

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2005289439 B2**

- (54) Title
Optimized vaccines to provide protection against Ebola and other viruses
- (51) International Patent Classification(s)
C07K 14/08 (2006.01) **C07K 16/10** (2006.01)
A61K 39/12 (2006.01) **C12N 15/861** (2006.01)
- (21) Application No: **2005289439** (22) Date of Filing: **2005.09.27**
- (87) WIPO No: **WO06/037038**
- (30) Priority Data
- | (31) Number | (32) Date | (33) Country |
|-------------------|-------------------|--------------|
| 60/679,767 | 2005.05.10 | US |
| 60/715,874 | 2005.09.09 | US |
| 60/613,883 | 2004.09.27 | US |
| 60/677,606 | 2005.05.03 | US |
| 60/701,694 | 2005.07.22 | US |
- (43) Publication Date: **2006.04.06**
- (44) Accepted Journal Date: **2011.12.01**
- (71) Applicant(s)
The Government of the United States of America, as represented by the Secretary, Department of Health and Human Services;Crucell Holland B.V.
- (72) Inventor(s)
Pau, Maria Grazia;Nabel, Gary J.;Chakrabarti, Bimal;Sullivan, Nancy;Yang, Zhi-Yong;Goudsmit, Jaap
- (74) Agent / Attorney
FB Rice, Level 23 44 Market Street, Sydney, NSW, 2000
- (56) Related Art
WO 2003092582

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 April 2006 (06.04.2006)

PCT

(10) International Publication Number
WO 2006/037038 A1

(51) International Patent Classification:

C07K 14/08 (2006.01) *C12N 15/861* (2006.01)
C07K 16/10 (2006.01) *A61K 39/12* (2006.01)

(21) International Application Number:

PCT/US2005/034798

(22) International Filing Date:

27 September 2005 (27.09.2005)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/613,883	27 September 2004 (27.09.2004)	US
60/677,606	3 May 2005 (03.05.2005)	US
60/679,767	10 May 2005 (10.05.2005)	US
60/701,694	22 July 2005 (22.07.2005)	US
60/715,874	9 September 2005 (09.09.2005)	US

(71) Applicants (for all designated States except US): **THE GOVERNMENT OF THE UNITED STATES OF AMERICA**, as represented by **THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; National Institutes Of Health, Office Of Technology, 6011 Executive Blvd., Suite 325, Rockville, Md Maryland 20852-3804 (US). **CRUCCELL HOLLAND B.V.** [NL/NL]; N-2301 CA Leiden (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SULLIVAN, Nancy** [US/US]; 4217 Colchester Drive, Kensington, MD 20895 (US). **CHAKRABARTI, Bimal** [IN/US]; 361 West Side Drive, #102, Gaithersburg, MD 20878 (US). **YANG, Zhi-Yong** [CN/US]; 10004 Penfold Court, Potomac, MD

20854 (US). **PAU, Maria, Grazia** [IT/NL]; Bargelaan 80, NL-2333 CW Leiden (NL). **GOUDSMIT, Jaap** [NL/NL]; Koninginneweg 4, NL-1075 CX Amsterdam (NL). **NABEL, Gary, J.** [US/US]; 2520 30th Street NW, Washington, DC 20008 (US).

(74) Agent: **ALTMAN, Daniel, E.**; Knobbe, Martens, Olson And Bear, LLP, 2040 Main Street, Fourteenth Floor, Irvine, CA 92614 (US).

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

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

Declaration under Rule 4.17:

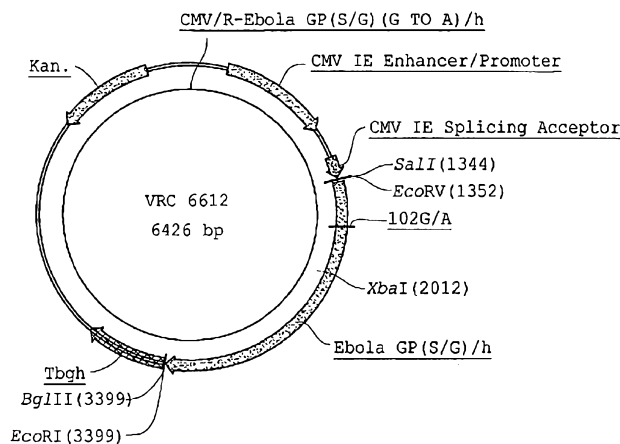
— of inventorship (Rule 4.17(iv))

Published:

— with international search report

[Continued on next page]

(54) Title: OPTIMIZED VACCINES TO PROVIDE PROTECTION AGAINST EBOLA AND OTHER VIRUSES



(57) Abstract: The invention is related to a nucleic acid molecule comprising a polynucleotide encoding a modified filovirus glycoprotein (GP) having at least one amino acid change located in a relatively conserved region of said GP that decreases *in vitro* cytotoxicity and retains immunogenicity when compared to *in vitro* cytotoxicity and immunogenicity of a wild type filovirus GP, and related modified filovirus GPs, plasmid DNAs, recombinant viruses, adenoviruses, pharmaceutical compositions, vaccine compositions, antibodies that are specifically reactive with the modified filovirus GPs, and related methods of making and using the same.

WO 2006/037038 A1



— *with amended claims*

Date of publication of the amended claims: 3 August 2006

(15) Information about Correction:

Previous Correction:

see PCT Gazette No. 21/2006 of 26 May 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

OPTIMIZED VACCINES TO PROVIDE PROTECTION AGAINST EBOLA AND OTHER VIRUSES

Related Applications

5 This application claims benefit of US Provisional Application No. 60/613,883 filed September 27, 2004, US Provisional Application No. 60/677,606 filed May 03, 2005, US Provisional Application No. 60/679,767 filed May 10, 2005, US Provisional Application No. 60/701,694 filed July 22, 2005, and US Provisional Application No. 60/715,874 filed September 9, 2005, all of which are hereby incorporated by reference in their entireties.

Field of the Invention

10 The present invention relates generally to viral vaccines and, more particularly, to filovirus vaccines and methods of eliciting an immune response against a filovirus or a disease caused by infection with filovirus.

Background of the Invention

15 Ebola virus and Marburg virus make up the family *Filoviridae*. The family is divided into two genera, currently designated as “Ebola-like viruses” and “Marburg-like viruses.” There are four subtypes within the “Ebola-like viruses,” Zaire (type species), Sudan, Reston, and Côte d’Ivoire (Ivory Coast). A single type, Marburg virus, makes up the “Marburg-like viruses.”

20 The glycoprotein (GP) is the sole structural protein making up the virion surface spikes that mediate virus entry into susceptible host cells through receptor binding. GP is the most studied of the filovirus proteins, not only for its importance in virus entry and pathogenesis but because it is a prime target for vaccine development. Research on filovirus GP has been facilitated through the use of recombinant DNA technology to permit
25 biochemical and functional assays without the constraints of working with the infectious filovirus.

GP expression in cultured human endothelial and epithelial cells causes cell rounding and detachment (Yang Z.-Y. et al. 2000 *Nat Med* 6:886-889). These effects require the presence of the mucin-like serine and threonine –rich domain of GP. The
30 cytotoxic effects of GP on macrophage and endothelial cell function disrupt inflammatory cell function and the integrity of the vasculature. In addition, by altering the cell surface expression of adhesion proteins and immune recognition molecules, Ebola virus may

disrupt processes critical to immune activation and cytolytic-T-cell function. These phenomena likely account for the dysregulation of the inflammatory response and the vascular dysfunction characteristic of lethal Ebola virus infection, providing a rationale for focusing on GP as a target for a preventative vaccine.

5

Summary of the Invention

The invention is related to a nucleic acid molecule comprising a polynucleotide encoding a modified filovirus glycoprotein (GP) having at least one amino acid change located in a relatively conserved region of said GP that decreases *in vitro* cytotoxicity and retains immunogenicity when compared to *in vitro* cytotoxicity and immunogenicity of a wild type filovirus GP.

10

In one embodiment, the amino acid change is positioned in the N-terminal domain, excluding the conserved cysteine residues, and is located at amino acid position 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299 or 300 in Ebola Zaire GP in an exemplary manner or corresponding thereto in other strains of said GP.

15

20

25

30

In another embodiment, the amino acid change is located at amino acid position 71 or 102 in Ebola Zaire GP in an exemplary manner or corresponding thereto in other strains of said GP.

In yet another embodiment, the amino acid change is E71D or G102A in Ebola Zaire GP in an exemplary manner or corresponding thereto in other strains of said GP.

In still another embodiment, the modified filovirus GP is encoded by the insert of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7 or SEQ ID NO:8, or sequence having at least 95% identity thereto.

In yet a further embodiment, the polynucleotide encoding the modified filovirus GP
5 has a sequence taken from the insert of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8, or sequence having at least 95% identity thereto.

Other embodiments of the invention are related to modified filovirus GPs encoded by the nucleic acid molecules, plasmid DNAs comprising the nucleic acid molecules,
10 recombinant viruses comprising the nucleic acid molecules, adenoviruses comprising the nucleic acid molecules, pharmaceutical compositions comprising the nucleic acid molecules or the modified filovirus GPs in a therapeutically effective dose, vaccine compositions comprising the nucleic acid molecules or the modified filovirus GPs in a prophylactically effective dose, antibodies that are specifically reactive with the modified
15 filovirus GPs, and related methods of making and using the same.

Brief Description of the Drawings

Figure 1. **A)** VRC6612 (pCMV/R-Ebola GP (S/G) (G to A)/h) construct map (see human codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of VRC6612 (SEQ ID NO: 1).

20 Figure 2. **A)** VRC6615 (pCMV/R-Ebola GP (Z) (full length G to A)/h) construct map (see human codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of VRC6615 (SEQ ID NO: 2).

Figure 3. **A)** VRC6613 (pCMV/R-Ebola GP (S/G) (E to D)/h) construct map (see human codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of
25 VRC6613 (SEQ ID NO: 3).

Figure 4. **A)** VRC6616 (pCMV/R-Ebola GP (Z) (full length E to D)/h) construct map (see human codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of VRC6613 (SEQ ID NO: 4).

Figure 5. **A)** VRC6712 (pCMV/R-Marburg/Angola GP/h) construct map (see human
30 codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of VRC6712 (SEQ ID NO: 5).

Figure 6. **A)** VRC6713 (pCMV/R Marburg/Angola GP (G102A)/h) construct map (see human codon-optimized Ebola/Marburg plasmids in Table 2). **B)** Nucleotide sequence of VRC6713 (SEQ ID NO: 6).

Figure 7. **A)** pAdApt.Ebo.GP.FL.(Z).E71D (adenoviral adaptor plasmid-Ebola/Zaire GP (full length E71D)/h). **B)** Nucleotide sequence of pAdApt.Ebo.GP.FL.(Z).E71D (SEQ ID NO: 7). Upper case is the coding sequence Ebo.GP.FL.(Z).E71D, the boxed region shows the E71D mutation and restriction site sequences used for cloning are bold and underlined.

5 Figure 8. **A)** pAdApt.Ebo.GP.FL.(S/G).E71D (adenoviral adaptor plasmid-Ebola/(Sudan/Gulu) GP (full length E71D)/h). **B)** Nucleotide sequence of pAdApt.Ebo.GP.FL.(S/G).E71D (SEQ ID NO: 8). Upper case is the coding sequence Ebo.GP.FL.(S/G).E71D, the boxed region shows the E71D mutation and restriction site sequences used for cloning are bold and underlined.

10 Figure 9. Schematic representation of Ebola virus GP. The GP1 and GP2 subunits of GP are drawn to scale (residue numbers are indicated below the diagram). The positions of the signal sequence, conserved cysteine residues (S), the mucin-like region (region of O-linked glycosylation), the furin cleavage site, the fusion peptide, the coiled-coil domain, and the membrane spanning domain are indicated.

15 Figure 10. Elimination of GP cytopathic effects and expression of transmembrane-deleted protein. **A)** Expression of GP(Δ TM) in 293 cells. Ebola GP proteins from supernatants and cell lysates in (A) were visualized by SDS- PAGE and Western blot using a polyclonal antibody against Ebola GP. **B)** Elimination of cell rounding by GP(Δ TM). 293 cells were transfected with a plasmid encoding vector control, Ebola GP or Ebola GP(Δ TM). Cell
20 monolayers were visualized under phase contrast using a Nikon 40X objective and photographed at 24 hours post transfection.

Figure 11. Comparative efficacy of GP and GP(Δ TM) for protection against Ebola virus challenge. **A)** Kaplan-Meier survival curve of macaques, immunized as indicated, and challenged with 1000 PFU of Ebola virus (1995 Zaire subtype) one month post immunization.
25 The x-axis indicates weeks post-challenge. n=3 in different immunization groups except for the GP(Z)+NP (10^{12}) group, n=4, and Control, n=1. **B)** Immune responses in immunized animals. Left and middle panels: intracellular flow cytometry was performed to quantify TNF- α production from Ebola-specific CD4⁺ or CD8⁺ lymphocytes, respectively, from animals immunized as indicated. Immune responses were measured at 3 weeks post-immunization.
30 Circle, diamond, square, triangle: responses for individual animals. Horizontal line: average of individual responses in the immunization group. Results represent the percent cytokine positive in the gated lymphocyte group and background stimulation (DMSO alone) has been subtracted from each sample. Right panel: ELISA titers of Ebola GP-specific antibodies in

serum of vaccinated animals collected at week 3 post-immunization. ELISA results represent endpoint dilution titers determined by optical density as described in Example 1.

Figure 12. Determination of lowest vaccine dose for immune protection against Ebola virus challenge by adenoviral vector vaccine. **A)** Kaplan-Meier survival curve of macaques: immunization and challenge were performed with the 1995 Zaire subtype Ebola virus as in Fig. 11A. **B)** Immune responses in immunized animals. Intracellular flow cytometry was performed to quantify TNF- α production from Ebola-specific CD4 (left panel) or CD8 (right panel) lymphocytes, respectively, from animals immunized as indicated. Immune responses were measured at 3 weeks post-immunization. Circle, diamond, square: responses for individual animals. Horizontal line: average of individual responses in the immunization group. Results represent the percent cytokine positive in the gated lymphocyte group and background stimulation (DMSO alone) has been subtracted from each sample (p-values obtained using unpaired Student's t-test. n. s. = not significant). **C)** Antibody responses in immunized animals. Anti-GP ELISA titers (left panel) and serum neutralizing antibody responses (right panel) were measured as described in Example 1.

Figure 13. Comparative efficacy of wild type and point mutant glycoprotein vaccines against lethal Ebola virus challenge. **A)** Kaplan-Meier survival curve of macaques: immunization and challenge were performed with the 1995 Zaire subtype Ebola virus as in Fig. 11A. **B)** Immune responses in immunized animals. Left and middle panels: intracellular flow cytometry was performed to quantify TNF- α production from Ebola-specific CD4 or CD8 lymphocytes, respectively, from animals immunized as indicated. Immune responses were measured at 3 weeks post-immunization. Circle, diamond, square: responses for individual animals. Horizontal line: average of individual responses in the immunization group. Results represent the percent cytokine positive in the gated lymphocyte group and background stimulation (DMSO alone) has been subtracted from each sample. Right panel: ELISA titers of Ebola GP-specific antibodies in serum of vaccinated animals collected at week 3 post-immunization. ELISA results represent endpoint dilution titers determined by optical density as described in Example 1.

Figure 14. Elimination of GP cytopathic effects with single point mutation. **A)** Expression of point mutants in 293 cells. Ebola GP proteins from supernatants and cell lysates in (A) were visualized by SDS- PAGE and Western blot using a polyclonal antibody against Ebola GP. **B)** Reactivity of point mutants with a conformation-dependent antibody. 293 cells were transfected with a control plasmid (Con), or plasmids expressing wild type (GP(Z)) or mutant (E71D(Z)) proteins. Eighteen hours post-transfection, cells were harvested,

stained with a GP-specific antibody and cell surface GP expression was analyzed by flow cytometry. C) Elimination of cell rounding by amino acid substitution at position 71. 293 cells were transfected with a plasmid encoding vector control, wild type Ebola glycoprotein from Zaire (GP(Z)) or Sudan-Gulu (GP(S/G)) or their respective point mutations (E71D(Z)),
5 E71D(S/G)). Cell monolayers were visualized under phase contrast using a Nikon 40X objective and photographed at 24 hours post transfection.

Figure 15. Cellular immune responses generated in NHP immunized with GP(Z) E71D/NP.

Figure 16. Humoral immune responses generated in NHP immunized with GP(Z)
10 E71D/NP.

Figure 17. Comparative efficacy of wild type GP vs. point mutant immunogens.

Figure 18. Elimination of GP cytopathic effects with single point mutation. Elimination of cell rounding by amino acid substitution at position 71 or 102. 293 cells were transfected with a plasmid encoding vector control, wildtype Ebola Zaire glycoprotein, WT
15 GP, or point mutations 71E/D, 102G/A, 138V/A or a mutation in which the transmembrane domain is deleted, Δ TM. Cell monolayers were visualized under phase contrast using a Nikon 40X objective and photographed at 24 hours post transfection. Cell rounding is eliminated with mutation at residue 71 or 102, and with transmembrane deletion.

Figure 19. Single point mutation in GP and retention of reactivity with neutralizing
20 antibodies. Reactivity of point mutants with a conformation-dependent antibody. 293 cells were transfected with a control plasmid, dashed line, or plasmids expressing wildtype Ebola GP (WT GP) or mutant (71E/D 102G/A), solid line, proteins. Eighteen hours post transfection, cells were harvested, stained with a GP-specific neutralizing antibody and cell surface GP expression was analyzed by flow cytometry.

25 Figure 20. Ebola Zaire GP amino acid sequence (SEQ ID NO: 9).

Table 1. Filovirus Genbank Accession Numbers

Virus	Subtype	Strain	"aa71"	"aa102"	Genbank Accession No.
Ebola	Ivory Coast	Coto d'Ivoire- Tai Forest, 1994	E	G	U28006
		Reston, 1989	E	G	U23152, NC 004161, AF034645
	Reston	Philippines, 1989	E	G	U23416
		Pennsylvania	E	G	AY769362, AF522874
		Siena, 1992	E	G	U23417
		Texas, 1996			
	Sudan	Boniface, 1976	E	G	U28134
		Maleo, 1979	E	G	U23069, Q66798
		Gulu- Uganda, 2000-2001	E	G	AY729654, AY344234, AY316199
	Zaire	Mayinga- Zaire, 1976	E	G	AY142960, AF499101, AF272001, AF086833, U23187, NC 002549
		Zaire- Zaire, 1976	E	G	AY354458, U28077, U31033, P87666
		Eckron-Zaire, 1976	E	G	U81161
		Bouee-96	E	G	AY058898
		Tandala- Zaire, 1977			
		Kikwit- Zaire, 1995			
		Gabon- Zaire, 1994-1997	E	G	U77384
		Mendamba A, 2001	G	G	AY526105 (G at 71)
		Mvoula, 2003	E	G	AY526104
		Yembelengoye, 2002		G	AY526103 (partial seq, unknown at 71)
Marburg		Entsiami, 2002	G	G	AY526102 (G at 71)
		Makoukou, 2001	G	G	AY526101 (G at 71)
		Etikangaye, 2001	G	G	AY526100 (G at 71)
		Olloba	G	G	AY526099 (G at 71)
		Mendamba B, 2001	G	G	AY526098 (G at 71)
		Musoke- Kenya, 1980		G	Z12132
		Ratayczak- West Germany, 1967		G	AF005735
		Popp- West Germany, 1967		G	NC 001608, Z29337, X68493
		Voegel- Yugoslavia, 1967			
		Ozolin- Zimbabwe, 1987		G	AY358025, AF005733

Virus	Subtype	Strain	"aa71"	"aa102"	Genbank Accession No.
		Ravn- Kenya, 1987		G	AF005734
		Angola, 2004-2005		G	(see Fig. 5 and SEQ ID NO: 5, human codon-optimized)
		pp3 guinea pig lethal variant		G	AY430365
		pp4 guinea pig nonlethal variant		G	AY430366

Table 2. Human Codon-Optimized Ebola/Marburg Plasmids

Construct	Construct Name/Description	Construct Map Name	SEQ ID NO	Figure
VRC6612	pCMV/R-Ebola GP (S/G) (G to A)/h*	CMV/R-Ebola GP (S/G) (G to A)/h	1	1
VRC6615	pCMV/R-Ebola GP (Z) (full length G to A)/h	CMV/R-Ebola GP (Z) (full length G to A)/h	2	2
VRC6613	pCMV/R-Ebola GP (S/G) (E to D)/h	CMV/R-Ebola GP (S/G) (E to D)/h	3	3
VRC6616	pCMV/R-Ebola GP (Z) (full length E to D)/h	CMV/R-Ebola GP (Z) (full length E to D)/h	4	4
VRC6712	pCMV/R-Marburg/Angola GP/h	CMV/R-Marburg/Angola GP/h	5	5
VRC6713	pCMV/R Marburg/Angola GP (G102A)/h	CMV/R Marburg/Angola GP (G102A)/h	6	6
pAdApt.Ebo.GP.FL.(Z).E71D	pAdApt.Ebo.GP.FL.(Z).E71D (adenoviral adaptor plasmid-Ebola/Zaire GP (full length E71D)/h)	AdApt.Ebo.GP.FL.(Z).E71D (adenoviral adaptor plasmid-Ebola/Zaire GP (full length E71D)/h)	7	7
pAdApt.Ebo.GP.FL.(S/G).E71D	pAdApt.Ebo.GP.FL.(S/G).E71D (adenoviral adaptor plasmid-Ebola/(Sudan/Gulu) GP (full length E71D)/h)	AdApt.Ebo.GP.FL.(S/G).E71D (adenoviral adaptor plasmid-Ebola/(Sudan/Gulu) GP (full length E71D)/h)	8	8

*h = human codon-optimized

Detailed Description of the Preferred Embodiment

In the absence of effective therapies for Ebola virus infection, the development of a vaccine becomes an important strategy to contain outbreaks. Immunization with DNA and/or replication-defective adenoviral (rAd) vectors encoding the Ebola glycoprotein (GP) and nucleoprotein (NP) has been previously shown to confer specific protective immunity in nonhuman primates (Sullivan, N. J. et al. 2000 *Nature* **408**:605-609; Sullivan, N. J. et al. 2003 *Nature* **424**:681-684). GP can exert cytopathic effects on transfected cells *in vitro* (Yang, Z.-Y. et al. 2000 *Nat Med* **6**:886-889) and multiple GP forms have been identified in nature, raising the question of which would be optimal for a human vaccine. To address this question, we have explored the efficacy of mutant GPs from multiple Ebola virus strains with reduced *in vitro* cytopathicity and analyzed their protective effects in the primate challenge model, with or without NP. Deletion of the GP transmembrane domain eliminated *in vitro* cytopathicity but reduced its protective efficacy by at least one order of magnitude. In contrast, single point mutations were identified that abolish *in vitro* cytopathicity but retained immunogenicity and conferred immune protection in the absence of NP. The minimal effective rAd dose was established at 10^{10} particles, two logs lower than used previously. Expression of specific GPs alone vectored by rAd are sufficient to confer protection against lethal challenge in a relevant nonhuman primate model, providing the basis for identification of a vaccine.

Definitions

Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. See, e.g., Singleton P and Sainsbury D., Dictionary of Microbiology and Molecular Biology 3rd ed., J. Wiley & Sons, Chichester, New York, 2001, and Fields Virology 4th ed., Knipe D.M. and Howley P.M. eds, Lippincott Williams & Wilkins, Philadelphia 2001.

Nucleic Acid Molecules

As indicated herein, nucleic acid molecules of the present invention may be in the form of RNA or in the form of DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) encoding a modified filovirus structural gene product; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode an ORF of a modified filovirus structural gene product. Of course, the genetic code is well known in the art. Degenerate variants optimized for human codon usage are preferred.

A filovirus structural gene product, *e.g.*, glycoprotein (GP), the sole structural protein making up the virion surface spikes that mediate virus entry into susceptible host cells through receptor binding, is modified by having at least one amino acid change that decreases *in vitro* cytotoxicity and retains immunogenicity when compared to *in vitro* cytotoxicity and immunogenicity of a wild type (*e.g.*, naturally occurring) filovirus GP. The amino acids of particular importance to the *in vitro* cytotoxicity are by no means limited to the exact position as defined for the, *e.g.*, Zaire strain of Ebola virus but are simply used in an exemplary manner to point out the preferred amino acids being at that position or corresponding to that position in other strains such as found in Sudan strain of Ebola virus or Angola strain of Marburg virus and filoviruses in general since they are highly conserved. For filoviruses other than the Ebola Zaire strain the numbering of the positions of the preferred amino acids is often different but an expert in the field of the molecular biology of filoviruses will easily identify these preferred amino acids by their position relative to the highly conserved amino acids of said glycoprotein.

The present invention is further directed to fragments of the nucleic acid molecules described herein. By a fragment of a nucleic acid molecule having the nucleotide sequence of an ORF encoding a modified filovirus structural gene product is intended fragments at least about 15 nt., and more preferably at least about 20 nt., still more preferably at least about 30 nt., and even more preferably, at least about 40 nt. in length. Of course, larger

fragments 50, 100, 150, 200, 250, 300, 350, 400, 450, or 500 nt. in length are also intended according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence of the ORF encoding a modified filovirus structural gene product. By a fragment at least 20 nt. in length, for example, is intended fragments which include 20 or
5 more contiguous bases from the nucleotide sequence of the ORF encoding a modified filovirus structural gene product.

Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the filovirus structural protein. In particular, such nucleic acid fragments of the present invention include nucleic acid
10 molecules encoding epitope-bearing domains of a filovirus structural protein, where the domain is the N-terminal domain, the mucin-like domain, the furin cleavage site, the fusion peptide domain, the coiled-coil domain, the membrane spanning domain, and the intracellular domain, and any combination thereof, for example, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the membrane spanning and
15 intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the coiled-coil domain and membrane-spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the furin cleavage site,
20 fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the mucin-like domain, furin cleavage site, fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain. Another example is a filovirus glycoprotein having an amino, internal, or carboxy deletion to delete the mucin-like domain, the furin
25 cleavage site, the fusion peptide domain, the coiled-coil domain, the membrane-spanning domain, or the intracellular domain.

In another aspect, the invention provides a nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above. By "stringent
30 hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5 x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml

denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 x SSC at about 65°C.

By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt.),
5 and more preferably at least about 20 nt., still more preferably at least about 30 nt., and even more preferably about 30-70 nt. of the reference polynucleotide.

By a portion of a polynucleotide of "at least 20 nt. in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide. Of course, a polynucleotide which hybridizes only to a poly A sequence or
10 a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly A stretch or the complement thereof (*e.g.*, practically any double-stranded cDNA clone).

As indicated herein, nucleic acid molecules of the present invention which encode a
15 filovirus structural gene product may include, but are not limited to those encoding the amino acid sequence of the full-length polypeptide, by itself, the coding sequence for the full-length polypeptide and additional sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro-protein sequence, the coding sequence of the full-length polypeptide, with or without the aforementioned additional coding
20 sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example, ribosome binding and stability of mRNA; and additional coding sequence which codes for additional amino acids, such as those which
25 provide additional functionalities.

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the filovirus structural gene product. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a
30 genome of an organism. (*Genes II*, Lewin, B., ed., John Wiley & Sons, 1985 New York). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions, which may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the filovirus structural gene product or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 95% identical, and more preferably at least 96%, 97%, 98% or 99% identical to a nucleotide sequence encoding a polypeptide having the amino acid sequence of a modified filovirus structural gene product or fragment thereof or a nucleotide sequence complementary thereto.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a filovirus structural gene product is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the Ebola virus structural gene product. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 95%, 96%, 97%, 98% or 99% identical to the reference nucleotide sequence can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). Bestfit uses the local homology algorithm of Smith and Waterman 1981 *Advances in Applied Mathematics* 2:482-489, to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for

instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

5 The present application is directed to nucleic acid molecules at least 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequences described herein which encode a polypeptide having Ebola or Marburg polypeptide activity. By "a polypeptide having Ebola or Marburg, polypeptide activity" is intended polypeptides exhibiting Ebola or Marburg polypeptide activity in a particular biological assay. For example, GP polypeptide activity
10 can be measured for changes in biological activity such as receptor binding activity, connection between GP1 and GP2, and contribution to the formation of the stalk structure of the virion peplomer, and modified GP polypeptide activity by decrease of *in vitro* cytotoxicity while retaining immunogenicity.

 Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art
15 will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 95%, 96%, 97%, 98%, or 99% identical to a nucleic acid sequence described herein will encode a polypeptide "having Ebola or Marburg polypeptide activity". In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above
20 described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having Ebola or Marburg polypeptide activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (*e.g.*, replacing one aliphatic amino acid with a second
25 aliphatic amino acid).

 For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al. 1990 *Science* **247**:1306-1310, wherein the authors indicate that proteins are surprisingly tolerant of amino acid substitutions.

Polypeptides and Fragments

The invention further provides a filovirus polypeptide having the amino acid sequence encoded by an open reading frame (ORF) of a modified filovirus structural gene, or a peptide or polypeptide comprising a portion thereof (*e.g.*, soluble GP).

5 It will be recognized in the art that some amino acid sequences of the filovirus polypeptides can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

10 Thus, the invention further includes variations of the filovirus polypeptide which show substantial filovirus polypeptide activity or which include regions of filovirus protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions. As indicated, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J.U. et al. 1990 *Science* **247**:1306-1310.

15 Thus, the fragment, derivative or analog of the polypeptide of the invention may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues include a substituent group, or (iii) one
20 in which additional amino acids are fused to the mature polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

25 As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table A).

Table A. Conservative Amino Acid Substitutions

Aromatic	Phenylalanine Tryptophan Tyrosine
Ionizable: Acidic	Aspartic Acid Glutamic Acid
Ionizable: Basic	Arginine Histidine Lysine
Nonionizable Polar	Asparagine Glutamine Selenocystine Serine Threonine
Nonpolar (Hydrophobic)	Alanine Glycine Isoleucine Leucine Proline Valine
Sulfur Containing	Cysteine Methionine

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of amino acid substitutions for any given filovirus polypeptide will not be more than 50, 40, 30, 20, 10, 5 or 3.

Amino acids in the filovirus polypeptides of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham & Wells 1989 *Science* **244**:1081-1085). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding activity, connection between GP1 and GP2, and contribution to the formation of the stalk structure of the virion peplomer, and modified GP polypeptide activity by decrease of *in vitro* cytotoxicity while retaining immunogenicity.

The polypeptides of the present invention are conveniently provided in an isolated form. By "isolated polypeptide" is intended a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention. Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially,

from a recombinant host cell or a native source. For example, a recombinantly produced version of the filovirus polypeptide can be substantially purified by the one-step method described in Smith and Johnson 1988 *Gene* 67:31-40.

5 The polypeptides of the present invention include a polypeptide comprising a polypeptide having the amino acid sequence of a modified filovirus structural gene product or portion thereof or encoded by a nucleic acid sequence described herein; as well as polypeptides which are at least 95% identical, and more preferably at least 96%, 97%, 98%, or 99% identical to those described above and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids.

10 By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of an filovirus polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the filovirus polypeptide. In other words, to
15 obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy
20 terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 95%, 96%, 97%, 98%, or 99% identical to a reference amino acid sequence can be determined
25 conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are
30 set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

In another aspect, the invention provides portions of the polypeptides described herein with at least 30 amino acids and more preferably at least 50 amino acids. Preferred portions of the present invention include polypeptides comprising an epitope-bearing portion of a filovirus structural protein. In particular, preferred portions of the present invention include polypeptides comprising an epitope-bearing domain of a filovirus structural protein, where the domain is the N-terminal domain, the mucin-like domain, the furin cleavage site, the fusion peptide domain, the coiled-coil domain, the membrane-spanning domain, and the intracellular domain, and any combination thereof, for example, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the membrane spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the coiled-coil domain and membrane spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain, a filovirus glycoprotein having a truncation at the carboxy terminus to delete the furin cleavage site, fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain, and a filovirus glycoprotein having a truncation at the carboxy terminus to delete the mucin-like domain, furin cleavage site, fusion peptide domain, coiled-coil domain, and membrane-spanning and intracellular domain. Another example is a filovirus glycoprotein having an amino, internal, or carboxy deletion to delete the mucin-like domain, the furin cleavage site, the fusion peptide domain, the coiled-coil domain, the membrane-spanning domain, or the intracellular domain.

The polypeptides of the invention may be produced by any conventional means (Houghten, R.A. 1985 *PNAS USA* **82**:5131-5135). The "Simultaneous Multiple Peptide Synthesis (SMPS)" process is described in U.S. Pat. No. 4,631,211 to Houghten et al. (1986).

The present invention also relates to vectors which include the nucleic acid molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of filovirus polypeptides or fragments thereof by recombinant techniques.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus,

it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will preferably include a translation initiation at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60, and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986).

The filovirus polypeptides can be recovered and purified from recombinant cell cultures by well known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Antibodies

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, humanized, human, and CDR-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) and other binding proteins specific for filovirus GP polypeptides or fragments thereof. The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind a filovirus GP polypeptide exclusively (*i.e.*, are able to distinguish a filovirus GP polypeptide from related polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Antibodies A Laboratory Manual, Harlow et al. (Eds), Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the filovirus GP polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, a filovirus GP polypeptide of the invention from which the fragment was derived. The specific antibodies of the invention are envisioned as having utility for diagnostic purposes and passive immunization.

Use of Recombinant Virus to Induce Immune Response to Antigen

The present invention relates to generation of a CD8+ T cell immune response against an antigen and also eliciting an antibody response. More particularly, the present invention relates to "accelerated" immunization regimes in which the immune response is induced by administration of a single dose form and "prime and boost" immunization regimes in which the immune response induced by administration of a priming composition is boosted by administration of a boosting composition. The present invention, in one embodiment, is based on the inventors' experimental demonstration that effective immunization can be achieved using recombinant virus, *e.g.*, adenovirus, optionally as boosting compositions following priming with any of a variety of different types of priming compositions.

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

The present invention, in one embodiment, employs recombinant virus, *e.g.*, adenovirus, which, as the experiments described below show, has been found to be an effective means for inducing a CD8+ T cell immune response, optionally as a boosting composition primed by antigen using any of a variety of different priming compositions, and for eliciting an antibody response.

Replication-deficient adenovirus derived from human serotype 5 has been developed as a live viral vector by Graham and colleagues (Graham & Prevec 1995 *Mol Biotechnol* 3:207-20; Bett et al. 1994 *PNAS USA* 91:8802-6). Adenoviruses are non-enveloped viruses containing a linear double-stranded DNA genome of around 36 kb. Recombinant viruses can be constructed by *in vitro* recombination between an adenovirus genome plasmid and a shuttle vector containing the gene of interest together with a strong eukaryotic promoter, in a permissive cell line which allows viral replication. High viral titres can be obtained from the permissive cell line, but the resulting viruses, although capable of infecting a wide range of cell types, do not replicate in any cells other than the permissive line, and are therefore a safe antigen delivery system. Recombinant adenoviruses have been shown to elicit protective immune responses against a number of

antigens including tick-borne encephalitis virus NS1 protein (Jacobs et al. 1992 *J Virol* 66:2086-95) and measles virus nucleoprotein (Fooks et al. 1995 *Virology* 210:456-65).

Remarkably, the experimental work described below demonstrates that use of embodiments of the present invention allows for recombinant adenovirus expressing an antigen to induce an immune response, optionally as a boosting composition primed by a DNA vaccine. The adenovirus was found to induce an immune a CD8+ T cell and antibody response after intramuscular immunization. In prime/boost vaccination regimes the recombinant virus, *e.g.*, adenovirus, is also envisioned as being able to prime an immune response that can be boosted by a different recombinant virus or recombinantly produced antigen.

Non-human primates immunized with recombinant virus, *e.g.*, adenovirus, optionally as a boosting composition following priming with plasmid DNA were protected against challenge. Both recombinant adenovirus and plasmid DNA are vaccines that are safe for use in humans. Advantageously, the inventors found that a vaccination regime using single dose immunization, optionally prime and boost immunization, can be employed, constituting a general immunization regime suitable for inducing an immune response, *e.g.*, in humans.

The present invention in various aspects and embodiments employs a recombinant virus, *e.g.*, adenovirus, encoding an antigen for inducing an immune response to the antigen, optionally for boosting an immune response primed by previous administration of the antigen or nucleic acid encoding the antigen.

A general aspect of the present invention provides for the use of a recombinant virus, *e.g.*, adenovirus, for inducing, optionally boosting an immune response to an antigen.

One aspect of the present invention provides a method of inducing, optionally boosting an immune response to an antigen in an individual, the method including provision in the individual of a recombinant virus, *e.g.*, adenovirus, including nucleic acid encoding the antigen operably linked to regulatory sequences for production of antigen in the individual by expression from the nucleic acid, whereby an immune response to the antigen is induced or an immune response to the antigen previously primed in the individual is boosted.

An immune response to an antigen may be primed by plasmid DNA immunization, by infection with an infectious agent, or by development of a recombinantly produced antigen.

A further aspect of the invention provides a method of inducing an immune response to an antigen in an individual, the method comprising administering to the individual a single dose of composition comprising the antigen or nucleic acid encoding the antigen or a priming composition comprising the antigen or nucleic acid encoding the antigen and then administering a boosting composition which comprises a recombinant virus, *e.g.*, adenovirus, including nucleic acid encoding the antigen operably linked to regulatory sequences for production of antigen in the individual by expression from the nucleic acid.

A further aspect provides for use of a recombinant virus, *e.g.*, adenovirus, as disclosed, in the manufacture of a medicament for administration to a mammal to induce, optionally to boost an immune response to an antigen. Such a medicament is optionally for administration in single dose form or following prior administration of a priming composition comprising the antigen or nucleic acid encoding the antigen.

The inducing, boosting, or priming composition may comprise any viral vector, including adenoviral, or other than adenoviral, such as a vaccinia virus vector such as a replication-deficient strain such as modified virus Ankara (MVA) (Mayr et al. 1978 *Zentralbl Bakteriol* **167**:375-90; Sutter and Moss 1992 *PNAS USA* **89**:10847-51; Sutter et al. 1994 *Vaccine* **12**:1032-40) or NYVAC (Tartaglia et al. 1992 *Virology* **118**:217-32), an avipox vector such as fowlpox or canarypox, *e.g.*, the strain known as ALVAC (Kanapox, Paoletti et al. 1994 *Dev Biol Stand* **82**:65-9), a herpes virus vector, a vesicular stomatitis virus vector, or an alphavirus vector.

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

In particular, prime and boost embodiments of the various aspects of the present invention, administration of a priming composition is followed by boosting with a boosting composition or first and second boosting compositions, the first and second boosting compositions being the same or different from one another. Still further boosting compositions may be employed without departing from the present invention. In one embodiment, a triple immunization regime employs DNA, then adenovirus as a first

boosting composition, and then MVA as a second boosting composition, optionally followed by a further (third) boosting composition or subsequent boosting administration of one or other or both of the same or different vectors. Another option is DNA then MVA then adenovirus, optionally followed by subsequent boosting administration of one or other
5 or both of the same or different vectors.

The antigen to be included in respective priming and boosting compositions (however many boosting compositions are employed) need not be identical, but should share epitopes. The antigen may correspond to a complete antigen in a target pathogen or cell, or a fragment thereof. Peptide epitopes or artificial strings of epitopes may be
10 employed, more efficiently cutting out unnecessary protein sequence in the antigen and encoding sequence in the vector or vectors. One or more additional epitopes may be included, for instance epitopes which are recognized by T helper cells, especially epitopes recognized in individuals of different HLA types.

Within the recombinant virus, *e.g.*, adenovirus, regulatory sequences for expression
15 of the encoded antigen will include a promoter. By "promoter" is meant a sequence of nucleotides from which transportation may be initiated of DNA operably linked downstream (*i.e.*, in the 3' direction on the sense strand of double-stranded DNA). "Operably linked" means joined as part of the same nucleic acid molecule, suitably positioned and oriented for transcription to be initiated from the promoter. DNA operably
20 linked to a promoter is "under transcriptional initiation regulation" of the promoter. Other regulatory sequences including terminator fragments, polyadenylation sequences, enhancer sequences, marker genes, internal ribosome entry site (IRES) and other sequences may be included as appropriate, in accordance with the knowledge and practice of the ordinary person skilled in the art: see, for example, *Molecular Cloning: a Laboratory Manual*, 2nd
25 edition, Sambrook et al. 1989 Cold Spring Harbor Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology*, Ausubel et al. eds., John Wiley & Sons, 1994.

30 Suitable promoters for use in aspects and embodiments of the present invention include the cytomegalovirus immediate early (CMV IE) promoter, with or without intron A, and any other promoter that is active in mammalian cells.

Adjuvants suitable for co-administration in accordance with the present invention should be ones that are potentially safe, well tolerated and effective in people including QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjuver, PG-026, GSK-1, GcMAF, B-aethine, MPC-026, Adjuvax, CpG ODN,
5 Betafectin, Alum, and MF59 (see Kim et al., 2000 *Vaccine* 18:597 and references therein).

Other contemplated adjuvants that may be administered include lectins, growth factors, cytokines and lymphokines such as alpha-interferon, gamma interferon, platelet derived growth factor (PDGF), granulocyte-colony stimulating factor (gCSF), granulocyte macrophage colony stimulating factor (gMCSF), tumor necrosis factor (TNF), epidermal
10 growth factor (EGF), IL-1, IL-2, IL-4, IL-6, IL-8, IL-10, and IL-12 or encoding nucleic acids therefore.

Administration of the boosting composition is generally weeks or months after administration of the priming composition, preferably about 2-3 weeks or 4 weeks, or 8 weeks, or 16 weeks, or 20 weeks, or 24 weeks, or 28 weeks, or 32 weeks.

15 Preferably, administration of single dose composition, boosting composition, or priming composition is intramuscular immunization.

Intramuscular administration of adenovirus vaccines or plasmid DNA may be achieved by using a needle to inject a suspension of the virus or plasmid DNA. An alternative is the use of a needleless injection device to administer a virus or plasmid DNA
20 suspension (using, e.g., Biojector™) or a freeze-dried powder containing the vaccine (e.g., in accordance with techniques and products of Powderject), providing for manufacturing individually prepared doses that do not need cold storage. This would be a great advantage for a vaccine that is needed in rural areas of Africa.

Adenovirus is a virus with an excellent safety record in human immunizations. The
25 generation of recombinant viruses can be accomplished simply, and they can be manufactured reproducibly in large quantities. Intramuscular administration of recombinant adenovirus is therefore highly suitable for prophylactic or therapeutic vaccination of humans against diseases which can be controlled by an immune response.

The individual may have a disease or disorder such that delivery of the antigen and
30 generation of an immune response to the antigen is of benefit or has a therapeutically beneficial effect.

Most likely, administration will have prophylactic aim to generate an immune response against a pathogen or disease before infection or development of symptoms.

Diseases and disorders that may be treated or prevented in accordance with the present invention include those in which an immune response may play a protective or therapeutic role.

Components to be administered in accordance with the present invention may be formulated in pharmaceutical compositions. These compositions may comprise a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may depend on the route of administration, *e.g.*, intravenous, cutaneous or subcutaneous, intramucosal (*e.g.*, gut), intranasal, intramuscular, or intraperitoneal routes.

As noted, administration is preferably intradermal, subcutaneous or intramuscular.

Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

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

A slow-release formulation may be employed.

Following production of adenoviral particles and optional formulation of such particles into compositions, the particles may be administered to an individual, particularly human or other primate.

Administration may be to another mammal, *e.g.*, rodent such as mouse, rat or hamster, guinea pig, rabbit, sheep, goat, pig, horse, cow, donkey, dog or cat.

Administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, *e.g.*, decisions on dosage

etc., is within the responsibility of general practitioners and other medical doctors, or in a veterinary context a veterinarian, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned
5 above can be found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed., 1980.

In one preferred regimen, DNA is administered (preferably intramuscularly) at a dose of 10 micrograms to 50 milligrams/injection, followed by adenovirus (preferably intramuscularly) at a dose of 5×10^7 - 1×10^{12} particles/injection.

10 The composition may, if desired, be presented in a kit, pack or dispenser, which may contain one or more unit dosage forms containing the active ingredient. The kit, for example, may comprise metal or plastic foil, such as a blister pack. The kit, pack, or dispenser may be accompanied by instructions for administration.

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

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

Specific Modifications of Ebola GP Optimize Vaccine Efficacy in Nonhuman Primates

20 To develop an optimal Ebola vaccine using rAd vectors, we first analyzed mutant forms of GP in which the transmembrane domain had been removed. Though we have previously reported that deletion of the mucin domain eliminates cytotoxicity (Yang, Z.-Y. et al. 2000 *Nat Med* 6:886-889), this deletion removes nearly 200 amino acids, eliminating many potential T- and B-cell epitopes. Previous data suggested that the *in vitro* cytopathic
25 effects of GP may be mediated at or near the cell surface and require transmembrane anchoring of the protein (Sullivan, N. J. et al. 2005 *J Virol* 79:547-553; Takada, A. et al. 2000 *Virology* 278:20-26; Chan, S. Y. et al. 2000 *J Gen Virol* 81:2155-2159). An alternative approach to the elimination of the GP-induced cytopathic effects was therefore explored by removal of the 26 amino acid putative transmembrane and cytoplasmic
30 domains.

Diminished immune protection of a mutant GP lacking a transmembrane anchor domain

GP protein was readily detected in the supernatants of cells transfected with the transmembrane-deleted vector Δ TM(Z), confirming its secretion, in contrast to supernatants from cells transfected with the wild type GP(Z) (Fig. 10A). Furthermore, synthesis of the two previously defined forms of GP, generated by post-translational processing (Volchkov, V. E. et al. 1995 *Virology* 214:421-430; Sanchez, A. et al. 1998 *J Virol* 72:6442-6447), was readily detected at comparable levels. Deletion of the transmembrane domain eliminated GP-induced cytopathicity in transfected 293 cells in contrast to wild type GP (Fig. 10B), but total Δ TM expression was equivalent to wild type protein levels (Fig. 10A). To determine whether the Δ TM mutant of the Zaire strain could protect against infectious Ebola challenge, cynomolgus macaques were immunized with rAd vectors encoding NP and either Δ TM(Z), or GP(Z). Immunization with GP(Z) + NP protected all animals vaccinated with either 10^{11} or 10^{12} adenoviral particles and challenged with 1000 pfu of the Zaire strain of Ebola virus 28 days later (Fig. 11A). In contrast, survival frequencies decreased in animals receiving the Δ TM(Z) vaccine. In the group vaccinated with 10^{12} adenoviral particles, protective immunity was decreased by 33% and at 10^{11} , by 66%, indicating a substantial decrease in efficacy in animals vaccinated with Δ TM + NP vs GP + NP ($p < 0.05$). In a separate experiment, 10^{11} particles of Δ TM alone failed to protect against infection. Analysis of cell-mediated immune responses showed that $CD4^+$ and $CD8^+$ T-cell responses were present in the majority of animals by 3 weeks post-immunization (Fig. 11B, left and middle panels, respectively) and did not correlate with the differences in survival: antigen specific cellular responses measured by intracellular cytokine (TNF- α) secretion were indistinguishable between GP(Z)- and Δ TM(Z)-vaccinated animals. Similarly, humoral immune responses measured by anti-Ebola GP ELISA IgG titers were comparable in all vaccinated animals (Fig. 11B, right panel). Neutralizing antibody titers were low, and were absent in some surviving animals. These results suggested that deletion of the GP transmembrane domain reduces vaccine efficacy, with no readily apparent correlates of protection.

Definition of minimal protective vaccine dose for protection with a rAd vaccine encoding GP and NP

In the previous experiment, a log decrease in dose of the GP(Z) + NP vaccine was protective, in comparison to previous studies using 10^{12} rAd particles. To establish the

lowest dose of adenoviral vectors that would afford protection against Ebola infection, a dose-response analysis was performed. Animals were immunized with rAd vectors encoding GP(Z) and NP at increasing doses from 10^9 to 10^{12} particles per animal. Survival was 100% in all groups receiving a dose of 10^{10} or greater, whereas challenge infection was uniformly lethal in the 10^9 dose group (Fig. 12A). Virus isolation by plaque assay on Vero cells was negative for all surviving animals. Pre-challenge CD4⁺ T-cell responses for TNF- α were unremarkable as reported previously for immunized cynomolgus macaques (Fig. 12B). CD8⁺ T-cell responses were similar across vaccine dose groups, except for the higher responder immunized at 10^{12} rAd particles. Antigen-specific IgG was also generated in immunized animals, and the levels were equivalent among animals in the groups that survived Ebola virus challenge (Fig. 12C, left). However, there was a difference in more than one log ($p=0.004$) in IgG levels between survivors immunized at 10^{10} and fatalities immunized at 10^9 rAd particles, suggesting that such levels may correlate with protection for this immunization regimen. Neutralizing antibody titers against GP did not differ significantly between survivors and fatalities (Fig. 12C, right). These results indicated that the threshold for immune protection lies at about $\sim 10^{10}$ rAd particles. Therefore, subsequent experiments were carried out using this dose to increase sensitivity to detect differences in antigenic strength between various immunogens.

Identification of GP point mutants with diminished *in vitro* cytotoxicity that confer effective immune protection

We sought to identify other mutants of GP that do not exhibit cytopathic effects yet retain native antigenic structures when expressed *in vitro*. Relatively conserved regions of GP were identified, and point mutations were systematically introduced. GP proteins bearing single amino acid changes were screened for decreased induction of cell rounding but wild type levels of expression and reactivity with conformation-dependent antibodies. Substitution of aspartic for glutamic acid at position 71 in Ebola GP from the Zaire or Sudan/Gulu subtypes (E71D(Z), E71D(S/G), respectively) abolished the cell rounding phenotype in transfected 293 cells but did not alter protein expression or reactivity with antibodies whose binding properties are sensitive to changes in protein conformation (Fig. 14).

The E71D mutants were evaluated for their ability to induce protective immunity alone or in combination with NP. When E71D from Zaire and Sudan-Gulu were combined with NP, survival of cynomolgus macaques immunized in these groups was diminished by

33% and 66%, respectively (Fig. 13A). In contrast, complete protection was achieved in animals immunized with E71D(Z) and E71D(S/G), as it was in animals receiving wild type GP(Z) plus NP. Ebola GP-specific responses in T-lymphocytes detected by intracellular staining of TNF- α did not show statistically significant differences in the CD4⁺ population
5 between different immunization groups (Fig. 13B, left panel). Similarly, individual differences in the CD8⁺ response did not correlate with survival, though there was a trend toward diminished survival in groups with lower antigen-specific CD8⁺ cellular responses (Fig. 13B, middle panel. Antigen-specific ELISA IgG was also stimulated in all immunized animals (Fig. 13B, right panel). The results of this experiment illustrate that NP
10 may not be necessary for protective immunity against Ebola infection and that it may diminish protection when combined with modified GP immunogens at the lower limits of protective vaccine doses.

Ebola virus outbreaks are associated with high lethality due to the absence of treatment options or a licensed vaccine. Both DNA priming and rAd vector boosting, as
15 well as rAd alone can confer protection to lethal challenge in an animal model that closely parallels human disease (Geisbert, T. W. et al. 2003 *Am J Pathol* **163**:2347-2370). The rAd vector vaccine conferred protection in an accelerated vaccine regimen in nonhuman primates (Sullivan, N. J. et al. 2003 *Nature* **424**:681-684). Although *in vitro* cytopathicity has been observed by over expression of Ebola GP, one of the vaccine components, we
20 have not seen toxicity in animals vaccinated by vectors expressing Ebola GP. However, because this hypothetical complication has been raised, we sought to modify GP to eliminate *in vitro* cytopathicity yet retain antigenic properties that are necessary for protective immunity. Here, the efficacies of different forms of GP were evaluated using doses at the threshold of protection in the accelerated vaccination model. We have
25 identified a vaccine with decreased *in vitro* cytopathicity that retained immunogenicity necessary to protect against Ebola infection.

We find that alternative forms of GP confer differential immune protection. Deletion of the GP transmembrane domain abolished cytopathic effects in transfected 293 cells, but the corresponding Δ TM(Z) vaccine was less efficacious than wild type GP(Z) in
30 protecting nonhuman primates against infection. Though cellular and humoral immune responses were indistinguishable between groups receiving the different immunogen forms, the inherent variability in quantitating the responses in outbred macaques may obscure our ability to identify immune responses responsible for higher survival. Alternatively,

ΔTM(Z) may differ from wild type GP(Z) in antigenic qualities that are not captured by measurements of total antigen-specific IgG or intracellular cytokine responses stimulated by a broad peptide pool. For example, the transmembrane-deleted protein is secreted and likely shows conformational differences from the membrane anchored protein. Subsequent
5 modifications of the glycoprotein to retain membrane attachment and a more native envelope structure yielded a mutant, E71D, with reduced *in vitro* cytopathicity. Recently, it has been suggested that this region of GP contributes to viral receptor binding (Manicassamy, B. et al. 2005 *J Virol* 79:4793-4805). It is noteworthy that the envelope glycoprotein cytopathicity of other viruses such as HIV is linked to receptor binding and
10 fusion (Cao, J. et al. 1996 *J Virol* 70:1340-1354), raising the possibility that Ebola GP shares similar properties.

Ongoing outbreaks of both Ebola and Marburg viruses illustrate the importance of developing a filovirus vaccine for human use. This report shows that protective immunity against Ebola infection is achieved in nonhuman primates by the generation of antigen-
15 specific immune responses to a single protein, GP, which has been modified to eliminate *in vitro* cytopathic effects. The accelerated vaccine strategy has since been repeated using vesicular stomatitis virus (VSV) vectors (Jones S. M. et al. 2005 *Nat Med*, published online June 5, 2005), validating the promise of vaccines for Ebola. However, there are concerns about the use of VSV as a human vaccine because it is replication-competent and
20 derives from a virus that is pathogenic in animals. In contrast, the rAd vector vaccine is non-replicating, can be manufactured to high yields, and safety data exist for this platform. Immunity follows a single injection with 10^{10} rAd particles, a dose that is two orders of magnitude lower than previously reported for this single modality vaccine. Such doses of rAd vectors have proven to be well-tolerated and immunogenic for other recombinant genes
25 *in vivo* and can be evaluated for the vectors reported here, alone or in DNA prime/rAd boost combinations. Immunization with 10^{10} rAd particles of E71D(Z) + E71D(S/G) was effective against infectious challenge with Ebola Zaire, and protection did not require NP. Elimination of NP from the vaccine and dose reductions to 10^{10} rAd particles do not diminish protection and simplify the vaccine for future development in human trials.

Example 1

Vector construction and transfections

E1/E3-deleted, replication-incompetent Ad5 vectors were generated in PER.C6® cells (Fallaux, F. J. et al. 1998 *Hum Gene Ther* 9:1909-1917) using a pBR322-based

adaptor plasmid pAdApt together with cosmid pWE.Ad.AfIII-rITRAE3 essentially as described elsewhere (Havenga, M. J. et al. 2001 *J Virol* 75:3335-3342). The adaptor plasmid contained the left portion of the Ad5 genome (nucleotides 1-454), followed by transcriptional control elements and the adaptor Ad5 DNA region (nucleotides 3511-6095 in Ad5). Ebola GP encoding genes were cloned into the expression cassette in the adaptor plasmids under transcriptional control of the human full-length immediate-early CMV promoter and the SV40 polyadenylation signal. Adenoviruses containing Ebola GP, GPΔTM, and point mutations were generated by cotransfection of linearized pAdApt-Ebola GP plasmids together with the linearized cosmid pWE.Ad.AfIII-rITRAE3 containing the right portion of the Ad5 genome to PER.C6[®] cells using Lipofectamine (Invitrogen). PER.C6[®] cells were cultured in DMEM supplemented with 10% fetal bovine serum (GIBCO) and incubated at 37°C under humidified atmosphere and 10% CO₂. Homologous recombination led to the generation of rAd5-Ebola GP viruses. Adenoviral vectors in crude lysates were plaque purified using limiting dilutions and agar overlays, and Ad vector clones were analyzed for presence and expression of the transgene. Positive clones were amplified for large-scale production using PER.C6[®] cells in 48 triple-layer 3x175 cm² flasks. Viruses were purified by standard two-step CsCl gradient ultracentrifugation and subsequently desalted and formulated by three consecutive dialysis steps into TRIS-Cl pH 8.0 containing 2.5% glycerol. Purified Ad vectors were stored as single use aliquots at -80°C. Virus particle (vp) titers were determined by anion-exchange high-performance liquid chromatography based on described procedures (Shabram, P. W. et al. 1997 *Hum Gene Ther* 8:453-465). Infectivity was assessed by TCID₅₀ using 911 cells. Ebola GP expression was assessed by infection of A549 cells followed by analysis of culture lysates on western blot. The identity of the purified vectors was confirmed by PCR. Expression vectors p1012, pGP and pΔTM and point mutants contain a CMV enhancer promoter that have been described previously (Sullivan, N.J. et al. 2000 *Nature* 408:605-609). The pΔTM contains a deletion from amino acid 651 to 676 and was created by digesting with BspMI/Klenow, and then fusing to TGA. The resulting plasmid also contained four extra amino acids at the C-terminus (MAAS). 293 human embryonal kidney cells were cultured in DMEM supplemented with 10% fetal bovine serum (GIBCO). Transfections to measure protein expression and cell rounding were performed in 293 cells with 2μg DNA per well of a 6-well plate using calcium phosphate (Invitrogen) according to the manufacturer's

instructions. Protein expression was evaluated by SDS-PAGE followed by Western blot with a GP-specific antibody kindly provided by A. Sanchez.

Animal study and safety

Cynomolgus macaques (*M. fascicularis*), 3-5 years old and weighing 2-3 kg, obtained from Covance, were used for immunization and challenge experiments. The monkeys, housed singly, were anesthetized with ketamine to obtain blood specimens and to administer vaccines. In conducting this research, the investigators adhered to the *Guide for the Care and Use of Laboratory Animals*, prepared by the Institute of Laboratory Animal Resources, National Research Council (National Academy Press, Washington, D.C., 1996). The facilities are fully accredited the Association for Assessment and Accreditation of Laboratory Animal Care International. They received regular enrichment according to the Guide for the Care and Use of Laboratory Animals (DHEW No. NIH 86-23). Before Ebola virus challenge and to the end of each experiment, the animals were maintained in the Maximum Containment Laboratory (BSL-4) and fed and checked daily.

15 Macaque immunization and challenge

Cynomolgus macaques were injected intramuscularly with a 1.0 ml equal mixture of immunogens at the doses indicated. Viral challenge was performed by inoculation of animals in the left or right caudal thigh with 0.5 ml of viral stock that contained a target dose of ~1000 PFU EBOV (Zaire species) at four weeks after the initial immunization. No adverse effects of the adenovirus vaccination were observed acutely. The Ebola virus used in this study was originally obtained from a fatally infected human from the former Zaire in 1995 (Jahrling, P. B. et al. 1996 *Arch Virol Suppl* 11:135-140). Collection of serum and blood for viral load and ELISA titers was performed as previously described (Sullivan, N.J. et al. 2000 *Nature* 408:605-609).

25 Flow cytometry and antibodies

Transfected cells were collected after incubation with PBS (3 mM EDTA) and incubated with control Ig or rabbit anti-sGP/GP serum (generously provided by Dr. A. Sanchez) for 30 minutes on ice. The cells were washed twice with ice-cold PBS containing 2.5% fetal bovine serum, incubated with FITC- or PE-conjugated secondary antibodies (Jackson ImmunoResearch Laboratories and Sigma, respectively) for 30 minutes on ice, followed by washing. Analysis was conducted using a Becton Dickinson 4-color Calibur flow cytometer and FlowJo analysis software (Tree Star, Inc).

ELISA

Nunc-Immuno Maxisorp plates (Nunc, Rochester, NY) were coated with Ebola GP from 293 cell supernatants and incubated at 4°C until use. All further incubations were carried out at room temperature. Plates were then washed six times with PBS containing
5 Tween 20. Test sera were diluted in PBS containing Tween 20 and 1% fetal calf serum and allowed to react with the Ag-coated wells for 60 minutes. After washing plates six times, goat anti-human IgG (H+L; Chemicon, Temecula, CA) conjugated to horseradish peroxidase was used as a detection antibody. Bound IgG was detected by Sigma Fast o-Phenylenediamine Dihydrochloride Tablet Sets (Sigma-Aldrich, St. Louis, MO) and the
10 optical density was determined. A panel of normal sera was run each time the assay was performed.

Neutralizing antibody analysis

Ebola GP(Z) pseudotyped lentiviral virions were produced as previously described (Yang, 2004 *J Virol* 78:5642-5650). Briefly, 293T cells were plated in 10-cm-diameter
15 tissue culture dishes and transfected the next day by calcium phosphate reagent (Invitrogen) with pCMVΔR8.2, pHR'CMV-Luc and CMV/R Ebola GP(Z) plasmid DNA. Cells were transfected overnight, washed, and replenished with fresh medium. Forty-eight hours later, supernatants containing pseudotyped virus were harvested, filtered through a 0.45-μm-pore-size syringe filter, and stored in aliquots at -80°C. Neutralization assays were performed on
20 HUVECs (Cambrex CC-2517) plated in a 24 well plate 1 day prior to infection. Virus stocks were incubated at 37°C for 1 hour in the presence of serum from immunized cynomolgus macaques. The culture media was removed from the cells and replaced with the virus/serum media in the presence of polybrene (Sigma-Aldrich, 107689) at a final concentration of 5 ug/mL. 72 hours post infection cells were lysed and assayed by
25 Luciferase Assay System (Promega, E1501/E1531). Luciferase activity was determined using a Veritas Microplate Luminometer from Turner Biosystems.

Intracellular cytokine analysis

Peripheral blood mononuclear cells (PBMC) were isolated from cynomolgus macaque whole blood samples by separation over Ficoll. Approximately 1×10^6 cells were
30 stimulated in 200 μl RPMI medium (GIBCO) for 6 hours at 37°C with anti-CD28 and -CD49d antibodies, brefeldin A, and either DMSO or a pool of 15-mer peptides spanning the Ebola GP Zaire (Mayinga strain) open reading frame. The peptides were 15-mers overlapping by 11 spanning the entire Ebola glycoprotein at a final concentration of 2

μg/ml. Cells were fixed and permeablized with FACS Lyse (Becton Dickinson) supplemented with Tween 20, and stained with a mixture of antibodies against lineage markers (CD3-PE, CD4-PerCP, CD8-FITC) and either TNF-APC. Samples were run on a FACS Calibur or FACS Aria and analyzed using the software FlowJo. Positive gating for lymphocytes using forward vs. side scatter was followed by CD3⁺/CD8⁻ and CD3⁺/CD4⁻ gating, and specific populations were further defined by anti-CD4 and anti-CD8 positivity, respectively. Cytokine positive cells were defined as a percentage within these individual lymphocyte subsets and at least 200,000 events were analyzed for each sample.

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

2005289439 11 Nov 2011

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A nucleic acid molecule comprising a polynucleotide encoding a modified filovirus glycoprotein (GP) having at least one amino acid change at amino acid position 71 or 102 in Ebola Zaire GP or in a corresponding position in the GP of other strains of filovirus, wherein said at least one amino acid change decreases *in vitro* cytotoxicity and retains immunogenicity when compared to *in vitro* cytotoxicity and immunogenicity of a wild type filovirus GP.
2. The nucleic acid molecule of Claim 1, wherein the amino acid change is E71D or G102A in Ebola Zaire GP.
3. The nucleic acid molecule of Claim 1 or 2, wherein the modified filovirus GP is encoded by the insert of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7 or SEQ ID NO:8, or sequence having at least 95% identity thereto.
4. The nucleic acid molecule according to any one of Claims 1 to 3, wherein the polynucleotide encoding the modified filovirus GP has a sequence taken from the insert of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8, or sequence having at least 95% identity thereto.
5. A modified filovirus GP encoded by the nucleic acid molecule according to any one of Claims 1 to 4.
6. A plasmid DNA comprising the nucleic acid molecule according to any one of Claims 1 to 4.
7. A recombinant virus comprising the nucleic acid molecule according to any one of Claims 1 to 4.
8. An adenovirus comprising the nucleic acid molecule according to any one of Claims 1 to 4.
9. A pharmaceutical composition comprising the nucleic acid molecule according to any one of Claims 1 to 4 or the modified filovirus GP of Claim 5 in a therapeutically effective dose.

2005289439 11 Nov 2011

10. A vaccine composition comprising the nucleic acid molecule according to any one of Claims 1 to 4 or the modified filovirus GP of Claim 5 in a prophylactically effective dose.

11. The composition of Claim 11 or 12 further comprising an adjuvant.

12. An isolated or recombinant antibody that is specifically reactive with the modified filovirus GP of Claim 5.

13. A method of boosting an immune response to an antigen in a primate, the method comprising provision in the primate of the modified filovirus GP of Claim 5 or the recombinant virus of Claim 7, whereby an immune response to the antigen previously primed in the primate is boosted.

14. A method of inducing an immune response to an antigen in a primate, the method comprising provision in the primate of the nucleic acid molecule according to any one of Claims 1 to 4 or the modified filovirus GP of Claim 5, whereby an immune response to the antigen in the primate is induced.

15. A method of inducing an immune response to an antigen in a primate, the method comprising provision in the primate of a priming composition comprising the nucleic acid molecule according to any one of Claims 1 to 4 or the modified filovirus GP of Claim 5 and then provision in the primate of a boosting composition comprising the antigen or a recombinant virus encoding the antigen.

16. The method according to any one of Claims 13 to 15, wherein the primate is a human.

17. A method of making the nucleic acid molecule according to any one of Claims 1 to 4, said method comprising preparing an adapter plasmid and achieving recombination with a virus genome to produce a recombinant virus composed of the nucleic acid molecule according to any one of Claims 1 to 4.

18. An isolated or recombinant nucleic acid molecule having SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8.

2005289439 11 Nov 2011

19. Use of the nucleic acid molecule according to any one of claims 1 to 4 or the modified filovirus GP according to claim 5 or the recombinant virus according to claim 7 or the adenovirus according to claim 8 in the preparation of a medicament for prevention or treatment of infection by an Ebola virus.

20. The nucleic acid molecule according to any one of claims 1 to 4 or 18 or the modified filovirus GP according to claim 5 or plasmid DNA according to claim 6 or the recombinant virus according to claim 7 or the adenovirus according to claim 8 or the composition according to any one of claims 9 to 11 or the isolated or recombinant antibody according to claim 12 or the method according to any one of claims 13 to 17 or the use according to claim 19 substantially as described herein with reference to the accompany drawings and/or examples.

DATED this ELEVENTH day of NOVEMBER 2011

**The Government of the United States of America, as represented by
the Secretary, Department of Health and Human Services**

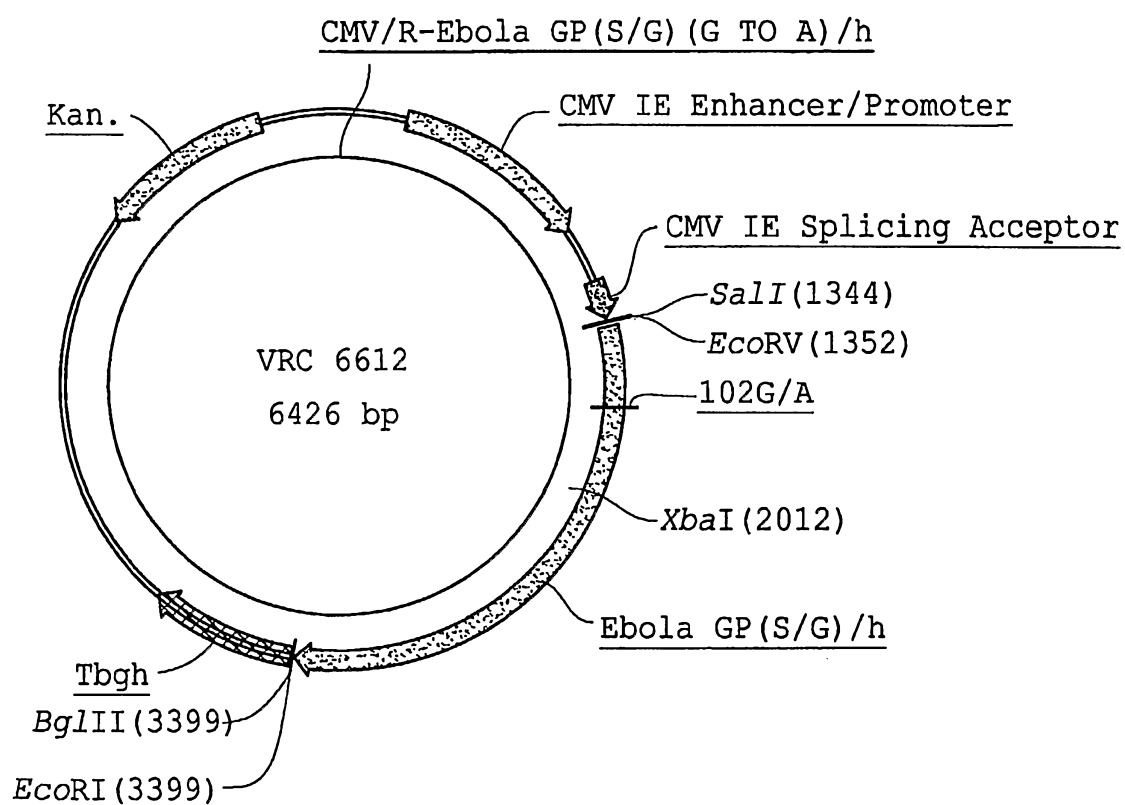
-and-

Crucell Holland B.V.

by the Patent Attorneys for the Applicant:

FB Rice

1/44

*FIG. 1A*

2/44

*FIG. 1B-1***VRC 6612 (SEQ ID NO: 1)**

1 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctccc gagacgggtca
 61 cagcttctgt gtaagcggat gccgggagca gacaagcccg tcaggggcgcg tcagcgggtg
 121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
 181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
 241 ctattggcca ttgcatacgt tctatccata tcataatag tacattata ttggtcatg
 301 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac
 361 ggggtcatta gtctatagcc catatatgga gtccgcgtt acataactta cggtaaatgg
 421 cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttc
 481 catagtaacg ccaataggga ctttcattg acgtcaatgg gtggagtatt tacggtaaac
 541 tgcccacttg gcagtacatc aagtgtatc tatgccaatg acgcccccta ttgacgtcaa
 601 tgacggtaaa tggccgcct ggcatatgc ccagtacatg acctatggg accttcctac
 661 ttggcagtac atctacgtat tagtcatgc tattaccatg gtgatgcgtt ttggcagta
 721 catcaatggg cgtggatagc ggttgactc acggggatti ccaagtctcc accccattga
 781 cgtcaatggg agttgtttt ggcacaaaa tcaacgggac ttccaaaat gtcgtaaca
 841 ctccgcccc ttgacgcaa tggcggtgag gcgtgtacgg tgggaggtct atataagcag
 901 agctcgttta gtaaccgtc agatcgctg gagacgccat ccacgtctt ttgacctca
 961 tagaagacac cgggaccgat ccagcctcca tcggctcgca tctctcttc acgccccgc
 1021 cgccctacct gaggccgcca tccacgcccg ttgagtcgcg ttctccgcc tccgcctgt
 1081 ggtgcctcct gaactgcgtc cgcgtctag gtaagtttaa agctcaggtc gagaccgggc
 1141 cttgtccgg cgtcccttg gagcctacct agactcagcc ggctctccac gcttgcctg
 1201 accctgcttg ctcaactcta gtaacgggtg gagggcagtg tagtctgagc agtactcgtt
 1261 gctccgcgc gcgccaccag acataatagc tgacagacta acagactgtt ccttccatg
 1321 ggtcttttct gcagtcaccg tcgtcgacga tatcgccgc atggaggggc tgagcctgt
 1381 gcagctgcc agggacaagt tcaggagag cagcttctc gtgtgggtga tcatcctgtt
 1441 ccagaaggcc ttgagcatgc cctgggctg ggtgaccaac agcaccttg aggtgaccga
 1501 gatcgaccag ctggtgtgca aggaccacct ggccagcacc gaccagctga agagcgtggg
 1561 cctgaacctg gagggcagcg gcgtgagcac cgacatccc agcgccacca agagggtggg
 1621 cttaggagc ggcgtgcctc ccaaggtggt gagctacgag gccgcccagt gggccgagaa
 1681 ctgtacaac ctgagatca agaagcccga cggcagcgag tgcctgcctc ctctctga
 1741 cggcgtgagg ggttcccca ggtgcaggtg cgtgcacaag gccaggggca cggccctg
 1801 cccggcgac tacgccttc acaaggacgg cgccttctc ctgtacgaca ggctggccag
 1861 caccgtgatc tacaggggcg tgaacttcgc cgaggggcgt atgccttc tgatcctggc
 1921 caagcccaag gagaccttc tgacagccc tccatcagg gaggcgtga actacaccga
 1981 gaaccaccag agtactacg ccaccagta tctagagtag gagatcgaga acttcggcgc
 2041 ccagcacagc accacctgt tcaagatga caacaacacc ttggtaggc tggacaggcc
 2101 ccacacctt cagttctgt tccagctga cgacaccatc cactgcacc agcagctgag
 2161 caacaccacc ggcaggtga tctggacct ggacgccaac atcaacgccc acatcgccga
 2221 gtgggccttc tgggagaaca agaagaacct gagcgagcag ctgaggggcg aggagctgag
 2281 cttagagcc ctgagcctga acgagaccga ggacgagcag gccgccagca gcaggatcac
 2341 caagggcagg atcagcgaca gggccaccag gaagtacagc gacctgggtc ccaagaacag
 2401 cccggcatg gtgcccctgc acatcccga gggcgagacc accctgccc gccagaacag
 2461 caccgagggc aggaggggtg gcgtgaacac ccaggagacc ataccgaga ccgcccacc
 2521 catcatcgcc accaacggca accacatgca gatcagcacc atcggeatca ggcccagcag
 2581 cagccagatc cccagcagca gcccaccac cgcctctagc cccaggccc agacccccc
 2641 caccacacc agcggaccca gcgtgatggc caccgaggag cccaccacc ctccggcag
 2701 cagccccgga cccaccacc aggccccctac cctgaccacc cctgagaaca tcaccaccg
 2761 cgtgaagacc gtgtgcccc aggagagcac cagcaacggc ctgatacca gcaccgtgac
 2821 cggcatcctg ggcagcctgg gcctgaggaa gaggagcagg aggcagacca acaccaaggc
 2881 caccggcaag tgcaaccca acctgcacta ctggaccgcc caggagcagc acaacggcg
 2941 cggcatcgcc tggattccct acttggccc cggcgccgag ggcatctaca ccgagggcct

SUBSTITUTE SHEET (RULE 26)

3/44

FIG. 1B-2

3001 gatgcacaac cagaacgccc tgggtgctgg cctgaggcag ctggccaacg agaccacca
 3061 ggccctgcag ctgttcctga gggccaccac cgagctgagg acctacacca tctgaacag
 3121 gaaggccatc gacttcctgc tgaggaggtg gggcggcacc tgcaggattc tgggcccga
 3181 ctgctgcatc gagcccacg actggaccaa gaacatcacc gacaagatca accagatcat
 3241 ccacgacttc atgcacaacc ctctgccaa ccaggacaac gacgacaact ggtggaccgg
 3301 ctggcggcag tggatacctg ccggcatcgg catcacgggc atcatcatcg ccacatcg
 3361 tctgtgtgc gtgtgcaagc tgctgtctg agaattcaga tctacagat ctgtgtgcc
 3421 tttagttgc cagccatctg ttgttgccc ctccccctg ccttcctga cctggaagg
 3481 tggcactccc actgtcctt cctaataaaa tgaggaaatt gcatcgatt gtctgagtag
 3541 gtgtcattct attctggggg gtgggggtgg gcaggacagc aagggggagg atgggaaga
 3601 caatagcagg catgctgggg atcggtggg ctctatgggt acccaggtgc tgaagaattg
 3661 acccggttcc tcttgggcca gaaagaagca ggcacatccc ctctctgtg acacaccctg
 3721 tccagcccc tggctcttag ttccagcccc actcatagga cactcatagc tcaggagggc
 3781 tccgcctca atcccaccg ctaaagtact tggagcgtc tctccctccc tcatagccc
 3841 accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag gctattaagt
 3901 gcaggaggag agaaaatgcc tccaacatgt gaggaagtaa tgagagaaat catagaattt
 3961 taaggccatg atttaaggcc atcatggcct taactctcg ctctctgct cactgactcg
 4021 ctgcgtctgg tctgtggct gcggcgagcg gtatcagctc actcaaggc ggtatacgg
 4081 ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag
 4141 gccaggaaac gtaaaaggc cgcgtgtctg gcgttttcc ataggctccg cccccctgac
 4201 gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga
 4261 taccaggcgt tccccctgg aagctccctc gtgcgtctc ctgtccgac cctgcccgtt
 4321 accggatacc tctccgctt tctccctcg ggaagcgtgg cgttttca tagctacgc
 4381 ttaggtatc tcagttcgt gttagtctgt cgtcccaagc tgggtgtgt gcacgaacc
 4441 ccgttcagc ccgaccgtg cgcctatcc gtaactatc gtcttgatc caaccggta
 4501 agacacgact tatcgccact ggcagcagcc actgtaaca ggattagcag agcgaggat
 4561 taggcgggtg ctacagagtt ctgaagtgg tggcctaact acggctacac tagaagaaca
 4621 gtaattgta tctgcgtct gctgaagcca gtaccttcg gaaaaagagt tggtagctt
 4681 tgatccggca acaaaaccac cgctggtagc ggtggtttt ttgttgcaa gcagcagatt
 4741 acgcgcagaa aaaaaggatc tcaagaagat ccttgatct ttctacggg gtctgacgt
 4801 cagtggaaac aaaactcagc ttaagggtt ttggtcatga gattatcaa aaggatctc
 4861 acctagatcc tttaaatga aaatgaagt tttaaatcaa tctaaagtat atatgagtaa
 4921 acttggctg acagttacca atgcttaac agtgaggcac ctatctcagc gatctgtcta
 4981 ttctgtcat ccatagttgc ctgactcggg gggggggggc gctgaggtct gcctcgtgaa
 5041 gaaggtgtg ctgactcata ccaggcctga atcgcccat catccagcca gaaagtgagg
 5101 gagccacggt tgatgagagc ttgtttag gtggaccagt tggtagttt gaactttgc
 5161 ttgccacgg aacggtctg gttgtcgga agatgcgta tctgacctt caactcagca
 5221 aaagttcgt ttattcaaca aagccgctg cccgtcaagt cagcgtaatg ctctgccagt
 5281 gttacaacca attaaccaat tctgattaga aaaactcatc gagcatcaaa tgaactgca
 5341 atttattcat atcaggatta tcaatccat attttgaaa aagccgttc tgtaatgaag
 5401 gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc
 5461 cgactcgtcc aacatcaata caacctatta atttccctc gtcaaaata aggttatcaa
 5521 gtgagaatc acctagagt acgactgaat ccggtgagaa tggcaaaagc ttatgcattt
 5581 ctctccagac ttgtcaaca ggccagccat tacgtctgic atcaaaatca ctgcaccaa
 5641 ccaaaccgtt attcattcgt gattgcgct gagcgagacg aaatcgcga tcgctgttaa
 5701 aaggacaatt acaaacagga atcgaaatga accggcgagc gaacactgcc agcgcatcaa
 5761 caatatttc acctgaatca ggatattctt ctatccctg gaatgctgt ttccggggga
 5821 tcgcagtgtg gattaaatc gcatcatcag gattacggat aaaatgctt atggtcgga
 5881 gaggcataaa ttccgtcagc cagtttagt tgaccatctc atctgaaca tcattggcaa

4/44

FIG. 1B-3

5941 cgctaccttt gccatgttc agaaacaact ctggcgcac gccgttccca tacaatcgat
6001 agattgtcgc acctgattgc ccgacattat cgcgagccca ttataccca tataaatcag
6061 catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgtga atatggctca
6121 taacaccctt tgtattactg ttatgtaag cagacagttt tattgttcat gatgatata
6181 tttatcttg tgcaatgtaa catcagagat ttgagacac aacgtggctt tcccccccc
6241 cccattattg aagcatttat caggggtatt gtctcatgag cggatacata ttgaatgta
6301 tttagaaaa taaacaata ggggttccgc gcacattcc ccgaaaagt ccacctgacg
6361 tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct
6421 ttcgtc

*FIG. 1B**FIG.1B-1**FIG.1B-2**FIG.1B-3*

5/44

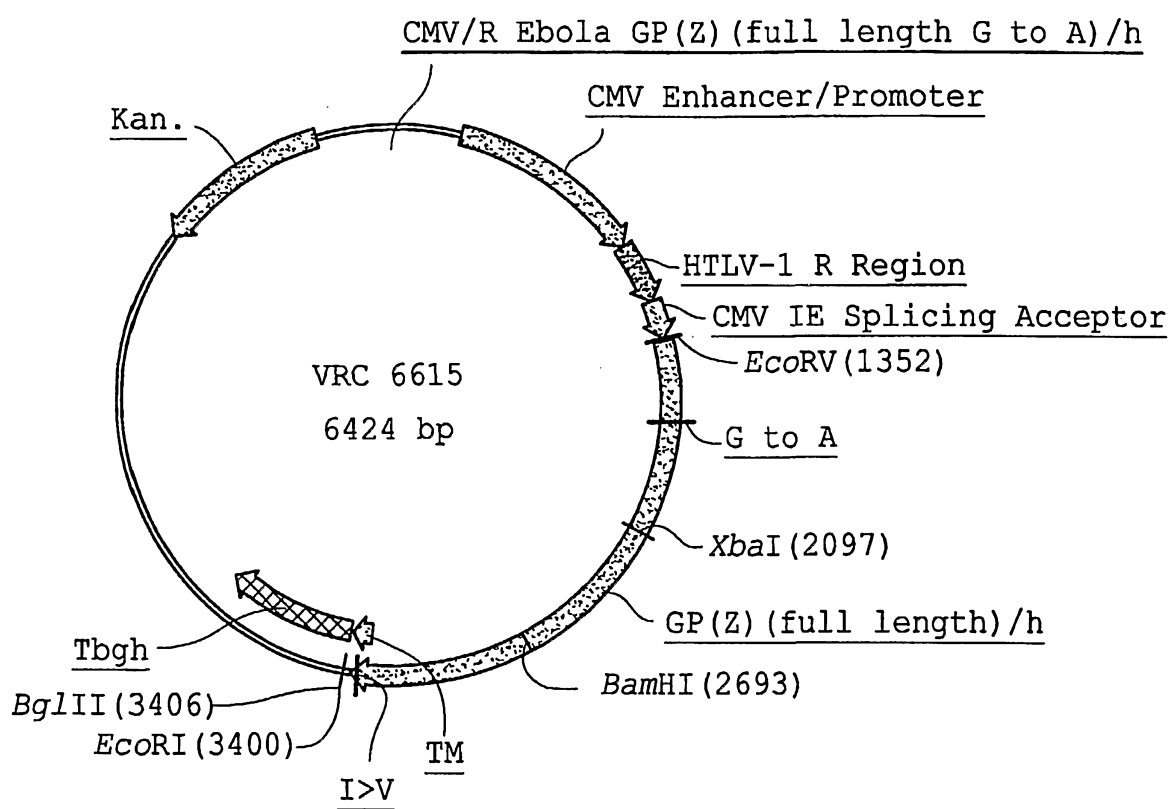


FIG. 2A

*FIG. 2B-1***VRC 6615 (SEQ ID NO: 2)**

1 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca
 61 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg
 121 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc
 181 accatatcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
 241 ctattggcca ttgcatacgt tctatccata tcataatatg tacattata ttggctcatg
 301 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac
 361 ggggtcatta gtcatagcc catatatgga gtccgcgtt acataactta cggtaaatgg
 421 cccgcctggc tgaccgcca acgaccccc ccatgacg tcaataatga cgtatgttc
 481 catagtaacg ccaataggga ctctcattg acgtcaatgg gtggagtatt tacggtaaac
 541 tgcccacttg gcagtacatc aagtgtatca tatccaagt acgccccctt ttgacgtcaa
 601 tgacggtaaa tggcccgctt ggcattatgc ccagtacatg acctatggg acttctctac
 661 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatcggtt ttggcagta
 721 catcaatggg cgtggatagc ggttgactc acggggattt ccaagtctcc accccattga
 781 cgtcaatggg agttgtttt ggcacaaaa tcaacgggac ttccaaaat gtcgtacaa
 841 ctccgcccc ttgacgcaa tggcggttag gcgtgtacgg tgggaggtct atataagcag
 901 agctcgttta gtgaaccgtc agatcgctg gagacgcat ccacgtgtt ttgacctca
 961 tagaagacac cgggaccgat ccagctcca tcggctcga tctctcttc acgcgccgc
 1021 cgccctacct gagggcgcca tccacgccc ttgagtcgcg tctgcgcc tcccgcctgt
 1081 ggtgcctcct gaactgcgtc cgcgtctag gtaagttaa agctcaggtc gagaccgggc
 1141 ctgttcggc gcgtccctt gagcttacct agatcagcc ggctctccac gcttgcctg
 1201 accctgctg ctcaactcta gtaacgggt gagggcagtg tagtctgagc agtactcgtt
 1261 gctgcgcgc gcgccaccag acataatagc tgacagacta acagactgtt ccttccatg
 1321 ggtctttct gcagtcaccg tcgtcgacga tatcgccgc atggcgctga ccggcatcct
 1381 gcagtgccc agggacaggt tcaagaggac cagctcttc ctgtgggtga tcatcctgtt
 1441 ccagaggacc ttacgatcc cctgggctg gatccacaac agcaccctgc aggtgagcga
 1501 cgtggacaag ctggtgtgca gggacaagct gagcagcacc aaccagctga ggagcgtggg
 1561 cctgaacctg gagggcaacg gcgtggccac cgactgccc agcgccacca agagggtggg
 1621 ctacaggagc ggcgtgcctc ccaagggtgt gaactacgag gccgcccagt gggccgagaa
 1681 ctgtacaac ctggagatca agaagcccga cggcagcgag tgctgccc cggcccctga
 1741 cggcatcagg ggttccccca ggtgcaggtc cgtgcacaag gtgagcgga cggccccctg
 1801 cgccggcgac ttgccttcc acaaggaggc gccttcttc ctgtacgaca ggctggccag
 1861 caccgtgatc tacaggggca ccacttcgc caggggcgtg gtggccttc tgatcctgcc
 1921 ccaggccaag aaggacttct tcagcagcca cctctgagg gagccctga acgccaccga
 1981 ggacccagc agcggtact acagcaccac catcaggtac caggccaccg gcttcggcac
 2041 caacgagacc gactacctgt tcgaggtgga caactgacc tacgtgcagc tggagtctag
 2101 attcaccctc cagttctgc tgcagctgaa cgagaccatc tacaccagcg gcaagaggag
 2161 caacaccacc ggcaagctga tctggaaggt gaaccccag atcgacacca ccatcggcga
 2221 gtgggccttc tgggagacca agaagaacct gaccaggaag atcaggagcg aggagctgag
 2281 ctacaccgtc gtgagcaacg gggccaagaa catcagcggc cagagcccc ccaggaccag
 2341 cagcgacccc ggcaccaaca ccaccaccga ggaccacaag atcatggcca gcgagaacag
 2401 cagcgccatg gtgcaggtgc acagccaggc cagggaggcc gccgtgagcc acctgaccac
 2461 cctggccacc atcagacca gccctcagtc ttaaccacc aagcccgcc ccgacaacag
 2521 caccacaac accctgtgt acaagctgga catcagcgag gccaccagg tggagcagca
 2581 ccacaggagg accgacaacg acagcaccgc cagcgacacc cttccgcca ccaccggc
 2641 cggccctccg aaggccgaga acaccaacac cagcaagagc accgacttc tggatccgc
 2701 caccaccacc agccctcaga accacagcga gaccgcccgc aacaacaaca ccaccacca
 2761 ggacaccggc gaggagagcg ccagcagcgg caagctgggc ctgatacca acaccatcgc
 2821 cggcgtggcc ggctgatca ccggcgccag gaggaccagg agggaggcca tctgaacgc
 2881 ccagcccaag tgcaaccca acctgcacta ctggaccacc caggacgagg gcgcccct
 2941 cggcctggcc tggatccct acttcggccc cgcccgag ggcatttaca tcgaggcct

7/44

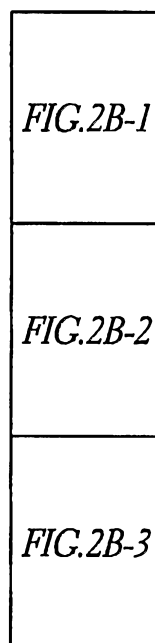
FIG. 2B-2

3001 gatgcacaac caggacggcc tgatctgcgg cctgaggcag ctggccaacg agaccacca
 3061 ggccctgcag ctgttcctga gggccaccac cgagctgagg acctcagca tctgaacag
 3121 gaaggccatc gacttcctgc tgcagagggtg gggcggcacc tggcacatcc tgggccccga
 3181 ctgctgcatc gagccccacg actggaccaa gaacatcacc gacaagatcg accagatcat
 3241 ccacgacttc gtggacaaga cctgccaga ccagggcgac aacgacaact ggtggaccgg
 3301 ctggcggcag tggatacctg ccggcatcgg cgtgaccggc gtggtgatcg ccgtgatcgc
 3361 tctgttctgc atctgcaagt tctgttctg aacacgtgga atcagatct gctgtgcctt
 3421 ctagtggcca gccatctgtt gttgccctt ccccggtgcc ttcttgacc ctggaagggtg
 3481 ccactccac tgcctttcc taataaaatg aggaaattgc atgcattgt ctgagtaggt
 3541 gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca
 3601 atagcaggca tctgggggat gcggtgggct ctatgggtac ccagggtgctg aagaattgac
 3661 ccggttcttc ctgggccaga aagaagcagg cacatccctt tctctgtgac acaccctgtc
 3721 cagccctcgt gttcttagtt ccagcccccac tcataggaca ctcatagctc aggagggtc
 3781 cgccttaaat cccaccgct aaagtacttg gagcgtgctc tccctccctc atcagccac
 3841 caaaccaaac ctgacctcca agagtgggaa gaaatfaag caagataggc tattaagtgc
 3901 agagggagag aaatgcctc caacatgta ggaagtaatg agagaaatca tagaattta
 3961 aggccatgat ttaaggccat catggccctt atcttcgct tctctgctca ctgactcgt
 4021 gcgctgggct gttcggctgc ggcgagcggc atcagctcac tcaaggcgg taatacgggt
 4081 atccacagaa tcaggggata acgcaggaaa gaacatgta gcaaaaggcc agcaaaaggc
 4141 caggaaccgt aaaaaggccg cgttgctggc gttttccat aggtccgcc cccctgacga
 4201 gcatcaciaa aatcgacgt caagtcagag gtggcgaaac ccgacaggac tataaagata
 4261 ccaggcgttt cccctggaa gctccctcgt gcgctctct gttccgacct tgcgcttac
 4321 cggatacctg tccgcttcc tccctcggg aagcgtggcg cttctcata gctcacgtg
 4381 taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaacccc
 4441 cgttcagccc gaccgtgctg cttatccgg taactatctt cttagtcca acccgtaag
 4501 acacgactta tgcacttg cagcagccac tggtaacagg attagcagag cgaggtagt
 4561 aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt
 4621 atttggtatc tgcgtctgc tgaagccagt tacctcggg aaaaagagttg gtatgtcttg
 4681 atccggcaaa caaacaccg ctggtagcgg tggtttttt gttgcaagc agcagattac
 4741 gcgcagaaaa aaaggatctc aagaagatcc ttgatctt tctacgggt ctgacgtca
 4801 gtggaacgaa aactcacgtt aaggatttt ggtcatgaga ttatcaaaaa ggtatctac
 4861 ctatgctctt taaataaa aatgaagttt taaatcaatc taaagtatat atgagtaaac
 4921 ttggtctgac agtaccat gctaatcag ttaggcacct atctcagcga tctgtctatt
 4981 tctgtcatcc atagtgcct gactcggggg gggggggcgc tgaagttctc ctgtgaaga
 5041 aggtgtgtct gactcatacc aggcctgaat cggccatca tccagccaga aagtgggga
 5101 gccaggttg atgagagctt tgttgtaggt ggaccagttg gtgatttga accttgcct
 5161 tgcacggaa cggctcgtg tgcgggaag atgcgtgac tgatcttca actcagcaaa
 5221 agttcgattt attcaacaaa gccgcgtcc cgtcaagta gcgtaatgt ctgcaggtg
 5281 tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat
 5341 ttattcatat caggattatc aataccatat tttgaaaaa gccgttctg taatgaagga
 5401 gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg
 5461 actcgtccaa catcaataca acctattaat ttccctcgt caaaaataag gttatcaagt
 5521 gagaaatcac catgagtac gactgaatcc ggtgagaatg gcaaaagctt atgcattct
 5581 ttccagactt gtcaacagg ccagccatta cgtcgtcat caaatcact cgcataacc
 5641 aaaccgttat tcatcgtga ttgcgctga gcgagacgaa atacgcgatc gctgttaaaa
 5701 ggacaattac aaacaggaat cgaatgcaac cggcgagga acactgccag cgcataaca
 5761 atattttcac ctgaatcagg atattcttct aatacctgga atgctgttt cccggggatc
 5821 gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgctgtat ggtcggaaga
 5881 ggcataaatt ccgtcagcca gttagtctg accatctcat ctgtaacatc attggcaacg
 5941 ctaccttgc catgttcag aaacaactct ggccatcgg gcttccata caatcgatg
 6001 attgtcgac ctgattgcc gacattatcg cgagccatt tataccata taaatcagca

8/44

FIG. 2B-3

6061 tccatgttg aattaatcg cggcctcgag caagacgttt cccgttgaat atggctcata
6121 acaccccttg tattactgtt tatgtaagca gacagtttta ttgtcatga tgatatattt
6181 ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttc cccccccccc
6241 cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt
6301 tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc
6361 taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt
6421 cgtc

FIG. 2B

9/44

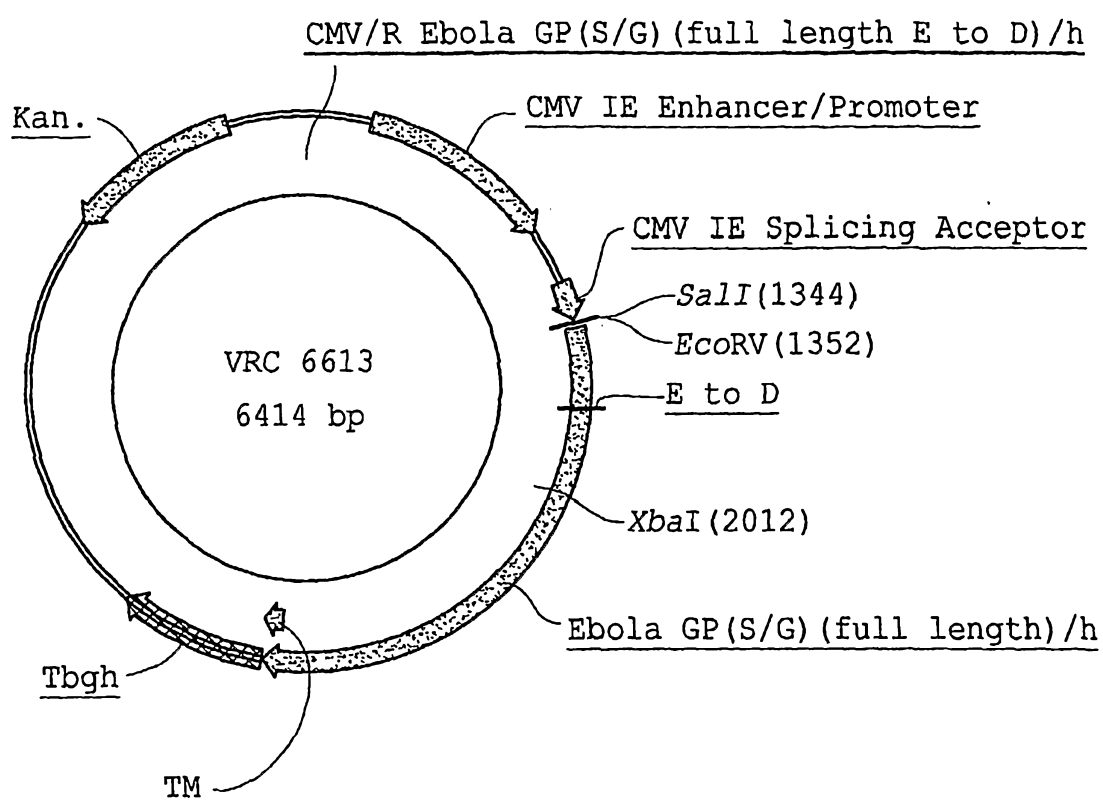
*FIG. 3A*

*FIG. 3B-1***VRC 6613 sequence (SEQ ID NO: 3)**

```

1  tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca
61  cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
121  ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc
181  accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
241  ctattggcca ttgcatagct tctatccata tcataatag tacattata ttggctcatg
301  tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac
361  ggggtcatta gtcatagcc catatatgga gtccgcgtt acataactta cggtaaatgg
421  cccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc
481  catagtaacg ccaataggga cttccattg acgtcaatgg gtggagtatt tacggtaaac
541  tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa
601  tgacggtaaa tggcccgctt ggcattatgc ccagtacatg acctatggg acttctac
661  ttggcagtac atctacgtat tagtcatgc tattaccatg gtgatgcggt ttggcagta
721  catcaatggg cgtggatagc ggttgactc acggggattt ccaagtctcc accccattga
781  cgtcaatggg agttgtttt ggcacaaaa tcaacgggac ttccaaaat gtcgtaacaa
841  ctccgcccc ttagcgaaa tggcggttag gcgtgtacgg tgggaggctc atataagcag
901  agtctgttta gtaaccgtc agatgcctg gagacgcat ccacgctgtt ttgacctca
961  tagaagacac cgggaccgat ccagctcca tcggtcga tctctctc acgcgccccg
1021  cgccctacct gaggccgcca tccacgccc ttgagtcgcg ttctgccgc tccgcctgt
1081  ggtgcctcct gaactgcgtc cgccgttag gtaagttaa agctcaggtc gagaccgggc
1141  cttgtccgg cgtcccttg gaccctac agactcagcc ggctctccac gcttgcctg
1201  accctgctt ctcaactta gtaacgggt gagggcagtg tagtctgagc agtactcgt
1261  gctgccgcgc gccaccag acataatagc tgacagacta acagactgtt ctttccatg
1321  ggtctttct gcagtcaccg tcgtcgacga tatcgccgc atggagggcc tgagcctgt
1381  gcagtcgcc agggacaagt tcaggaagag cagcttctc gtgtgggtga tcatcctgt
1441  ccagaaggcc ttcagcatgc cctgggctt ggtgaccaac agcacctgg aggtgaccga
1501  gatcgaccag ctggtgtgca aggaccact ggccagcacc gaccagctga agagcgtggg
1561  cctgaacctg gacggcagcg gcgtgagcac cgatcccc agcgccacca agagggtggg
1621  ctacaggagc ggcgtgcctc ccaagggtgt gagctacgag gccggcgagt gggccgagaa
1681  ctgtactaac ctggagatca agaagcccga cggcagcgag tgcctgcctc ctctcctga
1741  cggcgtgagg ggttcccca ggtgcagga cgtgcacaag gccaggggca cggccctg
1801  ccccgggcag tacgccttcc acaaggacgg cgccttctc ctgtacgaca ggctggccag
1861  caccgtgatc tacaggggcg tgaactcgc cgagggcgtg atgccttcc tgatctggc
1921  caagcccaag gagaccttcc tcagagccc tccatcagg gaggccgtga actacaccga
1981  gaaccaccag agtactacg ccaccagcta tctagatgac gagatcgaga acttcggcgc
2041  ccagcacagc accaccctgt tcaagatga caacaacacc ttcgtgaggc tggacaggcc
2101  ccacaccctc cagttctgt tccagctga cgacaccatc cactgcacc agcagctgag
2161  caacaccacc gccaggctga tctggacctt ggacgccaac atcaacgccg acatcgggca
2221  gtgggccttc tgggagaaca agaagaacct gagcgagcag ctgaggggcg aggagctgag
2281  ctgcaggccc ctgagcctga acgagaccga ggacgacgac gccgccagca gcaggatcac
2341  caagggcagg atcagcgaca gggccaccag gaagtacagc gacctggtgc ccaagaacag
2401  ccccggtatg gtgcccctgc acatccccga gggcgagacc accctgcca gccagaacag
2461  caccgagggc aggagggtgg gcgtgaacac ccaggagacc atcaccgaga ccgccccac
2521  catcatggc accaacggca accacatgca gatcagcacc atcggtacca ggcccagcag
2581  cagccagatc cccagcagca gcccaccac cggccctagc cccgaggccc agacccccac
2641  caccacacc agcggacca gcgtgatggc caccgaggag cccaccacc ctcccggcag
2701  cagccccgga cccaccacc agggccctac cctgaccacc cctgagaaca tcaccaccg
2761  ctgtaagacc gtgctgcccc aggagagcac cagcaacggc ctgatacca gcaccgtgac
2821  cggcatcctg ggcagcctgg gcctgaggaa gaggagcagg aggcagacca acaccaaggc
2881  caccggcaag tgcaaccca acctgcacta ctggaccgcc caggagcagc acaacggcg
2941  cggcatcgcc tggattccct acttggccc cggcgccgag ggcatctaca ccgaggccct

```

FIG. 3B-2

3001 gatgcacaac cagaacgccc tgggtgctgg cctgaggcag ctggccaacg agaccacca
 3061 ggccctgcag ctgttcctga gggccaccac cgagctgagg acctacacca tcctgaacag
 3121 gaaggccatc gacttcctgc tgaggaggtg gggcggcacc tgaggatc tgggccccga
 3181 ctgtctcatc gagccccacg actggacca gaacatcacc gacaagatca accagatcat
 3241 ccacgacttc atcgacaacc ctctgcccac ccaggacaac gacgacaact ggtggaccgg
 3301 ctggcggcag tggatactg ccggcatcgg catcaccggc atcatcatc ccatcatcgc
 3361 tctgtgtgc gtgtgcaagc tctgtgtcgt aacacgatct gctgtccct ctagtgcga
 3421 gccatctgtt gttgcccct ccccgctgcc ttcctgacc ctggaagggt ccactccac
 3481 tgccttttc taataaatg aggaattgc atgcattgt ctgagtaggt gtcattctat
 3541 tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca
 3601 tctgggggat gcggtgggct ctatgggtac ccagggtcgt aagaattgac ccggttctc
 3661 ctgggcccaga aagaagcagg cacatccct tctctgtgac acacctgtc cagccccgt
 3721 gttcttagtt ccagccccc tcataggaca ctcatagtc aggagggtc cgccttcaat
 3781 cccaccgct aaagtactg gagcggctc tccctccctc atcagccac caaaccaaac
 3841 ctgacctcca agagtgggaa gaaattaaag caagataggc tattaagtgc agagggagag
 3901 aaaatgctc caacatgtga ggaagtaag agagaaatca tagaattta aggccatgat
 3961 ttaaggccat catggccta atcttcgct tctcgtcga ctgactcgt gcgctcgtc
 4021 gttcggctgc ggcgagcgt atcagctcac tcaaaggcgg taatacgggt atccacagaa
 4081 tcaggggata acgaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaacct
 4141 aaaaaggccg cgttgcctgc gttttccat aggtccgccc cccctgacga gcatcaciaa
 4201 aatcgacgt caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt
 4261 cccctggaa gctccctcgt gcgctcctt gttccgccc tggcgcttac cggatacctg
 4321 tccgctttc tccctcggg aagcgtggcg ctttctcata gtcacgctg taggtatctc
 4381 agttcgggtg aggtcgtcgt ctccaagctg ggctgtgtgc acgaacccc cgttcagccc
 4441 gaccgctcgc ccttatccgg taactatcgt ctgagtcca acccggttag acacgactta
 4501 tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtagt aggcggtgct
 4561 acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt atttggtatc
 4621 tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtatgctctg atccggcaaa
 4681 caaaccaccg ctggtagcgg tggttttt gttgcaagc agcagattac gcgcagaaaa
 4741 aaaggatctc aagaagatcc ttgatctt tctacggggt ctgacgctca gtggaacgaa
 4801 aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac ctatgacctt
 4861 taaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac
 4921 agttaccaat gcttaactag tgggcacct atctcagcga tctgtctatt tcttcatcc
 4981 atagttgctt gactcggggg gggggggcgc tgaggctcgc ctctgtaaga aggtgtgtct
 5041 gactcatacc aggcctgaat gcgccatca tccagccaga aagttaggga gccacggttg
 5101 atgagagctt tgtgttaggt ggaccagttg gtgatttga actttgctt tggcacggaa
 5161 cgtctctcgt tctcgggaag atgcgtgac tgatcttca actcagcaaa agttcgattt
 5221 attcaacaaa gccgctgcc cgtcaagtca gcgtaatgct ctgccagtgt tacaaccaat
 5281 taaccaatc tgattagaaa aactcatcga gcatcaaatg aaactgcaat ttatcatat
 5341 caggattatc aataccatat tttgaaaaa gccgtttctg taatgaagga gaaaactcac
 5401 cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtccaa
 5461 catcaataca acctattaat tcccctcgt caaaaataag gttatcaagt gagaatcac
 5521 catgagtac gactgaatcc ggtgagaatg gcaaaagctt atgcattct tccagactt
 5581 gtcaacagg ccagccatta cgctcgtcat caaaactact cgcatacaac aaaccgttat
 5641 tcatcgtga ttgcgctga gcgagacgaa atacgcgac gctgttaaaa ggacaattac
 5701 aaacaggaat cgaatgcaac cggcgagga aactgcccag cgcatacaaa atattttcac
 5761 ctgaatcagg atattctct aatacctgga atgctgttt cccgggggac gcagtgggtga
 5821 gtaacctgc atcatcagga gtacggataa aatgcttgat ggtcgggaaga ggcataaatt
 5881 ccgtcagcca gtttagctg accatcicat ctgtaacatc attggcaacg ctacctttgc
 5941 catgtttcag aaacaactct ggcgcacgg gcttccata caatcgaatg attgtcgac
 6001 ctgattgccc gacattatc cgagcccat tataccata taaatcagca tccatgttg

12/44

FIG. 3B-3

6061 aatttaatcg cggcctcgag caagacgttt cccgttgaat atggctcata acacccttg
6121 tattactgtt tatgtaagca gacagtitta ttgttcata tgatatatt ttatctgtg
6181 caatgtaaca tcagagattt tgagacacaa cgtggcttcc ccccccccc cattattgaa
6241 gcatttatca gggttattgt ctcagagcg gatacatatt tgaatgtatt tagaaaaata
6301 aacaaatagg ggttcgcgc acatttccc gaaaagtgc acctgacgtc taagaaacca
6361 ttattatcat gacattaac tataaaaaata ggcgtatcac gaggcccttt cgtc

*FIG. 3B**FIG.3B-1**FIG.3B-2**FIG.3B-3*

13/44

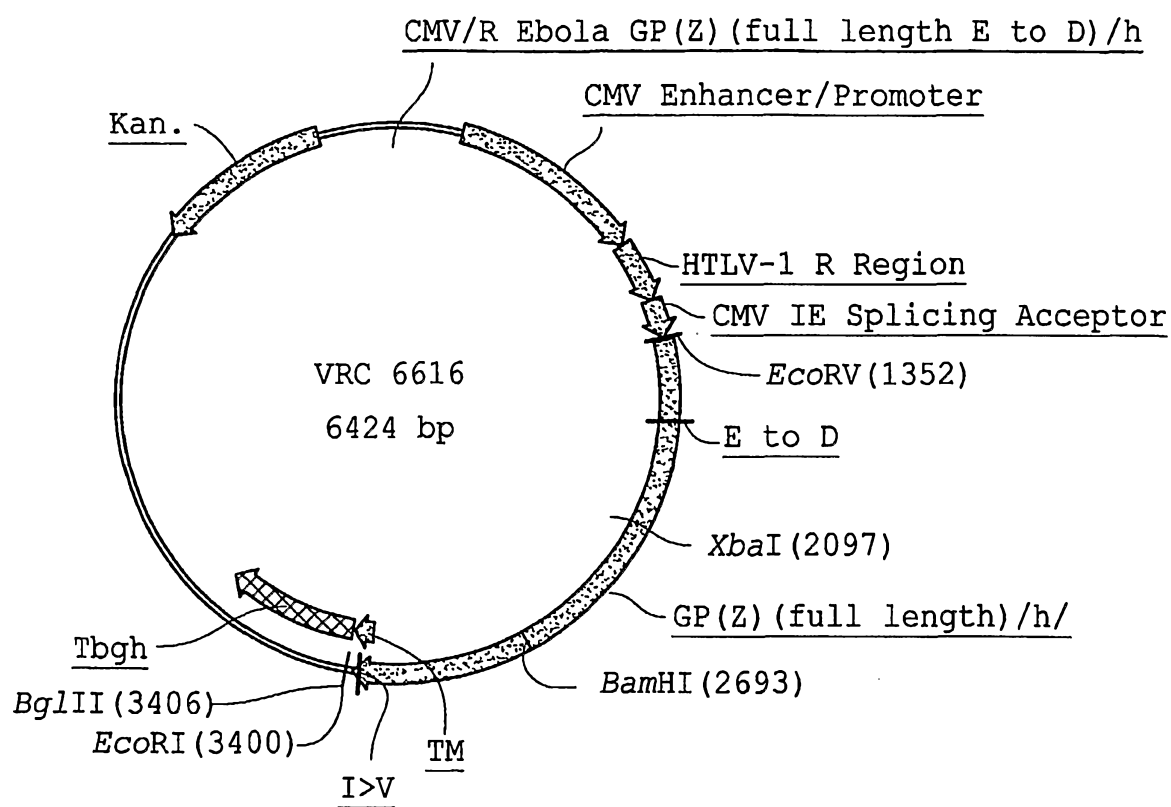
*FIG. 4A*

*FIG. 4B-1***VRC 6616 (SEQ ID NO: 4)**

1 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca
 61 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg
 121 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc
 181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
 241 ctattggcca ttgcatacgt tgtatccata tcataatag tacatttata ttggctcatg
 301 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac
 361 ggggtcatta gtcatagcc catatatgga gtccgcgtt acataactta cggtaaatgg
 421 ccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc
 481 catagtaacg ccaataggga ctttcattg acgtcaatgg gtggagtatt tacggtaaac
 541 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta ttgacgtcaa
 601 tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttctac
 661 ttggcagtac atctacgtat tagtcatgc tattaccatg gtgatgcgtt ttggcagta
 721 catcaatggg cgtggatagc gtttgactc acggggattt ccaagtctcc accccattga
 781 cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tticcaaaat gtcgtaacaa
 841 ctccgcccc ttgacgcaa tgggcggtag gcgtgtacgg tgggaggctt atataagcag
 901 agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctca
 961 tagaagacac cgggaccgat ccagcctcca tcgctcgca tctctcttc acgcgccccg
 1021 cgccctacct gaggccgcca tccacgcccg ttgagtcgag ttctgcccgc tcccgcctgt
 1081 ggtgcctcct gaactgcgtc cgccgtctag gtaagttaa agctcaggtc gagaccgggc
 1141 ctttgccg cgctcccttg gagcctacct agactcagcc ggctctccac gctttgcctg
 1201 accctgcttg ctcaactcta gttaacggtg gagggcagtg tagtctgagc agtactcgtt
 1261 gctgccgcgc gcgccaccag acataatagc tgacagacta acagactgtt ctttccatg
 1321 ggtctttct gcagtcaccg tcgtcgacga tatcgccgc atgggcgtga ccggcatcct
 1381 gcagctgccc agggacaggt tcaagaggac cagcttctc ctgtgggtga tcatctcgtt
 1441 ccagaggacc ttacagatcc cctgggcgt gatccacaac agcaccctgc aggtgagcga
 1501 cgtggacaag ctggtgtgca gggacaagct gagcagcacc aaccagctga ggagcgtggg
 1561 cctgaacctg gacggcaacg gcgtggccac cgacgtgccc agcggccacca agagggtggg
 1621 cttcaggagc ggcgtgccic ccaaggtggt gaactacgag gccggcgagt gggccgagaa
 1681 ctgtacaac ctggagatca agaagcccga cggcagcgag tgcctgcccg ccgcccctga
 1741 cggcatcagg ggcttcccc ggtgcaggta cgtgcacaag gtgagcggca ccggcccctg
 1801 cgccggcgac ttgccttcc acaaggaggc cgccctctc ctgtacgaca ggctggccag
 1861 caccgtgatc tacaggggca ccaccttcgc cgagggcgtg gtggccttcc tgatcctgcc
 1921 ccaggccaag aaggacttct tcagcagcca cctctgagg gagcccgtga acgcccacca
 1981 ggacccagc agcggtact acagcaccac catcaggtac caggccaccg gcttcggac
 2041 caacgagacc gattacctgt tcgaggtgga caacctgacc tacgtgcagc tggagtctag
 2101 attcaccctt cagttcctgc tgcagctgaa cgagaccatc tacaccagcg gcaagaggag
 2161 caacaccacc ggcaagctga tctggaaggt gaaccccgag atcgacacca ccatcgcgga
 2221 gtgggccttc tgggagacca agaagaacct gaccaggaag atcaggagcg aggagctgag
 2281 cttcaccgtc gtgagcaacg gggccaagaa catcagcggc cagagccccg ccaggaccag
 2341 cagcgacccc ggcaaccaaca ccaccaccga ggaccacaag atcatggcca gcgagaacag
 2401 cagcgccatg gtgcaggtgc acagccaggc cagggaggcc gccgtgagcc acctgaccac
 2461 cctggccacc atcagcacca gccctcagtc tttaaccacc aagccccgcc ccgacaacag
 2521 caccacaac acccctgtgt acaagctgga catcagcgag gccaccagg tggagcagca
 2581 ccacaggagg accgacaacg acagcaccgc cagcgacacc ccttcggcca ccaccggcg
 2641 cgccctccg aaggccgaga acaccaacac cagcaagagc accgacttic tggatccgc
 2701 caccaccacc agccctcaga accacagcga gaccgcccgc aacaacaaca cccaccacca
 2761 ggacaccggc gaggagagcg ccagcagcgg caagctgggc ctgatcaca acaccatgc
 2821 cggcgtggcc ggctgatca ccggcggcag gaggaccagg agggaggcca tegtgaacgc
 2881 ccagcccaag tgaaccca acctgacta ctggaccacc caggacgagg gcgcccat
 2941 cgccctggcc tggattcctt acttcggccc cgcccgag ggcactaca tcgaggcct

FIG. 4B-2

3001 gatgcacaac caggacggcc tgatctcggg cctgaggcag ctggccaacg agaccaccca
 3061 ggccctgcag ctgttctcga gggccaccac cgagctgagg accttcagca tctgaacag
 3121 gaaggccatc gacttctcgc tgcagagggt gggcggcacc tgccacatcc tgggccccga
 3181 ctgctgcacg gagccccacg actggaccaa gaacatcacc gacaagatcg accagatcat
 3241 ccacgacttc gtggacaaga ccctgccaga ccagggcgac aacgacaact ggtggaccgg
 3301 ctggcggcag tggatactg ccggcatcgg cgtgaccggc gtggtgatcg ccgtgatcgc
 3361 tctgttctgc atctgaagt tctgttctg aacacgtgga atcagatct gctgtgcctt
 3421 ctagtigcca gccatctgtt gttgcccct ccccgctgcc ttcctgacc ctggaagggt
 3481 ccactccac tctccttcc taataaatg aggaattgc atcgattgt ctgagtaggt
 3541 gtattctat tctggggggg ggggtggggc aggacagcaa gggggaggat tgggaagaca
 3601 atcagaggca tctgggggat gcgggtgggt ctatgggtac ccagggtgctg aagaattgac
 3661 ccggttctc ctgggccaga aagaagcagg cacatcccct tctctgtgac acaccctgc
 3721 cagccccctg gttcttagt ccagccccc tcataggaca tcatagctc aggagggtc
 3781 cgccttcaat ccaccctgct aaagtactg gagcggctc tccctccctc atcagccac
 3841 caaaccaac ctgacctcca agagtggga gaaattaaag caagataggc taftaagtgc
 3901 agaggagag aaaatgcctc caacatgtga ggaagtaatg agagaaatca tagaattta
 3961 aggccatgat taaaggcat catggccta atctccgct tctcgcica ctgactcgt
 4021 gcgtcgggtc gttcggctgc ggcgagcgg atcagctcac tcaaggcgg taatacgggt
 4081 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc
 4141 caggaaacct aaaaaggccg cgttgcggc gttttccat aggtccgcc cccctgacga
 4201 ccacacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata
 4261 ccaggcggtt cccctggaa gctccctcgt gcgtctcct gtccgacct tgcgcttac
 4321 cggatacctg tccgcttcc tccctcggg aagcgtggcg ctttccata gctcacgctg
 4381 taggtatctc agtctggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaacccc
 4441 cgttcagccc gaccgctcgc cttatccg taactatct cttagtcca acccggttag
 4501 acacgactta tgcgacttg cagcagccac tggtaacagg attagcagag cgaggatgt
 4561 aggcgggtgt acagagtct tgaagtggg gcctaactac ggctacata gaagaacata
 4621 atttggatc tgcgtctgc tgaagccagt tacctcggg aaaagagtgt gtagctctg
 4681 atccggcaaa caaacaccg ctggtagcgg tggttttt gtttcaagc agcagattac
 4741 gcgcagaaaa aaaggatctc aagaagatcc ttgatctt tctacgggt ctgacgctca
 4801 gtggaacgaa aactcacgtt aagggtattt ggtcatgaga ttatcaaaa ggatctcac
 4861 ctatgacctt taaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaac
 4921 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctacgca tctgtctatt
 4981 tcttcatcc atagtgcct gactcgggg gggggggcgc tgaggtctgc ctctgaaga
 5041 aggtgtgtct gactcatacc aggcctgaat cggccatca tccagccaga aagtgaggga
 5101 gccacgggtg atgagagctt tgtgtagggt ggaccagtgt gtagtttga actttgctt
 5161 tggcacggaa cgtctcgtg tctcgggaag atgcgtgac tgatcttca actcagcaaa
 5221 agttcgattt attcaacaaa gccgcccgtc cgtcaagta cgtaatgct ctgccagtgt
 5281 tacaaccaat taaccaatc tgaatagaaa aactcatcga gcatcaaatg aaactgcaat
 5341 ttattcatat caggattatc aataccatat tttgaaaaa gccgtttctg taatgaagga
 5401 gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg
 5461 actcgtccaa catcaatata acctattaat tcccctcgt caaaaataag gttatcaagt
 5521 gagaactcac catgagtac gactgaatcc ggtgagaatg gcaaaagctt atgatttct
 5581 ttccagactt gttcaacagg ccagccatta cgtcgtcat caaatcact cgcataacc
 5641 aaaccgttat tcttctgta ttgcgcctga gcgagacgaa atacgcgac gctgttaaaa
 5701 ggacaattac aaacaggaat cgaatgcaac cggcgaggga aactgcccag cgcataca
 5761 atatttcac ctgaatcagg atattcttct aatacctgga atgctgttt cccggggatc
 5821 gcagtgtgta gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcggga
 5881 ggataaatt ccgtcagcca gtttagtctg accatctcat ctgaacatc attggcaacg
 5941 ctaccttgc catgttccag aaacaactct ggcgcatcgg gcttccata caatcgatag
 6001 attgtcgac ctgattgccc gacattatcg cgagccatt tataccata taaatcagca

16/44

FIG. 4B-3

6061 tccatgttg aattaatcg cggcctcgag caagacgtt cccgttgaat atggctcata
6121 acacccttg tattactgtt tatgtaagca gacagttta ttgtcatga tgatatatt
6181 ttatctgtg caatgtaaca tcagagattt tgagacacaa cgtggcttcc ccccccccc
6241 cattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt
6301 tagaaaaata aacaaatagg ggtccgcgc acatttccc gaaaagtgcc acctgacgtc
6361 taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gaggcccttt
6421 cgtc

*FIG. 4B**FIG. 4B-1**FIG. 4B-2**FIG. 4B-3*

17/44

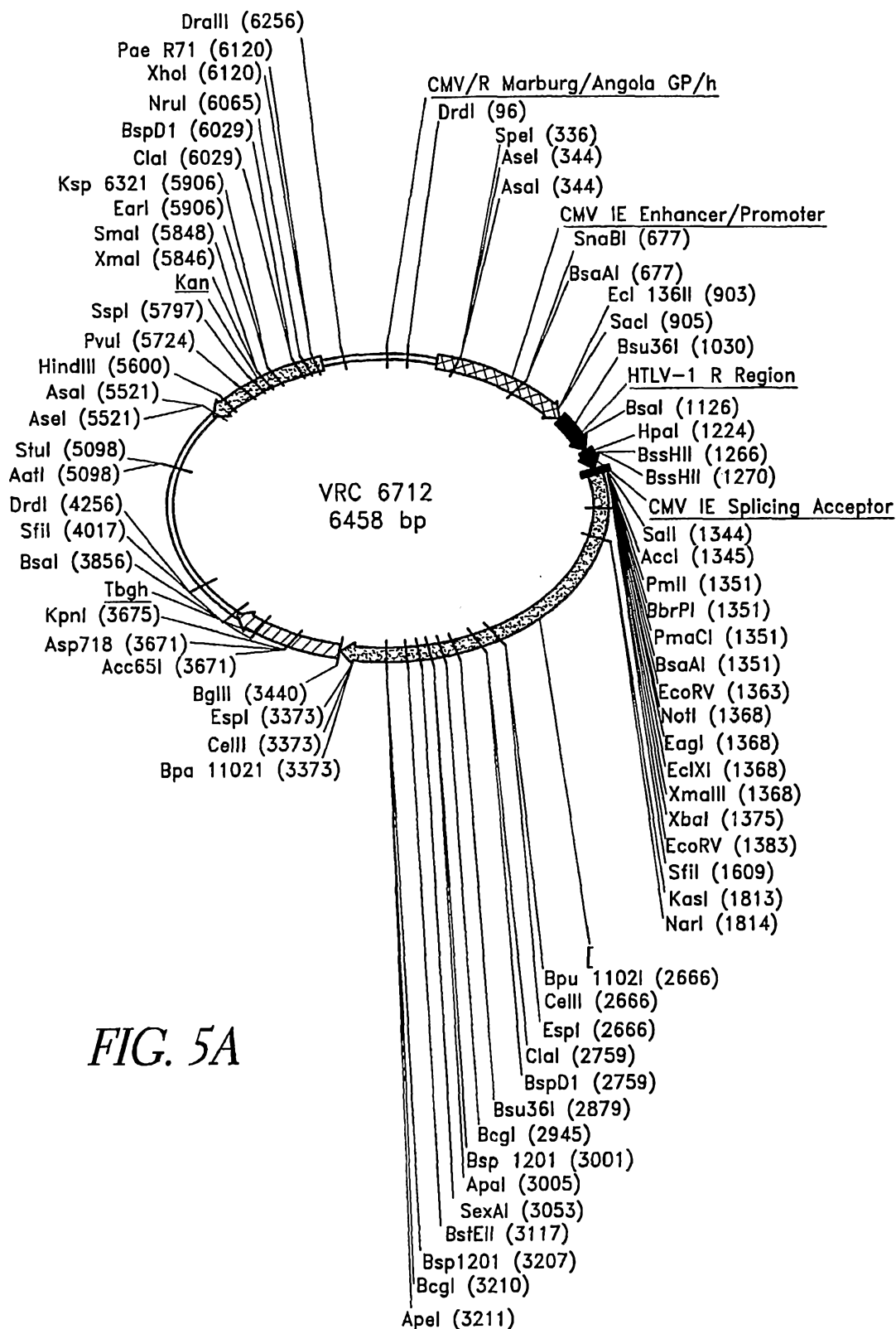


FIG. 5A

FIG. 5B-1

VRC 6712 (SEQ ID NO: 5)

1 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca
 61 cagcttgctt gtaagcggat gccgggagca gacaagcccc tcaggggcgcg tcagcgggtg
 121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc
 181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
 241 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg
 301 tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac
 361 ggggtcatta gticatatgcc catatatgga gticcgcgtt acataactta cggtaaatgg
 421 cccgcctggc tgaccgcca acgaccccc cccattgacg tcaataatga cgtatgttcc
 481 catagtaacg ccaataggga ctcttcattg acgtcaatgg gtggagtatt tacggtaaac
 541 tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa
 601 tgacggtaaa tgccccgcct ggcatatgc ccagtacatg accctatggg actttcctac
 661 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt ttggcagta
 721 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga
 781 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac ttccaaaat gtcgtaaaca
 841 ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggctc atataagcag
 901 agctcgttta gtgaaccgtc agatcgctg gagacgcat ccacgctgtt ttgacctca
 961 tagaagacac cgggaccgat ccagcctcca tcggctcgca tctctcttc acgcgccgc
 1021 cggcctacct gaggcgcca tccacgccg ttgagtcgcg ttctgccgc tcccgcctgt
 1081 ggtgcctcct gaactgcgtc cgccgtctag gtaagttaa agctcagtc gagaccgggc
 1141 cttgtccgg cgtcccttg gagcctacct agactcagc gctctccac gcttgcctg
 1201 accctgctg ctcaactcta gttaacgtg gagggcagt tagtctgagc agtactcgtt
 1261 gctgccgcgc gcgccaccag acataatagc tgacagacta acagactgtt cctttcatg
 1321 ggtcttttt gcagtcaccg tcgtcgacac gtgtgatcag atatcgccgc cgtctagag
 1381 atatcgccgc catgaagacc acctgcctgc tgatcagcct gatcctgac cagggcgtga
 1441 agacctgcc catcctggag atcgccagca acatccagcc ccagaacgtg gacagcgtgt
 1501 gcagcggcac cctgcagaag accgaggacg tgcacctgat gggcttacc ctgagcggcc
 1561 agaaggtggc cgacagccct ctggaggcca gcaagagggt ggcttcagg gccggcgtgc
 1621 cccccaagaa cgtggagtac accgaggcgc aggaggccaa gacctgtac aacatcagc
 1681 tgaccgccc cagcgccaag agcctgctgc tggacctcc caccaacatc agggactacc
 1741 ctaagtgaac gaccatccac cacatccagg gccagaaccc tcacgcccag ggcatcgccc
 1801 tgcacctgtg gggcgcttc ttctgtacg acaggatgc cagcaccacc atgtacagag
 1861 gaaaagtgtt cacagaggga aacatcgctg ctatgatcgt gaacaagacc gtgcataaga
 1921 tgatcttcag cagacaggga caggatata gacatatgaa cctgacatcc acaacaagt
 1981 actggacaag cagcaacgga acacagacaa acgatacagg atgtttgga acactgcagg
 2041 aatacaactc caccaagaac cagacatgtg ccctagcaa gaagcctctg cctctgcta
 2101 cagctcatcc tgaagtgaag ctgacatcca caagcacaga tgccacaag ctgaacacaa
 2161 cagatcctaa tagcgacgac gaggatctga caacaagcgg atccggatcc ggagaacagg
 2221 aaccttatac aacaagcgac gctgtctaaa aacagggact gtcctccaca atgcctccta
 2281 cacctagccc tcagcctagc acacctcagc agggaggcaa caacacaaac cattcccagg
 2341 gagtgggtgac agaacctgga aagacaaaca caacagccca gcctagcatg cctctcata
 2401 acacaacaac aatcagcaca aacaacacct ccaagcaca tctgagcaca cctagcgtgc
 2461 ctattcagaa tgccaccaac tacaacacac agtcacagc cctgaaaac gaacagacct
 2521 ccgccccctc caaacaacc ctgtgccta cagaaaacc tacaacagcc aagagcaca
 2581 acagacaaaa gagccctaca acaacagtgc ctaacacaa aaacaagtat agcacaagcc
 2641 ctagccctac acctattcc acagctcagc atctgggtga tttagaaga aagagaaca
 2701 tctgtggag agaaggagat atgtccctt tctggatgg actgatcaac gctcctatcg
 2761 attttgatcc tgtgcctaac acaagacaa tctttgatga aagcagcagc agcggagcct
 2821 ccgccaaga agatcagcat gccctcccta acatcagcct gacactgagc tatttcccta
 2881 aggtgaacga aacacagcc cattccggag aaaacgaaa cgattgtgat gccgaactga
 2941 gaatctggag cgtgcaggaa gatgatctgg ccgcccagc gagctggatc cctttttt

FIG. 5B-2

3001 ggcccggaaat tgaaggactg tacaccgccg gctgatcaa gaaccagaac aacctgggtg
 3061 gcaggctgag gaggctggcc aaccagaccg ccaagagcct ggagctgctg ctgaggggtga
 3121 ccaccgagga gaggaccttc agcctgatca acaggcacgc catcgacttc ctgctggcta
 3181 ggtggggcgg cactgcaag gtgctgggcc ccgactgctg catcgccatc gaggacctga
 3241 gcaggacaat cagcgagcag atcgaccaga tcaagaagga cgaacagaag gagggcaccg
 3301 gctggggcct gggcggaag tggtagacca gcgactgggg agtgcgaca aacctgggaa
 3361 tctgctgct gctgagcatt gccgtgtca ttgctgtc ctgtatctt agaatttta
 3421 ccaagtacat cggatgata gctgctgtg ccttctagt gccagccatc tgtgtttgc
 3481 cctccccgc tgccttctt gacctggaa ggtgccactc cactgtctt ttctaataa
 3541 aatgaggaat ttgcatgca ttgtctagt aggtgtcatt ctattctggg gggtaggggtg
 3601 gggcaggaca gcaaggggga ggattggaa gacaatagca ggcattctgg ggaatgcggtg
 3661 ggctctatgg gtaccaggt gctgaagaat tgaccgggt cctcctgggc cagaaagaag
 3721 caggcacatc ccttctctg tgacacacc tgtccagcc cctggttctt agttccagcc
 3781 ccactcatag gacactcata gctcaggagg gctccgcctt caatccacc cgttaaagta
 3841 ctggagcgg tcttccctc cctcatcag ccaccaaac aaacctagcc tccaagagt
 3901 ggaagaaatt aaagcaagat aggtattaa gtgcagagg agagaaaatg cctccaacat
 3961 gtgaggaagt aatgagagaa atcatagaat ttaaggcca tgatttaagg ccatcatggc
 4021 cttaatttc cgcttctcg ctactgact cgtgcgctc ggtcgttcgg ctgcggcgag
 4081 cggtatcagc tactcaaa gcgtaatac ggtatccac agaatacagg gataacgcag
 4141 gaaagaacat gtgagcaaaa ggccagcaaa agccaggaa cgtaaaaag gccgcgttg
 4201 tggcgtttt ccataggctc gccccctg acgagcatca caaaatcga cgctcaagtc
 4261 agaggtggcg aaaccgcga gactataaa gataccagc gttccccct ggaagctccc
 4321 tctgctgctc tctgttccg accctgccg ttaccggata cctgtccgc ttctccctt
 4381 cgggaagcgt ggcgttct catagctcac gctgtaggta tctcagttcg gtgtaggtc
 4441 ttcgtccaa gctgggctgt gtgcacgaac ccccggtca gccgaccgc tgcgcctat
 4501 ccgtaacta tctcttgag tccaaccgg taagacaga ctatcgcca ctggcagcag
 4561 ccactggtaa caggattagc agagcgaggt atgtaggagg tctacagag ttctgaagt
 4621 ggtggcctaa ctacggctac actagaagaa cagtatttg tatctgcgt ctgctgaagc
 4681 cagtacctt cggaaaaa gttgtagct ctgatccgg caaacaacc accgttggt
 4741 cgggtgttt ttgtttgc aagcagcaga ttacgcgag aaaaaagga tctcaagaag
 4801 atccttgat ctcttctac ggtctgacg ctgagtgaa cgaactca cgttaaggga
 4861 ttgtgcat gagattatca aaaaggatct taccatagat cttttaat taaaaatga
 4921 gtttaaatc aatctaaagt atatatagt aaacttggtc tgacagttac caatgctaa
 4981 tcatgaggc acctatctc gcgactgtc ttttcttc atccatagt gcctgactc
 5041 gggggggggg gcgctgaggt ctgcctctg aagaagggtg tctgactca taccaggcct
 5101 gaatgcccc atcatccagc cagaaagtga gggagccag gttgatgaga gcttgtgtg
 5161 aggtggacca gttggtgatt tgaacttt ctttgccac ggaacggtc cgttctcgg
 5221 gaagatgcgt gatctgatc ticaactcag caaaagtctg attattcaa caaagccgc
 5281 gtccctcaa gtcagcgtaa tgcctgcca gtgtacaac caattaacca attctgatta
 5341 gaaaaacta tcgagcatc aatgaaactg caatttattc atatcaggat tatcaatacc
 5401 atattttga aaaagccgt tctgtaata aggagaaaac tcaccgaggc agttccatg
 5461 gatggcaaga tcttggtatc ggtctgcgt tccgactct ccaacatca tacaacctat
 5521 taattcccc tctcaaaaa taaggttatc aagtgaagaa tcaccatgag tgacgactga
 5581 atccggtgag aatggcaaaa gcttatgcat tctttccag actgttcaa caggccagcc
 5641 attacgtcg tcatcaaat cactcgcatc aaccaaccg ttattcttc gtgattgcgc
 5701 ctgagcgaga cgaatacgc gatcgctgtt aaaaggacaa ttacaacag gaatcgaatg
 5761 caaccggcgc aggaacactg ccagcgcatc aacaatttt taccatgaat caggatattc
 5821 ttctaatacc tggaaatctg ttccccggg gatcgagtg gtgagtaacc atgcatcatc
 5881 aggagtagcg ataaaatgct tgatggtcgg aagaggcata aattccgtca gccagtttag
 5941 tctgaccatc tcatctgtaa catcattggc aacgtacct ttgcatgtt tcagaaacaa
 6001 ctctggcgca tgggcttcc catacaatcg atagattgtc gcactgatt gcccgacatt

20/44

FIG. 5B-3

6061 atcgcgagcc cattatacc catataaatc agcatccatg ttggaattta atcgcggcct
6121 cgagcaagac gtttccggtt gaatatggct cataacaccc ctgtattac tgtttatgta
6181 agcagacagt ttattgttc atgatgatat attttatct tgtgcaatgt aacatcagag
6241 attttgagac acaacgtggc ttcccccc ccccatfat tgaagcattt atcagggtta
6301 ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaaciaa taggggttcc
6361 gcgcacattt cccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt
6421 aacctataaa aataggcgta tcacgaggcc ctttcgtc

*FIG. 5B**FIG.5B-1**FIG.5B-2**FIG.5B-3*

21/44

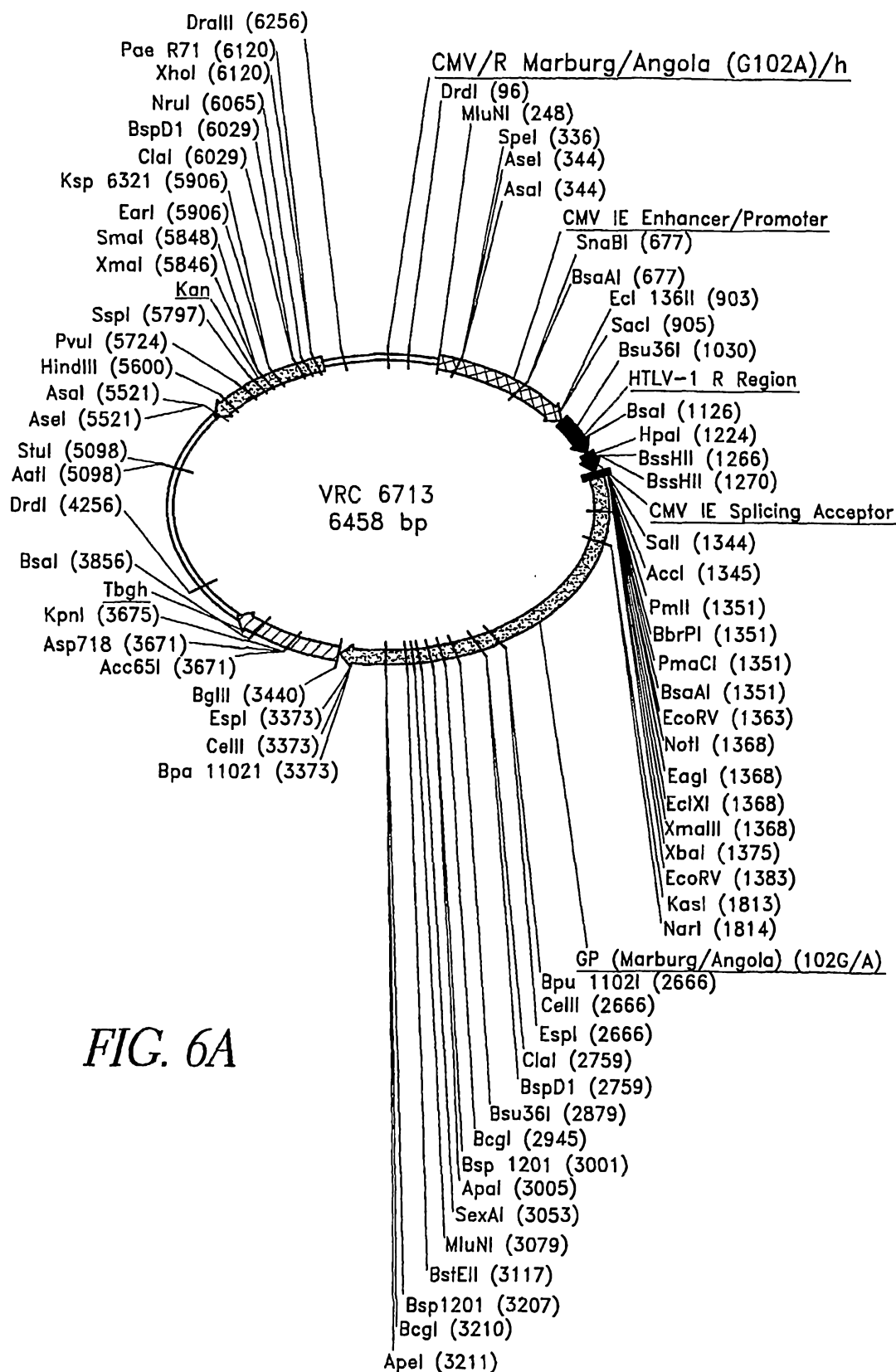


FIG. 6A

22/44

*FIG. 6B-1***VRC 6713 (SEQ ID NO: 6)**

1 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca
 61 cagcttgctt gtaagcggat gccgggagca gacaagcccg tcaggcgcg cgacgggggtg
 121 ttggcgggtg tcggggcttg cttactatg cggcatcaga gcagattgta ctgagagtgc
 181 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg
 241 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg
 301 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac
 361 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg
 421 cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga cgtatgttcc
 481 catagtaacg ccaataggga ctttcattg acgtcaatgg gtggagtatt tacggtaaac
 541 tgcccattg gcagtacac aagtgtatca tatgccaagt acgccccta ttgacgtcaa
 601 tgacggtaaa tggcccgctt gccattatgc ccagtacatg acctatggg actttctac
 661 ttggcaglac atctacgtat tagtcatcgc tattaccatg gtgatgcggt ttggcagta
 721 catcaatggg cgtggatagc ggttgactc acggggattt ccaagctcc accccattga
 781 cgtcaatggg agtttggtt ggacacaaaa tcaacgggac ttccaaaaa gtcgtaacaa
 841 ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag
 901 agctcgitta gtgaaccgtc agatcgctg gagacgcat ccacgtgtt ttgacctca
 961 tagaagacac cgggaccgat ccagcctcca tcggctcga tctctcttc acgcccgcg
 1021 cgcctacct gaggcgcca tccacgcgg ttgagtcg cgcttgcgc tcccgcctgt
 1081 ggtgcctct gaactgcgtc cgccgtctag gtaagttaa agctcaggtc gagaccgggc
 1141 cttgtccgg cgctccctt gagcctacct agactcagc ggctctccac gctttgcctg
 1201 acctgcttg ctcaactcta gttacgggtg gagggcagtg tagtctgagc agtactcgtt
 1261 gctgccgcg gcgccaccag acataatagc tgacagacta acagactgtt ctttccatg
 1321 ggtctttt cagtcaccg tcgtcgacac gtgtgatcag atatcgcg cgctctagag
 1381 atatcgccg catgaagacc acctgcctgc tgatcagcct gatcctgac cagggcgtga
 1441 agacctgcc catcctggag atcgccagca acatccagc ccagaacgtg gacagcgtgt
 1501 gcagcggcac cctgcagaag accgaggacg tgcacctgat gggcttcacc ctgagcggcc
 1561 agaaggtggc cgacagccct ctggaggcca gcaagagggt ggccctcagg gccggcgctg
 1621 ccccaagaa cgtggagtac accgaggccg aggaggccaa gacctgctac aacatcagcg
 1681 tgaccgccc cagcggcaag agcctgctgc tggacctcc caccaacatc agggactacc
 1741 ctatgtcaa gaccatccac cacatccagg gccagaaccc tcacgcccag ggcacgccc
 1801 tgcacctgtg gggcgcttc ttctgtacg acaggatgc cagcaccacc atgtacagag
 1861 gaaaagtgtt cacagaggga aacatcgtc ctatgatcgt gaacaagacc gtgcataaga
 1921 tgaattcag cagacaggga caggatata gacatatga cctgacatcc acaacaagt
 1981 actggacaag cagcaacgga acacagacaa acgatacagg atgttttga aactgcagg
 2041 aatacaactc caccaagaac cagacatgtg cccctagcaa gaagcctctg cctctgccta
 2101 cagctcatcc tgaagtgaag ctgacatcca caagcacaga tgccacaaag ctgaacaaa
 2161 cagatcctaa tagcgacgac gaggatcga caacaagcgg atccggatcc ggagaacagg
 2221 aaccttatac aacaagcgac gctgtacaa aacagggact gtccctcaca atgcctccta
 2281 cactagccc tcagcctagc acaccagc agggaggcaa caacacaaac cattcccagg
 2341 gagtgggtgac agaacctgga aagacaaa caacagccca gcctagcatg cctctcata
 2401 acacaacaac aatcagcaca aacaacacct ccaagcaca tctgagcaca cctagcggtc
 2461 ctattcagaa tgccaccaac tacaacacac agtccacag ccctgaaaac gaacagacct
 2521 ccgccccttc caaacaaccc ctgtgccta cagaaaaccc tacaacagcc aagagcaca
 2581 acagcacaaa gagccctaca acaacagtgc ctacacaa aaacaagtat agcacaagcc
 2641 ctagccctac acctaatcc acagtcagc atctggtgta tttagaaga aagagaacaa
 2701 tctgtggag agaaggagat atgtccctt tctggatgg actgatcaac gctcctatc
 2761 atttgatcc tgtgcctaac acaagacaa tcttgatga aagcagcagc agcggagcct
 2821 ccggcgaaga agatcagcat gcctccccta acatcagcct gacactgagc tatcttcta
 2881 aggtgaacga aaacacagcc cattccggag aaaacgaaa cgattgtat gccgaactga

23/44

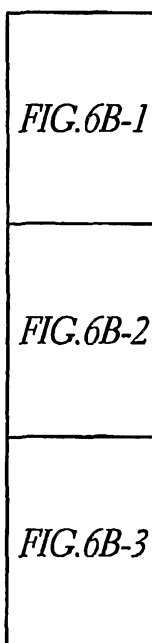
FIG. 6B-2

2941 gaatctggag cgtgcaggaa gatgatctgg ccgccggact gagctggatc ccttttttg
 3001 ggcccggaaat tgaaggactg tacaccgccg gcctgatcaa gaaccagaac aacctgggtg
 3061 gcaggctgag gaggctggcc aaccagaccg ccaagagcct ggagctgctg ctgagggtga
 3121 ccaccgagga gaggaccttc agcctgatca acaggcacgc catcgacttc ctgctggcta
 3181 ggtggggcgg cacctgcaag gtgctgggcc ccgactgctg catcggcacg gaggacctga
 3241 gcaggaaacat cagcgagcag atcgaccaga tcaagaagga cagcagaag gagggcaccg
 3301 gctggggcct gggcggcaag tggggacca gcgactgggg agtgcgaca aacctgggaa
 3361 tctgctgct gctgagcatt gccgtgctca ttgctgtgc ctgtatctgt agaattctta
 3421 ccaagtacat cggatgatag atctgctgtg ccttctagt gccagccatc tgtgtttgc
 3481 cctcccccg tgccttctt gacctggaa ggtgccactc cactgtctt ttctaataa
 3541 aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg ggtgtgggtg
 3601 gggcaggaca gcaaggggga ggaattggaa gacaatagca ggcatgctgg ggaatcggtg
 3661 ggctctatgg gtaccagggt gctgaagaat tgaccgggt cctctgggc cagaagaag
 3721 caggcacatc ccttctctg tgacacacc tgcacgcc cctggtctt agttccagcc
 3781 ccactcatag gacactcata gctcaggagg gctccgctt caatccacc cgtaaaagta
 3841 cttggagcgg tcttccctc cctcatcagc ccaccaaac aaacctagcc tccaagagt
 3901 ggaagaaatt aaagcaagat aggctatfaa gtgcagaggg agagaaatg cctccaacat
 3961 gtgagggaagt aatgagagaa atcatagaat ttaaggcca tgattlaagg ccatcatggc
 4021 ctaattctc cgttctctg ctactgact cgctgcgctc ggtgttcgg ctgcggcgag
 4081 cggtatcagc tactcaaaag gcggtaalac ggtatccac agaatcagg gataacgcag
 4141 gaaagaaat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttc
 4201 tggcgtttt ccatagctc cgccccctg acgagcatca caaaatcga cgctcaagtc
 4261 agagggtggc aaaccgaca ggactataaa gataccaggc gttccccct ggaagctccc
 4321 tctgctgctc tctgttctg accctgccg ttaccggata cctgtccgc ttctccctt
 4381 cgggaagcgt ggcgttctc catagctcac gctgtaggta tctcagttc gtgtaggtc
 4441 ttctccaa gctgggtgt gtgcacgaac cccccgtca gcccgaccg tgcgcctat
 4501 ccgtaacta tctgttgag tcaaccggg taagacacga ctatcgcca ctggcagcag
 4561 ccactgtaa caggattagc agagcgagg atgtaggcg tctacagag ttctgaagt
 4621 ggtggcctaa ctacggctac actagaaga cagtattgg tatctgcgt ctgctgaagc
 4681 cagttacct cggaaaaaga gttgtagct ctgatccg caaacaaacc accgtggtg
 4741 gcggtggtt tttgtttgc aagcagcaga ttacgcgag aaaaaagga tctaaagag
 4801 atcctttgat cttttctac ggtctgacg ctactggaa cgaaaactca cgtaaggga
 4861 tttggtcat gagattatc aaaaggatct tcacctagat cctttaaat taaaatgaa
 4921 gtttaaatc aatctaaagt atatatgagt aaactggtc tgacagtiac caatgctaa
 4981 tcagtgggc acctatcica gcgatctgtc tattcgttc atccatagt gccctactg
 5041 gggggggggg gcgtgaggt ctgctcgtg aagaagggtg tctgactca taccaggct
 5101 gaatcgcccc atatccagc cagaaagtga gggagccag gtgatgaga gcttgtgt
 5161 aggtggacca gttgtgatt tgaacttt gcttgccac ggaacggtc gcgtgtcgg
 5221 gaagatgct gatctgacc tcaactcag caaagttc attattcaa caagccgcc
 5281 gtccgctca gtcagcgtaa tgcctgcca gtttacaac caattaacca attctgatta
 5341 gaaaaactca tcgagcatc aatgaactg caattatc atacaggat tatcaatac
 5401 atattttga aaaagccgt tctgtaatga aggagaaaac tcaccaggc agttccatg
 5461 gatggcaaga tcttggatc ggtctgcgt tccgactgt ccaacatca tacaacctat
 5521 taatttccc tctcaaaaa taaggttatc aagtgaagaa tcaccatgag tgacgactga
 5581 atccggtgag aatggcaaaa gcttatgcat ttcttcag actgttcaa caggccagcc
 5641 attacgtctg tcatcaaat cactgcac aaccaaccg ttatcattc gtgattgcg
 5701 ctgagcgaga cgaatacgc gatcgctgtt aaaaggcaa ttacaacag gaatcgaatg
 5761 caaccggcgc aggaacactg ccagcgcatc acaatattt tcactgaat caggatattc
 5821 ttctaatac tgaatgctg ttctccggg gatcgagtg gtgagtaacc atgcatc
 5881 aggagtagcg ataaaatgt tgaatgctg aagggcata aattccgtc gccagtttag
 5941 tctgaccatc tcatctgaa catattggc aacgctacct ttgcatgtt tcagaacaa

24/44

FIG. 6B-3

6001 ctctggcgca tcgggcttcc catacaatcg atagattgtc gcacctgatt gcccgacatt
6061 atcgcgagcc catttatacc catataaatc agcatccatg ttggaattta atcgcggcct
6121 cgagcaagac gtttcccggt gaatatggct cataacaccc ctgtattac tgtttatgta
6181 agcagacagt tttattgtc atgatgatat atttttatct tgtgcaatgt aacatcagag
6241 atttgagac acaacgtggc ttccccccc cccccattat tgaagcattt atcaggggta
6301 ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaaca aa taggggttcc
6361 gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accatttata tcatgacatt
6421 aacctataaa aataggcgta tcacgaggcc ctttcgtc

FIG. 6B

25/44

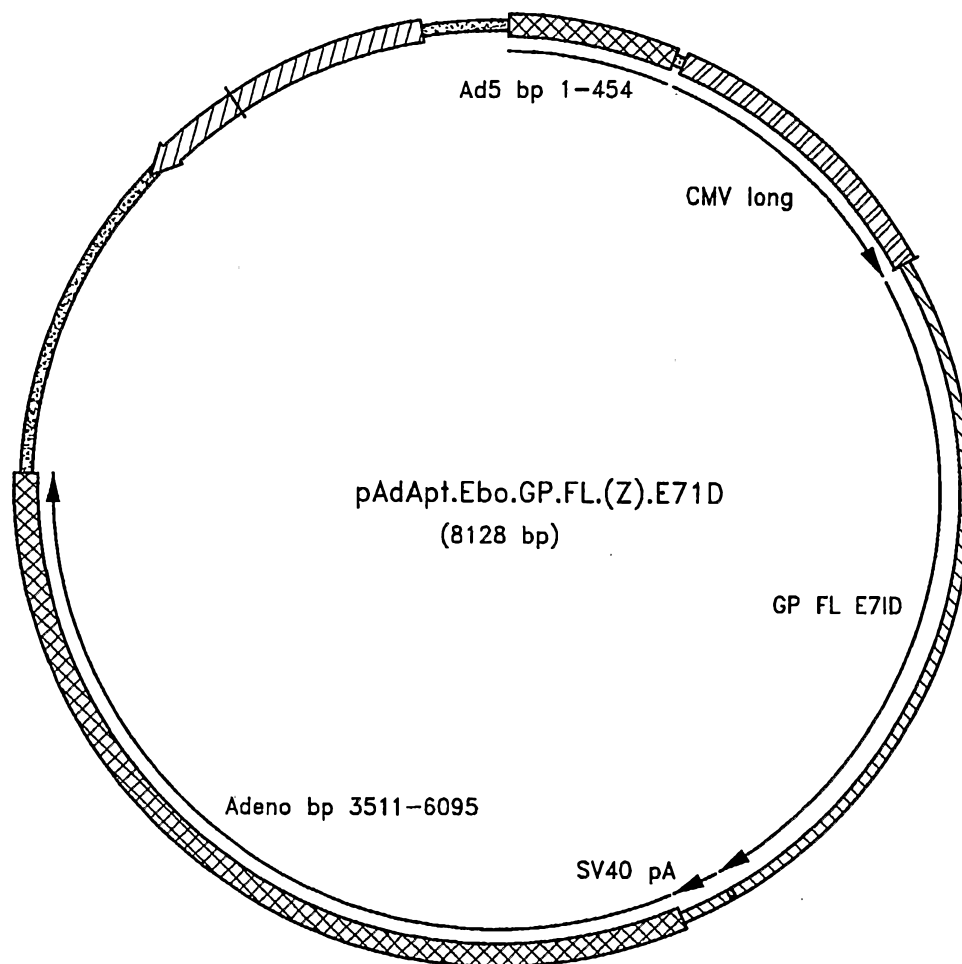


FIG. 7A

26/44

FIG. 7B



FIG. 7B-1

27/44

pAdApt.Ebo.GP.Fl.(Z).E71D (SEQ ID NO: 7)

ttaattaaccgcaattctcatgtttgacagcttatcatcatcaataataaccttattttggattgaagccaa
 tatgataatgagggggtggagtttgtgacgtggcgcgggcgctgggaacggggcggtgacgtagtagtggtg
 cggaagtgtgatgttgcaagtgtggcggaacacatgtaagcgacggatgtggcaaaagtacgtttttggtgt
 gcgcgggtgtacacaggaagtacaattttcgcgcggttttaggcggatgtttagtaaaatttggcgtaacc
 gtagaagatttggccattttcgcgggaaaactgaataagaggaagtgaatctgaataatttgggttactca
 tagcgcgtaaatatttctagggcgcggggactttgaccgtttacgtggagactcgccaggtgtttttctc
 aggtgtttttcgcggttccgggtcaaagttggcggttttattattatagtcagtagctaccagtgcaactggccta
 ggtggtcaatattggccattagccatattattcattggttatatagcataaatcaatattggctattggccat
 tgcatacgttgtatccatatcataatatgtacatttatattggctcatgtccaacattaccgccatgttgaca
 ttgattattgactagttattaatagtaatcaattacggggctcattagttcatagcccatataggagttccgc
 gttacataaacttacggtaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaataatga
 cgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagattttacggtaaatgc
 ccacttggcagtagcatcaagtgtatcatatgccagtagccccctattgacgtcaatgacggtaaatggccc
 gcctggcattatgccagtagcatgaccttatgggactttcctacttggcagtagcatctacgtattagtcacgc
 ctattaccatggtgatgcggttttggcagtagcatcaatggcggtggatagcggttgactcacggggatttcc
 aagctctccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt
 aacaactccgccccattgacgcaaatggcggttaggcgtgtacggtgggaggtctatataagcagagctcgtt
 tagtgaaccgtcagatcgctggagacgcatccacgctgttttgacctccatagaagacaccgggaccgatc
 cagcctccgcgccgggaacggtgcattggaagcttgcgcgc

ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTCTTTCTTTGGGTAATTA
 TCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAATAGCACATTACAGGTTAGTGATGTCGA
 CAACTAGTTTGTCTGTGACAACTGTCTCCACAATCAATTGAGATCAGTTGGACTGAATCTCGATGGCAAT
 GGAGTGGCAACTGACGTGCCATCTGCAACTAAAAGATGGGGCTTCAGGTCCGGTGTCCCACCAAGGTGGTCA
 ATTATGAAGCTGGTGAATGGGCTGAAAAGTGTCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAGTGCT
 ACCAGCAGCGCCAGACGGGATTTCGGGGCTTCCCCCGGTGCCGGTATGTGCACAAAGTATCAGGAACGGGACCG
 TGTGCCGGAGACTTTGCCTTCCATAAAGAGGGTGCTTTCTTCTGTATGATCGACTTGCTTCCACAGTTATCT
 ACCGAGGAACGACTTTTCGTGAAGGTGTCTGTTGCTTCTGATACTGCCCCAAGCTAAGAAGGACTTCTTCAG
 CTCACACCCCTTGAGAGAGCCGGTCAATGCAACGGAGGACCCGTCTAGTGGCTACTATTCTACCACAATTAGA
 TATCAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACCTACGTCCAAC
 TTGAATCAAGATTACACCACAGTTTCTGCTCCAGCTGAATGAGACAATATATACAAGTGGGAAAAGGAGCAA
 TACCACGGGAAAATAATTTGGAAGGTCAACCCGAAATTGATACAACAATCGGGGAGTGGGCCTTCTGGGAA
 ACTAAAAAAACCTCACTAGAAAAATTCGCAGTGAAGAGTTGTCTTTCACAGTTGTATCAAACGGAGCCAAAA
 ACATCAGTGGTCAGAGTCCGGCGCGAATTCTTCCGACCCAGGGACCAACACAACAACCTGAAGACCACAAAAT
 CATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCTGCAGTGTGCGATCTA
 ACAACCTTCCACAAATCTCCACGAGTCCCCAATCCCTCACAACCAACCAGGTCGGGACAACGACCCCATATA
 ATACCCCTGTGATAAATTCGATCTCTGAGGCAACTCAAGTTGAACAACATCACCAGCAACAGACAACGA
 CAGCACAGCCTCCGACACTCCCTCTGCCACGACCGCAGCCGGACCCCCAAAAGCAGAGAACACCAACACGAGC
 AAGAGCACTGACTTCTGGACCCCGCCACCACAACAAGTCCCCAAAACACAGCGAGACCGCTGGCAACAACA
 ACATCATACCAAGATACCGGAGAAGAGAGTGCCAGCAGCGGGAAGCTAGGCTTAATTACCAATACTATTGC
 TGGAGTCGACGAGTGTATCACAGGCGGGAGAAGAACTCGAAGAGAAGCAATTGTCAATGCTCAACCCAAATGC
 AACCCTAATTTACATTACTGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCCTGGATACCATATTTTCG
 GGCCAGCAGCCGAGGGAATTTACATAGAGGGGCTAATGCACAATCAAGATGGTTTAATCTGTGGGTTGAGACA
 GCTGGCCAACGAGACGACTCAAGCTCTTCAACTGTTTCTGAGAGCCACAACCTGAGCTACGCACCTTTTCAATC
 CTCAACCGTAAGGCAATTGATTTCTTGTGCTGACGATGGGGCGGCACATGCCACATTCTGGGACCGGACTGCT
 GTATCGAACCACATGATTGGACCAAGAACATAACAGACAAAATTGATCAGATTATTATGATTTTGTGATAA
 AACCCTTCGGGACCGGGGACAATGACAATTGGTGGACAGGATGGAGACAATGGATACCGGCAGGTATTGGA
 GTTACAGGCGTTGTAATTGACAGTTATCGCTTTTATTCTGTATATGCAAAATTTGTCTTT

tagtaattctagacgagatccgaacttgtttattgagcttataatgggttacaaataaagcaatagcatcaciaa
 atttcacaaataaagcatttttttcaactgcattctagttgtggtttgtccaaactcatcaatgtatcttatca
 tgtctagatctgtactgaaatgtgtggcggtggttaagggtgggaagaatataaaggtgggggtcttatg
 tagttttgtatctgttttgcagcagccgcgcgcgcctgagcaccactcggttgatggaagcattgtgagct
 catatttgacaacgcgcgtgccccatgggcccgggtgcgtcagaatgtgatgggctccagcattgatggctg
 ccccgctcctgcccgaactctactaccttgacctacgagaccgtgtctggaacgcggttgagactgcagcc
 tccgcgcgcgttcagccgctgcagccaccgcccgcgggattgtgactgactttgctttcctgagcccgttg
 caagcagtgagcttcccgttcatccgcccgcgatgacaagttgacggctcttttggcacaattggattcttt
 gaccgggaacttaatgtcgtttctcagcagctgttggatctgcgccagcaggtttctgcctgaaggcttcc
 tcccctcccaatgcggtttaaaacataaataaaaaaccagactctgtttggatttggatcaagcaagtgtctt

28/44

FIG. 7B-2

gctgtctttatatttaggggttttgcgcgcgcggtagggcccgaggaccagcgggtctcggtcggttaggggtcctgtg
tattttttccaggacgtggtaaaggtagctctggatgttcagatacatgggcataagcccgctctctgggggtgg
aggttagcaccactgcagagcttcatgctgcgggggtgggtgtttagatgatccagtcgtagcaggagcgcgtggg
cgtgggtgcctaaaaatgtctttcagtagcaagctgattgccaggggcaggcccttggtgtaagtgtttacaaa
gcgggttaagctgggatgggtgcatacgtggggatatgagatgcatcttgactgtatttttaggttggctatg
ttcccagccatatccctccggggattcatgttgtgcagaaccaccagcacagtgtatccggtgcacttgggaa
atttgtcatgtagcttagaaggaaatgcgtggaagaacttggagacgcccttgtgacctccaagattttccat
gcattcgtccataatgatggcaatgggcccacggggcgccggcctgggccaagatatcttgggatcactaacg
tcatagtgtgttccaggatgagatcgtcataggccatttttacaaggcgcgggcggagggtgccagactgcg
gtataatggttccatccggcccaggggcgtagttaccctcacagatttgcatttcccacgctttgagttcaga
tggggggatcatgtctacctgcggggcgatgaagaaaacgggttccggggtaggggagatcagctgggaagaa
agcaggttccctgagcagctgcgacttaccgcagccgggtgggcccgtaaatcacacctattaccggctgcaact
ggtagttaagagagctgcagctgccgtcatccctgagcagggggggccacttcgttaagcatgtccctgactcg
catgttttccctgaccaaaccggccagaaggcgtcgccgcccagcgatagcagttcttgcaaggaagcaaaag
tttttcaacgggttgagaccgtccgccgtaggcagctgttttgagcgtttgaccaagcagttccaggcgggtccc
acagctcggtcacctgctctacggcatctcgatccagcatactctcctcgtttcgcgggttggggcggttctcg
ctgtacggcagtagtcggtgctcgtccagacggggcagggtcatgtctttccacgggcgaggggtcctcgta
gcgtagctcgggtcacggtgaaggggtgcgtccgggctgcgcgctggccagggtgcgcttgaggctgggtcct
gctggtgctgaagcgtgcgggtcttcgcccgtgcgcgctcgccaggtagcatttgaccatgggtgtcatagtcc
agcccctccgcggcgtggcccttggcgcgcagcttgcccttggaggaggcgcgcagcaggggcagtgacagac
ttttgagggcgtagagcttgggcgcgagaaataccgattccggggagtaggcacccgcgcagggccccgca
gacggtctcgcattccacgagccaggtgagctctggccgttcggggccaacagggttcccccatgcttt
ttgatgcgtttcttacctctggtttccatgagccgggtgtccacgctcggtagcgaacagggtgtccgtgtccc
cgtatacagacttgagaggcctgtcctcgagcgggtgttccgcggtcctcctcgatatagaaactcggaccactc
tgagacaaaggctcgcgtccaggccagcacgaaggaggttaagtgggagggtagcgggtcgttgtccactagg
gggtccactcgtccagggtgtgaagacacatgtcgcctcttcggcatcaaggaaaggtagttgggtttagg
gttagggccacgtgaccgggtgttccctgaagggggcgtataaaaagggggtggggcgcggttcgtcctcactctc
ttccgcatcgctgtctgcgagggccagctgttggggtgagtcgacgcgaggtggatggccttccccattatg
attcttctcgttccggcgcatcgggatgcccgcttgaggccatgctgtccaggcaggtagatgacgacc
atcaggacagcttcaaggccagcaaaaggccaggaaccgtaaaaaggccggttgcgtggcgtttttccatag
gctccgccccctgacgagcatcaaaaaatcgacgctcaagtcagaggtggcgaacccgacaggactataa
agataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgacctgcccgttaccggatacc
tgtccgcctttctcccttcgggaagcgtggcgttttctcatagctcacgctgtaggtatctcagttcgggtga
ggctggtcgtccaagctgggtgtgtgcacgaacccccgttcagcccagccgtgcgccttatccggtaac
tatcgtcttagtccaacccggtgaagacagacttatcgccactggcagcagccactggtaacaggattagca
gagcaggttgcgttagggcgtgtacagagttcttgaagtggtggcctaactacggctacataagaaggacag
atttggtagctcgtcgtgtaagccagttaccttcggaagagagttggtagctcttgatccggcaaaacaa
accaccgctggtagcgggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaag
atcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggatttttggtcatgag
attatcaaaaaggatcttcacctagatcttttaaaatataaaatgaagttttaaatacaatctaaagtataat
gagtaaaacttggctgacagttaccaatgcttaatcagtgaggcacctatctcagcgtatctgtctatttcgtt
catccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtg
tgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaacagccagccggaagggcc
gagcgcagaagtgtcctgcaactttatccgcctccatccagctctattaattgttgccgggaagctagagtaa
gtagttcgccagtttaagtttgcgaacggttggccattgctgcaggcatcgtgggtgcacgctcgtcgtt
tggtatggcttattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaa
gcgggttagctccttcggctcctccgatcgttgcagaagtaagttggccgcagtggttatcactcaggttatgg
cagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaa
gtcattctgagaatagtgtatgcggcgaccgagttgctcttgcggcggtcaacacgggataataaccgcgcca
catagcagaactttaaagtgtcatattggaacggttcttcggggcgaaaactctcaaggatcttaccgc
tggttagatccagttcgatgtaaccactcgtgcaccaactgatcttcagcatcttttacttccaccagcgt
ttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaagggaataagggcgacacggaaatgttgaata
ctcactactcttctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatataatattg
aatgtatttagaaaaataacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaaga
aaccattattatcatgacattaacctataaaaaataggcgtatcacgaggcccttctgtcttcaagaattgtt

29/44

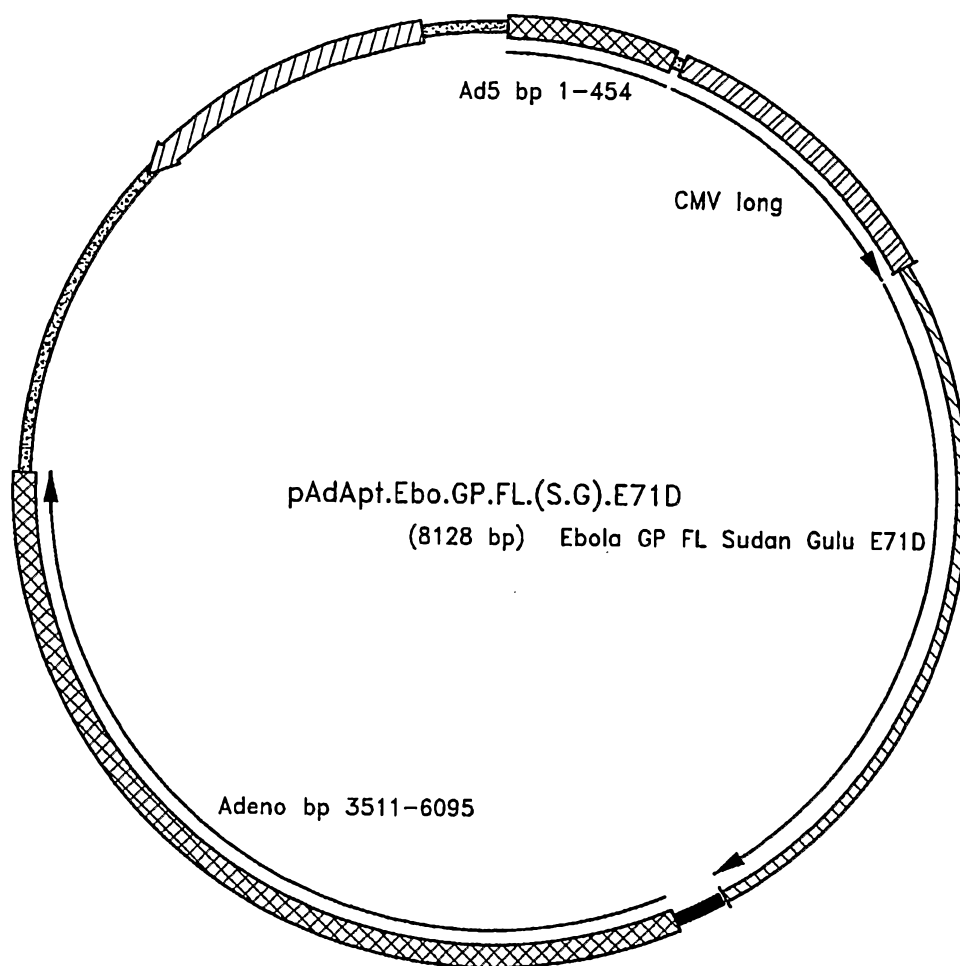


FIG. 8A

30/44

FIG. 8B

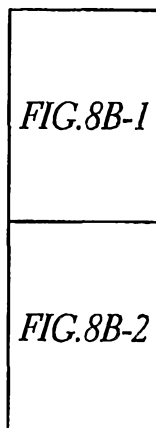


FIG. 8B-1

31/44

pAdApt.Ebo.GP.FL.(S/G).E71D (SEQ ID NO: 8)

ttaattaaccgcaattctcatgtttgacagcttatcatcatcaataatataccttattttggattgaagccaa
 tatgataatgagggggtggagtttgtgacgtggcgcgggcggtgggaacggggcggtgacgtagtagtggtg
 cggaagtgtgatgttgcaagtgtggcggaacacatgtaagcgacggatgtggcaaaagtacgtttttggtgt
 gcgccggtgtacacaggaagtacaattttcgcgcggttttagcgcgatgtttagtaaaatttgggcgtaacc
 gtagaagatttggccattttcgcgggaaaactgaataagaggaagtgaatctgaataattttgtgttactca
 tagcgcgtaaatattgtctagggcgcggggactttgaccgtttacgtggagactcgcccagggtgtttttctc
 aggtgtttttcgcggttccgggtcaaagttggcggtttttattattatagtcagtcagtcaccagtgcaactggccta
 ggtggtcaatattggccattagccatattattcattggttatatagcataaatcaatattggctattggccat
 tgcatacgttgtatcccatatcataatatgtacattttatattggctcatgtccaacattaccgccatgttgaca
 ttgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgc
 gttacataaacttacggtaaatggcccgctgggtgaccgccaacgacccccgcccattgacgtcaataatga
 cgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagttttacggtaaaactgc
 ccacttggcagtagcatcaagtgtatcatatgccagtagccccctattgacgtcaatgacggtaaatggccc
 gcctggcattatgccagtagcatgacctatgggactttcctacttggcagtagcatctacgtatttagtcatcg
 ctattaccattgggtgatgcggttttggcagtagcatcaatggcggtggatagcggtttagctcacggggatttcc
 aagtcctccaccattgacgtcaatgggagtttgggttggcaccaaaatcaacgggactttccaaaatgtcgt
 aacaactccgccccattgacgcaaatggcggttaggcgtgtacgggtgggaggtctatataagcagagctcgtt
 tagtgaaccgtcagatcgctggagacgcatccacgctgttttgacctcatagaagacaccgggaccgatc
 cagcctccgcgccgggaacgggtgcattggaagcttgcgcgc

ATGGGGGGTCTTAGCCTACTCCAATTGCCAGGGACAAATTTCCGAAAAGCTCTTTCTTTGTTGGGTCAATCA
 TCTTATTCCAAAAGGCCCTTTTCCATGCCTTTGGGTGTTGTGACTAACAGCACCTTTAGAAGTAACAGAGATTGA
 CCAGCTAGTCTGCAAGGATCATCTTGCATCTACTGACCAGCTGAAATCAGTTGGTCTCAACCTCGACGGGAGC
 GGAGTATCTACTGATATCCCATCTGCAACAAAGCGTTGGGGCTTCAGATCTGGTGTTCCTCCCAAGGTGGTCA
 GCTATGAAGCGGGAGAATTGGGCTGAAAATTGCTACAATCTTGAAATAAAGAAGCCGGACGGGAGCGAATGCTT
 ACCCCACCGCCAGATGGTGTGAGAGGCTTTTCCAAGGTGCGCGCTATGTTTCAAAAGCCCCAAGGAACCGGGCCC
 TGCCCAGGTGACTACGCCTTTTCAAGGATGGAGCTTTCTTCTCTATGACAGGCTGGCTTCAACTGTAATTT
 ACAGAGGAGTCAATTTTGTGAGGGGGTAATTGCATTCTTGATATTGGCTAAACCAAAAGAAACGTTCTTCA
 GTCACCCCCCATTCGAGAGGCAGTAACTACACTGAAATACATCAAGTTATTATGCCACATCCTACTTGGAG
 TATGAAATCGAAAATTTTGGTGTCAACACTCCACGACCCTTTTCAAATTGACAATAATACTTTTGTTCGTC
 TGGACAGGCCCCACACGCCTCAGTTCCTTTTCCAGCTGAATGATACCATTACCTTCACCAACAGTTGAGTAA
 TACAACTGGGAGACTAATTTGGACACTAGATGCTAATATCAATGCTGATATTGGTGAATGGGCTTTTGGGAA
 AATAAAAAAATCTCTCCGAACAACACTACGTGGAGAAGAGCTGTCTTTCGAAGCTTTATCGCTCAACGAGACAG
 AAGACGATGATGCGGCATCGTCGAGAATTACAAAGGGAAGAATCTCCGACCGGGCCACCAGGAAGTATTCGGA
 CCTGGTTCCAAAGAATTCCTTGGGATGGTTCCATTGCACATACCAGAAGGGGAACAACATTGCCGTCTCAG
 AATTGCAGAGAAGTTCGAAGAGTAGGTGTGAACACTCAGGAGACCATTACAGAGACAGCTGCAACAATTTATAG
 GCCTAACCGGCAACCATATGCAGATCTCCACATCGGGATAGAGACCGAGCTCCAGCCAAATCCCGATTTCTCT
 ACCGACCACGGCACCAAGCCCTGAGGCTCAGACCCCCACAACCCACACATCAGGTCCATCAGTAGTGCCACC
 GAGGAACCAACAACACCACCGGAAGCTCCCCCGGCCCAACAACAGAAGCACCCACTCTCACCACCCAGAAA
 ATATAACAACAGCGGTTAAACTGTCTGCCACAGGAGTCCACAAGCAACGGTCTAATAACTTCAACAGTAAC
 AGGGATTCTTGGGAGTCTTGGGCTTCGAAAACGCAGCAGAAGACAACTAACACCAAAGCCACGGGTAAGTGC
 AATCCCAACTTACACTACTGGACTGCACAAGAACAACATAATGCTGCTGGGATTGCCTGGATCCCGTACTTTG
 GACCGGGTGCAGGAAGGCATATACACTGAAGGCCTGATGCATAACCAAAATGCCTTAGTCTGTGGACTTAGGCA
 ACTTGCAAAATGAAACAACACTCAAGCTCTGCAGCTTTTCTTAAGAGCCACAACGGAGCTGCGGACATATACCATA
 CTCAATAGGAAGGCCATAGATTTCTTCTGCGACGATGGGGCGGGACATGCAGGATCCTGGGACCAGATTGTT
 GCATTGAGCCACATGATTGGACAAAAACATCACTGATAAAATCAACCAATCATCCATGATTTTATCGACAA
 CCCCTTACCTAATCAGGATAATGATGATAATTGGTGGACGGGCTGGAGACAGTGGATCCCTGCAGGAATAGGC
 ATTACTGGAATTATTATTGCAATTATTGCTCTTCTTCTTGGCTTTGCAAGCTGCTTTGC
 tgataatctagacgagatccgaactgtttattgagcttataatgggttacaataaagcaatagcatcacaa
 atttcacaaaataagcatttttttcaactgcattctagttgtggtttgtccaaactcaaatgtatcttatca
 tgtctagatctgtactgaaatgtgtggcggtggcttaagggtgggaagaatatataaggtgggggtcttatg
 tagttttgtatctgttttgcagcagccgcgcgcgcctgagcaccactcggttgatggaagcattgtgagct
 catatttgacaacgcgcgcgtcccccatgggcccgggtgctgcagaatgtgatgggctccagcattgatggctg
 ccccgctcctgcccgaactctactaccttgacctacgagaccgtgtctggaacgcggttgagactgcagcc
 tccgcgcgcgcttcagccgctgcagccaccgcccgcgggattgtgactgactttgctttcctgagcccgctt
 caagcagtgacgcttcccgcttcatccgcccgcgatgacaagttgacggctcttttggcacaattggattcttt
 gaccggggaacttaattgtcgtttctcagcagctgttggatctgcgccagcaggtttctgcctgaaggttcc
 tccctcccaatgcggtttaaataaataaaaaaccagactctgtttggatttggatcaagcaagtgtctt

32/44

FIG. 8B-2

gctgtctttatatttaggggttttgcgcgcgcggttaggccccgggaccagcgggtctcgggtcgttgaggggtcctgtg
tattttttccaggacgtggtaaaagggtgactctggatgttcagatacatgggcataagcccgtctctgggggtgg
aggtagcaccactgcagagcttcatgctgcgggggtgggtgtttagatgatccagtcgtagcaggagcgtggg
cgtgggtgcctaaaaatgtctttcagtagcaagctgattgccaggggcaggcccttggtgtaagtgtttacaaa
gcggttaagctgggatgggtgcatacgtggggatatgagatgcatcttggactgtatttttaggttggctatg
ttccagccatatccctccggggattcatgttgtgcagaaccaccagcacagtgtatccgggtgcacttgggaa
atattgtcatgtagcttagaaggaaatgcgtggaaagaacttggagacgcccttgtgacctccaagattttccat
gcattcgtccataatgatggcaatgggcccacgggcccggcctgggcgaagataatttctgggatcactaacg
tcatagtgtgttccaggatgagatcgtcataggccatttttacaagcgcgggcccggagggtgccagactgcg
gtataatgggttccatccggcccaggggcgtagttaccctcacagatttgcatttcccacgctttgagttcaga
tggggggatcatgtctacctgcggggcgatgaagaaaacgggtttccggggtaggggagatcagctgggaagaa
agcaggttccctgagcagctgcgacttaccgcagcgggtgggcccgtaaatcacacctattaccggctgcaact
ggtagttaagagagctgcagctgccgtcatccctgagcaggggggcccacttcgttaagcatgtccctgactcg
catgttttccctgaccaaactccgccagaaggcgtcgcgcgccagcgatagcagttcttgcaagggaagcaaag
tttttcaacgggttgagaccgtccgccgtaggcatgcttttgagcgtttgaccaagcagttccaggcgggtccc
acagctcgggtcacctgctctacggcatctcgatccagcatactcctcgtttcgcgggttggggcggccttcg
ctgtacggcagtagtcgggtgctcgtccagacggggccagggtcatgtctttccacgggcccaggggtcctcgtca
gcgtagctcgggtcacgggtgaaggggtgcgctccgggtgcgcgctggccagggtgcgcttgaggctgggtcct
gctgggtgctgaagcgtcgcgggtcttcgccctgcgcgctcgccaggttagcatttgaccatgggtgcatagtcc
agccccctccgcggcggtggcccttggcgcgagcttgccttggaggaggcgccgcacagggggcagtgacagac
ttttgagggcgtagagcttggggcgagaaaataccgattccggggagtaggcatccgcgcgcagggccccgca
gacgggtctcgcattccacgagccaggtgagctctggccgttcgggggtcaaaaaccaggtttcccccatgcttt
ttgatgcgtttcttacctctggtttccatgagccgggtgtccacgctcgggtgacgaaaaggctgtccgtgtccc
cgtatacagacttgagaggcctgtcctcgagcgggtgttcgcgggtcctcctcgtatagaaaactcggaccactc
tgagacaaaaggctcgcgtccaggccagcacgaaggaggctaaagtgggaggggtagcgggtcgttgtccactagg
gggtccactcgtccaggggtgtgaagacacatgtcgcctcttccggcatcaaggaagggtgattggtttgtagg
tgtaggccacgtgacccgggtgttccgaaggggggtataaaaaggggtggggggcggttcgtcctcactctc
ttccgcacgtcgtctgcgagggccagctgttgggggtgagtcgacgcgagggctggatggccttccccattatg
attcttctcgtctccggcggtcagggatgcccggttgcaggccatgctgtccaggcaggttagatgacgacc
atcaggggacagcttcaaggccagcaaaaaggccaggaaccgtaaaaaggccgcgttgctggcggtttttccatag
gctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaaccgcagaggactataa
agataaccaggcggtttccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttacccgataacc
tgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtagggtatctcagttcgggtgta
ggctcgttcgctccaagctgggtgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaac
tatcgtcttgagtcgaacccggtaagacacgacttctgcacactggcagcagcactggtaacaggattagca
gagcgagggttagtgagggtgtacagagttcttgaagtggtggcctaactacggctacactagaaggacaggt
atattggtatctgcgctctgctgaagccagttaccttcggaaaaagagtttggtagctcttgatccggcaaaaca
accaccgctggtagcgggtgggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaag
atcctttgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgttaagggtattttggtcatgag
attatcaaaaaggatcttcacctagatccttttaattaaaaatgaagttttaaatcaatctaaagtatatat
gagtaaaacttggtctgacagttaccaatgcttaatcagtgagggacctatctcagcgatctgtctatttcgtt
catccatagtttgctgactccccgtcgtgtagataactacgatacgggaggggttaccatctggccccagtg
tgcaatgataccgcgagacccacgctcacgggtccagatttatcagcaataaaccagccagccggaaggcc
gagcgcagaagtggctcgtcaactttatccgcctccatccagctctattaattggttgcgggaagctagagtaa
gtagttcggccagttaatagtttgcgcaacgttgggtgacattgctgcaggcatcgtgggtgcagctcgtcgtt
tggtatggcttcttcagctccgggttcccaacgatcaaggcgagttacatgatcccccatggttgcaaaaaa
gcgggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtggttatcactcatgggtatgg
cagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaa
gtcattctgagaatagtgtatgcggcgacccaggttgccttgcggcggtcaacacgggataataaccgcgcca
catagcagaactttaaaagtgtcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgc
tggttagatccagttcgatgtaaccactcgtgcacccaactgatcttcagcatcttttactttaccagcgt
ttctgggtgagcaaaaacaggaaggcaaaaatgccgcaaaaaagggaataaggggcgacacggaaatgttgaata
ctcatactcttcttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttg
aatgtatttagaaaaataaacaatataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaaga
aaccattattatcatgacattaacctataaaaaataggcgtatcacgaggcccttctcgtcttcaagaattggtt

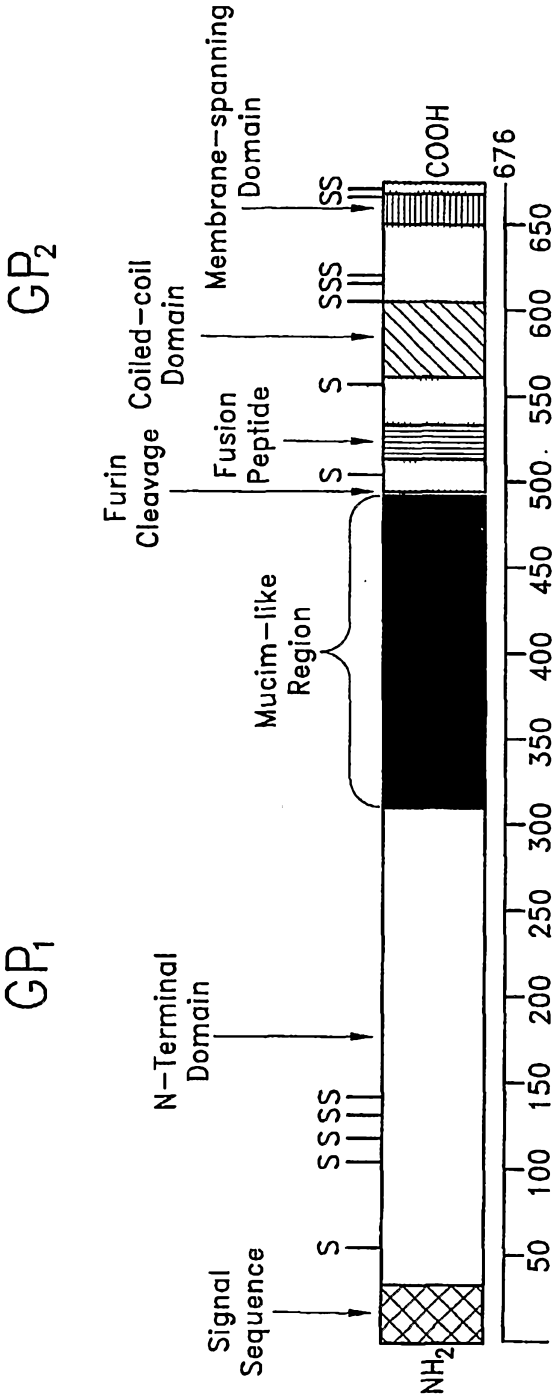
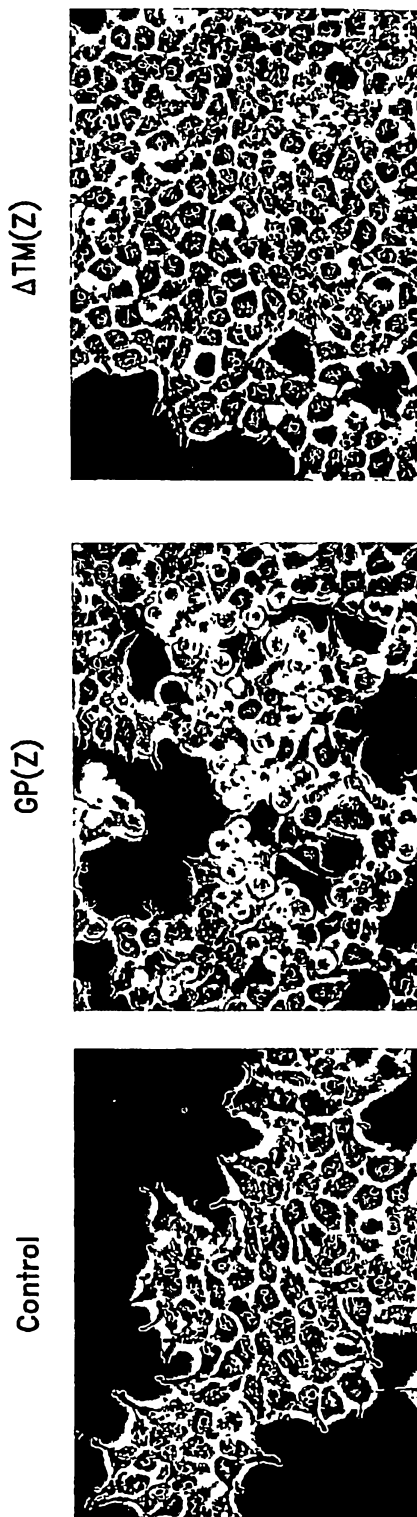


FIG. 9



FIG. 10B



35/44

FIG. 11A

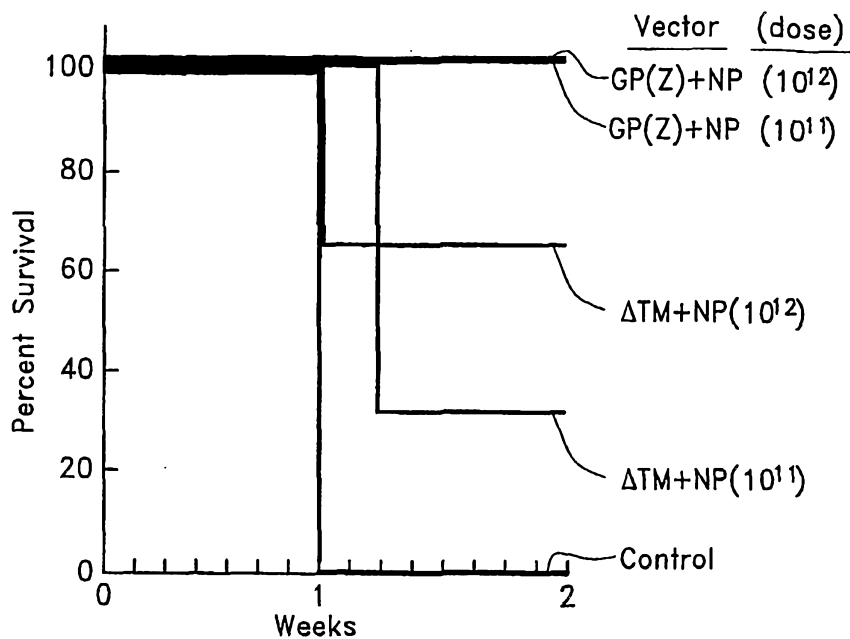


FIG. 11B

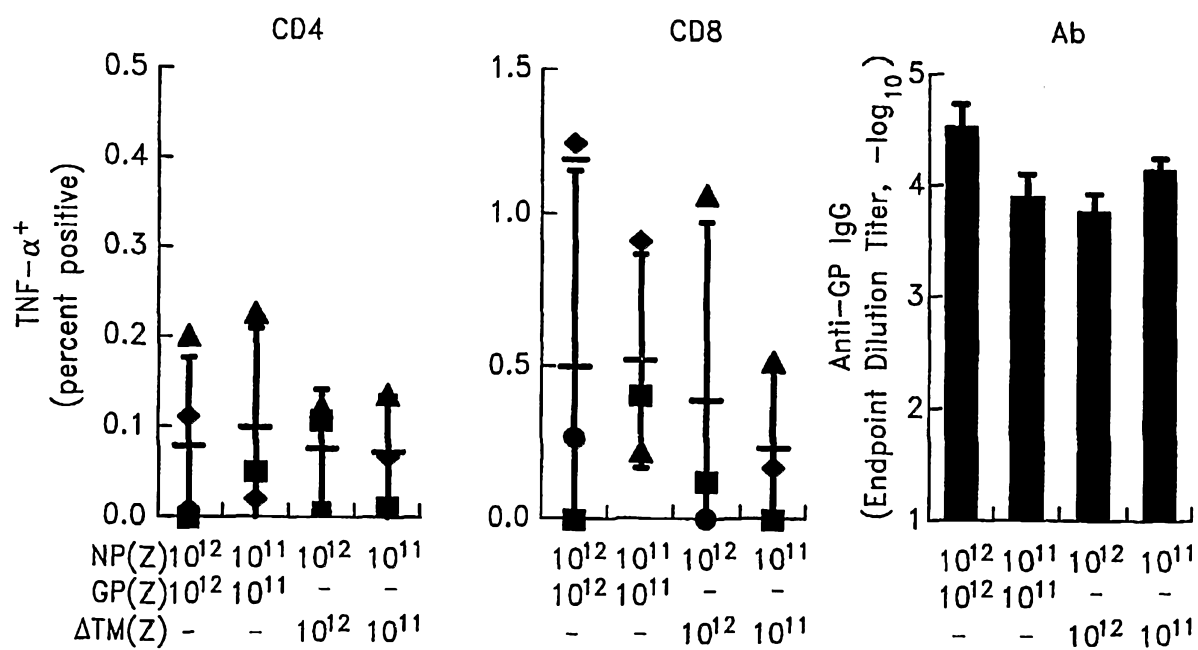


FIG. 12A

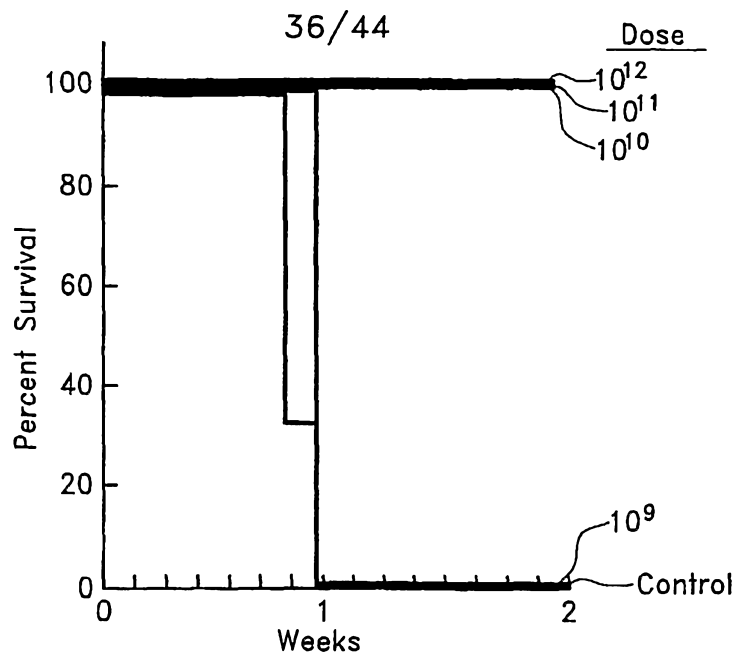


FIG. 12B

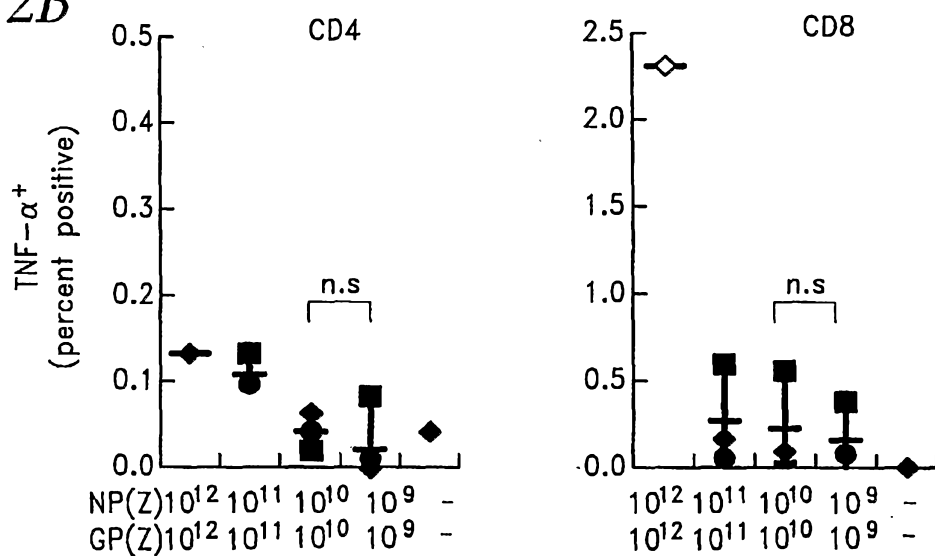
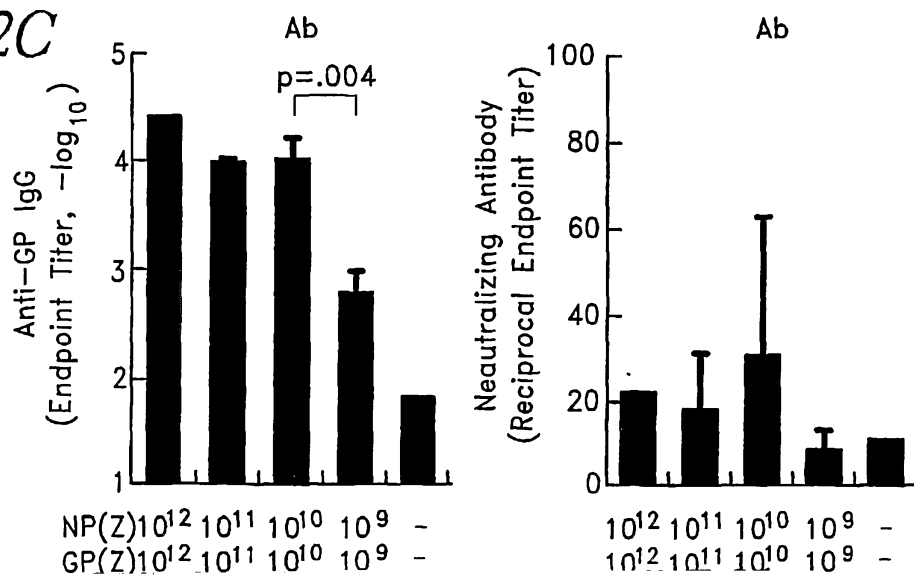


FIG. 12C



37/44

FIG. 13A

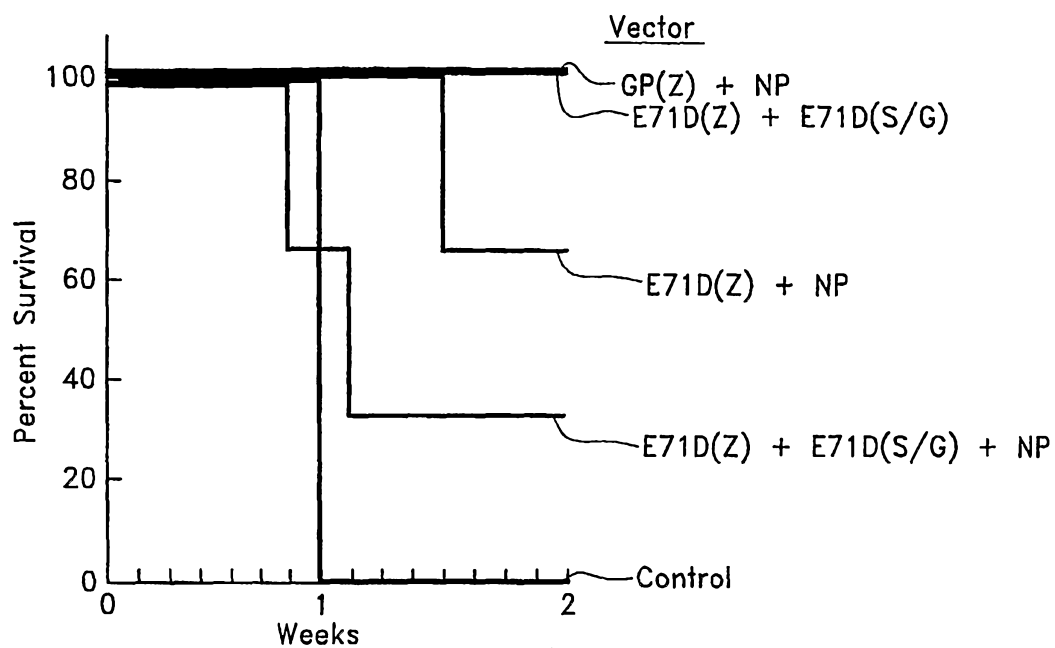


FIG. 13B

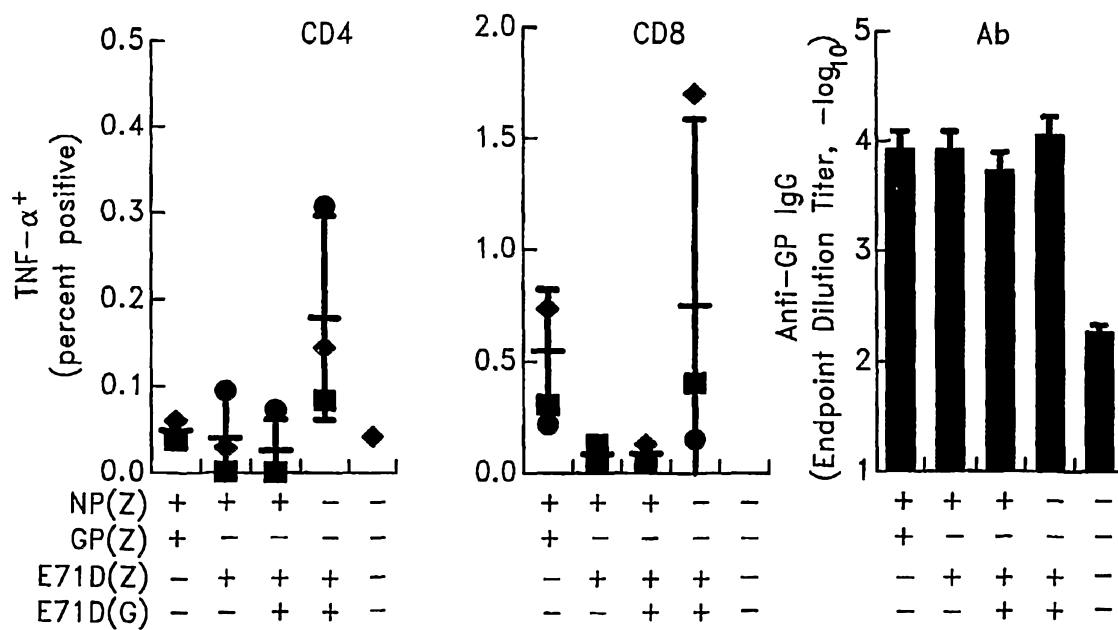


FIG. 14A

FIG. 14B

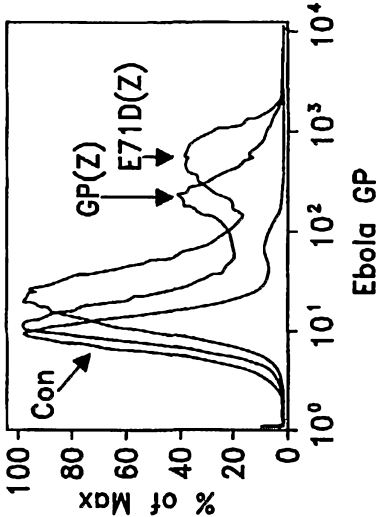
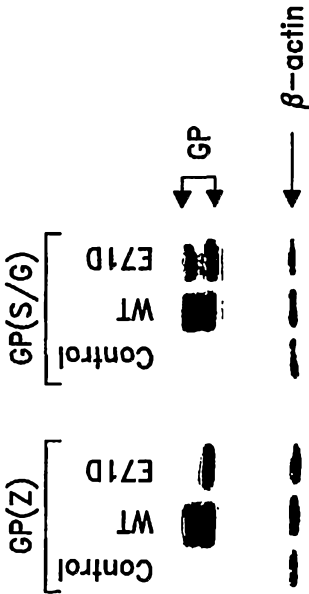
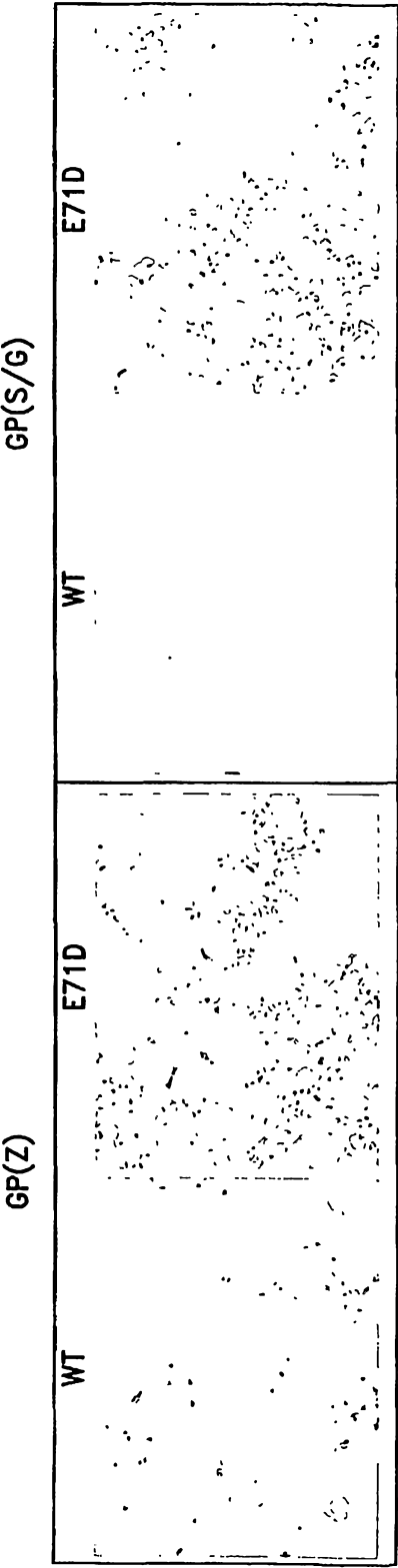


FIG. 14C



39/44

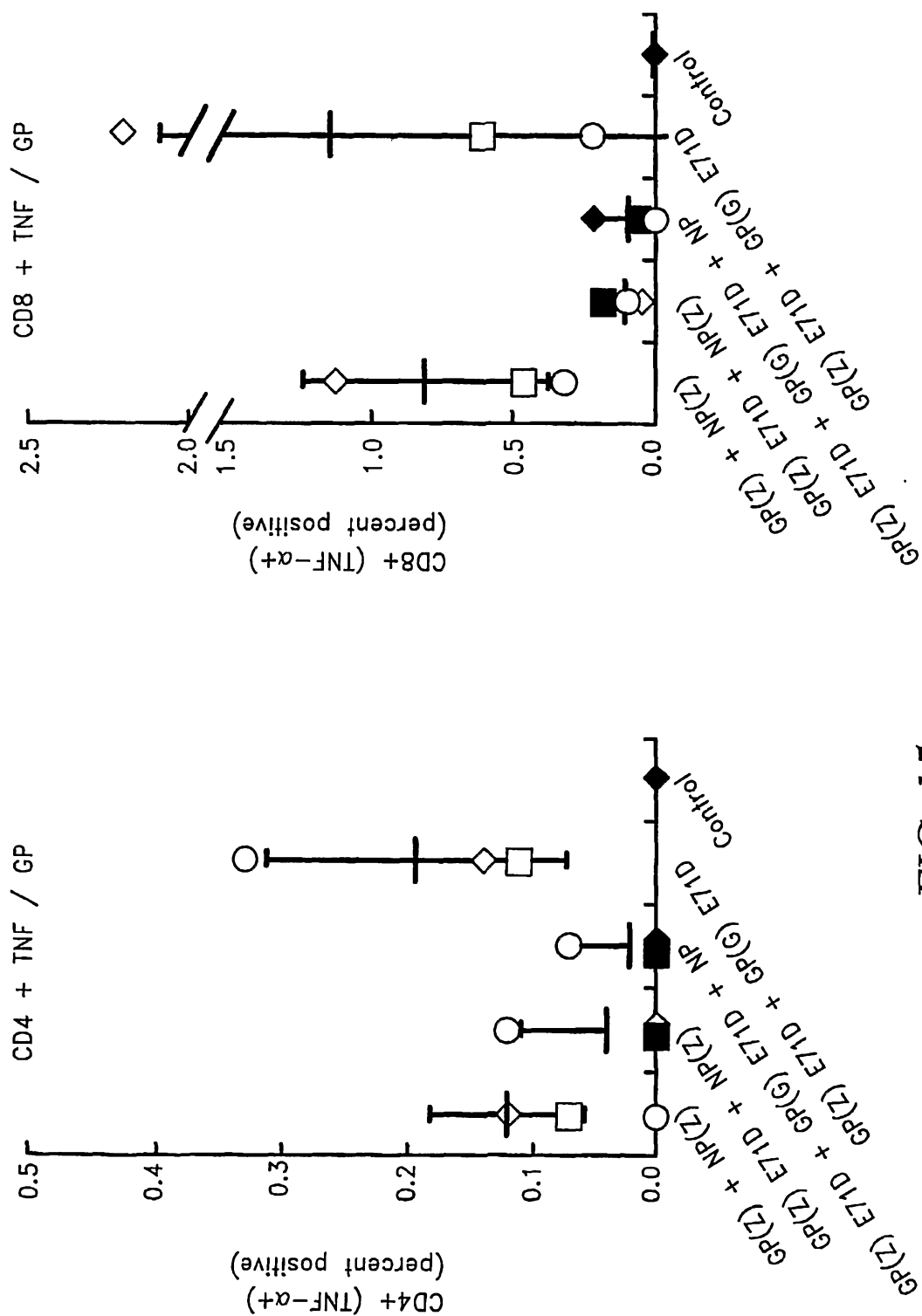


FIG. 15

40/44

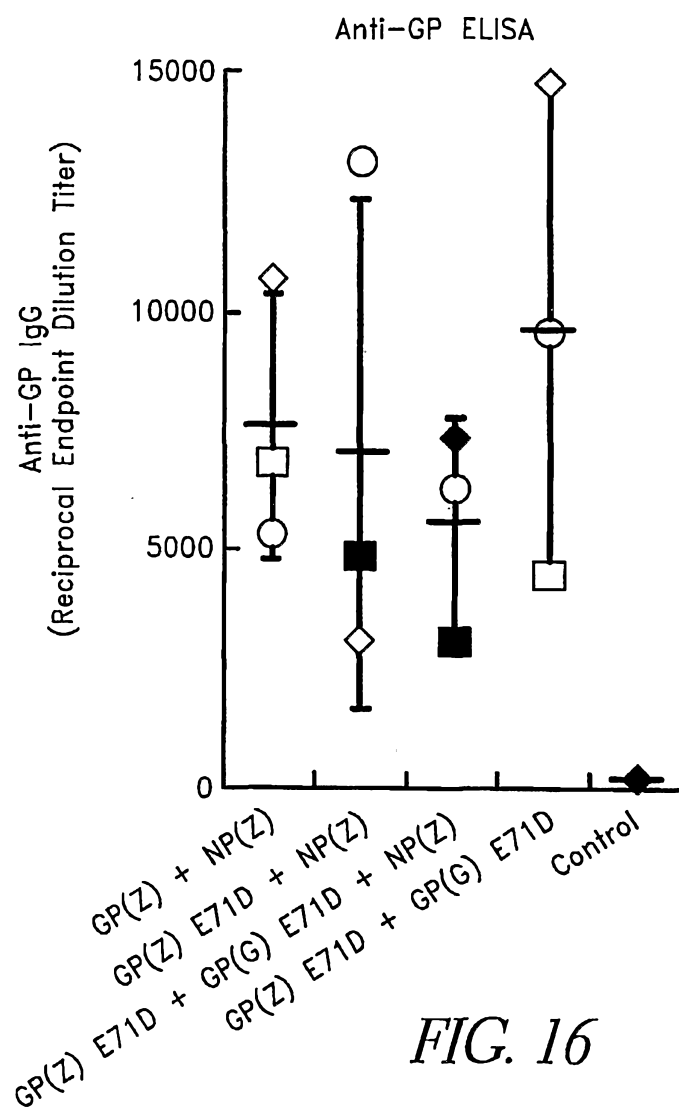


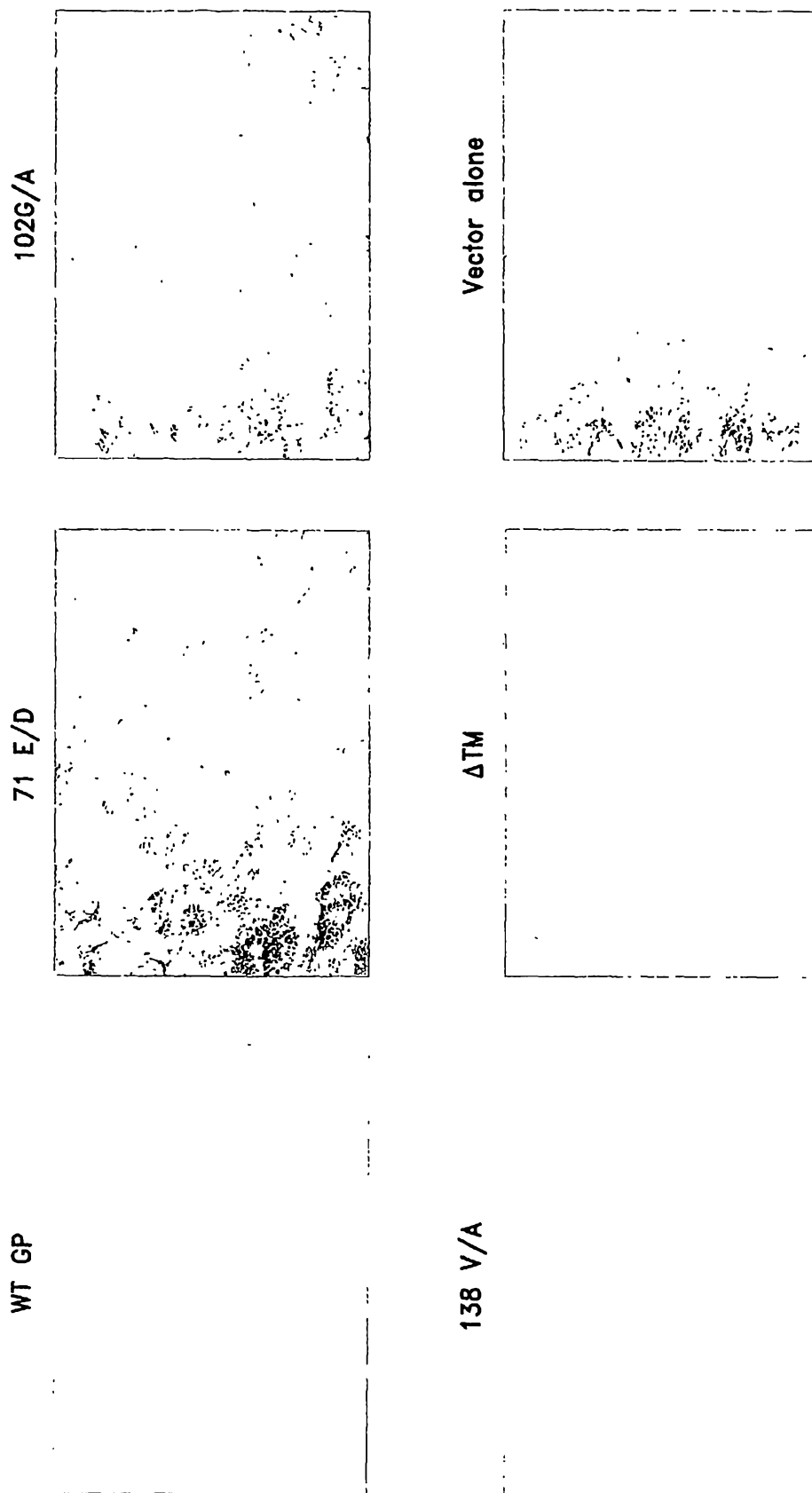
FIG. 16

41/44

Group	Immunogen	Dose (PU)	Survival
1	GP(Z) + NP(Z)	10^{10} each	3/3
2	GP(Z) E71D + NP(Z)	10^{10} each	2/3
3	GP(Z) E71D + GP(G) E71D + NP(Z)	10^{10} each	1/3
4	GP(Z) E71D + GP(G) E71D	10^{10} each	3/3
5	Control	0	0/1

FIG. 17

FIG. 18



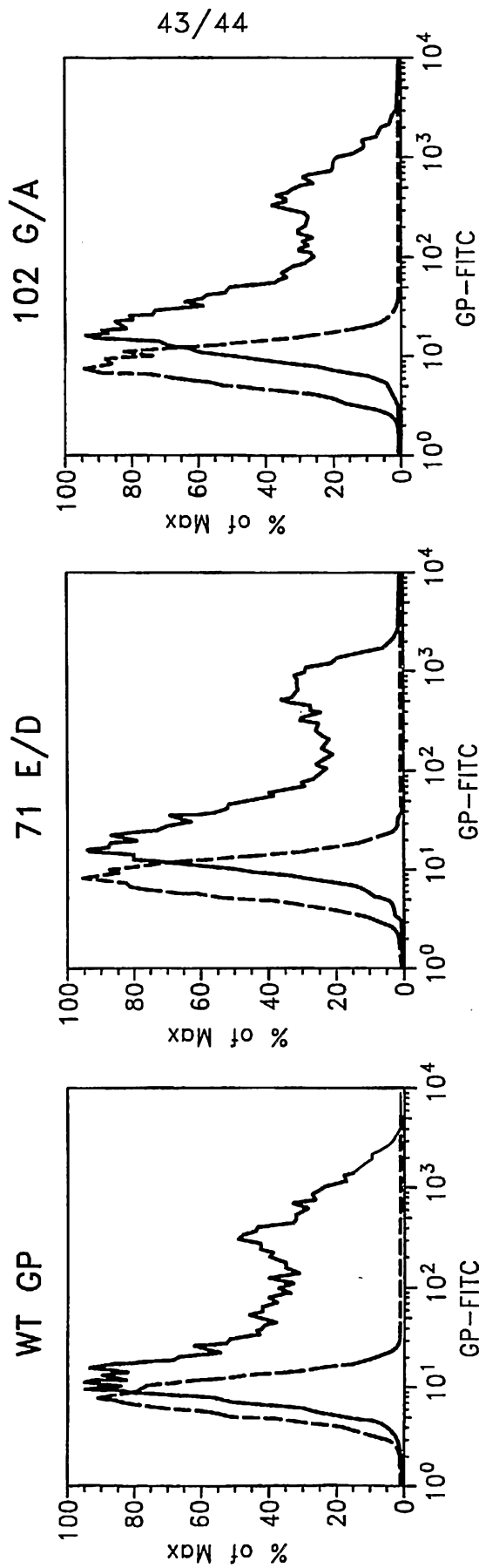


FIG. 19

*FIG. 20***Ebola Zaire Glycoprotein (SEQ ID NO: 9)**

MGVTGILQLPRDRFKRTSFFLWVILFQRTFSIPLGVIHNSTLQVSDVDKLVCRDKLS
STNQLRSVGLNLENGVATDVPSATKRWGFRSGVPPKVNYEAGEWAENCYNLE
IKKPDGSECLPAAPDGIRGFPCRYVHKVSGTGPCAGDFAFHKEGAFFLYDRLAST
VIYRGTTFAEGVVAFLLPQAKKDFSSHPLREPVNATEDPSSGYYSTTIRYQATGF
GTNETEYLFEDNLTYYVQLESRFTPQFLQLNETIYTSGKRSNTTGKLIWKVNPEID
TTIGEWAFWETKKNLTRKIRSEELSFTVVSNGAKNISGQSPARTSSDPGTNTTTEDH
KIMASENSSAMVQVHSQGREAAVSHLTTLATISTSPQSLTTKPGPDNSTHNTVPYK
LDISEATQVEQHRRRTDNDSTASDTPSATTAAAGPPKAENTNTSKSTDFLDPATTTSP
QNHSETAGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNAQ
PKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQLANE
TTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDWTKNITD
KIDQIHDFVDKTLDPQGDNDNWWTGWRQWIPAGIGVTGVVIAVIALFCICKFVF

NIH316.001VPC_SEQLIST.TXT

SEQUENCE LISTING

<110> GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE
 SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES, THE
 CRUCELL HOLLAND B.V.
 Sullivan, Nancy
 Chakrabarti, Bimal
 Yang, Zhi-yong
 Grazia Pau, Maria
 Goudsmit, Jaap
 Nable, Gary J.

<120> OPTIMIZED VACCINES TO PROVIDE PROTECTION
 AGAINST EBOLA AND OTHER VIRUSES

<130> NIH316.001VPC

<150> 60/613,883

<151> 2004-09-27

<150> 60/677,606

<151> 2005-05-03

<150> 60/679,767

<151> 2005-05-10

<150> 60/701,694

<151> 2005-07-22

<150> 60/715,874

<151> 2005-09-09

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 6426

<212> DNA

<213> Ebola virus (Sudan, Gulu)

<400> 1

tgcgcgcttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	gagacgggtca	60
cagcttgctt	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cgccatcaga	gcagattgta	ctgagagtgc	180
accatatgct	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacattttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaattg	420
cccgcctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tgggccgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccct	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagccctca	tcggctcgca	tctctccttc	acgcgcccgc	1020
cgccctacct	gaggccgcca	tccacgccgg	ttgagtcgcg	ttctgccgcc	tcccgcctgt	1080
ggtgcctcct	gaactgcgtc	cgccgtctag	gtaagtttaa	agctcaggtc	gagaccgggc	1140
ctttgtccgg	cgctcccttg	gagcctacct	agactcagcc	ggctctccac	gctttgcctg	1200
accctgcttg	ctcaactcta	gttaacggtg	gagggcagtg	tagtctgagc	agtactcgtt	1260
gctgcccgcg	gcgccaccag	acataatagc	tgacagacta	acagactgtt	cctttccatg	1320
ggtcttttct	gcagtcaccg	tcgtcgacga	tatcgccgcc	atggagggcc	tgagcctgct	1380
gcagctgccc	agggacaagt	tcaggaagag	cagcttcttc	gtgtgggtga	tcactcgtgt	1440

NIH316.001VPC_SEQLIST.TXT

ccagaaggcc	ttcagcatgc	ccctgggctg	ggtgaccaac	agcaccctgg	agggtgaccga	1500
gatcgaccag	ctggtgtgca	aggaccacct	ggccagcacc	gaccagctga	agagcgtggg	1560
cctgaacctg	gagggcagcg	gcgtgagcac	cgacatcccc	agcggccacca	agaggtgggg	1620
cttcaggagc	ggcgtgcctc	ccaaggtggt	gagctacgag	gccgcccaggt	gggcccagaa	1680
ctgctacaac	ctggagatca	agaagcccga	cggcagcgag	tgcttgcctc	ctcctcctga	1740
cggcgtgagg	ggcttcccca	ggtgcaggta	cgtgcacaag	gcccagggca	ccggccccctg	1800
ccccggcgac	tacgccttcc	acaaggacgg	cgcttcttcc	ctgtacgaca	ggctggccag	1860
caccgtgatc	tacaggggcg	tgaacttcgc	cgagggcggtg	atcgcccttcc	tgatcctggc	1920
caagcccaag	gagaccttcc	tgagagcccc	tcccatcagg	gaggccgtga	actacaccga	1980
gaacaccagc	agctactacg	ccaccagcta	tctagagtac	gagatcgaga	acttcggcgc	2040
ccagcacagc	accaccctgt	tcaagatcga	caacaacacc	ttcgtgaggc	tggacaggcc	2100
ccacacccct	cagttcctgt	tccagctgaa	cgacaccatc	cacctgcacc	agcagctgag	2160
caacaccacc	ggcaggctga	tctggaccct	ggacgccaac	atcaacgcgg	acatcggcga	2220
gtgggccttc	tgggagaaca	agaagaacct	gagcgagcag	ctgaggggag	aggagctgag	2280
cttcgaggcc	ctgagcctga	acgagaccga	ggacgacgac	gccgcccagca	gcaggatcac	2340
caagggcagg	atcagcgaca	gggcccaccag	gaagtacagc	gacctggtgc	ccaagaacag	2400
ccccggcatg	gtgcccctgc	acatccccga	ggcgagagac	accctgccca	gccagaacag	2460
caccgagggc	aggaggggtg	gcgtgaacac	ccaggagacc	atcaccgaga	ccgccgccac	2520
catcatcggc	accaacggca	accacatgca	gatcagcacc	atcggcata	ggcccagcag	2580
cagccagatc	cccagcagca	gccccaccac	cgcccctagc	cccgaggccc	agacccccac	2640
cacccacacc	agcggaccca	gcgtgatggc	caccgaggag	cccaccacc	ctcccggcag	2700
cagccccgga	ccacccaccg	aggcccctac	cctgaccacc	cctgagaaca	tcaccaccgc	2760
cgtgaagacc	gtgctgcccc	aggagagcac	cagcaacggc	ctgatcacca	gcaccgtgac	2820
cggcatcctg	ggcagcctgg	gcctgaggaa	gaggagcagg	aggcagacca	acaccaaggc	2880
caccggcaag	tgcaacccca	acctgcacta	ctggaccgcc	caggagcagc	acaacgccgc	2940
cggcatcgcc	tggattccct	acttcggccc	cggcgccgag	ggcatctaca	ccgaggggct	3000
gatgcacaac	cagaacggcc	tgggtgtcgg	cctgaggcag	ctggccaacg	agaccacca	3060
ggccctgcag	ctgttctctga	gggcccaccac	cgagctgagg	acctacacca	tcctgaacag	3120
gaaggccatc	gacttctctgc	tgaggagggtg	ggcgcgacac	tgacaggattc	tgggccccga	3180
ctgctgcatc	gagccccacg	actggaccaa	gaacatcacc	gacaagatca	accagatcat	3240
ccacgacttc	atcgacaacc	ctctgcccac	ccaggacaac	gacgacaact	ggtggaccgg	3300
ctggcggcag	tggatacctg	ccggcatcgg	catcaccggc	atcatcatcg	ccatcatcgc	3360
tctgtctgtc	gtgtgcaagc	tgctgtgtctg	agaattcaga	tctacacgat	ctgctgtgcc	3420
ttctagtgtc	cagccatctg	ttgtttgccc	ctccccctg	ccttccttga	ccctggaagg	3480
tgccactccc	actgtccttt	cctaataaaa	tgaggaaatt	gcategcatt	gtctgagtag	3540
gtgtcattct	attctggggg	gtgggggtgg	agaggacagc	aagggggagg	attgggaaga	3600
caatagcagg	catgctgggg	atgctggggg	ctctatgggt	acccagggtgc	tgaagaattg	3660
acccggttcc	tcctggggcca	gaaagaagca	ggcacatccc	cttctctgtg	acacaccctg	3720
tccacgcccc	tggttcttag	ttccagcccc	actcatagga	cactcatagc	tcaggagggc	3780
tcgcgcttca	atcccaccgc	ctaaagtact	tggagcggtc	tctccctccc	tcacgagccc	3840
accaaaccaa	acctagcctc	caagagtggg	aagaaattaa	agcaagatag	gctattaagt	3900
gcagagggag	agaaaatgcc	tccaacatgt	gaggaagtaa	tgagagaaat	catagaattt	3960
taaggccatg	atttaaggcc	atcatggcct	taatcttccg	cttcctcgct	cactgactcg	4020
ctgcgtctcg	tcgttcgggt	gcggcgagcg	gatcagctc	actcaaaggc	ggtaatacgg	4080
ttatccacag	aatcagggga	taacgcagga	caagaacatgt	gagcaaaaag	gagcaaaaag	4140
gccaggaacc	gtaaaaaggc	cgctttgctg	gcgtttttcc	ataggctccg	ccccctgac	4200
gagcatcaca	aaaatcgacg	ctcaagttag	agggtggcga	acccgacagg	actataaaga	4260
taccaggcgt	ttccccctgg	aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	4320
accgataacc	tgctcgcttc	tctcccttgc	ggaagcggtg	cgctttctca	tagctcacgc	4380
tgtaggtatc	tcagttcggt	gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaaccc	4440
cccgttcagc	ccgaccgctg	cgctttatcc	ggtaactatc	gtcttgagtc	caaccgggta	4500
agacacgact	tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	agcgagggtat	4560
gtaggcggtg	ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	4620
gtatttggtg	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	4680
tgatccggca	aacaaaccac	cgctggtagc	ggtggttttt	ttgtttgcaa	gcagcagatt	4740
acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	4800
cagtggaaacg	aaaactcacg	ttaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	4860
acctgatacc	ttttaaatta	aaaatgaagt	tttaaataca	tctaaagtat	atatgagtaa	4920
acttgggtctg	acagttacca	atgtttaatc	gttgaggcac	ctatctcagc	gatctgtcta	4980
tttcgttcat	ccatagttgc	ctgactcggg	gggggggggc	gctgaggtct	gcctcgtgaa	5040
gaaggtgttg	ctgactcata	ccaggcctga	atcgccccat	catccagcca	gaaagtgagg	5100
gagccacggt	tgatgagagc	tttgtttag	gtggaccagt	tgggtgattt	gaacttttgc	5160
tttggcacgg	aacggtctgc	ggtgtcggga	agatgcgtga	tctgatcctt	caactcagca	5220
aaagttcgat	ttattcaaca	aagccgccgt	cccgtaagt	cagcgtaagt	ctctgccagt	5280
gttacaacca	attaaccaat	tctgattaga	aaaactcatc	gagcatcaaa	tgaactgca	5340
atattattcat	atcaggatta	tcaataccat	atttttgaaa	aagccgtttc	tgtaatgaag	5400
gagaaaactc	accgaggcag	ttccatagga	tggcaagatc	ctggtatcgg	tctgcgattc	5460
cgactcgtcc	aacatcaata	caacctatta	atttcccctc	gtcaaaaaata	aggttatcaa	5520

NIH316.001VPC_SEQLIST.TXT

gtgagaaatc	accatgagtg	acgactgaat	ccggtgagaa	tggcaaaagc	ttatgcattt	5580
ctttccagac	ttgttcaaca	ggccagccat	tacgctcgtc	atcaaaatca	ctcgcacaa	5640
ccaaaccgtt	attcattcgt	gattgcgctt	gagcgagacg	aaatacgcga	tcgctgttaa	5700
aaggacaatt	acaaacagga	atcgaatgca	accggcgagc	gaacactgcc	agcgcatcaa	5760
caatatcttc	acctgaatca	ggatattctt	ctaatacctg	gaatgctgtt	ttcccgggga	5820
tcgcagtggt	gagtaaccat	gcatcatcag	gagtacggat	aaaatgcttg	atggctcgga	5880
gaggcataaa	ttccgtcagc	cagtttagtc	tgaccatctc	atctgtaaca	tcattggcaa	5940
cgctaccttt	gcatgtttc	agaaacaact	ctggcgcatc	gggcttccca	tacaatcgat	6000
agattgtcgc	acctgattgc	ccgacattat	cgcgagccca	tttataccca	tataaatcag	6060
catccatggt	ggaatttaat	cgcgccctcg	agcaagacgt	ttcccgttga	atatggctca	6120
taacaccctt	tgtattactg	tttatgtaag	cagacagttt	tattgttcat	gatgatatat	6180
ttttatcttg	tgcaatgtaa	catcagagat	tttgagacac	aacgtggctt	ttcccccccc	6240
cccattattg	aagcatttat	cagggttatt	gtctcatgag	cggtacata	tttgaatgta	6300
tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	6360
tctaagaaac	cattattatc	atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	6420
ttcgtc						6426

<210> 2
 <211> 6424
 <212> DNA
 <213> Ebola (Zaire)

<400> 2

tcgcgcgttt	cggtgatgac	gggtgaaacc	tctgacacat	gcagctcccg	gagacgggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccc	tcaggggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cgcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagatttg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgcatgttt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggctatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
ccgcctggc	tgaccgcca	acgacccccg	ccattgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaatagggg	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggccgcctt	ggcattatgc	gcagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	accccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctcca	tcggctcgca	tctctccttc	acgcgcccgc	1020
cgccctacct	gaggccgcca	tccacgcgcg	ttgagtcgcg	ttctgccgcc	tcccgcctgt	1080
ggtgcctcct	gaactgcgtc	cgccgtctag	gtaagttaa	agctcaggtc	gagaccgggc	1140
ctttgtccgg	gcctcccttg	gagcctacct	agactcagcc	ggctctccac	gctttgcctg	1200
accctgcttg	ctcaactcta	gttaacggtg	gagggcagtg	tagtctgagc	agtactcgtt	1260
gctgccgcgc	gcgccaccag	acataatagc	tgacagacta	acagactgtt	cctttccatg	1320
ggtcttttct	gcagtcaccg	tcgtcgacga	tatcgccgcc	atgggctgta	ccggcatcct	1380
ccagctgccc	agggacaggt	tcaagaggac	cagcttcttc	ctgtgggtga	tcatctgttt	1440
gcagaggacc	ttcagcatcc	ccctgggctg	gatccacaac	agcaccctga	aggtagcgga	1500
cgtggacaag	ctggtgtgca	gggacaagct	gagcagcacc	aaccagctga	ggagcgtggg	1560
cctgaacctg	gagggcaacg	gcgtggccac	cgacgtgccc	agcgccacca	agaggtgggg	1620
cttcaggagc	ggcgtgcctc	ccaaggtggt	gaactacgag	gccgcccagt	gggcccagaa	1680
ctgctacaac	ctggagatca	agaagcccga	cggcagcgag	tgccctgccg	ccgccccctga	1740
cggcatcagg	ggcttcccca	gggtgcaggt	cgtgcacaag	gtgagcggca	ccggccccctg	1800
cgccggcgac	ttcgcttcc	acaaggagg	cgcttcttc	ctgtacgaca	ggctggccag	1860
caccgtgatc	tacaggggca	ccaccttcgc	cgagggcggtg	gtggccttcc	tgatcctgcc	1920
ccaggccaag	aaggacttct	tcagcagcca	ccctctgag	gagcccgtga	acgccaccga	1980
ggacccccagc	agcggtact	acagcaccac	catcaggtac	caggccaccg	gcttcggcac	2040
caacgagacc	gagtacctgt	tcgaggtgga	caacctgacc	tacgtgcagc	tgaggtctag	2100
attcaccctt	cagttcctgc	tgcagctgaa	cgagaccatc	tacaccagcg	gcaagaggag	2160
caacaccacc	ggcaagctga	tctggaaggt	gaaccccag	atcgacacca	ccatcggcga	2220
gtgggccttc	tgggagacca	agaagaacct	gaccaggaag	atcaggagcg	aggagctgag	2280
cttcaccgtc	gtgagcaacg	gggccaagaa	catcagcggc	cagagcccg	ccaggaccag	2340
cagcgacccc	ggcaccaaca	ccaccaccga	ggaccacaag	atcatggcca	gcgagaacag	2400
cagcgccatg	gtgcaggtgc	acagccagg	cagggaggcc	gccgtgagcc	acctgaccac	2460
cctggccacc	atcagcacca	gccctcagtc	tttaaccacc	aagcccggcc	ccgacaacag	2520
caccacaaac	accctgtgtg	acaagctgga	catcagcgag	gccaccagg	tggagcagca	2580

NIH316.001VPC_SEQLIST.TXT

ccacaggagg	accgacaacg	acagcaccgc	cagcgacacc	ccttccgcc	ccaccgccgc	2640
cggccctccg	aaggccgaga	acaccaacac	cagcaagagc	accgactttc	tggatcccgc	2700
caccaccacc	agccctcaga	accacagcga	gaccgccggc	aacaacaaca	cccaccacca	2760
ggacaccggc	gaggagagcg	ccagcagcgg	caagctgggc	ctgatcacca	acaccatcgc	2820
cggcgtggcc	ggcctgatca	ccggcggcag	gaggaccagg	agggaggcca	tcgtgaacgc	2880
ccagcccaag	tgcaacccca	acctgcacta	ctggaccacc	caggacgagg	gcgcccgcct	2940
cggcctggcc	tggattccct	acttcggccc	cgccgccgag	ggcatctaca	tcgagggcct	3000
gatgcacaac	caggacggcc	tgatctgcgg	cctgaggcag	ctggccaacg	agaccaccca	3060
ggccctgcag	ctgttcctga	gggccaccac	cgagctgagg	accttcagca	tcctgaacag	3120
gaaggccatc	gacttcctgc	tgcaagagtg	ggcgccgacc	tgccacatcc	tgggccccga	3180
ctgctgcac	gagccccacg	actggacca	gaacatcacc	gacaagatcg	accagatcat	3240
ccacgacttc	gtggacaaga	ccctgccaga	ccagggcgac	aacgacaact	ggtggaccgg	3300
ctggcggcag	tggtatccctg	ccggcatcgg	ctgacccggc	gtggtgatcg	ccgtgatcgc	3360
tctgttctgc	atctgcaagt	tcgtgttctg	aacacgtgga	attcagatct	gctgtgcctt	3420
ctagttgcc	gccatctgtt	gtttgccctt	cccccgctgc	ttccttgacc	ctggaagggtg	3480
ccactcccac	tgctctttcc	taataaaatg	aggaaattgc	atcgcatgtt	ctgagtaggt	3540
gtcattctat	tctggggggg	gggggtgggg	aggacagcaa	gggggaggat	tgggaagaca	3600
atagcaggca	tgctggggat	gcgggtgggt	ctatgggtac	ccaggtgctg	aagaattgac	3660
ccggttcctc	ctgggccaga	aagaagcagg	cacatcccct	tctctgtgac	acaccctgtc	3720
cacgcccctg	gttcttagtt	ccagccccac	tcataggaca	ctcatagctc	aggagggctc	3780
cgcttcaat	cccacccgct	aaagtacttg	gagcggcttc	tccctccctc	atcagcccac	3840
caaaccaaac	ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	tattaagtgc	3900
agagggagag	aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	tagaatttta	3960
aggccatgat	ttaaggccat	catggcctta	atcttccgct	tcctcgctca	ctgactcgct	4020
gcgctcggtc	gttcggctgc	ggcgagcggg	atcagctcac	tcaaaggcgg	taatacgggt	4080
atccacagat	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	4140
caggaaacgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	4200
gcatacacia	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	4260
ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgacct	tgccgcttac	4320
cggtatccctg	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	4380
taggtatctc	agttcggtgt	aggtcggttc	ctccaagctg	ggctgtgtgc	acgaaccccc	4440
cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	4500
acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	cgaggatagt	4560
aggcgggtgt	acagagttct	tgaagtgggt	gcctaactac	ggctacacta	gaagaacagt	4620
atcttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	4680
atccggcaaa	caaaaccacg	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	4740
gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	4800
gtggaacgaa	aactcacggt	aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	4860
ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	4920
ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	4980
tcgttcattc	atagttgcct	gactcggggg	ggggggggcg	tgaggctctg	ctcgtgaaga	5040
agggtgtgct	gactcatacc	aggcctgaat	cgccccatca	tccagccaga	aagtgaggga	5100
gccacggttg	atgagagctt	tggtgtaggt	ggaccagttg	gtgattttga	acttttgctt	5160
tgccacggaa	cggtctgcgt	tgctcggaag	atgcgtgatc	tgatccttca	actcagcaaa	5220
agttcgattt	attcaacaaa	gccgccgtcc	cgtaagtcga	gcgtaatgct	ctgccagtgt	5280
tacaaccaat	taaccaattc	tgattagaaa	aactcatcga	gcatcaaatg	aaactgcaat	5340
ttattcatat	caggattatc	aataccatat	tttgaaaaaa	gccgtttctg	taatgaagga	5400
gaaaactcac	cgaggcagtt	ccataggatg	gcaagatcct	ggtatcgggtc	tgcgattccg	5460
actcgtccaa	catcaataca	acctattaat	ttccctcgt	caaaaataag	gttatcaagt	5520
gagaaatcac	catgagtgac	gactgaatcc	ggtgagaatg	gcaaaagctt	atgcattttc	5580
ttccagactt	gttcaacagg	ccagccatta	cgctcgtcat	caaaatcact	cgcatcaacc	5640
aaaccggtat	tcattcgtga	ttgcccctga	gcgagacgaa	atacgcgatc	gctgttaaaa	5700
ggacaattac	aaacaggaat	cgaatgcaac	cggcgaggga	acactgccag	cgcatcaaca	5760
atattttcac	ctgaatcagg	atattcttct	aatacctgga	atgctgtttt	cccggggatc	5820
gcagtggtga	gtaaccatgc	atcatcagga	gtacggataa	aatgcttgat	ggtcgggaaga	5880
ggcataaatt	ccgtcagcca	gtttagtctg	accatctcat	ctgtaacatc	attggcaacg	5940
ctacctttgc	catgtttcag	aaacaactct	ggcgcatcgg	gcttcccata	caatcgatag	6000
attgtcgcac	ctgattgccc	gacattatcg	cgagcccatt	tatacccata	taaatcagca	6060
tccatgttgg	aatttaacgc	cggcctcgag	caagacggtt	cccgttgat	atggctcata	6120
acacccttgc	tattactgtt	tatgtaagca	gacagtttta	ttgttcatga	tgatatattt	6180
ttatcttgtg	caatgtaaca	tcagagattt	tgagacacaa	cgtggctttc	ccccccccc	6240
cattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	6300
tagaaaaata	aacaaatagg	ggttccgcgc	acatttccc	gaaaagtgc	acctgacgtc	6360
taagaaacca	ttattatcat	gacattaacc	tataaaaaata	ggcgtatcac	gaggcccttt	6420
cgtc						6424

NIH316.001VPC_SEQLIST.TXT

<210> 3
 <211> 6414
 <212> DNA
 <213> Ebola virus (Sudan, Gu1u)

<400> 3
 tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcggggtg 120
 ttggcggggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
 accatatgag gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
 tccaacatta ccgcatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
 ggggtcatta gtccatagcc catatatgga gttccgctt acataactta cggtaaatgg 420
 cccgcctggc tgaccgccc acgaccccc cccattgacg tcaataatga cgtatgttcc 480
 catagtaacg ccaatagggg ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
 tgcccacttg cgagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600
 tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
 ttggcagtag atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
 ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
 agctcgttta gtgaaccgtc agatcgccgt gagacgccat ccacgctgtt ttgacctcca 960
 tagaagacac cgggaccgat ccagcctcca tcggctcgca tctctccttc acgcgcccgc 1020
 cgccctacct gaggccgcca tccacgcccg ttgagtcgag ttctgcccgc tccgcctgt 1080
 ggtgcctcct gaactgcgtc cgccgtctag gtaagttaa agctcaggtc gagaccgggc 1140
 ctttgcctgg cgctcccttg gagcctaccc agctcagcc ggctctccac actttgcctg 1200
 accctgcttg ctcaactcta gttaacgggt gagggcagtg tagtctgagc agtactcgtt 1260
 gctgcccgcg gcgccaccag acataatagc tgacagacta acagactgtt cctttccatg 1320
 ggtcttttct gcagtcaccg tcgtcgacga tatcgccgcc atggaggggc tgagcctgct 1380
 ccagctgccc agggacaagt tcaggaagag cagcttcttc gtgtgggtga tcatcctgtt 1440
 cgagaaggcc ttcatgcatg ccctgggctg ggtgaccaac agcaccctgg aggtgaccga 1500
 gatcgaccag ctggtgtgca aggaccacct ggccagcacc gaccagctga agagcgtggg 1560
 cctgaacctg gacggcagcg gcgtgagcac cgacatcccc agcgccacca agaggtgggg 1620
 cttcaggagc ggcgtgcctc ccaaggtggg gagctacgag gccggcgagt gggccgagaa 1680
 ctgtcataac ctggagatca agaagcccga cggcagcgag tgcctgcctc ctctcctga 1740
 cggcgtgagg ggttcccca ggtgcaggta cgtgcacaag gccaggggca ccggccccctg 1800
 ccccggcgac tacgccttcc acaaggacgg cgcttcttc ctgtacgaca ggctggccag 1860
 caccgtgatc tacaggggag tgaacttcgc cgagggcggt atcgcccttc tgatcctggc 1920
 caagcccaag gagaccttcc tgcagagccc tcccatcagg gaggccgtga actacaccga 1980
 gaacaccagc agctactacg ccaccagcta tctagagtac gagatcgaga acttcgggcg 2040
 ccagcacagc accacctgtt tcaagatcga caacaacacc ttcgtgaggc tggacaggcc 2100
 ccacaccctt cagttcctgt tccagctgaa cgacaccatc cacctgcacc agcagctgag 2160
 caacaccacc ggcaggctga tctggaccct ggacgccaac atcaacgccg acatcggcga 2220
 ttgggccttc tgggagaaca agaagaacct ctgagggcag ctgaggggag agggagctgag 2280
 cttcgaggcc ctgagcctga acgagaccga ggacgacgac gccgccagca gcaggatcac 2340
 caagggcagg atcagcgaca gggccaccag gaagtacagc gacctgggtc ccaagaacag 2400
 ccccggcagc gtgcccctgc acatccccga gggcgagacc accctgccc gccagaacag 2460
 caccgagggc agggaggtgg gcgtgaacac ccaggagacc atcaccgaga ccgcccacc 2520
 catcatcgcc accaacggca accacatgca gcctgacacc atcggcacac ggccaccagc 2580
 cagccagatc cccagcagca gccccaccac cggccctagc cccgaggccc agacccccac 2640
 caccacacc agcggaacca gcgtgatggc caccgaggag cccaccacc ctcccggcag 2700
 cagccccgga cccaccaccg agggccctac cctgaccacc cctgagaaca tcaccaccgc 2760
 cgtgaagacc gtgctgcccc aggagagcac cagcaacggc ctgatcaca gcaccgtgac 2820
 cggcatcctg ggcagcctgg gcctgaggaa gagggagcagg aggcagacca acaccaaggc 2880
 caccggcaag tgcaacccca acctgcacta ctggaccgcc caggagcagc acaacgccgc 2940
 cggcatcgcc tggattccct acttcggccc cggcgccgag ggcattctaca ccgagggcct 3000
 gatgcacaac cagaacgccc tgggtgtgagg cctgaggcag ctggccaacg agaccacca 3060
 ggccctgcag ctgttcctga gggccaccac cggagctgag accctacaca tcctgaacag 3120
 gaaggccatc gacttcctgc tgaggagggt ggcggccacc tgcaggattc tgggccccga 3180
 ctgctgcatc gagccccacg actggaccaaa gaacatcacc gacaagatca accagatcat 3240
 ccacgacttc atcgacaacc ctctgcccac ccaggacaac gacgacaact ggtggaccgg 3300
 ctggcggcag tggatacctg ccggcatcgg catcacgggc atcatcatcg ccatcatcgc 3360
 tctgtgtgct ggtgtgcaag tctgtgtgct aacacgatct gctgtgcctt ctggttccca 3420
 gccatctgtt ttttgccctt ccccggtgcc ttccttgacc ctggaagggt ccactcccac 3480
 tgtcctttcc taataaaatg aggaattgc atcgattgt ctgagtaggt gtcattctat 3540
 tctggggggg ggggtggggc aggcagcaa gggggaggat tgggaagaca atagcaggca 3600
 tgctggggat gcggtgggct ctatgggtac ccaggtgctg aagaattgac ccggttcctc 3660
 ctgggcccaga aagaagcagg cacatccctt tctctgtgac acaccctgtc cagccccctg 3720

NIH316.001VPC_SEQLIST.TXT

gttcttagtt	ccagccccac	tcataggaca	ctcatagctc	aggagggctc	cgcccttcaat	3780
cccacccgct	aaagtacttg	gagcgggtctc	tccctccctc	atcagcccac	caaaccaaac	3840
ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	tattaagtgc	agagggagag	3900
aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	tagaatttta	aggccatgat	3960
ttaaggccat	catggcctta	atcttccgct	tcctcgcctca	ctgactcgct	gcgctcggct	4020
gttcggctgc	ggcgaagcgt	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	4080
tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaaccgt	4140
aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	4200
aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	4260
ccccctggaa	gctccctcgt	gcgctctcct	gttcgcgacc	tgccgcttac	cggtacacct	4320
tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	gtcacgctg	taggtatctc	4380
agttcgggtg	aggtcgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cggtcagccc	4440
gacgcgtgcg	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	4500
tcgccactgg	cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	aggcgggtgt	4560
acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	gaagaacagt	atttgggtatc	4620
tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	4680
caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	4740
aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	gtggaacgaa	4800
aactcacggt	aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	4860
ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	ttgggtctgac	4920
agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcgttcatcc	4980
atagttgcct	gactcggggg	ggggggggcg	tgaggctctg	ctcgtgaaga	aggtgtttgct	5040
gactcatacc	aggcctgaat	cgccccatca	tccagccaga	aagtgagggg	gccacgggtg	5100
atgagagctt	tggttagagg	ggaccagttg	gtgattttga	acttttgctt	tgccacggaa	5160
cggtctgcgt	tgctgggaag	atgcgtgatc	tgatccttca	actcagcaaa	agttcgaatt	5220
attcaacaaa	gccgcgctcc	cgtaagtca	gcgtaagtct	ctgccagtgt	tacaaccaat	5280
taaccaattc	tgattagaaa	aactcatcga	gcatcaaatg	aaactgcaat	ttattcatat	5340
caggattatc	aataccatat	ttttgaaaaa	gccgtttctg	taatgaagga	gaaaactcac	5400
cgaggcagtt	ccataggatg	gcaagatcct	ggtatcggtc	tgcgattccg	actcgtccaa	5460
catcaatata	acctattaat	ttccctcgt	caaaaataag	gttatcaagt	gagaaatcac	5520
catgagtgc	gactgaatcc	ggtagaagt	gcaaaaagctt	atgcatttct	ttccagactt	5580
gttcaacagg	ccagccatta	cgctcgtcat	caaaatcact	cgcatcaacc	aaaccgttat	5640
tcattcgtga	ttgcgcctga	gcgagacgaa	atacgcgctc	gctgttaaaa	ggacaattac	5700
aaacaggaat	cgaatgcaac	cggcgcagga	acactgccag	cgcatcaaca	atattttcac	5760
ctgaatcagg	atattcttct	aatacctgga	atgctgtttt	cccggggatc	gcagtgggtg	5820
gtaaccatgc	atcatcagga	gtacggataa	aatgctttgat	ggctcggaaga	ggcataaatt	5880
ccgtcagcca	gtttagtctg	accatctcat	ctgtaacatc	attggcaacg	ctacctttgc	5940
catgtttcag	aaacaactct	ggcgcacatg	gcttcccata	caatcgatag	attgtcgcac	6000
ctgattgccc	gacattatcg	cgagcccat	tatacccata	taaatcagca	tccatgtttg	6060
aatttaaatcg	cggcctcgag	caagacgttt	cccgttgaat	atggctcata	acaccccttg	6120
tattactggt	tatgtaagca	gacagtttta	ttgttcatga	tgatatattt	ttatctttgtg	6180
caatgtaaca	tcagagattt	tgagacacaa	cggtggcttt	ccccccccc	cattattgaa	6240
gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	6300
aaacaaatag	ggttccgcgc	acatttcccc	gaaaagtgc	acctgacgtc	taagaaacca	6360
ttattatcat	gacattaacc	tataaaaaa	ggcgtatcac	gaggcccttt	cgctc	6414

<210> 4
 <211> 6424
 <212> DNA
 <213> Ebola virus (Zaire)

<400> 4						
tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	gagacgggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cgcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	cgcacagatg	gcgtaaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tacataatag	tacattttata	ttggctcatg	300
tccaacatta	ccgcatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggctatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
cccgcctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	cgtatgtttc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tcaggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaggt	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggccgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatcggtt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccatttga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacggggac	tttccaaaat	gtcgtatacaa	840

NIH316.001VPC_SEQLIST.TXT

ctccgccc	ttgacgcaa	tgggcggtag	gcgtgtacg	tgggaggtct	atataagcag	900
agctcgtt	gtgaaccgt	agatcgctg	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagac	cgggaccgat	ccagcctcca	tcggtctgca	tctctccttc	acgcgcccgc	1020
cgccctac	gaggccgcca	tccacgccc	ttgagtcg	ttctgcccgc	tcccgctgt	1080
ggtgcctc	gaactgcgt	cgccgtctag	gtaagttaa	agctcaggtc	gagaccgggc	1140
ctttgtcc	cgctcccttg	gagcctacct	agactcagcc	ggctctccac	gctttgcctg	1200
acctgcttg	ctcaactcta	gttaacgggt	gagggcagtg	tagtctgagc	agtactcggt	1260
gctgccgc	gcgcaccag	acataatag	tgacagacta	acagactgtt	cctttccatg	1320
ggtctttt	gcagtcaccg	tcgtcgacga	tatcgccgcc	atgggcgtga	ccggcatcct	1380
gcagctgcc	agggacaggt	tcaagaggac	cagcttcttc	ctgtgggtga	tcatcctgtt	1440
ccagaggac	ttcagcatcc	ccctgggctg	gatccacaac	agcaccctgc	aggtgagcga	1500
cgtggaca	ctgggtgtga	gggacaagct	gagcagcacc	aaccagctga	ggagcgtggg	1560
cctgaacct	gacggcaacg	gcgtggccac	cgagctgccc	agcgccacca	agaggtgggg	1620
cttcaggag	ggcgtgcctc	ccaaggtgg	gaactacgag	gccggcgagt	gggccgagaa	1680
ctgctaca	ctggagatca	agaagcccga	cggcagcgag	tgcttggccc	ccgcccctga	1740
cgccatcag	ggcttcccca	ggtgcaggt	cgtgcacaag	gtgagcggca	ccggccccctg	1800
cgccggcg	ttcgcttctc	acaaggagg	cgcttcttct	ctgtacgaca	ggctggccag	1860
caccgtgat	tacaggggca	ccaccttcgc	cgagggcg	gtggccttcc	tgatcctgcc	1920
ccaggcca	aaggacttct	tcagcagcca	ccctctgagg	gagcccgtga	acgccaccga	1980
ggaccccag	agcggtact	acagcaccac	catcaggtac	caggccaccg	gcttcggcac	2040
caacgagac	gagtacctgt	tcgaggtgga	caacctgacc	tacgtgcagc	tgaggtctag	2100
attcaccc	cagttcctgc	tcagctgaa	cgagaccatc	tacaccagcg	gcaagaggag	2160
caacaccac	ggcaagctga	tctggaaggt	gaaccccag	atcgacacca	ccatcggcga	2220
gtgggcctt	tgggagacca	agaagaacct	gaccaggaag	atcaggagcg	aggagctgag	2280
cttcaccgt	gtgagcaacg	gggccaagaa	catcagcggc	cagagccccg	ccaggaccag	2340
cagcgaccc	ggcaccaca	ccaccaccga	ggaccacaag	atcatggcca	cgagagagag	2400
cagcgccat	gtgcaggtgc	acagccaggg	caggggaggc	gccgtgagcc	acctgaccac	2460
cctggccacc	atcagcacca	gccctcagtc	tttaaccacc	aagcccggcc	ccgacaacag	2520
caccacaac	accctgtgt	acaagctgga	catcagcgag	gccaccacag	tgagagcagca	2580
ccacaggag	accgacaacg	acagcaccgc	cagcgacacc	ccttccgcca	ccaccgccc	2640
cgccctccg	aaggccgaga	acaccaacac	cagcaagagc	accgactttc	tgatccccgc	2700
caccaccac	agccctcaga	accacagcga	gaccgcccgc	aacaacaaca	cccaccacca	2760
ggacaccgg	gaggagagcg	ccagcagcgg	caagctgggc	ctgatcacca	acaccatcgc	2820
cgccgtggc	ggcctgatca	ccggcggcag	gaggaccagg	agggaggcca	tcgtgaaccg	2880
ccagccaa	tgaacccca	acctgcacta	ctggaccacc	caggacgagg	gcgcggccat	2940
cgccctggc	tggattccct	acttcggccc	cgccgcccag	ggcatctaca	tcgaggccct	3000
gatgcaca	caggacggcc	tgatctgcgg	cctgaggcag	ctggccaacg	agaccacca	3060
ggccctgc	ctgttcctga	gggccaccac	cgagctgagg	accttcagca	tcctgaacag	3120
gaaggccat	gacttcctgc	tgcaaggtg	ggcgccgacc	tgccacatcc	tgggccccga	3180
ctgctgcac	gagccccacg	actggacca	gaacatcacc	gacaagatcg	accagatcat	3240
ccacgactt	gtggacaaga	ccctgccaga	ccagggcgac	aacgacaact	ggtggaccgg	3300
ctggcgag	tgataacctg	ccggcatcgg	cgtgaccggc	gtggtgatcg	ccgtgatcgc	3360
tctgttctg	atctgcaagt	tcgtgttctg	aacacgtgga	attcagatct	gctgtgcctt	3420
ctagtgtcc	gccatctgtt	gtttgccctc	ccccgtgcc	ttccttgacc	ctggaagggt	3480
ccactccac	tgtcctttcc	taataaaatg	aggaaattgc	atcgcatatt	ctgagtaggt	3540
gtcattctat	tctggggggg	ggggtggggc	aggacagcaa	gggggaggat	tggaagaca	3600
atagcagg	tgctggggat	gcggtgggct	ctatgggtac	ccaggtgctg	aagaattgac	3660
ccggttcct	ctggggccaga	aagaagcagg	cacatcccc	tctctgtgac	acaccctgtc	3720
cagccccct	gttcttagtt	ccagccccac	tcataggaca	ctcatagctc	aggagggtc	3780
cgcttcaat	cccacccgct	aaagtacttg	gagcgggtct	tccctccctc	atcagccac	3840
caaaccaac	ctagcctcca	agagtgggaa	gaaattaaag	caagataggc	tattaagtgc	3900
agaggagag	aaaatgcctc	caacatgtga	ggaagtaatg	agagaaatca	tagaatttta	3960
aggccatgat	ttaaggccat	catggcctta	atcttccgct	tcctcgctca	ctgactcgct	4020
gcgctcggt	gttcggctgc	ggcgagcgg	atcagctcac	tcaaaggcgg	taatacgggt	4080
atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	4140
caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	4200
gcatacaca	aatcgacgct	caagtacag	gtggcgaaac	ccgacaggac	tataaagata	4260
ccaggcggtt	ccccctggaa	gctccctcgt	cgctctcctc	gttccgaccc	gtccgcttac	4320
cgagatac	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	4380
taggtatctc	agttcggtgt	aggtcggtcg	ctccaagctg	ggctgtgtgc	acgaaccccc	4440
cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	4500
acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	4560
aggcggtgt	acagagttct	tgaagtgggt	gcctaactac	ggctacacta	gaagaacagt	4620
atttgggtat	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	4680
atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	4740
gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggt	ctgacgctca	4800
gtggaacgaa	aactcacgtt	aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	4860
ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	4920

NIH316.001VPC_SEQLIST.TXT

ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	4980
tcgttcaccc	atagttgcct	gactcggggg	gggggggagc	tgaggctctg	ctcgtgaaga	5040
aggtgttgct	gactcatacc	aggcctgaat	cgccccatca	tccagccaga	aagttaggga	5100
gccacggttg	atgagagcct	tggtgttagt	ggaccagttg	gtgattttga	acttttgctt	5160
tgccacggaa	cggtctgcgt	tgtcgggaag	atgctgtatc	tgatccttca	actcagcaaa	5220
agttcgatct	attcaacaaa	gccgcccgtc	cgtcaagtca	gcgtaatgct	ctgccagtgt	5280
tacaaccaat	taaccaattc	tgattagaaa	aactcatcga	gcatacaatg	aaactgcaat	5340
ttattcatat	caggattatc	aataccatat	ttttgaaaaa	gccgtttctg	taatgaagga	5400
gaaaactcac	cgaggcagtt	ccataggatg	gcaagatcct	ggtatcggtc	tgcgattccg	5460
actcgtccaa	catcaatata	acctattaat	ttccctcgt	caaaaataag	gttatcaagt	5520
gagaaatcac	catgagtgc	gactgaatcc	ggtgagaatg	gcaaaagctt	atgcatttct	5580
ttccagactt	gttcaacagg	ccagccatta	cgctcgtcat	caaaatcact	cgcatacaac	5640
aaaccgttat	tcattcgtga	ttgcgcctga	gcgagacgaa	atacgcgatc	gtgttataaa	5700
ggacaattac	aaacaggaat	cgaatgcaac	cggcgcagga	acactgccag	cgcatacaac	5760
atattttcac	ctgaatcagg	atattcttct	aatacctgga	atgctgtttt	cccggggatc	5820
gcagtgggtg	gtaaccatgc	atcatcagga	gtacggataa	aatgcttgat	ggtcgggaag	5880
ggcataaaatt	ccgtcagcca	gtttagtctg	accatctcat	ctgtaacatc	attggcaacg	5940
ctacctttgc	catgtttcag	aaacaactct	ggcgcatcgg	gcttcccata	caatcgatag	6000
attgtcgcac	ctgattgccc	gacattatcg	cgagccatt	tatacccata	taaatcagca	6060
tccatgtttg	aatttaaatcg	cggcctcgag	caagacgttt	cccgttgaat	atggctcata	6120
acaccccttg	tattactggt	tatgtaagca	gacagtttta	ttgttcatga	tgatataatt	6180
ttatctttgtg	caatgtaaca	tcagagattt	tgagacacaa	cgtggctttc	ccccccccc	6240
cattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	6300
tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	6360
taagaaacca	ttattatcat	gacattaacc	tataaaaaata	ggcgtatcac	gaggcccttt	6420
cgctc						6424

<210> 5
 <211> 6458
 <212> DNA
 <213> Marburg virus (Angola)

<400> 5

tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccc	gagacgggtca	60
cagcttgctt	gtaagcggat	gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cgcatcaga	gcagattgta	ctgagagtgc	180
accatatgct	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataatatg	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
cccgcctggc	tgaccgcccc	acgacccccg	ccatttgacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaag	acgcccccta	ttgacgtcaa	600
tgacggtaaa	tggcccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	660
ttggcagtag	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccccatta	780
cgtcaatggg	agtttggttt	ggcaccaaaa	tcaacgggac	tttccaaaaa	gtcgtatacaa	840
ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggagggtc	atataagcag	900
agctcgttta	gtgaaccgtc	agatcgccct	gagacgccat	ccacgctgtt	ttgacctcca	960
tagaagacac	cgggaccgat	ccagcctcca	tcggctcgca	tctctccttc	acgcgcccgc	1020
ggccctacct	gaggccgcca	tccacgccgg	ttgagtcgcg	ttctgcccgc	tcccgcctgt	1080
ggtgcctcct	gaactgcgtc	cgccgtctag	gtaagttaa	agctcagggt	gagaccgggc	1140
ctttgtccgg	cgctcccttg	gagcctacct	agactcagcc	ggctctccac	gctttgcctg	1200
accctgcttg	ctcaactcta	gttaacgggt	gagggcagtg	tagtctgagc	agtactcggt	1260
gctgcccgcg	gcgccaccag	acataatagc	tgacagacta	acagactggt	cctttccatg	1320
ggtcttttct	gcagtcaccg	tcgtcgacac	tggtgatcag	atatcgcggc	cgctctagag	1380
atatcgccgc	catgaagacc	acctgcctgc	tgatcagcct	gatcctgatc	cagggcgtga	1440
agaccctgcc	catcctggag	atcgccagca	acatccagcc	ccagaacgtg	gacagcgtgt	1500
gcagcggcac	cctgcagaag	accgaggacg	tgacactgat	gggcttcacc	ctgagcggcc	1560
agaagggtgc	cgacagccct	ctggaggcca	gcaagagggt	ggccttcagg	gccggcggtg	1620
cccccaagaa	cgtggagtac	accgagggcg	aggaggccaa	gacctgctac	aacatcagcg	1680
tgaccgaccc	cagcggcaag	agcctgctgc	tggaccctcc	caccaacatc	agggactacc	1740
ctaagtgcaa	gaccatccac	cacatccagg	gccagaacct	tcacgcccag	ggcatcgccc	1800
tgacactgtg	gggcgccttc	ttcctgtacg	acaggatcgc	cagcaccacc	atgtacagag	1860
gaaaagtgtt	cacagagggg	aacatcgctg	ctatgatcgt	gaacaagacc	gtgcataaga	1920
tgatcttcag	cagacagggg	cagggatata	gacatatgaa	cctgacatcc	acaaacaagt	1980

NIH316.001VPC_SEQLIST.TXT

actggacaag	cagcaacgga	acacagacaa	acgatacagg	atgtttttgga	acactgcagg	2040
aataacaactc	caccaagaac	cagacatgtg	cccctagcaa	gaagcctctg	cctctgccta	2100
cagctcatcc	tgaagtgaag	ctgacatcca	caagcacaga	tgccacaaaag	ctgaacacaa	2160
cagatcctaa	tagcgacgac	gaggatctga	caacaagcgg	atccggatcc	ggagaacagg	2220
aaccttatac	aacaagcgac	gctgctacaa	aacagggact	gtcctccaca	atgcctccta	2280
cacctagccc	tcagcctagc	acacctcagc	agggaggcaa	caacacaaaac	cattcccagg	2340
gagtgggtgac	agaacctgga	aagacaaaac	caacagccca	gcctagcatg	cctcctcata	2400
acacaacaac	aatcagcaca	aacaacacct	ccaagcacaa	tctgagcaca	cctagcgtgc	2460
ctattcagaa	tgccaccaac	tacaacacac	agtccacagc	ccctgaaaac	gaacagacct	2520
ccgcccccttc	caaaaacaacc	ctgctgccta	cagaaaaccc	tacaacagcc	aagagcacia	2580
acagcacaaa	gagccctaca	acaacagtgc	ctaacacaa	aaacaagtat	agcacaagcc	2640
ctagccctac	acctaattcc	acagctcagc	atctggtgta	ttttagaaga	aagagaaaac	2700
tcctgtggag	agaaggagat	atgttccctt	ttctggatgg	actgatcaac	gtccttatcg	2760
atthttgatcc	tgtgcctaac	acaaagacaa	tctttgatga	aagcagcagc	agcggagcct	2820
ccgccgaaga	agatcagcat	gcctccccta	acatcagcct	gacactgagc	tattttccta	2880
aggtgaacga	aaacacagcc	cattccggag	aaaacgaaaa	cgattgtgat	gccgaactga	2940
gaatctggag	cgtgcaggaa	gatgatctgg	ccgccggact	gagctggatc	cctttttttg	3000
ggcccgggaat	tgaaggactg	tacaccgctg	gcctgatcaa	gaaccagaac	aacctggtgt	3060
gcaggctgag	gaggctggcc	aaccagaccg	ccaagagcct	ggagctgctg	ctgaggggtga	3120
ccaccgagga	gaggaccttc	agcctgatca	acaggcacgc	catcgacttc	ctgctgggcta	3180
ggtggggcg	cacctgcaag	gtgctgggct	ccgactgctg	catcggcatc	gaggacctga	3240
gcaggaacat	cagcgagcag	atcgaccaga	tcaagaagga	cgagcagaag	gagggcaccg	3300
gctggggcct	ggcgggcaag	tgggtggacca	gcgactgggg	agtgtgaca	aacctgggaa	3360
tcctgctgct	gctgagcatt	gccgtgctca	ttgtctgtgc	ctgtatctgt	agaatcttta	3420
ccaagtacat	cggatgatag	atctgtctgt	ccttctagtt	gccagccatc	tgttgttttg	3480
ccctcccctg	tgcttccctt	gacctggaa	agtgccactc	ccactgtctc	ttcctaataa	3540
aatgagga	ttgcatcgca	ttgtctgagt	gggtgtcatt	ctattctggg	gggtggggtg	3600
gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	ggcatgctgg	ggatgcggtg	3660
ggctctatgg	gtaccagagt	gctgaagaat	tgaccgggtt	cctcctgggc	cagaaagaag	3720
caggcacatc	cccttctctg	tgacacaccc	gtgccacgcc	cctggttcct	agttccagcc	3780
ccactcatag	gacactcata	gctcaggagg	gtcccgctt	caatcccacc	cgctaaagta	3840
cttggagcgg	tctctccctc	cctcatcagc	ccaccaaac	aaacctagcc	tccaagagt	3900
ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	agagaaaatg	cctccaacat	3960
gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	tgatttaagg	ccatcatggc	4020
cttaattcttc	cgcttccctg	ctcactgact	cgctgcgtc	ggcgttcgg	ctgcggcgag	4080
cggtatcagc	tactcaaaag	cggttaatac	ggttatccac	agaatcagg	gataacgcag	4140
gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	4200
tggtggtttt	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	4260
agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	4320
tcgtgctgct	tctgttccg	accctgccgc	ttaccggata	cctgtccgct	tttctccctt	4380
cggtgaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtagggtcg	4440
ttcgtctcaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccagaccg	tgcgcccttat	4500
ccgttaacta	tcgtcttgag	tccaaccgg	taagacacga	cttatcgcca	ctggcagcag	4560
ccactggtta	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	4620
ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	4680
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaacc	accgctggta	4740
gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	4800
atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgtaagggga	4860
ttttggtcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	4920
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	4980
tcagttaggc	acatatctca	gcgatctgtc	tatttctgtc	atccatagtt	gcctgactcg	5040
gggggggggg	gcgctgaggt	ctgcctcgtg	aagaagggtgt	tgctgactca	taccaggcct	5100
gaatcgcccc	atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	gctttgttgt	5160
aggtggacca	gttgggtgatt	ttgaactttt	gctttgccac	ggaacgggtc	gcgttgtcgg	5220
gaagatgcgt	gatctgatcc	ttcaactcag	caaaagtctg	atthattcaa	caaagccgcc	5280
gtcccgtcaa	gtcagcgtaa	tgctctgcca	gtgttacaac	caattaacca	attctgatta	5340
gaaaaactca	tcgagcatca	aatgaaactg	caattttattc	atatcaggat	tatcaatacc	5400
atatthtttg	aaaagccgtt	tctgtaatga	aggagaaaaac	tcaccgaggc	agttccatag	5460
tagggcaaga	tctgggtatc	ggtctgcgat	tcggactcgt	ccaacatcaa	tacaacctat	5520
taattttccc	tcgtcaaaaa	taaggttatc	aagtggagaa	tcaccatgag	tgacgactga	5580
atccggtgag	aatggcaaaa	gcttatgcat	ttctttccag	acttggtcaa	caggccagcc	5640
attacgctcg	tcattcaaaa	cactcgcac	aaacaaaccg	ttattcattc	gtgattgcgc	5700
ctgagcgaga	cgaataacgc	gatcgctgtt	aaaaggacaa	ttacaacacg	gaatcgaatg	5760
caaccgctgc	aggaacactg	ccagcgcac	aaaataattt	tcacctgaat	caggatattc	5820
ttctaatacc	tggaaatgctg	ttttcccggg	gatcgagtg	gtgagtaacc	atgcatcatc	5880
aggagtacgg	ataaaatgct	tgatggtcgg	aagaggcata	aattccgtca	gccagtttag	5940
tctgaccatc	tcattctgtaa	catcattggc	aacgctacct	ttgccatgtt	tcagaaacaa	6000
ctctggcgca	tcgggcttcc	catacaatcg	atagattgtc	gcacctgatt	gcccgcacatt	6060

NIH316.001VPC_SEQLIST.TXT

atcgcgagcc	catttataacc	catataaaatc	agcatccatg	ttggaattta	atcgcgggcct	6120
cgagcaagac	gtttcccgtt	gaatatggct	cataacaccc	cttgtattac	tgtttatgta	6180
agcagacagt	tttattgttc	atgatgatat	atttttatct	tgtgcaatgt	aacatcagag	6240
attttgagac	acaacgtggc	tttccccccc	ccccattat	tgaagcattt	atcaggggta	6300
ttgtctcatg	agcggataca	tatttgaatg	tatttagaaa	aataaacaaa	taggggttcc	6360
gcgcacattt	ccccgaaaag	tgccacctga	cgtctaagaa	accattatta	tcattgacatt	6420
aacctataaa	aataggcgta	tcacgaggcc	ctttcgtc			6458

<210> 6

<211> 6458

<212> DNA

<213> Marburg virus (Angola)

<400> 6

tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacgggtca	60
cagcttgtct	gtaagcggat	gccggggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cggcacacaga	gcagattgta	ctgagagtgc	180
accatatgcy	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagattgg	240
ctattggcca	ttgcatacgt	tgtatccata	tcataaatag	tacatttata	ttggctcatg	300
tccaacatta	ccgccatggt	gacattgatt	attgactagt	tattaatagt	aatcaattac	360
ggggctcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	cggtaaatgg	420
cccgcctggc	tgaccgcccc	acgacccccg	cccatggacg	tcaataatga	cgtatgttcc	480
catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	540
tgcccacttg	gcagtagatc	aagtgtatca	tatgccaaag	acgcccccta	ttgacgtcaa	600
tgacggtaaa	ttggcccgct	ggcattatgc	ccagtagatg	accttatggg	actttcctac	660
ttggcagtag	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggg	tttggcagta	720
catcaatggg	cgtggatagc	ggtttgactc	acgggggattt	ccaagtctcc	accccattga	780
cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaaaca	840
agctcgttta	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggagggtct	atataagcag	900
tagaagacac	gtgaaccgtc	agatcgctcg	gagacgccat	ccacgctgtt	ttgacctcca	960
cgccctacct	cgggaccgat	ccagcctcca	tcggctcgca	tctctccttc	acgcgcccgc	1020
ggtgcctcct	gaggcccgcca	tccacgcccg	ttgagtcgcy	ttctgccgcy	ttccgcctgt	1080
ctttgtccgg	gaactgcgtc	cgccgtctag	gtaagttaa	agctcaggtc	gagaccgggc	1140
accctgcttg	cgctcccttg	gagcctacct	agactcagcc	ggctctccac	gctttgcctg	1200
gctgcccgcg	ctcaactcta	gttaacggtg	gagggcagtg	tagtctgagc	agtactcgtt	1260
ggtcttttct	gcccaccag	acataatagc	tgacagacta	acagactgtt	cctttccatg	1320
atatacgccg	gcagtcaccg	tcgtcgacac	gtgtgatcag	atatcgcggc	cgctctagag	1380
agaccctgcc	catgaagacc	acctgcctgc	tgatcagcct	gattcctgat	cagggcgtga	1440
gcagcggcac	catcctggag	atcgccagca	acatccagcc	ccagaacgtg	gacagcgtgt	1500
agaaggtggc	cctgcagaag	accgaggacg	tgacactgat	gggcttcacc	ctgagcggcc	1560
cccccaagaa	cgacagccct	ctggaggcca	gcaagagggtg	ggccttcagg	gccggcggtg	1620
tgaccgaccc	cgtggagtac	accgaggccg	aggaggccaa	gacctgctac	aacatcagcg	1680
ctaagtgcaa	cagcggcaag	agcctgctgc	tggaacctcc	caccaacatc	agggactacc	1740
tgacactgtg	gaccatccac	cacatccagg	gccagaaccc	tcacgcccag	ggcatcgccc	1800
gaaaagtgtt	gggcgccttc	ttcctgtacg	acaggatcgc	cagcaccacc	atgtacagag	1860
tgatcttcag	cacagaggga	aacatcgctg	ctatgatcgt	gaacaagacc	gtgcataaga	1920
actggacaag	cagacaggga	caggatata	gacatatgaa	cctgacatcc	acaaacaagt	1980
aatacaactc	cagcaacgga	acaagacaaa	acgatacagg	atgttttggg	acactgcagg	2040
cagctcatcc	caaccaagaac	cagacatgtg	cccctagcaa	gaagcctctg	cctctgccta	2100
catatcctaa	tgaagtgaag	ctgacatcca	caagcacaga	tgccacaaag	ctgaacacaa	2160
aaccttatat	tagcgacgac	gaggatctga	caacaagcgg	atccggatcc	ggagaacagg	2220
cacctagccc	aacaagcgac	gctgctacaa	aacagggact	gtcctccaca	atgcctccta	2280
gagtgggtgac	tcagcctagc	acacctcagc	agggaggcaa	caacacaaac	cattcccagg	2340
acacaacaac	agaacctgga	aagacaaaca	caacagccca	gcctagcatg	cctcctcata	2400
ctattcagaa	aatcagcaca	aacaacacct	ccaagcaca	tctgagcaca	cctagcgtgc	2460
ccgccccttc	tgccaccaac	tacaacacac	agtcacacag	ccctgaaaa	gaacagacct	2520
acagcacaac	gagccctaca	acaacagtgc	cagaaaaacc	tacaacagcc	aagagacaaa	2580
ctagccctac	acctaattcc	acagctcagc	ctaacacaa	aaacaagtat	agcacaagcc	2640
tcctgtggag	agaaggagat	atgttccctt	ttctggatgg	ttttagaaga	aagagaaaca	2700
attttgatcc	tgtgcctaac	acaaagacaa	ttcttggatg	actgatcaac	gtccttatcg	2760
ccgccaagaa	agatcagcat	gcctccccta	acatcagcct	aagcagcagc	agcggagcct	2820
aggtgaacga	aaacacagcc	cattccggag	aaaacgaaaa	gacactgagc	tattttccta	2880
gaatctggag	cgtgcaggaa	gatgatctgg	ccgccggact	cgattgtgat	gccgaactga	2940
ggcccggaa	tgaaggactg	tacaaccgcc	gcctgatcaa	gagctggatc	cctttttttg	3000
gcaggctgag	gaggctggcc	aaccagaccg	ccaagagcct	gaaccagaac	aacctggtgt	3060
				ggagctgctg	ctgagggtga	3120

NIH316.001VPC_SEQLIST.TXT

ccaccgagga	gaggaccttc	agcctgatca	acaggcacgc	catcgacttc	ctgctggcta	3180
ggtggggcgg	cacctgcaag	gtgctgggcc	ccgactgctg	catcggcatc	gaggacctga	3240
gcaggaacat	cagcgagcag	atcgaccaga	tcaagaagga	cgagcagaag	gagggcaccg	3300
gctggggcct	ggcgggcaag	tgggtggacca	gcgactgggg	agtgctgaca	aacctgggaa	3360
tcctgctgct	gctgagcatt	gccgtgctca	ttgctctgtc	ctgtatctgt	agaatcttta	3420
ccaagtacat	cggatgatag	atctgctgtg	ccttctagtt	gccagccatc	tggtgtttgc	3480
ccctccccc	tgcttccctt	gacctggaa	ggtggcactc	ccactgtcct	ticctaataa	3540
aatgagga	ttgcatcgca	ttgtctgagt	aggtgtcatt	ctattctggg	gggtgggggtg	3600
gggcaggaca	gcaaggggga	ggattgggaa	gacaatagca	ggcatgctgg	ggatgcgggtg	3660
ggctctatgg	gtacccagggt	gctgaagaat	tgacccgggt	cctcctgggc	cagaaagaag	3720
caggcacatc	cccttctctg	tgacacacc	tgtccacgcc	cctggttcct	agttccagcc	3780
ccactcatag	gcactcata	gctcaggagg	gctccgcctt	caatcccacc	cgctaaagta	3840
cttgagcg	tctctccctc	cctcatcagc	ccaccaaacc	aaacctagcc	tccaagagtg	3900
ggaagaaatt	aaagcaagat	aggctattaa	gtgcagaggg	agagaaaatg	cctccaacat	3960
gtgaggaagt	aatgagagaa	atcatagaat	tttaaggcca	tgatttaaag	ccatcatggc	4020
cttaatcttc	cgcttccctg	ctcactgact	cgctgcgtc	ggcgttcgg	ctgcggcgag	4080
cggtatcagc	tcactcaaa	gcggtaatac	ggttatccac	agaatcagg	gataacgcag	4140
gaaagaacat	gtgagcaaaa	ggccagcaaa	agccaggaa	ccgtaaaaag	gccgcgttgc	4200
tggcggtttt	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	4260
agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	4320
tcgtgcgctc	tctgtttccg	accctgccgc	ttaccggata	cctgtccgcg	tttctccctt	4380
cggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggctg	4440
ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccgaccgc	tgcgctttat	4500
ccggtacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	4560
ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	4620
ggtggcctaa	ctacggctac	actagaagaa	cagtatttgg	tatctgcgt	ctgctgaagc	4680
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaacc	accgctggta	4740
gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	4800
atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgtaaaggga	4860
ttttggtcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	4920
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	4980
tcagtggagc	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcg	5040
gggggggggg	gcgctgaggt	ctgcctcgtg	aagaaggtgt	tgctgactca	taccaggcct	5100
gaatcgcccc	atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	gctttgttgt	5160
agggtggacca	gttgggtgatt	ttgaactttt	gctttgccac	ggaacggtct	gcgttgctcg	5220
gaagatgcgt	gatctgatcc	ttcaactcag	caaaagtctg	atatttcaa	caaagccgcc	5280
gtcccgtcaa	gtcagcgtaa	tgctctgcca	gtgttacaac	caattaacca	attctgatta	5340
gaaaaactca	tcgagcatca	aatgaaactg	caattttattc	atatcaggat	tatcaatacc	5400
atatttttga	aaaagccggt	tctgtaatga	aggagaaaaac	tcaccgaggc	agttccatag	5460
gatggcaaga	tcctgggtatc	ggtctgcgat	tccgactcgt	ccaacatcaa	tacaacctat	5520
taatttcccc	tcgtcaaaaa	taaggttatc	aagtggagaaa	tcaccatgag	tgacgactga	5580
atccggtgag	aatggcaaaa	gcttatgcat	ttctttccag	acttggtcaa	caggccagcc	5640
attacgctcg	tcatcaaaat	cactcgcac	aaacaaaccg	ttattcattc	gtgattgcgc	5700
ctgagcgaga	cgaaatacgc	gatcgctgtt	aaaaggacaa	ttacaaacag	gaatcgaatg	5760
caaccggcgc	aggaacactg	ccagcgcac	aacaatattt	tcacctgaat	caggatattc	5820
ttctaatacc	tggaaatgctg	ttttcccggg	gatcgcagtg	gtgagtaacc	atgcatcatc	5880
aggagtacgg	ataaaatgct	tgatggtcgg	aagaggcata	aattccgtca	gccagtttag	5940
tctgaccatc	tcactctgtaa	catcattggc	aacgctacct	ttgccatgtt	tcagaaacaa	6000
ctctggcgca	tcgggctttcc	catacaatcg	atagattgtc	gcacctgatt	gcccgcatt	6060
atcgcgagcc	catttatacc	catataaatc	agcatccatg	ttggaattta	atcgcggcct	6120
cgagcaagac	gtttcccgtt	gaatatggct	cataacaccc	cttgatttac	tgtttatgta	6180
agcagacagt	tttattgttc	atgatgatat	atttttatct	tgtgcaatgt	aacatcagag	6240
attttgagac	acaacgtggc	tttccccccc	ccccattat	tgaagcattt	atcaggggta	6300
ttgtctcatg	agcggataca	tatttgaatg	tatttagaaa	aataaacaaa	taggggttcc	6360
gcgcacattt	ccccgaaaag	tgccacctga	cgtctaagaa	accattatta	tcatgacatt	6420
aacctataaa	aataggcgta	tcacgaggcc	ctttcgtc			6458

<210> 7

<211> 8128

<212> DNA

<213> Ebola virus (Zaire)

<400> 7

ttaattaacc	gcaatttctca	tgtttgacag	cttatcatca	tcaataatat	accttatttt	60
ggattgaagc	caatatgata	atgagggggt	ggagtttgtg	acgtggcgcg	ggcggtggga	120
acggggcggg	tgacgtagta	gtgtggcgga	agtgtgatgt	tgcaagtgtg	gcggaacaca	180

NIH316.001VPC_SEQLIST.TXT

tgtaagcgac	ggatgtggca	aaagtgcagt	ttttggtgtg	cgccggtgta	cacaggaagt	240
gacaattttc	gcgcggtttt	aggcggtatg	tgtagtaaat	ttgggcgtaa	ccgagtaaga	300
tttggccatt	ttcgcgggaa	aactgaataa	gaggaagtga	aatctgaata	attttgtgtt	360
actcatagcg	cgtaatat	gtctagggcc	gcggggactt	tgaccgttta	cgtaggagact	420
cgcccaggtg	tttttctcag	gtgttttccg	cgttccgggt	caaagtggc	gttttattat	480
tatagttagt	acgtaccagt	gactggccct	agggtggtcaa	tattggccat	tagccatatt	540
attcattggt	tatatagcat	aaatcaatat	tggctattgg	ccattgcata	cgttgatcc	600
atatcataat	atgtacattt	atattggctc	atgtccaaca	ttaccgccat	gttgacattg	660
attattgact	agttattaat	agtaatcaat	tacgggggtca	ttagttcata	gcccataat	720
ggagttccgc	gttacataac	ttacggtaaa	tggcccgcct	ggctgaccgc	ccaacgaccc	780
ccgcccattg	acgtcaataa	tgacgtatgt	tcccatagta	acgccaatag	ggactttcca	840
ttgacgtcaa	tggttgaggt	atttacggta	aactgccac	ttggcagta	aaatgggagg	900
tcatacgccc	agtagcggc	ctattgacgt	caatgacggg	aaatggcccg	cctggcatta	960
tgcccagtag	atgaccttat	gggactttcc	tacttggcag	tacatctacg	tatttagtcat	1020
cgctattacc	atggtgatgc	ggttttggca	gtacatcaat	gggctgggat	agcggtttga	1080
ctcacgggga	tttccaagtc	tccaccccat	tgacgtcaat	gggagtttgt	tttggcacca	1140
aaatcaacgg	gactttccaa	aatgtcgtaa	caactccgcc	ccattgacgc	aaatgggagg	1200
taggcgtgta	cggtgggagg	tctatataag	cagagctcgt	ttagtgaacc	gtcagatcgc	1260
ctggagacgc	catccacgct	gttttgacct	ccatagaaga	caccgggacc	gatccagcct	1320
ccgcgccggg	gaacgggtgca	ttggaagctt	gccgccatgg	gcgttacagg	aatattgcag	1380
ttacctcgtg	atcgattcaa	gaggacatca	tcttttcttt	gggtaattat	ctttttccaa	1440
agaacatttt	ccatcccact	tggagtcatc	cacaatagca	cattacagggt	tagtgatgtc	1500
gacaaactag	tttgtcgtga	caaactgtca	tccacaaatc	aattgagatc	agttggactg	1560
aatctcgtatg	gcaatggagt	ggcaactgac	gtgccatctg	caactaaaag	atggggcttc	1620
aggtccggtg	ttcccacaaa	ggtggtcaat	tatgaagctg	gtgaatgggc	tgaaaactgc	1680
tacaatcttg	aaatcaaaaa	acctgacggg	agtgagtgtc	taccagcagc	gccagacggg	1740
attcggggct	ttccccgggtg	ccggtatgtg	cacaaagtat	caggaacggg	accgtgtgcc	1800
ggagactttg	ccttcataaa	agaggggtgt	ttcttcctgt	atgatcgact	tgcttccaca	1860
gttatctacc	gaggaacgac	tttcgctgaa	ggtgtcgttg	catttctgat	actgccccaa	1920
gctaagaagg	acttcttcag	ctcacacccc	ttgagagagc	cggtaaatgc	aacggaggac	1980
ccgtctagtg	gctactattc	taccacaatt	agatatcagg	ctaccgggtt	tggaaccaat	2040
gagacagagt	acttgttcga	ggttgacaat	ttgacctacg	tccaacttga	atcaagattc	2100
acaccacagt	ttctgctcca	gctgaatgag	acaatatata	caagtgggaa	aaggagcaat	2160
accacgggaa	aaactaattt	gaaggtcaac	cccgaattg	atacaacaat	cggggagttg	2220
gccttctggg	aaactaaaaa	aaacctcact	agaaaaattc	gcagtgaaga	gttgtctttc	2280
acagttgtat	caaacggagc	caaaaacatc	agtggtcaga	gtccggcgcg	aacttcttcc	2340
gaccacggga	ccaacacac	aactgaagac	cacaaaatca	tggttctaga	aaattcctct	2400
gcaatggttc	aagtgcacag	tcaaggaaag	gaagctgcag	tgctcgatct	aacaaccctt	2460
gccacaatct	ccacgagtcc	ccaatccctc	acaaccaaac	cagggtccga	caacagcacc	2520
cataatcac	ccgtgtataa	acttgacatc	tctgaggcaa	ctcaagttga	acaacatcac	2580
cgcagaacag	acaacgacag	cacagcctcc	gacactccct	ctgccacgac	cgcagccgga	2640
cccccaaaag	cagagaacac	caacacgagc	aagagcactg	acttcttgga	ccccgccacc	2700
acaacaagtc	cccaaaacca	cagcgagacc	gctggcaaca	acaacactca	tcaccaagat	2760
accggagaag	agagtgccag	cagcgggaag	ctaggcttaa	ttaccaatac	tactcttgga	2820
gtcgcaggac	tgatcacagg	cgggagaaga	actcgaagag	aagcaattgt	caatgctcaa	2880
cccaaatgca	accctaattt	acattactgg	actactcagg	atgaagggtg	tgcaatcgga	2940
ctggcctgga	taccatattt	cgggccaagca	gccgagggaa	ttacataga	ggggctaatt	3000
cacaatcaag	atggtttaat	ctgtgggttg	agacagctgg	ccaacgagac	gactcaagct	3060
cttcaactgt	tcctgagagc	cacaactgag	ctacgcacct	tttcaatcct	caaccgtaag	3120
gcaattgatt	tcttgctgca	gcatgggggc	ggcacatgcc	acattctggg	accggactgc	3180
tgtatcgaa	cacatgattg	gaccaagaac	ataacagaca	aaattgatca	gattattcat	3240
gattttgttg	ataaaaacct	tccggaccag	ggggacaatg	acaattggtg	gacaggatgg	3300
agacaatgga	taccggcagg	tattggagtt	acaggcggtg	taattgcagt	tatcgcttta	3360
ttctgtatat	gcaaattttg	cttttagtaa	tctagacgag	atccgaactt	gtttattgca	3420
gcttataatg	gttacaaata	aagcaatagc	atcacaattt	tcacaaataa	agcatttttt	3480
tcactgcatt	ctagtgtgtg	tttgtccaaa	ctcatcaatg	tatcttatca	tgtctagatc	3540
tgtactgaaa	tgtgtggggc	tggcttaagg	gtgggaaga	atatataagg	tgggggcttt	3600
atgtagtgtt	gtatctgttt	tgacgacggc	gccgcccga	tgagcaccac	ctcggttgat	3660
ggaagcattg	tgagctcata	tttgacaacg	cgcatgcccc	catgggcccg	ggtgctgcag	3720
aatgtgatgg	gctccagcat	tgatggctgc	cccgtcctgc	ccgcaaactc	tactaccttg	3780
acctacgaga	ccgtgtcttg	aacgcggtg	gagactgcag	cctccgccgc	cgcttcagcc	3840
gctgcagcca	ccgcccgcgg	gattgtgact	gactttgctt	tcctgagccc	gcttcaagc	3900
agtgacagct	ccggttcac	cgcccgcgat	gacaagttga	cggtctttt	ggcacaattg	3960
gattctttga	cccgggaact	taatgtcgtt	tctcagcagc	tggtggatct	gcgccagcag	4020
gtttctgccc	tgaaggcttc	ctcccctccc	aatgcggttt	aaaacataaa	taaaaaacca	4080
gactctgttt	ggatttggtg	caagcaagtg	tcttgctgtc	tttatttagg	ggttttgcgc	4140
gcgcggtagg	cccgggacca	gcggtctcgg	tcgttgaggg	tcctgtgtat	ttttccagg	4200
acgtggtaaa	ggtgactctg	gatgttcaga	tacatgggca	taagcccgtc	tctggggtgg	4260

NIH316.001VPC_SEQLIST.TXT

aggtagcacc	actgcagagc	ttcatgctgc	gggggtgggtg	tgtagatgat	ccagtcgtag	4320
caggagcgct	gggctgggtg	cctaaaaatg	tctttcagta	gcaagctgat	tgccaggggc	4380
aggcccttgg	tgtaagtgtt	tacaaaagcg	ttaagctggg	atgggtgcat	acgtggggat	4440
atgagatgca	tcttgactg	tatttttagg	ttggctatgt	tcccagccat	atccctccgg	4500
ggattcatgt	tgtgcagaac	caccagcaca	gtgtatccgg	tgcacttggg	aaatttgtca	4560
tgtagcttag	aaggaaatgc	gtggaagaac	ttggagacgc	ccttgtgacc	tccaagattt	4620
tccatgcatt	cgtccataat	gatggcaatg	ggcccacggg	cggcggcctg	ggcgaagata	4680
tttctgggat	cactaacgtc	atagttgtgt	tccaggatga	gatcgtcata	ggccattttt	4740
acaaagcgcg	ggcggagggg	gccagactgc	ggtataatgg	ttccatccgg	cccaggggcg	4800
tagttaccct	cacagatttg	catttcccac	gctttgagtt	cagatggggg	gatcatgtct	4860
acctgcgggg	cgatgaagaa	aacggtttcc	ggggtagggg	agatcagctg	ggaagaaagc	4920
aggttcccta	gcagctgcga	cttaccgcag	ccggtgggccc	cgtaaatcac	acctattacc	4980
ggctgcaact	ggtagttaag	agagctgcag	ctgcgctcat	ccctgagcag	ggggggccact	5040
tcgttaagca	tgtccctgac	tcgcatgttt	tccctgacca	aatccgccag	aaggcgctcg	5100
ccgcccagcg	atagcagttc	ttgcaaggaa	gcaaagtttt	tcaacggttt	gagaccgtcc	5160
gccgtaggca	tgcttttgag	cgtttgacca	agcagttcca	ggcgtccca	cagctcggtc	5220
acctgctcta	cggcatctcg	atccagcata	tctcctcggt	tcgagggttg	ggcgcgcttt	5280
cgctgtacgg	cagtagtcgg	tgctcgtcca	gacggggccag	ggtcatgtct	ttccacgggc	5340
gcagggctct	cgtagcgtga	gtctgggtca	cgggtgaagg	gtgcgtcccg	ggctgcgcgc	5400
tggccagggt	gcgcttgagg	ctggtcctgc	tgggtgtgaa	gcgctgccgg	tcttcgcctt	5460
gcgctcggc	caggtagcat	ttgaccatgg	tgctcatagt	cagccctcc	cgggcggtggc	5520
ccttggcgcg	cagcttgccc	ttggaggagg	cgccgcacga	ggggcagtg	agacttttga	5580
ggcgtagag	cttgggcgcg	agaaataccg	attccgggga	gtaggcatcc	gcgccgcagg	5640
ccccgcagac	ggtctcgcac	tccacgagcc	aggtgagctc	tggccgttcg	gggtcaaaaa	5700
ccaggtttcc	cccatgcttt	ttgatgcgtt	tcttacctct	ggtttccatg	agccggtgtc	5760
cacgctcggt	gacgaaaagg	ctgtccgtgt	ccccgtatac	agacttgaga	ggcctgtcct	5820
cgagcgggtg	tccgcggtcc	tcctcgtata	gaaactcgga	ccactctgag	acaaaggctc	5880
gcgtccaggc	cagcacgaag	gaggctaagt	gggaggggta	gcggtcgttg	tccactaggg	5940
ggtccactcg	ctccagggtg	tgaagacaca	tgtcgccttc	ttcggcatca	aggaaggtga	6000
ttggtttgta	ggtgtaggcc	acgtgaccgg	gtgttctga	aggggggcta	taaaaggggg	6060
tggggcgcg	ttcgtcctca	ctctcttcgg	catcgctgtc	tgcgaggggc	agctgttggg	6120
gtgagtcgac	gcgaggctgg	atggccttcc	ccattatgat	tcttctcgct	tccggcgcca	6180
tcgggatgcc	cgcggttgca	gccatgctgt	ccaggcaggt	agatgacgac	catcaggggac	6240
agcttcaagg	ccagcaaaa	gccaggaacc	gtaaaaaggc	cggttgctg	gcgtttttcc	6300
ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	6360
acccgacagg	actataaaga	taccaggcgt	ttccccctgg	aagctccctc	gtgcgtctct	6420
ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtgg	6480
cgctttctca	tagctcacgc	tgtaggatat	tcagttcggg	gtaggctcgt	cgctccaagc	6540
tgggtgtgtg	gcacgaacc	cccgttcagc	ccgacccgtg	cgcttatcc	ggtaactatc	6600
gtctttagtc	caaccggta	agacacgact	tatcgccact	ggcagcagc	actggttaaca	6660
ggattagcag	agcgagggtat	gtaggcggtg	ctacagagtt	cttgaagtgg	tggcctaact	6720
acggctacac	tagaaggaca	gtatttggta	tctgcgtctc	gctgaagcca	gttaccttcg	6780
gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	cgctggtagc	ggtggttttt	6840
ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaggatc	tcaagaagat	cctttgatct	6900
tttctacggg	gtctgacgct	cagtggaaag	aaaactcacg	ttaagggtat	ttggtcatga	6960
gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	aaaatgaagt	tttaaataca	7020
tctaaagtat	atatgagtaa	acttggctct	acagttacca	atgcttaatc	agtgaggcac	7080
ctatctcagc	gatctgtcta	tttcgttcac	ccatagttgc	ctgactcccc	gtcgtgtaga	7140
taactacgat	acgggagggc	ttaccatctg	gccccagtg	tgcaatgata	ccgcgagacc	7200
cacgctcacc	ggctccagat	ttatcagcaa	taaaccagcc	agccggaagg	gccgagcgca	7260
gaagtggctc	tgcaacttta	tccgcctcca	tccagttctat	taattgttgc	cggaagcta	7320
gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	tgccatttgt	gcaggcatcg	7380
tgggtgcacg	ctcgtcgttt	ggtagggctt	cattcagctc	cggttcccaa	gcatcaaggc	7440
gagttacatg	atcccccatg	ttgtgcaaaa	aagcggttag	ctccttcggg	cctccgatcg	7500
ttgtcagaag	taagttggcc	gcagtggtat	cactcatggg	tatggcagca	ctgcataatt	7560
ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	tggtagtagc	tcaaccaagt	7620
cattctgaga	atagtgtatg	cggcgaccga	tgtgtctttg	cccggcgctg	acacgggata	7680
ataccgcgcc	acatagcaga	actttaaaag	gttctcatcat	tggaaaacgt	tcttcggggc	7740
gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	actcgtgcac	7800
ccaactgatc	ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	aaaacaggaa	7860
ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	ctcatactct	7920
tcctttttca	atattattga	agcatttatc	ggggttattg	tctcatgagc	ggatacatat	7980
ttgaattgtat	ttagaataat	aggaataatg	aggttccgcg	cacattttcc	cgaaaagatg	8040
cacctgacgt	ctaagaaacc	attattatca	tgacattaac	ctataaaaa	aggcgtatca	8100
cgaggccctt	tcgtcttcaa	gaattgtt				8128

NIH316.001VPC_SEQLIST.TXT

<210> 8
 <211> 8128
 <212> DNA
 <213> Ebola (Sudan, Gulu)

<400> 8
 ttaattaacc gcaatttctca tgtttgacag cttatcatca tcaataatat accttatttt 60
 ggattgaagc caatatgata atgagggggg ggagtttgtg acgtggcgcg gggcggtggga 120
 acggggcggg tgacgtagta gtgtggcgga agtgtgatgt tgcaagtgtg gcggaacaca 180
 tgtaagcgac ggatgtggca aaagtgacgt ttttgggtgt cgccggtgta cacaggaagt 240
 gacaattttc gcgcgggttt aggcggatgt tgtagtaaat ttgggcgtaa ccgagtaaga 300
 tttggccatt ttcgcgggaa aactgaataa gaggaagtga aatctgaata attttgtgtt 360
 actcatagcg cgtaatatat gtctagggcc gcggggactt tgaccgttta cgtggagact 420
 cgcccaggtg tttttctcag gtgttttccg cgttccgggt caaagtgtgg gttttattat 480
 tatagttagt acgtaccagt gcaactggcct aggtgtgtcaa tattggccat tagccatatt 540
 attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata cgttgtatcc 600
 atatcataat atgtacattt atattggctc atgtccaaca ttaccgcat gttgacattg 660
 attattgact agttattaat agtaatcaat tacgggggtca ttagttcata gcccatatat 720
 ggagttccgc gttacataac ttacggtaaa tggcccgccg ggtgaccgc ccaacgacct 780
 ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca 840
 ttgacgtcaa ttgggtggag atttacggta aactgccac ttggcagta atcaagtgtg 900
 tcatatgcca agtacgccc ctattgacgt caatgacggt aaatggcccg cctggcatta 960
 tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac 1020
 cgctattacc atggtgatgc ggttttggca gtacatcaat ggcggtggat agcggtttga 1080
 ctacggggg tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcaca 1140
 aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatggcgcg 1200
 taggcgtgta cgggtggagg tctatataag cagagctcgt ttagtgaacc gtcagatcgc 1260
 ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct 1320
 ccgcgcccg gaacggtgca ttggaagctt gccgccatgg ggggtcttag cctactccaa 1380
 ttgcccaggg acaaatttcg gaaaagctct tcttttgttt ggtcatcat cttattccaa 1440
 aaggcctttt ccatgccttt ggggtgtgtg actaacagca ctttagaagt aacagagatt 1500
 gaccagctag tctgcaagga tcatcttgca tctactgacc agctgaaatc agttggctctc 1560
 aacctcgacg ggagcggagt atctactgat atcccatctg caacaaagcg ttggggcttc 1620
 agatctggtg ttcttcccaa ggtggtcagc tatgaagcgg gagaatggg tgaaaattgc 1680
 tacaatcttg aaataaagaa gccggacggg agcgaatgct taccgccacc gccagatggt 1740
 gtcagaggct ttccaagggt ccgctatgtt cacaagccc aaggaaccgg gccctgcca 1800
 ggtgactacg cttttcacia ggatggagct ttcttctct atgacaggct ggcttcaact 1860
 gtaatttaca gaggagtcaa ttttgctgag ggggtaattg cattcttgat attggctaaa 1920
 ccaaaagaaa cgttccctca gtcaccccc attcagagg cagttaaata cctgaaaat 1980
 acatcaagtt attatgccac atcctacttg agtagtgaat tcgaaaattt tgggtgctca 2040
 cactccacga cctttttcaa aattgacaat aatactttt ttctgtctga caggccccac 2100
 acgcctcagt tctttttcca gctgaatgat accattcacc ttaccaaca gttgagtaat 2160
 acaactggga gactaatttg gacactagat gctaataatc atgctgata tgggtgaatg 2220
 gctttttggg aaataaaaaa aatctctcc gacaactac gtggagaaga cgtgtctttc 2280
 gaagctttat cgctcaacga gacagaagac gatgatgcgg catcgctcag aattacaag 2340
 ggaagaatct ccgaccgggc caccaggaag tattcggaac tggttccaaa gaattcccct 2400
 gggatggttc cattgcacat accagaaggg gaaacaacat tgccgtctca gaattcgaca 2460
 gaaggtcgaa gagttagtgt gaacactcag gagaccatta cagagacagc tgcaacaatt 2520
 ataggcacta acggcaacca tatgcagatc tccaccatcg ggataagacc gagctccagc 2580
 caaatcccga gttcttcacc gaccacggca ccaagccctg aggtcagac cccacaacc 2640
 cacacatcag gtccatcagt gatggccacc gaggaaccaa caacaccacc gggagctcc 2700
 cccggcccaa caacagaagc acccactctc accaccccag aaaatataac aacagcggtt 2760
 aaaactgtcc tgccacagga gtccacaagc aacgggtctaa taacttcaac agtaacagg 2820
 attcttggga gtcttgggct tcgaaaacgc agcagaagac aaactaacac caaagccag 2880
 ggtaagtga atcccaactt acactactgg actgcacaag aacaacataa tgctgctggg 2940
 attgcttggg tcccgacttt tggaccgggt gcggaaggca tataactga aggcctgatg 3000
 cataaccaaa atgccttagt ctgtggactt aggcaacttg caaatgaaac aactcaagct 3060
 ctgcagcttt tcttaagagc cacaacggag ctgcggacat ataccatact caataggaa 3120
 gccatagatt tctttctgcg acgatggggc gggacatgca ggatcctgg accagattgt 3180
 tgcattgagc cacatgattg gacaaaaaac atcactgata aaatcaacca aatcatccat 3240
 gatttcatcg acaaccctt acctaactag gataatgatg ataattggtg gacgggctgg 3300
 agacagtggg tccctgcagg aataggcatt actggaatta ttattgcaat tattgtctct 3360
 ctttgccttt ttgctgataa tctgacgag atccgaactt gttttattga 3420
 gcttataatg gttacaaata aagcaatagc atcacaattt tcacaaataa agcatttttt 3480
 tcaactgatt ctagttgtgg tttgtccaaa ctcactaatg tatcttatca tgtctagatc 3540
 tgtactgaaa tgtgtgggcg tggcttaagg gtgggaaaga atatataagg tgggggtctt 3600
 atgtagtttt gtatctgttt tcgacgacc ccgcgccca tgagcaccia ctcgtttgat 3660
 ggaagcattg tgagctcata ttgacaacg ccatgcccc catgggccc ggtgcgtcag 3720

NIH316.001VPC_SEQLIST.TXT

aatgtgatgg	gctccagcat	tgatgggtcgc	cccgtcctgc	ccgcaaactc	tactaccttg	3780
acctacgaga	ccgtgtctgg	aacgccgttg	gagactgcag	cctccgccgc	cgcttcagcc	3840
gctgcagcca	ccgcccgcg	gattgtgact	gactttgctt	tcctgagccc	gcttgcaagc	3900
agtgacgctt	cccgttcac	cgcccgcat	gacaagttga	cggtctttt	ggcacaattg	3960
gattctttga	cccgggaact	taatgtcgtt	tctcagcagc	tggtggatct	gcgccagcag	4020
gtttctgccc	tgaaggcttc	ctccccctcc	aatgcggttt	aaaacataaa	taaaaaacca	4080
gactctgttt	ggatttggat	caagcaagtg	tcttgcctgc	tttatttagg	ggttttgcgc	4140
gcgcggtagg	cccgggacca	gcggtctcgg	tcgttgaggg	tcctgtgtat	tttttccagg	4200
acgtggtaaa	ggtgactctg	gatgttcaga	tacatgggca	taagcccgtc	tctggggtgg	4260
aggtagcacc	actgcagagc	ttcatgctgc	gggggtgggt	tgtagatgat	ccagtcgtag	4320
caggagcgct	gggcgtgggt	cctaaaaatg	tctttcagta	gcaagctgat	tgccaggggc	4380
aggcccttgg	tgtaatgtgt	tacaaagcgg	ttaaagctggg	atgggtgcat	acgtggggat	4440
atgagatgca	tcttggactg	tatttttagg	ttggctatgt	tcccagccat	atccccccgg	4500
ggattcatgt	tgtgcagaac	caccagcaca	gtgtatccgg	tgactttggg	aaattttgtca	4560
tgtagcttag	aaggaaatgc	gtggaagaac	ttggagacgc	ccttgtgacc	tccaagattt	4620
tcctatgcatt	cgtccataat	gatggcaatg	ggcccacggg	cggcggcctg	ggcgaagata	4680
tttctgggat	cactaacgtc	atagtttgtt	tccaggatga	gatcgtcata	ggccattttt	4740
acaaagcgcg	ggcggagggt	gccagactgc	ggtataatgg	ttccatccgg	cccaggggcg	4800
tagttaccct	cacagatttg	catttcccac	gctttgagtt	cagatggggg	gatcatgtct	4860
acctgcgggg	cgatgaagaa	aacggtttcc	ggggtagggg	agatcagctg	ggaagaaagc	4920
aggttcctga	gcagctgcga	cttaccgcag	ccggtggggc	cgtaaatcac	acctattacc	4980
ggctgcaact	ggtagttaag	agagctgcag	gtccgctcat	ccctgagcag	gggggacct	5040
tcgttaagca	tgtccctgac	tcgcatgttt	tccctgacca	aatccgccag	aaggcgctcg	5100
ccgccagcgc	atagcagttc	ttgcaaggaa	gcaaagtttt	tcaacggttt	gagaccgtcc	5160
gccgtaggca	tgctttttag	cgtttgacca	agcagttcca	ggcgggtcca	cagctcggtc	5220
acctgctcta	cggcatctcg	atccagcata	tctcctcggt	tcgcgggttg	ggcggtcttt	5280
cgctgtacgg	cagtagtcgg	tgctcgtcca	gacggggccag	ggtcatgtct	ttccacgggc	5340
gcagggtcct	cgtcagcgta	gtctgggtca	cggtgaaggg	gtgcgtcccg	ggctgcgcgc	5400
tgccagggt	gcgcttgagg	ctggctcctg	tggtgctgaa	gcgtgcggg	tcttcgcctt	5460
gcgcgtcggc	caggtagcat	ttgaccatgg	gtcctatagc	cagccccctc	gcggcggtgg	5520
ccttggcgcg	cagcttgccc	ttggaggagg	gcgcgcacga	ggggcagtg	agacttttga	5580
gggcgtagag	cttgggcgcg	agaaataccg	attccgggga	gtaggcatcc	gcgccgcagg	5640
ccccgcagac	ggtctcgcct	tccacgagcc	aggtgagctc	tgcccggttc	gggtcaaaaa	5700
ccaggtttcc	cccatgcttt	ttgatgcgtt	tcttacctct	ggtttccatg	agccggtgtc	5760
cacgctgggt	gacgaaaagg	ctgtccgtgt	ccccgtatac	agacttgaga	ggcctgtcct	5820
cgagcgggtg	tccgcggttc	tcctcgtata	gaaactcgga	ccactctgag	acaaaggctc	5880
gcgtccaggc	cagcacgaag	gaggctaagt	gggaggggta	gcggtcgttg	tccactaggg	5940
ggtccactcg	ctccagggtg	tgaagacaca	tgctgccttc	ttcggcatca	aggaagggtg	6000
ttggtttcta	ggtgtaggcc	acgtgaccgg	gtgttctcta	aggggggcta	taaaaggggg	6060
ttggggcgcg	ttcgtcctca	ctctcttccg	catcgctgtc	tgcgaggggc	agctgttggg	6120
gtgagtcgac	gcgaggctgg	atggccttcc	ccattatgat	tcttctcgct	tccggcgcca	6180
tcgggatgcc	cgctgttcag	gccatgctgt	ccaggcaggt	agatgacgac	catcaggggc	6240
agcttcaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	cgcttgcgtg	gcgtttttcc	6300
ataggctccg	ccccctgac	gagcatcaca	gaaactcgac	ctcaagtcag	aggtggcgaa	6360
acccgacagg	actataaaga	taccaggcgt	ttccccctgg	aagctccctc	gtgcgtcttc	6420
ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtag	6480
cgctttctca	tagctcacgc	tgtaggtatc	tcagttcggg	gtaggtcggt	cgctccaagc	6540
tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	cgcttatctc	ggttaactatc	6600
gtcttgagtc	caaccggta	agacacgact	tatcgccact	ggcagcagcc	actggttaaca	6660
ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	cttgaagtgg	tgccctaact	6720
acggctacac	tagaaggaca	gtatttggtg	tctgcgtctt	gctgaagcca	gttaccttcg	6780
gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	cgctggtagc	ggtggttttt	6840
ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaagggatc	tcaagaagat	cctttgatct	6900
tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	ttaagggtatt	ttggtcatga	6960
gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	aaaaatgaagt	tttaaatcaa	7020
tctaaagtat	atatgagtaa	acttgggtctg	acagttacca	atgcttaatc	agtgaggcac	7080
ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	ctgactcccc	gtcgtgtaga	7140
taactacgat	acgggagggc	ttaccatctg	gccccagtgc	tgcaatgata	cccgagacc	7200
cacgctcacc	ggctccagat	ttatcagcaa	ttaaccagcc	agccgggaagg	ggcgagcgca	7260
gaagtgggtc	tgcaacttta	tccgcctcca	tccagttctat	taattgttgc	cggaagcta	7320
gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	tgccattgct	gcaggcatcg	7380
tggtgtcacg	ctcgtcgttt	ggtatggctt	cattcagctc	cggttcccaa	cgatcaaggc	7440
gagttacatg	atcccccatg	ttgtgcaaaa	aagcgggttag	ctccttcggt	cctccgatcg	7500
ttgtcagaag	taagttggcc	gcagtgttat	cactcatggg	tatggcagca	ctgcataatt	7560
ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	tggtgagtac	tcaaccaagt	7620
cattctgaga	atagtgtatg	cggcgaccga	ggtgtctctg	ccggcgctca	acacgggata	7680
ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	tggaaaacgt	tcttcggggc	7740
gaaaactctc	aaggatctta	ccgctgttga	gatgtaaccc	actcgtgcac		7800

NIH316.001VPC_SEQLIST.TXT

```

ccaactgattc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 7860
ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 7920
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7980
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 8040
cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 8100
cgaggccctt tcgtcttcaa gaattgtt 8128

```

<210> 9
 <211> 676
 <212> PRT
 <213> Ebola virus (Zaire)

<400> 9
 Met Gly Val Thr Gly Ile Leu Gln Leu Pro Arg Asp Arg Phe Lys Arg
 1 5 10 15
 Thr Ser Phe Phe Leu Trp Val Ile Ile Leu Phe Gln Arg Thr Phe Ser
 20 25 30
 Ile Pro Leu Gly Val Ile His Asn Ser Thr Leu Gln Val Ser Asp Val
 35 40 45
 Asp Lys Leu Val Cys Arg Asp Lys Leu Ser Ser Thr Asn Gln Leu Arg
 50 55 60
 Ser Val Gly Leu Asn Leu Glu Gly Asn Gly Val Ala Thr Asp Val Pro
 65 70 75 80
 Ser Ala Thr Lys Arg Trp Gly Phe Arg Ser Gly Val Pro Pro Lys Val
 85 90 95
 Val Asn Tyr Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Glu
 100 105 110
 Ile Lys Lys Pro Asp Gly Ser Glu Cys Leu Pro Ala Ala Pro Asp Gly
 115 120 125
 Ile Arg Gly Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
 130 135 140
 Gly Pro Cys Ala Gly Asp Phe Ala Phe His Lys Glu Gly Ala Phe Phe
 145 150 155 160
 Leu Tyr Asp Arg Leu Ala Ser Thr Val Ile Tyr Arg Gly Thr Thr Phe
 165 170 175
 Ala Glu Gly Val Val Ala Phe Leu Ile Leu Pro Gln Ala Lys Lys Asp
 180 185 190
 Phe Phe Ser Ser His Pro Leu Arg Glu Pro Val Asn Ala Thr Glu Asp
 195 200 205
 Pro Ser Ser Gly Tyr Tyr Ser Thr Thr Ile Arg Tyr Gln Ala Thr Gly
 210 215 220
 Phe Gly Thr Asn Glu Thr Glu Tyr Leu Phe Glu Val Asp Asn Leu Thr
 225 230 235 240
 Tyr Val Gln Leu Glu Ser Arg Phe Thr Pro Gln Phe Leu Leu Gln Leu
 245 250 255
 Asn Glu Thr Ile Tyr Thr Ser Gly Lys Arg Ser Asn Thr Thr Gly Lys
 260 265 270
 Leu Ile Trp Lys Val Asn Pro Glu Ile Asp Thr Thr Ile Gly Glu Trp
 275 280 285
 Ala Phe Trp Glu Thr Lys Lys Asn Leu Thr Arg Lys Ile Arg Ser Glu
 290 295 300
 Glu Leu Ser Phe Thr Val Val Ser Asn Gly Ala Lys Asn Ile Ser Gly
 305 310 315 320
 Gln Ser Pro Ala Arg Thr Ser Ser Asp Pro Gly Thr Asn Thr Thr Thr
 325 330 335
 Glu Asp His Lys Ile Met Ala Ser Glu Asn Ser Ser Ala Met Val Gln
 340 345 350
 Val His Ser Gln Gly Arg Glu Ala Ala Val Ser His Leu Thr Thr Leu
 355 360 365
 Ala Thr Ile Ser Thr Ser Pro Gln Ser Leu Thr Thr Lys Pro Gly Pro
 370 375 380
 Asp Asn Ser Thr His Asn Thr Pro Val Tyr Lys Leu Asp Ile Ser Glu
 385 390 395 400
 Ala Thr Gln Val Glu Gln His His Arg Arg Thr Asp Asn Asp Ser Thr
 405 410 415
 Ala Ser Asp Thr Pro Ser Ala Thr Thr Ala Ala Gly Pro Pro Lys Ala

NIH316.001VPC_SEQLIST.TXT

```

      420      425      430
Glu Asn Thr Asn Thr Ser Lys Ser Thr Asp Phe Leu Asp Pro Ala Thr
      435      440      445
Thr Thr Ser Pro Gln Asn His Ser Glu Thr Ala Gly Asn Asn Thr
      450      455      460
His His Gln Asp Thr Gly Glu Glu Ser Ala Ser Ser Gly Lys Leu Gly
465      470      475      480
Leu Ile Thr Asn Thr Ile Ala Gly Val Ala Gly Leu Ile Thr Gly Gly
      485      490      495
Arg Arg Thr Arg Arg Glu Ala Ile Val Asn Ala Gln Pro Lys Cys Asn
      500      505      510
Pro Asn Leu His Tyr Trp Thr Thr Gln Asp Glu Gly Ala Ile Gly
      515      520      525
Leu Ala Trp Ile Pro Tyr Phe Gly Pro Ala Ala Glu Gly Ile Tyr Ile
      530      535      540
Glu Gly Leu Met His Asn Gln Asp Gly Leu Ile Cys Gly Leu Arg Gln
545      550      555      560
Leu Ala Asn Glu Thr Thr Gln Ala Leu Gln Leu Phe Leu Arg Ala Thr
      565      570      575
Thr Glu Leu Arg Thr Phe Ser Ile Leu Asn Arg Lys Ala Ile Asp Phe
      580      585      590
Leu Leu Gln Arg Trp Gly Gly Thr Cys His Ile Leu Gly Pro Asp Cys
      595      600      605
Cys Ile Glu Pro His Asp Trp Thr Lys Asn Ile Thr Asp Lys Ile Asp
      610      615      620
Gln Ile Ile His Asp Phe Val Asp Lys Thr Leu Pro Asp Gln Gly Asp
625      630      635      640
Asn Asp Asn Trp Trp Thr Gly Trp Arg Gln Trp Ile Pro Ala Gly Ile
      645      650      655
Gly Val Thr Gly Val Val Ile Ala Val Ile Ala Leu Phe Cys Ile Cys
      660      665      670
Lys Phe Val Phe
      675

```