0006] The use of a single HIV-1 group M consensus/ancestral Env sequence has been shown to elicit T-cell responses with greater breadth of cross-reactivity than single natural strains in animal models (see, e.g., Santra et al., *supra;* Weaver et al., *supra*). Such central sequences do not exist in nature, and phylogenetic ancestral reconstructions are an approximate model of an ancestral state of the virus (see, e.g., Gao et al., *Science*, *299*: 1517-1518 (2003)). Thus, central

sequence strategies have provided evidence that various informatically-derived gene products can elicit immune responses to T-cell epitopes found in diverse circulating strains. While consensus genes have been found to be superior to wild-type genes (see, e.g., Weaver et al., *supra;* Santra et al., *supra*), the ability of the most recent informatically-derived HIV-1 gene products (also known as "mosaics") to elicit immune responses to T-cell epitopes found in diverse circulating strains has not been defined.

[0007] Thus, there remains a need for vaccines against HIV-1 which improve, and desirably optimize, coverage of T-cell epitopes. This invention provides nucleic acid sequences for HIV-1 vaccination, as well a methods for producing such nucleic acid sequences.

BRIEF SUMMARY OF THE INVENTION

[0008] The invention provides a nucleic acid sequence encoding a HIV-1 Env polypeptide, which HIV-1 Env polypeptide comprises an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide.

[0009] The invention also provides a method of making a nucleic acid sequence encoding a modified HIV-1 Env polypeptide. The method comprises (a) identifying at least one T-cell epitope which is present in a natural Env polypeptide of a plurality of circulating HIV-1 strains, and (b) introducing a nucleic acid sequence encoding the at least one T-cell epitope into a nucleic acid sequence encoding an HIV-1 Env polypeptide which does not naturally contain the T-cell epitope.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S)

[0010] Figure 1A is a schematic representation of modified HIV Env genes and comparison of the potential epitope coverage (9-mer matches) between different trivalent Env cocktails and the Los Alamos HIV Database of M group Env sequences. The major structural features of the Env proteins used (gp160, gp160ΔVs, gp160ΔCFI, gp160ΔCFIΔVs, and gp145ΔCFIΔVs (F1, F2, F3, F4)) are shown. LS represents the leader sequence. V1, V2, V3, V4, and V5 indicate the respective variable regions. The transmembrane domain (TM) and other major structural domains also are shown. Various domains of Env were deleted and the modified Env genes

were inserted into the mammalian expression plasmid pCMV/R (Barouch et al., *J. Virol.*, *79*: 8828-8834 (2005), and Yang et al., *Science*, *317*: 825-828 (2007)) for immunization.

[0011] Figure 1B is a schematic representation of modified HIV Env genes and comparison of the potential epitope coverage (9-mer matches) between different trivalent Env cocktails and the Los Alamos HIV Database of M group Env sequences. The Env alignment used for these comparisons is from the 2007 Los Alamos database, and includes one sequence from each of 1531 people, collected worldwide and aligned, including the spectrum of diversity found among known M group clades and recombinants. Black shading indicates the fraction of all 9-mers in the database set that are found exactly matched in a given antigen cocktail. Dark gray shading indicates 8/9 positions are matched, while light gray shading indicates matches in 7/9 positions. The gp145ΔCFI-modified versions of these proteins, shown on the right, do not have as many matching potential epitopes, because sections of Env are deleted. The nat.3 set was optimized for gp160, and is not particularly well optimized for gp145ΔCFI, and gives comparable coverage to Env ABC in the retained regions. When the subset of sequences from clades A (82 sequences), B (454) or C (464) were each compared separately to these M group-designed antigens, each clade behaved roughly comparably (within a few percent) to the full dataset results shown in the figure.

[0012] Figure 2 is a graph depicting CD4 and CD8 T-cell responses in mice immunized with different gp160 mosaics, natural sequences, and gp145ΔCFI. The indicated variants of gp160 mosaics and natural strain, with or without ΔVs (depicted in Fig. 1A), were compared to gp145ΔCFI. Percent positive CD4+ (left) and CD8+ (right) cells of the total T-cells were measured using ICS for IFN-γ (solid bars) and TNF-α (open bars).

[0013] Figure 3A is a schematic demonstrating estimates of vaccine-induced immunity based on IFN-γ ICS responses to different peptide pools: Env ABC, PTE, and grouped PTE peptides. Plots of the mean and uncertainty in v_i; vaccine strength goes as $\exp(v_i)$. The vaccine effect measures the contribution of the particular vaccine antigen modification on observed signal, based on a model that factors in variation between experiments and random error. The vaccine effect is a relative measure of the reactivity of the peptide pools versus the negative control. The negative control is plotted on the left, and the error is centered around zero; the vaccine effect is relative to this point. The "X" plotted second from the left is a positive control, where an

autologous response to the natural strains unselected A, B and C clade strains is measured against peptides derived from the vaccine protein sequences. Error bars denote $\pm 2\sigma$ uncertainties. The open circle is a negative control; the rectangle denotes gp160; the filled circle denotes gp145ΔCFI; and "X" denotes gp160ΔCFI.

[0014] Figure 3B is a schematic demonstrating estimates of vaccine-induced immunity based on TNF-α ICS responses to different peptide pools: Env ABC, PTE, and grouped PTE peptides. Plots of the mean and uncertainty in v_i; vaccine strength goes as $\exp(v_i)$. The vaccine effect measures the contribution of the particular vaccine antigen modification on observed signal, based on a model that factors in variation between experiments and random error. The vaccine effect is a relative measure of the reactivity of the peptide pools versus the negative control. The negative control is plotted on the left, and the error is centered around zero; the vaccine effect is relative to this point. The "X" plotted second from the left is a positive control, where an autologous response to the natural strains unselected A, B and C clade strains is measured against peptides derived from the vaccine protein sequences. Error bars denote $\pm 2\sigma$ uncertainties. The open circle is a negative control; the rectangle denotes gp160; the filled circle denotes gp145ΔCFI; and "X"denotes gp160ΔCFI.

[0015] Figure 4A is a graph depicting IFN-γ ICS responses of CD8+ T cells to 78 PTE peptide pools for selected vaccine antigen designs. Data from all dates and modifications of basic vaccines are plotted. Different dates are indicated by different colors. Positive PTE pools, in which the median response of replicate experiments exceeds 0.1, are indicated by Xs. The median response for each pool is indicated by an open square. Negative control vaccines are not shown. The date and vaccine effects were both highly significant ($P < 10^{-9}$ for all PTE peptide pool models). Analysis of standardized residuals supported the Gaussian model for the errors. The residual standard errors ranged from 0.25 to 0.34.

[0016] Figure 4B is a graph depicting TNF-α ICS responses of CD8+ T cells to 78 PTE peptide pools for selected vaccine antigen designs. Data from all dates and modifications of basic vaccines are plotted. Different dates are indicated by different colors. Positive PTE pools, in which the median response of replicate experiments exceeds 0.1, are indicated by Xs. The median response for each pool is indicated by an open square. Negative control vaccines are not shown. The date and vaccine effects were both highly significant ($P < 10^{-9}$ for all PTE peptide

pool models). Analysis of standardized residuals supported the Gaussian model for the errors. The residual standard errors ranged from 0.25 to 0.34.

[0017] Figure 5A is a graph depicting IFN-γ ICS responses of CD4+ T cells to 78 PTE peptide superpools for selected vaccine antigen designs. The cutoff is 0.05, and the vertical axis extends only to 1.0, rather than 2.0. As in the small pool data, the date and vaccine effects for all models were highly significant ($P < 2\times10^{-7}$).

[0018] Figure 5B is a graph depicting TNF-α ICS responses of CD4+ T cells to 78 PTE peptide superpools for selected vaccine antigen designs. The cutoff is 0.05, and the vertical axis extends only to 1.0, rather than 2.0. As in the small pool data, the date and vaccine effects for all models were highly significant ($P < 2\times10^{-7}$).

[0019] Figure 6A is a schematic depicting a heatmap for CD8 bifunctional (IFN-γ or TNF-α) T-cell responses to the 78 PTE peptide pools with small numbers of peptides. For CD8 cells, light gray indicates no response, dark gray indicates a response to either IFN-γ or TNF-α, and black indicates a response to both IFN-γ and TNF-α. White squares indicate data that was not available. Rows represent particular experiments, and columns represent particular peptide pools. Each duplicated symbol indicates a repeated experiment on a different date. The number of antigens in the polyvalent vaccine is indicated at the left on the key. The histogram shows the total number of each class of response (i.e., 0, 1, or 2 responses). The six peptide pools tested that contained the 36 peptides that bear the most frequent and highly conserved 9-mers are indicated by arrows underneath the appropriate columns. Statistically robust clades in the dendrograms are indicated by a number at the branch point.

[0020] Figure 6B is a schematic depicting a heatmap for CD8 bifunctional (IFN-γ or TNF-α) T-cell responses to the 78 PTE peptide pools with small numbers of peptides. Light gray indicates no response, dark gray indicates a response to either IFN-γ or TNF-α, and black indicates a response to both IFN-γ and TNF-α. White squares indicate data that was not available. Only the 4 most conserved peptide pools demonstrated very low reactivity (arrows). Rows represent particular experiments, and columns represent particular peptide pools. Each duplicated symbol indicates a repeated experiment on a different date. The number of antigens in the polyvalent vaccine is indicated at the left on the key. The histogram shows the total number of each class of response (i.e., 0, 1, or 2 responses). The six peptide pools tested that

contained the 36 peptides that bear the most frequent and highly conserved 9-mers are indicated by arrows underneath the appropriate columns. Statistically robust clades in the dendrograms are indicated by a number at the branch point.

[0021] Figure 7 is a graph depicting CD4 and CD8 T-cell responses in mice immunized with gp160ΔCFI mos 2 and mos 3 mosaics, a gp160ΔCFI single mosaic, two gp160ΔCFI natural strains sets, and gp145ΔCFI clades A, B and C. Immunization and ICS were performed as described herein. The vaccination groups included: pCMV/R with no insert (control; 15 μg/animal), gp160 mos 1 (15 μg/animal), gp160ΔCFI mos 1 (15 μg/animal); gp160 mos 2.1 (7.5 μg/animal), gp160 mos 2.2 (7.5 μg/animal), gp160ΔCFI mos 2.1 (7.5 μg/animal), gp160ΔCFI mos 2.2 (7.5 μg/animal), gp160 mos 3.1 (5 μg/animal), gp160 mos 3.2 (5 μg/animal), gp160 mos 3.3 (5 μg/animal), gp160ΔCFI mos 3.1 (5 μg/animal), gp160ΔCFI mos 3.2 (5 μg/animal), gp160ΔCFI mos 3.3 (5 μg/animal); three gp160 natural strains (5 μg/animal), gp160ΔCFI natural strains (5 μg/animal); and gp145ΔCFI clades A, B and C (5 μg/animal). Percent positive CD4+ (left panel) and CD8+ (right panel) cells of the total T cell population was measured using ICS for IFN-γ (solid bars) and TNF-α (open bars). The minimal threshold response indicated by horizontal dashed lines was defined as two times the response from each group's un-stimulated samples (un). Gray bars indicate responses above the background, while black bars show the level of response to epitope pools below the background level of detection.

[0022] Figure 8 is a graph depicting CD4 and CD8 T-cell responses in mice immunized with gp160ΔCFIΔV mos 3 mosaics, gp160ΔVΔCFI natural strains, gp160 mos 2 mosaics, gp160 mos 3 mosaics, gp160ΔCFI mos 3 mosaics, gp160ΔCFI natural strains, and gp145ΔCFI clades A, B and C. Immunization and ICS were performed as described herein. Percent positive CD4+ (left panel) and CD8+ (right panel) cells of the total T cell population were measured using ICS for IFN-γ (solid bars) and TNF-α (open bars). gp160ΔVΔCFI mos 3 mosaics elicited similar breadth of CD8+ T-cell responses as gp160ΔCFI mos 3 mosaics, but less than gp160 mos 3 mosaics.

[0023] Figure 9 is a graph depicting CD4 (left panel) and CD8 (right panel) T-cell responses in mice immunized with gp145ΔCFI mos 1 mosaic, gp145ΔCFI mos 2 mosaics, gp145ΔCFI mos 3 mosaics, gp145ΔCFI natural strains, gp145ΔCFIΔVs natural strains, gp160 mos 2 mosaics, gp160 mos 3 mosaics, gp160ΔCFI mos 3 mosaics, and gp145ΔCFI clades A, B and C.

Immunization and ICS were performed as described herein. The gp145ΔCFI mos 1 mosaics, gp145ΔCFI mos 2 mosaics, and gp145ΔCFI mos 3 mosaics demonstrated minimal CD8 responses to PTE peptide pools.

[0024] Figure 10 is a graph depicting CD4 (left panel) and CD8 (right panel)T-cell responses in mice immunized with gp160ΔV mos 2 mosaics, gp160ΔV mos 3 mosaics, gp160 mos1 mosaics, gp160 mos2 mosaics, and gp160 mos 3 mosaics. Immunization, ICS, and labeling was performed as described herein. The gp160ΔV mos 2 mosaics and gp160ΔV mos 3 mosaics elicited fewer CD8 responses to PTE peptides than gp160 mos 2 or mos 3 mosaics.

DETAILED DESCRIPTION OF THE INVENTION

[0025] The invention provides an isolated or purified nucleic acid sequence encoding a HIV-1 Env polypeptide, which HIV-1 Env polypeptide comprises an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide. The nucleic acid sequence can encode any Env polypeptide known in the art. Suitable Env polypeptides are known in the art and include, for example, gp160, gp120, gp41, gp145, and gp140. The inventive nucleic acid sequence may include, but is not limited to, a nucleic acid sequence encoding a full-length Env polypeptide. Alternatively, the inventive nucleic acid sequence can encode any portion of an Env polypeptide, and preferably a portion of an Env polypeptide that is immunogenic (i.e., induces an immune response against HIV). In addition, the inventive nucleic acid sequence can encode a modified Env polypeptide that exhibits enhanced immunogenicity *in vivo*. For example, the nucleic acid sequence encoding an Env polypeptide can comprise mutations in the cleavage site, fusion peptide, or interhelical coiled-coil domains of the Env protein (ΔCFI Env proteins), which expose the core protein for optimal antigen presentation and recognition (see, e.g., U.S. Patent 7,470,430; Cao et al., *J. Virol.*, *71*: 9808-9812 (1997); Yang et al., *J. Virol.*, *78*: 4029-4036 (2004)). In addition, the Env polypeptide can lack the cytoplasmic domain of the Env protein. The Env polypeptide also can lack one or more variable loops of a wild-type Env polypeptide. For example, the inventive nucleic acid sequence preferably does not encode the variable loops 1, 2, 3, 4, or 5 of Env, or combinations thereof (see, e.g., International Patent Application Publication WO 2005/034992).

[0026] The nucleic acid sequence can encode an Env polypeptide from any group or clade of HIV. HIV-1 can be classified into four groups: the "major" group M, the "outlier" group O, group N, and group P. Preferably, the nucleic acid sequence encodes an Env polypeptide from group M. Within group M, there are several genetically distinct clades (or subtypes) of HIV-1. Thus, the nucleic acid sequence can encode an Env polypeptide from HIV-1 clade A, B, C, D, E, F, G, H, J, and K, and the like. HIV Gag, Env, and Pol proteins from the different HIV clades, as well as nucleic acid sequences encoding such proteins and methods for the manipulation and insertion of such nucleic acid sequences into vectors, are known (see, e.g., HIV Sequence Compendium, Division of AIDS, National Institute of Allergy and Infectious Diseases (2003); HIV Sequence Database (hiv-web.lanl.gov/content/hiv-db/mainpage.html); Sambrook et al., *Molecular Cloning, a Laboratory Manual*, 2d edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1989); Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing Associates and John Wiley & Sons, New York, N.Y. (1994)).

[0027] While the inventive nucleic acid sequence preferably encodes an Env polypeptide, in some embodiments the nucleic acid sequence encodes additional HIV polypeptides or antigens. Examples of other suitable HIV polypeptides include, but are not limited to, all or part of an HIV Gag, Pol, Tat, Reverse Transcriptase (RT), Vif, Vpr, Vpu, Vpo, Integrase, or Nef proteins. It will be appreciated that an entire, intact HIV protein is not required to produce an immune response. Indeed, most antigenic epitopes of HIV proteins are relatively small in size. Thus, nucleic acid sequences encoding fragments (e.g., epitopes or other antigenic fragments) of an HIV protein, such as any of the HIV proteins described herein, can be used in the context of the invention. Antigenic fragments and epitopes of the HIV Gag, Env, and Pol proteins, as well as nucleic acid sequences encoding such antigenic fragments and epitopes, are known (see, e.g., HIV Immunology and HIV/SIV Vaccine Databases, Vol. 1, Division of AIDS, National Institute of Allergy and Infectious Diseases (2003)).

[0028] The inventive nucleic acid sequence also may encode a fusion protein or polypeptide. A fusion protein can comprise all or part an HIV-1 Env polypeptide and all or part of a different HIV polypeptide fused together. The fusion protein can comprise all or part of any of the HIV polypeptides described herein. For example, all or part of an HIV Env protein (e.g., gp120 or gp160) can be fused to all or part of the HIV Pol protein, or all or part of HIV Gag protein can be

fused to all or part of the HIV Pol protein. Such fusion proteins effectively provide multiple HIV antigens in the context of the invention, and can be used to generate a more complete immune response against a given HIV pathogen as compared to that generated by a single HIV antigen. Similarly, polyproteins also can provide multiple HIV antigens. Polyproteins useful in conjunction with the invention include those that provide two or more HIV antigens, such as two or more of any of the HIV antigens described herein. Nucleic acid sequences encoding fusion proteins and polyproteins of HIV antigens can be prepared and inserted into vectors by known methods (see, e.g., U.S. Patents 5,130,247 and 5,130,248; Sambrook et al., *supra;* Ausubel et al., *supra*).

[0029] The terms "nucleic acid sequence," "nucleic acid," and "polynucleotide" are intended to encompass a polymer of DNA or RNA, i.e., a polynucleotide, which can be single-stranded or double-stranded and which can contain non-natural or altered nucleotides. In this respect, the terms include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs and modified polynucleotides such as, though not limited to, methylated and/or capped polynucleotides.

[0030] By "isolated" is meant the removal of a nucleic acid from its natural environment. By "purified" is meant that a given nucleic acid, whether one that has been removed from nature (including genomic DNA and mRNA) or synthesized (including cDNA) and/or amplified under laboratory conditions, has been increased in purity, wherein "purity" is a relative term and does not mean absolute purity. It is to be understood, however, that nucleic acids and proteins may be formulated with diluents or adjuvants and nevertheless for practical purposes be isolated.

[0031] As used herein a "codon" refers to the three nucleotides which, when transcribed and translated, encode a single amino acid residue or in the case of UUA, UGA, or UAG encode a termination signal. Codons encoding amino acids are well known in the art. The inventive nucleic acid sequence preferably comprises codons expressed more frequently in humans than in HIV. While the genetic code is generally universal across species, the choice among synonymous codons is often species-dependent. Infrequent usage of a particular codon by an organism likely reflects a low level of the corresponding transfer RNA (tRNA) in the organism. Thus, introduction of a nucleic acid sequence into an organism which comprises codons that are not frequently utilized in the organism may result in limited expression of the nucleic acid

sequence. One of ordinary skill in the art would appreciate that, to achieve an optimal immune response against HIV, the inventive nucleic acid sequence must be capable of expressing high levels of HIV-1 Env polypeptide in a human host. In this respect, the inventive nucleic acid sequence encodes an HIV-1 Env polypeptide, but comprises codons that are expressed more frequently in humans than in HIV-1. Such modified nucleic acid sequences are commonly described in the art as "humanized" or as utilizing "human-preferred" codons. Optimal codon usage is indicated by codon usage frequencies for expressed genes, as described in, for example, R. Nussinov, *J. Mol. Biol., 149*: 125-131 (1981).

[0032] The inventive nucleic acid sequence encodes an HIV-1 Env polypeptide which comprises an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide. A "T-cell epitope" is an amino acid sequence of an antigen that is recognized and bound by a T-cell receptor. A "potential T-cell epitope" is an amino acid sequence of an antigen that is hypothesized to be recognized and bound by a T-cell receptor. An "antigen" is a molecule that induces an immune response in a mammal. An "immune response" can entail, for example, antibody production and/or the activation of immune effector cells (e.g., T-cells). An antigen in the context of the invention can comprise any subunit, fragment, or epitope of any proteinaceous molecule, preferably a protein or peptide of HIV-1 which ideally provokes an immune response in mammal, preferably leading to protective immunity. By "epitope" is meant a sequence on an antigen that is recognized by an antibody or an antigen receptor. Epitopes also are referred to in the art as "antigenic determinants." T-cell epitopes typically are located in the inner, unexposed side of an antigen, and become accessible to the T-cell receptors after proteolytic processing of the antigen.

[0033] Potential T-cell epitopes (PTEs) of HIV-1 can be identified and/or synthetically generated using a variety of techniques known in the art (see, e.g., Sbai et al., *Curr. Drug Targets Infect. Disord., 1*(3): 303-313 (2001)). In one embodiment, T-cell epitopes of HIV-1 can be designed using a bioinformatic approach. For example, synthetic Env polypeptides can be designed to maximize the inclusion of common T-cell epitope peptides of nine amino acids ("9-mers"), and minimize the inclusion of rare epitopes likely to elicit strain-specific responses. Using routine recombinant DNA synthesis techniques, "mosaic" polypeptides can then be assembled containing the T-cell epitopes and natural HIV amino acid sequences. Mosaic

proteins resemble natural proteins, but maximize the coverage of T-cell epitopes (i.e., peptides of nine amino acids) for a viral population (see, e.g., Fischer et al., *Nat. Med.*, *13*(1): 100-106 (2007)).

[0034] In a preferred embodiment, the invention involves generating, *in silico*, recombinants of natural variants of HIV-1 Env group M polypeptides from, for example, the Los Alamos National Laboratory HIV database. Such recombinants can be scored and selected in combination to optimize the coverage of 9-mers in the global database for a given vaccine cocktail size. One of ordinary skill in the art will appreciate that, while mosaic polypeptides do not occur in nature, mosaic polypeptides align well to natural HIV proteins, and any short span of amino acids found in mosaic polypeptides will tend to be found repeatedly among natural strains. However, some of the hypervariable loop regions of Env are so variable that they are not repeated among circulating strains. Thus, in some embodiments of the invention it will be necessary to bridge such variable regions with amino acid sequences that are present in a single HIV strain. In one aspect of the invention, *in silico* recombination breakpoints in mosaic Env polypeptides are constrained to create fusion points found in natural sequences. It is possible to provide increased breadth of coverage with a single mosaic, providing maximum diversity coverage for stretches of nine amino acids. Alternatively, multiple mosaics can increase the breadth of representation and are encompassed by the invention.

[0035] The invention also provides a method of making a nucleic acid sequence encoding a modified HIV-1 Env polypeptide. The method comprises (a) identifying at least one T-cell epitope which is present in a natural Env polypeptide of a plurality of circulating HIV-1 strains, and (b) introducing a first nucleic acid sequence encoding the at least one T-cell epitope into a second nucleic acid sequence that encodes an HIV-1 Env polypeptide and that does not naturally encode the T-cell epitope. A modified Env polypeptide can be generated using standard recombinant DNA methodology (see, e.g., Sambrook et al., *supra*). For example, a nucleic acid sequence encoding a native (i.e., wild-type) Env polypeptide can be modified such that a nucleic acid sequence encoding one or more non-native (i.e., heterologous) T-cell epitopes (identified as described above) is inserted therein. In another embodiment, a recombinant nucleic acid sequence can be synthesized which encodes an Env polypeptide that comprises one or more T-cell epitopes that are frequently present in the HIV-1 strains circulating in the human population.

A T-cell epitope is frequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in at least 30% (e.g., at least 30%, at least 35%, or at least 40%) of circulating HIV-1 strains. More preferably, a T-cell epitope is frequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in at least 50% (e.g., at least 50%, at least 55%, or at least 60%) of circulating HIV-1 strains. Most preferably, a T-cell epitope is frequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in at least 70% (e.g., at least 70%, at least 75%, at least 80%, or at least 90%) of circulating HIV-1 strains. The inventive nucleic acid sequence can encode a HIV-1 Env polypeptide comprising one T-cell epitope or a plurality of T-cell epitopes (e.g., 2, 3, 5, 10, 20, or more T-cell epitopes). The Env polypeptides encoded by a nucleic acid sequence made in accordance with the invention are also referred to as "mosaic" polypeptides.

[0036] In order to accommodate an insertion of one or more nucleic acid sequences encoding non-native T-cell epitopes into an Env polypeptide, native (or natural) nucleic acids present in the wild-type Env nucleic acid sequence desirably are removed. For example, nucleic acid sequences encoding protein domains that do not contribute, or are detrimental, to the immunogenicity of the Env polypeptide can be removed. Alternatively and preferably, one or more nucleic acid sequences encoding a T-cell epitope that is native to the Env polypeptide but present infrequently in the Env proteins of HIV-1 strains circulating in the human population are removed, such that the Env polypeptide does not comprise at least one T-cell epitope that is present in the natural (i.e., wild-type) Env polypeptide. A T-cell epitope is infrequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in no more than 20% (e.g., no more than 20%, no more than 15%, or no more than 10%) of circulating HIV-1 strains. More preferably, a T-cell epitope is infrequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in no more than 5% (e.g., no more than 5%, or no more than 3%) of circulating HIV-1 strains. Most preferably, a T-cell epitope is infrequently present in HIV-1 strains circulating in the human population if the T-cell epitope is found in no more than 1% of circulating HIV-1 strains.

[0037] The amino acid sequence of the HIV-1 Env polypeptide encoded by the inventive nucleic acid sequence preferably comprises SEQ ID NOs: 1-6, SEQ ID NOs: 10-13, SEQ ID NOs: 17-22, SEQ ID NOs: 26-28, or SEQ ID NOs: 32-37.

[0038] The invention also provides a vector comprising the nucleic acid sequence described herein. A "vector" is a molecule, such as plasmid, phage, cosmid, liposome, molecular conjugate (e.g., transferrin), or virus, into which another nucleic acid sequence may be introduced so as to bring about the replication of the inserted sequence. Preferably, the vector is a plasmid or a viral vector. Suitable viral vectors include, for example, retroviral vectors, herpes simplex virus (HSV)-based vectors, parvovirus-based vectors, e.g., adeno-associated virus (AAV)-based vectors, AAV-adenoviral chimeric vectors, and adenovirus-based vectors. These viral vectors can be prepared using standard recombinant DNA techniques described in, for example, Sambrook et al., *supra*, and Ausubel et al., *supra*.

[0039] Retrovirus is an RNA virus capable of infecting a wide variety of host cells. Upon infection, the retroviral genome integrates into the genome of its host cell and is replicated along with host cell DNA, thereby constantly producing viral RNA and any nucleic acid sequence incorporated into the retroviral genome. As such, long-term expression of a therapeutic factor(s) is achievable when using retrovirus. Retroviruses contemplated for use in human gene transfer are relatively non-pathogenic, although pathogenic retroviruses exist. When employing pathogenic retroviruses, e.g., human immunodeficiency virus (HIV) or human T-cell lymphotrophic viruses (HTLV), care must be taken in altering the viral genome to eliminate toxicity to the host. A retroviral vector additionally can be manipulated to render the virus replication-deficient. As such, retroviral vectors are considered particularly useful for stable gene transfer *in vivo*.

[0040] An HSV-based viral vector is suitable for use as a vector to introduce a nucleic acid into numerous cell types. The mature HSV virion consists of an enveloped icosahedral capsid with a viral genome consisting of a linear double-stranded DNA molecule that is 152 kb. Most replication-deficient HSV vectors contain a deletion to remove one or more intermediate-early genes to prevent replication. Advantages of the HSV vector are its ability to enter a latent stage that can result in long-term DNA expression and its large viral DNA genome that can accommodate exogenous DNA inserts of up to 25 kb. HSV-based vectors are described in, for example, U.S. Patents 5,837,532, 5,846,782, 5,849,572, and 5,804,413, and International Patent Application Publications WO 91/02788, WO 96/04394, WO 98/15637, and WO 99/06583.

[0041] AAV vectors are viral vectors of particular interest for use in human gene transfer. AAV is a DNA virus, which is not known to cause human disease. The AAV genome is comprised of two genes, rep and cap, flanked by inverted terminal repeats (ITRs), which contain recognition signals for DNA replication and packaging of the virus. AAV requires co-infection with a helper virus (i.e., an adenovirus or a herpes simplex virus), or expression of helper genes, for efficient replication. AAV can be propagated in a wide array of host cells including human, simian, and rodent-cells, depending on the helper virus employed. An AAV vector used for administration of a nucleic acid sequence typically has approximately 96% of the parental genome deleted, such that only the ITRs remain. This eliminates immunologic or toxic side effects due to expression of viral genes. If desired, the AAV rep protein can be co-administered with the AAV vector to enable integration of the AAV vector into the host cell genome. Host cells comprising an integrated AAV genome show no change in cell growth or morphology (see, e.g., U.S. Patent 4,797,368). As such, prolonged expression of therapeutic factors from AAV vectors can be useful in treating persistent and chronic diseases.

[0042] In a preferred embodiment, the vector is an adenoviral vector. Adenoviruses are generally associated with benign pathologies in humans, and the 36 kilobase (kb) adenoviral genome has been extensively studied. Adenoviral vectors can be produced in high titers (e.g., about 10^{13} particle forming units (pfu)), and such vectors can transfer genetic material to nonreplicating, as well as replicating, cells; in contrast with, e.g., retroviral vectors, which only transfer genetic material to replicating cells. The adenoviral genome can be manipulated to carry a large amount of exogenous DNA (up to about 8 kb), and the adenoviral capsid can potentiate the transfer of even longer sequences (Curiel et al., *Hum. Gene Ther.*, *3*, 147-154 (1992)). Additionally, adenoviruses generally do not integrate into the host cell chromosome, but rather are maintained as a linear episome, thus minimizing the likelihood that a recombinant adenovirus will interfere with normal cell function. In addition to being a superior vehicle for transferring genetic material to a wide variety of cell types, adenoviral vectors represent a safe choice for gene transfer, a particular concern for therapeutic applications.

[0043] Adenovirus from various origins, subtypes, or mixture of subtypes can be used as the source of the viral genome for the adenoviral vector. While non-human adenovirus (e.g., simian, avian, canine, ovine, or bovine adenoviruses) can be used to generate the adenoviral vector, a

human adenovirus preferably is used as the source of the viral genome for the adenoviral vector of the invention. Adenovirus can be of various subgroups or serotypes. For instance, an adenovirus can be of subgroup A (e.g., serotypes 12, 18, and 31), subgroup B (e.g., serotypes 3, 7, 11, 14, 16, 21, 34, 35, and 50), subgroup C (e.g., serotypes 1, 2, 5, and 6), subgroup D (e.g., serotypes 8, 9, 10, 13, 15, 17, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 36-39, and 42-48), subgroup E (e.g., serotype 4), subgroup F (e.g., serotypes 40 and 41), an unclassified serogroup (e.g., serotypes 49 and 51), or any other adenoviral serotype. Adenoviral serotypes 1 through 51 are available from the American Type Culture Collection (ATCC, Manassas, VA). Preferably, the adenoviral vector is of human subgroup C, especially serotype 2 or even more desirably serotype 5. However, non-group C adenoviruses can be used to prepare adenoviral gene transfer vectors for delivery of gene products to host cells. Preferred adenoviruses used in the construction of non-group C adenoviral gene transfer vectors include Ad35 (group B), Ad26 (group D), and Ad28 (group D). Non-group C adenoviral vectors, methods of producing non-group C adenoviral vectors, and methods of using non-group C adenoviral vectors are disclosed in, for example, U.S. Patents 5,801,030, 5,837,511, and 5,849,561 and International Patent Application Publications WO 97/12986 and WO 98/53087.

[0044] The adenoviral vector of the invention can be replication-competent. For example, the adenoviral vector can have a mutation (e.g., a deletion, an insertion, or a substitution) in the adenoviral genome that does not inhibit viral replication in host cells. The adenoviral vector also can be conditionally replication-competent. Preferably, however, the adenoviral vector is replication-deficient in host cells.

[0045] By "replication-deficient" is meant that the adenoviral vector requires complementation of one or more regions of the adenoviral genome that are required for replication, as a result of, for example a deficiency in at least one replication-essential gene function (i.e., such that the adenoviral vector does not replicate in typical host cells, especially those in a human patient that could be infected by the adenoviral vector in the course of the inventive method). A deficiency in a gene, gene function, or genomic region, as used herein, is defined as a deletion of sufficient genetic material of the viral genome to obliterate or impair the function of the gene (e.g., such that the function of the gene product is reduced by at least about 2-fold, 5-fold, 10-fold, 20-fold, 30-fold, or 50-fold) whose nucleic acid sequence was deleted in

whole or in part. Deletion of an entire gene region often is not required for disruption of a replication-essential gene function. However, for the purpose of providing sufficient space in the adenoviral genome for one or more transgenes, removal of a majority of a gene region may be desirable. While deletion of genetic material is preferred, mutation of genetic material by addition or substitution also is appropriate for disrupting gene function. Replication-essential gene functions are those gene functions that are required for replication (e.g., propagation) and are encoded by, for example, the adenoviral early regions (e.g., the E1, E2, and E4 regions), late regions (e.g., the L1-L5 regions), genes involved in viral packaging (e.g., the IVa2 gene), and virus-associated RNAs (e.g., VA-RNA1 and/or VA-RNA-2).

[0046] The replication-deficient adenoviral vector desirably requires complementation of at least one replication-essential gene function of one or more regions of the adenoviral genome. Preferably, the adenoviral vector requires complementation of at least one gene function of the E1A region, the E1B region, or the E4 region of the adenoviral genome required for viral replication (denoted an E1-deficient or E4-deficient adenoviral vector). In addition to a deficiency in the E1 region, the recombinant adenovirus also can have a mutation in the major late promoter (MLP), as discussed in International Patent Application Publication WO 00/00628. Most preferably, the adenoviral vector is deficient in at least one replication-essential gene function (desirably all replication-essential gene functions) of the E1 region and at least one gene function of the nonessential E3 region (e.g., an Xba I deletion of the E3 region) (denoted an E1/E3-deficient adenoviral vector). With respect to the E1 region, the adenoviral vector can be deficient in part or all of the E1A region and/or part or all of the E1B region, e.g., in at least one replication-essential gene function of each of the E1A and E1B regions, thus requiring complementation of the E1A region and the E1B region of the adenoviral genome for replication. The adenoviral vector also can require complementation of the E4 region of the adenoviral genome for replication, such as through a deficiency in one or more replication-essential gene functions of the E4 region.

[0047] When the adenoviral vector is deficient in at least one replication-essential gene function in one region of the adenoviral genome (e.g., an E1- or E1/E3-deficient adenoviral vector), the adenoviral vector is referred to as "singly replication-deficient." A particularly preferred singly replication-deficient adenoviral vector is, for example, a replication-deficient

adenoviral vector requiring, at most, complementation of the E1 region of the adenoviral genome, so as to propagate the adenoviral vector (e.g., to form adenoviral vector particles).

[0048] The adenoviral vector of the invention can be "multiply replication-deficient," meaning that the adenoviral vector is deficient in one or more replication-essential gene functions in each of two or more regions of the adenoviral genome, and requires complementation of those functions for replication. For example, the aforementioned E1-deficient or E1/E3-deficient adenoviral vector can be further deficient in at least one replication-essential gene function of the E4 region (denoted an E1/E4- or E1/E3/E4-deficient adenoviral vector), and/or the E2 region (denoted an E1/E2- or E1/E2/E3-deficient adenoviral vector), preferably the E2A region (denoted an E1/E2A- or E1/E2A/E3-deficient adenoviral vector). An adenoviral vector deleted of the entire E4 region can elicit a lower host immune response.

[0049] In one embodiment of the invention, the adenoviral vector can comprise an adenoviral genome deficient in one or more replication-essential gene functions of each of the E1 and E4 regions (i.e., the adenoviral vector is an E1/E4-deficient adenoviral vector), preferably with the entire coding region of the E4 region having been deleted from the adenoviral genome. In other words, all the open reading frames (ORFs) of the E4 region have been removed. Most preferably, the adenoviral vector is rendered replication-deficient by deletion of all of the E1 region and by deletion of a portion of the E4 region. The E4 region of the adenoviral vector can retain the native E4 promoter, polyadenylation sequence, and/or the right-side inverted terminal repeat (ITR).

[0050] The adenoviral vector, when multiply replication-deficient, especially in replication-essential gene functions of the E1 and E4 regions, can include a spacer sequence to provide viral growth in a complementing cell line similar to that achieved by singly replication-deficient adenoviral vectors, particularly an E1-deficient adenoviral vector. In a preferred E4-deficient adenoviral vector of the invention wherein the L5 fiber region is retained, the spacer is desirably located between the L5 fiber region and the right-side ITR. More preferably in such an adenoviral vector, the E4 polyadenylation sequence alone or, most preferably, in combination with another sequence exists between the L5 fiber region and the right-side ITR, so as to sufficiently separate the retained L5 fiber region from the right-side ITR, such that viral

production of such a vector approaches that of a singly replication-deficient adenoviral vector, particularly a singly replication-deficient E1 deficient adenoviral vector.

[0051] The spacer sequence can contain any nucleotide sequence or sequences which are of a desired length, such as sequences at least about 15 base pairs (e.g., between about 15 base pairs and about 12,000 base pairs), preferably about 100 base pairs to about 10,000 base pairs, more preferably about 500 base pairs to about 8,000 base pairs, even more preferably about 1,500 base pairs to about 6,000 base pairs, and most preferably about 2,000 to about 3,000 base pairs in length. The spacer sequence can be coding or non-coding and native or non-native with respect to the adenoviral genome, but does not restore the replication-essential function to the deficient region. The spacer can also contain a promoter-variable expression cassette. More preferably, the spacer comprises an additional polyadenylation sequence and/or a passenger gene. Preferably, in the case of a spacer inserted into a region deficient for E4, both the E4 polyadenylation sequence and the E4 promoter of the adenoviral genome or any other (cellular or viral) promoter remain in the vector. The spacer is located between the E4 polyadenylation site and the E4 promoter, or, if the E4 promoter is not present in the vector, the spacer is proximal to the right-side ITR. The spacer can comprise any suitable polyadenylation sequence. Examples of suitable polyadenylation sequences include synthetic optimized sequences, BGH (Bovine Growth Hormone), Polyoma virus, TK (Thymidine Kinase), EBV (Epstein Barr Virus), and the papillomaviruses, including human papillomaviruses and BPV (Bovine Papilloma Virus). Preferably, particularly in the E4 deficient region, the spacer includes an SV40 Polyadenylation sequence. The SV40 polyadenylation sequence allows for higher virus production levels of multiply replication deficient adenoviral vectors. In the absence of a spacer, production of fiber protein and/or viral growth of the multiply replication-deficient adenoviral vector is reduced by comparison to that of a singly replication-deficient adenoviral vector. However, inclusion of the spacer in at least one of the deficient adenoviral regions, preferably the E4 region, can counteract this decrease in fiber protein production and viral growth. Ideally, the spacer comprises the glucuronidase gene. The use of a spacer in an adenoviral vector is further described in, for example, U.S. Patent 5,851,806 and International Patent Application Publication WO 97/21826.

[0052] Desirably, the adenoviral vector requires, at most, complementation of replication-essential gene functions of the E1, E2A, and/or E4 regions of the adenoviral genome for

replication (i.e., propagation). However, the adenoviral genome can be modified to disrupt one or more replication-essential gene functions as desired by the practitioner, so long as the adenoviral vector remains deficient and can be propagated using, for example, complementing cells and/or exogenous DNA (e.g., helper adenovirus) encoding the disrupted replication-essential gene functions. In this respect, the adenoviral vector can be deficient in replication-essential gene functions of only the early regions of the adenoviral genome, only the late regions of the adenoviral genome, both the early and late regions of the adenoviral genome, or all adenoviral genes (i.e., a high capacity adenovector (HC-Ad); see Morsy et al., *Proc. Natl. Acad. Sci. USA*, *95*: 965-976 (1998); Chen et al., *Proc. Natl. Acad. Sci USA*, *94*: 1645-1650 (1997); Kochanek et al., *Hum. Gene Ther.*, *10*: 2451-2459 (1999)). Suitable replication-deficient adenoviral vectors, including singly and multiply replication-deficient adenoviral vectors, are disclosed in U.S. Patents 5,837,511, 5,851,806, 5,994,106, 6,127,175, 6,482,616, and 7,195,896; U.S. Patent Application Publications 2001/0043922 A1, 2002/0004040 A1, 2002/0110545 A1, and 2004/0161848 A1; and International Patent Application Publications WO 94/28152, WO 95/02697, WO 95/16772, WO 95/34671, WO 96/22378, WO 97/12986, WO 97/21826, and WO 03/022311.

[0053] By removing all or part of, for example, the E1, E3, and E4 regions of the adenoviral genome, the resulting adenoviral vector is able to accept inserts of exogenous nucleic acid sequences while retaining the ability to be packaged into adenoviral capsids. The nucleic acid sequence can be positioned in the E1 region, the E3 region, or the E4 region of the adenoviral genome. Indeed, the nucleic acid sequence can be inserted anywhere in the adenoviral genome so long as the position does not prevent expression of the nucleic acid sequence or interfere with packaging of the adenoviral vector.

[0054] Replication-deficient adenoviral vectors are typically produced in complementing cell lines that provide gene functions not present in the replication-deficient adenoviral vectors, but required for viral propagation, at appropriate levels in order to generate high titers of viral vector stock. Complementing cell lines for producing the adenoviral vector include, but are not limited to, 293 cells (described in, e.g., Graham et al., *J. Gen. Virol.*, *36*, 59-72 (1977)), PER.C6 cells (described in, e.g., International Patent Application Publication WO 97/00326, and U.S. Patents 5,994,128 and 6,033,908), and 293-ORF6 cells (described in, e.g., International Patent

Application Publication WO 95/34671 and Brough et al., *J. Virol.*, *71*: 9206-9213 (1997)). Additional complementing cells are described in, for example, U.S. Patents 6,677,156 and 6,682,929, and International Patent Application Publication WO 03/20879. In some instances, the cellular genome need not comprise nucleic acid sequences, the gene products of which complement for all of the deficiencies of a replication-deficient adenoviral vector. One or more replication-essential gene functions lacking in a replication-deficient adenoviral vector can be supplied by a helper virus, e.g., an adenoviral vector that supplies in *trans* one or more essential gene functions required for replication of the desired adenoviral vector.

[0055] If the adenoviral vector is not replication-deficient, ideally the adenoviral vector is manipulated to limit replication of the vector to within a target tissue. The adenoviral vector can be a conditionally-replicating adenoviral vector, which is engineered to replicate under conditions pre-determined by the practitioner. For example, replication-essential gene functions, e.g., gene functions encoded by the adenoviral early regions, can be operably linked to an inducible, repressible, or tissue-specific transcription control sequence, e.g., promoter. In this embodiment, replication requires the presence or absence of specific factors that interact with the transcription control sequence. Conditionally-replicating adenoviral vectors are described further in U.S. Patent 5,998,205.

[0056] The coat protein of an adenoviral vector can be manipulated to alter the binding specificity or recognition of the virus for a viral receptor on a potential host cell. For adenovirus, such manipulations can include deletion of regions of the fiber, penton, or hexon, insertions of various native or non-native ligands into portions of the coat protein, and the like. Manipulation of the coat protein can broaden the range of cells infected by the adenoviral vector or enable targeting of the adenoviral vector to a specific cell type.

[0057] Any suitable technique for altering native binding to a host cell, such as native binding of the fiber protein to the coxsackievirus and adenovirus receptor (CAR) of a cell, can be employed (see, e.g., U.S. Patent Application Publication 2009/0148477, and U.S. Patent 5,962,311). In addition, the nucleic acid residues encoding amino acid residues associated with native substrate binding can be changed, supplemented, or deleted (see, e.g., International Patent Application Publication WO 00/15823; Einfeld et al., *J. Virol.*, *75*(23): 11284-11291 (2001); van Beusechem et al., *J. Virol.*, *76*(6): 2753-2762 (2002)) such that the adenoviral vector

incorporating the mutated nucleic acid residues (or having the fiber protein encoded thereby) is less able to bind its native substrate. In this respect, the native CAR and integrin binding sites of the adenoviral vector, such as the knob domain of the adenoviral fiber protein and an Arg-Gly-Asp (RGD) sequence located in the adenoviral penton base, respectively, can be removed or disrupted. In one embodiment, the adenoviral vector comprises a fiber protein and a penton base protein that do not bind to CAR and integrins, respectively. Alternatively, the adenoviral vector comprises fiber protein and a penton base protein that bind to CAR and integrins, respectively, but with less affinity than the corresponding wild type coat proteins. The adenoviral vector exhibits reduced binding to CAR and integrins if a modified adenoviral fiber protein and penton base protein binds CAR and integrins, respectively, with at least about 5-fold, 10-fold, 20-fold, 30-fold, 50-fold, or 100-fold less affinity than a non-modified adenoviral fiber protein and penton base protein of the same serotype.

[0058] The adenoviral vector also can comprise a chimeric coat protein comprising a non-native amino acid sequence that binds a substrate (i.e., a ligand), such as a cellular receptor other than CAR the αv integrin receptor. Such a chimeric coat protein allows an adenoviral vector to bind, and desirably, infect host cells not naturally infected by the corresponding adenovirus that retains the ability to bind native cell surface receptors, thereby further expanding the repertoire of cell types infected by the adenoviral vector. A "non-native" amino acid sequence can comprise an amino acid sequence not naturally present in the adenoviral coat protein or an amino acid sequence found in the adenoviral coat but located in a non-native position within the capsid. By "preferentially binds" is meant that the non-native amino acid sequence binds a receptor, such as, for instance, αvβ3 integrin, with at least about 3-fold greater affinity (e.g., at least about 5-fold, 10-fold, 15-fold, 20-fold, 25-fold, 35-fold, 45-fold, or 50-fold greater affinity) than the non-native ligand binds a different receptor, such as, for instance, αvβ1 integrin.

[0059] Desirably, the adenoviral vector comprises a chimeric coat protein comprising a non-native amino acid sequence that confers to the chimeric coat protein the ability to bind to an immune cell more efficiently than a wild-type adenoviral coat protein. In particular, the adenoviral vector can comprise a chimeric adenoviral fiber protein comprising a non-native amino acid sequence which facilitates uptake of the adenoviral vector by immune cells, preferably antigen presenting cells, such as dendritic cells, monocytes, and macrophages. In a

preferred embodiment, the adenoviral vector comprises a chimeric fiber protein comprising an amino acid sequence (e.g., a non-native amino acid sequence) comprising an RGD motif including, but not limited to, CRGDC (SEQ ID NO: 45), CXCRGDCXC (SEQ ID NO: 46), wherein X represents any amino acid, and CDCRGDCFC (SEQ ID NO: 47), which increases transduction efficiency of an adenoviral vector into dendritic cells. The RGD-motif, or any non-native amino acid sequence, preferably is inserted into the adenoviral fiber knob region, ideally in an exposed loop of the adenoviral knob, such as the HI loop. A non-native amino acid sequence also can be appended to the C-terminus of the adenoviral fiber protein, optionally via a spacer sequence. The spacer sequence preferably comprises between one and two-hundred amino acids, and can (but need not) have an intended function.

[0060] In another embodiment, the adenoviral vector can comprise a chimeric virus coat protein that is not selective for a specific type of eukaryotic cell. The chimeric coat protein differs from a wild-type coat protein by an insertion of a non-native amino acid sequence into or in place of an internal coat protein sequence, or attachment of a non-native amino acid sequence to the N- or C- terminus of the coat protein (see, e.g., U.S. Patent 6,465,253 and International Patent Application Publication WO 97/20051).

[0061] A non-native amino acid sequence can be conjugated to any of the adenoviral coat proteins to form a chimeric adenoviral coat protein. Therefore, for example, a non-native amino acid sequence can be conjugated to, inserted into, or attached to a fiber protein, a penton base protein, a hexon protein, proteins IX, VI, or IIIa, etc. The sequences of such proteins, and methods for employing them in recombinant proteins, are well known in the art (see, e.g., U.S. Patents 5,543,328; 5,559,099; 5,712,136; 5,731,190; 5,756,086; 5,770,442; 5,846,782; 5,962,311; 5,965,541; 5,846,782; 6,057,155; 6,127,525; 6,153,435; 6,329,190; 6,455,314; 6,465,253; 6,576,456; 6,649,407; and 6,740,525, and International Patent Application Publications WO 96/07734, WO 96/26281, WO 97/20051, WO 98/07877, WO 98/07865, WO 98/40509, WO 98/54346, WO 00/15823, WO 01/58940, and WO 01/92549). The chimeric adenoviral coat protein can be generated using standard recombinant DNA techniques known in the art. Preferably, the nucleic acid sequence encoding the chimeric adenoviral coat protein is located within the adenoviral genome and is operably linked to a promoter that regulates expression of the coat protein in a wild-type adenovirus. Alternatively, the nucleic acid sequence

encoding the chimeric adenoviral coat protein is located within the adenoviral genome and is part of an expression cassette which comprises genetic elements required for efficient expression of the chimeric coat protein.

[0062] Disruption of native binding of adenoviral coat proteins to a cell surface receptor can also render it less able to interact with the innate or acquired host immune system. Aside from pre-existing immunity, adenoviral vector administration induces inflammation and activates both innate and acquired immune mechanisms. Adenoviral vectors activate antigen-specific (e.g., T-cell dependent) immune responses, which limit the duration of transgene expression following an initial administration of the vector. In addition, exposure to adenoviral vectors stimulates production of neutralizing antibodies by B cells, which can preclude gene expression from subsequent doses of adenoviral vector (Wilson & Kay, *Nat. Med.*, *3*(9): 887-889 (1995)). Indeed, the effectiveness of repeated administration of the vector can be severely limited by host immunity. In addition to stimulation of humoral immunity, cell-mediated immune functions are responsible for clearance of the virus from the body. Rapid clearance of the virus is attributed to innate immune mechanisms (see, e.g., Worgall et al., *Human Gene Therapy*, *8*: 37-44 (1997)), and likely involves Kupffer cells found within the liver. Thus, by ablating native binding of an adenovirus fiber protein and penton base protein, immune system recognition of an adenoviral vector is diminished, thereby increasing vector tolerance by the host.

[0063] Another method for evading pre-existing host immunity to adenovirus, especially serotype 5 adenovirus, involves modifying an adenoviral coat protein such that it exhibits reduced recognition by the host immune system. Thus, the adenoviral vector preferably comprises such a modified coat protein. The modified coat protein preferably is a penton, fiber, or hexon protein. Most preferably, the modified coat protein is a hexon protein. The coat protein can be modified in any suitable manner, but is preferably modified by generating diversity in the coat protein. Preferably, such coat protein variants are not recognized by pre-existing host (e.g., human) adenovirus-specific neutralizing antibodies. Diversity can be generated using any suitable method known in the art, including, for example, directed evolution (i.e., polynucleotide shuffling) and error-prone PCR (see, e.g., Cadwell, *PCR Meth. Appl.*, *2*: 28-33 (1991); Leung et al., *Technique*, *1*: 11-15 (1989); Pritchard et al., *J. Theoretical Biol.*, *234*: 497-509 (2005)). Preferably, coat protein diversity is generated through directed evolution

techniques, such as those described in, e.g., Stemmer, *Nature*, *370*: 389-91 (1994); Cherry et al., *Nat. Biotechnol.*, *17*: 379-84 (1999); Schmidt-Dannert et al., *Nat Biotechnol.*, *18*(7): 750-53 (2000); U.S. Patent Application Publication 2009/0148477.

[0064] An adenoviral coat protein also can be modified to evade pre-existing host immunity by deleting a region of a coat protein and replacing it with a corresponding region from the coat protein of another adenovirus serotype, particularly a serotype which is less immunogenic in humans. In this regard, amino acid sequences within the fiber protein, the penton base protein, and/or the hexon protein can be removed and replaced with corresponding sequences from a different adenovirus serotype. Thus, for example, when the fiber protein is modified to evade pre-existing host immunity, amino acid residues from the knob region of a serotype 5 fiber protein can be deleted and replaced with corresponding amino acid residues from an adenovirus of a different serotype, such as those serotypes described herein. Likewise, when the penton base protein is modified to evade pre-existing host immunity, amino acid residues within the hypervariable region of a serotype 5 penton base protein can be deleted and replaced with corresponding amino acid residues from an adenovirus of a different serotype, such as those serotypes described herein. Preferably, the hexon protein of the adenoviral vector is modified in this manner to evade pre-existing host immunity. In this respect, when the adenoviral vector is of serotype 5, amino acid residues within one or more of the hypervariable regions, which occur in loops of the hexon protein, are removed and replaced with corresponding amino acid residues from an adenovirus of a different serotype. Preferably, amino acid residues within the FG1, FG2, or DE1 loops of a serotype 5 hexon protein are deleted and replaced with corresponding amino acid residues from a hexon protein of a different adenovirus serotype. An entire loop region can be removed from the serotype 5 hexon protein and replaced with the corresponding loop region of another adenovirus serotype. Alternatively, portions of a loop region can be removed from the serotype 5 hexon protein and replaced with the corresponding portion of a hexon loop of another adenovirus serotype. One or more hexon loops, or portions thereof, of a serotype 5 adenoviral vector can be removed and replaced with the corresponding sequences from any other adenovirus serotype, such as those described herein. The structure of Ad2 and Ad5 hexon proteins and methods of modifying hexon proteins are disclosed in, for example, Rux et al., *J. Virol.*, *77*: 9553-9566 (2003), and U.S. Patent 6,127,525. The hypervariable regions of a

hexon protein also can be replaced with random peptide sequences, or peptide sequences derived from a disease-causing pathogen (e.g., HIV).

[0065] Suitable modifications to an adenoviral vector are described in U.S. Patents 5,543,328; 5,559,099; 5,712,136; 5,731,190; 5,756,086; 5,770,442; 5,846,782; 5,871,727; 5,885,808; 5,922,315; 5,962,311; 5,965,541; 6,057,155; 6,127,525; 6,153,435; 6,329,190; 6,455,314; 6,465,253; 6,576,456; 6,649,407; 6,740,525; and 6,951,755, U.S. Patent Application Publication 2003/0099619 A1, and International Patent Applications WO 96/07734, WO 96/26281, WO 97/20051, WO 98/07865, WO 98/07877, WO 98/40509, WO 98/54346, WO 00/15823, WO 01/58940, and WO 01/92549.

[0066] The vectors of the invention comprise a nucleic acid sequence encoding any of the HIV-1 Env polypeptides described herein in a form suitable for expression of the nucleic acid sequence in a host cell, which means that the vectors include one or more sequences which regulate expression of the nucleic acid sequence. Such regulatory sequences are operatively linked to the nucleic acid sequence to be expressed. By "operably linked" is meant that the nucleic acid sequence is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleic acid sequence. The term "regulatory sequence" is intended to include promoters, enhancers, and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Gene Expression Technology: Methods in Enzymology,* 185, Academic Press, San Diego, Calif. (1990).

[0067] Any promoter or enhancer sequence can be used in the context of the invention, so long as sufficient expression of the inventive nucleic acid sequence is achieved and a robust immune response against the encoded polypeptide is generated. In this regard, the promoter can be a viral promoter. Suitable viral promoters include, for example, cytomegalovirus (CMV) promoters, such as the mouse CMV immediate-early promoter (mCMV) or the human CMV immediate-early promoter (hCMV) (described in, for example, U.S. Patents 5,168,062 and 5,385,839), Rous sarcoma virus (RSV) promoters, such as the RSV long terminal repeat, mouse mammary tumor virus (MMTV) promoters, HSV promoters, such as the Lap2 promoter or the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci., 78*: 144-145 (1981)), promoters derived from SV40 or Epstein Barr virus, an adeno-associated viral promoter, such as

the p5 promoter, and the like. Preferably, the promoter is the CMV immediate-early promoter (mouse or human).

[0068] Alternatively, the promoter can be a cellular promoter, i.e., a promoter that is native to eukaryotic, preferably animal, cells. In one aspect, the cellular promoter is preferably a constitutive promoter that works in a variety of cell types, such as cells associated with the immune system. Suitable constitutive promoters can drive expression of genes encoding transcription factors, housekeeping genes, or structural genes common to eukaryotic cells. Suitable cellular promoters include, for example, a ubiquitin promoter (e.g., a UbC promoter) (see, e.g., Marinovic et al., *J. Biol. Chem.*, *277*(19): 16673-16681 (2002)), a human β-actin promoter, an EF-1α promoter, a YY1 promoter, a basic leucine zipper nuclear factor-1 (BLZF-1) promoter, a neuron specific enolase (NSE) promoter, a heat shock protein 70B (HSP70B) promoter, and a JEM-1 promoter.

[0069] Many of the above-described promoters are constitutive promoters. Instead of being a constitutive promoter, the promoter can be an inducible promoter, i.e., a promoter that is up- and/or down-regulated in response to an appropriate signal. The use of a regulatable promoter or expression control sequence is particularly applicable to DNA vaccine development inasmuch as antigenic proteins, including viral and parasite antigens, frequently are toxic to cell lines in which they are produced. A promoter can be up-regulated by a radiant energy source or by a substance that distresses cells. For example, a promoter can be up-regulated by drugs, hormones, ultrasound, light activated compounds, radiofrequency, chemotherapy, and cyofreezing. Thus, a promoter sequence that regulates expression of the inventive nucleic acid sequence can contain at least one heterologous regulatory sequence responsive to regulation by an exogenous agent. Suitable inducible promoter systems include, but are not limited to, the IL-8 promoter, the metallothionine inducible promoter system, the bacterial lacZYA expression system, the tetracycline expression system, and the T7 polymerase system. Further, promoters that are selectively activated at different developmental stages (e.g., globin genes are differentially transcribed from globin-associated promoters in embryos and adults) can be employed.

[0070] The promoter can be a tissue-specific promoter, i.e., a promoter that is preferentially activated in a given tissue and results in expression of a gene product in the tissue where activated. A tissue-specific promoter suitable for use in the invention can be chosen by the

ordinarily skilled artisan based upon the target tissue or cell-type. Preferred tissue-specific promoters for use in the inventive method are specific to immune cells.

[0071] To optimize protein production, preferably the inventive nucleic acid sequence further comprises a polyadenylation site 3' of the coding sequence. Any suitable polyadenylation sequence can be used, including a synthetic optimized sequence, as well as the polyadenylation sequence of SV40 (Human Sarcoma Virus-40), BGH (Bovine Growth Hormone), mouse globin D (MGD), polyoma virus, TK (Thymidine Kinase), EBV (Epstein Barr Virus), and the papillomaviruses, including human papillomaviruses and BPV (Bovine Papilloma Virus). Also, preferably all of the proper transcription signals (and translation signals, where appropriate) are correctly arranged such that the nucleic acid sequence is properly expressed in the cells into which it is introduced. If desired, the nucleic acid sequence also can incorporate splice sites (i.e., splice acceptor and splice donor sites) to facilitate mRNA production.

[0072] The invention provides an isolated host cell comprising the nucleic acid sequence of the invention, or a vector comprising the nucleic acid sequence of the invention. For example, the nucleic acid sequence or vector can be expressed in prokaryotic cells, such as *E. coli*. Preferably, the nucleic acid sequence or vector is expressed in eukaryotic cells, such as insect cells (using baculovirus expression vectors), yeast cells, or mammalian cells (e.g., Chinese hamster ovary (CHO) cells, 293 cells, COS cells, or other human cells). Suitable host cells are discussed further in Goeddel, *supra*. Nucleic acid sequences and vectors comprising nucleic acid sequences can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. Such techniques include, for example, calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al., *supra*.

[0073] An isolated host cell, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) any of the HIV-1 Env polypeptides encoded by the nucleic acid sequences described herein Accordingly, the invention further provides methods for producing a modified HIV-1 Env polypeptide using the host cells of the invention.

[0074] The invention further provides a method of inducing an immune response against HIV-1 in a mammal (preferably a human). In one embodiment, the method comprises administering to a mammal the HIV-1 Env-encoding nucleic acid sequences described herein. In another embodiment, the method comprises administering to a mammal a composition comprising the HIV-1 Env-encoding nucleic acid sequence or vector described herein. In yet another embodiment, the method comprises administering to a mammal the HIV-1 Env polypeptide encoded by the nucleic acid sequence described herein.

[0075] The inventive nucleic acid sequence, or a vector comprising the inventive nucleic acid sequence, desirably is administered in a composition, preferably a pharmaceutically acceptable (e.g., physiologically acceptable) composition, which comprises a carrier, preferably a pharmaceutically (e.g., physiologically acceptable) carrier and the nucleic acid sequence, vector, or polypeptide. Therefore, the invention provides a composition capable of eliciting an immune response against HIV. The composition can be capable of eliciting a protective immune response against HIV when administered alone, or in combination with at least one additional immunogenic agent or composition. It will be understood by those of skill in the art that the ability to produce an immune response after exposure to an antigen is a function of complex cellular and humoral processes, and that different subjects have varying capacity to respond to an immunological stimulus. Accordingly, the compositions disclosed herein are capable of eliciting an immune response in an immunocompetent subject, that is a subject that is physiologically capable of responding to an immunological stimulus by the production of a substantially normal immune response, e.g., including the production of antibodies that specifically interact with the immunological stimulus, and/or the production of functional T-cells ($CD4^+$ and/or $CD8^+$ T-cells) that bear receptors that specifically interact with the immunological stimulus. It will further be understood that a particular effect of infection with HIV is to render a previously immunocompetent subject immunodeficient. Thus, with respect to the methods discussed herein, it is generally desirable to administer the compositions to a subject prior to exposure to HIV (that is, prophylactically, e.g., as a vaccine) or therapeutically at a time following exposure to HIV during which the subject is nonetheless capable of developing an immune response to a stimulus, such as an antigenic polypeptide.

[0076] Suitable formulations for the composition include aqueous and non-aqueous solutions, isotonic sterile solutions, which can contain anti-oxidants, buffers, and bacteriostats, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example, water, immediately prior to use. Extemporaneous solutions and suspensions can be prepared from sterile powders, granules, and tablets. Preferably, the carrier is a buffered saline solution. More preferably, the composition is formulated to protect the nucleic acid sequence or vector from damage prior to administration. For example, the pharmaceutical composition can be formulated to reduce loss of the nucleic acid or vector on devices used to prepare, store, or administer the composition, such as glassware, syringes, or needles. The composition can be formulated to decrease the light sensitivity and/or temperature sensitivity of the nucleic acid sequence or vector. To this end, the composition preferably comprises a pharmaceutically acceptable liquid carrier, such as, for example, those described above, and a stabilizing agent selected from the group consisting of Polysorbate 80, L-arginine, polyvinylpyrrolidone, trehalose, and combinations thereof. Use of such a composition will extend the shelf life of the nucleic acid sequence or vector, facilitate administration, and increase the efficiency of the inventive method.

[0077] A composition also can be formulated to enhance transduction efficiency of the nucleic acid sequence or vector. In addition, one of ordinary skill in the art will appreciate that the composition can comprise other therapeutic or biologically-active agents. For example, factors that control inflammation, such as ibuprofen or steroids, can be part of the composition to reduce swelling and inflammation associated with *in vivo* administration of the composition. Antibiotics, i.e., microbicides and fungicides, can be present to treat existing infection and/or reduce the risk of future infection, such as infection associated with gene transfer procedures.

[0078] The composition also can be formulated to contain an adjuvant in order to enhance the immunological response. Suitable adjuvants include, but are not limited to, mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, other peptides, oil emulsions, and potentially useful human adjuvants such as

Bacillus Calmette Guerin (BCG) and Corynebacterium parvum. Adjuvants for inclusion in the inventive composition desirably are safe, well tolerated, and effective in humans, such as QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjumer, PG-026, GSK-1, GcMAF, B-alethine, MPC-026, Adjuvax, CpG ODN, Betafectin, Alum, and MF59 (as described in, e.g., Kim et al., *Vaccine*, *18*: 597 (2000)). Other adjuvants that can be administered to a mammal include lectins, growth factors, cytokines, and lymphokines (e.g., alpha-interferon, gamma-interferon, platelet derived growth factor (PDGF), gCSF, gMCSF, TNF, epidermal growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10, and IL-12).

[0079] Any route of administration can be used to deliver the composition to the mammal. Indeed, although more than one route can be used to administer the composition, a particular route can provide a more immediate and more effective reaction than another route. Preferably, the composition is administered via intramuscular injection, for example, using a syringe or needleless delivery device. The pharmaceutical composition also can be applied or instilled into body cavities, absorbed through the skin (e.g., via a transdermal patch), inhaled, ingested, topically applied to tissue, or administered parenterally via, for instance, intravenous, peritoneal, or intraarterial administration.

[0080] The composition can be administered in or on a device that allows controlled or sustained release, such as a sponge, biocompatible meshwork, mechanical reservoir, or mechanical implant. Implants (see, e.g., U.S. Patent 5,443,505), devices (see, e.g., U.S. Patent 4,863,457), such as an implantable device, e.g., a mechanical reservoir or an implant or a device comprised of a polymeric composition, are particularly useful for administration of the composition. The composition also can be administered in the form of a sustained-release formulation (see, e.g., U.S. Patent 5,378,475) comprising, for example, gel foam, hyaluronic acid, gelatin, chondroitin sulfate, a polyphosphoester, such as bis-2-hydroxyethyl-terephthalate (BHET), and/or a polylactic-glycolic acid.

[0081] The dose of the composition administered to the mammal will depend on a number of factors, including the size of a target tissue, the extent of any side-effects, the particular route of administration, and the like. The dose ideally comprises an "effective amount" of the composition, i.e., a dose of composition which provokes a desired immune response in the

mammal. The desired immune response can entail production of antibodies, protection upon subsequent challenge, immune tolerance, immune cell activation, and the like. One dose or multiple doses of the composition can be administered to a mammal to elicit an immune response with desired characteristics, including the production of HIV specific antibodies, or the production of functional T-cells that react with HIV. In certain embodiments, the T-cells may be CD8 T-cells.

[0082] When the nucleic acid sequence encoding a HIV-1 Env polypeptide is administered to a mammal via an adenoviral vector, the composition desirably comprises a single dose of adenoviral vector comprising at least about $1x10^5$ particles (which also is referred to as particle units) of adenoviral vector. The dose preferably is at least about $1x10^6$ particles (e.g., about $1x10^6$-$1x10^{12}$ particles), more preferably at least about $1x10^7$ particles, even more preferably at least about $1x10^8$ particles (e.g., about $1x10^8$-$1x10^{11}$ particles or about $1x10^8$-$1x10^{12}$ particles), and most preferably at least about $1x10^9$ particles (e.g., about $1x10^9$-$1x10^{10}$ particles or about $1x10^9$-$1x10^{12}$ particles), or even at least about $1x10^{10}$ particles (e.g., about $1x10^{10}$-$1x10^{12}$ particles) of the adenoviral vector. Alternatively, the dose comprises no more than about $1x10^{14}$ particles, preferably no more than about $1x10^{13}$ particles, more preferably no more than about $1x10^{12}$ particles, even more preferably no more than about $1x10^{11}$ particles, and most preferably no more than about $1x10^{10}$ particles (e.g., no more than about $1x10^9$ particles). In other words, the composition can comprise a single dose of adenoviral vector comprising, for example, about $1x10^6$ particle units (pu), $2x10^6$ pu, $4x10^6$ pu, $1x10^7$ pu, $2x10^7$ pu, $4x10^7$ pu, $1x10^8$ pu, $2x10^8$ pu, $4x10^8$ pu, $1x10^9$ pu, $2x10^9$ pu, $4x10^9$ pu, $1x10^{10}$ pu, $2x10^{10}$ pu, $4x10^{10}$ pu, $1x10^{11}$ pu, $2x10^{11}$ pu, $4x10^{11}$ pu, $1x10^{12}$ pu, $2x10^{12}$ pu, or $4x10^{12}$ pu of adenoviral vector.

[0083] Administration of the inventive nucleic acid sequence, composition, or polypeptide can be one component of a multistep regimen for inducing an immune response against HIV in a mammal. In particular, the inventive method can represent one arm of a prime and boost immunization regimen. The inventive method, therefore, can comprise administering to the mammal any suitable "priming" composition prior to administering the inventive nucleic acid sequence, composition, or polypeptide. Thus, in this embodiment, an immune response is "primed" by administration of the priming composition, and is "boosted" by administration of the inventive nucleic acid sequence, composition, or polypeptide. Alternatively, the inventive

method can comprise administering to the mammal any suitable "boosting" composition following administration of the inventive nucleic acid sequence, composition, or polypeptide. Thus, in this embodiment, an immune response is "primed" by administration of the inventive nucleic acid sequence, composition, or polypeptide, and is "boosted" by administration of the boosting composition. The priming composition and/or boosting composition (when not the inventive nucleic acid sequence, composition, or polypeptide) desirably comprises one or more nucleic acid sequences that encode at least one HIV polypeptide that is the same as the HIV polypeptide (e.g., an HIV-1 Env polypeptide) encoded by the inventive nucleic acid sequence. However, in some embodiments, the priming composition or booster composition comprises one or more nucleic acid sequences encoding an HIV polypeptide that is different from the polypeptide(s) encoded by the inventive nucleic acid sequence (such as those described herein).

[0084] The one or more nucleic acid sequences of the priming composition or the boosting composition can be administered as part of a vector or as naked DNA. Any vector, such as those described herein, can be employed in the priming or boosting composition, including viral and non-viral gene transfer vectors. Examples of suitable viral vectors include, but are not limited to, retroviral vectors, adeno-associated virus vectors, vaccinia virus vectors, herpesvirus vectors, and adenoviral vectors. Examples of suitable non-viral vectors include, but are not limited to, plasmids, liposomes, and molecular conjugates (e.g., transferrin). Ideally, the vector is a plasmid or an adenoviral vector. Alternatively, an immune response can be primed or boosted by administration of the antigen itself, e.g., an antigenic protein, inactivated pathogen, and the like.

[0085] The priming composition is administered to the mammal to prime the immune response to HIV, while the boosting composition is administered to enhance or augment the immune response induced by the priming composition. More than one dose of the priming composition or boosting composition can be provided in any suitable timeframe. Administration of the priming composition and administration of the boosting composition desirably is separated by at least about 1 week (e.g., at least about 1 week, 2 weeks, 4 weeks, 8 weeks, 12 weeks, 16 weeks, or more). Preferably, the primer composition is administered to the mammal at least three months (e.g., three, six, nine, twelve, or more months) before administration of the boosting composition. Most preferably, the primer composition is administered to the mammal

at least about six months to about nine months before administration of the boosting composition.

[0086] The following examples further illustrate the invention but, of course, should not be construed as in any way limiting its scope.

EXAMPLE 1

[0087] This example demonstrates a method of producing a mosaic HIV Env polypeptide comprising an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide.

[0088] Plasmid DNAs were generated encoding: (1) the full length Env protein, gp160, (2) the full length Env protein with variable loops 1, 2, 4, and 5 deleted, Env gp160ΔVs, (3) the full length Env protein with deletion of the fusion domain, the cleavage domains, and shortened interspace between heptad 1 (H1) and heptad 2 (H2) (Chakrabarti et al., *J. Virol.*, *76*: 5357-5368 (2002)), Env gp160ΔCFI, (4) the full length Env protein with deletion of the fusion domain, the cleavage domains, a shortened interspace between heptad 1 (H1) and heptad 2 (H2), and also with variable loops 1, 2, 4, and 5 deleted, Env gp160ΔCFIΔVs, and (5) the Env protein lacking the cytoplasmic domain, the fusion domain, the cleavage domains, variable loops 1, 2, 4, and 5, and having a shortened interspace between heptad 1 (H1) and heptad 2 (H2), Env gp145ΔCFIΔVs (F1, F2, F3, F4). All of the modified HIV Env genes were synthesized using human-preferred codons (GeneArt, Regensburg, Germany) (Kong et al., *Proc. Natl. Acad. Sci. USA*, *103*: 15987-15991 (2006)) or by preparation of oligonucleotides of 75 base pairs (bp) overlapping by 25 bp, or of 60 bp overlapping by 20 bp, and assembled by Pwo (Boehringer Mannheim, Mannheim, Germany) and Turbo Pfu (Stratagene, La Jolla, CA) as described in Chakrabarti et al., *J. Virol.*, *76*: 5357-5368 (2002), and Kong et al., *J. Virol.*, *77*: 12764-12772 (2003). All deletions or other modifications were generated by site-directed mutagenesis using a QuickChange kit (Stratagene, La Jolla, CA). The cDNAs were cloned into a plasmid expression vector, pCMV/R, which mediates high level expression and immunogenicity *in vivo* (Barouch et al., *J. Virol.*, *79*: 8828-8834 (2005), and Yang et al., *Science*, *317*: 825-828 (2007)).

[0089] Mosaic proteins were designed using the methods described in Fischer et al., *Nat. Med.*, *13*: 100-106 (2007), which is a web-based suite of tools that enables generation of

candidate mosaic sequences for any set of variable pathogen proteins, and epitope length sequence coverage comparison of different vaccine antigen candidates (Thurmond et al., *Bioinformatics*, *24*: 1639-1640 (2008)). Mosaics were optimized as a set for a particular size of cocktail, and are designed separately for 1, 2, and 3 antigen combinations (i.e., the single mosaic is not found in the 2 or 3 mosaic set). The input data was an unaligned version of the full Env group M alignment from the Los Alamos National Laboratory HIV database, as of July 2006 (restricted to include a single sequence per person). Sequences were generated as recombinants of that set, and optimized for 9-mer coverage of that set. Unnatural breakpoints were excluded. The three natural sequences that in combination provided the optimal 9-mer coverage of that same data set, either with or after exclusion of the V-loops, were selected using the same software suite (Gao et al., *J. Virol.*, *79*: 1154-1163 (2005), and Thurmond et al., *supra*). The length of 9 amino acids was selected for the optimization criteria because it is the most common length of optimal CD8 epitopes. Nearby lengths (e.g., 10, 11, 12, etc.) also receive greatly enhanced coverage through the process of optimizing on 9-mers. The full length Env protein amino acid sequences of the three sets of mosaics, gp160, and other Env mutants, are represented by SEQ ID NOs 1-44.

[0090] Using the methods described above, mosaic Env proteins were designed, and *in silico* recombinants of natural Env proteins with breakpoints that do not disrupt the protein, and which optimize potential epitope coverage of a diverse population were assembled. All stretches of 9-mers) were considered potential epitopes, and the presence of rare 9-mers was minimized. The M group sequence alignment from the 2006 HIV database (www.hiv.lanl.gov) was used as a baseline. M designates the "main" group of HIV-1 sequences that includes all of the standard clades (A-K) and their recombinants, which is the over-arching set of diverse HIV sequences responsible for the global HIV pandemic.

[0091] Separate optimizations for 1, 2, or 3 mosaic gp160 DNA vaccine antigen combinations were performed. Each set was comprised of plasmid DNAs encoding distinct proteins, which when combined yield optimal coverage for a given number of antigens. These designs were subsequently modified to parallel Env modifications that have been previously explored as vaccine antigens, including ΔCFI deletions, where ΔCFI refers to deletions of the cleavage site, fusogenic domain, and spacing of heptad repeats 1 and 2 (Chakrabarti et al., *J.*

Virol., *76*: 5357-5368 (2002)), and gp145, to compare the impact of Env modifications relative to intact gp160 on T-cell responses to the various antigens (Fig. 1A).

[0092] In addition, the three natural strains were selected that, in combination, optimize coverage of the M group, either including or excluding the hypervariable regions V1, V2, V4, and V5 (ΔVs). The ΔVs were optimized independently, and are distinct sets of proteins. The basic antigen design strategies included the following sets: (1) three natural strains that have been previously studied as a polyvalent vaccine in the modified Env gp145ΔCFI, each from different clades (Env ABC), (2) one, two, or three gp160 mosaics (mos 1, mos 2, mos 3), and (3) three natural strains selected to provide, in combination, optimal M group coverage of gp160 9-mers (nat.3) (i.e., CRF01AE, FIN92168, AF219267; clade B, QH0908 AF277072; and clade C, 93IN101 AB023804 (listed as clade, sequence name, and NCBI accession number)), or to provide optimal M group coverage if the V regions were excluded (natΔV.3) (i.e., clade C, 99BW46424 AF443084; clade B, QH0908 AF277072; clade A, KNH1088 AF457063 (listed as clade, sequence name, and NCBI accession number)). These baseline sets were further modified to enable direct comparisons of T-cell responses of the full gp160 proteins to previously studied envelope modifications. Thus, gp160 responses for a given antigen set were compared to gp145ΔCFI and gp160ΔCFI modifications. A negative control (negc) consisting only of the CMV/R vector was included.

[0093] Several exploratory immunizations using mosaic constructs that had the hypervariable loops removed were performed; however, no particular benefit was conferred. The rationale for the deletion of the hypervariable regions is that the hypervariable regions are often unique, and thus would be strain-specific, such that responses to these regions might divert the vaccine-induced immune response away from more conserved and potentially cross-reactive regions of Env. The code for designing mosaics and selecting optimal natural sequences is available (see, e.g., www.hiv.lanl.gov/content/sequence/MOSAIC/, and Thurmond et al., *supra*). A comparison of the basic gp160 vaccine designs to the M group sequences from the 2006 HIV-1 database Env alignment (www.hiv.lanl.gov) is shown in Fig. 1A, and the impact on numbers of potential epitopes lost by deleting parts of the protein in gp145ΔCFI constructs is shown in Fig. 1B. For comparison, a polyvalent vaccine that included one A, one B, and one C clade Env gene (Env ABC) was included. These strains were not optimized for 9-mer coverage,

and in combination have previously been shown to elicit immune responses (Barouch et al., *J. Virol.*, *79*: 8828-8834 (2005); Catanzaro et al., *Vaccine*, *25*: 4085-4092 (2007); Chakrabarti et al., *supra*; Fischer et al., *supra*; Kong et al., *J. Virol.*, *77*: 12764-12772 (2003); Seaman et al., *J. Virol.*, *79*: 2956-2963 (2005)). When the subset of sequences from clades A (82 sequences), B (454), or C (464) were each compared separately to the M group-designed antigens, each clade behaved roughly comparably (within a few percent) to the full dataset results.

[0094] This example demonstrates a method of producing a nucleic acid sequence encoding a HIV-1 Env polypeptide comprising an insertion of at least one T-cell epitope that is not naturally present the Env polypeptide in accordance with the invention.

EXAMPLE 2

[0095] This example demonstrates a method of inducing an immune response against HIV-1 in a mammal using nucleic acid sequences encoding mosaic Env polypeptides.

Vaccination

[0096] Six- to eight-week-old B6D2F1/J (H2 Haplotype b/d) female mice (Jackson Laboratory, Bar Harbor, ME) were used for immunogenicity studies. Mice were immunized with plasmid DNA vaccines encoding gp160 mosaics encoded by 1, 2, or 3 plasmids, one natural strain, or a natural strain with deleted V regions, and compared to a mixture of clade A, B, and C gp145 with deletions in the cleavage, furin, and interhelical domains (ΔCFI) described previously (Chakrabarti et al., *supra*, and Kong et al., *J. Virol.*, *77*: 12764-12772 (2003)). Specifically, the groups included the plasmid pCMV/R with no insert (control), 15 μg of the gp160 mosaic "mos 1", 7.5 μg of each plasmid in the two-plasmid groups (i.e., gp160 mosaics "mos 2.1" and "mos 2.2"), or 5 μg of each plasmid in the three-plasmid groups (i.e., gp160 mosaics "mos 3.1," "mos 3.2," and "mos 3.3"), three gp160 natural strains, gp160ΔVs natural strains, and gp145ΔCFI clades A, B, and C. Mice (n=10 per group) were immunized with a total of 15 μg of DNA four times at two-week intervals. Immunizations were administered bilaterally into the muscle of the hind leg using needle and syringe. Two weeks after the last immunization, the splenocytes from different groups of three immunized mice were isolated, pooled, and stimulated by different stimulants.

Intracellular Cytokine Staining (ICS)

[0097] Splenocytes from immunized mice were analyzed by intracellular cytokine staining (ICS) for TNF-α and IFN-γ T-cell responses against approximately 100 different peptide pools. 492 Env peptides were used for ICS stimulation. For this ICS analysis, 15 mer vaccine-important peptides (VIPs) (Li et al., *Vaccine*, *24*: 6893-6904 (2006)) were based on 549 full-length HIV-1 genome sequences obtained from the Los Alamos National Laboratory (LANL) HIV sequence database as of February 2005 to evaluate the vaccines as the common standardized panel of HIV-1 peptides for T-cell-based vaccines. The 492 Env peptide sequence was designed to permit expression of the potential T-cell epitopes (PTE) found most frequently in the sequences of circulating worldwide HIV-1 strains. All synthesized peptides (New England Peptide, Gardner, MA) were 15 amino acids in length with naturally occurring 9-amino acid sequences that are T-cell determinants captured in an unbiased manner (Li et al., *supra*, and Malhotra et al., *J. Virol.*, *81*: 5225-5237 (2007)). The 9-amino acid sequences were defined first by frequency of their individual subtypes, and then the panel was assembled jointly using a forward stepwise algorithm to cover all of the subtype-specific PTEs in the smallest set possible. This algorithm selects first for the highly conserved PTEs, then the less conserved. 492 PTE peptides were generated (for a coverage threshold of 15% (Li et al., *supra*)), and grouped into 78 pools of 6-12 PTE 15-mer peptides such that the peptides that carried the highest frequency 9-mers were grouped in the first pool, continuing so that the peptides with the rarest 9-mers were in the 78th pool. All but the pools representing the rarest potential epitopes contained six peptides each; the three pools of rarest potential epitopes contained 10-12 peptides. These sets are referred to as "PTE pools." Four pools with larger numbers of peptides were also tested, with 114 for the first three large pools, and 148 for the fourth large pool. These sets are referred to as "PTE superpools." Pooled sets of peptides, 15-mers overlapping by 11 amino acids, corresponding to each of the three Envelopes included in the Env ABC polyvalent vaccine, were also used as previously described (Barouch et al., *supra*; Catanzaro et al., *supra*; Chakrabarti et al., *supra*; Fischer et al., *supra*; Kong et al., *supra*; Seaman et al., *supra*).

Cellular Immune Analysis

[0098] Single-cell suspensions of mouse splenocytes were produced, washed, and resuspended to a final concentration of approximately 10^7 cells/ml. All groups of harvested

spleen cells (maximum of 10^6 cells/peptide pool) were stimulated for five hours in the presence of 2 μg of anti-CD28 and anti-CD49d monoclonal antibodies/mL (BD PharMingen, San Diego, CA), and also with 10 μg/mL brefeldin A (Sigma, St. Louis, MO). Cells were stimulated for five hours with (a) 15-mers of 6-peptide pools or 12-peptide pools of PTE, as the target testing stimulating agents, (b) no stimulation for background control, (c) Ebola GP protein as the negative control, and (d) phorbol myristate acetate (PMA) with ionomycin as the positive control. Cells were then washed and stained with Vivid dye (Invitrogen, Carlsbad, CA) to determine their viability. FC block monoclonal antibodies were added to the cells followed by staining with surface antigens (rat anti-mouse cell surface antigens CD3-PerCP-Cy5.5, CD4-AlexaFluor700 and CD8-APC-Cy7 (BD PharMingen, San Diego, CA)). The cells were washed again, permeabilized, fixed with Cytofix/Cytoperm, and stained with monoclonal antibodies (rat anti-mouse cell surface antigens CD3-PerCP-Cy5.5, CD4-AlexaFluor700 and CD8-APC-Cy7, and rat anti-mouse cytokines IFN-γ-APC, IL-2-PE, and TNF-α-PE-Cy7 (BD PharMingen, San Diego, CA)) followed by multi-parametric flow cytometry analysis to detect the IFN-γ, IL-2, or TNF-α positive cells in the CD4+ or CD8+ T-cell population.

Flow Cytometry

[0099] Stained cells were assayed on the BD™ LSR-II Flow Cytometer using FACSDiva software (BD Biosciences, San Jose, CA). The data were analyzed with FlowJo 8.6.1 software (Tree Star, Ashland, OR). Another three mice in each group were subjected to the same analysis two days after the initial test to repeat the analysis.

Statistical Analysis

[00100] For replicate experiments, the overall strength of the response to a specific vaccine varied by up to a factor of six on different dates. This effect was corrected for in order to compare antigenic design strategies and Env modifications (different test vaccines) that were each tested multiple times, on different dates, and also because a fixed threshold could not be used to assess positivity.

[00101] The rank-orderings of responses to small peptide pools for the same test vaccine on different days were very similar, suggesting that the variation was a consequence of inter-assay experimental variation that resulted in a random scaling of the data. Accordingly, the following

"two-way layout" was adopted to model the average strength of the response to vaccine *i* on day *j*:

$$l_{ij} = v_i + d_j + \varepsilon_{ij}, \qquad (1)$$

where l_{ij} is the (natural) logarithm of the strength of the responses for the vac/mod i on day j, v_i is the vaccine effect, i.e., the contribution to the observed response due to the vaccine, d_j is the date effect, and ε_{ij} are identical and independently distributed Gaussian random errors, to account for variation due to mouse-to-mouse natural variation and stochastic effects. The date effects also depended on the T-cell type and cytokine, so separate models were fitted to all combinations of CD4 and CD8, and of IFN and TNF. However, for the same T-cell type, cytokine, and date, different parts of the data gave very similar estimates for d_j. The date effect likely represents variations in solutions and reagents, etc., used for different cytokines and T-cell types on different days. It could also be called a "batch" effect.

[00102] The model disentangles the effects of the date and of the vaccine, which are of comparable magnitude in this experiment. The model thus facilitates the comparison of the average response strength of different vaccines, even when they were measured on different days. A specific date *j'* was chosen arbitrarily as the fiducial date, and the data for all other dates *j* were rescaled as follows:

$$\tilde{l}_{ij} = l_{ij} + \bar{d}_{j'} - \bar{d}_j, \qquad (2)$$

where the $\bar{d}_j$ are unbiased estimates of the d_j. Then

$$E(\tilde{l}_{ij} - \tilde{l}_{i'j}) = \tilde{v}_i - \tilde{v}_{i'}, \qquad (3)$$

where E is the expectation. Thus, expected differences in the rescaled strength data depend only on the vaccine. Generally, vaccine *i* is stronger than vaccine *i'* by a factor $\exp(\tilde{v}_i - \tilde{v}_{i'})$. The rescaled data is referred to as the "strength-corrected" data. In general, the model only provides estimates of differences. Thus, the date effects are always measured relative to a specific fiducial date, and the vaccine effects are measured relative to the negative control.

[00103] The log strength, l_{ij}, was computed in three ways: (i) averaging the logarithm of the largest ten responses from the small peptide pools for the given experiment (the average was restricted to the top ten responses in order to reduce the effect of noise, which dominates the smaller responses), (ii) averaging the logarithm of responses from the four large peptide pools,

and (iii) averaging the logarithm of the responses for the EnvA, EnvB, and EnvC pools. In all situations, l_{ij} is the logarithm of the raw measurement, including the noise, which is assumed to scale in the same way as the signal. Singularities that would occur if the measurement went to zero were therefore avoided. The background from the unstimulated counts was estimated, and replaced any smaller measurement by this estimate. The uncertainty in l increases as the measurements approach zero.

[00104] In order to assess breadth, the number of positive responses to the 78 peptide pools were counted. The data were corrected for the date effect, as described above, using strength estimates based on the small peptide pools. A pool was judged to produce a positive response if the strength-corrected response exceeded a threshold, which was set separately for each combination of T-cell type and cytokine, but was otherwise held constant. When combining data from identical experiments on different days, the combined data was deemed to produce a positive response if the median response exceeded the threshold. The threshold was chosen by examining the small-peptide pool responses. There were two patterns of response: (i) pools that were clearly positive, in that the pools consistently showed elevated responses, and (ii) pools that showed either no responses above background, or sporadic responses that might have resulted from random noise or rare actual responses. The threshold was chosen to discriminate between these two patterns. A separate threshold was determined for each T-cell type/cytokine combination.

[0100] Background counts were measured on each microtiter plate, as well as responses to the negative control. However, these measurements were too noisy to be useful, so the threshold to assess positive response directly was used, without first subtracting estimated background.

[0101] To assess functionality, a matrix was computed whose rows denote particular experiments and whose columns denote small peptide pools. For each element of the matrix, the number 0, 1, or 2 was assigned, depending on the number of positive responses for TNF-α and IFN-γ observed for the corresponding experiment and peptide pool. Some experiments were also performed testing IL-2 responses, but the results were weak and sporadic and thus excluded from further analyses. Between 7 and 10 vaccine antigen/protein modifications (vac/mod) configurations were tested on 12 separate days. The 12 sets of experiments were grouped into six pairs. In each pair, the same set of antigen configurations was tested. The magnitude of the

overall responses varied by a factor of up to about six on different days, and this effect was corrected through statistical methods, as described below. Not all configurations of vaccine plus modification were tested, and the number of times a particular configuration was repeat tested ranged from 2 to 12.

[0102] A standard agglomerative clustering algorithm was employed (Ward, *J. Am. Stat. Assoc.*, *58*: 236-244 (1963)) using Euclidean distances, to cluster the experiments (row vectors) and the peptide pools (column vectors) (www.hiv.lanl.gov/content/sequence/HEATMAP/heatmap.html, based on the R package heatmap.2). These cluster patterns are shown on the margins of the heatmaps (Figs. 6A and 6B), which were generated by shading responses to indicate those patterns that generated no response (light gray), one response to either TNF-α or IFN-γ (dark gray) or responses to both (black). Statistical support for the various clusters is indicated on the dendrogram branch points, based on the approximately unbiased test of multisetp-multiscale bootstraps (Shimodaira, *Ann. Statist.*, *32*: 2616-2641 (2004)).

[0103] The data in these experiments came from a total of 384 microtiter plates, each of which measured IFN or TNF responses to $CD4^+$ or $CD8^+$ T-cells for a particular vaccine modality (vac/mod) on a particular day. By "vac/mod" is meant the DNA vaccine antigen cocktails (including one, two, or three mosaics; three natural strains selected to provide in combination optimal 9-mer coverage; three natural strains, one each from clades A, B, and C), and the Env modifications (including gp160, gp145ΔCFI, gp160ΔCFI, and gp145ΔCFIΔV, where ΔV refers to removal of the hypervariable loops, and ΔCFI refers to deletions of the cleavage site, fusogenic domain, and spacing of heptad repeats 1 and 2) (Chakrabarti et al., *supra*). In some cases, all or part of the data from a given plate was clearly affected by systematic error, as indicated by trends or consistently elevated responses from pools in contiguous regions of the plate. Such plates, of which 17 involved CD8 and 2 involved CD4, were left out of the analysis. Thus, a total of 365 plates were used. Among these plates, there was also a very small fraction of small peptide pools (98/28,470, 0.3%) for which data was unavailable.

Results

[0104] The CD4 and CD8 cell responses in immunized mice were measured using ICS for IFN-γ and TNF-α (Fig. 2, left and right, solid and open bars respectively). The minimal threshold response indicated by horizontal dashed lines was defined as 2 times the negative control. The CD8 response of the one-plasmid (mos 1) gp160 mosaics had the lowest response comparing to the mos 2 and mos 3 mosaics. Some peptide pools were shared in common among the three different mosaic combinations, while other peptide pools were unique. The responses elicited by the natural strain plasmids, with or without V region deletions, or gp145ΔCFI trivalent plasmids, were decreased relative to the PTE peptide pools. Similar results were seen in CD4 responses (Fig. 2, left panel), though the wild type natural strain plasmid induced a larger number of CD4 epitope responses. The additional animals in each group were sacrificed and analyzed in the same way, and statistical analysis was applied to the responses measured from the entire group (as described herein). The gp160 mosaic plasmids elicited detectable T-cell responses against VIPs PTE peptide pools after immunization.

[0105] The immunogenicity of alternative mosaic plasmids, with or without the variable region and/or ΔCFI ,were compared using the methods described herein. These vectors were in turn compared to the multiclade Env gp145ΔCFI vectors described herein which improve the breadth of the Env T-cell response compared to single strain Env immunogens (Kong et al., *J. Virol.*, *77*: 12764-12772 (2003); Seaman et al., *supra*). The resultant immune responses for the complete groups of ten animals were then analyzed to determine their comparative magnitude and breadth of response. Comparing the number of responses to PTE peptides, the responses in these groups immunized with deleted mosaics was lower than with the wild type gp160 mosaic, particularly for CD8 responses.

[0106] The strength and breadth of the immune responses induced by the mosaic constructs was evaluated by analyzing the IFN-γ and TNFα ICS responses of CD4 and CD8 cells to different peptide pools as described above. The Env ABC vaccine elicited strong responses to the Env ABC peptide pools. Thus, the overall vaccine response tested using autologous peptides suggested a robust response to the matched vaccine. The response of CD4 cells to the vaccine was similar in both the small size and superpools, and it was highly significant for both IFN-γ and TNF-α responses ($p < 10^{-4}$). For CD8 responses to the Env ABC immunogens, the IFN-γ and TNF-α ICS responses were highest against B clade pooled peptides. Among the individual

A, B, and C pools, the responsiveness of CD8 T-cells was reflected in a rank order, B > A ~ C (paired Wilcoxon B > A p = 0.0005, B > C, p = 0.003). CD4 T-cell responses were more consistent across all three clade pools, with B ~ A > C, paired Wilcoxon B > C p = 0.009, A > C p = 0.052.

[0107] Vaccine strength was also assessed by measuring T-cell responses to subgroups of the PTE peptide pools, each containing 10 peptides. For CD8 cells, the strength of the response of the mosaics was much greater than that of the naturals or the positive control (Fig. 4A). Among the mosaics, the mos 2 and mos 3 mosaic set responses were stronger than the single mosaic. The overall strength of the response was much greater than the response to the Env A, B, and C pools. The gp145ΔCFI immunogens induced a consistently weak response, which may have resulted from a loss of epitopes due to the deletions in the open reading frame (see Fig. 1B), which then caused mismatch with a number of peptides made on the basis of the wild-type sequence. The response profile for CD4 is quite different from that of CD8 (Fig. 4B). Here, the strength of the mosaics and the naturals was roughly comparable, and there was less variation in the response of the different vaccines than in CD8 cells. The data suggests that the ΔCFI modification may reduce the response, possibly because of less matching to the ICS assay peptides, but the 145ΔCFI modification, which suppressed the CD8 response, does not do the same for the CD4 response. Finally, vaccine-induced immunity was determined by ICS with PTE superpools. Vaccine strength was assessed by measuring T-cell responses to the four peptide pools including large numbers of peptides (Fig. 5). The strength estimates for the large peptide pools were very similar to those obtained with the small peptide pools, despite the fact that they were obtained from separate experiments.

[0108] Vaccine breadth was also assessed using strength-corrected responses to the 78 small peptide PTE pools. Data for identical vaccines or vaccine variants on particular dates were pooled, and responses were deemed positive if the median response exceeded a fixed threshold.

[0109] The ICS IFN-γ responses in CD8 cells were compared among six different vaccines. The breadth of the mosaic responses was dramatically greater than that of the naturals, including both Env ABC and the three naturals selected to optimize epitope coverage (Fig. 4A). The mosaics all showed clear spikes for numerous peptide pools. The number of such pools increased from 4 to 10 in the monovalent versus the trivalent mosaics. The best natural set of

Envs showed only two weak spikes, at pools 7 and 13. The results for the TNF-α ICS responses in CD8 cells were very similar, both in magnitude of the response and in the specific peptide pools that test positive (Fig. 4B).

[0110] Analogous plots for CD4 IFN-γ and TNF-α ICS revealed similar patterns of responsiveness (Figs. 5A, B). The mosaics generated more positive responses than the naturals, although the difference was less striking. For example, the mos 2 and mos 3 mosaic sets generated a vigorous IFN-γ response to 15 and 14 pools, respectively, but three natural proteins generate an adequate IFN-γ response (10 pools each for Env ABC, NaturalΔV, and NaturalN1,N2,N3).

[0111] Although the strength correction for inter-assay variation improved the experimental consistency between experiments, the identification of positive responses was not particularly sensitive to uncertainty in the strength correction. Indeed, the greater breadth of CD8 response in the mosaics was clear, even if no strength correction was used. With regard to the threshold, changes in the threshold affected the number of positive responses, and these changes can affect different vaccines differently, depending on how many positives are near the threshold. The CD8 breadth comparisons were very robust to changes in the threshold; the CD4 comparisons were not as robust. Inasmuch as the breadth estimate depends on a choice of threshold, and there is uncertainty as to the correct value, the evidence for increased CD4 breadth in the mosaics should be viewed as suggestive but not conclusive.

[0112] To evaluate the qualitative nature of the T-cell response, an analysis as to whether these cells synthesized either IFN-γ and/or TNF-α in response to vaccination was performed, the latter being a surrogate for a multifunctional cytokine production. For this analysis, the responses were evaluated using a heatmap representation of functional CD8 responses to IFN-γ and TNF-α, with a threshold of 0.1 (Fig. 6A). No response, a response to either IFN-γ or TNF-α, or a response to both IFN-γ and TNF-α were evaluated. All gp160 mos 2 and mos 3 mosaics are contained in a single large cluster; i.e., they displayed consistent patterns of responses to particular peptide pools. The gp160 mos 1 responses occupied a single cluster with intermediate reactivity. A cluster for the natural gp160ΔCFI set was also evident. All of the mosaics elicited many bifunctional responses; however, the pattern was less apparent in the naturals. Analysis of the CD4 responses with this heatmap revealed that the mosaics, as well as the naturals, were

broadly bifunctional (Fig. 6B). The peptides pools were generated so that the most conserved 9-mers were found in the first pool, and as the number of the pool increased, the relative frequency of the 9-mers contained in the pool diminished. The pools that contained most conserved peptides tended not to elicit responses. Indeed, the number of bifunctional responses exceeded the number of monofunctional responses. The CD4 responses of the mosaics and the naturals were quite similar, not only in magnitude and breadth, but in the specific peptide pools generating a positive response.

[0113] This example demonstrates the immunogenicity of a nucleic acid sequence encoding a HIV-1 Env polypeptide comprising an insertion of at least one T-cell epitope that is not naturally present the Env polypeptide.

[0114] All references, including publications, patent applications, and patents, cited herein are hereby incorporated by reference to the same extent as if each reference were individually and specifically indicated to be incorporated by reference and were set forth in its entirety herein.

[0115] The use of the terms "a" and "an" and "the" and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms "comprising," "having," "including," and "containing" are to be construed as open-ended terms (i.e., meaning "including, but not limited to,") unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

Preferred embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those preferred embodiments may

become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

CLAIMS:

1. An isolated or purified nucleic acid sequence encoding a HIV-1 Env polypeptide, which HIV-1 Env polypeptide comprises an insertion of at least one T-cell epitope that is not naturally present in the Env polypeptide.

2. The nucleic acid sequence of claim 1, wherein the HIV-1 Env polypeptide comprises an insertion of a plurality of T-cell epitopes.

3. The nucleic acid sequence of claim 1 or claim 2, wherein the at least one T-cell epitope is found in at least 30% of circulating HIV-1 strains.

4. The nucleic acid sequence of any one of claims 1-3, wherein the at least one T-cell epitope is found in at least 70% of circulating HIV-1 strains.

5. The nucleic acid sequence of any one of claims 1-4, wherein the nucleic acid sequence comprises human-preferred codons.

6. The nucleic acid sequence of any one of claims 1-5, wherein the nucleic acid sequence encodes a full length Env polypeptide.

7. The nucleic acid sequence of any one of claims 1-5, wherein the nucleic acid sequence does not encode the variable loops 1, 2, 4, or 5 of Env, or combinations thereof.

8. The nucleic acid of any one of claims 1-5, wherein the Env polypeptide does not comprise the fusion domain, the cleavage domains, and the interspace between heptad 1 and heptad 2.

9. The nucleic acid sequence of any one of claims 1-5, wherein the Env polypeptide does not comprise (a) the fusion domain, (b) the cleavage domains, (c) the interspace between heptad 1 and heptad 2, and (d) one or more of variable loops 1, 2, 4, and 5.

10. The nucleic acid sequence of any one of claims 1-5, wherein the Env polypeptide does not comprise (a) the cytoplasmic domain, (b) the fusion domain, (c) the cleavage domains, (d) the interspace between heptad 1 and heptad 2, and (e) one or more of variable loops 1, 2, 4, and 5.

11. The nucleic acid sequence of any one of claims 1-5 and 7-10, wherein the Env polypeptide does not comprise at least one T-cell epitope that is present in a natural Env polypeptide.

12. The nucleic acid sequence of claim 11, wherein the Env polypeptide does not comprise a plurality of T-cell epitopes that are present in a natural Env polypeptide.

13. The nucleic acid sequence of claim 11 or claim 12, wherein the T-cell epitopes that are present in a natural Env polypeptide are present in no more than 20% of circulating HIV-1 strains.

14. The nucleic acid sequence of claim 13, wherein the T-cell epitopes that are present in a natural Env polypeptide are present in no more than 1% of circulating HIV-1 strains.

15. The nucleic acid sequence of any one of claims 11-14, wherein the T-cell epitopes that are present in a natural Env polypeptide are present in no more than three of the members of the M group of circulating HIV-1 strains.

16. The nucleic acid sequence of claim 1, wherein the nucleic acid sequence encodes a HIV-1 Env polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs: 1-6, SEQ ID NOs: 10-13, SEQ ID NOs: 17-22, SEQ ID NOs: 26-28, and SEQ ID NOs: 32-37.

17. A polypeptide encoded by a nucleic acid sequence of any one of claims 1-16.

18. The polypeptide of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 1-6, SEQ ID NOs: 10-13, SEQ ID NOs: 17-22, SEQ ID NOs: 26-28, and SEQ ID NOs: 32-37.

19. A vector comprising the nucleic acid sequence of any one of claims 1-16, wherein the vector is suitable for expressing the modified HIV-1 Env polypeptide.

20. The vector of claim 19, wherein the vector is a viral vector.

21. The vector of claim 20, wherein the viral vector is an adenovirus vector.

22. The vector of claim 21, wherein the adenovirus vector is an adenovirus 5 vector or an adenovirus 26 vector.

23. An isolated host cell comprising the vector of any one of claims 19-22, wherein the host cell is suitable for expressing the modified HIV-1 Env polypeptide.

24. A composition capable of eliciting an immune response against HIV-1 comprising (a) the nucleic acid sequence of any one of claims 1-16 or the vector of any of claims 19-22 and (b) a pharmaceutically acceptable carrier.

25. The composition of claim 22, wherein the composition comprises a plurality of nucleic acid sequences of any one of claims 1-16 or a plurality of vectors of any of claims 19-22.

26. A syringe comprising the composition of claim 24 or claim 25.

27. A needless delivery device comprising the composition of claim 24 or claim 25.

28. A method of inducing an immune response against HIV-1 in a mammal, which method comprises administering the nucleic acid sequence of any one of claims 1-16 to a mammal, whereupon an immune response against HIV-1 is induced in the mammal.

29. A method of inducing an immune response against HIV-1 in a mammal, which method comprises administering the polypeptide of claim 17 to a mammal, whereupon an immune response against HIV-1 is induced in the mammal.

30. A method of inducing an immune response against HIV-1 in a mammal, which method comprises administering the composition of claim 24 or claim 25 to a mammal, whereupon an immune response against HIV-1 is induced in the mammal.

31. A method of making a nucleic acid sequence encoding a modified HIV-1 Env polypeptide, which method comprises:

(a) identifying at least one T-cell epitope which is present in a natural Env polypeptide of a plurality of circulating HIV-1 strains, and

(b) introducing a nucleic acid sequence encoding the at least one T-cell epitope into a nucleic acid sequence encoding an HIV-1 Env polypeptide which does not naturally encode the T-cell epitope, whereupon a nucleic acid sequence encoding a modified HIV-1 Env polypeptide is produced.

32. The method of claim 31, wherein the at least one T-cell epitope is present in at least 30% of circulating HIV-1 strains.

33. The method of claim 32, wherein the at least one T-cell epitope is present in at least 70% of circulating HIV-1 strains.

FIG. 1A

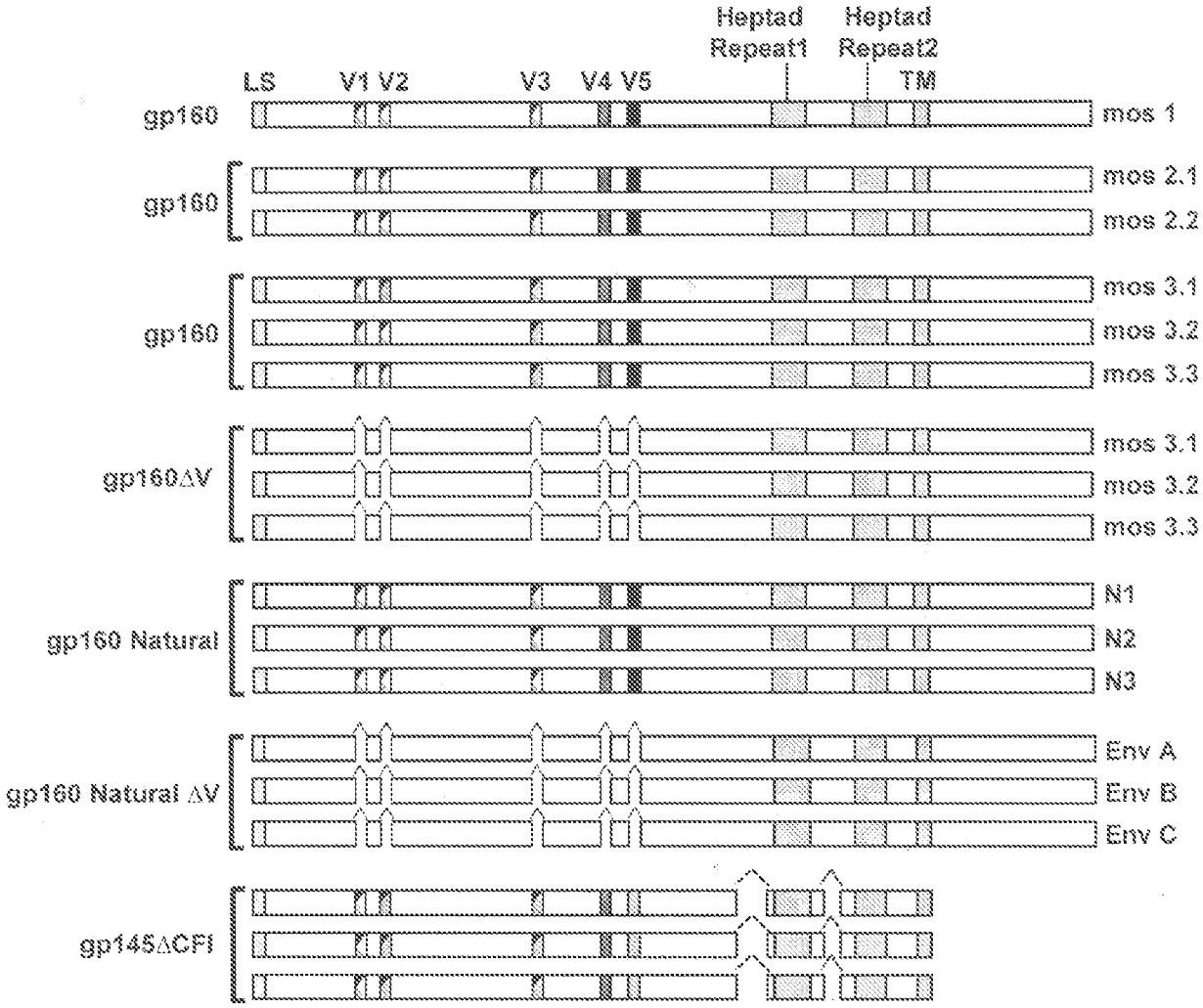

Heptad Repeat1
Heptad Repeat2
LS
V1 V2
V3
V4 V5
TM
gp160
mos 1
gp160
mos 2.1
mos 2.2
gp160
mos 3.1
mos 3.2
mos 3.3
gp160ΔV
mos 3.1
mos 3.2
mos 3.3
gp160 Natural
N1
N2
N3
gp160 Natural ΔV
Env A
Env B
Env C
gp145ΔCFI

FIG. 1B

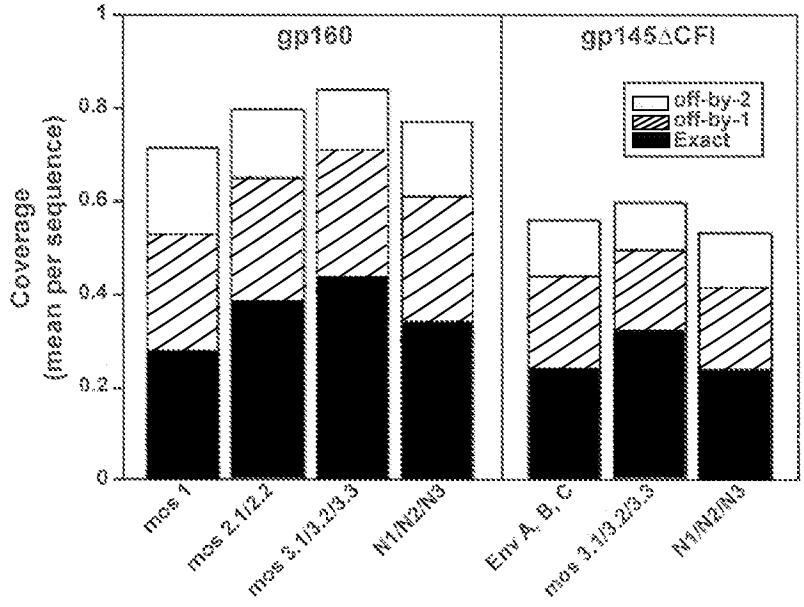

gp160
gp145ΔCFI
off-by-2
off-by-1
Exact
Coverage (mean per sequence)
1
0.8
0.6
0.4
0.2
0
mos 1
mos 2.1/2.2
mos 3.1/3.2/3.3
N1/N2/N3
Env A, B, C
mos 3.1/3.2/3.3
N1/N2/N3

FIG. 2

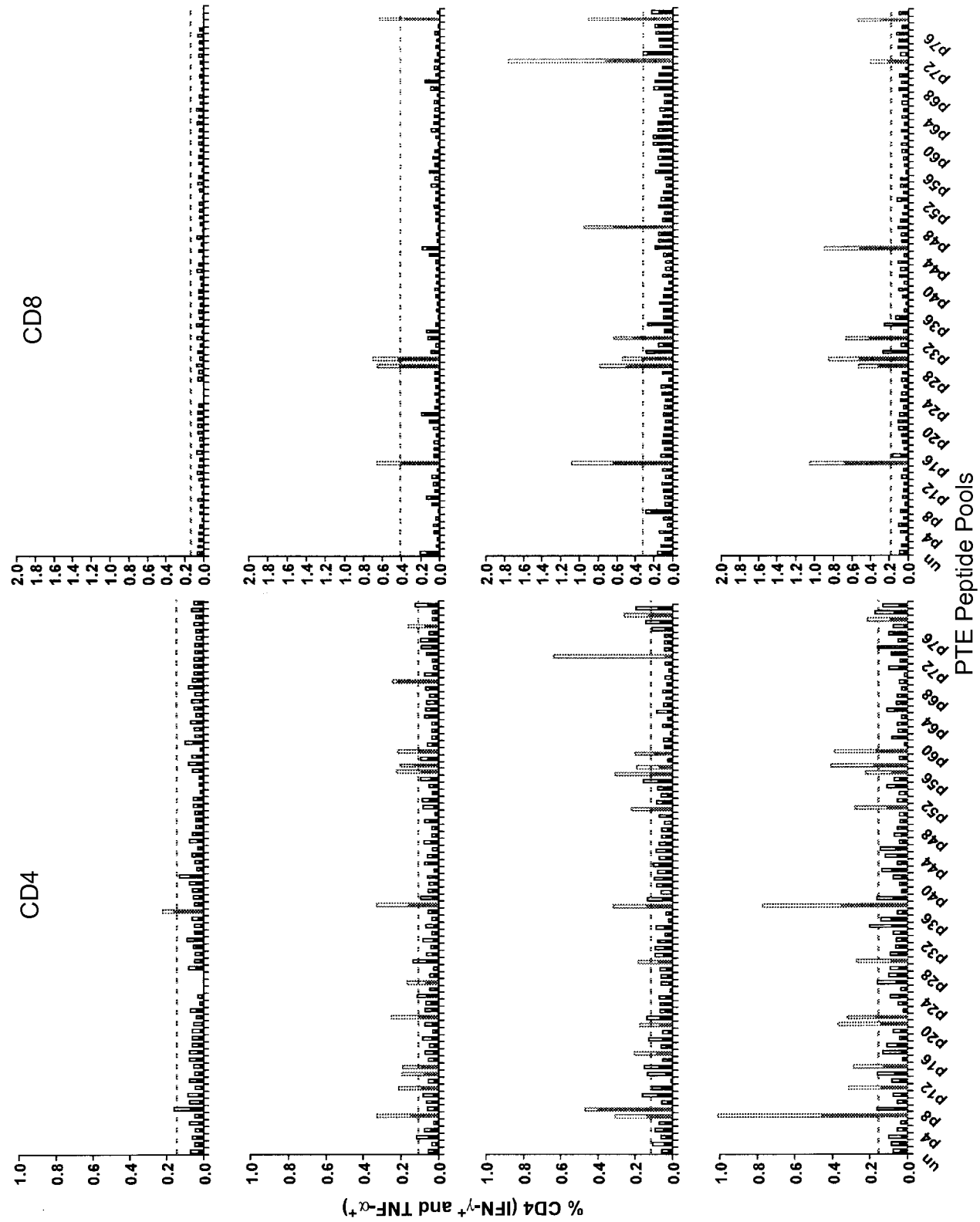

CD4
CD8
Control
gp160
mos 1
gp160
mos 2.1
mos 2.2
gp160
mos 3.1
mos 3.2
mos 3.3
% CD4 (IFN-γ+ and TNF-α+)
PTE Peptide Pools
un
p4
p8
p12
p16
p20
p24
p28
p32
p36
p40
p44
p48
p52
p56
p60
p64
p68
p72
p76

FIG. 2 Continued

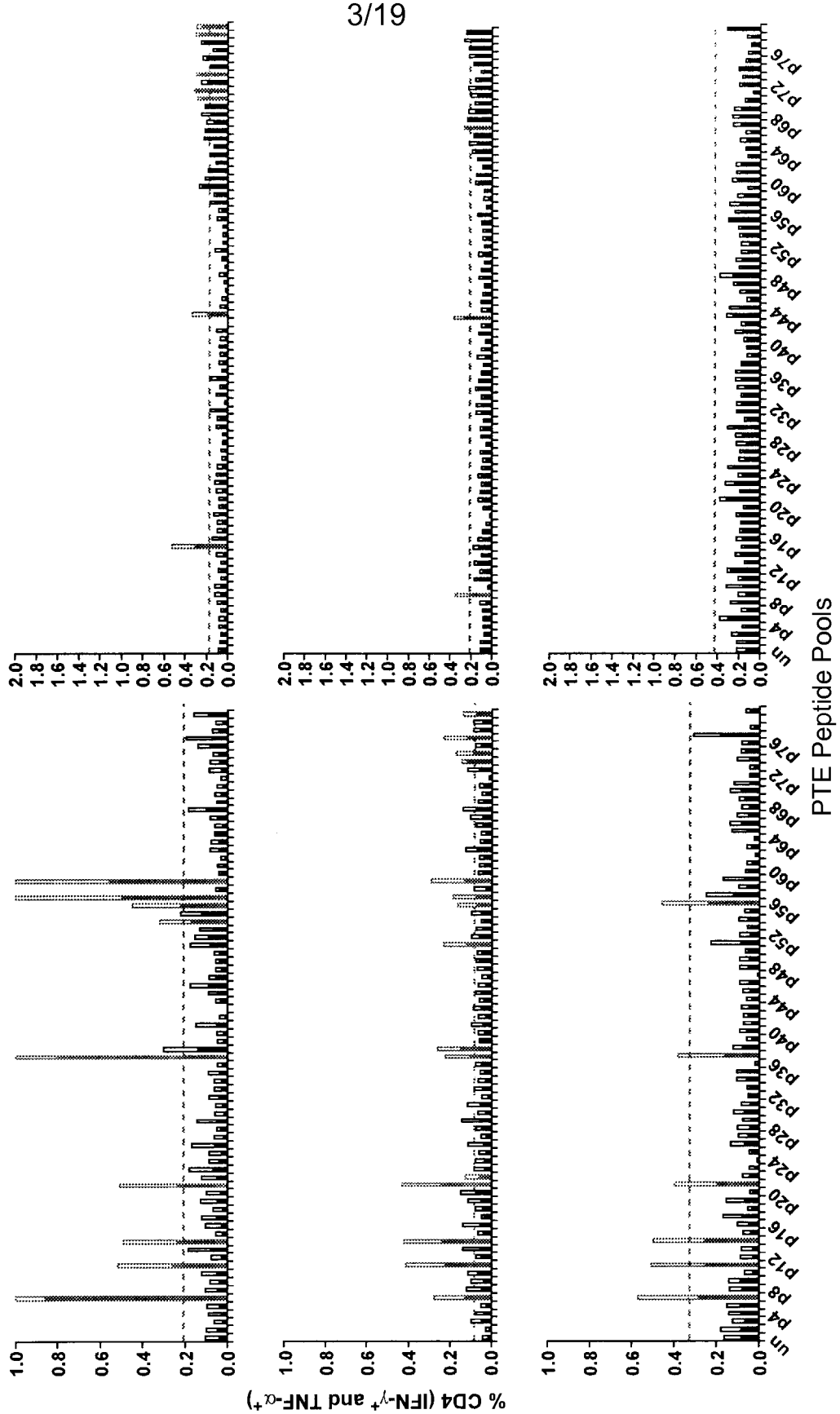


FIG. 3A

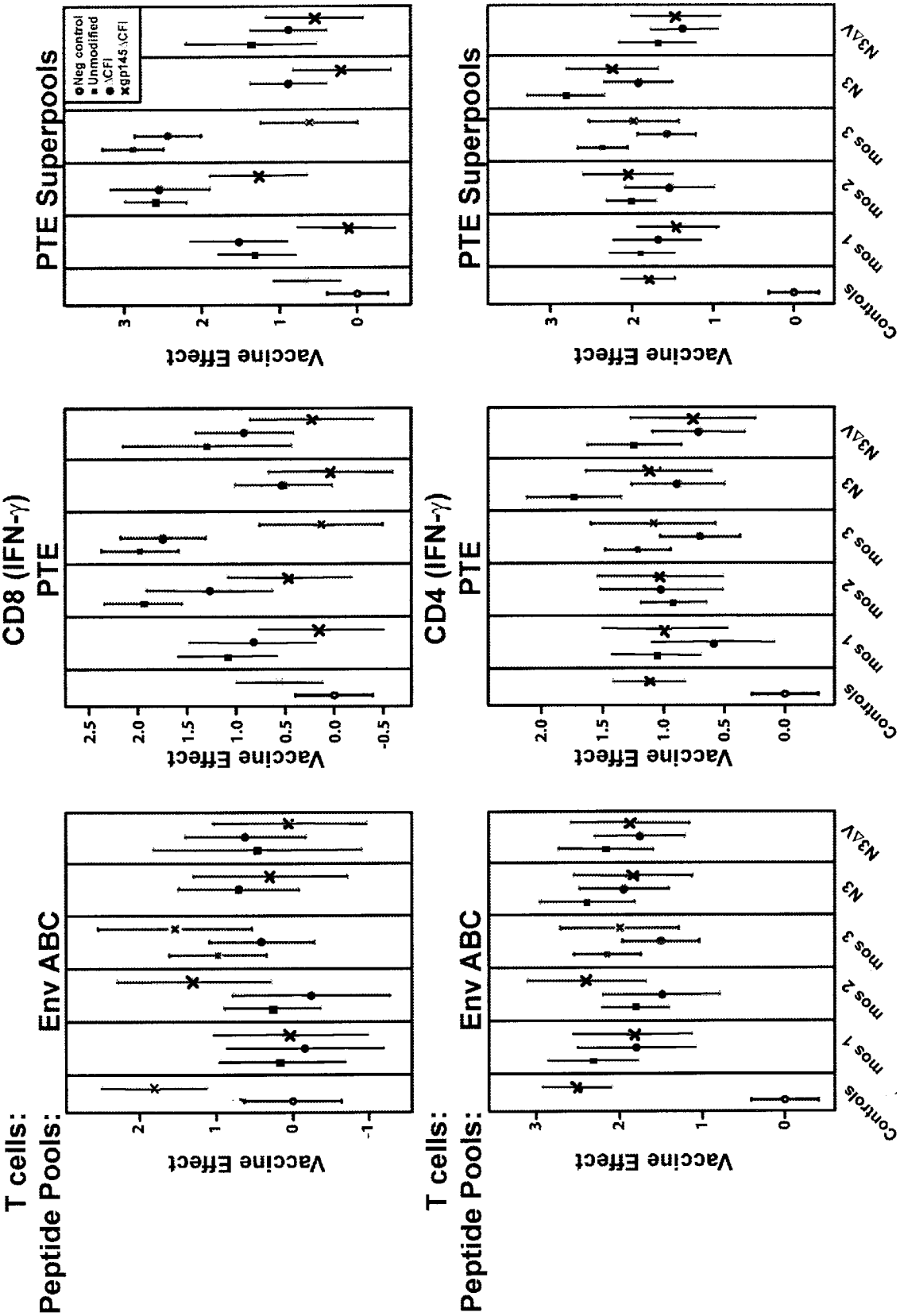

T cells:
Peptide Pools:
Env ABC
CD8 (IFN-γ)
PTE
CD4 (IFN-γ)
PTE
PTE Superpools
Neg control
Unmodified
ΔCFI
gp145ΔCFI
Vaccine Effect
Controls
mos 1
mos 2
mos 3
N3
N3ΔV

FIG. 3B

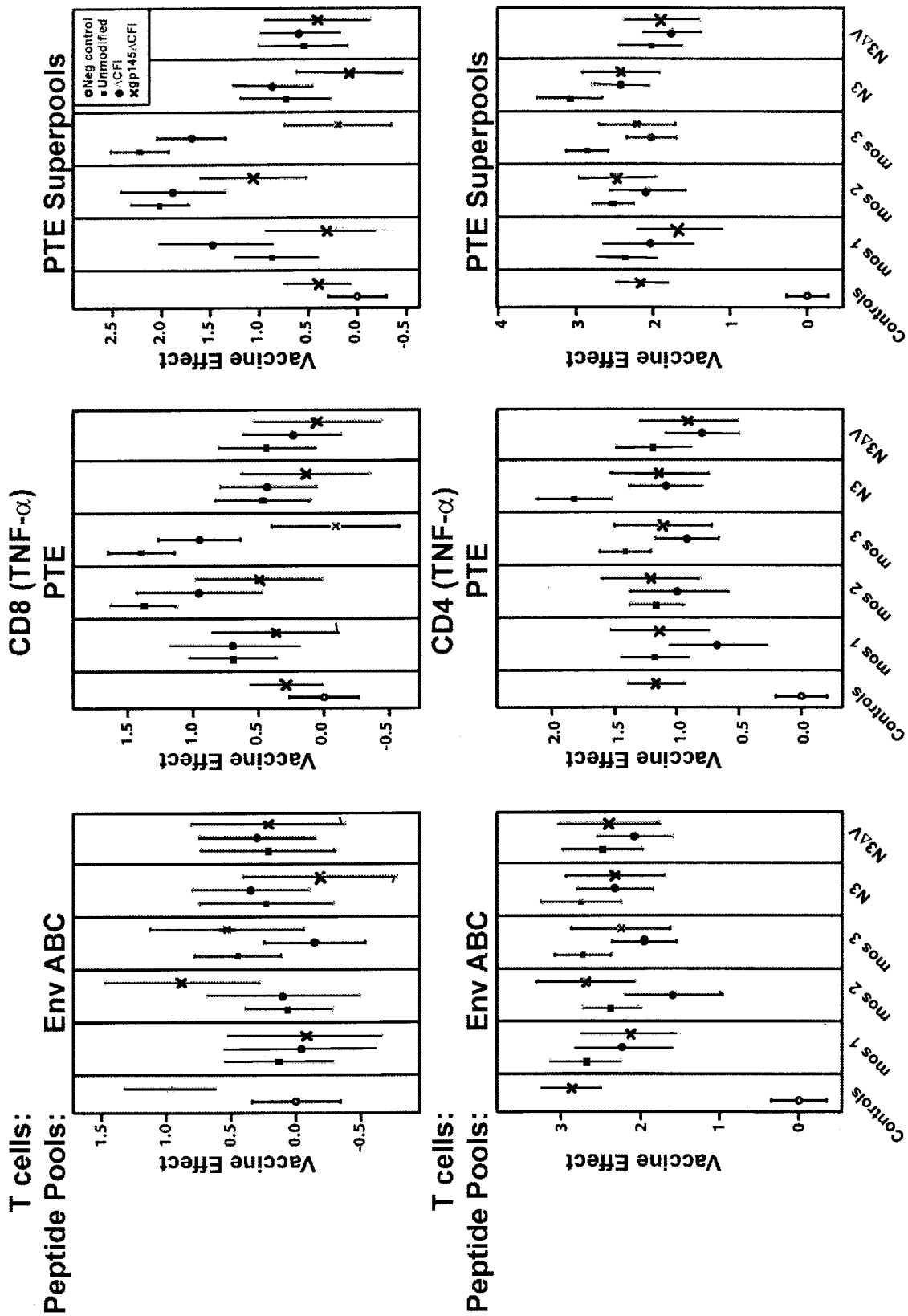

T cells:
Peptide Pools:
Env ABC
CD8 (TNF-α)
PTE
CD4 (TNF-α)
PTE
PTE Superpools
Neg control
Unmodified
ΔCFI
gp145ΔCFI
Vaccine Effect
Controls
mos 1
mos 2
mos 3
N3
N3ΔV

FIG. 4A

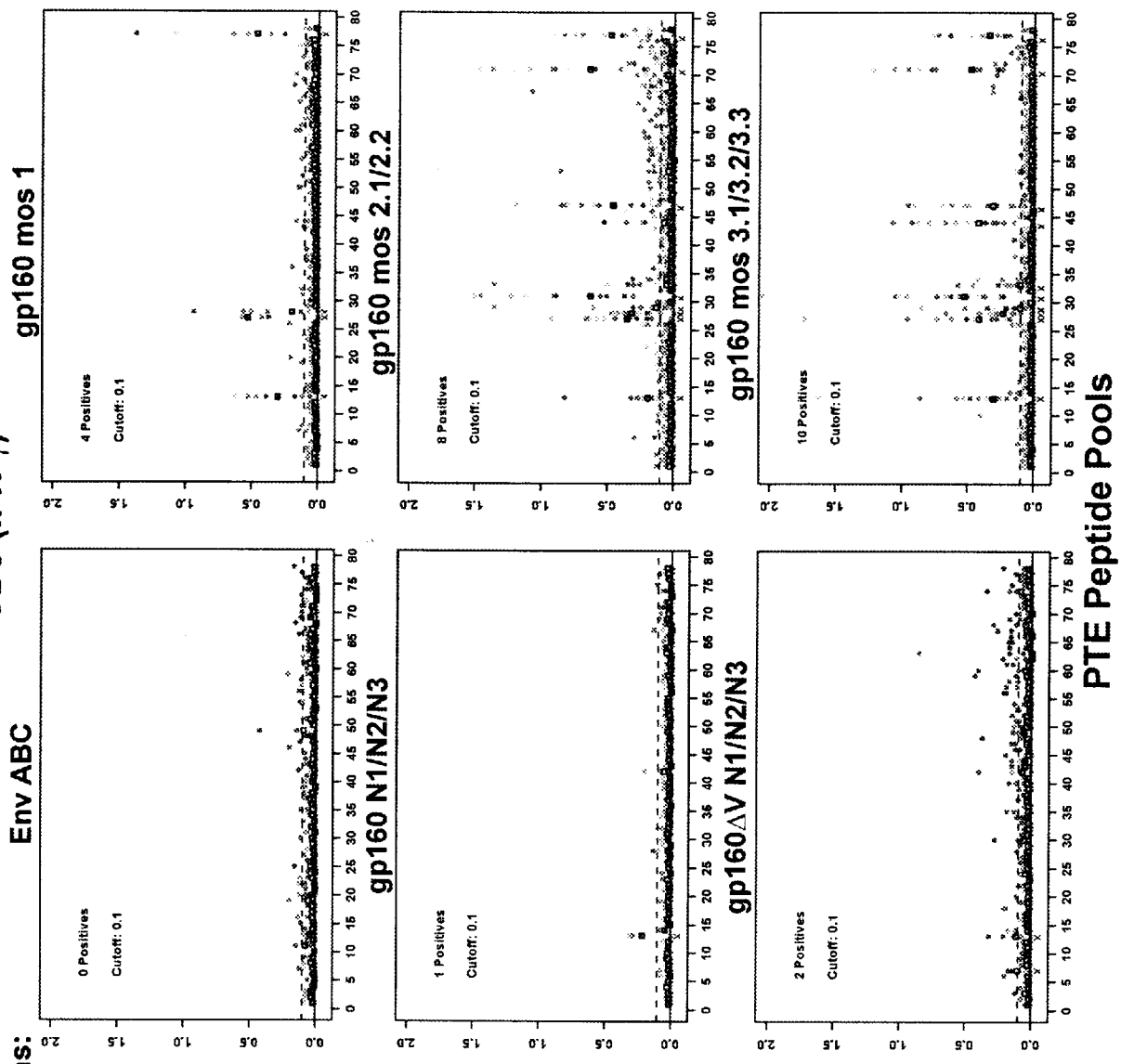

CD8 (IFN-γ)
Env ABC
Immunogens:
gp160 mos 1
gp160 mos 2.1/2.2
gp160 mos 3.1/3.2/3.3
gp160 N1/N2/N3
gp160ΔV N1/N2/N3
0 Positives
1 Positives
2 Positives
4 Positives
8 Positives
10 Positives
Cutoff: 0.1
PTE Peptide Pools

FIG. 4B

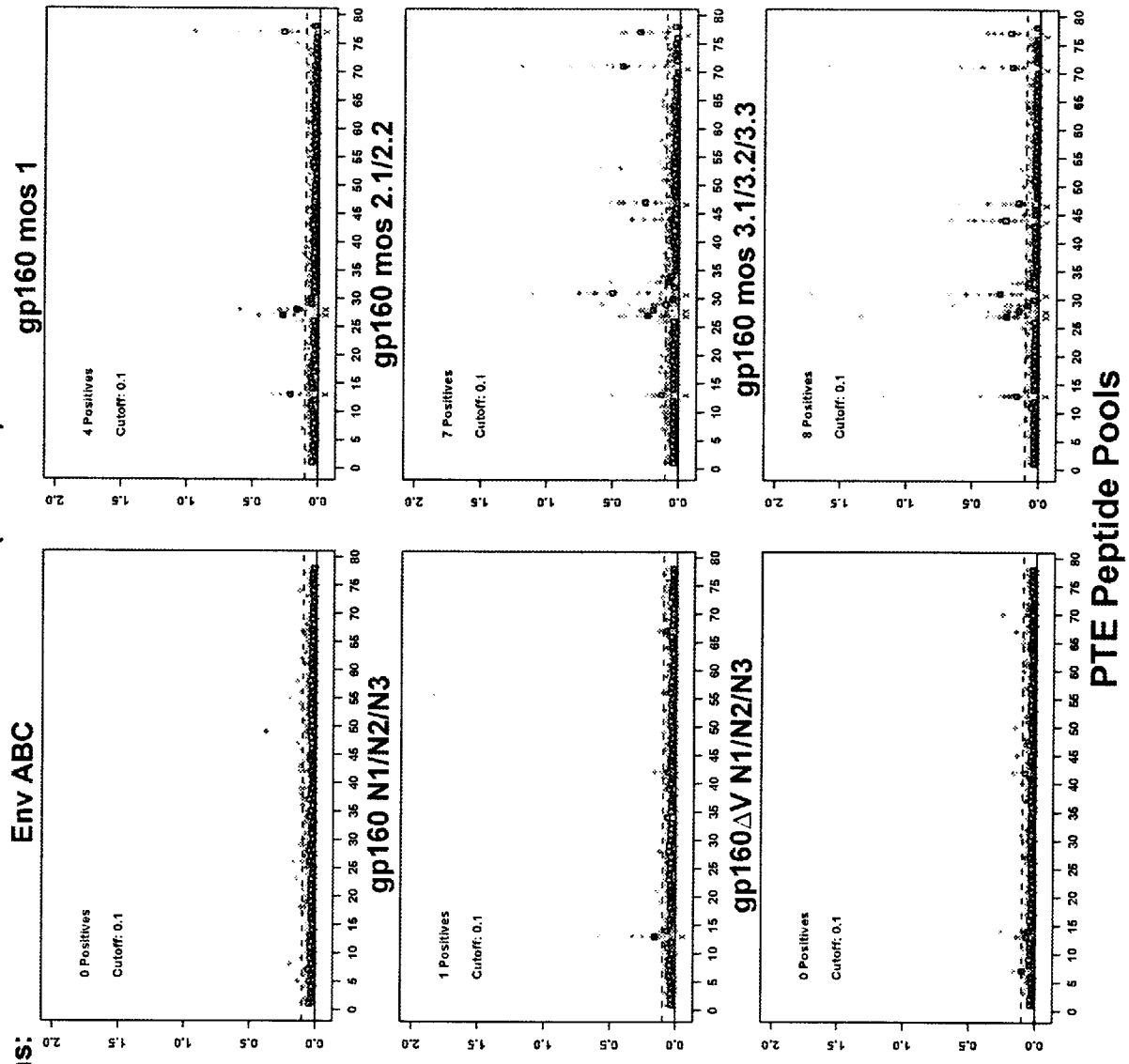

CD8 (TNF-α)
Immunogens:
Env ABC
gp160 mos 1
gp160 mos 2.1/2.2
gp160 mos 3.1/3.2/3.3
gp160 N1/N2/N3
gp160ΔV N1/N2/N3
4 Positives
7 Positives
8 Positives
1 Positives
0 Positives
Cutoff: 0.1
PTE Peptide Pools

FIG. 5A

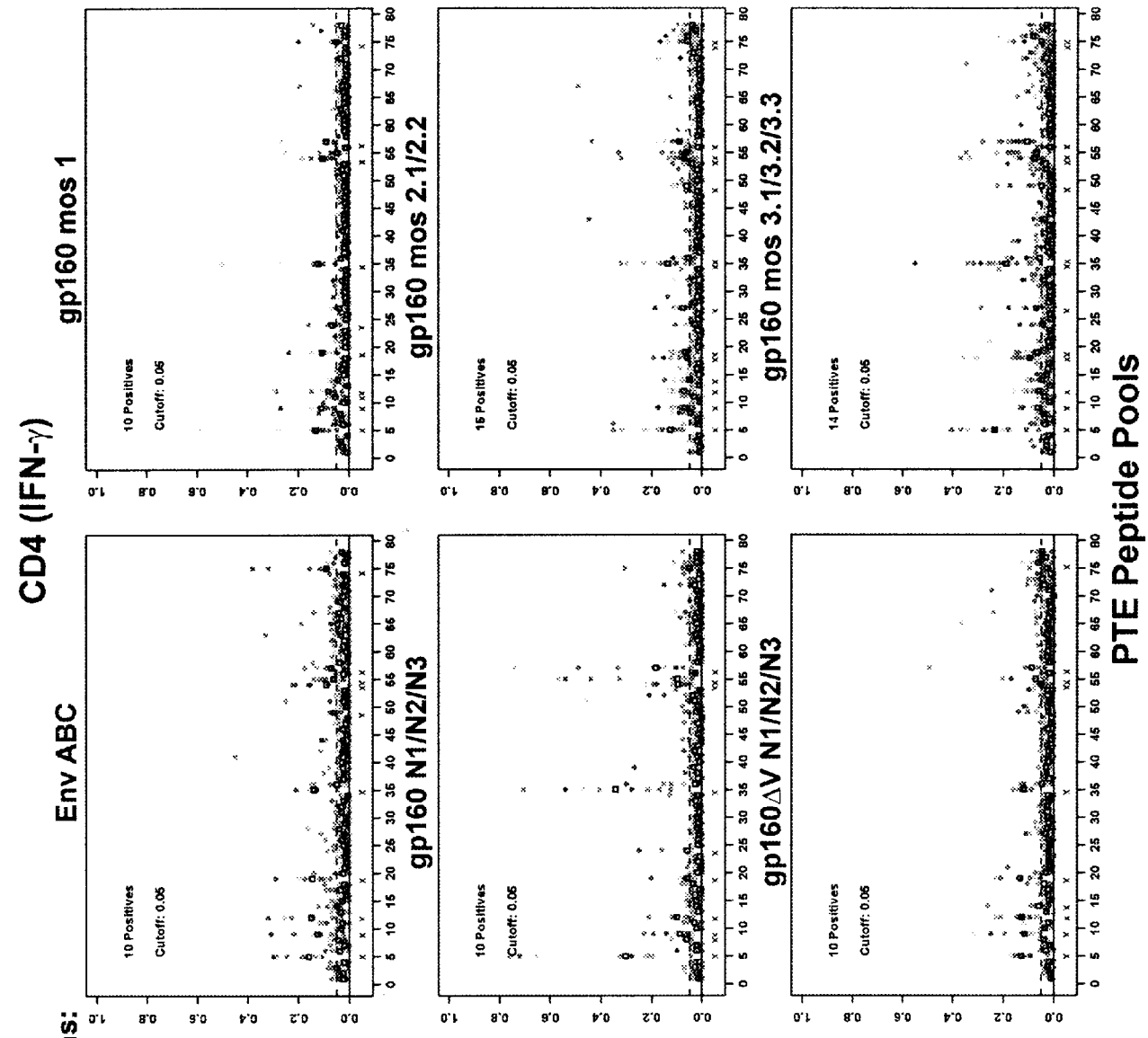

CD4 (IFN-γ)
Immunogens:
Env ABC
gp160 mos 1
gp160 N1/N2/N3
gp160 mos 2.1/2.2
gp160ΔV N1/N2/N3
gp160 mos 3.1/3.2/3.3
10 Positives
Cutoff: 0.05
15 Positives
14 Positives
PTE Peptide Pools

FIG. 5B

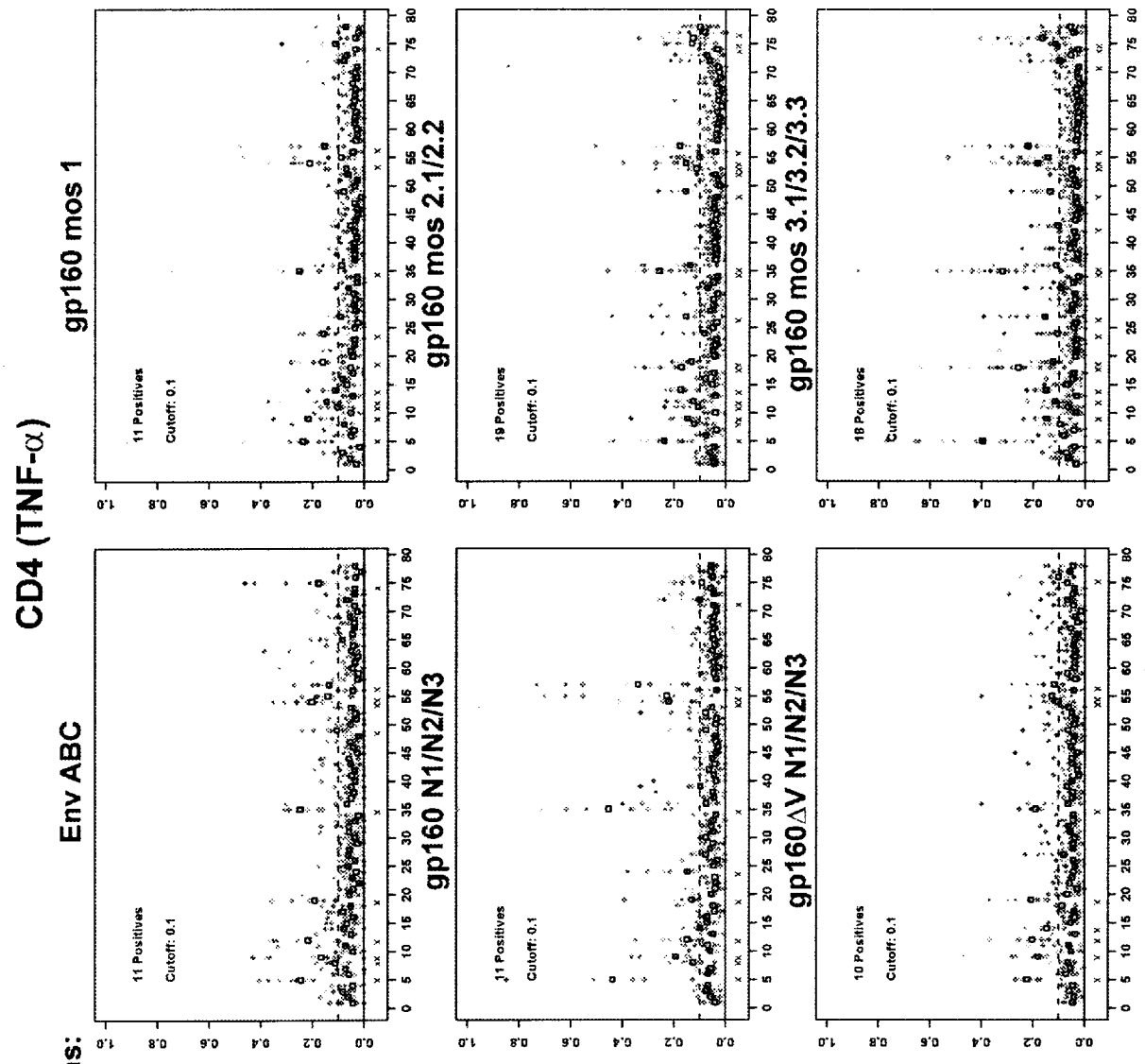
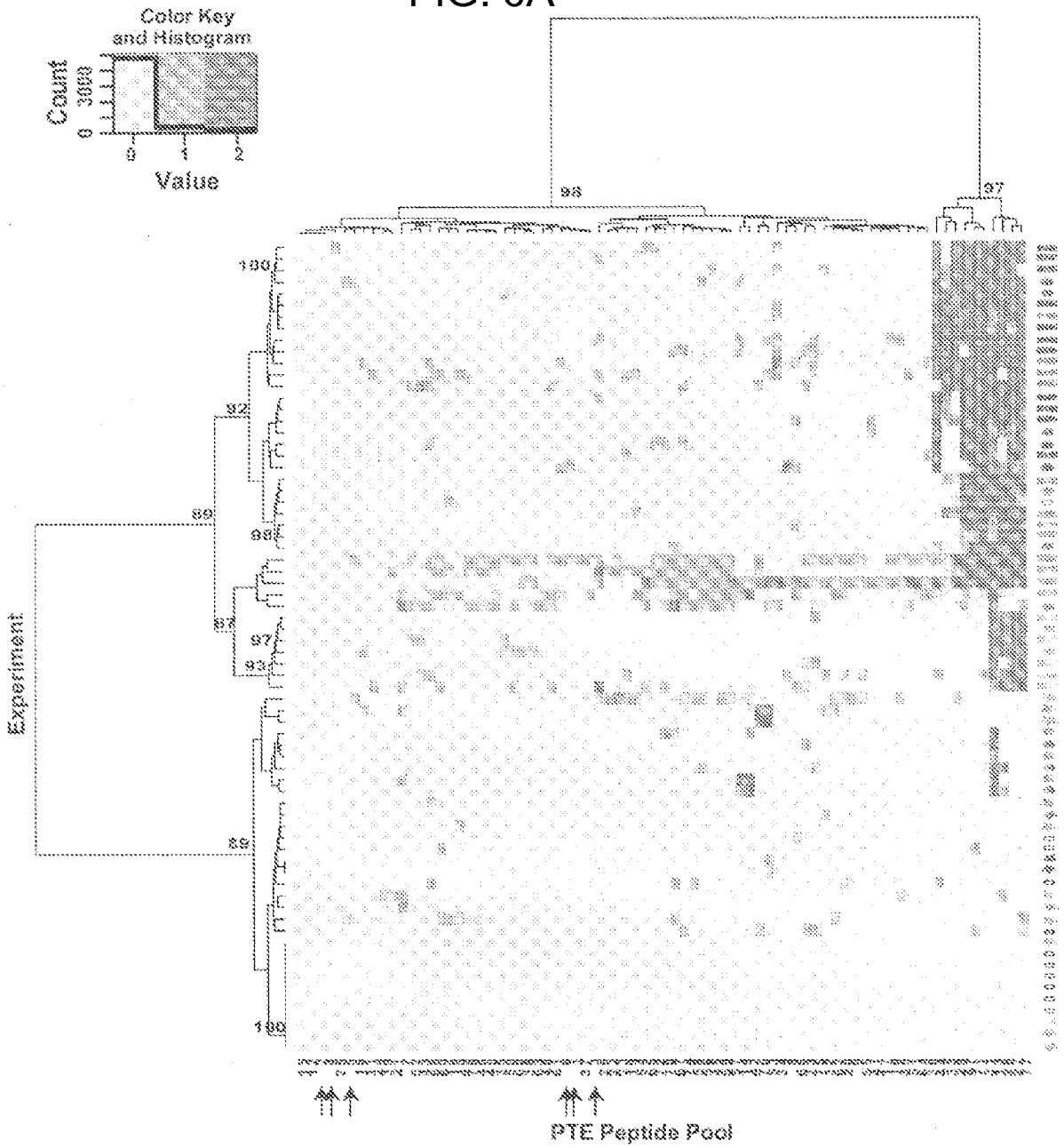

CD4 (TNF-α)
Immunogens:
Env ABC
gp160 mos 1
gp160 mos 2.1/2.2
gp160 mos 3.1/3.2/3.3
gp160 N1/N2/N3
gp160ΔV N1/N2/N3
11 Positives
19 Positives
18 Positives
10 Positives
Cutoff: 0.1
PTE Peptide Pools

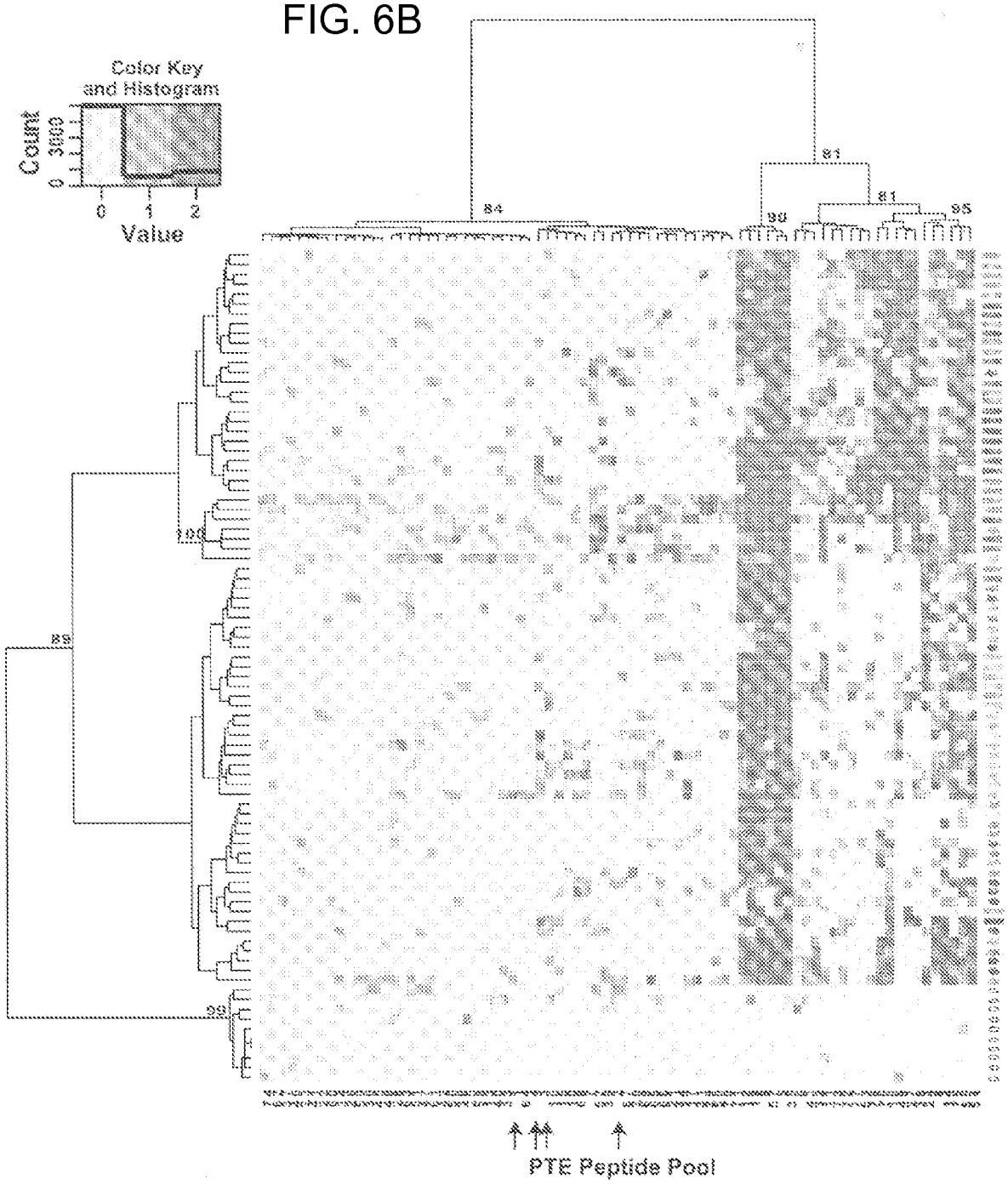
FIG. 6A



PTE Peptide Pool

Mosaics		
gp160	gp160ΔCFI	gp145ΔCFI
3		
2		
1		
Best Natural		
gp160	gp160ΔCFI	gp145ΔCFI
Best Natural ΔV		
gp160	gp160ΔCFI	gp145ΔCFI
EnvABC		
gp145ΔCFI		
Negative Cont		

FIG. 6B


Color Key
and Histogram
Count
3660
0
0
1
2
Value
84
81
81
95
98
100
89
99
PTE Peptide Pool

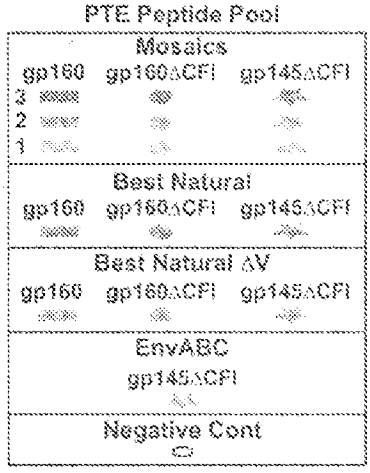

PTE Peptide Pool
Mosaics
gp160
gp160ΔCFI
gp145ΔCFI
3
2
1
Best Natural
gp160
gp160ΔCFI
gp145ΔCFI
Best Natural ΔV
gp160
gp160ΔCFI
gp145ΔCFI
EnvABC
gp145ΔCFI
Negative Cont

FIG. 7

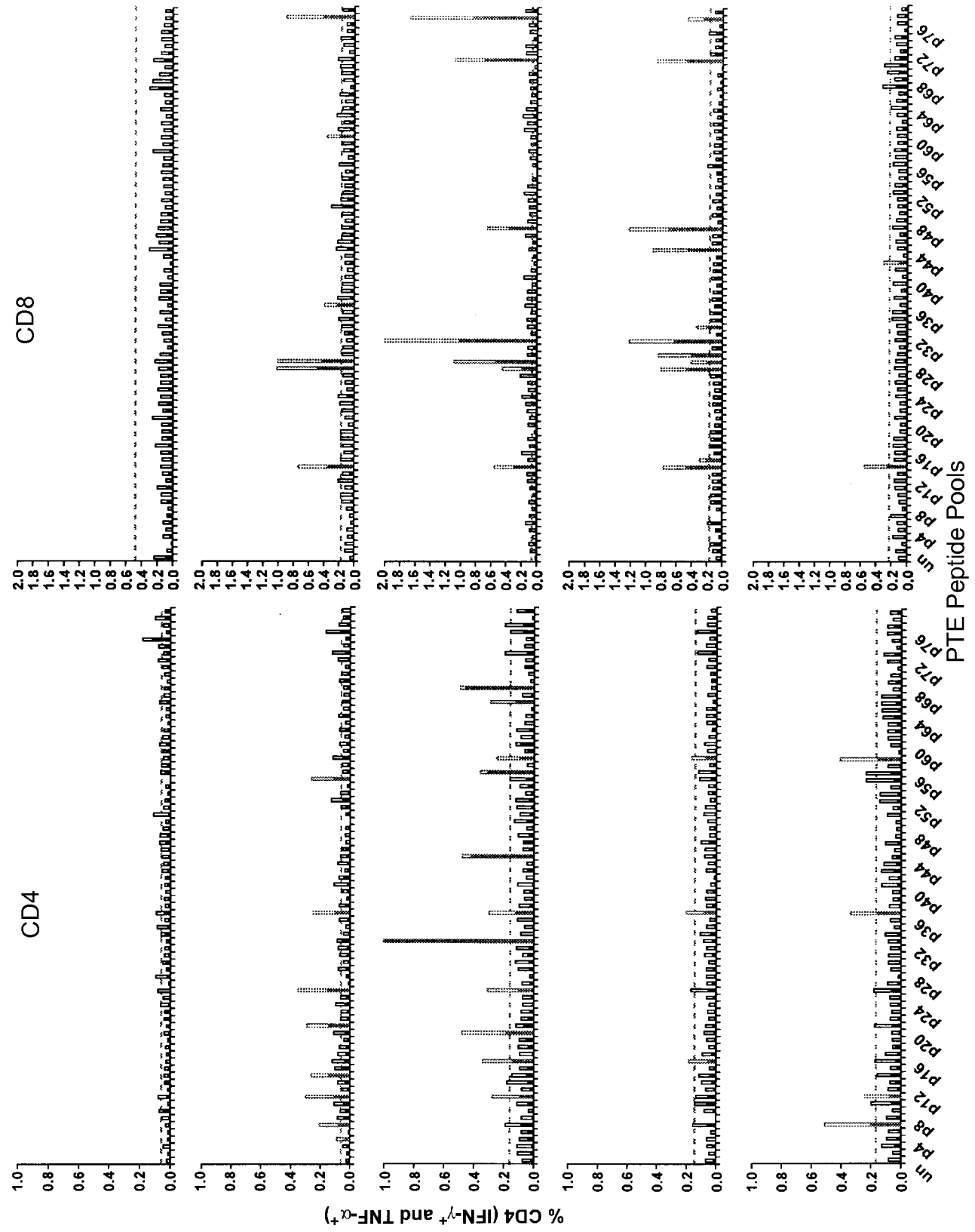

CD4
CD8
Control
gp160ΔCFI mos 1
gp160ΔCFI mos 2.1 mos 2.2
gp160ΔCFI mos 3.1 mos 3.2 mos 3.3
gp160ΔCFI N1 N2 N3
% CD4 (IFN-γ+ and TNF-α+)
PTE Peptide Pools

FIG. 7 Continued

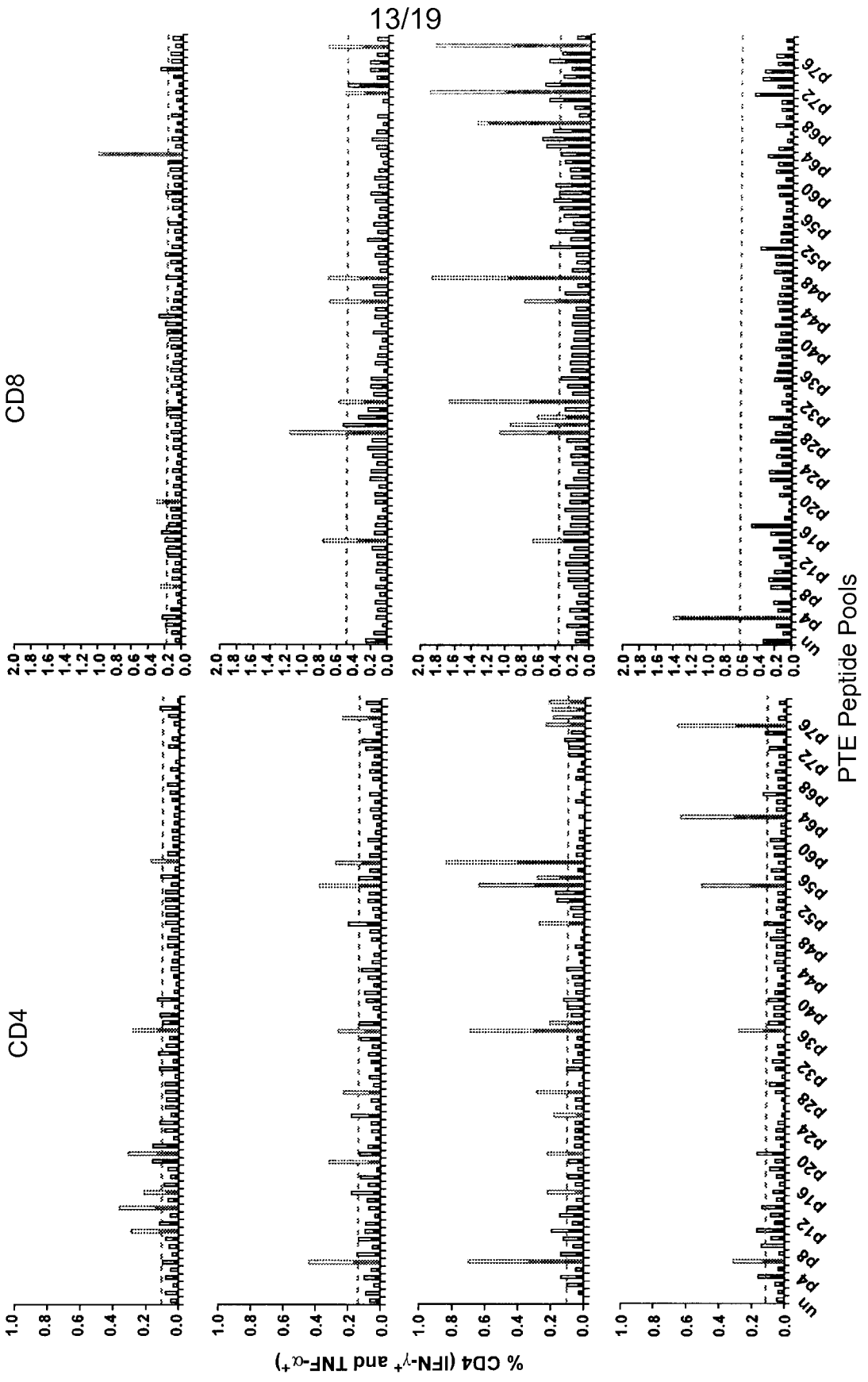


FIG. 8

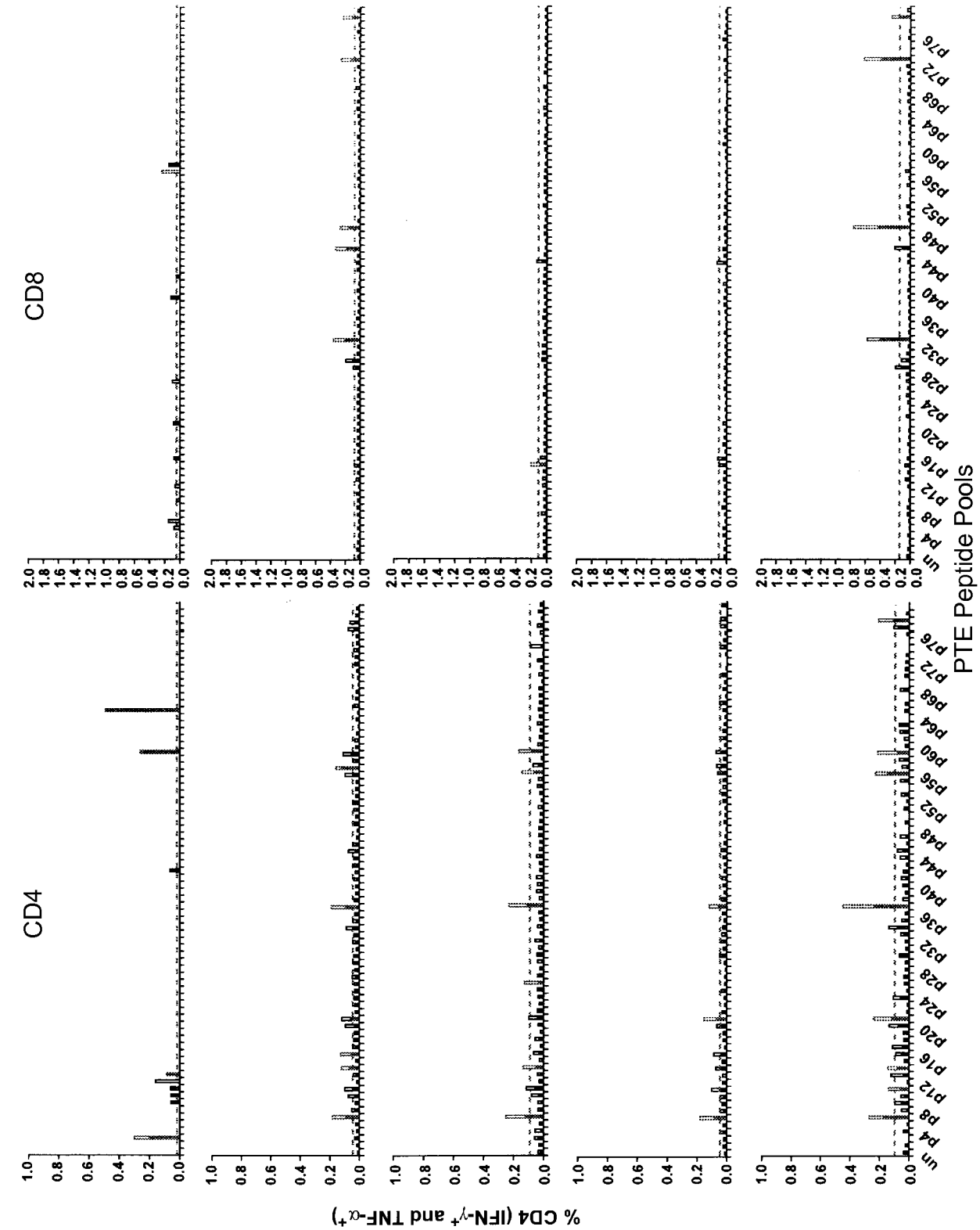

CD4
CD8
Control
gp160ΔCFI
mos 3.1
mos 3.2
mos 3.3
gp160ΔCFI
N1
N2
N3
gp160ΔVΔCFI
N1
N2
N3
gp160ΔVΔCFI
mos 3.1
mos 3.2
mos 3.3
% CD4 (IFN-γ+ and TNF-α+)
PTE Peptide Pools

FIG. 8 Continued

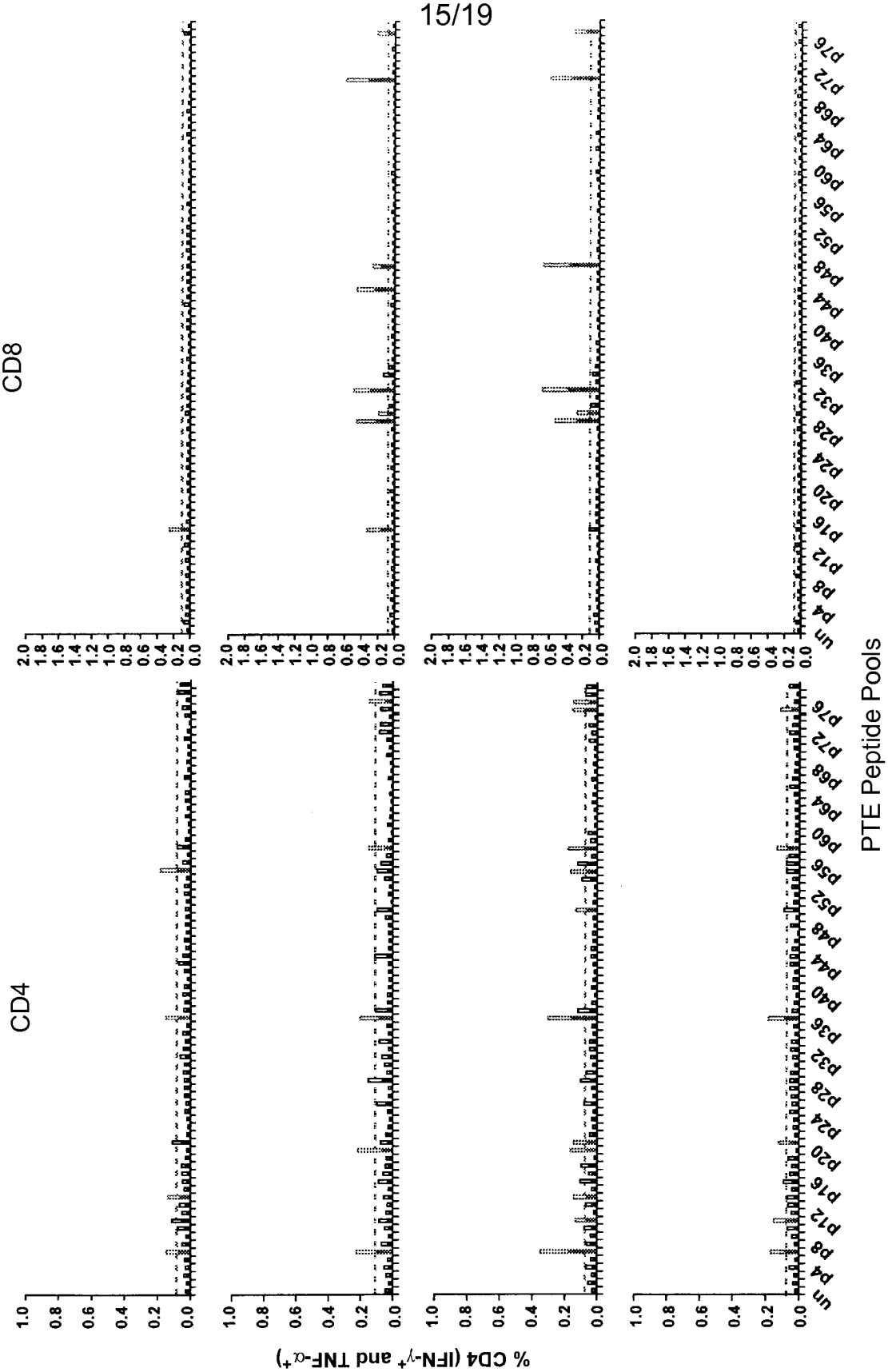


FIG. 9

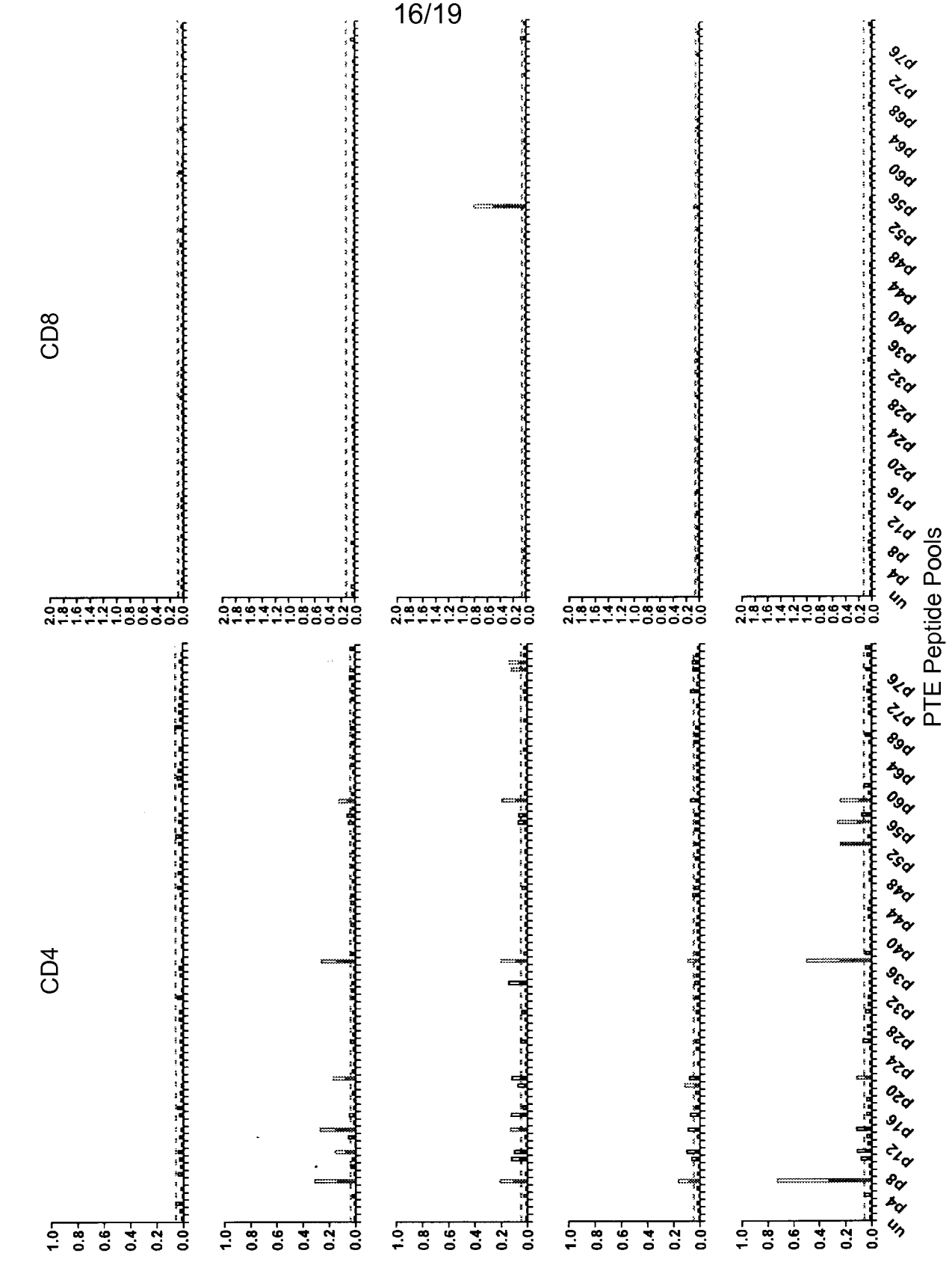

CD4
CD8
Control
gp145ΔCFI mos 1
gp145ΔCFI mos 2.1
mos 2.2
gp145ΔCFI mos 3.1
mos 3.2
mos 3.3
gp145ΔCFI N1
N2
N3
% CD4 (IFN-γ+ and TNF-α+)
PTE Peptide Pools

FIG. 9 Continued

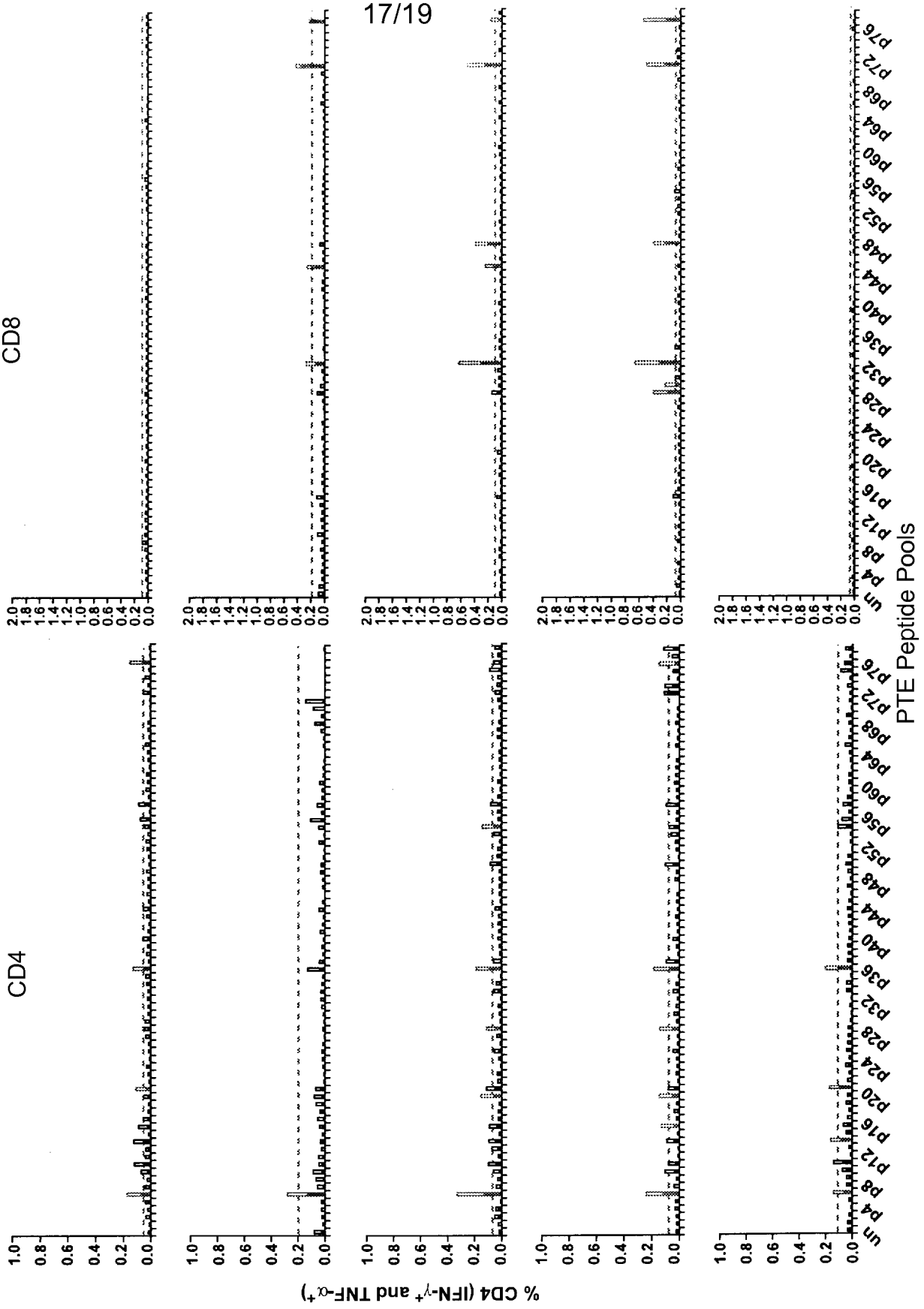


FIG. 10

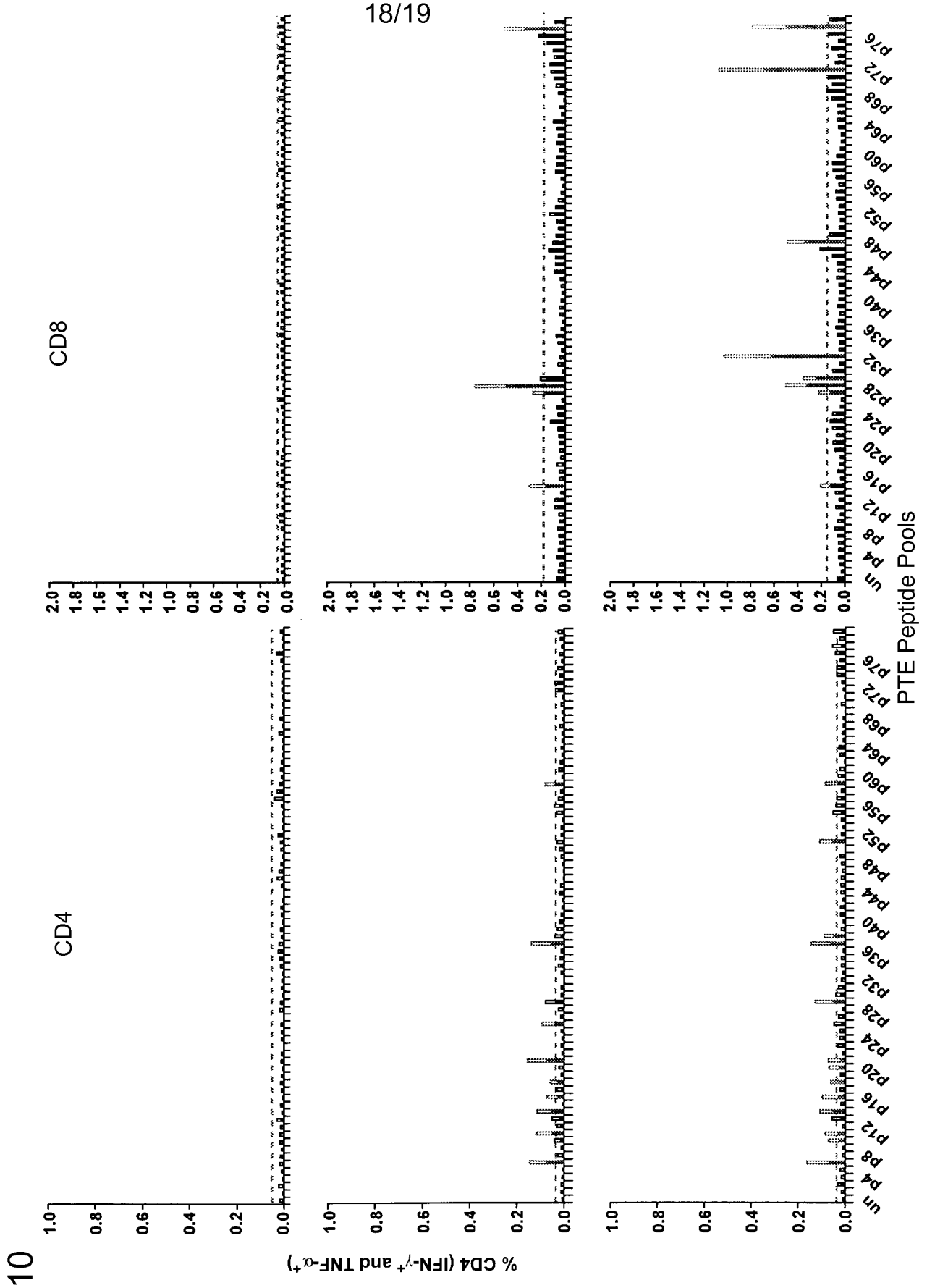
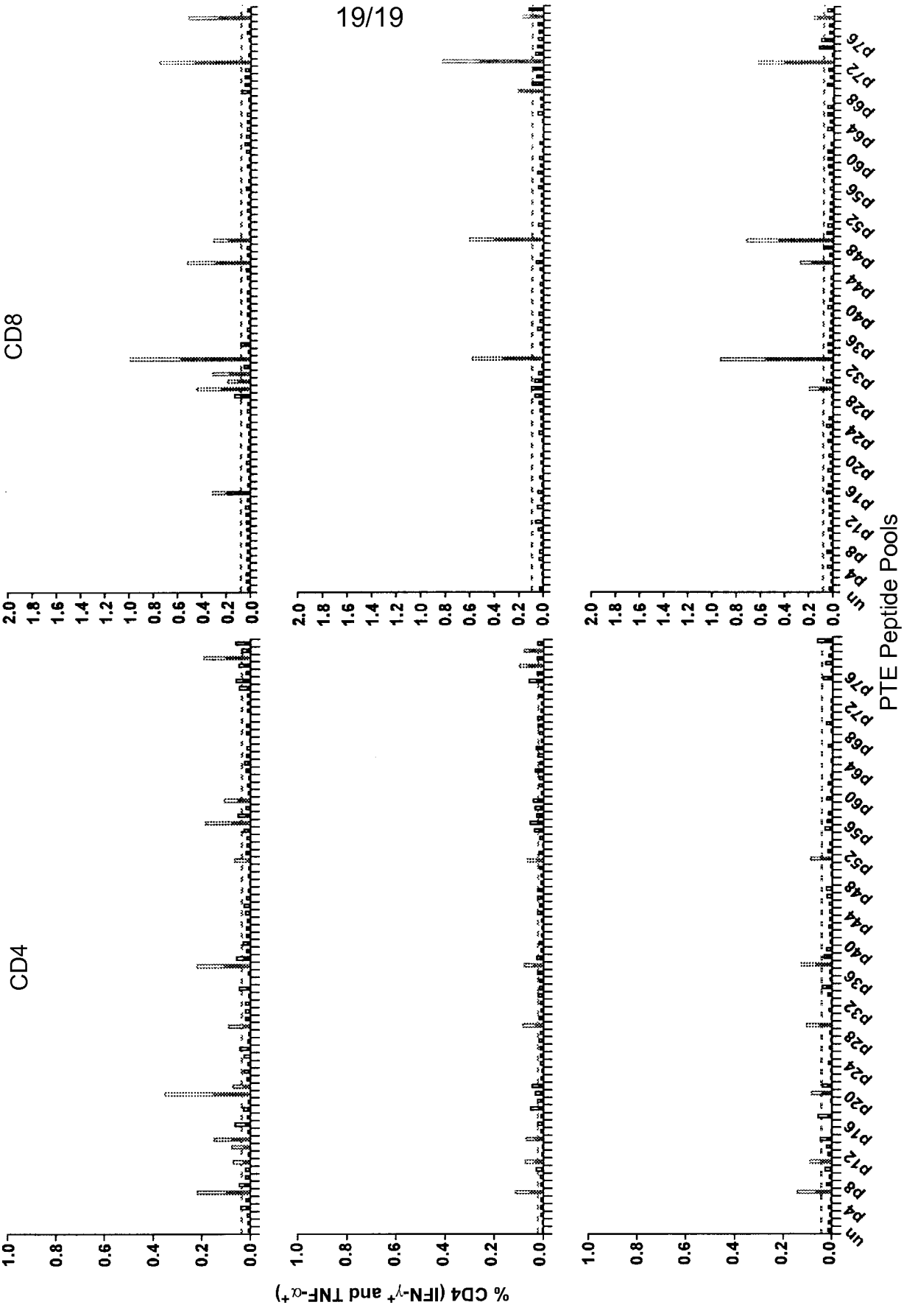

CD8
CD4
Control
gp160 mos 1
gp160 mos 2.1 mos 2.2
% CD4 (IFN-γ+ and TNF-α+)
PTE Peptide Pools

FIG. 10 Continued



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/060167

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K39/21 A61K39/295 A61P31/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CHAKRABARTI B K ET AL: "Modifications of the Human Immunodeficiency Virus Envelope glycoprotein enhance immunogenicity for genetic immunization" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 76, no. 11, 1 June 2002 (2002-06-01), pages 5357-5368, XP002210878 ISSN: 0022-538X abstract page 5357, right-hand column, paragraph 2 - page 5360, left-hand column, last paragraph figure 1 ----- -/--	1-15,17, 19-30

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 February 2010	15/02/2010

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Fax: (+31–70) 340–3016	Ulbrecht, Matthias

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/060167

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2005/034992 A2 (US GOVERNMENT [US]; NABEL GARY J [US]; CHAKRABARTI BIMAL [US]; KONG WI) 21 April 2005 (2005-04-21) page 84, line 3 - page 89, line 2716 page 39, line 5 - page 48, line 29 claims 1-3	1-15,17, 19-30
X	MALM MARIA ET AL: "CROSS-CLADE PROTECTION INDUCED BY HUMAN IMMUNODEFICIENCY VIRUS-1 DNA IMMUNOGENS EXPRESSING CONSENSUS SEQUENCES OF MULTIPLE GENES AND EPITOPES FROM SUBTYPES A, B, C, AND FGH" VIRAL IMMUNOLOGY, MARY ANN LIEBERT, INC., NEW YORK, US, vol. 18, no. 4, 1 January 2005 (2005-01-01), pages 678-688, XP009081289 ISSN: 0882-8245 abstract page 679, right-hand column, paragraph 1 - page 680, right-hand column, paragraph 2	1-6, 11-15, 17,19-30
X	SANTRA SAMPA ET AL: "A centralized gene-based HIV-1 vaccine elicits broad cross-clade cellular immune responses in rhesus monkeys." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 29 JUL 2008, vol. 105, no. 30, 29 July 2008 (2008-07-29), pages 10489-10494, XP007911455 ISSN: 1091-6490 abstract page 10489, right-hand column, paragraph 3 - page 10490, left-hand column, paragraph 1	1-6,8, 11-15, 17,19-30
X	WO 2007/024941 A2 (UNIV CALIFORNIA [US]; BETH ISRAEL DEACONNESS MEDICAL [US]; UNIV DUKE [) 1 March 2007 (2007-03-01) the whole document	1-6, 11-33
X,P	KONG WING-PUI ET AL: "Expanded breadth of the T-cell response to mosaic human immunodeficiency virus type 1 envelope DNA vaccination." JOURNAL OF VIROLOGY MAR 2009, vol. 83, no. 5, 24 December 2008 (2008-12-24), pages 2201-2215, XP007911460 ISSN: 1098-5514 the whole document	1-33

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/060167

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KIM YOUNG B ET AL: "Immunogenicity and ability of variable loop-deleted human immunodeficiency virus type 1 envelope glycoproteins to elicit neutralizing antibodies." VIROLOGY 5 JAN 2003, vol. 305, no. 1, 5 January 2003 (2003-01-05), pages 124-137, XP007911477 ISSN: 0042-6822 abstract	1-33
A	WO 2006/086284 A2 (MERCK & CO INC [US]; BETT ANDREW J [US]; CASIMIRO DANILO R [US]; SHIVE) 17 August 2006 (2006-08-17) the whole document	19-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2009/060167

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2005034992 A2	21-04-2005	AU 2004279362 A1	21-04-2005
		CA 2539068 A1	21-04-2005
		CN 1913919 A	14-02-2007
		EP 1675613 A2	05-07-2006
		US 2006275897 A1	07-12-2006
		US 2008286306 A1	20-11-2008
		ZA 200602858 A	28-05-2008
WO 2007024941 A2	01-03-2007	AU 2006283101 A1	01-03-2007
		CA 2620874 A1	01-03-2007
		EP 1917040 A2	07-05-2008
		US 2009324631 A1	31-12-2009
WO 2006086284 A2	17-08-2006	AU 2006212885 A1	17-08-2006
		CA 2597404 A1	17-08-2006
		EP 1851343 A2	07-11-2007
		JP 2008538894 T	13-11-2008
		US 2008254059 A1	16-10-2008